Study Guide 1: How to Devise a Research Question and Choose a Study Design

Dr David Chinn,
Research, Innovation & Knowledge (RIK) Office,
Queen Margaret Hospital, Dunfermline, Fife.
david.chinn@nhs.scot
Alternative contact: Prof Frances Quirk frances.quirk@nhs.scot 01383 623623 (ext 20941)

Contents Page

1 Overview and learning outcomes 1
2 Introduction 2
3 Devising the research question 2
4 Study types 4
5 Qualitative studies 5
6 Quantitative studies 7
6.1 Cross-sectional 8
6.2 Case control 9
6.3 Cohort 10
6.4 Cross-over 11
6.5 Randomised controlled trial 12
7 Bias 13
8 Exercises 15
9 Further reading 15
Appendix: answers to exercises 16
Glossary 18

(1) Overview and learning outcomes

This guide is directed at novice researchers who are in the early stages of planning a research project, in particular to anyone undertaking an MSc course. It is a vitally important step in the overall process as the study design is critically dependent on the research question. After reading this guide you should be able to:

- Identify the features of a good study
- Describe the components of a valid research question
- Understand the difference between a qualitative and quantitative approach
- Describe the strengths and weaknesses of different study designs
- Choose a study design appropriate to a research question
- Identify potential sources of bias which will be a threat to success of the project

Associated NHS Fife study guides:

2 How to write a protocol
3 How to critically appraise the literature
4 How to Apply for R&D Management Approval and for a 'Favourable' Opinion from an Ethics Committee
9 An introduction to qualitative research
10 An introduction to medical statistics
11 How to calculate sample size and statistical power
16 How to achieve success with your dissertation

(2) Introduction

Most health care practitioners will at some time in their career have been involved in undertaking an audit or service evaluation. These activities are defined as:

> Audit: *a quality improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of change.*
>
> Service evaluation: *a review process undertaken solely to define or judge current service with the intention of benefiting those who use it.*

Research does not compare activity against explicit criteria or judge current service provision but, instead, seeks to derive generalisable new knowledge *'by addressing clearly defined questions with systematic and rigorous methods.'*

Undertaking a research study is a daunting prospect and one that should not be taken on lightly. Studies may arise out of a general interest in a subject following an audit, for example, as part of career development or as a requirement for a post-graduate qualification (Doctoral or MSc degree). For MSc students the purpose of the project is to demonstrate to the supervisor and University examiners that the student has learned from the modules attended by reviewing the literature, identifying a gap in evidence, devising a focussed research question, designing a study to address it, collecting, analysing and interpreting data and making recommendations based on the findings. This applies whether the study is qualitative or quantitative in nature. Common elements in the early stages are devising the research question and choosing a study design.

(3) Devising the research question

A quote from "Alice in Wonderland" by Lewis Carroll

Alice was walking through a wood when she came across a fork in the path. A cat was sitting in a nearby tree.

"Which way ought I to walk from here?" asked Alice.
"That depends a good deal on where you want to get to," said the Cheshire cat.
"I don't much care where" said Alice.
"Then it doesn't matter which way you walk," said the cat.

The moral of this story is:

'The direction you take is more important than the distance you travel.'

The meaning: You must know where you want to get to when you set out. Otherwise you can spend a lot of time 'travelling' (in many directions) before you reach your destination, if ever. Hence, the research question is critically important in determining the 'direction' you need to take.

The research question is traditional, it provides the focus for the study, helps set it in context, identifies the methodology, sets the aims and objectives, and contributes to the development of any hypothesis (if relevant). Funding bodies, University postgraduate committees, ethics committees, Research and Development offices

and journal editors (if you ever get that far) will expect you to present a focused research question. The topic should be important (whether clinically or non-clinically) and the question should be researchable, ethical and answerable with the resources available. The aims should be achievable and any hypothesis testable.

A general question such as 'what is a good doctor?' is not answerable because there is no simple or single answer; there will be many aspects and perceptions to consider. However, it would be acceptable to ask 'what do patients consider to be important attributes of being a good doctor' as this is answerable in the context of the patients' perspectives.

Another general question such as 'How can asthma care be improved in Fife?' is also poorly constructed and lacks specification. To improve this, we devise a focused question containing four elements collectively referred to as PICO:

Patient / Population / Problem
Intervention
Comparison
Outcome

For example, what are the benefits of patient education in asthma? The concepts are benefits (O), patient education (I) and asthma (P). Each element needs an operational definition. 'Benefits' may relate to a general increase in health status for which specific outcomes might be a reduction in symptoms, a reduction in days lost from work (or school), a reduction in the diurnal variation in Peak Flow or an improvement in quality of life. The 'intervention' may involve a one-to-one interview with a 'specialist asthma nurse', provision of leaflets or use of a video on asthma management. The 'patient' may involve school-children, adolescents, adults (of whatever age) or focus on other subgroups such as ethnic minorities. Hence, an appropriate, focussed research question may be:

"What is the impact on quality of life of a video on asthma and its management in adolescents aged 12-18 years?"

It is implicit in this question that the 'comparison' (C of PICO) is likely to be 'usual care', which itself needs defining as 'usual care' is likely to differ between settings.

In general, if your question is such a good one then it is likely that someone has already thought of it, researched it and published their findings. Hence, the next step is to conduct a focused literature research using the key words from PICO and combining the searches. From the example above, key words would be 'asthma', 'management', 'quality of life', 'adolescents', 'intervention' (subcategory 'video'). Papers identified will need to be subjected to a critical appraisal with poor papers discarded and better papers retained. Following this process, it may be that sufficient evidence already exists on the topic though that should not stop you considering another study on the same subject. After all, what works in one health care setting may not work in another. Hence, previous studies can be replicated with potentially useful findings to improve clinical practice in your setting. However, at this stage and, only if felt necessary, the research question may be revised to investigate some other aspect of the topic (for example, to study patients in a different age group).

More examples:

"In women of child-bearing age does pregnancy result in increased tooth loss?"

P: Women of child bearing age
I: Pregnancy
C: Not pregnant, or women having a different number of pregnancies
O: Tooth loss

"What is the incidence of side effects using drug 'A' compared with drug 'B' in patients being treated for opiate withdrawal?"

P: Patients being treated for opiate withdrawal
I: Drug A
C: Drug B
O: Side effects

On occasions it may not be possible to specify these four elements. Examples include qualitative or quantitative studies where there is no intervention, such as a study involving the use of a questionnaire survey to determine a prevalence of a disease, symptom or condition in a defined population.

(4) Study types

The research question determines the study design of which the principle types are qualitative and quantitative in nature. Some important features of each are compared in Table 1.

Table 1. Some features of qualitative and quantitative study designs

	Qualitative	Quantitative
Role:	Observational To explore beliefs / experiences / knowledge To generate hypotheses	Observational / experimental To investigate relationships / causal pathways To generate and test hypotheses
Research question:	What is it like living with disease 'X'?	How many patients have disease 'X'?
Standard methods:	Document analysis Observation Interviews / focus groups	Structured data collection Instruments / protocols Measurement scales
Data:	Words: Subjective ('rich')	Numbers: Subjective / objective
Analysis:	Thematic / interpretive	Statistics
Results:	Not generalisable (usually)	Generalisable

Qualitative and quantitative methods are considered complementary by some, but not all authorities. Some consider the two approaches are radically different with quantitative methods appropriate to investigate 'fixed' knowledge on issues such as cause and effect that endure over time whereas qualitative methods address issues

that are socially constructed and subject to constant change and variation between settings and participants.

Qualitative research is concerned with developing explanations of social phenomena, including people's lived experiences, their views and attitudes. The data are non-numerical and typically relate to words. Quantitative research includes estimation of a numerical value such as a proportion (prevalence) or testing of a hypothesis. A study which seeks to answer the question 'How many women are recalled following a cervical smear' is clearly quantitative in nature, and that which asks the question 'What are the concerns of women recalled following a cervical smear' is then qualitative in nature. A research study can include both qualitative and quantitative methods.

(5) Qualitative studies

Qualitative studies have an important role particularly when first approaching a topic about which little is known. They use standard, observational methods to explore people's beliefs, experiences and knowledge. Techniques include document analysis (nursing notes, emails, minutes of meetings, diaries etc), participant observation (of a parent's interaction with their child, for example), one to one interviews (semi-structured, unstructured) and focus groups where small groups of individuals are asked open questions by a facilitator who records and interprets the conversations. Each method of data collection has its own strengths and weaknesses (Table 2).

Table 2. Strengths and weaknesses of common data collection methods in qualitative research.

Document analysis	
Strengths	• Low cost • Convenient (assuming documents are accessible) • Potential for unbiased data collection • Good for prospective studies (e.g. diaries of symptoms, medication adherence) • Potentially comprehensive records • May allow retrospective review of change over time in populations if source material has been collected rigorously and to high standards of completion (e.g. care home nursing notes) • Potential source of contemporary, independent evidence • May be the only source of evidence for long-term historical research
Weaknesses	• Missing documents or content a threat • Possible restricted accessibility (confidentiality) • If multiple observers involved be aware that writing styles and content may vary • Potential ineligibility of written content • Accuracy and authenticity of content not guaranteed • Selective reporting (e.g. of unfavourable events) • Potential change in standards/practice over time (historical studies) • Context in which content is recorded may not be appropriate for the research • Information recorded may not be germane to research question • Volume of data may be excessive (hence, collection and analysis time consuming)
Direct observation	
Strengths	• Can provide objective evidence on behaviours and interactions, verbal and non-verbal, in a natural setting if participants unaware they are being observed

	• Observations made within context and environment under study • Use of film and / or audio can provide remote and independent data unbiased by the presence of an observer • Researcher can be a participant (provides further in-depth analysis of context)
Weaknesses	• Hawthorne effect – participants may alter their behaviour, knowingly or unknowingly, if aware they are being observed. • Can be time consuming • May be subject to practical constraints • Rigorous training of multiple observers necessary • Potential conflict of interest if observer notes unethical or unprofessional behaviour
1:1 interviews	
Strengths	• Can be semi-structured or unstructured • Gives opportunity to probe in-depth using 'open questions' • Interviewer can clarify any uncertainty over question wording • Question sequence can be varied to suit interviewee • Questions can be left out if considered irrelevant • Can use less precise wording suited to the interviewee • Potential use of audio or video recording to collate data
Weaknesses	• Potentially expensive • Can be lengthy and collection of data, and its analysis time consuming • Not anonymous, though interviewer can give reassurance • Results subject to response bias but also to observer bias (training an important issue if using multiple interviewers) • Consent to record interview may be withheld and then the need to record field notes can be distractive • Does not provide evidence of interaction between participants
Focus groups	
Strengths	• Can offer more efficient data collection than 1:1 interviews • Groups can be made up of participants who know one other (e.g. work colleagues) who share an experience or participants who are total strangers (to elicit 'social, group norms') • Provides evidence on the interaction between participants • Improved access to 'hard to engage' groups • Allows interaction between respondents to explore similarities and differences in views • Replicates the cultural context in which people discuss issues, particularly sensitive ones • Venues can be chosen to offer a 'safe' environment • Can study how opinions are formed from the flow of conversations within the group
Weaknesses	• All must consent to audio record interview as any dissenting participant may risk accuracy of data collection • Analysis time consuming • Possible lack of disclosure of sensitive attitudes in a group setting • Not anonymous, maintaining confidentiality between participants can be uncertain • Potential discord between participants from disclosure of 'unsavoury' attitudes in a group setting • Potential suppression of views from 'power' relationships in groups where participants are known to one another • Risk of loss of control of the group and direction of conversations by the facilitator from (i) extraneous distractions at the venue and (ii) dominance of a single participant

The non-numerical data are described as 'rich'. Any hypothesis emerges from the data collected unlike quantitative research where an hypothesis is established first and subjected to challenge by experiment.

The analysis of qualitative data can be onerous, time consuming and subject to bias from the person undertaking the analysis. In general, the analysis is best undertaken by the person collecting the data as review of any transcripts of audio recordings can be misunderstood by those not privy to the way a patient's views may have been expressed. As an example, the simple statement in a transcript "she was alright" may be interpreted differently according to any emphasis made on individual words, or pauses made during its expression. Hence, "she was alright" is different from "she was alright" which, in turn is different from "she was [pause] alright [with the latter word expressed as a question]".

Qualitative studies can be made before, during or after a quantitative study. For example, when a new intervention or service is introduced qualitative studies may be used (1) beforehand, to interview staff to identify their concerns or potential barriers about the new service, (2) once in operation to interview service users about their experiences of using the new service, and (3) once the service has been embedded for some time to interview staff on residual or new, unanticipated problems arising from any new working arrangements.

As stated before any hypothesis emerges from the data. Qualitative studies use a number of different analytical and theoretical approaches. These include discourse analysis, grounded theory, ethnographic and phenomenological approaches. A detailed discussion of the many varied different approaches is beyond the scope of this guide and the reader is referred to one of the many textbooks on the subjects. (*see also,* Study guide 9: An Introduction to Qualitative Research).

(6) Quantitative studies

Quantitative studies can be of two types, broadly, observational and experimental. Observational studies are descriptive; the subjects do not receive any treatments or experimental interventions. The measures of interest are recorded with no attempt to influence the measurement. Experimental studies, by comparison, involve some intervention to change a variable and monitor the effect of this change on some function, for example the effect of a drug on blood pressure. The drug may be compared with another drug or with a placebo though the latter may not be possible, or ethical if seen to be withholding a proven treatment without patient consent. The need for a placebo may be a particular difficulty for community interventions in, for example, studies promoting a change in diet or exercise behaviour.

Experimental studies can be conducted using separate groups for treatment, control and placebo conditions (independent groups design) or by using the same group to receive all conditions (repeated measures design – crossover design). Cross-over studies have added benefits with regard to statistical power and require fewer participants compared with using two or more independent groups.

Quantitative studies have a role to play in investigating relationships, causal pathways and testing hypotheses. They use structured data collection methods including questionnaires (structured, semi-structured) and instruments to record physiological attributes with data being collected in a measurement scale and based on a robust protocol. Data may be subjective, for example from self-report, or objective, for example from a physiological measurement. The data may be subjected to simple descriptive statistical analysis or to complex analyses to

compare groups using parametric and non-parametric techniques depending on the distribution of the data (bell-shaped or skewed) (*see* Study guide 10: An Introduction to Medical Statistics).

The research question determines the study design each of which has its own strengths and weaknesses. The designs include:

(i) cross-sectional,
(ii) case control (also known as a case referent design),
(iii) cohort or longitudinal,
(iv) cross-over,
(v) randomised controlled trial (considered the 'gold standard' design for investigating hypotheses).

(6.1) Cross-sectional

In a cross-sectional study each participant is examined at one point in time. Such studies are relevant for estimating the prevalence of a disease, symptom or risk factor, or for investigating associations between a disease and putative causal factors. Cross-sectional studies can identify associations but cannot be used to investigate causal pathways as it is unknown which came first, the disease or the exposure to the putative causal factor. An early example from the 1940s was the observation that many patients with lung cancer happened to be tobacco smokers. The fact that two features are associated does not imply they are causally related. Consider the observation that many blind people happen to own a dog, and not just any dog but a Labrador! Could there be a link between dog ownership leading to blindness? In this example, the reality is the other way around, an effect called 'reverse causality'.

The results of a cross-sectional study can be used to generate a hypothesis. In the case of smoking and lung cancer the hypothesis that tobacco smoking leads to the development of lung cancer can be tested in a cohort (or longitudinal, follow-up) study of smokers and non-smokers to compare disease frequency in the two groups. Such a causal association was firmly established in the classic work of Richard Doll and Austin Bradford-Hill in the 1950s.

One problem with cross-sectional studies concerns the participants studied because they are seen at only one point in time. A study was designed by the Health and Safety Executive to investigate respiratory ill health associated with occupational exposure at cotton mills dealing with waste cotton where the exposure to cotton dust in the atmosphere was well above the accepted 'safe' levels (called the threshold limit value, TLV). The 60 workers were examined once but no evidence of disease due to their occupational exposure (byssinosis) was noted. However, it was clear that the turnover of staff at each mill was high and anyone who had difficulty working in those conditions simply left. The workers who remained were (apparently) unaffected by the adverse conditions and, effectively, were 'survivors'. Accordingly, any association between respiratory ill health and exposure to cotton dust was missed (*see* References, Chinn et al, 1976).

The strengths and weaknesses of the cross-sectional design are given in Table 3.

Table 3 Strengths and weaknesses of the cross-sectional study design

Strengths:	• Convenient, as carried out at one point in time • Often low cost • Can be easily set up and results obtained quickly • Useful for generating hypotheses by determining associations • Can be repeated in different settings
Weaknesses:	• Cannot study cause and effect relationships (temporality issues) • Prone to bias if studying only 'survivors' • Not good for rare conditions as numbers needed to study will be excessive • Cannot predict future health outcomes as any associations identified may be spurious rather than causally related

(6.2) Case-control

Case control studies are appropriate when studying factors associated with the development of rare conditions. Individuals with the condition of interest (cases) are identified and 'matched' with one or more individuals without the condition (controls). Features such as lifestyle factors and exposures can then be compared between cases and controls to identify suspect causative agents. An example from the 1990s was the study of the relationship between diet (specifically beef consumption) and the rare condition Bovine Spongiform Encephalitis ('Mad Cow disease').

Case control studies can be subject to selection bias. For example, a population-based study of patients with upper aero-digestive tract (UAT) cancers relied on recruiting patients from a regional radiotherapy centre. Patients with a UAT cancer attending for radiotherapy were identified and their personal, lifestyle and occupational exposures were compared with control patients. However, some patients with UAT elected not to undergo radiotherapy, were not offered radiotherapy, or died before attending for radiotherapy, so were unidentified from amongst the population and hence data capture was incomplete.

Another form of selection bias is called 'Berkson's fallacy' that can occur with hospital-based studies when cases and controls differ systematically in their risk of admission to hospital due to a combination of exposure and disease. The combination may increase or decrease the exposure rate amongst the cases that will distort the statistical results relating the exposure to disease occurrence. An example is the admission criteria applied before patients become eligible for surgery; some surgeons will only consider patients ready for coronary artery bypass operations after they have discontinued smoking.

The strengths and weaknesses of the case control design are given in Table 4.

Table 4 Strengths and weaknesses of a case control study.

Strengths:	• Good for studying rare conditions • Cases should be easily identifiable (and presumably available) • Relatively cheap • Can be done from hospital setting • Can be easily set up and results obtained quickly • May be statistical considerations in that fewer subjects required compared with cross-sectional and cohort studies • Can look at several potentially causative factors in the same study
Weaknesses:	• Highly dependent on suitable controls • A need for careful matching for known confounders, e.g. age, gender, etc • The greater the number of matching criteria the greater is the difficulty of finding suitable controls • Results can only support, not prove, causal associations (problem of temporality – which came first, the disease or the exposure?) • Subject to reporting bias, e.g. from patient's memory or notes. Cases can have selective memory e.g. mothers of children with autism may have greater recall of past events, which might be considered causative, compared to mothers of controls • Cases recruited from hospital may not be 'representative' of all cases with the disease (selective survival)

(6.3) Cohort

A cohort study seeks to follow-up a group of individuals over time to measure some aspect of change. Several groups may be involved with different exposure to a putative risk factor. Cohort studies may be prospective or retrospective. In a prospective study a group (cohort) of individuals are followed over time to investigate the development of a disease or relapse of symptoms, for example. Cohorts may include occupationally exposed individuals, infants and children (as in growth studies), patients discharged from hospital etc. In a retrospective study a cohort is defined from the past and the individuals followed-up to the present day (also called an historical cohort study). Such studies include those looking at the association between birth weight, early life exposures and subsequent health outcomes in adulthood (e.g. heart disease, stroke, diabetes).

Cohort studies are able to identify causal associations as, unlike cross-sectional studies they can address temporal relationships by recording which came first, the exposure or the disease. They can quantify the attributable risk of developing a disease (for example, the development of lung cancer in cigarette smokers) and hence the impact on population health status from eliminating the causative factor.

Cohort studies can extend over many years and can suffer bias in data collection due to selective loss to follow-up if individuals move away, die or drop out for reasons associated with the condition being investigated. Hence, patients who develop symptoms may decline to participate in a follow-up examination thereby distorting the measures of relative risk between groups. However, an assessment can be made in the data analysis to estimate the effects of bias from unbalanced loss to follow-up.

The strengths and weaknesses of the cohort design are given in Table 5.

Table 5 Strengths and weaknesses of a cohort (longitudinal) study design

Strengths	• Addresses issue of temporality (which came first, the exposure or the disease) • Good for studies of causation (identification of putative risk factors) • Can quantify the risk of developing a condition • Can quantify attributable risk and, therefore, the likely impact on health status from eliminating the causative factor. • Less prone to observer bias in data collection at the start of the study (investigators will not know which participants are likely to develop the condition under investigation) • Can assess multiple outcomes in the same study
Weaknesses	• Requires long-term commitment to maintain standards (quality control) • Can be expensive though not necessarily • Results may not be available for years, during which time exposure conditions may have changed (e.g. in industry) • Serious threat of bias from incomplete follow-up due to selective loss from the cohort • No control over changes (e.g. in the environment) which may affect the relationship between the disease and putative risk factor being investigated e.g. change in tobacco taxation or legislative changes such as the introduction of seat belts • Not relevant for rare diseases because follow-up must be prolonged to capture enough cases to make comparisons meaningful (threat to statistical power)

(6.4) Cross-over study

In a cross-over study each participant is subjected to both interventions being compared. A participant receives one intervention then, after a suitable washout period is switched to the second intervention. The order in which participants receive the interventions is randomised. One advantage of this study design is that, effectively, each participant acts as their own control. In consequence, the number of participants required to achieve a given statistical power is less than that required for other randomised designs involving parallel groups of different participants. This is because variability *within-patients* is less than that *between-patients.*

There are limitations, however, and cross-over trials are only really useful when the effect sought is short-term and the washout period is short. The strengths and weaknesses of cross-over studies are given in Table 6.

Table 6 Strengths and weaknesses of a cross-over study design

Strengths	• Useful for studies of short-acting drugs in chronic (stable) diseases • Allows for a randomised design, hence reducing potential bias • Convenient design where each participant acts as their own control • Requires fewer participants than a traditional randomised controlled trial involving parallel groups.
Weaknesses	• Requires a washout period between treatments • Residual effects from first treatment may interact with second treatment • Possible ethical and clinical concerns regarding withdrawal of treatment during the washout period • Less suitable for long-term drug effect studies • Less suitable for acute diseases if the condition varies naturally between treatments • Cannot be used for diseases / conditions that can be cured • Potential for bias in analysis failing to identify treatment order effects

(6.5) Randomised Controlled Trial

The randomised controlled trial (RCT) is considered best evidence (when it works). Participants are randomised to receive one of two or more treatments. Randomisation works in the long-term to smooth out differences between groups but cannot guarantee balanced groups when the number of participants is low. An RCT is not always possible because of ethical issues if assigning patients to what may be considered an inferior treatment (for example, use of a placebo drug in patients with asthma), or when there is potential to do harm (for example, when studying the effect of alcohol intake on pregnancy outcome).

RCTs work best when the number of patients recruited and followed-up to completion satisfies the power calculation to test the hypothesis. Measurements should be objective, valid, reproducible, and made contemporaneously in both groups, as well as double-blinded (i.e. neither the patient nor the researcher assessing the treatment effect is aware of which treatment the patient is on). Groups should be balanced at the start of treatment (by comparing baseline data) and the data should be analysed using an *'intention to treat'* analysis whereby participants remain in the groups to which they were allocated. Analyses where data are analysed according to the treatment participants actually received is called a *'per-protocol'* analysis and allows for the situation where participants may have been switched between groups.

RCTs are prone to errors in design from inappropriate randomisation strategies. Each participant should have an equal chance of being allocated to either group. Methods abound regarding random assignment. This can include randomisation in blocks to guarantee equal numbers in groups after, say, 20 recruits. 'Alternate assignment' is not the same as randomisation.

The strengths and weaknesses of RCTs are given in Table 7.

Table 7 Strengths and weaknesses of a randomised controlled trial

Strengths	• Considered best evidence of effectiveness • Provides better control over known (and unknown) confounders • Limits bias through double-blinding, where possible • Allows evaluation of a single intervention / drug on an outcome
Weaknesses	• Prone to problems of inappropriate randomisation • Double-blinding, or single-blinding not always possible • For drug trials the assumption that participants do take the medication according to the instructions • Can be expensive • Requires considerable resource through project management

The research question determines the study design (Table 8).

Table 8. Research questions and quantitative study designs

Research question	*Study design*
What is the prevalence of asthma in school aged children in Fife?	Cross-sectional
What is the association between barriers to physical activity and socio-economic position in adults aged 40-65?	Cross-sectional
Is there an association between shipyard welding and respiratory symptoms?	Cross-sectional
To what extent is the development of respiratory symptoms related causally to shipyard welding?	Prospective cohort
What is the incidence of laryngeal cancer in former steel workers?	Retrospective cohort
Is laryngeal cancer associated with past exposure to acid mists in steel mills?	Case control
Is maternal obesity a risk factor for stillbirth?	Case control
What is the frequency of occurrence of anaemia in relation to the diagnosis of colorectal cancer and site of tumour?	Retrospective cohort
What is the impact of a primary care-based dermatology nurse intervention on the quality of life of children with atopic eczema?	Randomised controlled trial
Is drug X better than placebo in treating fatigue in patients with multiple sclerosis?	Cross-over (or parallel group RCT)

(7) Bias

Bias is the unequal distribution of error. It is the greatest threat to any research study and potential sources should be identified early in the planning process so that efforts can be made to eliminate or reduce its presence. There are many sources of bias (Table 9).

Table 9. Principal sources of bias in research studies.

Source	Comment
Design	Any aspect of study design, for example, faulty sampling, incorrect randomisation, temporal differences in examination of subgroups, inappropriate calibration of instruments, poor statistical analysis with failure to account for confounding, use of wrong statistical tests.
Assumption	Faulty logic of investigator, which can lead to faulty conceptualisation of the research problem, faulty interpretations and conclusions.
Selection	Faulty selection when the characteristics of the sample differ from those of the wider, target population. All potential subjects should have an equal chance of being chosen. Can be a problem when, for example, a written invitation is sent to a person who cannot read or who cannot fully understand English.
Ascertainment	Variation in diagnostic criteria used between or within studies (for example, criteria to define hypertension). Criteria may change with time.
Response	A major source of bias leading to a systematic error from differences in characteristics between those who accept and those who decline an invitation to take part in the research. It is not always possible to compare the characteristics of responders and non-responders but it should be done where there is a source of independent data (e.g. if available from GP records).
Measurement	Systematic error from poor calibration regimes, measurement errors, change of instruments between repeated assessments, different instruments used to collect data from different subgroups, data handling procedures, digit preference.
Measurement decay	Error from a change in the measurement process over time due to a change in instrument performance or from change in technique by an observer.
Classification	Categorisation of the results. For example, definition of an ex-smoker (abstinent for one day, one week, one month, six months, one year, ten years?)
Recall	Recall by respondents may be selective or otherwise different between groups with different rates of cognitive decline.
Reporting	Respondents may be apprehensive about being interviewed and give the responses they think the interviewer wants. Respondents may under-report or over-report symptoms depending on any vested interest, for example, occupational surveys of back injury, with denial to avoid being made redundant, or over reporting to get compensation. Bias can arise with postal questionnaires when it may be uncertain who has filled in the questionnaire and if they have had help.
Social desirability	People may wish to present themselves at their best and will respond to questions accordingly.
Acquiescence response set ('yes-saying')	Respondents will more frequently endorse a statement than disagree with its opposite.
Observer	Differences in measurement techniques between observers, and within observers over time (measurement decay). Different interviewers may show systematic differences in asking questions and recording responses. Interviewers may ask questions in a manner which encourages respondents to answer in a desired way. Initial training and inter-observer assessments are very important to eliminate differences in techniques. Inter-observer assessments may need to be repeated throughout the study.
Follow-up	Loss of follow-up. Bias due to systematic differences in characteristics between those who return compared to those who decline to attend, or are otherwise lost to follow-up measurements in a cohort study.
Lead time	Failure to follow-up two or more comparison groups at the same time.
Analysis	Inappropriate use of statistical methods, for example, different treatment of outliers, missing data, incorrect tests of significance and neglect of confounders.
Interpretation	Errors in inferences drawn from the statistical analyses, for example, over aspects of association versus causation.
Publication	Reports of negative findings are less likely to be selected by editors for publication. Authors may have over-emphasised any positive findings to encourage acceptance of their paper. Publication bias may lead to a researcher believing that his/her contribution is unique and original.

(8) Exercises

Consider each of these scenarios, then check your answers in the Appendix.

(1) Treatment with the lipid-lowering drug Simvastatin decreases the risk of having a cardiovascular event by 30% but compliance with medication is low. This may be because Simvastatin is taken last thing at night and other cardioprotective drugs are taken in the morning. Devise a research question and choose a study design to test the effects of changing medication timing from evening to morning dosing.

(2) The British Association of Dermatologist's guidelines on the management of children with eczema stress the need for adequate time to discuss treatment issues with the child and parent. Such time is not always available in the GP consultation. It's believed that a nurse could do the job more efficiently. Devise a research question and choose a study design to assess the impact of a nurse intervention on the management of eczema in children.

(3) In recent years there has been a general increase in the number of persons attending for voluntary counselling and testing (VCT) for HIV in Malawi. People usually attend because they believe they have been exposed to the virus. Services are established for those testing positive but little follow-up support is available for those testing negative. A negative result may give false reassurance to an individual who may then discount future symptoms or continue with their risk-taking behaviour. It is unclear what support, if any, such persons may want, or need to help them avoid future risk of infection. Devise a research question and choose a study design to investigate the perceived needs of people testing HIV-negative.

(9) Further reading

Books:

Handbook of Health Research Methods: Investigation, Measurement and Analysis. Bowling A, Ebrahim S (Editors), 2005, Open University Press.

Designing Clinical Research. 4th ed. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB, 2013. Philadelphia: Lippincott Williams & Wilkins

Qualitative Inquiry and Research Design: Choosing Among Five Approaches. 2nd ed. Creswell JW, 2007. London: Sage.

Papers:

Chinn DJ, Cinkotai FF, Lockwood MG, Logan SHM. Airborne dust, its protease content and byssinosis in 'Willowing' mills. Annals of Occupational Hygiene. 1976;**19**:101-8.

O'Brien MJ, DeSisto MC. Every study begins with a query: how to present a clear research question. *NASN School Nurse* 2013: **28** ; 83-85.

Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *BMJ* 2000; **320**: 114-116.

Appendix: suggested answers to exercises (but be aware that other solutions will exist)

(1) Research question: Is the efficacy of Simvastatin affected by morning or evening dosing in patients with hypercholesterolemia stable on 20 mg daily?

P: Patients with hypercholesterolemia stable on 20mg Simvastatin daily
I: Shift to morning dosing
C: Evening dosing
O: Change in fasting blood cholesterol concentrations (Total, LDL and HDL) after a suitable period (say, 8 weeks).

Study design (1): *randomised controlled trial*

Patients: those currently taking 20 mg Simvastatin at night, randomised to either continue taking medication as normal (evening) or to switch to morning dosage.

Null hypothesis: There is no significant difference in the *change* in blood cholesterol after XX weeks following a switch from evening to morning administration of Simvastatin.

Think about the ethics: there is a possibility that morning dosage of Simvastatin may result in an increase in plasma cholesterol concentration possibly putting patients at risk of an adverse event. However, if the risk of this is considered very small, and the patient has given fully informed consent having had the risk identified then the study can go ahead following a favourable opinion from the ethics committee.

Study design (2): Note this could also be done as a *cross-over study* with each patient being randomised to either continue with their evening dosing or switched to morning dosing. Then they would be switched over to the alternate dosing regimen. The outcomes would involve a paired analysis of the *change* in cholesterol concentrations from baseline to the end of the dosing period chosen (e.g. 8 weeks, morning and evening). Each patient would then act as their own control.

(2) Research question: What is the impact of a single dermatology nurse consultation in primary care on the quality of life of children with atopic eczema?

P: Children (aged x to y years) with atopic eczema
I: A single 30-minute consultation in primary care with a dermatology nurse
C: Usual care (i.e. GP consultation only)
O: Change in quality of life assessed at baseline and after XX weeks.

Study design: randomised controlled trial

Patients: children with atopic eczema registered with a general practice randomised to receive the nurse intervention or ‘usual care’.

Null hypothesis: There is no significant difference in the *change* in quality of life after XX weeks in children receiving a nurse intervention compared with usual care.

(3) Research question: What are the perceived support needs of people who test negative after attending voluntary counselling and testing (VCT) for HIV?

P: Clients attending VCT who test negative
I: - *Note: no intervention*
C: - *Note: no comparison*
O: Perceived needs of clients and change in sexual behavior post VCT

Study design: qualitative study, involving longitudinal, follow-up interviews and/or focus groups.

Participants: persons testing negative following an HIV test and (possibly) counselling staff available to deliver services to these individuals.

Participants testing negative may be interviewed once to enquire about their perceived needs and subsequently followed-up to ask what changes, if any, they made to risk-taking behaviour, particularly sexual behaviour.

Glossary

Tip: search Google for an on-line glossary of research terms not included here

Audit	A quality improvement process that seeks to improve patient care and outcomes through systematic review of care against explicit criteria and the implementation of change.
Bias	The unequal distribution of error leading to a deviation from the truth.
Blinding	The process by which participants and researchers are made unaware of the treatment received in a clinical trial. Blinding can be 'single' when either the participant or researcher is naive or 'double' when both the participant and researcher are naive to the treatment assigned.
Case control	A study that begins with the identification of patients with a disease (or condition) of interest and a suitable control group without the disease. Cases and controls are 'matched' for important features and compared to measure the relative frequency of occurrence of a characteristic believed to be associated with the disease (or condition) in question.
Causality	The relating of causes to the effects they produce.
Clinical trial	An experiment that involves the administration of a test regime to evaluate its efficacy and safety to participants who are patients.
Cohort study	An observational study in which a group or groups of individuals are followed-up with repeated measures over time to determine the relative frequency of occurrence of a disease or condition. The cohort may be studied prospectively or defined in the past and followed-up to the present day (retrospectively).
Confounding	A source of error that occurs when groups being compared differ with regard to an important characteristic related to both the disease in question and the feature under study but which has not been controlled for in the study design. An example is a study comparing a drug with placebo to treat hypertension where one group is significantly older than the other group. Hypertension is age-related and the difference in study outcome (blood pressure) between the drug and placebo may be a consequence of confounding due to the failure to account for the difference in age rather than the effect of the drug.
Critical appraisal	A systematic method of assessing the strengths and weaknesses of a research study by considering issues of validity, accuracy, bias and clinical relevance.
Cross-over study	A design in which study participants are given all treatments under investigation but in a sequence with a suitable washout period between treatment periods. Each participant then acts as

	their own control.
Cross-sectional	An observational study to determine the frequency of a particular disease, characteristic or condition measured in a defined population at one point in time.
Data saturation	Data collection in a qualitative study is continued until the analysis reveals no new themes emerging.
Digit preference	A tendency to preferentially record measurements to a certain level of accuracy particularly when rounding up. Measurements may be rounded to the nearest whole number, even number or to multiples of 5 or 10. Examples include recordings of blood pressure, height, body weight, waist circumference etc.
Discourse analysis	The analysis of speech and text to gain an understanding behind the words people use.
Document analysis	Systematic analysis of document contents to answer a research question in a qualitative study.
Ethnography	A qualitative research methodology studying people in their natural settings to describe their social interactions and culture. The method is commonly used by anthropologists.
Focus group	A qualitative research method in which participants are questioned by a researcher in a small group allowing interaction between members of the group to elicit views.
Grounded theory	A method of analysis of qualitative data in which the researcher identifies issues that emerge from the data to establish theories that can be tested against further emerging evidence as the analysis progresses.
Hawthorne effect	An effect when participants change their behaviour, consciously or unconsciously, as a result of knowing they are being observed.
Incidence	The number of new events (e.g. new cases of a disease) in a defined population within a specified time period. The term 'incidence' is sometimes used to denote 'incidence rate' which is the rate at which new events occur in a defined population. The numerator is the number of new events that occur in a defined period (year, month, week) and the denominator is the population at risk of experiencing the event during that period.
Intention to treat analysis	A method of analysis in a randomised controlled trial whereby all participants are followed-up whether or not they actually received or completed the intervention and their outcome measures are analysed in the group to which they were assigned.
Intervention	A treatment, service or policy intended to improve health status or welfare of an individual, family or community.

Non-parametric	Statistical method of data analysis that makes no assumptions about the distribution of the data. The method is appropriate when the distribution of the data is skewed (not bell-shaped).
Parametric methods	Statistical method of data analysis that assumes the distribution of the data is bell-shaped (also called Normal or Gaussian), or approximately so. Examples include the t-test, and Pearson's correlation.
Per-protocol analysis	A method of analysis in a randomised controlled trial whereby participants' outcome measures are analysed according to the treatment they received and not in the group to which they were originally assigned (*see* Intention to treat analysis).
Phenomenology	A research methodology which has its roots in philosophy and which focuses on the lived experiences of individuals.
Power	The probability of rejecting the null hypothesis when it is false.
Power calculation	A method of calculating the number of subjects needed for the results of a study to be considered statistically significant.
Prevalence	The proportion of a defined population that has a disease or condition at one point of time ('point prevalence') or during a defined period (e.g. a year, called the 'annual prevalence', or during a lifespan, called the 'lifetime prevalence')
Qualitative research	A method of studying the meanings people give to their lived experiences, attitudes, expectations and how they make sense of their world. Data may be collected by interview (personal or in a focus group), by participant observation or by reading what they have written. The analysis is non-statistical.
Quantitative research	A method to measure and investigate the relationship between one thing (independent variable) and another (dependent variable). It seeks to quantify relationships between variables. Results can be expressed in simple descriptive terms or as tests of statistical significance between groups.
Randomised controlled trial	A clinical trial to compare one or more treatments with a control condition. Participants are assigned to a group (treatment or control) by random allocation to minimise bias in the study design.
Research	The attempt to derive generalisable new knowledge by addressing clearly defined questions with systematic and rigorous methods.
Semi-structured interview	An interview where the researcher has a set of questions to ask but which can be varied in the order given and where the interviewer can depart from the question set to explore emerging themes.
Service	A review process undertaken solely to define or judge current

evaluation	service with the intention of benefiting those who use it.
Structured interview	An interview where the researcher has a set of questions to ask each participant but in which the order and wording is fixed.
Triangulation	The use of more than one method, theory, data source in a research study to affirm the study results.
Unstructured interview	An interview where the researcher asks participants very general questions without any predetermined plan to allow the participant to shape the interview in whichever way they prefer.