Cochrane LA & Puvaneswaralingam S (2012). Clinical Research for Beginners – The Importance of Planning. Scottish Universities Medical Journal. 1 (2). p.154-164

# Clinical Research for Beginners - The Importance of Planning

**Dr. Lynda A Cochrane (Dundee Epidemiology and Biostatistics Unit) & Shobitha Puvaneswaralingam (4th MBChB; BMSc)**

Correspondence to: Lynda Cochrane : l.cochrane@dundee.ac.uk

## ABSTRACT

While the identification of the individual who said "Failing to plan is planning to fail!" is in doubt, the wisdom of the words is not especially in the context of undertaking clinical research. There appears to be a view in some quarters that collecting a set of data, for example on people with a particular condition, and carrying out a series of analyses to try to find something of statistical significance constitutes a scientific investigation - it does not! A clinical research study must have a clearly defined aim with a comprehensive plan including a statement of the population of interest, which data are to be collected, how they will be measured, by whom, when and with what, as well as details of how subjects are to be allocated to interventions (if any). The purpose of this article is to give an introduction to some of the issues involved in constructing such a plan. It is by no means a definitive guide and further details can be found in the reference list.

**Key Words**: clinical study; design; outcomes; sample

## Planning a study

Most statisticians will have encountered the following situation: a researcher, having collected a set of data, consults the statistician for the first time and asks "Which test do I use?". What, you might enquire, is the problem? Isn't that what statisticians do - advise on and carry out data analysis? Well, yes, that's true. But it is only part of the service they provide. One of the most important roles of a statistician is to ensure that a research study is designed so that its aims can be achieved. Without such planning, there is a possibility that the study will be inadequate and thus fail to meet the proposed aims.

Would anyone construct a bridge by gathering some materials and then asking a construction engineer which tools to use in order to put them together? Of course not: we all understand the need to plan the project and use scientific principles to determine which type of design will best meet the requirements of the structure, such as the load it is to carry, the environment in which it is situated and the funds available. The same principles apply to a clinical research study. We must identify a clear aim, indicate the population of interest and the inclusion and exclusion criteria for potential study participants, define the measurements we need to make and when we will make them, establish the methods of data analysis and so on. The importance of involving a statistician in the initial design, as well the analysis, of a research study is highlighted by Professor Douglas Altman[1].

The main steps involved in planning a study are listed in Figure 1 and discussed individually below.

Cochrane LA & Puvaneswaralingam S (2012). Clinical Research for Beginners – The Importance of Planning. Scottish Universities Medical Journal. 1 (2). p.154-164

Figure 1: Components of a Study Plan

- Aim(s)
- Population
- Interventions
- Outcomes
- Data collection: measuring outcomes (when, by whom, how, to what level of accuracy)
- Confounding factors
- Inclusion and exclusion criteria
- Sampling strategy
- Study design
- Sample size
- Compliance
- Data storage and management
- Analysis

## Aims

The first step in planning a study is to identify a clear, achievable and ethical aim. All studies need to have a purpose and aim to develop knowledge or understanding in a particular area.

Examples include:

A1) Compare changes in the blood pressure of hypertensive patients achieved with drug therapy alone with those with drug therapy and exercise
A2) Compare the outcomes of conservative and surgical treatment of scaphoid lunate fracture
A3) Describe the natural history of Duchenne muscular dystrophy (DMD)
A4) Identify the risk factors for liver cirrhosis
A5) Investigate factors linked to lower limb amputations in patients with diabetes mellitus (DM)

Examples of inappropriate, vague or unethical aims include:

A6) Monitor changes to the brain prior to the onset of Alzheimer's disease
A7) Identify the outcomes of regularly exceeding the recommended dose of paracetamol
A8) Assess prescribing patterns in Dundee

## Identifying the population

This is the set of patients about which we wish to make an inference. For example, in A2) above, the population is patients who have suffered a scaphoid lunate fracture and in A3) it is patients with a diagnosis of DMD. Identifying the population is not always so straightforward. For example, will patients with DM (A5) include Type 1, Type 2, gestational, MODY, LADA, Type 3, Type 1.5 etc? These different groups are likely to differ in, for example, age, drug treatment and co-morbidities and this heterogeneity would complicate the investigation.

## Define the interventions

An intervention is any action that is done to the subject, or to his or her environment. This can include, for example, a drug treatment (including placebo), surgery, wearing a support device, counselling or a combination of two or more treatments. The interventions for the hypertension study could be a six-month programme of bendroflumethiazide, 2.5mg early in

the morning (the drug therapy) and 30 minutes of walking at a steady pace (3mph) five days each week (the exercise programme).

## Identifying the outcome

Outcomes are endpoints or measures of the response to an intervention. The outcomes for aims A1), A4) and A5) above are clear: systolic & diastolic blood pressure, liver cirrhosis and the occurrence of a lower limb amputation, respectively. For A2), it would be useful to have a score constructed from several aspects of wrist function, including subjective (e.g. pain) and objective (e.g. flexion) measures such as the Modified Gartland and Werley Score[2]. The natural history of a disease such as DMD could be described by the different aids required (such as limb supports and a wheelchair), drug therapies and the time to these events. The occurrence, severity and time of onset of complications such as chest infection and osteoporosis would also be of interest.

## Data collection: measuring outcomes

There are several issues to consider when measuring the outcome. How will it be measured? When will it be measured and by whom? What is the level of accuracy and how valid and reliable is the measurement of the outcome? How will it be recorded and the data stored? Who will take responsibility for data management? These are particularly important matters when data are collected by more than one person and/or at more than one site.

Wherever possible, data should be measured and recorded as accurately as possible. It is tempting to group observations but this can be misleading and limiting. Suppose, for example, a researcher wishes to test his hypothesis that high heel height leads to back pain. Should he classify heels simply as '*high*' and '*not high*", as '*high*', '*medium*' and '*low*' or something else? Ideally, shoe heel height should be measured with a tape measure, at the back of the heel and recorded in millimetres. Judgements such as high and low are subjective: someone who regularly wears flat shoes might regard 30mm to be a high heel, whereas a stiletto-heel wearer might regard this as low.

The measuring device itself must be accurate. For example, a set of scales that give a reading that is higher than the true weight should not be used. Also of interest are the validity of the measuring instrument (does it measure what it claims to measure) and its reliability (how consistent is it when measurements are repeated).

## Confounding factors

Variables that are related to both the outcome of a study and the intervention can distort the effect of the intervention. These are known as confounding factors. There is some evidence, for example, that ethnicity can affect the response to enalapril, with black patients responding least well[3]. Therefore, the effectiveness of this anti-hypertensive drug would be reduced if there were a preponderance of black patients in the population studied. It is important to identify any such confounding factors during the planning phase and include them as independent variables.

## Inclusion and exclusion criteria

As the names suggest, inclusion and exclusion criteria identify who will be included or excluded from the sample. Patients who could benefit from the intervention are described by the inclusion criteria. Those for whom the intervention is inappropriate or could be dangerous, or who have co-morbidities that could mask its effect, are identified by the exclusion criteria. Inclusion criteria for the investigation of hypertensive patients (A1) would be patients with hypertension defined by the NICE Clinical Healthcare Guideline 34[4]. It would be reasonable to exclude those with isolated systolic hypertension and those with secondary causes of hypertension such as renovascular disease, genetic causes and

pregnancy. Also excluded would be patients with any of the contraindications to taking the intervention drug as specified in the British National Formulary (BNF)[5].

## Sampling the population

Usually it is impossible to study all of the members of a population and a subset or sample of the population is therefore chosen[6]. The simplest technique is random sampling. First, we define the population (e.g. patients in Tayside with cerebral palsy) then give every member of the population a number (from 1 to N inclusive, where N is the size of the population). For a sample size n, we select n random numbers. The corresponding members of the population are selected.

Sampling is a vital step in any research and governs any inferences that can be made. For example, if we investigate factors linked to lower limb amputations in patients with Type 2 DM, then we can make conclusions only about this group. The findings cannot be extrapolated to patients with other types of DM. If the population is to include other DM patients then the sample should be chosen to reflect the prevalence of the respective types in the general population. This is known as stratified sampling.

Often it is either not possible or not practical to select a random sample (e.g. if the population cannot be enumerated). In such cases, a clinician might choose to study a sample of patients in his/her clinic. Even if this sample itself is selected randomly, this does not constitute a truly random sample of the population: it is a random sample of a subset of the population that has not itself been chosen randomly. Such selections are referred to as convenience samples.

## Types of study design

There are many different types of study and each has merits in particular situations[7]. In a prospective study, subjects are selected from a population and analysed for a defined future outcome. In contrast, a retrospective study is an analysis of existing data. A study is said to be experimental if the effect of an intervention (e.g. a drug treatment or exercise program) otherwise it is an observational study. A study is described as cross-sectional if measurements are made at only one time point whilst a longitudinal study analyses multiple time points. An analytical study is one in which the aim is to analyse the data gathered in order to make an inference about the effect of an intervention on an outcome variable. In a descriptive study, the data are summarised using descriptive statistics (e.g. measures of centre and spread, frequencies) without consideration of the effects of one or more of the variables on the others.

One of the most widely known designs is the randomised controlled trial (RCT). A sample of subjects is selected from the population and allocated randomly to one of two or more groups (or arms) of the trial. One of the treatments is a control, which could be an existing treatment, a placebo or no treatment. Wherever possible, trials should be double blinded such that both the subjects and the researchers are unaware of the treatment allocations. However, although ideal, this may be impossible, for example when one of the treatments is counselling and the other is a drug therapy.

A parallel group design is a RCT in which subjects are allocated randomly to either the treatment or the control group. By allocating subjects completely randomly, the expectation is that any known or unknown factors that could affect the outcome – other than the treatment(s) – would be equally distributed between each arm of the trial. This is not necessarily the case and one way of dealing with this is to use a matched design[8] in which the subjects in each arm are matched for the factors known to affect the response to the treatment (e.g. age, BMI). This design requires fewer subjects than a random design

because some of the variability of the response has been controlled for. Further efficiency can be achieved by using a within-subjects design, in which individuals are allocated to both arms of the trial (simultaneously or consecutively). As a result, the inter-subject variability is eliminated because each subject acts as his/her own control. Interventions that can be applied simultaneously include topical treatments applied to each leg. If treatments are consecutive (e.g. the comparison of two drugs to relieve chronic pain), care should be taken to avoid a carry-over effect between treatments by allowing a washout period. In addition, the order in which treatments are applied should be randomised to avoid any order effects.

Cross sectional studies provide information about a population of interest at a particular moment in time. Examples include surveys to estimate the prevalence of a disease and studies to investigate the reliability of a measuring instrument.

## Identifying risk factors

Some of the most commonly reported studies involve identifying risk factors for disease. It would be unethical to deliberately subject individuals to something that could be harmful, although instances have been known (e.g. in the 1960s, French soldiers were deliberately exposed to radiation during nuclear tests in order to investigate their physiological and psychological responses)[9]. There are thus two primary ways of assessing risk factors for various diseases: prospective cohort and retrospective case-control studies. In a prospective cohort study, a group of healthy individuals are monitored until they develop the disease under investigation. These tend to be long, large and therefore expensive but provide the most reliable results. Case-control studies involve comparing subjects with the disease (cases) with individuals who do not have the disease (controls) but otherwise are similar (e.g. same gender, age, co-morbidities etc.). These are shorter studies and less expensive but less reliable than prospective cohort studies. Despite its shortcomings, this type of design has generated some important findings, most notably the association between tobacco smoking and lung cancer found by Professor Richard Doll and his team[10].

## Sample size

Another question frequently asked is how many subjects are needed in a study. There is no simple answer to this and the sample size depends on several factors[11]. First, how variable is the outcome? The required sample size required for the study increases according to the variability of the data. Estimates of likely variability of the data can be obtained either from existing literature or by carrying out a pilot study, which tests the feasibility of the main experiment and provides useful information about measures of centre and spread. Second, there is the effect size. This is a measure of the size and direction of the effect of a treatment (intervention). For continuous outcomes, usually this is expressed as a proportion of the standard deviation of the response (SD): that is to say, it is calculated as (change in outcome with treatment - change in outcome with control) ÷ SD. This removes the effect of scale and allows comparisons to be made between different studies. When the outcome is binary (e.g. did the patient develop a hospital acquired infection: Yes / No), one measure of effect size is the number of subjects that would need to be treated to prevent one outcome (NNT, e.g. the occurrence of one infection) and this is known as the 'number needed to treat'. Another measure of effect size in studies with binary outcomes is the odds ratio (OR). This is the ratio of the odds of the outcome observed with one treatment divided by the odds observed with another e.g. the odds of survival to 1 year with two regimes of chemotherapy in patients with pancreatic cancer. Provision should also be made for patients who drop out of the study.

## Compliance

Compliance, or lack of it, is one of the hazards of clinical studies: patients do not always follow the instructions they are given. This is especially likely if the intervention is inconvenient or unpleasant. There are two approaches to the subsequent analysis of the data: per intention to treat (ITT) or per protocol (PP, sometimes referred to as modified intention to treat). In the former, data are analysed according to what the intention (plan) stated and in the latter patients who do not adhere to the protocol are omitted from the analysis. For example, suppose there are two arms of a trial in which Group 1 follow a low-fat diet and walk for 20 minutes each day and Group 2 follow a low-fat diet plan. If a patient in Group 1 follows the diet but does not exercise then in the ITT analysis the patient would be included in Group 1 whereas the per protocol analysis would exclude him/her from the analysis. There are some repercussions that can arise with PP analyses[12] and many statisticians prefer the ITT option[13].

## Data storage and collection

Unless data are accurate, valid and reliable, the results of a medical research study will be unreliable. Security, including the protection of patient identifiable data, is of critical importance when dealing with clinical information. Many institutions have a specialised unit which co-ordinates the collection, storage and management of research data and this is the preferred option.

## Analysis

Details of the analyses to be undertaken and the statistical tools to be used should be specified in the study plan. This will be the subject of a subsequent article.

## Guidelines

Two sets of guidelines are available for reporting observational studies (STROBE[14]) and RCTs (CONSORT[15]). You might find they help you to think through your study design.

## Examples

It is important to become familiar with the above clinical research terms/concepts when planning any research project. Indeed, one of the best ways to achieve this is by reading and looking at examples of real clinical research program. Outlined in Figure 2 is a suggested design for a study to investigate the relationship between exercise and blood pressure in patients with hypertension.

For her BMSc project, SP carried out a study entitled “Does weight loss improve cardiovascular morbidity in patients with type 2 diabetes mellitus?” (Figure 3). This was a retrospective cohort study of linked data extracted from databases held by the University of Dundee’s Health Informatics Centre (HIC). It became clear during the course of the analysis that there were a number of problems with the design (Figure 4). Suggested ways of overcoming these are shown in Figure 5.

## Figure 2: Design of Exercise and Hypertension Study

### Aim(s)

To investigate the relationship between exercise and blood pressure in Tayside patients with hypertension

### Population

Patients with hypertension, defined as in the NICE clinical guideline 34[3]

### Interventions

Counselling relating to the benefits of exercise and the provision of a pedometer to measure the number of steps taken per day

### Outcomes

Systolic and diastolic blood pressure

Total number of steps taken

### Data collection

1) Measurements to be made by Researcher A, a trained HCP, at the start of the intervention and 6 months later:
   a) Systolic and diastolic blood pressure (according to NICE clinical guideline 34[3])
   b) Weight: kg to 1 dp
   c) Height: m to 1 dp
2) Number of steps to be recorded daily by the patient, on forms produced by Researcher A. Total taken over 6 months to be calculated by Researcher A
3) BMI: calculated by Researcher A according to the formula mass (kg)÷(height (m))$^2$
4) High and low density lipids: mmol/L to 2 dp obtained from hospital laboratory assay, at the start of the intervention and 6 months later (blood to be collected by Researcher A).
5) Gender, date of birth, date of diagnosis as noted in the patients' records

### Confounding factors

Secondary causes of hypertension (e.g. renal failure, hyperthyroidism[?])

### Inclusion and exclusion criteria

Exclude: Cases of secondary hypertension[?]

Include: Patients with hypertension (defined as in the NICE clinical guideline 34[3]) aged 20-65 years.

### Sampling strategy

A simple random sample of eligible cases will be selected from the database of all Tayside patients with hypertension.

### Study design

Prospective experimental

### Sample size

Assuming SDs of systolic and diastolic blood pressure to be 20 and 10mmHg respectively[16], the sample size required to detect reductions of 5 and 2.5mmHg, respectively, at the 5% significance level with 80% power, is 128. Assuming a drop-out rate of 30%, the number of cases required is 192.

### Data storage and management

Data storage, extraction, anonymisation and record linkage to be performed by XYZ Data Solutions. All data will be stored in a data safe haven and accessed via a password-protected system.

[?] These would be determined from research and identified by WHO ICD10 codes in the patient record

## Figure 3: Design of Weight Loss and CVD Study

**Aim(s):** To investigate the relationship between weight loss and cardiovascular morbidity in Tayside patients with type 2 DM.

**Population:** Patients in Tayside diagnosed with type 2 DM

**Interventions:** None

**Outcomes:** CVD

**Data collection**

NICE guidelines[11] recommend annual risk factor assessment for CVD; therefore, values for each of the following should be available on a yearly basis from diabetic clinic appointments:

Weight: kg to 1 dp as measured by a trained HCP at outpatient appointments

Height: m to 1 dp as measured by a trained HCP at outpatient appointments

BMI: calculated by researcher according to the formula mass (kg)÷(height (m))$^2$

Systolic and diastolic blood pressure: mmHg to 1 dp as measured by a trained HCP according to NICE guidelines[1]

Cholesterol: mmol/L to 2 dp (measured lipid biochemistry results)

Drugs: extracted from electronic records of prescriptions dispensed in the community

CVD: As recorded in inpatient records. Identified through ICD-10 codes of CVD

**Confounding factors:** Unintentional weight loss

**Inclusion and exclusion criteria**

Include: Patients aged 35-65 years. Those younger than 35 are at risk of having type 1 DM rather than type 2 DM. Patients over the age of 65 are at higher risk of cardiovascular disease due to natural ageing, which would obstruct the effect of any weight loss, and are also at higher risk of unintentional weight loss due to ill health.

Exclude: Patients with unintentional weight loss, patients prescribed insulin within 1 year of diagnosis because this may suggest possible type 1 DM.

**Sampling strategy:** All eligible Tayside patients with type 2 DM diagnosed in 2005 to be selected. This would allow observation of both short- and long-term outcomes over the most recent five-year period.

**Study design:** Retrospective cohort study

**Sample size:** All who meet the criteria

**Data storage and management**

Electronic data including hospital admissions, dispensed medication and laboratory results from immunology, virology and biochemistry will be extracted and anonymised by HIC according to HIC Standard Operating Procedures. Ethical approval will be obtained from the Tayside research ethics committee and databases will be used in accordance with procedures approved under the Data Protection Act and Caldicott Guardian.

**Figure 4: Problems with Weight Loss and CVD Study**

**Measuring outcomes**

Outcomes found in electronic patients records rely upon the completeness and accuracy of the record keeping of others. Since we were working with anonymised patient data, the source documents could not be traced to clarify outcomes and data, and records with dubious data were thus deleted.

**Confounding factors**

In addition to intentional weight loss, drugs used to treat hypertension and hypercholesterolaemia should have been included: weight gain can make it difficult to control blood pressure and cholesterol and therefore these patients might be prescribed drugs to manage their condition. This could lead to an observed reduction in blood pressure and cholesterol with weight gain.

**Inclusion and exclusion criteria**

Excluding conditions that do not necessarily cause weight loss (such as benign neoplasms) and failing to exclude conditions that can cause weight loss (such as inflammatory bowel disease and cirrhosis) resulted in a specifically selected healthier sample of younger age in which cardiovascular outcomes were unlikely to develop.

**Sample size**

No sample size calculation was performed. All eligible cases were included and complete data were available for 570 patients.

## Figure 5: Revised Design for Weight Loss and CVD Study

**Aim**
To investigate the relationship between weight loss and cardiovascular morbidity in Tayside patients with type 2 DM.

**Population**
Patients in Tayside diagnosed with type 2 DM.

**Interventions**
None

**Outcomes**
CVD as recorded in outpatient and inpatient records of primary and secondary care.

**Data collection**
NICE guidelines[17] recommend annual risk factor assessment for CVD; therefore, values for each of the following should be available on a yearly basis from diabetic clinic appointments:
**Weight**: kg to 1 dp as measured by a trained HCP at outpatient appointments
**Height**: m to 1 dp as measured by a trained HCP at outpatient appointments
**BMI**: calculated by researcher according to the formula mass (kg)÷(height (m))$^2$
**Systolic and diastolic blood pressure**: mmHg to 1 dp as measured by a trained HCP according to NICE guidelines[1]
**Cholesterol**: mmol/L to 2 dp (measured lipid biochemistry results)
**Drugs**: extracted from electronic records of prescriptions dispensed in the community
**CVD**: As recorded in inpatient records. Identified through ICD-10 codes of CVD

**Confounding factors**
Drug therapy for CVD risk factors (blood pressure and cholesterol).

**Inclusion and exclusion criteria**
Include: Patients diagnosed with type 2 DM from 1/1/2002 to 31/12/2012
Exclude: Patients prescribed insulin within 1 year of diagnosis as it may suggest possible type 1 DM, and those with co-morbidities causing unintentional weight loss such as inflammatory bowel disease and cirrhosis. A full list of exclusion conditions to be established in discussion with a senior clinician.

**Sampling strategy**
All patients with type 2 DM in the Tayside area, who meet the inclusion and exclusion criteria, will be included in the sample.

**Study design**
Retrospective cohort

Cochrane LA & Puvaneswaralingam S (2012). Clinical Research for Beginners – The Importance of Planning. Scottish Universities Medical Journal. 1 (2). p.154-164

## References


1. Altman DG. The scandal of poor medical research. *BMJ* 1994;308(6924): 283-284.
2. Gartland JJ Jr, Werley CW. Evaluation of healed Colles' fracture. *J Bone Joint Surg Am* 1951;33:895–907.
3. Exner DV, Dries DL, Domanski MJ, Cohn JN. Lesser response to angiotensin-converting-enzyme inhibitor therapy in black as compared with white patients with left ventricular dysfunction. *N Engl J Med* 2001;344:1351-1357.
4. National Institute for Health and Clinical Excellence. Hypertension: management of hypertension in adults in primary care (partial update). NICE Clinical Guideline 34. London: NICE; 2006.
5. British National Formulary. London: Joint Formulary Committee; March 2012.
6. Altman DG. Practical Statistics for Medical research. London: Chapman and Hall; 1991.
7. Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley & Sons; 1986.
8. Bland JM, Altman DG. Matching. *BMJ* 1994;309: 1128.
9. FRANCE 24 International News 24/7[homepage on the Internet]. France [updated 2010 Feb 16; cited 2012 May 13]. Available from: http://www.france24.com/en/20100216-soldiers-deliberately-exposed-nuclear-tests-says-report.
10. Doll R, Hill AB. "Smoking and Carcinoma of the Lung". *BMJ* 1950;2 (4682): 739–748.
11. Chow S-C, Shao J, Wang H. Sample size calculations in clinical research. New York: Marcel Dekker Inc; 2003.
12. Abraha I, Montedori A, Romagnoli C. Modified intention to treat: frequency, definition and implication for clinical trials [abstract]. Sao Paulo, Brazil: XV Cochrane Colloquium, 2007;86-7.
13. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670-4.
14. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996; 276(8):637-639.
15. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, for the STROBE Initiative: Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:806-8.
16. Haider A, Larson M, Franklin S, Levy D. Systolic blood pressure, diastolic blood pressure, and pulse pressure as predictors of risk for congestive heart failure in the Framingham Heart Study. *Ann Intern Med* 2003;138:10-16.
17. National Institute for Health and Clinical Excellence. Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease. Clinical Guideline 67. London: NICE; 2008.