Risks in Randomized Controlled Trials: Role of Interim Analysis and Data Monitoring Committee

Haider A. Naqvi

Randomised Controlled Trials (RCTs) are “controlled” experiments intended to obtain answers relevant for clinical practice. The three features of RCTs which confer the advantage over other (i.e. uncontrolled) experiments are randomization, placebo-control group and blinding. The important aspect of blinding or masking is that the participants, investigators and assessors are masked to treatment allocation and other important aspect of the trial conduct. This is intended to minimize the bias associated with the pre-conceived ideas.1 It is important to provide an oversight on the outcome measure since experimentation involves risks. The risk can be associated with the intervention or fatality of the illness, which is further enhanced when the intervention involves the use of placebo and mortality as an outcome. The new intervention too is not without its own perils. Since the product is new, the side-effects are mostly unknown. The investigational product developed through various stages of animal and human subject research also stands in trial, with researchers’ careers and funding money at stake. This editorial intends to describe the measures taken to mitigate the risk in RCTs.

In the initial instance, there should be an ethically and scientifically valid reason for recruiting seriously ill patients in to the randomised trial and subjugating them to rigorous monitoring procedures. There has to be a clinical equipoise, among the community of experts, in order to carry out a controlled experimentation. This is not only a scientific but also an ethical pre-requisite for conducting an RCT - uncertainty should lead to inquiry in the domain of science. However, the (human) rights of individuals involved in the trial should not take precedence over intellectual pursuits and good of society through discovery, alleviation of suffering and greater good of many individuals. In certain conditions, patients are also desperate to seek new treatment, given the fact that illnesses are chronic and effective treatment is not available. With this background, patients are enrolled in the trial with the intent to keep a balance between therapeutic benefits entailed in participation in the trial and risks associated with trying out a new intervention. Monitoring the results by the physicians involved in the trial conduct is expected to disturb their initial equipoise of preferring one treatment over another. Given this state of affairs, the participating physicians might not be inclined to recruit patients to placebo or the treatment group. As in general matters of life, someone has to be in charge of the situation. This responsibility is generally relegated to a committee of scientific experts who has the responsibility to examine the ethical, statistical and operational matters of the research according to a pre-defined schedule.2

Members of the Data Monitoring Committee (DMC) not only have to be informed regarding the matters of science, but also regarding matters outside the (accumulating data) trial. In fulfilling this task, they have to be independent, with no overt of covert conflict of interest. As ideal as it may sound, the financial, academic or intellectual and emotional conflict of interest has to be minimal and non-interfering with the conduct of trial.3 Charged with this responsibility, the DMC should provide an oversight to the trial by reviewing the interim data. They intervene by advising the trial executive committee whether to continue or stop the trial early. The three conditions in which trial can be stopped early are called efficacy, safety and futility rule. In the efficacy rule the trial is discontinued as there is a clearly demonstrated efficacy of one intervention over another with future recruitment unable to change the results. The safety rule pertains to clearly demonstrated harm to the trial participants which was unanticipated before the start of the trial. The futility rule of stopping the trial has to do with no favorable benefit of one intervention over another thereby reducing the harm of exposing the participants to inferior intervention or withholding otherwise effective treatment. The DMC can also advise protocol changes to the design, analysis or conduct of the study since they have an access to the accumulating data and interim results on efficacy or harm of the intervention.4

DMC has to have mix of representation from subject specialists, trialists, bioethicists and other professionals, who are independent, have no conflict of interest and have access to unblinded trial data, on the primary endpoint and safety outcomes, as it accrues. There are statistical methods, based on group sequential interim analysis, of trial data as it accrues.5 However, the decision to stop the trial is not based on statistical tests.

---

Department of Psychiatry, The Ziauddin University, Karachi.

Correspondence: Dr. Haider A. Naqvi, B-9, Block 16, Gulistan-e-Jauhar, Karachi.
E-mail: haiderinalaqvi9@gmail.com

Received: September 04, 2014; Accepted: July 04, 2015.
alone. There are corollaries which need to be considered. Phase III clinical trials are expected to provide evidence on the efficacy of therapeutic agent/device beyond reasonable doubt, leading to change in clinical practice. Therefore, sufficient evidence needs to be accumulated to establish the efficacy and long-term safety of the therapeutic agent/device. A premature termination may raise concerns related to long-term safety of the compound. There is no agreed definition of 'efficacy beyond reasonable doubt' in the context of regulatory approvals and litigations by current or future patients. However, statistics provides guidance in term of objective decision making. A collective wisdom of the DMC is required to tackle the challenge facing the early termination of trial based on crossing of group sequential boundaries in the interim analysis for efficacy or harm, review of external evidence from other on-going trials or meta-analysis of existing data related to the trial. There are four general approaches to statistical monitoring of the interim trial results, allowing for early termination. The group sequential boundary approach does interim analysis on 'pre-defined time' of the trial as the data on primary (and secondary) endpoint accrues. The approach looks to plot the effect size on a graph with standardized Z-scores on the Y-axis and fraction of trial information on the X-axis. The sequential probability ratio allows for the continuous examination of the trial results. A triangular plot boundary examines the interim results, contingent on the null and alternative hypothesis, to lie in a range of values. Group sequential boundaries represent the accumulated wisdom of monitoring interim results in clinical trials given the nature of ethical/safety concerns related to participants in clinic trial. The conditional power (stochastic) approach takes in to consideration the (original) power calculation and its ability to answer the trial question given the results from interim analysis.

The paradoxical problem with interim analysis is that the data tends to be subjected to random variation and fluctuation of results. The baseline characteristics of the recruited subjects and the observed effect estimates are also expected to be variable given the initial assumptions. Possibility of high-risk subjects entering the trial early and leading to biased estimate of higher efficacy cannot be ruled out. Repeated testing can and dose lead to false positive error rates. The possibility of falsely rejecting the null hypothesis is classically kept at 0.05%. When multiple tests are done this possibility is increased, if we do not take into account the statistical assumptions related to repeated testing, thereby falsely rejecting the null hypothesis. The group sequential boundaries are constructed in order to reduce this risk especially early in the trial when data fluctuations are substantial. The false positive error rate is adjusted to be conservative early on with future increment to the industry standard (0.05) conventional to trials. This enables meaningful analysis keeping in view the statistical considerations of frequentists approach to statistics. Practical considerations is that group sequential approach guard against inadvertent false positive rejection of null hypothesis, based on random variation in data. The boundary philosophy also protects the alpha spending function at various stages of the interim analysis of data. Plotting of results on a graph with specified boundaries also informs the DMC on the number of positive events required to have meaningful efficacy/harm results in remaining part of the trial. This not only checks on the underlying sample size assumptions but also enables the researchers to calculate the power required to check the difference in the two groups. If the initially hypothesized power (1-β) is unable to detect the meaningful difference, beyond the constructed (triangular) boundaries, then the trial can be stopped for futility. This consideration is equally applicable in the domain of harm/efficacy. The group sequential approach enables the DMC to re-calculate the power, given the underlying assumption on the event rate, during the conduct of the trial.

There are three statistical tests based on alpha-spending function in group sequential boundary approach. These are Peto, Pocock and O'Brien-Fleming tests (Figure 1). Among the three, the O'Brien-Fleming is by far the most robust approach. The O'Brien-Fleming boundary, divided over for four analyses, would estimate the false positive error rate to be 0.0001, 0.004, and 0.019 for 3 interim analyses while the final false positive error rate will be 0.043 which is nearly equal to the 0.05 standard significance level. They can serve as predefined statistical stopping guidelines. If the results of interim analysis cross these predefined statistical stopping guidelines then DMC could deliberate on the possibility of stopping the trial early. It also needs to look at the secondary outcomes which could warrant

Figure 1: Three stopping guidelines in the case of five analyses (four interim and one final. Each of the guidelines keeps the overall type I error at 5% by Dr Neal Alexander, London School of Hygiene and Tropical Medicine (LSHTM).
continued accrual of participants to offset any initial findings. If the aggregated data on two arms of the intervention has no major concern on the primary and secondary outcomes then case can be made for early termination based on the efficacy. Decision to stop the trial early should only be made if there is reasonable evidence, beyond doubt, on the efficacy of the intervention. Additionally, evidence external to the trial should also be kept in mind during the DMC meeting examining the interim data.

There can be positive as well as negative consequences to more frequent interim analysis during the trial duration. The frequent monitoring of results has implications on cost, complexity of trial coordination and inadvertent unmasking related to more frequent monitoring. Contrarily, it can be argued that multiple analysis would divide the data in to smaller bits, thereby increasing the random variation in subgroup of participants and haste of conclusion. Frequent monitoring, however, is expected to inform the trial executives on the quality of study conduct and other operational measure which could be corrected thereby enhancing the trial integrity. The Chairperson of the DMC needs to be well versed in matters of science as well as management. He/she needs to be a person who has background experience of conducting and supervising a clinical trial. Additionally, he/she should have no major conflict of interest. The Chairperson of DMC is expected to take various actions based on the conduct of the trial as well as the contextual factors of the trial. Ideally, he/she should base his actions on the rules set out in the DMC Charter. All major phase III trial should have written down charter, which should be signed by the principle investigator, the trial sponsor and the hospital administration or the University leadership in case of an academic setting. The course of action would have been set-out or agreed upon during the initial meeting with the trial stakeholders.

The DMC can be internal or external to the trial executive committee or investigators. The Federal Drug Authority in USA requires that Phase III trials which require licensing approval should have an external, independent DMC. If the risks are low as the drug has known toxicity or efficacy then case can be made for an internal DMC. The reason for having an external/ independent DMC in phase III trial is that interventions are looking to test new investigational product in a serious medical condition with probably high morbidity and mortality. The safety concerns dictate that there should be an independent monitoring. Additionally, the use of placebo and blinding makes the participating clinicians unaware of the segregated data on serious adverse events. Sometimes the follow-up is spread over course of months, after the initial intervention, favoring monitoring by external DMC. Argument against the external DMC can be that at times there are no known toxicity or life threatening complication associated with the intervention. DMCs can add complexity (and cost) to the data monitoring process especially in multinational trial setting. In such cases an internal DMC is constituted. The Internal DMC could comprise of executives, researchers and trialists who are employed with the same organization but are working on other projects. Care has to be taken that these individuals should not have any direct conflict of interest related to the task at hand.

The clinical trials research is increasingly getting organized in recent times with oversight from various regulatory authorities. The donor organizations, academic institutions and regulatory bodies are setting up their own guidelines for monitoring risks in research. It is time that certifying bodies like College of Physicians and Surgeons Pakistan develop policies, procedures and operational capabilities to monitor trials conducted under its ambit.

REFERENCES