Chapter 1

Clinical Trials

1. History of Clinical Trials

The evaluation of a medical treatment or therapeutic procedure has a very long history dating back thousands of years[1] but it was not until the middle of the 20th century that the clinical trial as we know it today was first developed. Earlier trials, such as trial of treatments for scurvy conducted by James Lind,[2] were flawed in one vital respect, they failed to include a robust system of randomisation. It is widely acknowledged that the first properly conducted randomised trial was an evaluation of streptomycin in pulmonary tuberculosis conducted by the British Medical Research Council.[3] The trial was different from trials that had gone before because of the concealment of allocation before randomisation. Knowledge of the randomisation schedule before enrolment could lead to bias which would seriously compromise the trial and invalidate the results[4], see Section 5.

A total of only 107 patients were randomised in the streptomycin trial, but this was sufficient to demonstrate a significant benefit in mortality and radiological improvement at six months. In addition to the concealment of the

allocation there were other features of this trial that are worthy of note. The assessment of the chest radiographs was performed by a radiologist who was blind to the allocated treatment as were the laboratory technicians performing the bacteriological assessments. Of particular note was the five year follow-up which demonstrated that the early beneficial effects were not sustained as a consequence of the emergence of streptomycin resistance in the treated arm, Table 1, [3] illustrating the importance of long term follow-up.

Table 1: Mortality at 6-months and 5-years in the Streptomycin trial

	6-month results[3]				5-year results[5]			
	Streptomycin		Bed rest		Streptomycin		Bed rest	
	N	%	N	%	N	%	N	%
Mortality	4	7%	14	27%	32	58%	35	67%
N assessed	55	100%	52	100%	55	100%	52	100%

Reflecting on his involvement in the streptomycin trial many years later, Sir John Crofton commented that "For many of those of us who had been involved in the MRC

streptomycin trial, randomised trials became a way of life, and provided much of the evidence upon which rational treatment policies came to be based." [6] The streptomycin trial became the model on which the assessment of subsequent new drugs and regimens for tuberculosis were evaluated, a model that spread to many other areas of medicine and has become the gold standard whereby new treatments are expected to be assessed. However, the presence of randomisation is not sufficient in itself if there are flaws in the design or conduct of the study,[7] see Box 1. As a result of the historical importance of tuberculosis in clinical trials and the many challenges posed by the natural history and prolonged treatment required for tuberculosis, examples from tuberculosis trials are used to describe the concepts in this chapter.

2. The importance of RCTs

The limitations of observational studies

Observational studies might appear to be an attractive alternative to randomised trials since they may be more representative of the patient population, are much simpler and require fewer resources to conduct, and can be usually be completed in a shorter space of time than is required for a randomised trial. They do,

however, have serious limitations and need to be interpreted with considerable caution because of potential biases.[8] An example of this is a case study of a randomised trial and an observational study in which the same two anti-retroviral regimens were compared using death as the endpoint.[9] Unexpected results in the observational study were probably due to patients with poorer prognosis being given one regimen in preference to the other.[9] Guyatt et al summarise potential limitations of nonrandomised studies noting common sources of bias include failure to develop appropriate eligibility criteria, differences in measurement of exposure and outcome, failure to adequately control for confounding and incomplete follow-up[7], Box 1. The difference between the limitations in randomised trials and observational studies is that the former can be avoided by careful planning whereas it is often not possible to avoid those in observational studies.

Box 1: Possible limitations in randomised trials and observational studies (based on Guyatt et al. [7])

Randomised trials	**Observational studies**
1 Lack of allocation concealment	1 Failure to apply appropriate

2 Lack of blinding 3 Incomplete accounting of patients and outcomes 4 Selective outcome reporting bias 5 Other including early stopping, unvalidated outcomes	eligibility criteria 2 Flawed measurement of exposure and outcome 3 Failure to adequately control confounding 4 Incomplete follow-up

A criticism often made of randomised trials is that they do not reflect real life[10]. An advantage of observational data, provided it is systematically collected, is the potential completeness of coverage, since patients who do not satisfy the eligibility requirements or those who are unwilling to participate in a randomised trial are not excluded. Valuable information on outcomes of different interventions may be obtained from medical records of high quality but the limitation of lack of randomisation will always mean potential differences in the patient populations cannot be fully discounted or allowed for.

Pragmatic and explanatory trials

A trial is described as pragmatic if the objective is to evaluate the intervention under conditions that are close to usual care and as

explanatory if the objective is to evaluate the intervention under optimal conditions, testing the principle of whether an intervention actually works. This distinction is more a continuum than a dichotomy with trials having features that could be considered more or less pragmatic or explanatory.[11] Very loose eligibility criteria for entry into the study, including all patients who would require treatment irrespective of other comorbidities, for example, would make a trial more pragmatic whereas regular follow-up visits outside of usual care would make a trial more explanatory. Both types of trials are of value – explanatory trials provide proof of concept that an intervention is efficacious and can be delivered safely while pragmatic trials show what is likely to happen when the intervention is implemented in practice.

The initial trials of short course chemotherapy for tuberculosis conducted by the British Medical Research Council in East Africa were more explanatory than pragmatic, conducted under strictly controlled conditions; patients were hospitalised throughout treatment and followed intensively throughout treatment and for 24 months of follow-up.[12] Subsequently, a trial conducted in Algeria under programme conditions with patients seen mostly as outpatients with limited supervision of their

treatment gave results consistent with the findings of earlier studies[13], thus strrengthening the evidence base for the regimen.

A comparison of results from a controlled trial and those from the tuberculosis programme, both conducted in Kenya, gave contrasting results, higlighting the limitations of highly explanatory trials. Whilst the survey confirmed the poor results obtained in the trial with a regimen of thiacetazone and isoniazid alone, the results of a regimen with an initial supplement of streptomycin were substantially better in the trial, 96% culture negative at one year, compared to only 78% in the programme. The difference in outcomes was attributed to poorer adherence to treatment in the continuation phase of treatment under routine conditions.[14]

3. Types, phases and designs of trials

There are traditionally considered to be four separate phases of clinical trials which are described in Box 2.

Box 2. Phases of clinical trials

Phase I	The first time a drug has been given to a human, participants are healthy volunteers. The objective is to assess safety and achieve some indication of a maximum tolerated dose of the drug.
Phase II	Participants are patients with the disease being studied. The objective is to achieve a preliminary evaluation of efficacy and further explore safety over a longer period of time in a larger group of patients. Phase II is often split into phase IIA with more focus on dose selection and safety and phase IIB with more focus on efficacy, often on an intermediate endpoint.
Phase III	The pivotal phase III trial involves treatment with the new drug for the intended duration with sufficient numbers of patients to allow for an unequivocal demonstration of efficacy on a definitive patient-relevant outcome.
Phase	The objective of the phase IV trial is

IV	to collect longer-term safety data in much larger numbers of patients than were enrolled in previous trials; often embedded into routine practice post-licensing.

In TB drug development, a short (14-day) early bactericidal activity (EBA) study to evaluate safety and compare doses in 12-15 patients per arm would be a phase IIA trial, whereas a study evaluating decline in TB bacilli over 8-12 weeks in 60-100 patients per arm would be a phase IIB trial. A phase III trial would evaluate treatment failure and relapse over 18-24 months of follow-up as the primary endpoint with hundreds of patients per arm, with usual standard of care are the comparator.

Classifying trials into phases can be overly restrictive. The development pathway may not always the same as an intervention may move from a small early-phase trial directly into a large confirmatory trial if the results are promising.

Superiority and Non-inferiority

The commonest type of type of late-phase (II-III) trial has a superiority design where the objective of the trial is to evaluate whether the intervention has superior efficacy to a standard

treatment control. However, non-inferiority designs are increasingly used, where the objective is to evaluate whether the intervention has efficacy that is as good as that of the control. To demonstrate non-inferiority, it is necessary to show that the intervention is no more inferior than a pre-specified amount. This difference is called the margin of non-inferiority. A non-inferiority trial is only appropriate when the intervention has additional benefits over the control such as being less toxic, less costly or of shorter duration. The use of non-inferiority trials for interventions with no additional benefit has led to some over-reactive criticism, with some describing any non-inferiority trial as unethical [15-20]. The current treatment of Multi-Drug Resistant TB (MDR-TB) is an area where non-inferiority trials are appropriate. MDR-TB is currently treated with combinations of toxic drugs daily for 20-24 months, severely limiting a patient's ability to return to work and other daily activities. Trials are evaluating substantially shorter, less toxic regimens that would result in major patient benefit even if efficacy is only at least as good as that of the current treatment[21].

Phase III trials for new treatments for drug sensitive TB commonly have a non-inferiority design since the standard 6-month regimen has

excellent efficacy in clinical trials that does not always translate well into clinical practice due to the long duration of treatment. A shorter 2 month regimen, for example, that had efficacy not much worse than the standard of care would likely translate into much improved outcomes in practice due to improved adherence. The remainder of this chapter focuses on superiority trials.

Adaptive trial designs

In a traditional fixed-sample trial, analysis occurs only at the end of the trial once all patients have completed follow-up and all data accrued. An alternative to this approach is an adaptive design where there are one or more interim analyses during the course of the study and these interim results used to adapt the study design. Possible adaptations include changing the sample size, stopping a trial early for overwhelming efficacy or lack of benefit, or dropping arms in a multi-arm study. Importantly, such adaptation should never be used to attempt to salvage a failing trial and procedures for adaptation must be pre-specified in the study protocol before the study begins (see Trial Monitoring section below). Consideration must be given of the impact of any interim analyses on the overall type I and type II errors (see section 4 Power and Sample

Size). Where there are several promising new interventions but limited resources, a particularly attractive design is the Multi-Arm Multi-Stage (MAMS) design where multiple arms are compared simultaneously with a single control. The MAMS design allows for multiple interim analyses to facilitate the early termination of poorly performing arms in order that resources can be focused on the more promising arms. This methodology was developed in cancer trials[22,23] but is being adapted for use in trials of new TB drugs[24,25].

Other trial designs

In most randomised controlled trials, individual participants are randomly allocated to treatment arms. An alternative is a cluster randomised trial which evaluate an intervention that is applied to a 'cluster' of individual participants, at the community- or health system-level, where clusters rather than individual participants are randomised to receive one of the interventions[26]. An intervention such as improved methods for tuberculosis case-finding, for example, would be better suited to a cluster-randomised trial.

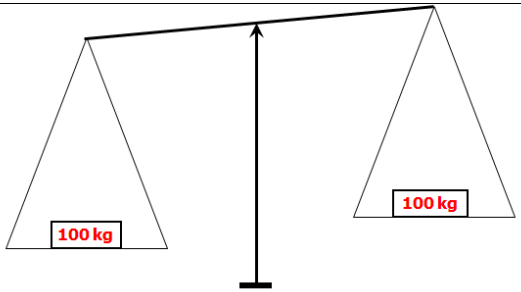
Other trial designs that are not covered here include step wedge cluster randomised trials, factorial trials and sequential multiple assignment randomized trials (SMARTs).

4. Power and sample size

The sample size of a trial is the minimum number of participants required to be enrolled in a clinical trial to achieve the trial objective. It is unethical to enrol either insufficient patients to achieve the trial objective or more patients than are necessary.. Sample size calculations are driven by minimising the probability of Type I and Type II errors. In a superiority trial, the Type I error is the probability of demonstrating a difference when no difference exists while the Type II error is the probability or failing to show a difference when there actually is a real difference between interventions (see figure).

It is common to maintain the type I error rate low at 5%, but usually acceptable to allow a type II error rate of 10-20%. The Power of a trial is the probability that an effective intervention will be demonstrated to be so and is calculated by subtracting the type II error rate from 100%; a power of 90% corresponds to a type II error rate of 10%.

Figure: Type I and II errors



Type I error. Wrongly declaring a difference when one doesn't exist.

Type II er
one does e

When a trial includes multiple interventions, multiple primary endpoints or when multiple analyses have been conducted (such as in an adaptive design), there is an increased chance of falsely finding a statistically significant difference if a conventional significance level of 5% is used for each comparison. In this context, it is appropriate to either adjust the individual significances levels to maintain the overall type I error rate at an acceptable rate (such as 5%) or take multiple testing into account when interpreting the results[27,28].

Having decided on values for the Type I error rate and power, the targeted effect size or difference between treatments is another key driver for the sample size. A trial designed with adequate power to detect a small effect size will be larger than one designed to detect a

large effect size. It is good practice to link the targeted effect size to the minimum clinically important difference (MCID) so that the trial is designed to detect effect sizes at least as large as the MCID and would only miss effects not considered clinically important. This highlights the distinction between statistical significance when there is evidence that a difference, however small, does exist and clinical significance when the difference is considered large enough for the intervention to change practice.

An important factor that influences the sample size is the expected outcome in the control arm. If this turns out to be very different to what was assumed when the trial was designed, the trial may turn out to be underpowered. It is therefore recommended to use a conservative estimate from previous trials, if available.

It is usually necessary to enrol more patients than is required for the analysis to account for loss to follow-up. It is important to make realistic estimates about the expected rate of loss to follow-up when designing a trial; these need to be kept as low as possible when conducting the trial since the true outcome of such patients remains unknown. Differential losses to follow-up may indicate that one

treatment regimen is less acceptable or more toxic than another and results in bias in interpretation of the data. Other sources of bias are described in the next section.

5. Methods for avoidance of bias

The estimated effect of an intervention from a clinical trial is said to be biased if there is systematic error such that it does not reflect the true intended effect of the intervention. There are a number of sources of bias and corresponding measures to avoid bias.

Randomisation is the most critical method in an RCT to prevent bias as, if properly implemented, it ensures that any known and unknown factors that might affect outcomes are balanced between arms with any imbalance occurring by chance. Proper implementation involves adequate allocation concealment, ensuring that neither the patient nor the investigator is aware of which arm a patient will be allocated to before consent has been obtained for enrolment in the trial. This avoids selection bias where patients are selected for particular arms such as, for example, sicker participants being allocated to the control arm by an investigator concerned about a novel treatment.

Blinding (sometimes also called masking) is a well established method in clinical trials to avoid bias by keeping secret which treatment a patient is on. The strongest form of blinding is known as double-blind when neither the patient, clinicians, investigators nor any other individuals carrying out assessments (such as laboratory technicians) know what arm the patient has been allocated to until completion of the trial. The purpose of blinding is to ensure that every aspect of patient management and data collection is unaffected by the knowledge of which arm a patient has been allocated to. Blinding is particularly important when the primary endpoint has a subjective element such as a patient reported quality of life measure but less important when the primary endpoint is objective such as all-cause mortality.

Blinding is achieved in a treatment trial by giving patients on the control arm identical inactive tablets (placebo) at the same frequency as on the intervention arm. A double-blind trial can be very difficult; it is challenging to produce an inactive perfectly-matched placebo for the TB drug rifampicin, for example, which turns a patient's urine and bodily fluids an orange colour. Blinding will also substantially increase the complexity and

cost of a clinical trial due to the manufacture of matching placebo and central packing facilities separate from the trial sites.

Even when it is not feasible or desirable to blind patients and clinicians (in which case the trial is sometimes designated as being open label), it is still important to limit knowledge of patient allocation for endpoint assessors wherever possible and also ensure that only the Independent Data Monitoring Committee (IDMC, see Section 6) are aware of aggregated data by treatment arm. If the primary endpoint involves some clinical judgement (identifying AIDS-defining illness, for example) an endpoint review committee of experts independent to the trial could be convened to review the data blinded to treatment allocation and classify outcomes.

To ensure that no trial procedures change during the course of a trial as a response to accruing data (particularly important in an open label study) key aspects of the trial such as trial objectives, primary endpoint, primary methods of analysis should be clearly pre-specified in the protocol and remain unchanged once the first patient has been enrolled. The statistical analysis plan which provides details of the analysis of the primary and secondary

endpoints should also be finalised and signed off early in the trial.

Publication bias occurs where there is selective reporting of trial outcomes to favour the intervention or where whole trials with unfavourable outcomes are not reported. For this reason, it is now expected by most clinical trials funders that a trial is registered, namely key details are lodged with publically available registries (such as ClinicalTrials.gov or International Standard Randomised Controlled Trial Number Register, ISRCTN) before the study starts. In this way, individuals conducting systematic reviews or those wanting to find out about a particular disease or intervention can search registries to find a more comprehensive picture of which trials are being or have been conducted.

6. Trial Monitoring

Unless the quality of the data obtained in a randomised controlled clinical trial can be relied upon, the results of the trial will be of no value. If a trial is being conducted according to the principles of the International Conference on Harmonisation Good Clinical Practice (ICH GCP) regulations[29], regular monitoring is essential. This can be performed in a number

of ways and to varying degrees. A GCP requirement is that monitors should be appropriately trained and should have the scientific and/or clinical knowledge needed to monitor a trial adequately.[29] In the past, monitoring was often only done by monitors visiting study sites and reviewing data and documentation with little regard to a strategy of prioritisation within the process. More recently, alternative approaches have been advocated[30] and these include central statistical monitoring [31,32] and the use of a risk assessment conducted at the start of a trial to identify the most appropriate monitoring strategies for the trial and for individual sites[30]. A risk assessment should be reviewed on an annual periodic basis throughout a clinical trial and the monitoring techniques employed considered and updated accordingly. Monitoring, whether done on site, centrally or both, may highlight the need for additional site staff training or changes in trial procedures.

Safety monitoring and expedited reporting

In addition to the determination of efficacy, an essential assessment in clinical trials is the safety of the interventions being studied. It is recommended that patients should be asked at each trial visit about any disability or incapacity or adverse events that have occurred as well as hospitalisations amd consultations

with other medical practitioners . ICH GCP regulations set out the responsibilities for the notification of adverse events by investigators to sponsors.[29] These include the reporting of any defined serious adverse events (SAEs) within an agreed time frame with particular expedited reporting requirements by sponsor to regulatory authorities for Suspected Unexpected Serious Adverse Reactions (SUSARs).

The Independent Data Monitoring Committee and Trial Steering Committee

All trials of medicinal products are expected to have an Independent Data Monitoring Committee (IDMC), its purpose is to protect the safety of the trial participants, the credibility of the study and the validity of the study results.[33] The membership, which is often no more than three to five people, should be totally independent of the trial and include clinical trial, statistical and relevant clinical expertise. The IDMC is expected to meet regularly, commonly every 6 months during the trial to review study progress and unblinded data on safety and efficacy. At the conclusion of each meeting the IDMC will make its recommendations to an executive decision making body such as the Trial Steering Committee (TSC) as to whether the trial should continue as designed or whether

modifications should be made. These could include early termination of one or more study arms on account of safety concerns or proof beyond reasonable doubt of differences in efficacy between one of the study arms and its comparator.

The TSC is a committee which provides expert oversight of the trial, monitoring progress on a regular basis and receiving the recommendations of the IDMC. The majority of members of the TSC, including the chair, should be independent of the trial although additional observers may be present at TSC meetings. In addition to deciding on the appropriate response following receipt of the IDMC recommendations, the TSC may be required to attend to issues of concern regarding trial conduct. These include poor recruitment or poor data quality, the approval of proposed protocol amendments or new trial sub-studies, the approval of requests for early release of data or external applications for the use of stored samples, and the approval of study reports or presentations.

7. Ethical approval and informed consent

Before commencing a trial, approval needs to be obtained from an independent research ethics committee to protect the rights and interests of the trial participants. This approval will usually be from more than one committee and will typically include a central ethics committee, often based in the same country as the trial sponsor, and ethics committees in each participating country or site. In the United Kingdom, for example, only the central committee approval is required for all sites. Any amendments to be made to the study protocol need to be approved by the ethics committee(s) before they can be introduced.

Informed consent must be obtained from all persons being considered for enrolment to the trial before any investigations are performed; this includes investigations to assess the eligibility for admission to the trial. Key information that needs to be conveyed to the patient includes the rationale for the study, potential risks and benefits, the trial treatments, the randomisation process, the follow-up schedule and right of participants to withdraw at any time. Before enrolling in the study a patient consent form needs to be signed. If the person cannot read patient information

documentation, an independent witness should be present during the consent process.

8. Dissemination and impact

It is important that the results of a clinical trial are published in peer-reviewed journals and disseminated in scientific conferences but this is insufficient as this will only reach the scientific community. There is an ethical obligation that the results of the trial should also be shared with the trial participants and the communities where the trial was conducted including the investigators and other site staff. This can sometimes take the form of a community meeting as a celebration of trial completion.

The impact of a clinical trial is often measured by how high profile the journal that the paper is published in is and how many times it is subsequently cited in other scientific publications. This is only one component of impact which can also be measured by, among other things, the extent to which national and international treatment guidelines are changed as a result of the trial, whether practitioners are actually using the new intervention to treat their patients or resulting advocacy activities

by patient and community groups. Funding agencies, such as the British Medical Research Council, now require that groups conducting clinical trials include steps to increase impact beyond just publication of the results in a peer-reviewed journal.

To have impact broader than just to the scientific community, additional methods of dissemination can therefore include press releases, videos posted on YouTube, policy briefing documents for governments and health ministries and direct contact with organisations such as the World Health Organisation that produce treatment guidelines. Having reports in journals published as open access, such that the publication is freely available to everyone without subscription, will also increase dissemination.

References

1. Bull J. The historical development of clinical therapeutic trials. **Journal of Chronic Diseases** 1959;10:218-48.
2. Lind J. A Treatise on the Scurvy: In Three Parts, Containing an Inquiry Into the Nature, Causes, and Cure, of that Disease: A. Millar; 1757.

3. Medical Research Council. Streptomycin Treatment of Pulmonary Tuberculosis: a Medical Research Council investigation. **BMJ** 1948;2:769.
4. Student. The Lanarkshire Milk Experiment. **Biometrika** 1931:398-406.
5. Fox W, Sutherland I, Daniels M. A five-year assessment of patients in a controlled trial of streptomycin in pulmonary tuberculosis: Report to the Tuberculosis Chemotherapy Trials Committee of the Medical Research Council. **QJM** 1954;23:347.
6. Crofton J. The MRC randomized trial of streptomycin and its legacy: a view from the clinical front line. **Journal of the Royal Society of Medicine** 2006;99:531-4.
7. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). **Journal of clinical epidemiology** 2011;64:407-15.
8. Grimes DA, Schulz KF. Bias and causal associations in observational research. **Lancet** 2002;359:248-52.
9. Dunn D, Babiker A, Hooker M, Darbyshire J. The dangers of inferring treatment effects from observational data: a case study in HIV infection. **Control Clin Trials** 2002;23:106-10.

10. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. **J Clin Epidemiol** 2009;62:499-505.
11. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, Tunis S, Bergel E, Harvey I, Magid DJ, Chalkidou K. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. **Journal of clinical epidemiology** 2009;62:464-75.
12. Fox W, Ellard G, Mitchison D. Studies on the treatment of tuberculosis undertaken by the British Medical Research Council tuberculosis units, 1946-1986, with relevant subsequent publications. **Int J Tuberc Lung Dis** 1999;3:S231-S79.
13. Algerian Working Group/British Medical Research Council. Controlled clinical trial comparing a 6-month and a 12-month regimen in the treatment of pulmonary tuberculosis in the Algerian Sahara. **Am Rev Respir Dis** 1984;129:921-8.
14. Kent P, Fox W, Miller A, Nunn A, Tall R, Mitchison D. The therapy of pulmonary tuberculosis in Kenya: a comparison of the results achieved in controlled clinical trials with those achieved by the routine treatment services. **Tubercle** 1970;51:24-38.
15. Garattini S, Bertele V. Non-inferiority trials are unethical because they disregard patients' interests. **Lancet** 2007;370:1875-7.

16. Soliman EZ. The ethics of non-inferiority trials. **Lancet** 2008;371:895; author reply 6-7.
17. Nunn AJ, Meredith SK, Spigelman MK, Ginsberg AM, Gillespie SH. The ethics of non-inferiority trials. **Lancet** 2008;371:895; author reply 6-7.
18. Menten J, Boelaert M. The ethics of non-inferiority trials. **Lancet** 2008;371:896; author reply -7.
19. Gandjour A. The ethics of non-inferiority trials. **Lancet** 2008;371:895; author reply 6-7.
20. Chuang-Stein C, Beltangady M, Dunne M, Morrison B. The ethics of non-inferiority trials. **Lancet** 2008;371:895-6; author reply 6-7.
21. Nunn AJ, Rusen I, Van Deun A, Torrea G, Phillips PP, Chiang CY, Squire SB, Madan J, Meredith SK. Evaluation of a standardized treatment regimen of anti-tuberculosis drugs for patients with multi-drug-resistant tuberculosis (STREAM): study protocol for a randomized controlled trial. **Trials** 2014;15:353.
22. Royston P, Barthel F, Parmar M, Choodari-Oskooei B, Isham V. Designs for clinical trials with time-to-event outcomes based on stopping guidelines for lack of benefit. **Trials** 2011;12:81.

23. Sydes MR, Parmar MK, James ND, Clarke NW, Dearnaley DP, Mason MD, Morgan RC, Sanders K, Royston P. Issues in applying multi-arm multi-stage methodology to a clinical trial in prostate cancer: the MRC STAMPEDE trial. **Trials** 2009;10:39.
24. Phillips PP, Gillespie SH, Boeree M, Heinrich N, Aarnoutse R, McHugh T, Pletschette M, Lienhardt C, Hafner R, Mgone C, Zumla A, Nunn AJ, Hoelscher M. Innovative trial designs are practical solutions for improving the treatment of tuberculosis. **Journal of Infectious Diseases** 2012;205 Suppl 2:S250-7.
25. Bratton DJ, Phillips PP, Parmar MK. A multi-arm multi-stage clinical trial design for binary outcomes with application to tuberculosis. **BMC Med Res Methodol** 2013;13:139.
26. Hayes RJ, Moulton LH. Cluster randomized trials. Boca Raton, Fla. ; London: Chapman & Hall/CRC; 2009.
27. Freidlin B, Korn EL, Gray R, Martin A. Multi-arm clinical trials of new agents: some design considerations. **Clinical Cancer Research** 2008;14:4368-71.
28. Wason JM, Stecher L, Mander AP. Correcting for multiple-testing in multi-arm trials: is it necessary and is it done? **Trials** 2014;15:364.

29. ICH Guidelines. 2014. at http://www.ich.org/products/guidelines.html.)
30. Baigent C, Harrell FE, Buyse M, Emberson JR, Altman DG. Ensuring trial validity by data quality assurance and diversification of monitoring methods. **Clin Trials** 2008;5:49-55.
31. Venet D, Doffagne E, Burzykowski T, Beckers F, Tellier Y, Genevois-Marlin E, Becker U, Bee V, Wilson V, Legrand C, Buyse M. A statistical approach to central monitoring of data quality in clinical trials. **Clin Trials** 2012;9:705-13.
32. Bakobaki JM, Rauchenberger M, Joffe N, McCormack S, Stenning S, Meredith S. The potential for central monitoring techniques to replace on-site monitoring: findings from an international multi-centre clinical trial. **Clin Trials** 2012;9:257-64.
33. Ellenberg SS, Fleming TR, DeMets DL. Data monitoring committees in clinical trials: a practical perspective: John Wiley & Sons; 2003.