

EXTRA READING

Lesson 5:

2.2. How to calculate the sample size for randomized controlled trials

Formulas for sample size calculation differ depending on the type of study design and the studies outcome(s). These calculations are particularly of interest in the design of randomized controlled trials (RCTs). In general, sample size calculations are performed based on the primary outcome of the study.

An example of how to calculate sample size using the simplest formulas for an RCT comparing two groups of equal size is given in the following.

Suppose one wished to study the effect of a new hypertensive medicine on systolic blood pressure (SBP) (measured in mmHg) as a continuous outcome.

The simplest formula for a continuous outcome and equal sample sizes in both groups, assuming: $\alpha = 0.05$ and power = 0.80 ($\beta = 0.20$, therefore $1-\beta=0.8$).

$$n = \frac{2[(a + b)^2 \sigma^2]}{(\mu_1 - \mu_2)^2}$$

n = the sample size in each of the groups

μ_1 = population mean in treatment Group 1

μ_2 = population mean in treatment Group 2

$\mu_1 - \mu_2$ = the difference the investigator wishes to detect

σ^2 = population variance (SD)

a = conventional multiplier for alpha* when alpha is 0.05

b = conventional multiplier for power* when beta is 0.80

When the significance level alpha is chosen at 0.05, one should enter the value 1.96 for a in the formula. Similarly, when beta is chosen at 0.20, the value 0.842 should be filled in for b in the formula.

Suppose the investigators consider a difference in SBP of 15 mmHg between the treated and the control group ($\mu_1 - \mu_2$) as clinically relevant and specified that such an effect should be detected with 80% power (0.80) and a significance level alpha of 0.05. Past experience with similar experiments, with similar measuring methods, and with similar subjects, suggests that the

data will be approximately normally distributed with an SD of 20 mmHg. Now we have all of the specifications needed for determining sample size using the approach as summarized in the formula above. Entering the values in the formula yields:

$$n = \frac{2[(1.96 + 0.842)^2 20^2]}{(15)^2} = 27.9$$

This means that a sample size of 28 subjects per group is needed to answer the research question.

*These values are looked up in a statistical table by the researchers. The table values are based on the normal distribution of these errors.

Lesson 8: Interpretation of Clinical Trial Data Results

Table of contents

1. Introduction
2. Bias
3. Where bias can be introduced
 - 3.1. During patient recruitment
 - 3.2. Information (or measurement) bias
 - 3.3. During trial conduct
 - 3.4. During analysis and interpretation of results
 - 3.5. At time of reporting
4. Correlation vs causation
5. Data tampering
6. Data transformation
7. Data merging
8. Assessing the value of a clinical trial

1. Introduction

Performing a clinical trial is a very complex and challenging activity. Bias may come in at different levels before, during and after the trial. Therefore, it is important for researchers to be able to interpret the trial results and to be able to identify potential bias in the design, conduct and analysis of a trial which could invalidate the trial analysis and ultimately the value of the clinical trial itself.

2. Bias

In research, bias occurs when systematic error is introduced into data sampling or hypothesis testing by selecting or encouraging one outcome or answer over others. Of note, bias is not always introduced intentionally. It can also be caused unintentionally for example by calibration error or unknown confounding variables. Bias may affect the results of a clinical trial by causing a deviation between the observed effect and its true value: estimates of association can be systematically larger or smaller than the true association. Bias may also take the form of systematic favoritism in the way results are reported or in the way they are interpreted in the discussion and conclusion on clinical trial results.

3. Where bias can be introduced

Bias can occur at any phase of research, including trial design, data collection, as

well as in the process of data analysis and publication. The bias that can occur at different stages during a clinical trial are described in the sections that follow.

3.1. During patient recruitment

Selection bias

Selection bias can occur at different levels. Some examples are given below:

1. If the patient eligibility criteria used in a trial are so stringent that only a small percentage of the intended patient population qualifies. For example, those with better physical condition who would be expected to be able to manage toxic treatments.
2. At treatment assignment, the eligibility criteria used to recruit and enroll patients into the treatment arms may be applied differently. This may result in some patient characteristics being over- expressed in one arm compared to the other. For example, if patients were selected differently according to their age or health status. The treatment outcomes may be stronger in the arm where the patients are younger and in a better shape. Therefore, any difference in outcome between the two treatment arms can no longer be attributed only to the received treatment.
3. The use of random patient allocation techniques to two or more treatment arms is key to avoiding this type of bias. Randomization aims to ensure that the treatment arms are comparable both in terms of known and unknown prognostic factors especially over a large number of patients. Well-performed patient randomization will allow the researcher to consider the observed treatment effects (response rate, survival, etc.) to be related to the treatment itself and not due to other factors. The treatment allocation should be independent from researchers' beliefs and expectations. The optimal way is by using a central randomization algorithm to balance patient characteristics over the treatment arms.

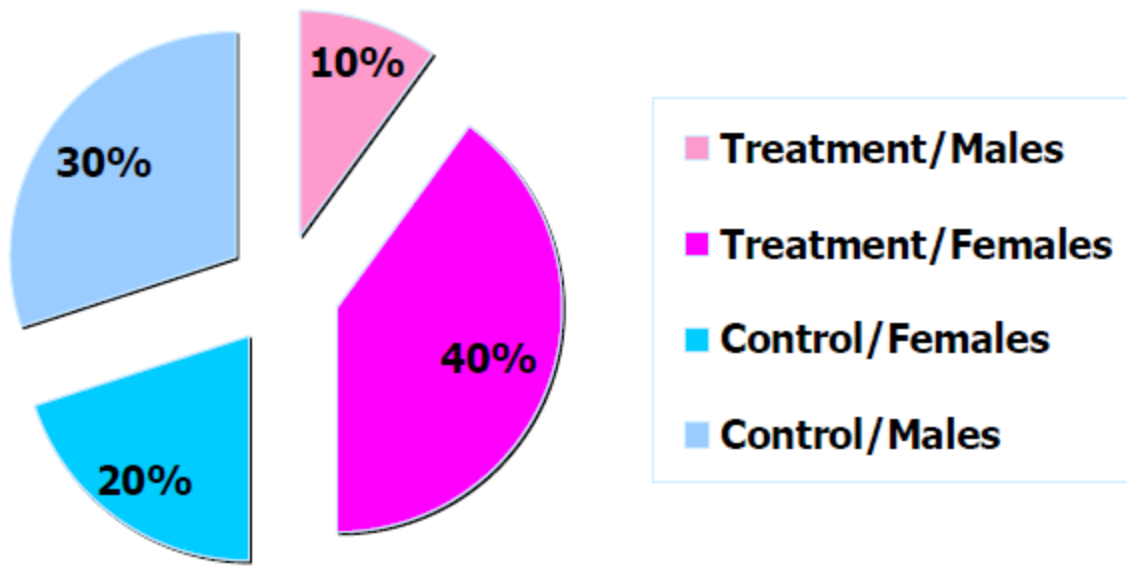


Figure 1: Example distribution of trial participants by arm and sex.

Note that in a trial of limited sample size, there is no guarantee that the randomization will effectively prevent imbalances in important prognostic factors (patient age or disease status). In this case, the randomization can be stratified according to a number of known important prognostic factors. The trial population will be split in groups according to these stratification factors and the randomization will ensure that the balance between the treatment arms within these strata is maintained.

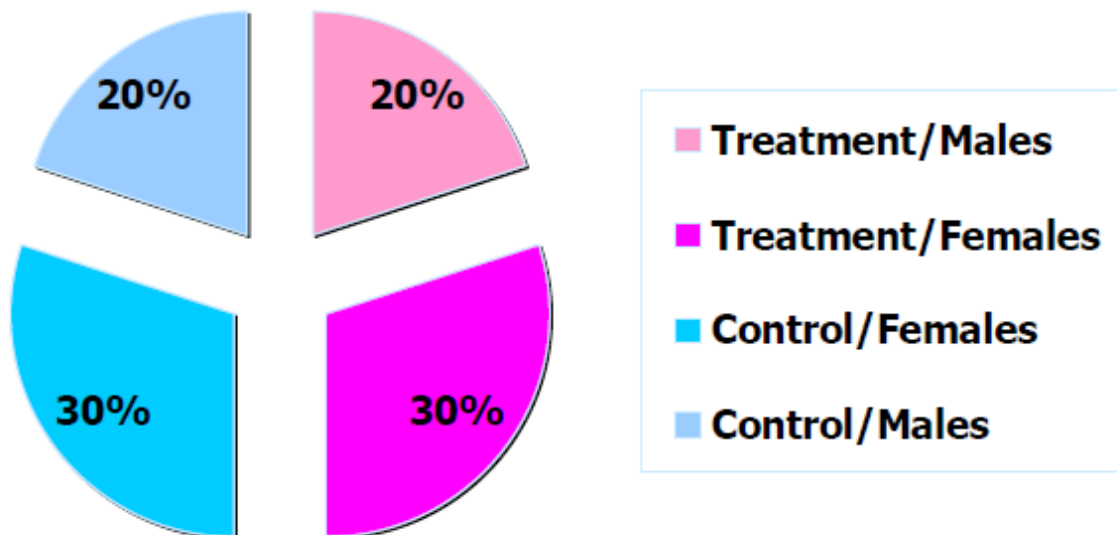


Figure 2: Example distribution of trial participants by arm and sex.

3.2. Information (or measurement) bias

This type of bias refers to a systematic error in the measurement of patient characteristics, exposure or treatment outcome, for example:

Patients may be wrongly classified as exposed to being at higher risk of disease when they are not.

The disease may be reported as progressing when it is stable.

This might be due to:

- Inaccuracies in the measurement tool.

- Expectations of trial participants - a patient could be more optimistic because they are assigned to the new treatment group.

Expectations of the investigators - also called the 'observer bias'. This occurs when:

- Investigators are very optimistic about the effect of the new treatment and interpret more favorably any clinical signal.

- Investigators monitor the adverse effects of the new treatment more carefully than for the standard treatment.

Blinding the allocated treatment to the patients and/or the investigators may prevent such bias. Blinding is of special interest when the trial outcome is subjective, like the reduction in pain, or when an experimental treatment is being compared to a placebo. However, while a blinded randomized trial is considered the gold standard of clinical trials, blinding may not always be feasible:

- Treatments may cause specific adverse effects that make them easy to identify.

- Treatments may need different procedures for administration or different treatment schedules.

3.3. During trial conduct

There are a few common problems that may arise during the course of a trial related to patient adherence (compliance) to the protocol and to the described treatment schedule. For example:

- Patients may turn out to be ineligible after randomization.

- Treatment may have been interrupted or modified but not according to the rules specified in the protocol.

- Disease assessments may have been delayed or not performed at all.

- A patient may decide to stop taking part in the trial, etc.

This can be problematic if not properly accounted for in the analysis. For example, consider the setting of a clinical trial comparing a new experimental treatment to the standard of

care. In this trial some patients taking the experimental treatment are too sick to go to the next visit within the allotted time due to side effects. A possible approach would be to include only patients with complete follow-up, so to exclude these sick patients from the analysis. However, by doing so, one selects a sub-group of patients whom, by definition, will present an artificially positive picture of the treatment under trial. This is again an example of selection bias.

One potential solution to this problem is a statistical concept called an intention-to-treat (ITT) analysis. ITT analysis includes every randomized patient whether they received the treatment or not. As such, ITT analyses maintain the balance of patients' baseline characteristics between the different trial arms obtained from the randomization. 'Protocol deviations' such as non-compliance to the assigned treatment (schedule, dosing, etc.) are part of daily practice. Therefore, treatment-effect estimates obtained from ITT analysis are considered to be more representative of the actual benefit of a new treatment in real life.

Bias can also happen when measuring the endpoint(s) of interest. For instance, time to disease progression or relapse can be severely affected by the hospital visit schedule. The frequency of disease assessment should be adequate and frequent enough to capture correctly the events that could not be observed by means other than by medical examination at the hospital. The disease assessment schedules should also be the same in both arms.

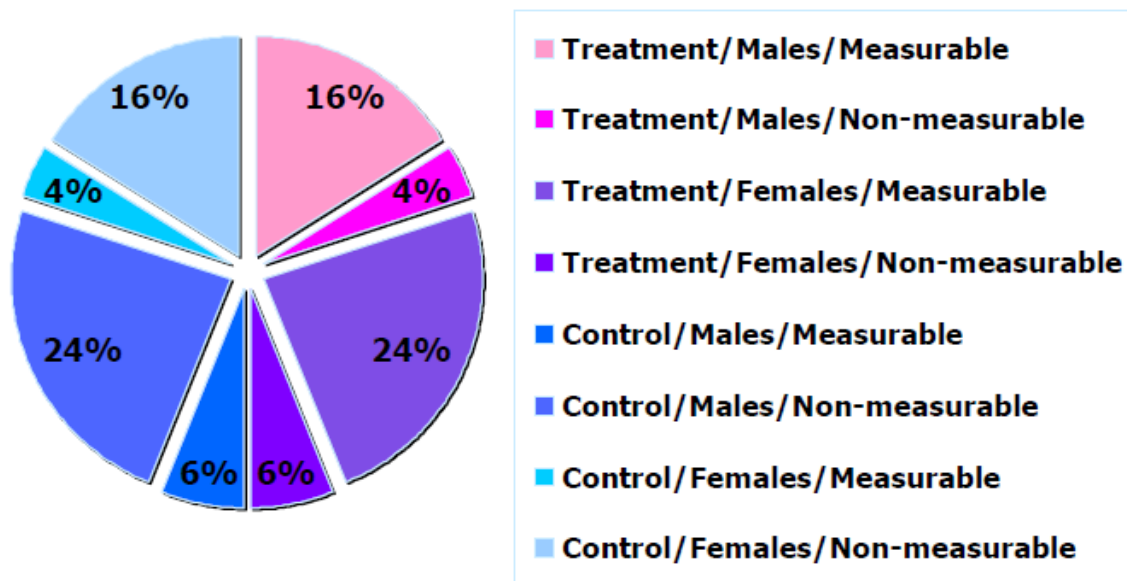


Figure 3: Example of sample distribution by arm and sex.

3.4. During analysis and interpretation of results

In a clinical trial sample, it is possible to find sub-groups of patients which respond better to treatment. Sub-group analyses involve splitting the trial participants into sub-groups. This could be based on:

Demographic characteristics (e.g. sex, age).

Baseline characteristics (e.g. a specific genomic profile).

Use of concomitant therapy.

Findings from sub-group analyses might be misleading for several reasons:

Firstly, sub-group analyses are observational (sub-groups are defined on observed patients' characteristics) and not based on randomized comparisons. The hindsight bias, also known as the 'I-knew-it-all-along'-bias, is the inclination to see events that have already occurred as being more predictable than they were before they took place.

Secondly, when multiple sub-group analyses are performed, the risk of finding a false positive result (i.e. a type I error) increases with the number of sub-group comparisons. Multiplicity issues are in general related to repeated looks at the same data set but in different ways until something 'statistically significant' emerges. With the wealth of data sometimes obtained, all signals should be considered carefully. Researchers must be cautious about possible over-interpretation. Techniques exist to protect against multiplicity, but they mostly require stronger evidence for statistical significance to control the overall type I error of the analysis (e.g. the Holm–Bonferroni method and the Hochberg procedure).

Thirdly, there is a tendency to conduct analyses comparing sub-groups based on information collected while on trial. A typical example is looking at the difference in survival between patients responding (yes/no) to treatment. Patients who are responding to treatment are by definition patients who are able to spend sufficient time on treatment to allow a response. Therefore, again by definition, they may simply represent a sub-group of patients of better prognosis and may therefore bias the analysis. This is an example of what is often referred to as lead-time bias or guarantee-time bias. One way of dealing with this is using a landmark as a starting point for the time-to-event analysis, and creating the categories based on the patients' characteristics at the time of this landmark (e.g. did a patient respond at three months, yes/no).

3.5. At time of reporting

Publication Bias

Publication bias refers to the fact that positive results are more likely to be published (rapidly) than negative results. Researchers may not be interested or less motivated in publishing negative trial results, fearing that such findings may negatively reflect on their professional abilities, the image of their company and perhaps their product. Hence positive results are more likely to be submitted rapidly for publication than negative results.

Publication bias is detrimental since it is preventing access to clinical trial results. Beyond single trial results, the efficacy and safety profile of a treatment needs to be assessed globally taking into account all the data and the results available from clinical trials investigating that treatment. Researchers planning new experimentation may be limited by the information available in published results. Negative results may inform about the lack of efficacy of a

treatment and the absence of justification for continuing with further development. The conclusions of a meta-analysis may be flawed if based only on published data.

Initiatives are ongoing for reducing the publication bias. One of them is to promote the registration of clinical trials for medicines. For instance, the International Committee of Medical Journal Editors (ICMJE) will not publish trials that are not registered in public registries such as clinicaltrials.gov, created and operated by the US National Institutes of Health (NIH). With such registries, researchers know what the existing clinical trials are, even if their results were never published, and may contact the trial sponsor in order to gain access to the results.

Despite these measures, publication bias has not been completely eliminated. While trial registries provide medical researchers with information about unpublished trials, researchers may be left to only speculate as to the results of these trials. Various organizations are currently engaged in initiatives to encourage or require the registration and disclosure of clinical trial information. In Europe, EudraCT, the European Clinical Trials Database of the European Medical Agency collects information on all clinical trials of medicines performed in Europe. As of July 2014, this database also makes trial summary results available to the public. The World Health Organization (WHO), through its International Clinical Trials Registry Platform (ICTRP), is setting international standards for registering and reporting on all clinical trials. In the US, the registry clinicaltrials.gov is doing similarly.

Reducing publication bias will also occur by increasing the willingness of scientific journals to publish trials with negative results. Manuscripts should be reviewed on the basis of the quality of the methodology used, not on the apparent success of the trial. In addition, funding agencies should take a more active role in the dissemination of clinical trials they fund.

4. Correlation vs causation

When analyzing the results from a trial it is important to remember that correlation is not the same thing as causation. Correlation is when two variables are linked in some way however this does not mean that one will cause the other. An example of this involves hormone replacement therapy (HRT) and coronary heart disease (CHD) where women taking HRT were at less risk from CHD. This however was not due to the actual HRT process but rather due to the fact that the group of people receiving HRT tended to belong to a higher socio-economic group, with better-than-average diet and exercise regime. This is why it is important to record as much information as possible about the subjects of trials.

5. Data tampering

Data tampering is the practice of selectively reporting data incorrectly or creating false results. An example of this would be when data that disagree with the expected result are discarded when there is a proportion of the results that would confirm the hypothesis. While it can be important to remove outliers from results it is important that those results are truly outliers and not just information that disagrees with expected or wanted results. Another example would be when a

data collector randomly generates a whole set of data out of a single measurement collected.

6. Data transformation

Data transformation recognized application of a formula to the data gained through a trial. This is often used to make the presentation of data clearer or easier to understand. For example, if measuring fuel efficiency for cars, it is natural to measure efficiency in the form of kilometers per litre. However, if you are assessing how much additional fuel would be required to increase the distance travelled it would be expressed as litres per kilometer. Applying an incorrect formula in this case, would affect the overall data.

7. Data merging

Data merging is the act of combining data from multiple experiments in order to gain a better understanding of the situation. One of the most common forms of this is meta-analysis where a person or group compares results from several different experiments whose results have been published. It is important whilst doing this to carefully check that the experiments are the same or comparable. Any differences need to be taken into account, so that there are hidden variables. An example might be the species of mice in an animal test.

8. Assessing the value of a clinical trial

Not all clinical trials are of equal validity. In deciding how much weight to give to the results of a trial, it is worth asking a few key questions:

1. How well designed was the trial? There is no one ‘correct’ design for a clinical trial – it is more a question of whether the design used was appropriate to the circumstances. While large trials are generally more reliable than small ones, this must be interpreted with common sense. For example, a trial of a rare inherited enzyme deficiency is never going to include the 5,000 patients which are frequently seen in trials for heart-attack medicines. A follow-up period of a few weeks is perfectly adequate for a pneumonia trial but would be inappropriate for a contraceptive pill. Placebo control groups, while very helpful in the interpretation of results, are clearly unethical in some situations (e.g. life-threatening illnesses for which effective treatments exist). While comparative trials are the best way of assessing efficacy, larger and longer open-label trials may offer more insight into real-life medicine safety.

Each trial design must be approached from the standpoint of ‘Was this the best way to do things in these circumstances?’

2. Does the population studied correspond to the one I am interested in? Information from a trial conducted in adults aged 18 to 65 may be of limited relevance to very elderly patients and will almost certainly be inadequate for guiding treatment of small babies. Similarly, people with severe or very advanced disease may respond quite differently to those with milder or earlier illness.

3. How relevant are the endpoints? Some diseases and symptoms lend themselves more readily than others to study in a clinical trial. If a new cancer medicine is increasing median survival by a year, there can be little doubt that this is a relevant measurement. A new painkiller used to treat the same patients will be much more difficult to assess because there are no clear 'standard units of pain'. Again, all that can be done is to ask whether the approach taken is appropriate to the circumstances.
4. Were the effects of the medicine clinically valuable? Generally, the bigger the effect of the medicine the better. All medicines come with costs in terms of both money and side effects. We are looking for the greatest possible benefit in return for those costs. It is worth remembering, however, that a result that is modest overall may be made up of a dramatic improvement in some patients and no change in others. If further research can help to identify the sub- group likely to do especially well, the new medicine may have much to offer this target population.
5. How do these results fit into the pattern of previous knowledge? It is very unusual for a clinical trial to stand alone as the only information available in a particular area of medicine. When this does happen, it usually represents the first use of a radically new approach to treatment and all one can do is to note the results with interest and wait to see whether subsequent trials support them. Much more commonly, there will have been previous trials with the same medicine or others of the same class in the same illness or in related diseases. One can then view the new results in the light of the previous body of knowledge. Findings that mesh well with what is already known are generally easier to accept than those, which directly contradict earlier results.

Lesson 9: Clinical Interpretation of Trial Results

Table of contents

1. Measurements in clinical trials
 - 1.1. Demographic and baseline information - who took part in the trial?
 - 1.2. Efficacy – how well did the trial medicine work?
 - 1.3. Safety – what side effects did the medicine have?
2. Important aspects of clinical trials
3. Recording clinical trial results
 - 3.1. European public assessment reports (EPARs)
 - 3.2. Clinical trial registries
 - 3.3. Marketing authorization product information
 - 3.4. Journal papers
 - 3.5. Conferences
 - 3.6. Patient organization websites
 - 3.7. Popular news media

1. Measurements in clinical trials

When clinical trials are conducted, medical details of participants (but not their identities) are collected for the purpose of statistical analyses in a computer database together with the results of any measurements made**. Statistical analyses are then conducted to formally assess the outcomes of the trial.

These analyses cover three areas of interest:

Descriptive statistics: Demographic and baseline information.

Inferential statistics: Efficacy.

Safety.

**In cases of emergency it must be possible to identify a participant (even after the completion of the trial).

1.1. Demographic and baseline information - who took part in the trial?

The effects of a medicine may differ considerably between different groups of people. It is therefore important to know details of the trial participants such as:

Age.

Sex.

Ethnic origin.

Severity of their illness.

In general, the closer the match between a trial group and a population of interest, the more relevant the findings will be.

1.2. Efficacy – how well did the trial medicine work?

Efficacy is often the main objective of the trial and is usually the aspect of highest interest. This part of the analysis is based on pre-defined ‘endpoints’. These are specific measurements related to the illness in question that have been specified in advance in the protocol (the document which describes in detail how the trial is going to be performed).

Endpoints in general can be categorized as:

‘Hard’ endpoints - those that take the form of numerical facts with intrinsic clinical importance. For example, how long the participant survived or what proportion of participants recovered from an infection.

‘Soft’ endpoints - those which are potentially influenced by the measurement process or with questionable reproducibility. For example, a quality-of-life questionnaire or the description of the participant’s mood at a given moment. In order to be analyzed statistically, soft endpoints have to be converted into a numerical format. This process can be controversial as it is subjective and potentially open to inconsistencies.

‘Surrogate’ endpoints - those that are not in themselves part of the patient’s experience of the illness but may be closely related to it. For example, the results of laboratory tests.

In general, hard endpoints are preferable to soft and surrogate endpoints. Soft and surrogate endpoints need to be assessed carefully in the light of how close or not they are linked to the actual illness.

Often, choosing which endpoints to use depends heavily on the nature of the illness being studied. Cancer offers obvious hard endpoints in the form of survival, whereas evaluation of depression must inevitably involve softer endpoints. Other illnesses, such as diabetes, are associated with well-established surrogate endpoints such as blood sugar levels.

1.3. Safety – what side effects did the medicine have?

Whenever a participant is seen by the doctor conducting a clinical trial, they are asked if anything undesired has happened. The ‘adverse event’ information collected in this way is analyzed to give an insight into possible side effects of the medicine. Particular attention is paid to ‘serious’ adverse events - those which are life-threatening or associated with death, hospitalization or birth abnormalities.

2. Important aspects of clinical trials

Clinical trials vary considerably in size, duration and design. These factors play a major part in the interpretation of trial results.

The most informative design is the ‘double-blind randomized comparison’ in which some patients receive the new medicine while others receive an alternative. The alternative treatment, sometimes called the ‘control’, may be either:

A placebo - inactive ‘dummy’ treatment.

An active comparator - generally consisting of a well-established treatment for the illness being studied.

Participants are allocated to each study group randomly, i.e. by chance. The trial is set up in such a way that while it is going on, neither the doctor nor the participant knows who is receiving which medicine, i.e. it is double-blind. This reduces the potential for bias in the results.

In such trials, the results are presented in terms of the difference between the group receiving the new medicine and the control group:

Where the comparison is against a placebo, this difference is a measure of the real effect of the new medicine.

Where the comparison is with an active comparator, the difference gives an insight into how the new medicine compares with current medical practice.

In both cases, two aspects of the difference are likely to be reported:

1. Size: This is often reported as a ‘point estimate’ (the actual difference recorded in this particular trial) together with a ‘95% confidence interval’. This is the range within which we can be 95% sure that the true difference would be represented in the population (all patients having the disease being studied). Although you may detect a statistical significance, it may not be clinically relevant. Generally speaking, the larger this difference, the more likely it is to be clinically relevant (to increase survival by a year is of more clinical relevance than to increase it by a day).

2. Statistical significance: Because some individuals respond better than others to treatment, there is always a risk that the difference between groups seen in a clinical trial may have arisen by chance. For example, if all the inherently good responders were randomized to one group and the bad responders to the other. Statisticians can calculate how likely it is for this scenario to have occurred in a particular clinical trial and they express their result as a ‘p-value’. This can be defined as the probability that a difference at least as large as the one observed could have arisen by chance if in reality there was no difference between the two treatments. A p-value of 0.05 means that there is a 5% or 1 in 20 chance of the difference happening by chance. It is conventionally taken as the threshold for accepting results as ‘statistically significant’. It is important to realize that the word ‘significant’ used in this sense says nothing about the clinical importance of the results – it merely offers reassurance that the result is unlikely to be accidental. For example, a one-meter increase in a six-minute walk distance might, in a large enough trial, be shown to be statistically significant (i.e. unlikely to have arisen by chance) but it would never be regarded by a heart-failure patient or his doctor as being of

any clinical value.

A second important group of clinical trials, often conducted to investigate long-term safety, takes the form of observational trials. In these there is no control group – everyone is treated with the new medicine and their experience is recorded. No differences between groups can arise (either accidentally or through genuine therapeutic effects) and hence there is no place for significance testing. Balanced against these shortcomings, open-label trials often include large numbers of patients (up to several thousand) studied for long periods of time (several years in some cases). These trials therefore make it easier to detect rare side effects and those that take a long time to develop.

The results of such trials list different adverse events and how frequently they were seen.

3. Recording clinical trial results

Following each clinical trial, the sponsor will compile a detailed clinical study report (CSR) which follows a format laid down by the Regulatory Authorities. It usually contains several hundred pages. Access to this report is usually limited to the sponsor themselves and the Regulatory Authorities assessing an application for marketing authorization.

Summarized information from the CSR is, however, likely to come into the public domain via a number of routes.

3.1. European public assessment reports (EPARs)

When a new medicine is approved centrally by the European Commission (EC), a technical assessment report is written by the Regulatory Authorities and is placed on the EMA website. This does not happen when the medicine is approved locally by one of the member states. In case of approval via the Mutual Recognition Procedure (MRP) or Decentralized Procedure (DCP) a similar public assessment report is published on the Heads of Medicines Agencies (HMA) website. The EPAR in the efficacy part contains a summarized version of the information in the corresponding CSR. The EPAR is intended for a professional audience and is therefore technical in language. It is, however, accompanied by a summary for the public, which presents the key facts in non-technical language.

3.2. Clinical trial registries

In Europe, EudraCT, the European Clinical Trials Database of the European Medical Agency collects information on all clinical trials of medicines performed in Europe. As of July 2014, this database also makes trial summary results available to the public. For trials taking place in the EU starting after January 1st, 2015, the results must be published whether negative or positive. The World Health Organization (WHO), through its International Clinical Trials Registry Platform (ICTRP), is setting international standards for registering and reporting on all clinical trials. In the US, the registry clinicaltrials.gov is taking a similar approach.

3.3. Marketing authorization product information

Although specific results of individual clinical trials are rarely presented in this way, an overall summary of the information available on a particular medicine is available from its summary of product characteristics (SmPC). This is a document aimed at healthcare professionals, however it forms the basis of the package leaflet (PL) (previously known as the patient information leaflet, (PIL)). This is aimed at the patient or user. These documents are normally available on the Internet depending on national regulations both from individual manufacturer's sites or from regulatory authorities or sites run by independent organizations.

3.4. Journal papers

The classic route for publication of clinical trial results is a research paper in a specialist medical journal. Virtually all modern journals have a peer-review process under which independent experts in the field review the manuscript and challenge any weak aspects of it before publication.

3.5. Conferences

A large number of international medical conferences are held every year, some with a quite general theme but many focus on a narrow specialist area. Clinical trial results are often presented at these conferences either as oral presentations or as posters displayed in public areas of the conference venue. Unfortunately, access to this information is often restricted to those who attend the conference and is not easily available to those who do not, unless the conference presentations are made available online. In many cases, however, the same trial will also be the subject of a journal paper.

3.6. Patient organization websites

Many patient organizations provide help to patients with specific illnesses and many have websites that publish reports of relevant clinical trials. Interpretation by experts working with the organization, together with the use of patient-friendly language, tend to make these reports particularly useful to patients.

3.7. Popular news media

The accuracy and understanding with which television, radio and newspaper accounts present the results of clinical trials varies a great deal. However, as a general rule it is wise to approach such reports with the understanding that a sensational story is more likely than a sober account to sell newspapers.

Lesson 10: Relevance of the Statistical Analysis Plan

Table of contents

1. Introduction
2. What is the statistical analysis plan (SAP)?
3. SAP contents
 - 3.1. Trial primary and secondary endpoints
 - 3.2. Survival curve
 - 3.3. Analysis methods
 - 3.4. (Primary) analysis set
 - 3.5. Sub-group analysis
 - 3.6. Interim analysis
4. Trial maturity
5. Unit summary

1. Introduction

Clinical trials are complex scientific experiments aiming to generate evidence regarding the safety and efficacy of treatments. Data generated by clinical trials are used for regulatory submissions and/or publications in scientific journals. It is therefore important that the way the data will be analyzed and presented are well defined upfront.

2. What is the statistical analysis plan (SAP)?

Several clinical trial documents are prepared to support the trial design, data collection, analysis and reporting. The statistical analysis plan (SAP) is one of these documents. The SAP provides details on the scope of planned analyses, population definitions and methodology. The SAP is crucial for guiding the data analyses and should therefore be created prior to the data analyses. The SAP is often integrated in the study protocol and must be submitted with the clinical trial application before the trial begins.

The SAP is submitted to Regulatory Authorities also as part of the submission package. The SAP is also an appendix of the clinical study report. The SAP is stored in the trial master file and it is used during audits to check if statistical analyses are performed as planned. The role of the SAP is explained in the International Council for Harmonization (ICH) E9 guideline 'Statistical principles for clinical trials'.

For each trial, one must specify upfront the planned statistical analysis in the SAP. This should include:

- Primary and secondary endpoints.
- Analysis methods.
- (Primary) analysis set.
- Pre-defined comparisons and significance levels.
- Exploratory analyses.
- Trial maturity.

3. SAP contents

The following are the main elements described in the SAP.

3.1. Trial primary and secondary endpoints

The trial endpoints can be of several types: Continuous, binary or time-to-event endpoints.

Continuous:

Is made up of measured data (e.g. height, weight).

Examples of continuous endpoints are blood pressure measurements or body temperature.

Binary:

Response Rate - how often a response occurs.

Absolute frequency - how many times a response occurs.

Relative frequency - for what percentage of participants does a response occur?

Time-to-event (TTE):

Time to progression (TTP): time from randomization to progression of disease.

Overall survival (OS): time from randomization to death from any cause.

‘Event’ free survival (EFS): time from randomization to either ‘event’ or death from any cause.

‘Event’ free interval (EFI): time from randomization to either ‘event’ or death related to the disease evolution or treatment.

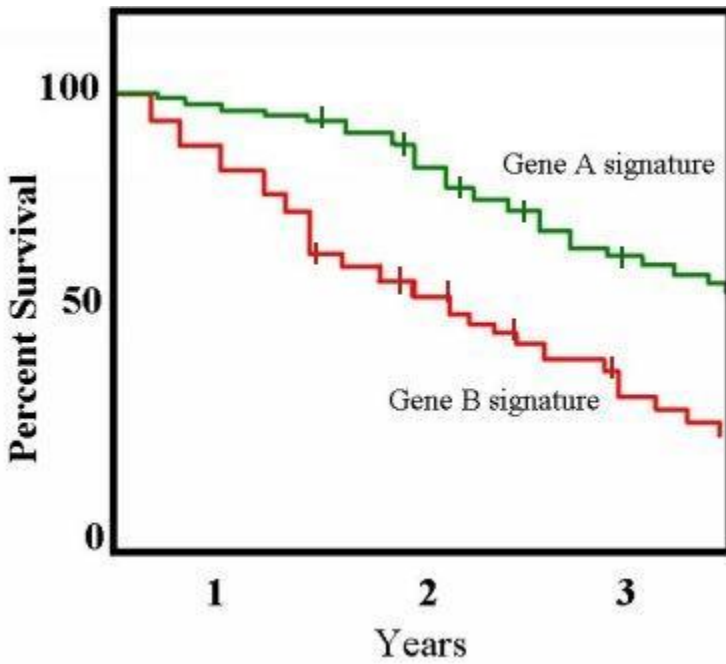
In this case, the ‘event’ should be completely unambiguous.

3.2. Survival curve

The survival probability = probability of surviving beyond a specified time point t

Estimated as:

$$S(t) = \frac{\text{Number of individuals with survival times } \geq t}{\text{Number of individuals at risk at time } t}$$



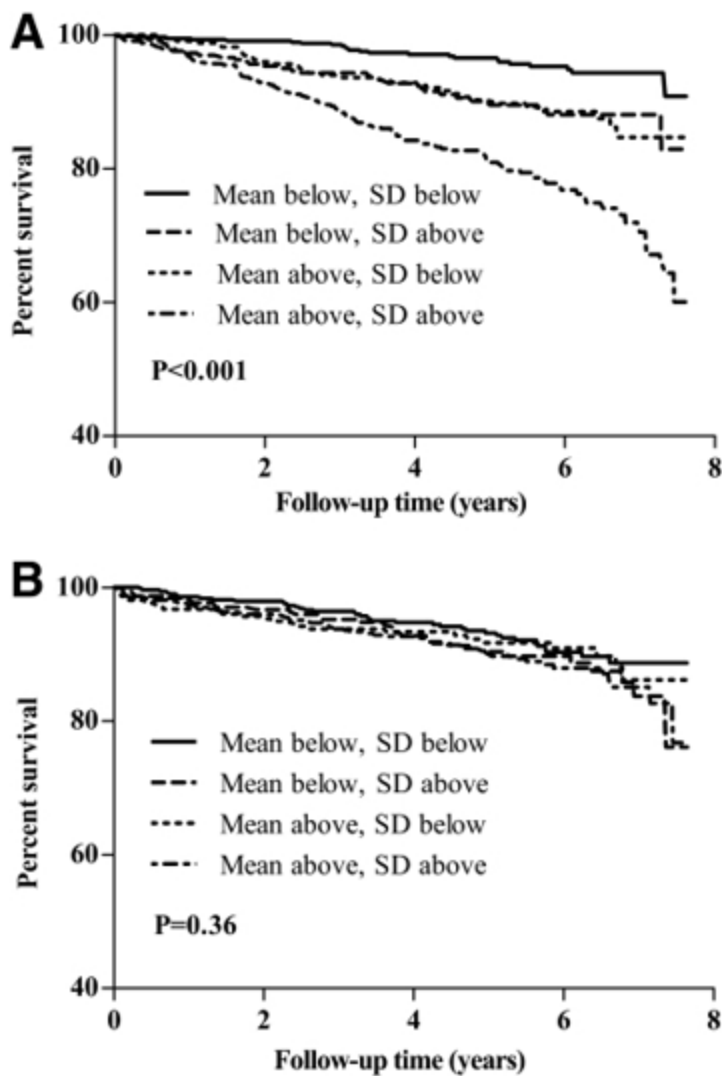


Figure 1: Example of a Kaplan-Meier survival curve. Taken from:

http://openi.nlm.nih.gov/detailedresult.php?img=2768180_zdb0110958890002&req=4

3.3. Analysis methods

The SAP describes which statistical methods are to be used to analyze the data. The following aspects need to be covered where applicable:

Main/primary analysis: to obtain the main clinical trial results on the specified trial endpoint(s).

Supportive/ sensitivity analyses: analyses on different sets of patients or using different analysis techniques than for the main analysis. These are used to confirm the conclusions of the main analysis.

Exploratory analyses: all other analyses e.g. further data exploration.

3.4. (Primary) analysis set

This section is designed to identify which of the recruited patients are to be included in the different analyses:

- Intention-to-treat (ITT).
- Per protocol (PP).
- Safety population.

The criteria typically relate to when the intended protocol could not be or was not followed. For example, if a patient who did not fulfil eligibility criteria was wrongly included. It's important to identify these 'protocol violations' and to deal with them appropriately in the analysis. This is because they may bias the final results of a trial or impact the power of the final analysis.

Case study: Consider the setting of a clinical trial comparing a new experimental treatment to the standard of care. However, some patients taking the experimental treatment are too sick, because of side effects, to go to the next visit within the allotted time. A possible approach would be to include only patients with complete follow-up (all the visits), so to exclude these patients with incomplete follow-up (missing visits) from the analysis. However, by doing so, one selects a sub-group of patients whom, by definition, will present an artificially positive picture of the treatment under investigation.

One potential solution to this problem is a statistical concept called intention-to-treat (ITT) analysis. ITT analysis includes every randomized patient and will consider that every patient received the treatment assigned by the randomization. As such, ITT analyses maintain the balance of patients' baseline characteristics between the different trial arms obtained from the randomization. 'Protocol deviations' such as non-compliance to the assigned treatment (schedule, dosing, etc.) are part of daily practice. Therefore, treatment-effect estimates obtained from ITT analysis are considered to be more representative of the actual benefit of a new treatment in real life.

Per protocol analysis (PP): This analysis population is restricted to the participants who strictly fulfil the protocol requirements in terms of patient eligibility criteria, treatment compliance and outcome assessment. PP analyses usually exclude patients who have not had at least one dose of the allocated treatment, all ineligible patients, patients with major protocol violations and sometimes patients with incomplete data for the targeted endpoint. A PP analysis is useful in determining the biological effect of a treatment. However, the value of the treatment may not be shown in a real-life situation since PP analysis is restricted to a highly selected patient subgroup corresponding to an 'ideal' setting.

Safety population: All randomized patients who have started their allocated treatment (at least one dose of the trial medicine). This analysis population is often used to describe the safety profile of a treatment.

3.5. Sub-group analysis

This section of the SAP aims to detail which sub-group analyses will be performed. Controlled

clinical trials are designed to investigate the effect of a treatment in a given population of patients. Sub-group analyses involve splitting the trial participants into sub-groups. This could be based on:

- demographic characteristics (e.g. sex, age)
- baseline characteristics (e.g. a specific genomic profile)
- use of concomitant therapy.

The principle is to look at the effects of treatment separately in different types of patients in order to collect information on who will benefit most from the investigated treatment. Sometimes, sub- group analyses are used to clarify heterogeneous treatment effects, e.g. when certain patient characteristics are driving the response to treatment.

Findings from sub-group analyses might be misleading for several different reasons. Firstly, sub-group analyses are observational (sub-groups are defined on observed patients' characteristics) and not based on randomized comparisons. The 'hindsight bias', also known as the 'I-knew-it-all-along' bias, is the inclination to see events that have already occurred as being more predictable than they were before they took place. This is why sub-group analysis should be pre-planned.

Even when pre-planned, they are still open to criticism of 'multiplicity'. When multiple sub-group analyses are performed, the risk of finding a false positive result (i.e. a type I error) increases with the number of sub-group comparisons. Multiplicity issues are in general related to repeated 'looks' at the same data set but in different ways until something 'statistically significant' emerges. With the wealth of data sometimes obtained, all signals should be considered carefully. Researchers must be cautious about possible over-interpretation. Techniques exist to protect against multiplicity, but they mostly require stronger evidence for statistical significance to control the overall type I error of the analysis. Here is a list of some of the common methods:

- Bonferroni's method.
- Holm's procedure.
- Hochberg procedure.

Finally, there is a tendency to conduct analyses comparing sub-groups based on information collected during the trial. A typical example is looking at the difference in survival between patients responding (yes/no) to treatment. Patients who are responding to treatment are by definition patients who are able to spend sufficient time on treatment to allow a response. Therefore, again by definition, they may simply represent a sub-group of patients of better prognosis and may therefore bias the analysis. This is an example of what is often referred to as 'lead-time bias's or 'guarantee-time bias'. One way of dealing with this is using a landmark as a starting point for the survival analysis, and creating the categories based on the patients' characteristics at the time of this landmark (e.g. did a patient respond at three months, yes/no).

3.6. Interim analysis

This is the analysis of early data accumulated in a clinical trial before all of the patients have been enrolled. The aim is to detect eventual trends with regard to the safety and/or efficacy of the tested treatment before the primary analysis. If the early data show a clear benefit with the tested treatment, it is unethical to continue exposing patients to less effective standard therapy. Similarly, if there is clearly no benefit with the tested treatment, the protocol should be modified or the trial stopped.

The SAP should describe:

- The planned interim analysis, including when it will take place.
- How the unblinding of data will be handled (in blinded trials).
- The scope of decisions on the trial conduct and on which criteria they will be based.

Repeated looks at the efficacy data during the course of a trial suffer from the same issue of multiplicity as explained previously. They can inflate the overall type I error rate of the trial (i.e. the probability of having a false positive outcome). In the case of interim analyses, specific methodology has been developed to establish thresholds for the significance of statistical tests supporting the decision to stop or not to stop a clinical trial. Examples of some of the methods are the Pocock, Haybittle–Peto and O’Brien- Fleming.

4. Trial maturity

Sufficient data need to be available in order to perform the planned statistical analysis. Rules for assessing the so-called maturity of the data need to be specified upfront in the SAP. Depending on the type of endpoint, this may refer to a certain number of patients with a pre-specified follow-up time (e.g. one year after the last patient was registered) or when the pre-specified number of events needed for the primary outcome analysis is reached, as is often done in the case of a time-to-event endpoint.

5. Unit summary

The statistical analysis plan (SAP) is a crucial trial document describing the planned analysis of clinical trial data. The SAP provides the researchers with relevant information and details on the scope of the primary, supportive and interim analyses, population definitions, and methodology.

Lesson 11: (excerpts)

3. Possible approaches in adaptive design

The term 'adaptive' covers a varied set of designs, but most of them follow a simple structure. Within an adaptive trial, there are learning and confirming stages, which follow a similar approach to the overall clinical development process across multiple trial settings (Phase I, Phase II, and Phase III). As a result, changes might be made to hypotheses or the design parameters.

Learning stages:

Major design elements may be changed (for instance dropping treatment arms). Statistical uncertainty (for instance bias, variability, incorrect selection). Estimation of the treatment effects (beneficial or adverse).

How the 'learning' part is done is crucial to the integrity and the validity of the trial results. Modifications based on blinded interim results will have less impact on the trial operating characteristics. However, adaptations based on unblinded interim comparative analyses can introduce all sorts of bias.

Confirming stages:

Control of statistical errors and operational biases are of utmost importance. Strong control of Type I errors is required (for example, finding a treatment efficient when it, in fact, is not).

The most commonly used adaptive design is trials with early stopping rules for futility (when the treatment or trial is not producing any useful results) or efficacy.

These rules are predetermined and are verified by one or more interim analyses. They prevent the participants from taking medicines that will not provide a beneficial effect or are unsafe. Most importantly, if it is found that the trial medicine is clinically more effective than the control, it would be unethical to continue administering the less-effective control medicine. Early stopping rules for futility allow a halt in the administration of a less-effective control medicine.

There are also designs where treatment arms are modified or dropped over the course of a trial, or where a sub-population is selected based on a biomarker of interest (pick the winner).

Some designs allow for sample size re-estimation, for instance an increase in the patient population if the results appear promising; or to maintain overall statistical power. At the interim analysis, a check is done for efficacy or futility. At this point the trial can be stopped in the presence of overwhelming evidence of efficacy or futility. If not, the conditional power (CP) is determined. The CP is the probability that the final study results will be statistically significant given the data currently observed. If this CP is either high or low then the trial is continued as planned. If the CP is somewhere in the middle, then the sample size is further increased.

During trial recruitment, if the expected number of participants under the original eligibility

criteria cannot be enrolled, modifications to non-critical eligibility criteria can be made based on examination of baseline characteristics.

Adaptive randomization is another example of an intuitively appealing design. In this design, a higher proportion of patients would be treated with the 'better' arm (if there is one). These adaptive trial designs are mostly based on unblended interim analyses that estimate the treatment effects – meaning that the analysts are aware of which treatment participants have been allocated to.

3.1. Example 1: Group sequential design

A group sequential design is a typical example of a Phase III trial with rules for early stopping for futility or efficacy. In the example trial depicted in the diagram below, patients were randomized between the first line treatment with either one medicine alone or two medicines in combination.

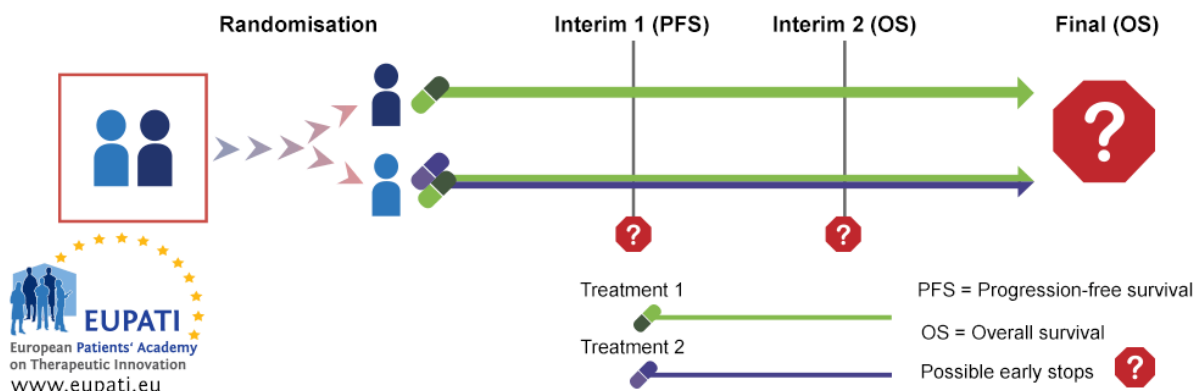
There were two interim stages where it was possible to stop the trial early and performing analysis before all the trial results are gathered. The trial could have been stopped:

- At Interim 1, for futility based on progression-free survival (PFS) – whether the patient stays free of any progression of a specific cancer or not.
- At Interim 2, for futility or efficacy based on overall survival.

Group sequential design is a classic example that is often forgotten when thinking about adaptive design, as it was already in use before other adaptive designs became more commonplace. Adaptation opportunities are planned upfront in the trial design, this results in the power and Type I error or sequential tests to be relatively easy to adjust when conducting multiple tests. This maintains the overall power and Type I error.

Group sequential design

An example trial using group-sequential design



Group sequential design Group sequential design allows for early stops on the basis of progression-free survival or overall survival. In this example, participants were randomized onto one of 2 arms, and received either Treatment 1, or combination of Treatment 1 and

Treatment 2.

3.2. Example 2: Multi-arm, multi-stage design (MAMS)

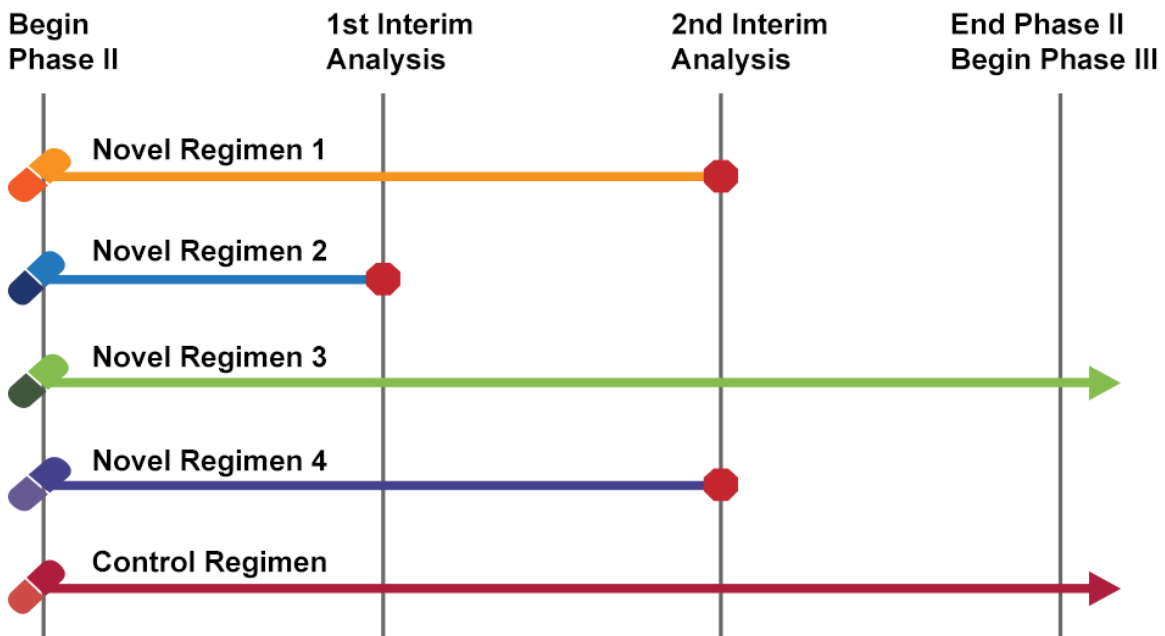
The multi-arm, multi stage (MAMS) trial is a new paradigm for conducting randomized controlled trials which makes use of an interesting adaptive design.

MAMS trials allow the simultaneous assessment of a number of research treatments against a single control arm. MAMS trials provide earlier answers and are potentially more cost-effective than a series of traditionally designed trials.

In this example, we see a design that uses multiple arms and stages at the same time.

The MAMS design requires a definitive primary and intermediate primary outcome measure. The definitive outcome measure is the one upon which the final conclusions should be based; the intermediate outcome measure provides a means of screening for emerging evidence of evidence.

Multi-arm multi-stage (MAMS) design



Multi-arm multi-stage design

The multi-arm multi-stage design (MAMS) allows multiple treatments to be tested simultaneously against a single control.



At the first interim analysis in the example above, Novel Regimen 2 is considered to lack sufficient benefit compared with the control and is not taken forward to stage 2. At the second interim analysis, recruitment to Novel Regimens 1 and 4 is stopped, and only the control regimen and Novel Regimen 3 are continued to the end of trial and advanced into Phase III studies.

Advantages of the MAMS design:

- **Fewer participants**

In this design, several trials are performed at once, which helps reduce the number of participants randomized to the control arm.

- **Less overall time required for medicine discovery**

The intermediate steps of the MAMS design replace the separate Phase II step. The decision on whether the medicine is sufficiently active is incorporated as a pilot phase into this trial.

- **Fewer applications and approvals required**

Regulatory work is done for one trial instead of for multiple trials.

- **Flexible**

Uninteresting Arms can be dropped and new arms can be added. Reduced cost This trial design requires fewer participants, fewer regulatory applications, and less overall time, all of which help to save on development costs.

Disadvantages in MAMS design:

- **Operating characteristics**

Because of the complexity of this approach, it may be difficult to manage and requires a lot of simulations during the design process.

- **Required number of participants**

This depends on the operating characteristics, but if treatment arms are added during the course of the trial, it may be difficult to predict budget and regulatory issues.

- **Trial duration**

If treatment arms are added, it becomes difficult to predict when the trial will naturally end.

- **Continued accrual (recruitment) to control arm**

In order to avoid a time bias when new treatment arms are added, recruitment to the control arm must continue throughout the course of the trial. Consideration must also be given to what happens if a new standard of care becomes available during the course of the trial – is the control still relevant?

- **Comparison between experimental arms**

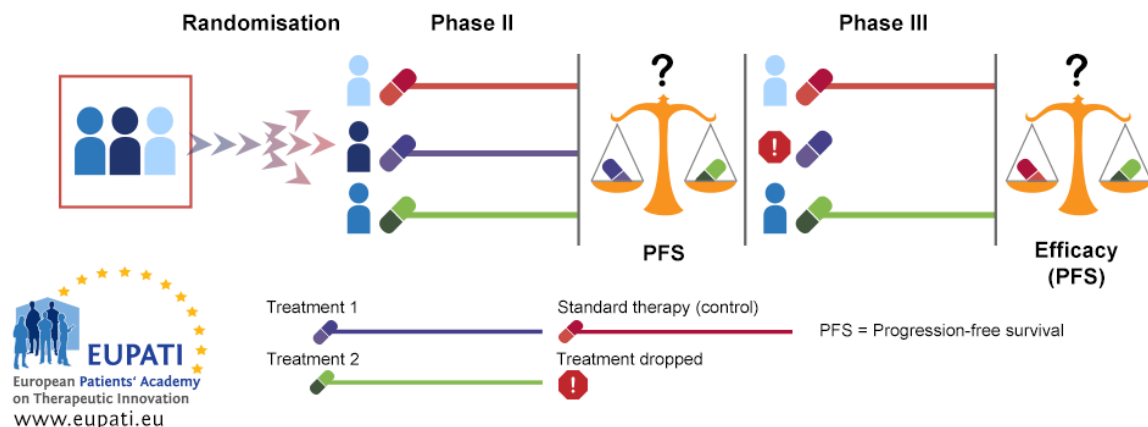
The MAMS design only allows for comparisons between individual treatment arms and the control arm; it does not allow for comparison between individual treatment arms themselves.

3.3. Example 3: Seamless Phase II/Phase III design

Seamless Phase II/Phase III design is often used in the case of rare diseases; it is also called a ‘combination test’. In the example below, patients are randomized between three treatment arms in the first stage of the design (Phase IIb). The first treatment arm is the control arm, where patients receive the standard of care therapy. Patients on the second and third treatment arms receive different treatments, Treatment 1 or Treatment 2.

At the end of the first stage (Phase IIb), Treatment 1 and Treatment 2 are compared based on best progression-free survival (PFS). The least effective treatment arm is dropped. The other treatment arm is then continued in the second stage (Phase III). In this stage, an efficacy comparison is performed against the standard of care treatment.

Seamless Phase II/III design



Seamless Phase II/III design

The seamless Phase II/III design allows Phase II and Phase III to be performed in the context of one trial.

Advantages in Seamless Phase II/Phase III design

- **Helps to mitigate bias**
Both steps are conducted independently and the results of both steps are combined in the end in an overall test result.
- **Shortens time and participant exposure**
Phase II and Phase III are performed within the context of one trial.
- **Relatively flexible**

The way that the treatment arm for final comparison is chosen in the Phase II part and merged with the Phase III part is relatively flexible. Efficient use of resources
Participants from Phase II and Phase III both contribute data to the final results.

Disadvantages

- **Complicated statistical analyses**

This design requires statistical aspects that are not so straightforward.

- **Recruitment gaps**

There is a gap in the recruitment between the two phases while waiting for enough data to be gathered in order to perform the interim analysis that decides whether to continue or not.

- **Logistic challenges**

This design is logistically challenging – it requires a quick flow of data so that the number of events in the analysis can be followed up on.

- **Difficulties arising from long-term endpoints**

This design requires information on PFS to be available relatively quickly. This becomes more difficult when the endpoints are long-term.

- **Risk of lost information**

Combining two arms risks the loss of information.

3.4. Example 4: Response-adaptive randomization

In this design, patients are hierarchically randomized based on their biomarker profile into one of four treatment arms. They start with a set of initial randomization probabilities and then based on the observed eight-week outcome, new randomization probabilities are derived. This maximizes the chance that the patient receives the treatment that is most effective for him/her.

To be able to make use of response-adaptive randomization, a few pre-requirements are necessary:

First, such a trial requires a fast dataflow. Data on the eight-week endpoint needs to be rapidly available in the data center so that they can update the randomization probabilities guiding new participants entering the trial. This is not always straightforward in the context of a large multicenter trial.

This example had a short endpoint (eight-week). It would not work so well if a Progression Free Survival (PFS) at six months was needed to guide randomization of new participants.

It is difficult to interpret results beyond estimation. It is difficult to perform comparisons when no longer working with the classical concept of two independent samples. In addition, it can't be conclusive for an arm that was prematurely terminated and for which little data is available.

When investigators start to realize that more participants are being randomized into a particular treatment arm, recruitment patterns may change during the course of a trial. This introduces operational bias, e.g. sicker participants could enroll earlier and healthier ones could decide to wait for a higher chance of receiving better treatment. In such a setting, blinding is essential but may not always be feasible.

4. Patient Involvement

Patient input into adaptive design can help researchers identify the most appropriate design by helping to define and understand the needs and requirements of the patient population. Patients can also be involved in the Data Safety Monitoring Board.

5. Further Reading

Clinical Trials Planned with an Adaptive Design (2007) at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf

Chow SC, Chang M. Adaptive design methods in clinical trials - a review. Orphanet J Rare Dis. 3(11) May 2008. <http://www.ojrd.com/content/pdf/1750-1172-3-11.pdf>

I. Judson, J. Verweij, H. Gelderblom, et al. Results of a randomized phase III trial (EORTC 62012) of single agent doxorubicin versus doxorubicin plus ifosfamide as first line chemotherapy for patients with advanced or metastatic soft tissue sarcoma: a survival study by the EORTC Soft Tissue and Bone Sarcoma Group. Sydes MR, Parmar MK, James ND, et al. Issues in applying multi-arm multi-stage methodology to a clinical trial in prostate cancer: the MRC STAMPEDE trial. *Trials*. 10 (39) Jun 2009 <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2704188/pdf/1745-6215-10-39.pdf>

Kairalla JA, Coffey CS, Thomann MA, Muller KE. Adaptive trial designs: a review of barriers and opportunities. *Trials*. 13 (145) Aug 2012 <http://www.trialsjournal.com/content/pdf/1745-6215-13-145.pdf> Cytel Webinar for East@SurvAdapt. October 28, 2010 http://www.cytel.com/pdfs/East-SurvAdapt-Webinar_10.10.pdf

Lesson 12: Clinical Trials in Non-Standard Situations

Clinical trials in non-standard situations by Markku Toivonen, MD, PhD Scientific Director [NDA](#) Regulatory Science Ltd.

Click on link below. First video is Markku Toivonen

Provided unmodified from: [EURORDIS Training Module on METHODOLOGY](#)



[Methodology](#) **Module:**

EURORDIS Training Module on METHODOLOGY by [EURORDIS](#) is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License](#)

Lesson 13: Investigator Brochure

Table of contents

1. Introduction
2. Content of the Investigator's Brochure
 - 2.1. Table of Contents
3. Regulation of Investigator's Brochure
4. Further Reading

1. Introduction

An Investigator's Brochure (IB) is a compilation of data on the investigational medicinal product (IMP) (the medicine being studied) which is relevant to the study of this product in humans. It includes clinical and non-clinical data, i.e. any data from studies with patients (clinical) and data from other sources such as laboratory tests (non-clinical).

The purpose of the IB is to provide the investigator and others (e.g. clinical trial coordinators, study nurses) with background information to help them work in line with the protocol.

The IB is prepared by the sponsor. They also control the distribution of the document. This is because it is the single most comprehensive document summarizing the properties of the IMP. It provides the clinician or potential investigator with information they need to assess the appropriateness of the trial, including the benefit-risk relationship. It allows them to do this in an independent and unbiased way. It provides insight to support the clinical management of study participants during the clinical trial. This includes information about doses, dose frequency, methods of administration and safety monitoring procedures. The IB is submitted with the Clinical Trial Application (CTA). This is sent to the Competent Authority for approval.

2. Content of the Investigator's Brochure

According to the EU requirements for good clinical practice in clinical trials (Note for guidance on Good Clinical Practice (CPMP/ICH/135/95)

http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002874.pdf), the information in an IB should be:

‘... presented in a concise, simple, objective, balanced and non-promotional form that enables a clinician or potential investigator to understand it and make an unbiased benefit-risk assessment of the appropriateness of the proposed clinical trial’.

The content should be approved by the scientists that generated the data in different disciplines (pharmacology, toxicology, etc.). It should also be reviewed and updated at least once a year, or upon receipt of significant new data. The IB must include:

Information about the sponsor's name and identity of each IMP (trial number, international non-

proprietary name -INN (generic name) and trade name). Confidentiality statement, with instructions to treat the document as confidential for the exclusive use of the investigator's team and review boards and ethics committees. Compilation of results gathered from non-clinical and clinical studies of the medicine. Background information on the properties and history of the medicine.

2.1. Table of Contents

The IB must include the following sections:

1. Table of contents.
2. Summary - guidance for the investigator, highlighting the important information relevant to the stage of clinical development of the product.
3. Introduction - background information, containing the chemical and international non-proprietary name - INN (generic name) of the active substance, and the trade name (when approved) of the IMP. This should also include the base for performing research and what the medicine is intended to treat (the indication). Additionally, it should provide the general approach to evaluate the IMP.
4. Physical, chemical and pharmaceutical properties and formulation - relevant properties of the investigational product (active substance and excipients (inactive ingredients)), including any similarity to any other known compound. Instructions for the storage and handling of the dosage forms should also be given.
5. Non-clinical studies - including the results of pharmacology, toxicology, pharmacokinetic and metabolism studies. The methodology used, results and relevance of findings should be explained, and when appropriate, there should also be information about animals and doses.
6. Effects in humans - including toxicological results, pharmacokinetics, metabolism, safety, efficacy and when possible, summaries of each clinical trial with the IMP. This section should identify any information from previous marketing experience, including the countries where the investigational product has or has not been approved.
7. Summary of data and guidance for the investigators – an overall discussion of the non-clinical and clinical data so they have the most informative interpretation of available data. This will help investigators to anticipate adverse drug reactions (ADRs) or other problems in clinical trials.

3. Regulation of Investigator's Brochure

Regulatory Authorities (EMA, NCA, etc.) require an up-to-date IB for any medicine in development or on the market (if applicable). They also review the updates to the IB and compare it to previous versions to ensure it is accurate, complete, and impartial.

If the investigational product has already been issued with a marketing authorization and its pharmacology is widely understood by medical practitioners, an extensive IB may not be necessary. The Summary of Product Characteristics (SmPC) may be used instead -

this is the document containing product information for healthcare professionals that is provided with an authorized medicine. Where permitted by Regulatory Authorities, a basic product information brochure, package leaflet, or labelling may be an appropriate alternative. However, it must include current, comprehensive, and detailed information on all aspects of the investigational product that might be of importance to the investigator.

If a marketed product is being studied for a new use (i.e. a new indication), an IB specific to that new use should be prepared. When relevant new information is available, the investigators and Research Ethics Committees (RECs) should be informed. Where possible, these parties should be told before this information is included in the revised IB.

In cases where preparation of a formal IB is impractical, the sponsor should provide, as a substitute, an expanded background information section.

4. Further Reading

Page 34-38. International Conference on Harmonization (ICH) Guidance for Industry E6: Good Clinical Practice, Section 7:

http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf

Lesson 15: Fraud and Misconduct in Biomedical Research and Clinical Development

Table of contents

1. Introduction
2. Examples of fraud in clinical research
 - 2.1. Example 1: USA, University of Vermont, Eric T. Poehlman
 - 2.2. Example 2: Korean stem-cell case, Woo-Suk Hwang
 - 2.3. Example 3 Jon Sudbø, Norwegian Radium Hospital, Oslo
3. Prevention of misconduct
4. Detection of fraud and misconduct

1. Introduction

The word misconduct is used differently in English. Therefore, there are different definitions. According to the UK Medical Research Council and the Wellcome Foundation, misconduct is defined as:

‘The fabrication, falsification, plagiarism or deception in proposing, carrying out or reporting results of research or deliberate, dangerous or negligent deviations from accepted practices in carrying out research.’

Alternatively, the Joint Consensus Conference, Edinburgh, (1999) defined it as:

‘Behavior by a researcher, intentional or not, that falls short of good ethical and scientific standards.’

In contrast, a definition of clinical research fraud is clearer:

‘The generation of false data with the intent to deceive.’

2. Examples of fraud in clinical research

Sponsors of clinical trials and doctors participating in clinical trials (investigators) put a lot of emphasis on planning, performing and evaluating clinical trials. This is to ensure that the trial results in:

- Trustworthy data.
- Clear results and treatment recommendations.

The data are generated by the investigators who enroll the participants into the trial they perform the measurements and observations. Findings are documented in Case Report Forms (CRFs) prepared by the sponsor. The Good Clinical Practice (GCP) standard defines the

conditions that sponsors and investigators have to apply to ensure good quality data. Auditing the trial performance at the investigator's site ensures that the trial is carried out properly. However, some investigators may decide to falsify or even invent data.

This topic explains the following:

- the difference between 'misconduct' and 'fraud' – examples are provided,
- ways to prevent, detect and handle fraud and misconduct.

The conduct of most clinical research is honest and adheres to ethical principles. Occasionally the sponsor or organizer of a clinical research project may be faced with data which are questionable, such data may or may not be fraudulent.

Fraud should not occur at all. However, we do not live in an ideal world...Crimes happen everywhere. The following three examples of fraud from different regions in the world demonstrate what happens when fraud is committed.

2.1. Example 1: USA, University of Vermont, Eric T. Poehlman

http://en.wikipedia.org/wiki/Eric_Poehlman

Eric T. Poehlman (born c. 1956), a scientist in the field of human obesity and ageing, was the first academic in the United States to be jailed for falsifying data in a grant application. His notorious crime was to publish utterly fraudulent research alleging hormone replacement injections as a therapy for menopause, when in fact it had no proven medical benefits at all. He joined the University of Vermont (UVM) College of Medicine in 1987 as an assistant professor, later working for three years at the University of Maryland in Baltimore. He eventually returned to UVM as a full professor.

Poehlman built a reputation as one of the leading authorities on the metabolic changes that come with ageing, particularly during menopause; he published more than 200 journal articles over two decades of research. His papers included research on the genetics of obesity and the impact of exercise, often following human participants over time to document changes in their physiology. However, his stellar career began to fall apart when Poehlman's misconduct was detected and exposed by a former University of Vermont lab technician, Walter DeNino, who once viewed Poehlman as his mentor. Poehlman was accused of scientific misconduct and on March 17, 2005 pleaded guilty to the charges, acknowledging falsifying 17 grant applications to the National Institutes of Health and fabricating data in 10 of his papers that were submitted between 1992 and 2000.

On June 28, 2006, Poehlman was ordered to serve a year and a day in federal prison for using falsified data in federal research grants that he had submitted for funding. An official with the National Institutes of Health said that this was the first case where an academic research scientist was given prison time for falsifying data in grant submissions. In a plea bargain that he made with the prosecutors, Poehlman pleaded guilty with one \$542,000 grant; the government prosecutors stated that Poehlman had defrauded agencies out of \$2.9 million.

‘Dr. Poehlman fraudulently diverted millions of dollars,’ said David V. Kirby, the US attorney for Vermont. ‘This in turn siphoned millions of dollars from the pool of resources available for valid scientific research proposals. As this prosecution proves, such conduct will not be tolerated.’

Before imposing the sentence, Judge William Sessions III said, ‘I generally think deterrence is significant, perhaps more so in this case. The scientific community may be watching’. Sessions reprimanded Poehlman for his misconduct, saying he had ‘violated the public trust’.

In addition to jail time, Poehlman is permanently barred from getting more federal research grants and was ordered by the court to write letters of retraction and correction to several scientific journals.

2.2. Example 2: Korean stem-cell case, Woo-Suk Hwang

<http://stemcellbioethics.wikischolars.columbia.edu/The+Cloning+Scandal+of+Hwang+Woo-Suk>

Between 2004 and 2005, Professor Hwang WooSuk, a highly regarded, highly funded South Korean researcher at Seoul National University, achieved international fame for his work on embryonic stem cells and the promises his findings offered. Considered a national hero, he had surprised the world with his report of creating 11 patient specific stem cell lines. His reputation was quickly destroyed, however, and his research activities were halted when his success in somatic cell nuclear transfer (SCNT – a method of ‘cloning’) became mired in scandal, particularly when it emerged that many of his data on SCNT were made up. He lost his university position and his two important papers on embryonic stem cell research had to be retracted from the journal ‘Science’.

Ethics violations

Several ethics violations were committed by his team members during the course of their research. In 2009, Hwang was convicted of misusing research funds and illegally buying human eggs for his research. Among many transgressions was the dubious manner in which the team persuaded women to donate their eggs for their SCNT research. Investigations revealed that many of the women who provided eggs had not given valid, informed consent, and nearly 75% of them reported that they were given cash or offered various financial incentives.

Some of the women who provided eggs were infertile patients who had agreed to donate any excess eggs following their fertility treatment. What they weren’t told, however, was that their eggs were initially assigned a quality grade, and the higher marked eggs were set aside for research while the lower graded ones were used for their treatment. Others who agreed to donate for the cause of research alone were not fully informed of the potential risks and harms involved in the egg donation process or the nature of the research for which their eggs would be used.

Concerns about probable coercion later surfaced when it became clear that at least two of the egg donors were junior members of Hwang’s research team. One, a PhD student, was listed as a coauthor of the 2004 Science paper. The other, apparently reluctant, was escorted to the donor

clinic by Hwang himself. Given the precarious position in which they presumably found themselves, the reported pressure to donate seems obvious.

The other chief concern raised by the method of gaining eggs was the payment that many of the women received. Some eggs were purchased directly, while in other cases women received compensation in the form of discounted fertility treatment. Though the concerns of egg trafficking – i.e. that women will be unduly pressured to donate despite the inherent risks are well agreed upon, there is no international consensus on the acceptability of selling eggs. There were no legal restrictions in place at the time of Hwang's actions.

The large number of eggs Hwang used in his SCNT experiments was staggering. The Ministry of Health and Welfare reported that Hwang acquired 2,221 eggs from 119 women while the Prosecutors' Office reported that 2,236 eggs were acquired from 122 women. Additionally, there were eggs that were retrieved from excised ovaries. The total number of eggs purchased or traded was 1,649, approximately 75% of the total number of eggs the Hwang lab used for research.

2.3. Example 3 Jon Sudbø, Norwegian Radium Hospital, Oslo



The screenshot shows a newspaper article from *Aftenposten* (NEWS FROM NORWAY) dated 17 Jan 2006, 11:32. The article is titled "Research scam makes waves" and discusses a Norwegian doctor's fabrication of cancer research. It mentions that the fraudulent research has led to faulty treatment of cancer patients and international investigations. The article quotes Richard Horton, editor of *The Lancet*, who is furious that a Norwegian doctor duped his publication. The article also mentions that the doctor's 13 co-authors and colleagues on the fraudulent cancer research project could have been duped as well. The article is accompanied by a black and white photograph of Richard Horton.

Aftenposten
NEWS FROM NORWAY

Celebra
JARLSBERG

First published: 17 Jan 2006, 11:32

Research scam makes waves

A Norwegian doctor's fabrication of cancer research is making waves far beyond Norway's borders. The fraudulent research may have led to faulty treatment of cancer patients, international investigations have been launched into how the fraud could have occurred, and top Norwegian officials all the way up to the ministerial level are desperately trying to control the damage.

The editor of the respected magazine, *The Lancet*, in which the fabricated article was published, calls the fraud "the worst the research world has seen."

Richard Horton told Oslo newspaper *Aftenposten* that he also can't understand how the Oslo doctor's 13 co-authors and colleagues on the fraudulent cancer research project could have been duped as well.

Horton claims at least six of the doctor's co-authors corresponded with *The Lancet*, and were highly involved with the substance of the article.

Richard Horton, editor of *The Lancet*, is furious that a Norwegian doctor duped his publication. He insists all editorial controls were in place.

PHOTO: THE LANCET

Jon Sudbø, a Norwegian investigator, fabricated results over a period of several years in the

Materials modified and shared with permission from <https://www.eupati.eu>



field of oncology. These results were then published in leading medical journals. The article that led to his downfall, was published in ‘The Lancet’, a top reputation journal. It was based on 900 patients that Sudbø had made up entirely. The editor of The Lancet, pictured above, described this as the biggest scientific fraud conducted by a single researcher ever.

3. Prevention of misconduct

Tackling research misconduct

- Prevention.
- Detection.
- Investigation.
- Prosecution.

For research involving clinical trials, standards were set in the ICH Good Clinical Practice (GCP) guidelines.

In Europe, these standards are set in the EC Directives on Clinical Trials (2001/20/EC) and on Good Clinical Practice (GCP) (2005/28/EC), which encompass ICHGCP. These are now enacted throughout Europe.

Misconduct and fraud destroys the trust of the public in clinical research and harms the patients treated on the basis of invented or unreliable study results. Therefore, misconduct and fraud must be prevented if ever possible. If it happens, it needs to be detected, carefully investigated and rigorously prosecuted where appropriate.

4. Detection of fraud and misconduct

Unintentional misconduct can be best prevented by training people and explaining to them why it is so important that they adhere to the quality standards for the conduct of clinical trials.

The prevention of misconduct must come first. But, if misconduct or fraud occurs despite the setting of these standards and training in their operation, then corrective procedures must be in place.

When misconduct is detected:

- The problem must be discussed with the person(s) concerned.
- The reason for the misconduct must be identified.
- Measures must then be taken to avoid the misconduct happening again.

Who finds fraud and misconduct?

In clinical trials misconduct and fraud are often identified by monitors checking the case report forms (CRFs) and trial performance at the sites. These people are called ‘Clinical Research Associates’ (CRAs). But also, the ‘Auditors’ sent to the sites by the sponsors to check the overall study performance and the GCP Inspectors sent by the competent authorities detect these problems. Sometimes suspicious data only become obvious once the statistician checks the

overall set of data and detects special ‘patterns’, less variability or lack of outliers.

An important source of fraud and misconduct detection are collaborators and team members. This might be a very difficult situation for them as they may have to blame a colleague, friend or boss of having done something wrong. It is important that these ‘whistle blowers’ receive the required attention and protection.

In an ideal situation, misbehavior or errors are detected by the researcher himself. People should be encouraged to announce their mistakes. They will only do so if they work in an environment where people are not punished for their mistakes. Instead the reason(s) for the misconduct should be detected and the team should work together to prevent this from happening again.

Lesson 17: Participant compensation (compensation schemes) and travel expense reimbursement in clinical trials

Table of contents

1. Introduction
 - 1.1. Task 1
2. What is compensation?
 - 2.1. Compensation for participation
 - 2.2. Compensation for damage suffered (insurance/indemnity)
3. Ethical considerations
 - 3.1. Vulnerable populations
4. How much compensation?
5. Summary
6. References

1. Introduction

Although not always standard, it is customary in many clinical studies that participants receive some form of compensation for their participation. This may take the form of money, the reimbursement of travel expenses, food or food vouchers, or other services. This lesson will introduce the current standards for such compensation, and prompt you to think about a few key aspects such as ethics, questions around vulnerable populations, and the involvement of different stakeholders.

1.1. Task 1

Take a few minutes and think about any possible considerations and questions you may have around the issue of compensation to clinical study participants. Do you have any reservations? Have you ever come across a compensation scheme? What types of compensation can you think of? Write down all your comments and questions on a piece of paper.

2. What is compensation?

Compensation can mean two distinct things:

- 1) When participants receive money or other benefits for their participation in the clinical trial, or
- 2) When they receive a payment or other services if they suffer any harm from a clinical trial (insurance/indemnity).

Both are called compensation and described in this lesson.

The practice of compensating participants for taking part in a clinical trial has been in place for more than 200 years. However, as one study points out (1), this has always remained a matter of debate with no universally accepted standards – mainly for ethical and moral reasons. So, why then is compensation paid in research? Compensation can be useful to recruit the required number of trial participants in the given time frame. It can be paid for relieving participants of financial burden or recognizing time sacrifice or as appreciation of their contribution to science; it can also be useful to motivate healthy volunteers to take part in Phase I clinical trials.

2.1. Compensation for participation

It depends on the sponsor and the given study whether or not compensation is paid to participants. Compensation schemes are more common in Phase I clinical trials. According to one report by the European Forum for Good Clinical Practice (EFGCP) (2), legislation and practice in Europe vary widely from country to country. Some countries exclude compensation entirely, but the most common practice requires that any compensation is reviewed and approved by the ethics committee associated with the given study. The EU Clinical Trial Directive (2001/20/EC) and Regulation 536/2014 (3) specifically prohibit the payment of any compensation for participation - other than the reimbursement of expenses or income lost - to participants who are incapacitated and cannot sign an informed consent form (ICF), or who are pregnant, or minors. Otherwise this EU legislation only provides that 'no undue influence, including that of a financial nature, [must be] exerted on subjects to participate in the clinical trial'.

For example, one clinical research center in the UK provides detailed guidelines for healthy volunteers who want to participate in clinical trials (4). This assures that a small payment is ensured alongside the reimbursement of travel expenses.

Some research suggests that the amount of money received by participants influences their decision to participate. Do you agree that financial compensation may influence participants or patients to take part in a clinical trial and not consider potential risks? Share your thoughts in the forum

2.2. Compensation for damage suffered (insurance/indemnity)

The EU Clinical Trials Directive introduced an 'obligatory insurance/indemnity' which has substantially increased the costs and administrative burden of conducting clinical trials. The new regulation recognizes that clinical trials do not always pose additional risk to the participants over normal clinical treatment. Therefore, in cases where there is no additional risk, or such risk is negligible, no specific damage compensation (insurance or indemnity) will be required. With respect to trials where there is additional risk and the sponsor is obliged to ensure adequate insurance coverage, the new regulation puts EU member states under an obligation to set up a national indemnification mechanism on a not-for-profit basis. This should be of particular value to non-commercial sponsors, for whom it has been difficult to obtain insurance

coverage. (5) The EU also requires all sponsors and CRO's to be completely transparent about financial transactions made with participants or trial sites.

The informed consent signed by the participant must contain specific references to any compensation schemes, and the insurance coverage offered to the participants should they suffer any injury or harm. The informed consent should also be specific about how the insurer can be contacted, so that participants are not necessarily required to arrange their claims through the study personnel or the CRO.

3. Ethical considerations

Payments in clinical trials have raised ethical concerns for many years. In 2005 Christine Grady (6) from the National Institutes of Health summarized the problem:

‘Several ethical concerns arise regarding the payment of research participants. The most commonly expressed concern is that payment could be coercive or serve as undue inducement to research participants. By definition, coercion is understood to involve a threat of physical, psychological, or social harm in order to compel an individual to do something, such as participate in research. However, money for research participation is an offer or an opportunity and not a threat and therefore cannot be perceived as coercion. But can money be considered an undue inducement? Existing guidelines warn against undue inducement and its potential to compromise informed consent, although there is disagreement about what exactly constitutes undue inducement and consequently disagreement about the extent to which it is a valid problem in research.’

3.1. Vulnerable populations

Compensation is always a special concern in case of vulnerable populations. It is a particularly difficult issue in case of children and the mentally challenged. They do not or cannot make their own decisions but their parents/legal guardians will decide for them, while the risk is not divided the same way: the member of the vulnerable group carries the risk, but the parent or guardian gets the compensation. This is one of the reasons why the EU does not allow compensation to such vulnerable groups beyond the reimbursement of their expenses. ‘For the poor, illiterate and the unaware, monetary inducements can easily be enticing’, state Pandya and Desai (1). It may happen that the participant does not understand completely what the implications of participation may be. Patient advocates and patient organizations may play a key role in mediating in these situations, and in flagging any irregular practices in this field to the authorities.

4. How much compensation?

Pandya and Desai give an accurate description of the currently existing models for setting the amount of compensation (1):

‘There are several proposed models of making payment to subjects for trial participation. Some of the ways are more ethically acceptable than the others. The common models are:

The **market model** is based on the principle of supply and demand, which decides when and what is to be paid to the research subjects for a particular study in a particular location. This means that compensation is paid to the subjects for the studies that offer little or no benefits or the studies for which the target population is difficult to reach. Also, this implies that in case of studies that offer benefits or have a huge target population, little or no compensation is paid. This model has advantages like targeted number of subject recruitment achieved in the required time frame, decreased financial sacrifice by the subjects and high completion bonus ensures protocol compliance. However, on the flip side, this model leads to very high compensation in few of the hard-to-find-subject studies, which could serve as undue inducement and could unnecessarily commercialize the research participation. High payment can lead to subjects not paying attention to the risks involved in the study, as well as leading them to hide important data that could deem them ineligible for the study. It could also create situations where the investigators are competing for subjects by paying higher amounts.

The **wage model** is based on the concept that research participation requires little or no skill, but it does involve consideration of the time and effort of the subject and also discomfort that is faced by subjects. The model is in alignment with egalitarianism. This model suggests that the subjects engaged in similar activities be paid similarly. Thus, here, the subjects are paid on a scale parallel with that of the unskilled but essential jobs. The advantages of this model could include minimization of the issue of undue inducement, reduced inter-study competition as seen in the market model that would also encourage investigators to minimize the risks involved, decreased financial sacrifice by the subjects and prevention of discrimination between high-income and low-income groups (like the reimbursement model described below) as subjects of the same study receive equal compensation. However, it creates difficulty in achieving the targeted number of subject recruitment in the required time frame and it usually attracts the low-income population. It views subject's research participation as an unskilled job and many believe it to be inappropriate commercialization of the research participation.

The **reimbursement model** is also in alignment with the egalitarianism principle. This model suggests that compensation should only recover the costs incurred by the subject for participating in the trial. Also, the time spent away from work can be reimbursed proportional to their earning capacity. This model helps in resolving the issue of undue inducement to a certain extent. Subjects are less likely to hide information or overlook the risks involved in the study. The model also decreases the financial sacrifice by the subjects. On the other hand, the issues with this model could be difficulty in achieving the targeted number of subject recruitment in the required time span. Also, different subjects have different earning capacities based on their qualifications, which leads to either preference for the low-income group or high cost of study if subjects from the high-income group are selected.

The **appreciation model** suggests compensation at the time of study completion as a token of gratitude or appreciation. This has no impact on the study recruitment as it is given at the end of the study. However, this model could have an impact on subject retention and may act as an inducement to prevent a patient from discontinuing. It needs to be used along with

one of the above-mentioned models. The researcher needs to carefully weigh the pros and cons of each of the above models and decide which one is best suited for the study on hand. It is also best to decide and document the mode of compensation before the trial is initiated, taking the stakeholders and the Ethics Committee in confidence and with the mandatory approval obtained from these.’

5. Summary

There are no clear, standard guidelines to follow for CRO’s or sponsors regarding compensation to research participants for taking part in a trial. Compensation is not always monetary, it can also take the form of health services, food, or other benefits. There are different models of compensation that studies apply, and often no compensation is given for participation alone, except to healthy volunteers in Phase I studies. Patient activists and advocates should be prepared to know more about the different regulatory frameworks in order to judge compensation related issues. In the European Union, it is mandatory to include explicit references to insurance coverage for study-related injuries in the informed consent documentation. Patient advocates should always be aware of the special needs and issues of vulnerable populations.

6. References

1. Pandya, M., & Desai, C. (2013). Compensation in clinical research: The debate continues. *Perspectives in Clinical Research*, 4(1), 70–74 <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3601710/>
2. <http://www.efgcp.eu/Downloads/EFGCPReportFiles/EFGCP%20ECs%20Report%202012%20-20Question%2026%20Updated.pdf>
3. http://ec.europa.eu/health/files/eudralex/vol-1/reg_2014_536/reg_2014_536_en.pdf
4. <https://crc.surrey.ac.uk/volunteers/participant-faqs>
5. <http://www.mondaq.com/x/238818/Life+Sciences+Biotechnology/Clinical+Trials+European+Commission+Proposal+To+Replace+The+Clinical+Trials+Directive>
6. Grady, C. (2005). Payment of clinical research subjects. *Journal of Clinical Investigation*, 115(7), 1681–1687 <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1159153/>

Lesson 18: Within-trial decisions

Table of contents

1. Introduction
2. Code-breaking (unblinding)
3. Premature termination
4. Data safety monitoring boards (DSMB)
 - 4.1. Function of Data Safety Monitoring Boards (DSMB)
5. References

1. Introduction

There are many different types of decision that may need to be made once a clinical trial has begun. Some are routine decisions in managing a trial, but especially in case of safety issues, some decisions may have a major impact on the conduct of the trial and could even result in the termination of the trial.

These decisions can also result in the modification of the study protocol, e.g. the trial rationale, procedural aspects, inclusion and exclusion criteria, medication and dose to be used, as well as many other details.

There are many different issues that may arise during the conduct of the trial that may necessitate immediate decision-making, such as premature trial termination. Such issues may be related to the emergence of new scientific evidence, as the trial progresses, e.g. interim trial results may give an indication on whether an intervention is beneficial or not. Trial data, and their interim analysis, often necessitate a reassessment of the scientific validity of the trial, along with an examination of what is clinically meaningful and ethically sound. This reassessment may result in the modification or termination of the trial, when evidence arises that the original assessment of benefit-risk to the patient is no longer favorable, or when the beneficial effect is so evident that it is ethically unsound not to give the treatment to all the patients.

This requires continuous observation of the clinical trial participants and a general oversight of the study conduct, which is ensured by trial monitoring, implemented and maintained by the sponsor. Monitoring ensures that the rights and overall wellbeing of the patients are safeguarded, including the setup of reporting processes that indicate if there is a serious safety matter requiring immediate attention, for example, an unexpected serious safety signal.

A safety signal, according to the Council for International Organizations of Medical Sciences (CIOMS), is information that arises during the trial suggestive of a causal association between the intervention (e.g. the medicinal product being tested) and an event or set of related events, either adverse or beneficial, which is judged to be sufficient to justify further action.

2. Code-breaking (unblinding)

In most pharmaceutical companies and contract research organizations (CRO) there is a procedure for unblinding an individual participant's treatment allocation while the trial is still ongoing. This happens if there is a medical emergency or serious medical condition during a clinical study when clinicians (investigators) cannot manage treatment without knowledge of the participant's study treatment.

To provide the documentation of the unblinding process is a regulatory requirement for study sponsors. This is usually done in a standard operating procedure (SOP) and there may be a guidance document or process flow to support the SOP.

In some organizations there is a call-center or automated system to manage the process on behalf of the study sponsor. In many organizations the study staff who may receive a call from a healthcare provider, particularly out-of-hours, will be trained to access the system, identify the treatment allocation and complete the required documentation of the case.

There may be situations where special product-specific circumstances exist, such as a known risk for overdose or a potentially harmful interaction with other medicines. These cases require the availability of an on-call doctor. The process would then be modified to ensure that a procedure is in place to ensure that an on-call doctor discusses the request for unblinding, prior to its execution. This is to protect the patient from any serious impact of removing a treatment that requires dose titration, or from a potential treatment that may react badly with the study medicine that has been taken.

Trial sites are provided with specific training on the unblinding procedure and 'patient cards' (or equivalent) are provided to the patient to enable emergency unblinding by anyone who may happen to see the patient such as clinicians not involved with the trial in emergency rooms. The patient cards carry information about the clinical trial in which the patient is taking part, including clinical trial name and number, the principal investigator's and institutions contact details, and the number to call for emergency unblinding. The patient is instructed to offer the card to any health worker other than those related to the trial in which they are participating.

Once the patient blind is broken, many patients will be required to discontinue the trial because of the possible bias. If there are too many cases where the treatment allocation has been disclosed during the study, the statistical integrity of the trial may be jeopardized. As additional safeguard, the trial-related clinical staff are often requested to discuss the need for knowing the patient's treatment in case they have to treat an emergency with the medical monitor/on-call doctor, prior to making the call to the automated system or helpdesk to break the blind.

3. Premature termination

There are many reasons why a study may be terminated early, but in most cases, these are pre-determined scientific or medical reasons. For example:

- emergence of a serious adverse reaction (ADR) where the risk to the patient

Materials modified and shared with permission from <https://www.eupati.eu>



is considered to be too great to continue, or

- a pre-planned interim analysis that clearly identifies superiority of one treatment or lack of effect of a treatment.

In most cases stopping a clinical trial requires the approval of senior medical management (in most pharmaceutical companies this is the chief medical officer). There may be cases where recruitment is proving so challenging that it is considered unethical to continue the trial. In these situations, it is unlikely that the scientific question will ever be answered. The regulatory authorities that approved the study need to agree to the termination.

In most organizations conducting clinical trials there are detailed instructions within the protocol on how interim analyses should be approached and what the statistical and decision-making processes will be to evaluate the data. Under no circumstances can an interim analysis be planned to have a quick check of the data during a trial. All interim analyses have to be clearly defined; including the time point when the data will be analyzed and why.

In the majority of cases an interim analysis is usually conducted to:

- identify whether there is an imbalance between the safety profile of the groups and therefore potentially placing some patients at risk,
- or to assess whether there is an imbalance in efficacy against a comparator treatment.

In all cases where the possibility exists to terminate a trial following an interim analysis, there has to be clearly defined stopping rules detailed in the protocol. Stopping rules need to be described in the informed consent documentation as well. To decrease any bias, the decision on whether the data produced from an interim analysis meets the stopping criteria should be made by an independent committee.

4. Data safety monitoring boards (DSMB)

Data safety monitoring boards (DSMB) also go under different names like Data Monitoring Committee (DMC), Data Monitoring Board or Data Safety Monitoring Committee. EU legislation uses the term DMC, but EUPATI will use the term DSMB based on the frequency of use seen, even if the original document used a different term.

A DSMB is a group of independent experts external to a study assessing the progress, safety data and, if needed critical efficacy endpoints of a clinical study. In order to do so a DSMB may review unblinded study information (on a patient level or treatment group level) during the conduct of the study. Based on its review the DSMB provides the sponsor with recommendations regarding study modification, continuation or termination. (1)

A DSMB may be set up for different reasons, here are some:

- In case of life-threatening diseases from an ethical point of view.
- In case of long-term trials even in non-life-threatening diseases for monitoring safety.
- In case of prior knowledge or strong suspicion that a treatment under consideration has

the potential to harm patients.

In case of a specific study design, e.g. in the context of pre-planned interim analyses for early stopping (either for futility or for positive efficacy) or in case of complex study designs where a possible modification of the study design based on unblinded interim data is intended.

4.1. Function of Data Safety Monitoring Boards (DSMB)

Although DSMB members are selected and appointed by the sponsor, all members should be completely independent of any ties to the trial that may affect their objectivity. Possible conflicts of interest (e.g. political, market, financial influences) should be taken into account. Any compensation paid to DSMB members should be reasonable to ensure there is no conflict of interest for members. Thus, any recommendation made by the DSMB should be free from bias.

DSMB operating procedures (how it works and communicates with other study participants – data centers, sponsor, etc.) must be established and described before the start of the trial.

Operating procedures should also describe how the integrity of the study with respect to preventing dissemination of unblinded study information is ensured. The size and composition of the DSMB depends on the type of trial to be executed. Members with clinical and statistical experience must always be included, and additional expertise in ethics and the specific disease area is often required. This is where community or patient representatives might be invited. Terms and conditions of appointment should be transparent and procedures should be clearly defined and well documented.

Although it seems a very logical step in protecting the participants' or patients' best interest, the inclusion of expert patients or other representatives of patient organizations in DSMB's is a relatively recent development. E.g. the European Community Advisory Board (on HIV/AIDS) has been dispatching members into DSMBs regularly since 2010. Patient representatives are equal participants of DSMBs, their work is also bound by strict confidentiality, however, they will import their experience of living with the given disease from a very special aspect – that of the patient themselves.

While in general, safety monitoring is the major task for a DSMB, other aspects of a clinical trial (e.g. trial integrity, design aspects) might also be assessed.

The DSMB will convene both open and closed meetings on a frequency laid out in the operating procedures at the start of the trial; this can be time based or when the pre-determined analysis points have been met, e.g. 50% patients reaching six months treatment. Usually only a sub-set of data that is most pertinent to the question being asked of the DSMB is analyzed and often still in blinded fashion, if the study is blinded. The operating procedures will contain rules that detail when the DSMB may request further data to be analyzed or the blind to be broken.

A report by the sponsor to the DSMB, (along with the full safety and efficacy data) may contain an open and a closed section, containing blinded and non-confidential data, and unblinded confidential data, respectively. The operating procedures should also clearly state the parties allowed to have access to any unblinded data.

Finally, the process of how the DSMB will arrive at a recommendation, must be documented. Ideally, recommendations should be made after reaching consensus, wherever possible, otherwise by vote. A recommendation that will result in modification, suspension or termination of a trial must be clearly supported by the reasons why the DSMB reached such a decision.

The proper communication of its recommendations is a major responsibility for a DSMB. If changes in the study conduct are recommended by a DSMB, sufficient information should be provided to allow the sponsor to decide whether and how to implement these recommendations. The implementation of any DSMB recommendation is solely the responsibility of the sponsor who has no obligation to follow them (in whole or in part).

Safety monitoring is an essential and integral part of any trial; nevertheless, not all clinical studies require a DSMB. DSMBs may be critical for studies intended to save lives, prevent serious disease progression or reduce the risk of a major adverse health outcome.

EMA Guideline on data monitoring committees Doc. Ref. EMEA/CHMP/EWP/5872/03 (1) describes the process for assessing the need for a DSMB. DSMBs are particularly important in studies where interim data analysis is required to ensure the safety of research participants.

As there are currently no accepted standards on what constitutes conclusive evidence, transparency in the decision-making process is paramount. A way to achieve this would be for the DSMB, after the trial

5. References

1. EMA guideline on data monitoring committees Doc. Ref. EMEA/CHMP/EWP/5872/03 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003635.pdf
2. US Food and Drug Administration. Guidance for Clinical Trial Sponsors - Establishment and Operation of Clinical Trial Data Monitoring Committees. Last accessed June 15, 2014: www.fda.gov/RegulatoryInformation/Guidances/ucm127069.htm

Lesson 21: Options for data collection and patient-reported outcomes (PROs)

Table of contents

1. Types of data capture
2. Paper CRFs
3. Electronic CRFs
 - 3.1. Regulatory requirements
4. Examples of direct data capture (DDC)

1. Types of data capture

Data in a clinical trial are generated and collected by the investigator, the study staff or directly by patients (called patient-reported outcomes - PROs).

This can occur in the traditional way on paper or in electronic form:

- CRF - Case Report Form - on paper including patient diaries or questionnaires.
- eCRF - Case Report Form – electronic.
- DDC - Direct Data Capture i.e. through electronic devices, data entered automatically into a database.
- PRO - Patient-Reported Outcomes - patient questionnaires and diaries - on paper.
- ePRO - Patient questionnaires and diaries – electronic via hand-held instruments like mobile phones or tablets.

2. Paper CRFs

Paper CRFs are designed for entering handwritten data. They are cheap to produce and allow the creation of direct copies through carbonless copy papers and faxing. Through new technology like optical character recognition (OCR) computers are able to ‘read’ the data written by the staff and to enter them automatically into a database.

Advantages:

- Low-tech - can be accessed anytime without a computer.
- Easier to correct if problems are discovered during study conduct.

Disadvantages:

- Volume of paper to store during and after the study.
- No automatic warnings concerning faulty data entry.
- Additional step required to enter data into database (one more step where errors could be made).

3. Electronic CRFs

Electronic CRFs are increasingly popular but are much more complicated to produce and need to adhere to strict regulations in Europe and the US:

- Europe: ICH GCP E-6, Section 5.5.3.
- US: FDA - 21CFR Part 11 and Guidance for Industry - Computerized Systems Used in Clinical Trials.

<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm125125.pdf>

The computer systems and the eCRF software must be validated and it must be possible to see every correction that has been made to the data entered (audit trail). It must be ensured that only authorized persons have access to the program and data. A back-up of the data has to happen regularly and automatically.

eCRFs in a study require that all investigator sites have sufficient and reliable access to computers and Internet. Intensive training of the site staff in using the eCRF is required, often also supported by a help- desk.

Advantages:

- The eCRF has automatic checks to reduce data entry errors and gives a warning if an error is made when entering the data (e.g. body temperature 370.6°C instead of 37.6°C). This also helps to avoid violation of protocol requirements.
- The entered data is immediately available to the sponsor who can see whether the site has entered data correctly and when this took place.
- Questions for further information on the entered data can be raised faster than in a paper-based system.

Disadvantages:

- potential technical problems related to eCRFs or to the study site's technical infrastructure
- study staff may need specific training or be capable of using a particular electronic system

3.1. Regulatory requirements

The EMA issued a reflection paper in 2010 (link no longer valid) summarizing what GCP inspectors will accept as electronic data capture. This document states, among many recommendations, that:

- There must be criteria under which NCAs will consider electronic records and signatures to be trustworthy, reliable and generally equivalent to paper records and handwritten signatures executed on paper.
- There must be a system validation to ensure accuracy, reliability, consistent intended performance (i.e. validation of the eCRF software), and the ability to discern

Materials modified and shared with permission from <https://www.eupati.eu>



invalid or altered records.

Likewise the FDA has issued very detailed and demanding rules under which conditions they accept electronic data capture. In a guidance for industry the FDA has explained how to use eCRFs in clinical trials. The Guidance for Industry recommends that:

- The protocol should identify when a computerized system will be used to create, modify, maintain, archive, retrieve, or transmit data.
- Documentation of software and hardware used should be kept with study records (in the Trial Master File (TMF)). An electronic record is the source document if the original observation is entered directly into a computerized system.

4. Examples of direct data capture (DDC)

Instruments that create data in a digital format are a usual source of information that gets directly linked to the study database. This sounds attractive but such connection must also follow strict rules and may be difficult to achieve. For example, when different laboratories are involved, they all have to adhere to the same data entry standards, normal ranges of their analytical methods, etc.

Some examples of direct data capture (DDC) are:

- Laboratory data –generated either by central or local labs.
- Electrocardiogram data (ECG).
- Central image reading – e.g. Magnetic Resonance Imaging (MRI) results.
- Electronic patient questionnaires / diaries.

More recently, mobile apps can provide data by monitoring for example physical activity or quality of sleep.

Lesson 22: Concept of Study Documentation

Table of contents

1. Study Documentation
2. What is source data?
3. What are Source Documents?
 - 3.1. Example of a Source Document
 - 3.2. Characteristics of source documents
4. What is a Case Report Form (CRF)?
 - 4.1. Example of a CRF
 - 4.2. CRF Content
 - 4.3. Regulatory requirements for CRFs
5. Source data verification
6. Essential Documents 6.1. Safe-Keeping
7. Clinical trial master file (TMF)
8. Archiving

1. Study Documentation

This topic describes the process for documenting a clinical trial. This includes:

- How all documentation generated for a study (at the site of the investigator and of the sponsor) gets compiled for a potential inspection by the competent authorities.
- How the data is collected.
- Where the data is recorded by the investigator.

A clinical trial is organized to generate information to enhance the knowledge about a new or existing treatment. Information means data. The quality of the data needs to be good to ensure that the conclusions drawn from this data are reliable. Therefore, data collection, handling and storage needs to be controlled. This means that there must be a process in place that ensures reliable data collection and control. In clinical trials, ‘source documents’ and ‘case report forms’ (CRFs) are used for this purpose.

2. What is source data?

The ICH-GCP Guideline defines ‘source data’ as:

‘All information in original records and certified copies of original records of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies)’.

This means that there are various types of data that are classed as source data. For example:

- Information which the investigator writes down in the patient’s record.
- Values in a lab result sheet or direct printout from a measuring device (e.g. analysis scale, spectrophotometer).
- Answers which a patient enters into a questionnaire.

3. What are Source Documents?

Source documents contain source data. They are original documents, data and records in which these original (source) data are noted. Examples include:

- Hospital records and clinical charts.
- Lab result sheets.
- Memos.
- Patients’ diaries or evaluation checklists.
- Medication dispensing records.
- Recorded data from automated instruments, ECG, X-ray.
- Informed consent form.

Copies are only recognized as ‘source documents’ if they are certified after verification as being accurate.

3.1. Example of a Source Document

This example of a source document is a form that can be used when a study participant (subject) gives a urine sample:

Urine sample collection and transfer sheet (Requisition Form for semiquantitative urinalysis and microscopy as per study protocol)

Group _____ Protocol Day: _____ Date: _____

		Urine collection	Urine transfer to laboratory	Urine transfer acknowledgement	Urine transfer acknowledgement	Comments
Initials	Subject No	Time	Time	Initials deliverer	Initials receiver	

- The study nurse will enter their initials, subject number, and time of urine

Materials modified and shared with permission from <https://www.eupati.eu>



collection. This information is noted for the first time - it is source data.

- The person transporting the urine samples to the lab will note the time of transport and sign with their initials so that it is clear who did this and when.
- The lab technician who receives the urine samples signs the form and notes the time.

In this way the process is completely documented including information on who had responsibility for the individual steps along the way (providing an audit trail). The comment column allows specific information to be entered when something unusual occurred. This may help to better interpret the results of the urine analyses.

3.2. Characteristics of source documents

Having good source documents in a clinical trial is vital for the quality of the study. Therefore, they need to fulfil certain characteristics:

- **Attributable** - it must be clear to which participant they belong.
- **Legible** - they must be readable.
- **Contemporaneous** - they must be noted immediately after the data is generated.
- **Original** - they must be the original. Accurate - they must be reliably correct.
- **Originator information** – who entered the data and when.

If data are collected via an electronic mean, the machine (and software) must be identifiable and be calibrated and validated.

4. What is a Case Report Form (CRF)?

The ICH-GCP Guideline glossary defines a ‘case report form’ (CRF) as:

‘A printed, optical, or electronic document designed to record all of the protocol required information to be reported to the sponsor on each trial participant.’

For an inspection the source data must be made available in a computerized format compliant with the EMA Q&A: Good clinical practice (GCP)

(<http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation>

[/q_and_a/q_and_a_detail_000016.jsp](http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/q_and_a/q_and_a_detail_000016.jsp)). (link expired) For this purpose, CRFs are prepared. The investigator and study staff are requested to enter the required data from a patient’s source documents into a patient-specific CRF. It is very important that no mistakes are made in this transcription process and that the information provided in the CRFs is complete.

4.1. Example of a CRF

This is an example of a CRF in which information like the weight of a patient, their body temperature, breathing frequency, blood pressure, and pulse frequency are captured.

1.1											
Date	<input type="text"/>	¹	24-hr Time	<input type="text"/>	²						
Weight	<input type="text"/>	³	kg	<input type="text"/>	⁴						
Temperature	<input type="text"/>	⁵	C	<input type="text"/>	⁶	Oral	<input type="text"/>	⁷			
Respiratory Rate	<input type="text"/>	⁸	breaths/min	<input type="text"/>	⁹						
Data Group #2											
2.1											
Date	<input type="text"/>	¹⁰	24-hr Time	<input type="text"/>	¹¹						
Systolic/Diastolic	<input type="text"/>	¹²	/	<input type="text"/>	¹³	mm Hg	<input type="text"/>	¹⁴	Supine	<input type="text"/>	¹⁵
Pulse	<input type="text"/>	¹⁶	beats/min	<input type="text"/>	¹⁷	Supine	<input type="text"/>	¹⁸			
2.2											
Date	<input type="text"/>		24-hr Time	<input type="text"/>							
Systolic/Diastolic	<input type="text"/>		/	<input type="text"/>		mm Hg	<input type="text"/>		Standing	<input type="text"/>	
Pulse	<input type="text"/>		beats/min	<input type="text"/>		Standing	<input type="text"/>				

4.2. CRF Content

Here are some examples of data that are typically collected in a CRF:

- Inclusion and exclusion criteria (features a patient needs to fulfil to get enrolled into the study).
- Demographics (the patient's personal information in an anonymized way).
- The patient's medical history and the result of the physical examination by the doctor.
- A description of the diseases that the patient currently has and the medicine they are taking.
- Efficacy and safety parameters (measurements that show the effect of the medicine and the potential side effects the patient is experiencing).
- Which medicine, usually in a blinded fashion, was administered. Termination of trial (the result of the final examinations and confirmation of the patient's completion of the study).

4.3. Regulatory requirements for CRFs

CRFs are prepared once the protocol is agreed and before the study starts (i.e. before the first participant is enrolled). CRFs must provide the format for capturing all information that is expected to be collected in the study, as outlined in the protocol.

The ICH Guidance on Statistical Principles for Clinical Trials

(http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf) establishes that:

- Data collected must be in full accordance with the protocol.
- CRF must be established before study starts.
- Data collected should enable the analysis to be performed.
- Identification of protocol compliance. - The CRF is also a good tool to check how well the study staff are following the protocol. If data is missing, measurement dates are not in line with the requested protocol timelines and data differ between the source documents and the CRF, this indicates that errors were made and need to be corrected.
- Participant codes must allow identification of all data reported for each participant – unambiguously.

The sponsor needs to ensure that there is a numbering system in place that gives a unique number to each patient who is screened for a study and another number for those patients that can ultimately get enrolled into the study. This number does not allow the identification of the patient. Only the investigator has a list that connects this number with the name and address of the patient. The study monitor, checks the performance of the study site, has the ability to look into the source documents.

5. Source data verification

The study monitor has to compare the content of the CRFs with the content of the source documents. This is called ‘source data verification’. The monitor also checks if all the required data of the participant is present in the CRF entries, and that data is consistent with the source documents. Previously, the rule was that 100% of all CRF entries were to be compared with the source data. Nowadays, the intensity of source data verification can be reduced. The concept of ‘risk-based monitoring’ allows the sponsor to decide to which degree the monitor needs to check individual types of data. Typically, the data that is checked most thoroughly is:

- The signed Informed Consent documentation (often called Informed Consent Form – ICF).
- The measurements that are most critical for answering the study’s scientific question(s),
- The adverse events.
- The inclusion/exclusion criteria adherence

6. Essential Documents

According to the ICH Guideline, Section 1.23, ‘essential documents’ are defined as:

‘Documents which individually and collectively permit evaluation of the conduct of a study and the quality of the data produced’.

The GCP Guideline gives under section 8 (Essential documents for the conduct of a clinical trial)

a minimum list of all documents that need to be generated, filed and archived by the investigator and by the sponsor. Some of these documents need to be filed in both places like the protocol or the Investigator's Brochure (IB). Some other documents are only to be filed by one party, for example:

The investigator files:

- The participant identification code list to document that investigator/institution keeps a confidential list of names of all participants allocated to trial numbers on enrolling in the trial. Allows investigator/institution to reveal identity of any participant.
- The signed Informed Consent documentation.
- Advertisement for participants recruitment (if used) to document that recruitment measures are appropriate and not coercive.

The sponsor files:

- Documents concerning the label(s) attached to investigational medicinal product container(s) to document compliance with applicable labelling regulations and appropriateness of instructions provided to the participants.
- Samples of the labels.

6.1. Safe-Keeping

The investigator must ensure that the 'Investigator Site File' (ISF) (officially the 'Investigator Trial Master File (TMF)') is safely archived in such a way that the documents and data can be found and read at a later time if questions come up or an inspection by the competent authorities takes place.

Systems must be in place to ensure the safe long-term storage of essential documents:

- At investigator sites, and
- by, or on behalf of, the sponsor.

Different EU directives/regulations as well as international guidance documents give different retention periods for trial related documentation.

7. Clinical trial master file (TMF)

The EMA Reflection paper on GCP compliance in relation to trial master files (paper and/or electronic) for management, audit and inspection of clinical trials (EMA/INS/GCP/636736/2012) (Link no longer valid) defines the minimum time the sponsor and investigator need to retain the essential documents relating to a clinical trial. The duration varies by situation and member state.

Trial master files (TMF) should be established at the beginning of the trial, both at the investigator/institution site and at the sponsor's office. A TMF is the collection of documentation that allows the conduct of the clinical trial, the integrity of the trial data and the compliance of the trial with GCP to be evaluated (monitoring by the sponsor (audits) and inspection by member

states). The TMF is normally composed of a sponsor TMF, held by the sponsor organization, and an investigator TMF held by the investigator(s). The TMF kept by the investigator and that kept by the sponsor may have a different content if this is justified by the different nature of the responsibilities of the investigator and the sponsor. The sponsor needs to make sure that no documents in the TMF contain information that could be used to identify the participant.

The sponsor's TMF contains:

- The 'study file' - this includes all of the sponsor's documents.
- The 'site file' - this is documentation that demonstrates the sponsor's involvement in getting study authorization for the site and overseeing the trial activities.

For every investigator site, the sponsor collects a copy of all documents in the 'Investigator Site File' (ISF), with exception of those that contain patient identifying information. The ISF is kept at the investigator's site.

You can download an excerpt from the minimum list of essential documents (for a TMF) in section 8 ICH- GCP guideline. For the complete list see: (specific link no longer valid) website www.ich.org

The various documents are grouped in three sections according to the stage of the trial during which they will normally be generated:

- 1) before the clinical phase of the trial commences,
- 2) during the clinical conduct of the trial, and
- 3) after completion or termination of the trial.

A description is given of the purpose of each document, and whether it should be filed in either the investigator/institution or sponsor files, or both.

8. Archiving

Archiving is an important means to ensure that data and documents can be retrieved at a later stage. All essential records must be appropriately archived, in conditions that include:

- Adequate and suitable space.
- Access restricted to authorized named individuals.
- Secure storage facilities with appropriate environmental controls and protection from physical damage.

Storage media must ensure documents remain complete, legible and available throughout the required period of retention. Electronic, magnetic, optical or non-indelible media must be controlled to ensure any alterations can be traced (audit trail). For electronic storage media it is particularly important to consider the lifespan of such media (e.g. tapes, CD, SSD drives) and the accessibility of digitally stored data which were stored using specific formats, software versions or technical equipment (downward compatibility of newer systems/techniques/software). For example: are you still able to access the information you stored only a few years ago

on your floppy discs?

Systems used for transfer of original documents to another media for storage must be validated (their reliability must be confirmed), and the availability of the equipment to convert records to readable format must be ensured.