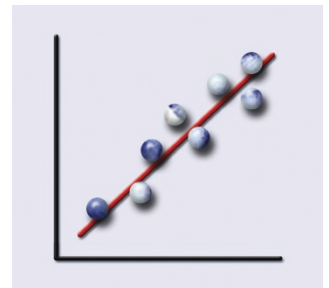**IN BRIEF**

- Choice of measures in research is an important, though at times confusing, step. This paper provides a guide to the properties that need to be considered in choosing a measure.
- Different types of measure are described and examples of each given. Broadly these include: clinical outcome measures, patient-based measures and laboratory tests.
- The steps involved in devising a questionnaire measure are described in detail.

Research in primary dental care Part 4: Measures

A. C. Williams[1], E. J. Bower[2] and J. T. Newton[3]

An important part of the design process for any study is deciding how to assess and measure the phenomena you are aiming to research. In this article we will use the term outcome measure to indicate the tool or index that can be used to quantify, in an accurate and predicable way, the phenomenon under investigation. This might, for example, be a patient's opinion, the effect of a treatment, the outcome of an intervention, or the prevalence of an oral disease. The paper presents an overview of outcome measures which can be used in research. We will discuss the properties to look for in a measure, suggest suitable evidence-based measures and briefly introduce how to go about developing new measures, in particular, questionnaires.

RESEARCH IN PRIMARY DENTAL CARE

1. Setting the scene
2. Developing a research question
3. Designing your study
4. **Measures**
5. Devising a proposal, obtaining funding and ethical considerations
6. Data analysis
7. Writing up your research

[1]Consultant Senior Lecturer in Orthodontics, Department of Child Dental Health, University of Bristol Dental School, Lower Maudlin St, Bristol BS1 2LY; [2]GDP, Staff Dental Service, Eastman Dental Hospital, 256, Gray's Inn Road, London WC1X 8LD; [3]*Professor of Psychology as Applied to Dentistry, Department of Dental Public Health & Oral Health Services Research, GKT Dental Institute, Caldecot Road, London SE5 9RW
*Correspondence to: Prof. J. T. Newton
Email: Tim.Newton@kcl.ac.uk

So how do you go about choosing your measure? The choice of measure will depend on the question that you are aiming to answer, and practicalities, such as the environment in which the data will be collected. The expertise and experience of your data collectors, and the time and funding that you have available for the study also influence the choice of measure. Ethical considerations dictate that you should use non-invasive methods wherever possible to measure the outcomes for your study. Figure 1 summarises the properties of an ideal outcome measure.[1,2]

For the results of a study to be meaningful, you should be able to demonstrate that the outcome measure you have chosen is *valid* (it measures what it claims to measure) and *reliable* (it measures the same thing in the same way on two separate occasions). For example, if an index is supposed to be measuring dental caries, it must only measure caries and not enamel hypoplasia *(validity)*. The index should yield the same caries score if a patient is examined by two different examiners *(inter-examiner reliability)* or by the same examiner on two different occasions *(intra-examiner reliability)*. The measure should also be able to detect small changes in the condition. This is known as the measure's *precision*.[2]

Sometimes, a diagnostic measure is required to demonstrate the presence or absence of a condition. The validity of a diagnostic measure is gauged by its sensitivity and specificity. The *sensitivity* of a diagnostic measure is its ability

Fig. 1 Properties of an ideal outcome measure

- Valid
- Reliable
- Acceptable to the patient
- Non-invasive
- Sensitive to changes in clinical status (precision)
- Clinically meaningful
- Easy to learn
- Quick to use
- Requires a minimum of equipment
- Can be used in the clinical setting
- Easy to record
- Evidence-based

to detect the condition when present, and the *specificity* of such a measure is its ability to exclude the condition when absent. For example, it is vital that an oral cancer screening test is accurate enough to pick up the earliest mucosal changes (sensitivity) without giving a false positive (specificity).
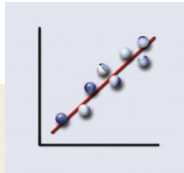
The measure should also be *easy* and *cheap* to use in the clinical situation and *acceptable* (non-traumatic) to the subject. There is no point in choosing a measure if it is difficult to use. It will take a long time to train your data collectors and your examiner reliability is likely to be low. It may also take a long time to collect the data, which will be unacceptable to the study participants.

Judging the *appropriateness* of the measure

for your study involves putting all these factors together. The perfect measure does not exist and a decision has to be taken about what aspects of the measure are most important for your study. For example, when attempting to accurately record changes in periodontal disease activity in a randomised controlled trial assessing the effectiveness of a particular periodontal therapy, it may be appropriate to record loss of attachment or the presence of bleeding on probing (high precision but slow to administer). However, when surveying the periodontal health of a large number of people in a population survey, you might choose the Basic Periodontal Examination (BPE) (reasonable sensitivity and specificity as a screening test, low precision but quick to administer).[1] Assessing the required reliability of the measure is also important. For example, if you wish to compare light-curing times for two different types of composite restoration, then a difference of 5 seconds in recording times between examiners will be unacceptable. If, however, you are recording the time that it takes to remove a resin-retained restoration, then you may decide to accept a wider margin of inter- or intra-examiner variability. Always ask yourself how realistic and meaningful the measures are *clinically* and assess the potential impact of the results on your daily clinical practice.

When deciding which outcome measure to use for your study you have two choices — either to use an existing measure or to develop a new measure. There is a wide array of measures available that have already been tested for validity, reliability, ease of use etc. Wherever possible, these *evidence-based* measures should be used in your study. Not only will the authors of the measure have tested its properties (although it is worth confirming that this is the case from the literature), but there may also be comparable data available from previous studies so that you can put the results of your own study into context. If you decide to develop your own measure, for example a questionnaire, then a considerable amount of time will need to be allocated at the beginning of your study to developing and testing it. Ethics committees are more likely to approve a study in which an appropriate evidence-based measure is chosen, if available, rather than a measure which you have developed.

It is important to undertake a thorough review of the literature to identify suitable evidence-based measures for your study. You should also study the details (if they are given) of the data collection process. This will allow you to decide whether it will be practical to use that particular measure in your study situation and to gauge the type of equipment, personnel and training required.

It is also important to appreciate that if you decide to modify a measure to suit your study objectives, then, strictly speaking, this invalidates the measure and so you will not be able to compare your results with previous studies. This is a particular problem with questionnaire studies where researchers often 'cherry-pick' questions from existing questionnaires or alter the wording of questions to suit their study. If you

> It is important to undertake a thorough review of the literature to identify suitable evidence–based measures for your study

---

Fig. 2 Examples of outcome measures

CLINICAL OUTCOME MEASURES

Measuring caries

DMFT/dmft[8]

British Association for the Study of Community Dentistry (BASCD) criteria for caries[9]

Clinical caries diagnostic criteria[10]

Measuring gingivitis and periodontal disease

Community Periodontal Index of Treatment Needs (CPITN) or Basic Periodontal Examination (BPE)[11]

Gingival Index[12]

Periodontal pocket probing depth[13]

Clinical attachment level[13]

The Extent and Severity Index[14]

Measuring enamel defects or fluorosis

Modified Dental Enamel Defects Index (DDE)[15]

Tooth Surface Index of Fluorosis (TSIF)[16]

Thylstrup-Fejerskov Index of Fluorosis (TF)[17]

Measuring dental erosion or tooth wear

Larsen et al erosion index[18]

Lussi erosion index[19]

Smith and Knight Tooth Wear Index[20]

Measuring orthodontic treatment need

Index of Orthodontic Treatment Need (IOTN)[21]

Modified Index of Orthodontic Treatment Need for oral health surveys (Modified IOTN)[22]

Measuring dental trauma

Trauma index[7]

PATIENT BASED OUTCOME MEASURES

Quality of life measures

Oral Health Impact Profile (OHIP)[23]

Shortened Oral Health Impact Profile (OHIP 14)[24]

Oral Impacts on Daily Performance measure (OIDP)[25]

Geriatric Oral Health Index of Assessment (GOHAI)[26]

Patient satisfaction

Dental Visit Satisfaction Scale[27]

do decide to do this then the data you collect will not be scientifically meaningful unless you have gone through the process of demonstrating that your new 'hybrid' questionnaire is both valid and reliable.

TYPES OF OUTCOME MEASURES

Outcome measures can be divided into three broad categories:

- Laboratory measures
- Clinical measures
- Patient-based measures (including quality of life measures)

Laboratory tests

Laboratory tests can take the form of biological tests on cell cultures or mechanical tests on clinical materials. Although you are unlikely to be involved directly in laboratory testing as part of research study in general practice, you may be interested in testing a material in the clinical setting. It is beyond the scope of this paper to discuss laboratory tests in detail. It is worth finding out however, how a material has been tested in the laboratory, for example, before you proceed to use it in a clinical study.

Clinical outcome measures

Clinical outcome measures tend to vary in their sensitivity, precision, ease of use, cost etc. depending on whether they have been developed for clinical trials, where the precision of the measure might be more important than the time taken to use it or its cost, or epidemiological studies, where speed and ease of use might be more important. However, as some standard epidemiological indices have been successfully used for clinical trials, what really matters is the suitability of the measure for *your* study. This is dictated by the aims of the study, the conditions in which you propose to collect the data and the funds and expertise available etc. as discussed previously.

It is beyond the scope of this paper to discuss such outcome measures in detail however, the papers by Pitts,[3] Azzopardi *et al.*[4] and Philstrom[5] give useful summaries of tools that have been developed for clinical trials regarding caries, erosion and periodontal disease respectively. The outcome measures detailed in the *1998 Adult Dental Health Survey*[6] and the *1993 Child Dental Health Survey*[7] are all suitable for primary care research. Some examples of clinical outcome measures are given in Figure 2.

Training and calibration

Once you have chosen your clinical outcome measure you will need to arrange for your data collectors (examiners) to undergo a period of training and calibration in the use of the measure. If your data collectors are not calibrated at the beginning of your study, then you will be unable to compare the results of your study with previous or future studies using the same measure. Calibration with an expert should also be used to check for possible drift in diagnostic levels from the beginning to the end of a study. A training programme is usually arranged prior to a calibration exercise, otherwise calibration is extremely difficult to achieve. Some authors run regular calibration courses for their measure and so it is worth finding out when and where these courses are held at an early stage in the study design process. If there are no suitable courses available, then it is worth contacting your local dental school or consultant in dental public health to find out whether there is an experienced examiner available locally who could provide a 'gold standard' for your data collectors to calibrate to.

Undertaking a calibration exercise

As an example, you may wish to use a test to identify caries. First we ask the 'gold standard' examiner to examine a sample of people using the test. You then arrange for the new examiner to examine the same sample of people under the same conditions and cross-tabulate their answers with those given by the gold standard clinician. Some theoretical data for this are shown in Table 1.

One measure of agreement is to calculate the percentage agreement (80% in this case) or the percentage disagreement (20%). The difficulty with this is that we might get high levels of agreement by chance particularly if the phenomenon we are interested in does not happen often. For example, if the new examiner had simply said 'No' to all cases, we would still get 70% agreement (since the 'gold standard' clinician said that 70% of people did not have caries). We can correct for this by using the kappa statistic. Kappa measures the proportion of agreement over and above that which might be expected by chance alone.[28] It can be used to assess the extent of agreement between examiners, or between alternative classification or diagnostic methods. It is worked out as follows:

$$\kappa = \frac{P_a - P_c}{1 - P_c}$$
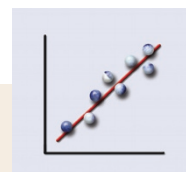
where:

P_a = observed proportion of agreement
P_c = proportion of agreement expected by chance

Kappa ranges from 0 (no agreement) to 1 (perfect agreement). Different values of kappa can be assigned labels as follows:[29]

0–0.6   Poor agreement
0.6–0.8   Satisfactory agreement
0.8–1.0   Excellent agreement

In the example given in Table 1, kappa = 0.47 which shows poor agreement when correcting for chance, despite the initial impression of reasonable agreement between the examiners.

In the example above we have shown the simplest case where the outcome (caries/no caries) is nominal (categorical). It is also possi-

Once you have chosen your clinical outcome measures you will need to arrange a period of training and calibration in the use of the measure

Table 1 Measuring agreement. Theoretical data comparing clinician's gold standard with caries identification by a new examiner

	Gold Standard Clinician's opinion		
	+ (Caries present)	− (Caries absent)	Total
New examiner	+ (Caries present) 15	5	20
	− (Caries absent) 15	65	80
Total	30	70	200

Percentage agreement = 80%
Percentage disagreement = 20%
kappa = 0.47

Patient perception of the delivery and outcome of healthcare is increasingly being used to assess the effectiveness of treatments in the NHS

ble to calculate agreement for ordinal (ordered categorical) outcomes. In this case we can use a similar statistic, the weighted kappa. This again calculates agreement corrected for chance but also weights disagreements according to how far apart the scores are on the scale. Some computer statistical packages, for example *SPSS* (see Part 6 in this series), have the facility to calculate kappa scores. It is important to consult a statistician or an experienced researcher for expert advice if you need to undertake a calibration exercise for your study.

Testing intra- and inter-examiner agreement
Even if your examiner(s) are experienced data collectors and have been calibrated in the use of a measure you will still need to test whether they are recording an outcome reliably. This involves testing both intra- and inter-examiner reliability. If you are using categorical data to record your outcomes then you should use kappa or weighted kappa as described previously. If the outcome or effect to be measured is continuous data, for example, tooth-brushing time, then you will need to calculate the variation between each set of readings compared with their mean.[30] Again, it is advisable to seek expert help at this stage.

Patient based outcome measures
Patient perception of the delivery and outcome of healthcare is increasingly being used to assess the effectiveness of treatments in the NHS.[31] It is therefore important, not only to assess the clinical impact of an intervention, but also to record the acceptability of the treatment process to the patient and its impact on their quality of life.

A number of specific quality of life measures for use in dental research have been developed.[32] The Oral Health Impact Profile (OHIP)[22] has been very widely used for assessing the impact of dental and oral disease on quality of life in adults, and a shortened 14-item version of the measure (OHIP 14)[23] was included in the most recent *Adult Dental Health Survey.*[6] OHIP 14 which can be very useful in primary care settings.[33] Some other examples of patient based outcome measures are given in Figure 2.

Such measures can be used in cost-effectiveness studies (*see* later) to compare the value of different types of dental treatments. Since special skills are required to administer the measures and analyse the data, it is worth seeking expert advice at an early stage if you are interested in undertaking this type of study.

It is possible that because a suitable measure has not yet been developed, you will wish to design your own measure to assess the opinions of patients and/or staff as part of your study. This measure will usually take the form of a questionnaire.
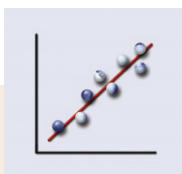
DESIGNING A QUESTIONNAIRE
Questionnaire design is a science all of its own. It is not acceptable just to scribble a few questions on a sheet of paper and then to post this out to a selected sample of patients whom you hope will give you positive feedback about your practice! Ideally your questionnaire should be thoroughly tested to show that it is both valid and reliable. This is a lengthy process and requires advice from an experienced researcher. As with clinical measures, rather than reinventing the wheel, search the literature to find examples of existing questionnaires that might be suitable for your study. If you have no choice but to develop your own questionnaire then Figure 3 lists the steps that you need to follow.

Specify your research question
As with choosing your clinical measures you should always refer back to the aims and objectives of your study to ensure that your questionnaire remains focused on the research question. This will ensure that the questionnaire remains as short as possible which is important for maximising response and completion rates.[34]

Decide the method of administration
Questionnaires may either be self-completed by the subject or filled in by the interviewer in a structured face-to-face or telephone interview with the subject. Self-completed questionnaires are cheaper to administer than structured interviews because they do not require the skills of a trained interviewer. There is no opportunity however, for subjects to explain the reasons behind their responses. Furthermore, you cannot be sure that people have interpreted your questions as you intended and there is no opportunity for further explanation if a question is ambiguous or unclear.

Self-complete questionnaires may either be posted to subjects for completion at home, or the respondents can be asked to complete the questionnaire in the waiting room, for example. The latter method is cheaper and likely to give you a better response rate. However, patients may feel intimidated in the practice setting and may give you the responses that they think you would like to hear, rather than giving their own opinions. Postal surveys are likely to give you more realistic data but are more expensive and complicated to administer.

Write your questions

Previous studies suggest that the wording of a question is an important influence on the response given.[35] Subjects are also more likely to complete a questionnaire if the issues covered are relevant to them. Questions need to be short, unambiguous and specific. It is particularly important to avoid asking leading questions or two questions in one. It is also unwise to ask about events that happened more than 6 months in the past since subjects are unlikely to be able to recall events.

The order of the questions in a questionnaire is also very important. You should aim to relax respondents at the start of the questionnaire by asking general questions first. Personal questions relating to age, sex, ethnic group, social class etc. should be kept to the minimum required to answer the research question and left until the end. This has the advantage that if people do not want to give you personal details they will have at least answered most of your questions by this point and so you will have some data to work on.

It is also very important to give clear navigation instructions so that people can find their way easily from one question to another. It is a good idea to use a different font size or type for the instructions so that they stand out separately from the questions.

Formulate your answers

Questions can be either 'open-ended', in which the subject gives or writes the answer in their own words, or 'closed', where the respondent chooses from a list of possible responses. It is a good idea to use closed questions wherever possible in self-complete questionnaires because these are quicker to complete and easier to code and analyse. You will need to spend some time pre-piloting (*see* later) closed questions to ensure that you have covered the complete range of possible answers. The answers to closed questions can be presented in a variety of ways, for example, tick boxes (*see* Q.11 below), Likert scales or visual analogue scales (*see* Fig. 4). It is well worth seeking advice from an experienced researcher about the most appropriate format for each of your questions.

It is also important to consider in the design stage how the data from your completed questionnaires will be entered into your database and by whom. Each question should correspond with a variable in the database; the sim-

Fig. 3 Questionnaire design

1. Specify your research question and the associations that you wish to test
2. Decide how you are going to administer the questionnaire
3. Write your questions
4. Formulate your answers
5. Decide how you are going to record and analyse your data
6. Design questionnaire, cover sheet and letter of explanation
7. Pre-pilot
8. Amend
9. Obtain ethical approval
10. Test validity
11. Pilot
12. Test reliability
13. Amend
14. Retest if necessary

Fig. 4 Lickert and visual analogue scales

Example of a Lickert scale

Below is a list of statements about working in a dental practice. Please indicate whether you agree or disagree with each statement by circling the number that best represents your answer:

'I feel valued as a member of the dental team'

Strongly agree	Agree	Undecided	Disagree	Strongly disagree
1	2	3	4	5

Example of a visual analogue scale

Please make a mark on the line below to indicate what your pain feels like <u>at this moment</u>:

No pain at all_____**Worst pain ever**

plest way to do this is to use the question number as the title for your variable. You should also assign a code number to each possible response to a question and mark this on the questionnaire. For example:

Q.11 Have you been to the dentist this year?

Yes	❑	1
No	❑	0
Can't remember	❑	2

In this question, the variable name is 'Q11' and the possible responses are coded as Yes = 1, No = 0 and Can't remember = 2.

Adding codes to the questionnaire from the beginning means that you can enter data directly into the database from the completed questionnaires. If you are entering your data onto a spreadsheet such as *Microsoft Excel©*, it is important to assign a code number, for example, '9', for a non-response so that it is clear in the analysis that the respondent failed to answer the question rather than a box being left empty due to a data-entry error. Responses to questionnaires can also be scanned into a database using an optical mark reader. This is expensive and

Table 2

Useful texts and source books

Outcome measures

1. Burt B, Eklund S. *Dentistry, Dental Practice, and the Community.* Philadelphia: WB Saunders Company, 1999. (particularly Chapters 13-17)

Questionnaire design and administration

1. Bowling A. *Research methods in health: Investigating health and health services.* Buckingham: Open University Press, 2002. (particularly Chapters 6, 11 and 12).
2. Frazer L, Lawley M. *Questionnaire design and administration.* Milton, Australia: Wiley, 2000. (very accessible!)
3. Dillman D. Mail and telephone surveys: the total design method. New York: Wiley, 1978.
4. Salant P, Dillman D. *How to conduct your own survey.* New York: Wiley, 1994.
5. Bulman J. *Questionnaires.* London: Faculty of General Dental Practitioners (UK), 1997.

Economic evaluation

1. Jefferson T, Demicheli V, Mugford M. *Elementary Economic Evaluation in Health Care.* London: BMJ Publishing Group, 1996.

only suitable for simple 'tick-box' responses. The expense of optical mark scanning can therefore only be justified for large surveys using very simple questionnaires.

Design the questionnaire layout, cover sheet and a letter of explanation

A subject is more likely to complete a questionnaire that is attractive, well-presented and easy to answer. The questions should flow in a logical order with clear directions about how to move from question to question. It is worth using brightly coloured paper for the cover to capture the subject's attention. White paper should be used for the questions, however, because this is less tiring to the eyes. Clear instructions should be given on the inside of the front cover about how to answer the questions. It is also important to emphasise that all the data collected will be anonymised. Do not forget to include a box for the subject's ID number on the front of the questionnaire.

You will also need to formulate a letter of explanation about the study for your subjects. Again this should be based on simple, specific sentences. You should explain the purpose of the study, emphasise its importance and its relevance to them and stress that all the data collected will be anonymised. The letter should also include a contact number and state that if a patient does not wish to participate in the study this will not affect their future dental care. You will also need to obtain written consent from study participants. Most ethics committees provide standard consent forms that you can adapt for your study. The questionnaire and any accompanying paper work will need to be approved by the local ethics committee, and again it is worth seeking advice from an experienced researcher about the most appropriate wording for letters and other documentation.

Maximising the response rate for postal questionnaire surveys

As we discussed previously, the response rate for postal questionnaire surveys is often lower than that for interview surveys.[2] Nevertheless, the evidence suggests that the response rate

> There are specific steps you can take to maximise the response rate to your postal survey

for postal surveys can be maximised by sending the questionnaires by first class post, using short questionnaires, using coloured ink, personalising the letters (addressed to a named person and signed personally), using follow-up reminders and sending second copies of the questionnaire to non-respondents, and keeping questions of a sensitive nature to a minimum.[34]

Pre–pilot

Before you undertake a formal pilot (which will need ethical approval) it is worth asking friends and colleagues to read through your questionnaire to identify any mistakes or ambiguities. If your measure is aimed at patients than ask a non-dentally qualified friend to check the questionnaire for medical jargon. It is worth spending the time testing, amending and retesting before you proceed to a formal pilot.
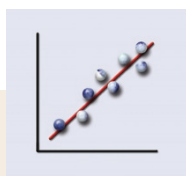
Pilot study

Ideally the questionnaire should be piloted by subjects similar to those who will be included in your study. If you will be surveying patients then ethical approval will be required for the pilot as well as the main study. During the pilot you should assess whether the questionnaire is easy to complete, valid and reliable. To test validity you need to be able to establish whether the respondents interpret the questions in the way that you intended. If the questionnaire asks for factual data, for example about attendance patterns, then this can be validated, with the patient's permission, from their dental records

The reliability of a questionnaire should be tested in two ways. Firstly, the internal consistency of the measure can be assessed by asking a question twice in two different ways in the questionnaire and then comparing the answers. This is only important for long questionnaires and is measured using a statistic known as Cronbach's alpha.[28] Secondly, test-retest reliability can be assessed by asking subjects in the pilot to complete the questionnaire on two separate occasions and comparing their responses (you can use the same statistical tests as described for the calibration study). It is worth remembering that short, specific and simple questions are more likely to be valid and reliable than long, rambling ambiguous ones. If you discover major problems with the validity or reliability of your questionnaire during the pilot then you will need to amend the questionnaire and repeat the pilot with a second group of subjects before proceeding to the main study.

INPUT MEASURES

If the aim of your study is to examine the effect of a treatment or other factors on an outcome then you will also need to decide and specify how these issues, which are termed 'input factors', will be recorded in your study and treated

in the analysis. Input factors can be divided into treatment factors and demographic factors. It is likely that you will wish to record treatment factors from the patient's clinical records. As stated previously, you must obtain the patient's written consent for this. It is worth spending some time defining exactly how these data are going to be measured and recorded, and then piloting the process to avoid observer bias.

Demographic factors are features of the participants such as their age, sex, marital status, ethnic group and socio-economic status. Socio-economic status can be assessed in several ways, for example, using the subject's occupation to classify their social class. The Registrar General's Classification of Occupation,[36] or the more recent Socio-Economic Classification[37] are examples of measures of socio-economic status based on the subject's occupation. Another way to assess a subject's social class is to use their postcode. Several such indices have been developed such as ACORN,[38] the Carstairs Index,[39] the Townsend Index,[40] and the Jarman Underpriviledged Area Score.[41] Using postcodes to assess social status is attractive because this information is readily available from patient records and so it is not necessary to ask subjects for personal information. Locker[42] gives an excellent review of deprivation measures that have been used in dental health studies.

MEASURING COSTS

In addition to assessing the effect that a new treatment or method of delivering care has on clinical outcome you may also wish to compare the costs of one procedure in comparison with another. This is known as an economic evaluation. Several different types of economic analyses are used in healthcare depending on the methods that are used to measure the outcomes of two treatments (*see* Jefferson *et al.*)[43]. Put simply, you should only undertake an economic evaluation if you expect there to be significant differences in the costs of delivering two types of treatment. For example, if you are comparing one type of composite restoration with another, since most of the costs of placing the restoration will be clinician-time which will be similar in each group, there is no need to include these costs in your analysis. If however, one type of restoration then needs to be replaced on several occasions during your study period this could be recorded in terms of clinician-time, the costs of the materials and also the cost to the patient of attending the practice for additional appointments. As with other potential impacts on the results of your study you should try to list these costs at the beginning of your study and make provision to collect these data as the study proceeds. Trying to assess costs retrospectively is difficult and open to bias. Again you should seek expert advice if you are interested in undertaking an economic evaluation as part of your study.

SUMMARY

In this article we have described the different types of outcome measure that are available to researchers and the factors that you need to consider when choosing a measure for your study. We have also given details of how to design and test your own measure, in this case, a questionnaire. When deciding which outcome measures to use it is important to remember to continually refer back to the aims and objectives of your study to ensure that your measures are appropriate for answering your research question.

1. Daly B, Watt R, Batchelor P, Treasure E. *Essential Dental Public Health.* Oxford: Oxford University Press, 2002.
2. Bowling A. *Research methods in health: Investigating health and health services.* Buckingham: Open University Press, 2002.
3. Pitts N B. Diagnostic tools and measurements — impact on appropriate care. *Community Dent Oral Epidemiol* 1997; **25:** 24-35.
4. Azzopardi A, Bartlett D W, Watson T F, Smith B G. A literature review of the techniques to measure tooth wear and erosion. *Eur J Prosthodont Restor Dent* 2000; **8:** 93-97.
5. Philstrom B. Issues in the evaluation of clinical trials of periodontitis: a clinical perspective. *J Periodont Res* 1992; **27:** 473-441.
6. Kelly M, Steele J, Nuttall N *et al. Adult dental health survey: Oral health in the United Kingdom 1998.* London: The Stationary Office, 2000.
7. O'Brien M. *Children's Dental Health in the United Kingdom 1993.* London: HMSO, 1994.
8. Klein H, Palmer C E, Knutson J W. Studies on dental caries: I. Dental status and dental needs of elementary school children. *Public Health Rep* (Washington), 1938; **53:** 751-765.
9. Pitts N B, Evans D J, Pine C M. British Association for the Study of Community Dentistry (BASCD) diagnostic criteria for caries prevalence surveys — 1996/7. *Community Dent Health* 1997; **14:** 6-9.
10. Nyvad B, Machiulskiene V, Baelum V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. *Caries Res* 1999; **33:** 252-260.
11. Ainamo J, Barmes D, Beagrie G *et al.* Development of the World Health Organisation (WHO) Community Periodontal Index of Treatment Needs (CPITN). *Int Dent J* 1982; **32:** 281-291.
12. Loe H, Silness J. Periodontal disease in pregnancy. I. Prevalence and severity. *Acta Odontol Scand* 1963; **21:** 533-551.
13. Heasman P A, McCracken G I, Steen N. Supportive periodontal care: the effect of periodic subgingival debridement compared with supragingival prophylaxis with respect to outcomes. *J Clin Periodontol* 2002; **29** (Suppl): 163-172.
14. Carlos J P, Wolfe M D, Kingman A. The extent and severity index: a simple method for use in epidemiological studies of periodontal disease. *J Clin Periodontol* 1986; **13:** 500-505.
15. Clarkson J, O'Mullane D. A modified DDE Index for use in epidemiological studies of enamel defects. *J Dent Res* 1989; **68:** 445-450.
16. Horowitz H S, Driscoll W S, Meyers R J *et al.* A new method for assessing the prevalence of dental fluorosis-the Tooth Surface Index of Fluorosis. *J Am Dent Assoc* 1984; **109:** 37-41.
17. Thylstrup A, Ferjerskov O. Clinical appearance of dental fluorosis in permanent teeth in relation to histologic changes. *Community Dent Oral Epidemiol* 1978; **6:** 315-328.
18. Larsen I B, Westergaard J, Stoltze K *et al.* A clinical index for evaluation and monitoring dental erosion. *Community Dent Oral Epidemiol* 2000; **28:** 211-217.
19. Lussi A. Dental erosion: clinical diagnosis and history taking. *Eur J Oral Health Sci* 1996; **104:** 191-198.
20. Smith B G, Knight J K. An index for measuring the wear of teeth. *Br Dent J* 1984; **156:** 435-438.
21. Brook P H, Shaw W C. The development of an index of orthodontic treatment priority. *Eur J Orthodont* 1989; **11:** 309-320.
22. Burden D J, Pine C M, Burnside G. Modified IOTN: an orthodontic treatment need index for use in oral health surveys. *Community Dent Oral Epidemiol* 2001; **29:** 220-225.

23. Slade G D and Spencer A J. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health* 1994; **11:** 3-11.

24. Slade G D. Derivation and validation of a short form Oral Health Impact Profile. *Community Dent Oral Epidemiol* 1997; **25:** 284-290.

25. Adulyanon S, Sheiham A. Oral impacts on daily performance. *In:* Slade G D (ed) *Measuring oral health and quality of life.* Chapel Hill: University of North Carolina; 1997.

26. Atchison K, Dolan T. Development of the geriatric oral health assessment index. *J Dent Educ* 1990; **54:** 680-687.

27. Corah N L, O'Shea R M, Pace L F, Seyrek S K. Development of a patient measure of satisfaction with the dentist: the Dental Visit Satisfaction Scale. *J Behav Med* 1984; **7:** 367-373.

28. Banerjee A. *Medical statistics made clear: an introduction to basic concepts.* London: Royal Society of Medicine Press Ltd, 2003.

29. Altman D G. *Practical statistics for medical research.* London: Chapman & Hall/CRC, 2003.

30. Bulman J S, Osbourn J F. *Statistics in Dentistry.* London: BDJ Books, 1989.

31. Department of Health. *The new NHS - modern, dependable.* London: HMSO, 1997.

32. Slade G D (Ed.) *Oral Health and Quality of Life.* Chapel Hill: University of North Carolina Dental Ecology, 1997.

33. McGrath C, Bedi R. The value and use of 'quality of life' measures in the primary dental care setting. *Prim Dent Care* 1999; **6:** 53-57.

34. Edwards P, Roberts I, Clarke M *et al.* Increasing response rates to postal questionnaires: systematic review. *Br Med J* 2002; **324:** 1183-1185.

35. Cartwright A. Some experiments with factors that might affect the response of mothers to a postal questionnaire. *Stat Med* 1986; **5:** 607-617.

36. Office of Population Censuses and Surveys. *Classification of Occupations.* London: HMSO, 1980.

37. Chandola T, Jenkinson C. The new UK National Statistics socio-economic classification (NS-SEC); investigating social class differences in self-reported health status. *J Public Health Med* 2000; **22:** 182-190.

38. Morgan M, Chinn S. ACORN group, social class and dental health. *J Epidemiol Community Dent Health* 1983; **37:** 196-203.

39. Carstairs V, Morris R. Deprivation and mortality: an alternative to social class? *Community Med* 1989; **11:** 210-219.

40. Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality and the North.* London: Croom Helm, 1988.

41. Jarman B. Identification of underprivileged areas. *Br Med J* 1983; **286:** 1705-1709.

42. Locker D. Deprivation and oral health: a review. *Community Dent Oral Epidemiol* 2000; **28:** 161-169.

43. Jefferson T, Demicheli V, Mugford M. *Elementary Economic Evaluation in Health Care.* London: BMJ Publishing Group, 1996.