Int. J. Drug Res. Tech. 2016, Vol. 6 (3), 193-208 ISSN 2277 - 1506

International Journal of Drug Research and Technology

Available online at http://www.ijdrt.com

Review Article

BIOSTATISTICS & RESEARCH METHODOLOGY WITH AN OVERVIEW ON CLINICAL RESEARCH

Abhishek Kumar Yadav*

Department of Pharmaceutical Sciences, Dibrugarh University, Dibrugarh-786004, Assam, India

ABSTRACT

Clinical Research is a systematic study for new drugs in human subjects to generate data for discovering or verifying the Clinical, Pharmacological (including pharmacodynamic and pharmacokinetic) or adverse effects with the objective of determining safety and efficacy of the new drug. Clinical Research is conducted in 4 Phases. Phase I trials - the new drug is administered to a small number, around 20-80 healthy, informed volunteers under the close supervision of a doctor is to determine whether the new compound is tolerated by the patient's body and behaves in the predicted way. Phase II trials – the medicine is administered to a group of approximately 100-300 informed patients to determine its effect and also to check for any unacceptable side effects. Phase III trials - between 1000 and 5000 patients use statistics to analyze the results. If the results are favourable, the data is presented to the licensing authorities for a commercial licence. Phase IV trials - the medicine is make available to doctors, who start prescribing it & effects are monitored on thousands of patients to help identify any unforeseen side effects. The role of statics is to investigate proposed medical treatments, assessing benefits of competing therapies and establishing optimal treatment combination. Clinical Trial Methodology emphasizes the importance of statistical thinking in clinical research and presents the methodology as a key component of clinical research. Selection of the proper statistical test depends on the type and number of variables and whether parametric conditions are met. Various methods include t-tests, analysis of variance, repeated measures ANOVA, linear regression, analysis of covariance, non-parametric tests, binomial tests, chi-square test, Fisher's exact test, McNamara's test, Cochran-Mantel-Haenszel test, logistic regression, log-rank test, and Cox proportional hazards model. Bio statistical analysis is an important part of clinical trials because it allows researchers to determine whether or not their results were meaningful on a statistical level. Thus Biostatisticians are those who perform statistical programming, design, and analysis for clinical trial projects. Planning, coordinating and providing statistical analyses, summaries and reports of studies is also a part of their job profile. They are also responsible for New Drug Applications and Biological License Applications submissions.

Keywords: Biostatistics, Clinical research, ANOVA, Research methodology.

INTRODUCTION

Background and Purpose

The efficacy and safety of medicinal products should be demonstrated by clinical trials which follow the guidance in 'Good Clinical Practice: Consolidated Guideline' (ICH E6) adopted by the ICH, 1 May 1996. The role of statistics in clinical trial design and analysis is acknowledged as essential in that ICH guideline. The proliferation of statistical research in the area of clinical trials coupled with the critical role of clinical research in the drug approval process and health care in general necessitate a succinct document on statistical issues related to clinical trials. This guidance is written primarily to attempt to

harmonise the principles of statistical methodology applied to clinical trials for marketing applications submitted in Europe, Japan and the United States. As a starting point, this guideline utilised the CPMP (Committee for Proprietary Medicina Products) Note for Guidance entitled "Biostatistical Methodology in Clinical Trials in Applications for Marketing Authorisations for Medicinal Products" (December, 1994). It was also influenced by "Guidelines on the Statistical Analysis of Clinical Studies" (March, 1992) from the Japanese Ministry of Health and Welfare and the U.S. Food and Drug Administration document entitled "Guideline for the Format and Content of the Clinical and Statistical Sections of a New Drug Application" (July, 1988).[1] Some topics related to statistical principles and methodology is also embedded within other ICH guidelines, particularly those listed below. The specific guidance that contains related text will be identified in various sections of this document.

- E1A: The Extent of Population Exposure to Assess Clinical Safety.
- E2A: Clinical Safety Data Management: Definitions and Standards for Expedited Reporting.
- E2B: Clinical Safety Data Management: Data Elements for Transmission of Individual Case Safety Reports.
- E2C: Clinical Safety Data Management: Periodic Safety Update Reports for Marketed Drugs.
- E3: Structure and Content of Clinical Study Reports.
- E4: Dose-Response Information to Support Drug Registration.
- E5: Ethnic Factors in the Acceptability of Foreign Clinical Data.
- E6: Good Clinical Practice: Consolidated Guideline.
- E7: Studies in Support of Special Populations: Geriatrics.
- E8: General Considerations for Clinical Trials.
- E10: Choice of Control Group in Clinical Trials.
- M1: Standardisation of Medical Terminology for Regulatory Purposes.
- M3: Non-Clinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals.

This guidance is intended to give direction to sponsors in the design, conduct, analysis, and evaluation of clinical trials of an investigational product in the context of its overall clinical development. The document will also assist scientific experts charged with preparing application summaries or assessing evidence of efficacy and safety, principally from clinical trials in later phases of development.

Scope and Direction

The focus of this guidance is on statistical principles. It does not address the use of specific statistical procedures or methods. Specific procedural steps to ensure that principles are implemented properly are the responsibility of the sponsor. Integration of data across clinical trials is discussed, but is not a primary focus of this guidance. Selected principles and procedures related to data management or clinical trial monitoring activities are covered in other ICH guidelines and are not addressed here. This guidance should be of interest to individuals from a broad range of scientific disciplines. However, it is assumed that the actual responsibility for all statistical work associated with clinical trials will lie with an appropriately qualified and experienced statistician, as indicated in ICH E6. The role and responsibility of the trial statistician (see Glossary), in collaboration with other clinical trial professionals, is to ensure that statistical principles are applied appropriately in clinical trials supporting drug development. Thus, the trial statistician should have a combination of education/training and experience sufficient to implement the principles articulated in this guidance.[2]

For each clinical trial contributing to a marketing application, all important details of its design and conduct and the principal features of its proposed statistical analysis should be clearly specified in a protocol written before the trial begins. The extent to which the procedures in the protocol are

followed and the primary analysis is planned a priori will contribute to the degree of confidence in the final results and conclusions of the trial. The protocol and subsequent amendments should be approved by the responsible personnel, including the trial statistician. The trial statistician should ensure that the protocol and any amendments cover all relevant statistical issues clearly and accurately, using technical terminology as appropriate. The principles outlined in this guidance are primarily relevant to clinical trials conducted in the later phases of development, many of which are confirmatory trials of efficacy. In addition to efficacy, confirmatory trials may have as their primary variable a safety variable (e.g. an adverse event, a clinical laboratory variable or an electrocardiographic measure), a pharmacodynamic or a pharmacokinetic variable (as in a confirmatory bioequivalence trial).

Furthermore, some confirmatory findings may be derived from data integrated across trials, and selected principles in this guidance are applicable in this situation. Finally, although the early phases of drug development consist mainly of clinical trials that are exploratory in nature, statistical principles are also relevant to these clinical trials. Hence, the substance of this document should be applied as far as possible to all phases of clinical development. Many of the principles delineated in this guidance deal with minimising bias (see Glossary) and maximising precision. As used in this guidance, the term 'bias' describes the systematic tendency of any factors associated with the design, conduct, analysis and interpretation of the results of clinical trials to make the estimate of a treatment effect (see Glossary) deviate from its true value. It is important to identify potential sources of bias as completely as possible so that attempts to limit such bias may be made. The presence of bias may seriously compromise the ability to draw valid conclusions from clinical trials.[3]

Some sources of bias arise from the design of the trial, for example an assignment of treatments such that subjects at lower risk are systematically assigned to one treatment. Other sources of bias arise during the conduct and analysis of a clinical trial. For example, protocol violations and exclusion of subjects from analysis based upon knowledge of subject outcomes are possible sources of bias that may affect the accurate assessment of the treatment effect. Because bias can occur in subtle or unknown ways and its effect is not measurable directly, it is important to evaluate the robustness of the results and primary conclusions of the trial. Robustness is a concept that refers to the sensitivity of the overall conclusions to various limitations of the data, assumptions, and analytic approaches to data analysis. Robustness implies that the treatment effect and primary conclusions of the trial are not substantially affected when analyses are carried out based on alternative assumptions or analytic approaches. The interpretation of statistical measures of uncertainty of the treatment effect and treatment comparisons should involve consideration of the potential contribution of bias to the p-value, confidence interval, or inference. Because the predominant approaches to the design and analysis of clinical trials have been based on frequentist statistical methods, the guidance largely refers to the use of frequentist methods when discussing hypothesis testing and/or confidence intervals. This should not be taken to imply that other approaches are not appropriate: the use of Bayesian and other approaches may be considered when the reasons for their use are clear and when the resulting conclusions are sufficiently robust.

CONSIDERATIONS FOR OVERALL CLINICAL DEVELOPMENT

Trial Context

Development plan

The broad aim of the process of clinical development of a new drug is to find out whether there is a dose range and schedule at which the drug can be shown to be simultaneously safe and effective, to the extent that the risk-benefit relationship is acceptable. The particular subjects, who may benefit from the drug, and the specific indications for its use, also need to be defined. Satisfying these broad aims usually requires an ordered programme of clinical trials, each with its

own specific objectives (see ICH E8). This should be specified in a clinical plan, or a series of plans, with appropriate decision points and flexibility to allow modification as knowledge accumulates. A marketing application should clearly describe the main content of such plans, and the contribution made by each trial. Interpretation and assessment of the evidence from the total programme of trials involves synthesis of the evidence from the individual trials. This is facilitated by ensuring that common standards are adopted for a number of features of the trials such as dictionaries of medical terms, definition and timing of the main measurements, handling of protocol deviations and so on. A statistical summary, overview or meta-analysis may be informative when medical questions are addressed in more than one trial. Where possible this should be envisaged in the plan so that the relevant trials are clearly identified and any necessary common features of their designs are specified in advance.[4] Other major statistical issues (if any) that are expected to affect a number of trials in a common plan should be addressed in that plan.

Confirmatory trial

A confirmatory trial is an adequately controlled trial in which the hypotheses are stated in advance and evaluated. As a rule, confirmatory trials are necessary to provide firm evidence of efficacy or safety. In such trials the key hypothesis of interest follows directly from the trial's primary objective, is always pre-defined, and is the hypothesis that is subsequently tested when the trial is complete. In a confirmatory trial it is equally important to estimate with due precision the size of the effects attributable to the treatment of interest and to relate these effects to their clinical significance. Confirmatory trials are intended to provide firm evidence in support of claims and hence adherence to protocols and standard operating procedures is particularly important; unavoidable changes should be explained and documented, and their effect examined. A justification of the design of each such trial, and of other important statistical aspects such as the principal features of the planned analysis, should be set out in the protocol. Each trial should address only a limited number of questions. Firm evidence in support of claims requires that the results of the confirmatory trials demonstrate that the investigational product under test has clinical benefits. The confirmatory trials should therefore be sufficient to answer each key clinical question relevant to the efficacy or safety claim clearly and definitively. In addition, it is important that the basis for generalisation to the intended patient population is understood and explained; this may also influence the number and type (e.g. specialist or general practitioner) of centres and/or trials needed. The results of the confirmatory trial(s) should be robust. In some circumstances the weight of evidence from a single confirmatory trial may be sufficient.[5]

Exploratory trial

The rationale and design of confirmatory trials nearly always rests on earlier clinical work carried out in a series of exploratory studies. Like all clinical trials, these exploratory studies should have clear and precise objectives. However, in contrast to confirmatory trials, their objectives may not always lead to simple tests of pre-defined hypotheses. In addition, exploratory trials may sometimes require a more flexible approach to design so that changes can be made in response to accumulating results. Their analysis may entail data exploration; tests of hypothesis may be carried out, but the choice of hypothesis may be data dependent. Such trials cannot be the basis of the formal proof of efficacy, although they may contribute to the total body of relevant evidence[6].

Any individual trial may have both confirmatory and exploratory aspects. For example, in most confirmatory trials the data are also subjected to exploratory analyses which serve as a basis for explaining or supporting their findings and for suggesting further hypotheses for later research. The protocol should make a clear distinction between the aspects of a trial which will be used for confirmatory proof and the aspects which will provide data for exploratory analysis.

Scope of Trials

Population

In the earlier phases of drug development the choice of subjects for a clinical trial may be

heavily influenced by the wish to maximise the chance of observing specific clinical effects of interest, and hence they may come from a very narrow subgroup of the total patient population for which the drug may eventually be indicated. However by the time the confirmatory trials are undertaken, the subjects in the trials should more closely mirror the target population. Hence, in these trials it is generally helpful to relax the inclusion and exclusion criteria as much as possible within the target population, while maintaining sufficient homogeneity to permit precise estimation of treatment effects. No individual clinical trial can be expected to be totally representative of future users, because of the possible influences of geographical location, the time when it is conducted, the medical practices of the particular investigator(s) and clinics, and so on. However the influence of such factors should be reduced wherever possible, and subsequently discussed during the interpretation of the trial results.[7]

Primary and secondary variables

The primary variable ('target' variable, primary endpoint) should be the variable capable of providing the most clinically relevant and convincing evidence directly related to the primary objective of the trial. There should generally be only one primary variable. This will usually be an efficacy variable, because the primary objective of most confirmatory trials is to provide strong scientific evidence regarding efficacy. Safety/tolerability may sometimes be the primary variable, and will always be an important consideration. Measurements relating to quality of life and health economics are further potential primary variables. The selection of the primary variable should reflect the accepted norms and standards in the relevant field of research. The use of a reliable and validated variable with which experience has been gained either in earlier studies or in published literature is recommended. There should be sufficient evidence that the primary variable can provide a valid and reliable measure of some clinically relevant and important treatment benefit in the patient population described by the inclusion and exclusion criteria.[8] The primary variable should generally be the one used when estimating the sample size.

In many cases, the approach to assessing subject outcome may not be straightforward and should be carefully defined. For example, it is inadequate to specify mortality as a primary variable without further clarification; mortality may be assessed by comparing proportions alive at fixed points in time, or by comparing overall distributions of survival times over a specified interval. Another common example is a recurring event; the measure of treatment effect may again be a simple dichotomous variable (any occurrence during a specified interval), time to first occurrence, rate of occurrence (events per time units of observation), etc. The assessment of functional status over time in studying treatment for chronic disease presents other challenges in selection of the primary variable. There are many possible approaches, such as comparisons of the assessments done at the beginning and end of the interval of observation, comparisons of slopes calculated from all assessments throughout the interval, comparisons of the proportions of subjects exceeding or declining beyond a specified threshold, or comparisons based on methods for repeated measures data. To avoid multiplicity concerns arising from post hoc definitions, it is critical to specify in the protocol the precise definition of the primary variable as it will be used in the statistical analysis. In addition, the clinical relevance of the specific primary variable selected and the validity of the associated measurement procedures will generally need to be addressed and justified in the protocol. The primary variable should be specified in the protocol, along with the rationale for its selection. Redefinition of the primary variable after unblinding will almost always be unacceptable, since the biases this introduces are difficult to assess. When the clinical effect defined by the primary objective is to be measured in more than one way, the protocol should identify one of the measurements as the primary variable on the basis of clinical relevance, importance, objectivity, and/or other relevant characteristics, whenever such selection is feasible.

Secondary variables are either supportive measurements related to the primary objective or measurements of effects related to the secondary objectives. Their pre-definition in the protocol is also important, as well as an explanation of their relative importance and roles in interpretation of trial results. The number of secondary variables should be limited and should be related to the limited number of questions to be answered in the trial.[9]

Composite variables

If a single primary variable cannot be selected from multiple measurements associated with the primary objective, another useful strategy is to integrate or combine the multiple measurements into a single or "composite" variable, using a pre-defined algorithm. Indeed, the primary variable sometimes arises as a combination of multiple clinical measurements (e.g. the rating scales used in arthritis, psychiatric disorders and elsewhere). This approach addresses the multiplicity problem without requiring adjustment to the type I error. The method of combining the multiple measurements should be specified in the protocol, and an interpretation of the resulting scale should be provided in terms of the size of a clinically relevant benefit. When a composite variable is used as a primary variable, the components of this variable may sometimes be analysed separately, where clinically meaningful and validated. When a rating scale is used as a primary variable, it is especially important to address such factors as content validity, inter- and intra-rater reliability and responsiveness for detecting changes in the severity of disease.[10]

Global assessment variables

In some cases, 'global assessment' variables (see Glossary) are developed to measure the overall safety, overall efficacy, and/or overall usefulness of a treatment. This type of variable integrates objective variables and the investigator's overall impression about the state or change in the state of the subject, and is usually a scale of ordered categorical ratings. Global assessments of overall efficacy are well established in some therapeutic areas, such as neurology and psychiatry. Global assessment variables generally have a subjective component. When a global assessment variable is used as a primary or secondary variable, fuller details of the scale should be included in the protocol with respect to:

- The relevance of the scale to the primary objective of the trial;
- The basis for the validity and reliability of the scale;
- How to utilise the data collected on an individual subject to assign him/her to a unique category of the scale;
- How to assign subjects with missing data to a unique category of the scale, or otherwise evaluate them.

If objective variables are considered by the investigator when making a global assessment, then those objective variables should be considered as additional primary, or at least important secondary, variables. Global assessment of usefulness integrates components of both benefit and risk and reflects the decision making process of the treating physician, who must weigh benefit and risk in making product use decisions. A problem with global usefulness variables is that their use could in some cases lead to the result of two products being declared equivalent despite having very different profiles of beneficial and adverse effects. For example, judging the global usefulness of a treatment as equivalent or superior to an alternative may mask the fact that it has little or no efficacy but fewer adverse effects[11]. Therefore it is not advisable to use a global usefulness variable as a primary variable. If global usefulness is specified as primary, it is important to consider specific efficacy and safety outcomes separately as additional primary variables.

Multiple primary variables

It may sometimes be desirable to use more than one primary variable, each of which (or a subset of which) could be sufficient to cover the range of effects of the therapies. The planned manner of interpretation of this type of evidence should be carefully spelled out. It should be clear whether an impact on any of the variables, some minimum number of them, or all of them, would be considered necessary to achieve the trial

objectives. The primary hypothesis or hypotheses and parameters of interest (e.g. mean, percentage, and distribution) should be clearly stated with respect to the primary variables identified, and the approach to statistical inference described. The effect on the type I error should be explained because of the potential for multiplicity problems (see above section); the method of controlling type I error should be given in the protocol. The extent of intercorrelation among the proposed primary variables may be considered in evaluating the impact on type I error. If the purpose of the trial is to demonstrate effects on all of the designated primary variables, then there is no need for adjustment of the type I error, but the impact on type II error and sample size should be carefully considered.[1]

Surrogate variables

When direct assessment of the clinical benefit to the subject through observing actual clinical efficacy is not practical, indirect criteria (surrogate variables - see Glossary) may be considered. Commonly accepted surrogate variables are used in a number of indications where they are believed to be reliable predictors of clinical benefit. There are two principal concerns with the introduction of any proposed surrogate variable. First, it may not be a true predictor of the clinical outcome of interest. For example it may measure treatment activity associated with one specific pharmacological mechanism, but may not provide full information on the range of actions and ultimate effects of the treatment, whether positive or negative. There have been many instances where treatments showing a highly positive effect on a proposed surrogate have ultimately been shown to be detrimental to the subjects' clinical outcome; conversely, there are cases of treatments conferring clinical benefit without measurable impact on proposed surrogates. Secondly, proposed surrogate variables may not yield a quantitative measure of clinical benefit that can be weighed directly against adverse effects.[13] Statistical criteria for validating surrogate variables have been proposed but the experience with their use is relatively limited. In practice, the strength of the evidence for surrogacy depends upon (i) the biological plausibility of the relationship, (ii) the demonstration in epidemiological studies of the prognostic value of the surrogate for the clinical outcome and (iii) evidence from clinical trials that treatment effects on the surrogate correspond to effects on the clinical outcome. Relationships between clinical and surrogate variables for one product do not necessarily apply to a product with a different mode of action for treating the same disease.

Categorised variables

Dichotomisation or other categorisation of continuous or ordinal variables may sometimes be desirable. Criteria of 'success' and 'response' are common examples of dichotomies which require precise specification in terms of, for example, a minimum percentage improvement (relative to baseline) in a continuous variable, or a ranking categorised as at or above some threshold level (e.g., 'good') on an ordinal rating scale. The reduction of diastolic blood pressure below 90mmHg is a common dichotomisation. Categorisations are most useful when they have clear clinical relevance. The criteria for categorisation should be pre-defined and specified in the protocol, as knowledge of trial results could easily bias the choice of such criteria. Because categorisation normally implies a loss of information, a consequence will be a loss of power in the analysis; this should be accounted for in the sample size calculation.[14]

Design Techniques to Avoid Bias

The most important design techniques for avoiding bias in clinical trials are blinding and randomisation, and these should be normal features of most controlled clinical trials intended to be included in a marketing application. Most such trials follow a double-blind approach in which treatments are pre-packed in accordance with a suitable randomisation schedule, and supplied to the trial centre(s) labelled only with the subject number and the treatment period so that no one involved in the conduct of the trial is aware of the specific treatment allocated to any particular subject, not even as a code letter[15]. This approach will be assumed in above section and

most of above section, exceptions being considered at the end. Bias can also be reduced at the design stage by specifying procedures in the protocol aimed at minimising any anticipated irregularities in trial conduct that might impair a satisfactory analysis, including various types of protocol violations, withdrawals and missing values. The protocol should consider ways both to reduce the frequency of such problems, and also to handle the problems that do occur in the analysis of data.

Blinding

Blinding or masking is intended to limit the occurrence of conscious and unconscious bias in the conduct and interpretation of a clinical trial arising from the influence which the knowledge of treatment may have on the recruitment and allocation of subjects, their subsequent care, the attitudes of subjects to the treatments, the assessment of end-points, the handling of withdrawals, the exclusion of data from analysis, and so on. The essential aim is to prevent identification of the treatments until all such opportunities for bias have passed. A double-blind trial is one in which neither the subject nor any of the investigator or sponsor staff who are involved in the treatment or clinical evaluation of the subjects are aware of the treatment received. This includes anyone determining subject eligibility, evaluating endpoints, or assessing compliance with the protocol. This level of blinding is maintained throughout the conduct of the trial, and only when the data are cleaned to an acceptable level of quality will appropriate personnel be unblinded. If any of the sponsor staff who are not involved in the treatment or clinical evaluation of the subjects are required to be unblinded to the treatment code (e.g. bioanalytical scientists, auditors, those involved in serious adverse event reporting), the sponsor should have adequate standard operating procedures to guard against inappropriate dissemination of treatment codes. In a single-blind trial the investigator and/or his staff are aware of the treatment but the subject is not, or vice versa. In an open-label trial the identity of treatment is known to all. The double-blind trial is the optimal approach. This requires that the treatments to be applied during the trial cannot be distinguished (appearance, taste, etc.) either before or during administration, and that the blind is maintained appropriately during the whole trial.[16]

Randomisation

Randomisation introduces a deliberate element of chance into the assignment of treatments to subjects in a clinical trial. During subsequent analysis of the trial data, it provides a sound statistical basis for the quantitative evaluation of the evidence relating to treatment effects. It also tends to produce treatment groups in which the distributions of prognostic factors, known and unknown, are similar. In combination with blinding, randomisation helps to avoid possible bias in the selection and allocation of subjects arising from the predictability of treatment assignments. The randomisation schedule of a clinical trial documents the random allocation of treatments to subjects. In the simplest situation it is a sequential list of treatments (or treatment sequences in a crossover trial) or corresponding codes by subject number. The logistics of some trials, such as those with a screening phase, may make matters more complicated, but the unique pre-planned assignment of treatment, or treatment sequence, to subject should be clear. Different trial designs will require different procedures for generating randomisation schedules. The randomisation schedule should be reproducible (if the need arises).[17] Although unrestricted randomisation is an acceptable approach, some advantages can generally be gained by randomising subjects in blocks. This helps to increase the comparability of the treatment groups, particularly when subject characteristics may change over time, as a result, for example, of changes in recruitment policy. It also provides a better guarantee that the treatment groups will be of nearly equal size. In crossover trials it provides the means of obtaining balanced designs with their greater efficiency and easier interpretation.[18] Care should be taken to choose block lengths that are sufficiently short to limit possible imbalance, but that are long enough to avoid predictability towards the end of the sequence in a block.

Investigators and other relevant staff should generally be blind to the block length; the use of two or more block lengths, randomly selected for each block, can achieve the same purpose. (Theoretically, in a double-blind trial predictability does not matter, but the pharmacological effects of drugs may provide the opportunity for intelligent guesswork.)

TRIAL DESIGN CONSIDERATIONS

Design Configuration

Parallel group design

The most common clinical trial design for confirmatory trials is the parallel group design in which subjects are randomised to one of two or more arms, each arm being allocated a different treatment. These treatments will include the investigational product at one or more doses, and one or more control treatments, such as placebo and/or an active comparator. The assumptions underlying this design are less complex than for most other designs. However, as with other designs, there may be additional features of the trial that complicate the analysis and interpretation (e.g. covariates, repeated measurements over time, and interactions between design factors, protocol violations, dropouts.[19]

Crossover design

In the crossover design, each subject is randomised to a sequence of two or more treatments, and hence acts as his own control for treatment comparisons. This simple manoeuvre is attractive primarily because it reduces the number of subjects and usually the number of assessments needed to achieve a specific power, sometimes to a marked extent. In the simplest 2×2 crossover design each subject receives each of two treatments in randomised order in two successive treatment periods, often separated by a washout period. The most common extension of this entails comparing n(>2) treatments in n periods, each subject receiving all n treatments. Numerous variations exist, such as designs in which each subject receives a subset of n(>2) treatments, or ones in which treatments are repeated within a subject[20].

Crossover designs have a number of problems that can invalidate their results. The chief difficulty concerns carryover, that is, the residual influence of treatments in subsequent treatment periods. In an additive model the effect of unequal carryover will be to bias direct treatment comparisons. In the 2×2 design the carryover effect cannot be statistically distinguished from the interaction between treatment and period and the test for either of these effects lacks power because the corresponding contrast is 'between subject'. This problem is less acute in higher order designs, but cannot be entirely dismissed[21].

Multicentre Trials

Multicentre trials are carried out for two main reasons. Firstly, a multicentre trial is an accepted way of evaluating a new medication more efficiently; under some circumstances, it may present the only practical means of accruing sufficient subjects to satisfy the trial objective within a reasonable time-frame. Multicentre trials of this nature may, in principle, be carried out at any stage of clinical development. They may have several centres with a large number of subjects per centre or, in the case of a rare disease, they may have a large number of centres with very few subjects per centre.[22]

Trials to Show Dose-response Relationship

How response is related to the dose of a new investigational product is a question to which answers may be obtained in all phases of development, and by a variety of approaches (see ICH E4). Dose-response trials may serve a number of objectives, amongst which the following are of particular importance: the confirmation of efficacy; the investigation of the shape and location of the dose-response curve; the estimation of an appropriate starting dose; the identification of optimal strategies for individual dose adjustments; the determination of a maximal dose beyond which additional benefit would be unlikely to occur. These objectives should be addressed using the data collected at a number of doses under investigation, including a placebo (zero dose) wherever appropriate.[23] For this purpose the application of procedures to estimate the relationship between dose and response,

including the construction of confidence intervals and the use of graphical methods, is as important as the use of statistical tests. The hypothesis tests that are used may need to be tailored to the natural ordering of doses or to particular questions regarding the shape of the dose-response curve (e.g. monotonicity). The details of the planned statistical procedures should be given in the protocol.

Group Sequential Designs

Group sequential designs are used to facilitate the conduct of interim analysis. While group sequential designs are not the only acceptable types of designs permitting interim analysis, they are the most commonly applied because it is more practicable to assess grouped subject outcomes at periodic intervals during the trial than on a continuous basis as data from each subject become available. The statistical methods should be fully specified in advance of the availability of information on treatment outcomes and subject treatment assignments (i.e. blind breaking. An Independent Data Monitoring Committee (see Glossary) may be used to review or to conduct the interim analysis of data arising from a group sequential design. While the design has been most widely and successfully used in large, long-term trials of mortality or major nonfatal endpoints, its use is growing in other circumstances. In particular, it is recognised that safety must be monitored in all trials and therefore the need for formal procedures to cover early stopping for safety reasons should always be considered.[24]

Data Capture and Processing

The collection of data and transfer of data from the investigator to the sponsor can take place through a variety of media, including paper case record forms, remote site monitoring systems, medical computer systems and electronic transfer. Whatever data capture instrument is used, the form and content of the information collected should be in full accordance with the protocol and should be established in advance of the conduct of the clinical trial. It should focus on the data necessary to implement the planned analysis, including the context information (such as timing assessments relative to dosing) necessary to confirm protocol compliance or identify important protocol deviations. 'Missing values' should be distinguishable from the 'value zero' or 'characteristic absent'. The process of data capture through to database finalisation should be carried out in accordance with GCP (see ICH E6). Specifically, timely and reliable processes for recording data and rectifying errors and omissions are necessary to ensure delivery of a quality database and the achievement of the trial objectives through the implementation of the planned analysis.[25]

TRIAL MONITORING AND INTERIM ANALYSIS

Careful conduct of a clinical trial according to the protocol has a major impact on the credibility of the results (see ICH E6). Careful monitoring can ensure that difficulties are noticed early and their occurrence or recurrence minimised. There are two distinct types of monitoring that generally characterise confirmatory clinical trials sponsored by the pharmaceutical industry. One type of monitoring concerns the oversight of the quality of the trial, while the other type involves breaking the blind to make treatment comparisons (i.e. interim analysis).[26] Both types of trial monitoring, in addition to entailing different staff responsibilities, involve access to different types of trial data and information, and thus different principles apply for the control of potential statistical and operational bias.

EVALUATION OF SAFETY AND TOLERABILITY

Scope of Evaluation

In all clinical trials evaluation of safety and tolerability constitutes an important element. In early phases this evaluation is mostly of an exploratory nature, and is only sensitive to frank expressions of toxicity, whereas in later phases the establishment of the safety and tolerability profile of a drug can be characterised more fully in larger samples of subjects. Later phase controlled trials represent an important means of exploring in an unbiased manner any new potential adverse effects, even if such trials generally lack power in this respect. Certain trials may be designed with the purpose of making

specific claims about superiority or equivalence with regard to safety and tolerability compared to another drug or to another dose of the investigational drug.[27] Such specific claims should be supported by relevant evidence from confirmatory trials, similar to that necessary for corresponding efficacy claims.

Choice of Variables and Data Collection

In any clinical trial the methods and measurements chosen to evaluate the safety and tolerability of a drug will depend on a number of factors, including knowledge of the adverse effects of closely related drugs, information from non-clinical and earlier clinical trials and possible consequences of the pharmacodynamic/pharma-cokinetic properties of the particular drug, the mode of administration, the type of subjects to be studied, and the duration of the trial. Laboratory tests concerning clinical chemistry and haematology, vital signs, and clinical adverse events (diseases, signs and symptoms) usually form the main body of the safety and tolerability data. The occurrence of serious adverse events and treatment discontinuations due to adverse events are particularly important to register (see ICH E2A and ICH E3). Furthermore, it is recommended that a consistent methodology be used for the data collection and evaluation throughout a clinical trial program in order to facilitate the combining of data from different trials. The use of a common adverse event dictionary is particularly important. This dictionary has a structure which gives the possibility to summarise the adverse event data on three different levels; system-organ class, preferred term or included term. The preferred term is the level on which adverse events usually are summarised, and preferred terms belonging to the same system-organ class could then be brought together in the descriptive presentation of data (see ICH M1).

Set of Subjects to be Evaluated and Presentation of Data

For the overall safety and tolerability assessment, the set of subjects to be summarised is usually defined as those subjects who received at least one dose of the investigational drug. Safety and tolerability variables should be collected as comprehensively as possible from these subjects, including type of adverse event, severity, onset and duration (see ICH E2B). Additional safety and tolerability evaluations may be needed in specific subpopulations, such as females, the elderly (see ICH E7), the severely ill, or those who have a common concomitant treatment. These evaluations may need to address more specific issues (see ICH E3).

All safety and tolerability variables will need attention during evaluation, and the broad approach should be indicated in the protocol. All adverse events should be reported, whether or not they are considered to be related to treatment. All available data in the study population should be accounted for in the evaluation. Definitions of measurement units and reference ranges of laboratory variables should be made with care; if different units or different reference ranges appear in the same trial (e.g. if more than one laboratory is involved), then measurements should be appropriately standardised to allow a unified evaluation. Use of a toxicity grading scale should be prespecified and justified. The incidence of a certain adverse event is usually expressed in the form of a proportion relating number of subjects experiencing events to number of subjects at risk. However, it is not always self-evident how to assess incidence. For example, depending on the situation the number of exposed subjects or the extent of exposure (in person-years) could be considered for the denominator. Whether the purpose of the calculation is to estimate a risk or to make a comparison between treatment groups it is important that the definition is given in the protocol. This is especially important if long-term treatment is planned and a substantial proportion of treatment withdrawals or deaths are expected. For such situations survival analysis methods should be considered and cumulative adverse event rates calculated in order to avoid the risk of underestimation.[28]

Statistical Evaluation

The investigation of safety and tolerability is a multidimensional problem. Although some

specific adverse effects can usually be anticipated and specifically monitored for any drug, the range of possible adverse effects is very large, and new and unforeseeable effects are always possible. Further, an adverse event experienced after a protocol violation, such as use of an excluded medication, may introduce a bias. This background underlies the statistical difficulties associated with the analytical evaluation of safety and tolerability of drugs, and means that conclusive information from confirmatory clinical trials is the exception rather than the rule. In most trials the safety and tolerability implications are best addressed by applying descriptive statistical methods to the data, supplemented by calculation of confidence intervals wherever this aids interpretation. It is also valuable to make use of graphical presentations in which patterns of adverse events are displayed both within treatment groups and within subjects. The calculation of p-values is sometimes useful either as an aid to evaluating a specific difference of interest, or as a 'flagging' device applied to a large number of safety and tolerability variables to highlight differences worth further attention. This is particularly useful for laboratory data, which otherwise can be difficult to summarise appropriately. It is recommended that laboratory data be subjected to both a quantitative analysis, e.g. evaluation of treatment means, and a qualitative analysis where counting of numbers above or below certain thresholds are calculated.[29] If hypothesis tests are used, statistical adjustments for multiplicity to quantify the type I error are appropriate, but the type II error is usually of more concern. Care should be taken when interpreting putative statistically significant findings when there is no multiplicity adjustment.

In the majority of trials investigators are seeking to establish that there are no clinically unacceptable differences in safety and tolerability compared with either a comparator drug or a placebo. As is the case for non-inferiority or equivalence evaluation of efficacy the use of confidence intervals is preferred to hypothesis testing in this situation. In this way, the considerable imprecision often arising from low frequencies of occurrence is clearly demonstrated.

Integrated Summary

The safety and tolerability properties of a drug are commonly summarised across trials continuously during an investigational product's development and in particular at the time of a marketing application. The usefulness of this summary, however, is dependent on adequate and well-controlled individual trials with high data quality. The overall usefulness of a drug is always a question of balance between risk and benefit and in a single trial such a perspective could also be considered, even if the assessment of risk/benefit usually is performed in the summary of the entire clinical trial program.[30]

REPORTING

Evaluation and Reporting

As stated in the Introduction, the structure and content of clinical study reports is the subject of ICH E3. That ICH guidance fully covers the reporting of statistical work, appropriately integrated with clinical and other material. The current section is therefore relatively brief. During the planning phase of a trial the principal features of the analysis should have been specified in the protocol as described in above section. When the conduct of the trial is over and the data are assembled and available for preliminary inspection, it is valuable to carry out the blind review of the planned analysis. This pre-analysis review, blinded to treatment, should cover decisions concerning, for example, the exclusion of subjects or data from the analysis sets; possible transformations may also be checked, and outliers defined; important covariates identified in other recent research may be added to the model; the use of parametric or non-parametric methods may be reconsidered. Decisions made at this time should be described in the report, and should be distinguished from those made after the statistician has had access to the treatment codes, as blind decisions will generally introduce less potential for bias. Statisticians or other staff involved in unblended interim analysis should not participate in the blind review or in making modifications to the

statistical analysis plan[31]. When the blinding is compromised by the possibility that treatment induced effects may be apparent in the data, special care will be needed for the blind review. Many of the more detailed aspects of presentation and tabulation should be finalised at or about the time of the blind review so that by the time of the actual analysis full plans exist for all its aspects including subject selection, data selection and modification, data summary and tabulation, estimation and hypothesis testing. Once data validation is complete, the analysis should proceed according to the pre-defined plans; the more these plans are adhered to, the greater the credibility of the results. Particular attention should be paid to any differences between the planned analysis and the actual analysis as described in the protocol, protocol amendments or the updated statistical analysis plan based on a blind review of data. A careful explanation should be provided for deviations from the planned analysis.

Summarising the Clinical Database

An overall summary and synthesis of the evidence on safety and efficacy from all the reported clinical trials is required for a marketing application (Expert report in EU, integrated summary reports in USA, Gaiyo in Japan). This may be accompanied, when appropriate, by a statistical combination of results. Within the summary a number of areas of specific statistical interest arise: describing the demography and clinical features of the population treated during the course of the clinical trial programme; addressing the key questions of efficacy by considering the results of the relevant (usually controlled) trials and highlighting the degree to which they reinforce or contradict each other; summarising the safety information available from the combined database of all the trials whose results contribute to the marketing application and identifying potential safety issues. During the design of a clinical programme careful attention should be paid to the uniform definition and collection of measurements which will facilitate subsequent interpretation of the series of trials, particularly if they are likely to be combined across trials.[32] A common dictionary for recording the details of medication, medical history and adverse events should be selected and used. A common definition of the primary and secondary variables is nearly always worthwhile, and essential for meta-analysis. The manner of measuring key efficacy variables, the timing of assessments relative to randomisation/entry, the handling of protocol violators and deviators and perhaps the definition of prognostic factors, should all be kept compatible unless there are valid reasons not to do so. Any statistical procedures used to combine data across trials should be described in detail. Attention should be paid to the possibility of bias associated with the selection of trials, to the homogeneity of their results, and to the proper modelling of the various sources of variation. The sensitivity of conclusions to the assumptions and selections made should be explored.[33]

Efficacy Data

Individual clinical trials should always be large enough to satisfy their objectives. Additional valuable information may also be gained by summarising a series of clinical trials which address essentially identical key efficacy questions. The main results of such a set of trials should be presented in an identical form to permit comparison, usually in tables or graphs which focus on estimates plus confidence limits. The use of meta-analytic techniques to combine these estimates is often a useful addition, because it allows a more precise overall estimate of the size of the treatment effects to be generated, and provides a complete and concise summary of the results of the trials. Under exceptional circumstances a meta analytic approach may also be the most appropriate way, or the only way, of providing sufficient overall evidence of efficacy via an overall hypothesis test. When used for this purpose the meta-analysis should have its own prospectively written protocol.[34]

Safety Data

In summarising safety data it is important to examine the safety database thoroughly for any indications of potential toxicity, and to follow up any indications by looking for an associated

supportive pattern of observations. The combination of the safety data from all human exposure to the drug provides an important source of information, because its larger sample size provides the best chance of detecting the rarer adverse events and, perhaps, of estimating their approximate incidence. However, incidence data from this database are difficult to evaluate because of the lack of a comparator group, and data from comparative trials are especially valuable in overcoming this difficulty. The results from trials which use a common comparator (placebo or specific active comparator) should be combined and presented separately for each comparator providing sufficient data.[35] All indications of potential toxicity arising from exploration of the data should be reported. The evaluation of the reality of these potential adverse effects should take account of the issue of multiplicity arising from the numerous comparisons made. The evaluation should also make appropriate use of survival analysis methods to exploit the potential relationship of the incidence of adverse events to duration of exposure and/or follow-up. The risks associated with identified adverse effects should be appropriately quantified to allow a proper assessment of the risk/benefit relationship.

CONCLUSION

Well designed and conducted investigation alone can prove or negate a hypothesis that is being tested. Statistical techniques cannot rectify mistakes due to careless or dishonest recording of data or faulty planning. The data should be collected honestly and sincerely without pre-conceived ideas about the outcome of interest.

REFERENCES

1. Altman, DG (1980), "***Designing Research: Practical Statistics for Medical Research***", Chapman and Hall, London, 74-103.
2. Armitage, P (1983), "***Statistical Methods in Medical Research***", Blackwell Scientific Publications, Oxford, 167-88.
3. Hennekens, CH and Buring, IE (1987), "***Design Strategies in Epidemiologic Research***", In: Mayrent SL (Ed), Epidemiology in medicine. Boston: Little, Brown and Co, 16-20.
4. Juneja, A; Agarwal, SS and Murthy, NS (1996), "***Designing of Case-Control Studies***", Rao V. (Ed), Biostatistics: A manual of statistical methods for use in health, nutrition and anthropology, Jaypee Brothers, New Delhi, 404-18.
5. Rao, NSN (1989), "***Elements of Health Statistics***", Tara Book Agency, Varanasi, 89-101.
6. (1988),"***International Committee of Medical Journal Editors***", Ursiform requirement for manuscript submitted to biomedical journals, 296,401-3
7. Curran-Everett, D and Benos, DJ (2007), "Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. Advan", ***Physiol. Edu.***,31, 295–298.
8. Curran-Everett, D and Benos, DJ (2007), "Guidelines for reporting statistics in journals published by the American Physiological Society", ***Am J Physiol Endocrinol Metab.***, 287, E189-191.
9. Greenhalgh, T (2001), "How to read a paper: The basics of evidence based medicined", ***BMJ Books***, London.
10. Yurdusev, AN (1993), "Level of Analysis" and "Unit of Analysis": A Case for Distinction", ***Millennium Journal of International Studies***, 22,77-88.
11. Khochbin, S; Chabanas, A; Albert, P and , JJ(1989), "Flow cytofluorimetric determination of protein distribution throughout the cell cycle", ***Cytometry***, 10,484-489.
12. Asssociation, WM (2008), "***WMA Declaration of Helsinki- Ethical Principles for Medical Research Involving Human Subjects***".
13. Concato, J; Shah, N and Horwitz, RI (2000), "Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs", ***N Engl J Med.***,342,1887-1892.

14. Moore, A; McQuay, H; Deery, S and Moore, M (2004), "***Number Needed to Treat (NNT)***", Bandolier: Evidence Based Thinking About Healthcare.
15. Eldridge, S; Kerry, S and Torgerson, DJ (2009), "Bias in identifying and recruiting participants in cluster randomised trials: what can be done?", BMJ,339, 4006–4010.
16. Von, Elm E; Altman, DG; Egger, M; Pocock, SJ; Götzsche, PC and Vandenbroucke, JP (2007), "The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies", ***Annals of Internal Medicine***, 147, 573-577.
17. Moher, D; Schulz, KF and Altman, DG (2001), "The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials", ***Lancet***, 357, 1191-1194.
18. (2010),"Consolidated Standards of Reporting Trials".
19. Kyrgidis, A; Triaridis, S; Constantinides, J and Antoniades, K (2009), "Current Research in Quality of Life in Head and Neck Cancer", In: Mandell S, *et al.*, editors, ***Current Research in Cancer,*** Research Media;, Trivandrum.
20. Kyrgidis, A Trivandrum: Printza, A (2009), "Evaluation of heartburn and hoarseness with the Reflux Symptom Index", ***Otolaryngology - Head and Neck Surgery***, 141,796.
21. (2006), "Trochim WMK", ***Research Methods Knowledge Base***.
22. Whitley, E and Ball, J(2002), "Statistics review 1: presenting and summarising data", ***Crit Care***,6,66-71.
23. Altman, DG (1997), "***Practical Statistics for Medical Research***", Chapman & Hall, London.
24. Kirkwood, BR (1997), "***Essentials of Medical Statistics***", Blackwell Science Ltd, London.
25. Whitley, E and Ball, J (2002), "Statistics review 2: Samples and populations", ***Critical Care***, 6,143-148.
26. Morton, JP (2009), "Reviewing scientific manuscripts: how much statistical knowledge should a reviewer really know?", ***Advan. Physiol.Edu.***,33,7-9.
27. Perel, P and Roberts, I (2007), "Colloids versus crystalloids for fluid resuscitation in critically ill patients", ***Cochrane Database Syst Rev.***, CD000567.
28. Whitley, E and Ball, J (2002), "Statistics review 3: hypothesis testing and P values", ***Crit Care***, 6,222–225
29. (2010), "Text Book SES, Power Analysis".
30. Kyrgidis, A; Verrou, E; Kitikidou, K; Andreadis, C; Katodritou, E and Vahtsevanos, K(2010), "***J Clin Oncol***", 28,e145-e147.
31. Rangachari, PK (2007), "Statistics: notaconfidencetrick. A commentary on "Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel", ***Advan. Physiol. Edu.***, 31,300-301.
32. Goodman, SN (1999), "Toward evidence-based medical statistics, 1: The P value fallacy", ***Ann Intern Med***., 130,995-1004.
33. Whitley, E and Ball, J (2002), "Statistics review 4: sample size calculations", ***Crit Care***, 6, 335-341.
34. Alderson, P (2004), "Absence of evidence is not evidence of absence", ***BMJ***, 328,476-477.
35. Bacchetti, P; Wolf, LE; Segal, MR and McCulloch, CE (2005), "Ethics and sample size", ***Am J Epidemiol.***, 161, 105-110.

Correspondence Author:

Abhishek Kumar Yadav*

Department of Pharmaceutical Sciences, Dibrugarh University, Dibrugarh-786004, Assam, India

Cite This Article: Abhishek, Kumar Yadav (2016), "Biostatistics & research methodology with an overview on clinical research", ***International Journal of Drug Research and Technology***, Vol. 6 (3), 193-208.

INTERNATIONAL JOURNAL OF DRUG RESEARCH AND TECHNOLOGY