So you want to do research? 4: An introduction to quantitative methods

Keith A Meadows

Common in clinical and biomedical research, quantitative research is an essential part of health services research. Perhaps the most obvious example of this is the randomized controlled trial, with its emphasis on experimentation and large sample sizes. However, quantitative research encompasses a much broader spectrum of activity, which can include small-scale descriptive studies through to more complex studies where relationships between variables are explored.

In contrast to qualitative research which is often viewed as being unscientific and lacking rigour (Mays and Pope, 1995), quantitative research – with its focus on hypothesis testing, reliability and validity, enumeration and statistical inference – has been considered as the epitome of the scientific approach. Despite the antithetical portrayal of the qualitative and quantitative approaches, but acknowledging that there are some clear and distinguishing features between them, they should be seen as complimentary to each other. For example, qualitative research can provide a good foundation for quantitative research, particularly in areas where little is known about the subject. Equally quantitative research can provide insight into areas in need of further and in-depth investigation.

This article introduces some of the key issues in quantitative research, starting with a description of quantitative research and the types of research designs, the strengths and weaknesses of the different methods of data collection, an overview of the different sampling methods for collecting data, through to an overview on the analysis of quantitative data. The article also discusses considerations for ensuring rigour and the reporting of the research findings.

Choosing the correct methodology

What is quantitative research?

Turning your research question into a research project and selecting the most appropriate research design are the crucial parts of any enquiry. When choosing the research design a number of issues need to be borne in mind:

- Purpose(s) – What is the study trying to achieve and why is it being done?
- Theory – What (if any) theory will guide or inform the study? How will the study's findings be understood?
- Research question(s) – What is the research question asking?
- Methods – What methods (e.g. questionnaire, interviews, observation) will be used to collect the information?
- Sampling – From whom, where and when will the data be collected? (Robson, 2002).

In earlier articles of this series (Meadows, 2003a,b) different research strategies have been generally classified as either qualitative or quantitative. The aim of qualitative research is to help us to understand social phenomena in a natural rather than an experimental setting with emphasis on the meanings, experiences, attitudes and views of the participants and the focus on to determine 'why?' rather than 'how many?' (Hoinville and Jowell, 1978; Pope and Mays, 1995). Data obtained from qualitative research is usually in the form of words rather than numbers and these

Keith A Meadows is Sector R&D Manager, North East London Consortium for Research and Development, Mile End Hospital, London

Email: keith.meadows@thpct.nhs.uk

ABSTRACT

This fourth article of a series of six focuses on some of the key aspects of quantitative research methods. Starting with an overview of what quantitative research is, the distinguishing characteristics of experimental and non-experimental research strategies, the different approaches for collecting data including, self-completion questionnaires, interviews and scales, together with their respective strengths and weaknesses are discussed. The differences between probability and non-probability sampling and the different methods for selecting a sample are described. Aspects of quantitative data analysis are briefly reviewed and the concepts of reliability and validity are described in the context of ensuring rigour in the research design. Finally, some guidance on the reporting the findings from quantitative research is provided.

words are based on observation, interviews or documents (Miles and Huberman, 1994).

Quantitative (or 'fixed-design') research involves experiments (such as randomized control trials (RCTs)) and surveys, where data are collected using standardized methods such as questionnaires and structured interviews. The data are in the form of numbers from which statistical generalizations can be made. Key characteristics of quantitative research are that much of it is pre-specified in terms of what and how is going to be done (Robson, 2002) and that the approach is deductive (where data are specifically collected for the purpose of testing ideas and hypotheses) rather than inductive (where ideas and generalizations emerge from the data). In other words, quantitative research should be theory-driven, and the variables to be measured (i.e. properties which can vary and be measured, such as disease duration, age, gender, depression score) and the procedures to be followed in the collection and analysis of the data, should all be pre-specified. In contrast to qualitative research where findings are reported from the perspective of the individual, findings from quantitative research are reported in terms of aggregates and group properties and averages.

Types of fixed quantitative research designs

Quantitative research strategies can be broadly classified as experimental or non-experimental (descriptive). Experimental designs are characterized by the manipulation or introduction of some variable – such as treatment – and comparing the outcome with a control group, which has not received the treatment. The most common example of an experimental design is the RCT, in which study participants are randomly allocated to either the experimental or control group. In contrast, in non-experimental designs the data are obtained from existing groups, for example, to look at the relationship between a number of variables such as the scores on a depression scale and sex or age (relational design); to identify the type and frequency of diseases in a specific group of people (descriptive); to determine why a particular group is affected while another is not (analytical). *Box 1* lists examples of quantitative studies.

Box 1. Examples of a range of quantitative studies

A comparative study of self-perceived health of parents of children with Down's syndrome and a randomized control group of parents from the Swedish SF-36 norm population.
Hedov G, Anneren G, Wikblad K (2000) Self-perceived health in Swedish parents of children with Down's syndrome. *Qual Life Res* **9**(4): 415–22

An assessment of the impact of cancer on the psychological well being of newly diagnosed cancer patients before and during a course of radiotherapy.
Chandra PS, Chaturvedi SK, Channabasavanna SM et al (1998) Psychological well-being among cancer patients receiving radiotherapy–a prospective study. *Qual Life Res* **7**(6): 495–500

A prospective cohort study of 7726 adults to determine whether poverty and unemployment increase the likelihood of delaying the recovery from common mental disorders.
Weich S, Lewis G (1998) Poverty, unemployment, and common mental disorders: population based cohort study. *BMJ* **317**(7151): 115–9

A cross-section study to examine the role of health status, personality and coping style on the reporting of health-related quality of life of HIV seropositive gay and bisexual men
Burgess AP, Carretero M, Elkington A, Pasqual-Marsettin E, Lobaccaro C, Catalan J (2000) The role of personality, coping style and social support in health-related quality of life in HIV infection. *Qual Life Res* **9**(4): 423–37

Methods of data collection

Surveys and questionnaires

One of the most common ways of collecting information in quantitative studies is the survey, which almost always uses self-completion questionnaires, face-to-face or telephone interviews or tests and scales.

Surveys

Most people have participated in a survey. A survey is a system for collecting information on a range of topics including health, education, psychology, law etc. (Fink, 1995a) of which the main characteristics are that data are collected from a number of individuals using a systematic and standardized approach (e.g. questionnaire/structured interview schedule/scales or tests) and that these individuals are a representative sample of the population under study.

The requirements of a good survey, whether small or large, are that:

- Its objectives must be specific and measurable
- The design must be sound
- The sample(s) or population studied must be appropriate for the study
- The questionnaires, scales and tests used must be reliable and valid
- The most appropriate analysis must be applied to meet the objectives
- Findings must be reported accurately (Fink, 1995a).

There are a number of ways in which questionnaires can be administered in a survey:

- Self-completion, where respondents fill in the answers themselves and the questionnaire is either sent out by post or handed to the respondent to complete.
- A face-to-face interview, where an interviewer asks the respondent the question and also completes the questionnaire. In quantitative studies the interviewer is more likely to use a fully structured interview schedule which has pre-set questions in a set order with fixed wording and responses for the respondent to choose from.

- Telephone interviews, where respondents are contacted by telephone and are asked the questions and the interviewer records the answer as in a face-to-face structured interview.

Tests and scales

While tests and scales are included under the category of questionnaires, they differ somewhat from what we traditionally understand a questionnaire to be. Questionnaires might ask, for example, 'How long did you have to wait before being seen by a doctor?' or 'How satisfied were you with the treatment you received?'. Tests and scales by contrast have been developed to assess people's abilities, attitudes, opinions, physical and psychological functioning etc. A typical example of a test might be the person's IQ or intelligence. Scales are intended to gain an insight into the individual's physical and psychological functioning, e.g. the Hospital Anxiety and Depression (HAD) scale (Zigmond and Snaith, 1983). Whether the objective is to assess the ability of the individual or gain some insight into the individual's opinions, attitudes or physical functioning, the purpose is to quantitatively scale the person on the measured attribute.

Uses of surveys and questionnaires

Surveys using questionnaires and standardized interviews can be used in both non-experimental and experimental designs as a method for collecting data (Fink, 1995a). Common uses for a survey in an experimental design include the measurement of change during, and outcome at the end of the experiment. For example, we might want to compare at two weeks, one and two months after attending a one-day education programme, the changes in the social, psychological and physical functioning of patients who had experienced hypoglycaemic episodes compared to a control that had not. Or we might want to interview patients who took part in a counselling programme two months after finishing the course and compare the outcomes with a control group. Surveys can also be used in experimental studies for other purposes, such as selecting study participants and checking on the comparability of the experimental and control groups, e.g. by age, sex, level of education etc.

Exmples of non-experimental studies that involve the use of a questionnaire, interview, scale or test, might include a cross-sectional design using a postal survey to find out the perceptions of the quality of the services provided to outpatients of a diabetes clinic, or a telephone interview with postoperative patients to find out what has happened since surgery. For a cohort study we might be interested in monitoring over time the reported health-related quality of life of a group of patients with severe hypoglycaemia. With a case-control study, we might want to examine the attitudinal, social and demographic variables of people to help understand why some people did not continue to use a particular insulin injection device compared to matched controls who did.

Strengths and weaknesses of questionnaires, interviews and tests

Every method of data collection has its strengths and weaknesses. Self-completion questionnaires are often sent by post enabling large samples to be reached, but response rates can be low, they are inappropriate for populations with high levels of illiteracy, they need to be short and the questions asked need to be in simple language. However, with careful design of the study and the questionnaire, self-completion questionnaires can provide very useful and representative information as well as overcome some of the barriers in the collection of sensitive data (see Oppenhiem, 1992).

Face-to-face interviews address a number of the limitations of the self-completion questionnaire: they can be used in populations with high levels of illiteracy, and interviewers can provide clarification and deal with misunderstandings as well as ensure that information is collected. However, face-to-face interviewing often requires training and cost more in time and resources, there are risks of interviewer bias and the collection of sensitive data can be problematic. Compared to postal questionnaires the size of the sample reached can be limited unless there is more than one interviewer.

Telephone surveys combine the advantages and disadvantages of the self-completion questionnaire and face-to-face interview. Telephone interviews enable large samples to be reached, interviewers can provide clarification and address misunderstandings, and levels of literacy are not such a significant problem. Limitations include confidentiality, bias and the asking of sensitive questions. Telephone surveys may also have to be conducted in the evenings when respondents can be reluctant to answer certain questions if other people are present. Results can also be biased because only respondents with a telephone have been included in the sample.

The importance of careful thought in the design of a questionnaire, interview schedule, scale or test cannot be stressed enough and will be discussed in detail in the next article of this series.

Sampling in surveys

Using an example from the first article of this series (Meadows, 2003a), we might want to carry out a survey of the smoking behaviour of all men aged between 35 and 60 years in a particular geographical area. Depending on the size of the area, and for our

'The importance of careful thought in the design of a questionnaire, interview schedule, scale or test cannot be stressed enough.'

'Non-probability sampling involves choosing samples not so much to be representative of the target population, but on the characteristics of the target population. '

findings to have some meaning, this would most likely entail the completion of perhaps many hundreds or thousands of questionnaires. To overcome this difficulty we would use sampling techniques that would give us a degree of confidence and the same information, but from a smaller group of men.

A sample is a portion or subset of the population we wish to study. There are two types of sampling methods we can use to select our sample.

Probability sampling

The first of these is probability sampling where every sampling unit (i.e. each man aged between 35-60 years) has an equal chance of being selected from the target population (i.e. all men aged 35 –60 years in the geographical area). The underlying rationale of probability sampling is random selection, which removes subjectivity in choosing the sample (Fisk, 1995b). If done correctly, it provides us with some confidence that our sample is representative of the target population and that our survey findings can be extrapolated to the target population within certain limits of confidence. There are several ways in which a probability sample can be selected:

- Simple random sampling – Selection is made by chance alone such as drawing numbers from a hat, or, for example, from a list of men 35 to 60 years of age with type 2 diabetes attending a diabetes centre. Each number or man has an equal chance of being selected. While this is simple to do, lists are not always available, and this method can be costly in practice if sampled units are geographically widely dispersed.
- Stratified random sampling – The target population is divided into different subgroups – e.g. by sex, treatment type, age – from which a random sample is selected. This is more likely to reflect the target population and reduce sampling variations, but can be time consuming and costly.
- Systematic sampling – You want to select 200 patients from a list of 1000 names to take part in a survey on attitudes to the service provided by a health centre, or interview the head of the household from 1000 addresses. Dividing 1000 by 200 gives 5, meaning that 1 in every 5 patients or addresses is selected. Although convenient, this method is not suitable where there is possible repetition in the list, e.g. names starting with a certain letter or dates of birth.
- Cluster sampling – Naturally occurring units such as GP practices, health centres, hospitals, schools, are randomly selected and the unit of interest is included in the sample. For example to explore the attitudes of GPs across London to the new contract, you could first randomly select geographical areas across London and from these randomly select GP practices and then interview all the GPs in those practices. This is a convenient method because it uses existing groups, and particularly with larger sampling units such as hospitals, schools, lists to be more likely to be available.

Non-probability sampling

Non-probability sampling involves choosing samples not so much to be representative of the target population, but on the characteristics of the target population. As a consequence we are unable to say whether the findings from the study are or are not applicable to the target population.

The most common application of non-probability sampling is in flexible or qualitative research designs, where the emphasis is to select purposively the sample so that it represents a wide spectrum of views and experiences as well as covering the full range of individuals to identify, explore and explain variations in the nature of views and experiences between them. Non-probability sampling can also be applied to quantitative designs, for example when we need to survey hard-to-identify groups. The methods used in non-probability sampling are:

- Convenience sample – Using a group of individuals that are readily available and are willing to be surveyed. For example, we are interested in finding out what health services older people use. To answer this question we could post interviewers or give out questionnaires at positions where we are likely to recruit older people, such as at specific outpatient clinics or health centres.
- Snowball sampling – Members of the group are requested to identify other member of the target population. This is most commonly used when dealing with a hard-to-identify group, or when no listing is available, e.g. illicit drug users.
- Quota sampling – Divides the target population into subgroups, e.g. by sex or age, based on known estimates and then selects the proportion of people in each of the subgroups from the target population. If we are interested in looking for any differences in dental health care between boys and girls aged 10–15 years at a particular school. Estimates from the school tell us that 24% of boys and 19% of girls are 10 years old, 21% boys and 20% girls are 11 years old, and so on. From these figures we would then select boys and girls in these proportions to be surveyed.

For a detailed discussion on sampling procedures, see Fink (1995b); Argyrous (2000); Barnett (2002).

Deciding on the sample size

When considering the sample size it is important that

you are clear what the objectives of the study are and what is the research question or hypothesis (Fink, 1995b). As the size of the sample increases the sampling variability or error decreases. However, the larger the sample the more costly the data collection and analysis will be, so making sure you have the optimal sample size is important. The sample needs be selected so it is as representative of the target population as possible, is of a sufficient size to detect effects or changes in the variable studied and it is as free as possible from sampling errors. Determining the appropriate sample size requires the use of statistical calculations, which will involve answering a number of statistical questions, e.g. what chances should there be of finding a significant difference between the groups investigated? (i.e. the 'power' of the test) (see Siegal and Castellan, 1988); what differences between our groups would be considered to be important (e.g. the mean difference in age between those patients reporting improvement in quality of life following psychotherapy treatment and those not)?

A health warning

Survey sampling is a complex and important part of the research process and while a number of issues have been briefly discussed in this article, it is beyond the scope of the series to discuss in detail the strengths and weaknesses of the different methodologies. It is always strongly advised that a statistician is consulted at the earliest stage of the project design. For a well written, comprehensive and simple guide on sampling in surveys see Fink (1995b). Barnett (2002) contains a more detailed and mathematical approach to sampling.

Analysis and interpretation of quantitative data

Statistical significance

The analysis and interpretation of quantitative data is very different from that of qualitative data. With quantitative research we are dealing with numbers rather than narrative and quantitative analysis is practically synonymous with significance testing (e.g. is the difference in mean ages statistically significantly different, or in other words is it likely that this finding was not due to chance?).

Looking for statistical significance in findings is controversial. One problem is that statistical significance is not related to the size or the importance of the effect or relationship at which we are looking (Robson, 2002). For example we might find that the mean difference in patients' quality of life scores is significantly higher in the group of patients who underwent some psychosocial intervention compared to those who did not. Although significant (i.e. the finding was not due apparently to chance), what does this finding really mean? Firstly, the chance of finding a statistically significant result increases as the size of the sample increases. Secondly, although statistically significant, the differences between the means might be marginal. Thirdly, the observed difference might have very little clinical significance despite being statistically significant. For a detailed discussion on the significance test controversy see Robson (2002), pages 400–2.

Analysis

All analysis of quantitative data will involve some statistical manipulation, which can range from organizing the data, to provide a descriptive account of the findings, such as the mean age and range of the sample, or the percentage of men and women etc, through to the very complex statistical analyses involving multivariate analysis.

While a detailed discussion on how to conduct this analysis is beyond the scope of this article, a brief overview of the main approaches is given below. For a more detailed discussion of the analysis of quantitative data see Siegal and Castellan (1988); Fink (1995c); Argyrous (2000); Robson (2002).

A simple approach to quantitative data analysis is to first report on each of the important or individual variables. For example we could, using frequency distributions and graphical displays (bar chart, histograms and pie charts), report on the composition of our sample, highlighting characteristics such as age, sex, duration of disease, types of treatment, the percentage of patients reporting they felt very satisfied with the service provided by the clinic etc. We could also provide summary statistics which include measures of central tendency such as the mean, mode and median, measures of variability including the range (the difference between the lowest and highest value or score), variance and standard deviation (measures of the spread of the scores around the mean) and confidence intervals (CI) which provide us within a given statistical probability, the limits within which our mean score can lie.

At a more complex level we can analyse the relationships between two or more variables. When looking at the relationship between two continuous variables, e.g. age and the level of physical functioning, the strength and the direction of the relationship are given by the correlation coefficient. Other measures of relationship include the Chi-square (χ^2) test, which shows the degree of association between two or more non-continuous variables (such as sex) or continuous variables which have been categorized (e.g. age in years to age groups). More complex statistical procedures when looking for relationships

'All analysis of quantitative data will involve some statistical manipulation, from organizing the data, to providing a descriptive account of the findings… through to the very complex statistical analyses involving multivariate analysis.'

'Validity is about the confidence we have that we are measuring what we think we are; the accuracy of our results. Do our results actually reflect what is happening or are they due to something else?'

between three or more variables include multiple regression and multivariate analysis, factor analysis and structural equation modelling.

A traditional task of analysing quantitative data is exploring whether differences exist between variables under different conditions or in different groups. For example, is there a difference in the measured outcome of psychological distress between groups of patients undergoing different forms of psychosocial intervention? Or do women score higher on quality-of-life scale compared to men? There are a number of statistical tests available to test whether such differences exist and to make inferences that the findings are statistically significant. The choice of test depends on whether we are looking for differences between one, two, three or more groups and the nature of the data (Argyrous, 2000). But bearing in mind the controversy regarding the value of a statistical test that simply tells us whether the finding is by chance or not, there should be greater emphasis on looking at the effect size which is independent of the size of the sample.

Choosing the most appropriate method of analysing data is an essential part of the research process. Assuming that the chosen method is the right approach the interpretation of the findings becomes one of deciding whether our findings provides the evidence we are looking for, be it a relationship between two variables or the difference between two patient groups.

Ensuring rigour in quantitative research

When undertaking research of any kind the strategy must be to ensure that the rigour in the research is systematic and self-conscious. In doing so the researcher should seek to provide an account of the methods and data, which are plausible and can stand independently, and coherent explanations of the topic under study (Miles and Huberman, 1994).

The two key concepts when establishing the rigour of quantitative research are reliability and validity (see Robson, 2002).

Reliability is about how well the data we collect can be reporoduced using the same measure. Unreliability can stem from a number of different sources. For example, we might be measuring the level of diabetes-related knowledge in people with diabetes using a specifically-designed questionnaire. If the scale were reliable then we would expect the knowledge score obtained from each person to be very much the same whether we gave them the test on a Monday or a Friday. Of course there are likely to be some random fluctuations in the scores between the two occasions perhaps due to tiredness or other short-term effects. Of more importance however, are the more systematic causes leading to unreliability for example, if a question has been written in such a way that it can be misinterpreted. Other causes of unreliability include observer bias, where the observer consciously or unconsciously biases their reporting.

While reliability is an essential requirement of whatever it is we are measuring, it is insufficient to ensure the validity of what we are measuring. Something can be reliable without being valid. Our diabetes knowledge questionnaire for example, may provide consistent scores between different administrations but fail to discriminate between those with better diabetes-related knowledge.

Validity is about the confidence we have that we are measuring what we think we are; the accuracy of our results. Do our results actually reflect what is happening or are they due to something else?

Taking the diabetes-related knowledge example further, we might want to implement an education programme and test whether this does in fact improve the level of diabetes-related knowledge. To test the effectiveness of the programme we could use a diabetes knowledge questionnaire to see whether there were overall improvements in knowledge following the programme's implementation. So the first thing we would be interested in is the validity of the questionnaire itself. There are different types of validity we need to consider, which are briefly discussed here.

Construct validity is most important form of validity and often the most difficult to determine. Construct validity is a measure of how well you are measuring what you think you are measuring. In the case of a scale or test it is how meaningful the scale or test is. Construct validity can generally only be determined after much practical use where evidence supporting the usefulness of the scale or test is accrued over time. Evidence in support of construct validity could be the scale's ability to discriminate between different levels of the measured attribute, e.g. knowledge, and that the scale's score's would show the predicted relationship with other concepts, for example age or disease duration.

Criterion validity is a measure of how well the scale or test predicts future outcomes or how well the scale's score correlate with some 'gold standard' of the same variable. For example, the ability of a test of school educational achievement to predict entry to university, or scores of a diabetes knowledge questionnaire to show a strong relationship with another measure of diabetes knowledge which is considered as the 'gold standard'.

Content validity is the extent to which the items in the scale or test reflect the measured concept. For example, if we were developing a scale to measure

diabetes-related knowledge in general, the content should represent the different areas such as blood monitoring, diet, complications, foot care etc.

Face validity is the least important form of validity and is a casual review of how good the items of the test appear. At its simplest, if we were asking patients about their level of anxiety, then we would expect the questionnaire to comprise relevant items relating to anxiety. If is does not then the measure does not have face validity. Face validity is often confused with content validity and it is essential that the distinction between the two is made.

Internal validity

Having established that our scale of diabetes-related knowledge has demonstrable reliability and validity, the next aspect of validity we need to establish is whether the outcome is related to the intervention. If we are able to demonstrate the causal relationship between improvement in diabetes-related knowledge (the outcome) and the implementation of the education programme (the treatment), the study is said to have internal validity (Campell and Stacey, 1963).

Just because we have a finding that the level of diabetes-related knowledge improves following the implementation of the education programme, there are a number of possible reasons why it might be still be unwise to conclude this the case. Changes in outcomes following some intervention can also be due to other factors outside the study itself. Campbell and Stanley (1963) have referred to these factors as 'threats' and have suggested eight possible threats to achieving internal validity (*Box 2*). (Robson, 2002 has a more detailed discussion on internal validity).

Generalizability

In addition to validity of a study, we need to identify how applicable our findings are to other situations, settings and people, in other words the 'generalizability' of the study. Campbell and Stanley (1963) refer to this as external validity. For example, our study on the effectiveness of an education programme on people with diabetes might have been carried out on people ages between 18 and 35 years of age with type 1 diabetes. If we want to know how effective the education programme would be with an older group of people with type 1 diabetes, we should then be concerned with the generalizability of the study. The factors which limit the ability to generalize from a particular study include: the specific group studied; the setting or context in which the study took place; the uniqueness of the specific experience or history of the group studied; and the construct or concepts studied which may be specific to the group studied (LeCompte and Goetz, 1982). So in our example study it would be less likely that we could generalize the outcomes of the study to people with tablet-treated type 2 diabetes aged over 55 years, because of differences in the specific and unique experiences resulting from the different treatments as well as the different knowledge requirements between type 1 and type 2 diabetes.

Box 2. Threats to internal validity

1. History – Things changing in the environment of the participants not related to the study
2. Testing – Changes occurring as a result of experience and practice gained from pre-tests
3. Instrumentation – Changes in the way participants were assessed pre and post testing
4. Regression – Study participants chosen for being atypical (e.g. high scores) are less likely to be less atypical at later testing
5. Mortality – Participants dropping out of the study, loss to follow-up
6. Maturation – Growth, change or development in participants unrelated to the study
7. Selection – Initial differences between groups before involvement in the study
8. Selection by maturation interaction – Groups naturally growing apart or together

Source: Campbell and Stanley (1963)

Reporting the results

As discussed in an earlier article of this series (Meadows, 2003c) the aim of writing is both communication and persuasion (Gilbert, 2001). Reporting the findings from research is an essential stage of the research process. Dissemination of research is essential if the findings are to be of benefit to others, be open to critical examination by peers and promote service development based on sound evidence.

While there is little consensus on how qualitative research should be reported (Robson, 2002), there is a conventional model for quantitative research reports (*Box 3*). While the format might vary slightly according to the type of journal the research report is being submitted to, the format in which the research is written up will by and large be the conventional approach. However, the mode in which the results are being presented (e.g. oral or poster etc) and the audience, may call for alternative forms to aid the communication, but can also be along the lines required by scientific journals. When reporting quantitative research in a scientific journal, it should be in a manner and of sufficient detail to enable someone else to replicate it.

Ethics and consent to participate

Last but not least are the issues of ethics and consent. All research involving the use of human subjects either directly or indirectly, must receive ethical approval before commencing. This will generally mean submitting to the ethics committee, details

Box 3. The key areas of a research report for a scientific journal

Title: Describes in summary the main purpose of the study.

Abstract: A concise summary of the research approx 150 –500 words.

Introduction: Providing a background to the study, previous research in the area, purpose of the research, research question or hypothesis to be tested.

Methods: A detailed description of the procedures including: the number of study participants and their characteristics and how selected; equipment and materials used including a description of the questionnaires/interview schedule and tests used; how these were scored; the reliability and validity of scales and tests used; description of the setting where the study took place; statistical methods used to analysis data; duration of the study.

Results: Number of participants and their descriptive statistics (e.g. age, sex etc); description of the quantitative data analysis findings using where possible tables, graphs and figures.

Discussion: Was the research question answered or hypothesis supported? What was the relationship between the study's findings and previous research? What are the implications of the research findings? What questions has the research raised and what are the suggestions for further research? What are the limitations of the research and how could the research have been improved?

Conclusion: A summary of the purpose of the research and its key findings and implications.

References: All references cited in the report in standard format or as requested by the scientific journal.

For a simple guide on reporting of survey results see Fink (1995d)

of the research, including who will take part in the research, what the research will comprise, what measures will the research use, storage of data etc., as well as copies of any questionnaires scales, tests and interviews schedules that will be used. It is also now common practice for journals before publishing research studies to be assured that ethics approval has been granted for the research. Gaining informed consent from the study participants is essential, and in doing so the researcher must consider such issues as the ability of the participant to give consent, including mental ability, age, level of literacy and language.

KEY POINTS

- Quantitative research strategies can be broadly classified into experimental and non-experimental (descriptive) designs.
- Experimental designs are characterized by the manipulation or introduction of some variable such as treatment and comparing the outcome with a control group.
- One of the most common ways of collecting information in quantitative studies is the survey, which almost always uses self-completion questionnaires, face-to-face or telephone interviews, or tests and scales.
- Every approach to the analysis of quantitative data will involve some statistical manipulation, which can range from organizing the data, to providing a simple descriptive account of the findings, to very complex analysis.
- The key concepts for establishing the trustworthiness of quantitative research are reliability and validity.

Conclusion

In describing some of the key considerations and methods for undertaking quantitative research, this article has sought to convey a single important message: that, as with qualitative research, careful planning and a well formulated research strategy are essential before a study can commence. The next article in this series will examine the process of formulating a questionnaire. ■

Argyrous G (2000) *Statistics for Social and Health Research*. Sage, London

Barnett V (2002) *Sample Survey Principles and Methods*. Arnold Hodder Hadline Group, London.

Burgess AP, Carretero M, Elkington A et al (2000) The role of personality, coping style and social support in health-related quality of life in HIV infection. *Qual Life Res* **9**(4): 423–37

Campbell DT, Stanley JC (1963) Experimental and quasi-experimental designs for research on teaching. In: Gage NL, ed. *Handbook of Research on Teaching*. Rand McNally, Chicago

Chadra Ps, Chaturvedi SK, Channabasavanna et al (1998) Psychological well-being among cancer patients receiving radiotherapy – a prospective study. *Qual Life Res* **7**(6): 495–500

Fink A (1995a) *The Survey Kit. Vol 5. How to Design Surveys*. Sage Publications, Thousand Oaks, CA

Fink A (1995b) *The Survey Kit. Vol 6. How to Sample in Surveys*. Sage Publications, Thousand Oaks, CA

Fink A (1995c) *The Survey Kit. Vol 8. How to Analyze Survey Data*. Sage Publications, Thousand Oaks, CA

Fink A (1995d) *The Survey Kit. Vol 9. How to Report on Surveys*. Sage Publications, Thousand Oaks, CA

Gilbert N (2001) *Researching Social Life*. 2nd edn. Sage Publications, London

Hedov G, Anneren G, Wikblad K (2000) Self-perceived health in Swedish parents of children with Downe's syndrome. *Qual Life Res* **9**(4) 415-414

Hoinville G, Jowell R, and associates (1978) *Survey Research Practice*. Heinmann Educational Books, London

Litwin MS (1995) *The Survey Kit. Vol 7. How to Measure Survey Reliability and Validity*. SAGE Publications, Thousand Oaks, CA

Mays N, Pope C (1995) Qualitative research: rigour and qualitative research. *BMJ* **311:** 109–12

Meadows KA (2003a) So you want to do research? 1: An overview of the research process. *Br J Community Nurs* **8**(8): 369–75

Meadows KA (2003b) So you want to do research? 2: Developing the research question. *Br J Community Nurs* **8**(9): 397–404

Meadows KA (2003) So you want to do research? 3: An introduction to qualitative methods. *Br J Community Nurs* **8**(10): 464–9

Miles MB, Huberman AM (1994) *Qualitative Data Analysis: An Expanded Source Book*. 2nd edn. Sage Publications, Inc, Thousand Oaks, CA

Oppenheim AN (1992) *Questionnaire Design, Interviewing and Attitude Measurement*. Pinter Publishers, London and New York

Pope C, Mays N (1995) Qualitative research: reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ* **311:** 42–5

Robson C (2002) *Real World Research: A Resource for Social Scientists and Practitioners–Researchers*. 2nd edn. Blackwell, Oxford

Siegel S, Castellan NJ (1988) *Nonparametric statistics for the behavioural sciences*. McGraw-Hill Inc

Weich S, Lewis G (1998) Poverty, employment and common mental disorders: population based cohort study. *BMJ* **317:** 115–19

Zigmond AS, Snaith RP (1983) The Hospital Anxiety and Depression Scale. *Acta Psyhiatr Scan* **305:** 361–70