Family Practice
© Oxford University Press 2000

Vol. 17, No. 2
Printed in Great Britain

The design and analysis of a randomized controlled trial to evaluate computerized decision support in primary care: the COGENT study

Martin Eccles, Jeremy Grimshaw[a], Nick Steen, David Parkin[b], Ian Purves[c], Elaine McColl and Nikki Rousseau

Eccles M, Grimshaw J, Steen N, Parkin D, Purves I, McColl E and Rousseau N. The design and analysis of a randomized controlled trial to evaluate computerized decision support in primary care: the COGENT study. *Family Practice* 2000; **17:** 180–186.

Introduction

Despite the current interest in guidelines, there remains uncertainty about how best to introduce them into practice. A systematic review of rigorous published evaluations of the introduction of clinical guidelines[1] identified 91 studies covering a wide range of clinical areas including patient management, prevention, test ordering and prescribing; 14 of the studies were conducted in the UK. Eighty-one of the 87 studies measuring the process of medical care identified improvements in the process of medical care; 13 of the 17 papers that measured the outcome of care reported improvements. Nevertheless, there were considerable variations in the range of improvement that possibly was associated with differences in any or all of the following: settings, targeted professional groups, targeted clinical activities, and how the guidelines were developed, disseminated and implemented. The review suggested that one promising method to implement guidelines involved utilizing patient-specific prompts at the time of a consultation. One means of achieving this is with computerized decision support (CDS).

CDS is a "system that compares patient characteristics with a knowledge base and then guides a health provider by offering patient-specific and situation-specific advice",[2] it offers more than merely a summary of patient data. There have been two systematic reviews[3,4] and one meta analysis[5] examining the effectiveness of CDS. Shea's meta analysis[5] focused on randomized controlled trials (RCTs) of computerized reminder systems for preventive care in ambulatory care settings. The 16 trials showed improved preventive practice for vaccination, breast cancer screening, colorectal cancer screening and cardiovascular risk reduction, but not for cervical cancer screening or 'other' preventive activities. The most recent systematic review,[4] which is an update of Johnston's earlier review,[3] identified 68 controlled trials. These showed benefit in: nine of 15 trials evaluating systems to improve drug dosing; one of five trials evaluating diagnostic aids; 14 of 19 trials evaluating systems to improve preventive care; and 19 of 26 trials of CDS in 'other medical care'. Of the 14 studies that measured patient outcomes, there were improvements in six. However, there were few studies that did not have design or analysis flaws that meant that the study had to be interpreted with caution. Therefore, whilst CDS may have considerable potential as a method of implementing guidelines, there are few UK NHS-based studies, and there are almost no studies of CDS in chronic disease management or of CDS integrated into routine computer systems.

If policy makers are to make evidence-based decisions about guideline implementation, they need information on the cost-effectiveness of different interventions (in different settings and for different targeted clinicians and behaviours), the factors likely to modify the effect of the intervention and the resources needed to deliver interventions. To provide such information, researchers need to use rigorous designs and methods to yield reliable and valid estimates of the likely effects of alternative interventions. It is then possible to have confidence that the observed effects are both attributable to the interventions studied and generalizable to other contexts. However, many existing studies use weak designs or are methodologically flawed with potentially major threats to validity, thereby limiting their value to inform decision making.[6] However, well designed, rigorous studies of guideline implementation strategies are both complex and methodologically challenging. This article describes

Received 7 September 1999; Accepted 26 October 1999.
Centre for Health Services Research, University of Newcastle Upon Tyne, Newcastle Upon Tyne NE2 4AA, [b]Department of Epidemiology and Public Health and [c]Sowerby Centre for Health Informatics, University of Newcastle Upon Tyne and [a]Health Services Research Unit, University of Aberdeen, Aberdeen, UK.

the design and proposed analysis of an RCT evaluating the use of CDS to implement clinical guidelines for the primary care management of asthma in adults and angina (computerized guidelines evaluation in the north of England; COGENT).

Study design

The study design is a before and after cluster RCT utilizing a 2 × 2 incomplete block design. Campbell and colleagues have discussed the issues of study design specific to implementation studies.[7]

Cluster randomization

As with other health care innovations, an RCT is the optimum design when evaluating behaviour change interventions. However, when evaluating changes in clinician behaviour, there is a risk that the treatment offered to control patients will be contaminated by doctors' experiences of applying the intervention to patients receiving the experimental management. The solution to this problem is to randomize groups (or clusters) of professionals while collecting data about the process and outcome of care at the individual patient level. However, as patients within any one cluster are more likely to respond in a similar manner, such a design violates the assumption that the outcome for an individual patient is completely independent of that for any other patient. Therefore, a cluster randomized design is not as statistically efficient as a patient randomized design; it has lower statistical power than a patient randomized trial of equivalent size,[8] and sample sizes need to be inflated to compensate for this. Despite the added complexity, cluster randomized trials provide the optimal design for guideline implementation studies.

Balanced incomplete block designs

In research studying human behaviour, of which implementation research is one example, there are a number of non-specific effects that may influence the estimate of the effect of an intervention. Currently, these are grouped together and termed the 'Hawthorne effect'. If such effects are imbalanced across study groups in guideline implementation trials, the resulting estimates of the effect of the intervention may be biased. Balanced incomplete block designs can be used to equalize such non-specific effects and thereby minimize their impact.[9] In the COGENT trial, study practices are allocated randomly to two groups. One receives computerized guidelines for the management of asthma and provides control data for the management of angina. The other receives computerized guidelines for the management of angina and provides control data for the management of asthma (see Table 1). Thus, doctors in both groups are subject to the same level of intervention and the Hawthorne effect is equalized across the two groups.

TABLE 1 *Balanced incomplete block design*

	Asthma	Angina
Doctor group 1	Intervention	Control
Doctor group 2	Control	Intervention

Before and after design

The patients in the trial have chronic illnesses and these may worsen with time. Therefore, patients may be more ill at the end of the study than they were at the start. To compensate as far as possible for this natural deterioration, the study evaluates change over two periods of 12 months. In addition, the ability to use antecedent data in the overall analysis will provide increased statistical power and also allows the adequacy of randomization to be checked.

Study conditions

The conditions were chosen for both clinical and methodological reasons. The conditions (asthma in adults and angina) were chosen as chronic illnesses, predominantly cared for in primary care and important because of their associated morbidity and mortality. For both conditions, there is evidence of underuse of effective treatments.[10–12] In addition, within a block design, it is important that the management of patients with the control condition is not influenced by the introduction of the guidelines for the intervention condition. For asthma and angina, the elements of care for the two conditions are quite discrete with no overlap in terms of actions performed. This would not be the case with conditions such as diabetes and ischaemic heart disease.

The clinical guidelines

Guideline development

Clinical guidelines are "systematically developed statements to assist practitioner decisions about appropriate health care for specific clinical circumstances".[13] It is important that guidelines are valid in the sense that, when followed, they actually lead to the health gains and costs predicted for them.[13,14] It is therefore important to ensure that guidelines are rigorously developed, and thus consistent with the available scientific evidence or, in the absence of such evidence, with best clinical judgement. Evidence-based guidelines for the two clinical conditions have been developed[15–17] using current optimum methods.[14]

Guideline dissemination

The guidelines will be disseminated through educational workshops discussing the content of the guidelines and the use of computer-generated prompts. Every clinician

(doctor or practice nurse) in the study will receive a paper copy of the summary version of both guidelines. In addition, each practice will receive a paper copy of the full version of both guidelines.

The computerized guidelines will present contextualized prompts to guide the consultation. These will be generated in the light of what is already known about the patient from within their electronic medical record and will both suggest appropriate actions and request the entry of appropriate data. Information for both clinician and patient may be generated on screen and in a printed form. The computerized guidelines can be used interactively within consultations in two ways. Firstly, the clinician can decide to access one or more task-related module within a routine consultation. As clinicians are often presented with two or more problems by a patient, allowing them to access chosen task-related modules (e.g. for prescribing or referral) may facilitate the support the computer can provide without compromising the conduct of the consultation. Secondly, the clinician can access the guideline within a formalized review process (e.g. a disease management clinic).

Study practices

Practices eligible to participate in the study are those with AAH Meditel or EMIS computing systems, in the Northern and Yorkshire region and with at least 50% of the GPs reporting using their practice computer system to view clinical data and for acute prescribing. In addition, they had not to be in other studies using the clinical guidelines and not using any other sophisticated CDS system. Single-handed practices are excluded from the study as their list sizes would be too small to yield the requisite number of patients. Practices were stratified by computer system and vocational training status (as a proxy for practice development[18]) prior to randomization.

In order to provide contextual information, practice managers complete a questionnaire that gathers details of the practice computer system and its use, practice characteristics and organization, and investigation and referral arrangements. The GPs and practice nurses complete a composite questionnaire composed of both previously validated instruments and new questions that gather data on: their attitudes to clinical guidelines;[19] their use of and attitudes to computers; aspects of team functioning;[20] and demographic details.

Sample size

There are no published tables for calculating the sample size needed for a 'before-and-after incomplete block'. The simplest approach is to regard the design as two embedded randomized trials (one of guidelines for angina, the other of guidelines for asthma in adults) and determine the sample size for each trial separately.

Changes in process of care

It is felt that for both conditions, doctors in the practices receiving the computerized prompt will adhere more closely to the guidelines than those in practices receiving the paper versions. It is required that each trial should have 80% power to detect a 10% difference in adherence (e.g. between 45 and 55%) using a significance level of 5%. Adherence to the guidelines will be determined by measures of process recorded in the patients' medical records. As the unit of randomization is the practice rather than the patient, it is necessary to take into account the lack of independence between observations on different patients within a single practice. It is estimated that the intraclass correlation coefficients (ICCs) for measures of process will be 0.06.[21] Using methods proposed by Donner,[8] it can be demonstrated that it is necessary to sample records from 57 patients with each condition in each of 60 practices. Different measures of process may have different ICCs. If the average ICC were 0.1, we would need to sample an average of 86 patients with each condition in each of 60 practices to give us 80% power to detect a difference of 12% (44–56%). Were the ICC 0.15, we would need to sample 40 patients in each of 60 practices to give us 80% power to detect a difference of 15% (42.5–57.5%).

Changes in outcome of care

Patient-based outcomes will be assessed using a range of generic and condition-specific measures (see Measurement of patient-based outcomes, below). Typically, these measures take the form of a summated Likert scale and can be considered as continuous variables with a normal distribution. Again it is necessary to take into account the clustering of patients within practices. An intraclass correlation coefficient of 0.06 is assumed.[21] Application of standard methods[8] indicates that if we sample 35 patients from each of 60 practices we would have 90% power to detect an effect size of 0.25 with a significance of 5% (or 80% power to detect an effect size of 0.22); 61 patients from each of 60 practices would give 90% power to detect an effect size of 0.23 with a significance of 5%.

Identification of patients

Patients aged 18 years or over and with angina or asthma are identified using a computerized query which searches for relevant morbidity, management and drug Read Codes,[22] either alone or in combination. For both conditions, the morbidity codes chosen are those that are highly suggestive of the index condition. For patients with angina, these are supplemented with Read Codes for nitrates. Searching on nitrates alone is estimated to identify >70% of patients with angina.[23] Patients with Read Codes for nitrates alone are excluded if they also have Read Codes indicative of heart failure. These are supplemented with Read Codes for inhaled β-2 agonists,

inhaled corticosteroids and asthma prophylaxis. Patients with Read Codes for asthma drugs alone are excluded if they also have morbidity Read Codes suggestive of chronic obstructive airways disease. In an effort to identify patients with current symptoms for both conditions, the drug codes are restricted to entries within the preceding 12 months. In addition, due to its more intermittent and evanescent nature, the asthma morbidity and management codes are also restricted to the preceding 12 months. The study has no upper age limit, but for both conditions any patients with Read Codes suggestive of dementia are excluded on the basis that they will not be able to complete the patient-based outcome questionnaires.

Medical record-based data collection

Data collection from the medical record will be conducted in two ways, electronically and by abstraction of data from notes. Data will be abstracted retrospectively from computerized medical records by a database query[24] for the identified patients. However, computerized records alone cannot be regarded as a complete record of care delivered. Therefore, these data will be validated and augmented by abstracting data from the patients' paper records. The choice of data items will reflect the underlying evidence base and appropriateness of the associated clinical activities.[25] These data will include both data on the process of care and data on intermediate outcomes.

In addition, a guideline usage log will record when the guidelines are used and by whom. These data will be used in two ways: firstly to indicate rates of uptake to inform the choice of interviewees in the case study (see below); and secondly as an explanatory variable in the main analysis.

Measurement of patient-based outcomes

For the measurement of patient-based outcomes, it is generally agreed that a combination of generic and condition-specific measures of outcome are required.[26] Generic measures facilitate comparisons of outcomes across conditions and are important for cost-effectiveness studies, while condition-specific measures are likely to be more responsive to small changes in outcome. The generic measures in the study are the SF36[27,28] and the EQ-5D,[29] the latter yielding a utility value. The condition-specific measure for angina is the Seattle Angina Questionnaire;[30,31] for asthma the measures of choice are the Newcastle Asthma Symptoms Questionnaire[32] and the asthma quality of life questionnaire.[33,34] The patient outcome questionnaires will be administered at three points in time: approximately a year before the intervention; just before the intervention; and approximately a year after the intervention.

Economic evaluation

The economic evaluation takes a health service perspective, but also covers costs to patients and their families. Data for the economic evaluation will be collected using standard techniques for assessing the costs and benefits of technological innovations in health.[35,36] Three types of cost data will be collected. The costs of development and implementation will be measured by questionnaires to, and interviews with, key participants in collaboration with finance departments of relevant organizations. Changes in the costs of care will be measured by linking process of care data to unit cost data. Changes in costs to patients and their families will be assessed by incorporating appropriate questions in the outcome and process data sets. Because there are both process and outcome data, and the latter include a means of assessing changes in Quality Adjusted Life Years,[29] it will be possible to carry out both cost-effectiveness and cost-utility analyses, using modelling techniques, which incorporate sensitivity analyses.

Principles of data analysis

In general, measures of process and outcome of care will be observations on individual patients (e.g. whether a measurement of blood pressure has been recorded in the patient's notes, or a patient's score on a health status measure). The unit of randomization in this study is the general practice. With these types of data, there is a choice of methods of analysis. It is possible to analyse the data using the practice as the unit of analysis or to analyse the data at the patient level allowing for the correlation of responses from patients within each practice. For example, if the variable of interest is whether an item of information is recorded in the medical notes, we could generate a practice statistic, i.e. the proportion of patients in that practice for which the item was present, and analyse this summary statistic. Alternatively, it is possible to use a range of alternative methods that use the patient observation as the unit of analysis but take into account the hierarchical structure of the data. When the design is balanced (i.e. the number of patients in each practice is the same), these two methods are equivalent. When the design is unbalanced, it is possible to make adjustments to an analysis of practice-based summary measures, but often it is more natural to use models that take into account both variation between practices and variation between individuals within practices.

The major advantage of making the patient the unit of analysis is that there are natural methods of analysing repeated measures made at the patient level. The patient outcome questionnaires will be administered at three points in time: approximately a year before the intervention; just before the intervention; and approximately a year after the intervention. The most important of

these measurements, for the purpose of the evaluation, is the one made after the intervention. However, a patient with more severe symptoms at one time point is likely to have more severe symptoms at all time points. We can take this into account by including the two initial measurements as covariates in the analysis of final outcome. It is proposed to do this using multilevel modelling. It will be assumed that outcome will vary randomly between patients about some practice mean and that practice means will vary randomly about some overall mean. The two previous measures will be included as fixed effects at the patient level. The intervention effect will be a fixed effect at the practice level.

The embedded case studies

The COGENT trial will produce a precise estimate of the effect of computerized guidelines on process and outcome of care but, beyond the impact of pre-defined explanatory variables, will offer little insight into how or why change was achieved. To gain some understanding of these issues, the study incorporates embedded qualitative case studies.[37] The case studies involve a very detailed study of an individual or organization (the 'case'). Multiple sources of data commonly are used to draw up a comprehensive picture of the case, including various types of qualitative data collection (observational work, interviews), and often also incorporate some quantitative data (though in a qualitative case study, analysis is predominantly qualitative in nature).

Whilst much variation in the use of computerized guidelines appears to be at the level of the individual, the social and organizational setting (the most important aspect of which is the practice) provides important context for the intervention. The use of case studies will allow us to gather information about the context of the evaluation in a way that would not be explicit in a design where the focus was on individual GPs selected from different practices. The practice itself is a case, and the individual health professionals within the practice are also cases (an embedded design[38]). The case study practices and their members will be followed over the lifetime of the COGENT study.

This longitudinal and embedded design was considered most appropriate for two main reasons. Firstly,

TABLE 2 *Purposive sampling frame for general practices*

Criteria	Theoretical basis	Sampling strategy
Similarity to RCT practices	Primary purpose of case study is to illuminate findings of RCT; therefore, it is important that practices are not too dissimilar to RCT practices in the following areas:	Broadly similar proportions to the RCT practices in the areas outlined below.
Computer system	We thought that the implementation of computerized guidelines might prove to be technically neater on one or other system or that there might be some other effects to do with the computer (training and back up, self-selection effects) which would enhance implementation	At least two of each computer system in our sample of 6.
Fundholding wave	There is some evidence that fundholding practices might be more responsive to change and positively inclined towards innovations than other general practices. There might be additional features associated with fundholding that would encourage practices to use the computerized guidelines	At least one non-fundholding practice and at least one phase one fundholding practice
Training practice	As above. Vocational training practices should have a positive attitude to teaching and research and will have to meet certain organizational standards.	Not more than half of our sample were training practices.
Number of partners	We felt that practice organization and communication between partners might differ between small and large practices and that this might affect the uptake of the guidelines.	A mixture of larger and smaller practices.
Part-time partners	We felt that part-time partners might find it more difficult to adopt the changes.	Some practices having part-time partners.
Already have guidelines for asthma and or angina.	These practices might be more positively inclined towards guidelines. Alternatively, differences between the two guidelines might be a disincentive to use the computerized guidelines.	A mixture of practices having and not having guidelines
Have some partners not using computer interactively	We were interested to see the impact of a computerized guideline system on those partners not currently making as much use of computers as their colleagues.	Some (but not all) practices having some partners not making interactive use of the computer.

implementation is not a one-off event but a process, with several stages that take place over time. Unless we investigated the implementation over the whole period, there was a danger we might miss something important. Secondly, health professionals do not act in isolation. While behaviour change ultimately has to take place at the level of the individual, interventions often take place at the practice level; indeed it may be difficult to stop an individual level intervention spreading to the practice. Similarly, social and organization factors beyond the individual will affect the success of an intervention. The individual health professional will be acted on by various forces both within and outside the practice; similarly, they will exert forces on their environment.

The five practices in the case study were identified from the main pool of practices both eligible and willing to take part in the RCT. The practices were selected purposively on the criteria in Table 2.

At the individual practitioner level, sampling in the case studies is using several approaches (Table 3).

While patients are clearly an important component in health services research, for the purposes of the COGENT case studies the primary focus was the health professional. We might observe patients in the consultation but we would not interview them directly.

Conclusions

Here we describe the design and analysis issues within a complex cluster RCT. Design issues such as the appropriate unit of analysis and allowance for clustering are essential for the robust evaluation of implementation studies.[21] The combination of the most robust trial design, the rigorous economic analysis and the embedded case study[37] will result in policy-relevant results that will inform the use of CDS within the NHS.

TABLE 3 *Sampling strategy for individual clinicians*

Pre-intervention
All contact persons (the person the practice identified as 'best to contact about COGENT issues') within practice
People identified as useful data sources during a preliminary interview with contact person
Results of an attitudes and experience questionnaire completed by GPs and practice nurses (including both typical and extreme cases)
Other people who become relevant as theories develop (e.g. interviews with health professionals may lead us to hypothesize that practice managers are key and they could then be interviewed).
Post-intervention
Follow-up of original cases
People mentioned during interviews
Computer use log files (recording amount and type of interactions with the computerised guidelines) to identify new cases (as well as informing discussion with existing cases being followed).
Post-RCT
Individuals from the RCT could be used to explore whether experiences in the case study practices were shared in the RCT sites.

Acknowledgements

The COGENT study is funded by the NHS R&D Programme 'Methods to promote the uptake of research findings'. The Health Services Research Unit, University of Aberdeen, is funded by the Chief Scientist Office of the Scottish Office Department of Health. The Centre for Health Services Research and the Department of Epidemiology and Public Health, University of Newcastle Upon Tyne and the Health Services Research Unit, University of Aberdeen are part of the MRC Health Services Research Collaboration. The views expressed are those of the authors and not necessarily those of the funding bodies.

References

[1] Grimshaw J, Freemantle N, Wallace S *et al*. Developing and implementing clinical practice guidelines. *Qual Health Care* 1995; **4:** 55–64.

[2] Friedman CP, Wyatt JC. *Evaluation Methods in Medical Informatics*. New York: Springer-Verlag, 1997.

[3] Johnston ME, Langton KB, Haynes B, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome: a critical appraisal of research. *Ann Intern Med* 1994; **120:** 135–142.

[4] Hunt DL, Haynes B, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. *J Am Med Assoc* 1998; **280:** 1339–1346.

[5] Shea S, DuMouchel W, Bahamonde L. A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventative care in the ambulatory setting. *J Am Med Informatics Assoc* 1996; **3:** 399–409.

[6] Bero L, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote implementation of research findings by health care professionals. *Br Med J* 1998; **317:** 465–468.

[7] Campbell MK, Steen IN, Grimshaw JM, Eccles M, Mollison J, Lombard C. Design and statistical issues in implementation research. In Thorsen T, Makela M (eds). *Changing Professional Practice*. Copenhagen: DSI: Danish Institute for Health Services Research and Development, 1999.

[8] Donner A. Some aspects of the design and analysis of cluster randomization trials. *Appl Stat* 1998; **47:** 95–113.

[9] Cochran WG, Cox GM. *Experimental Design*. New York: Wiley, 1957.

[10] Woods KL, Ketley D, Lowy A *et al*. Beta-blockers and antithrombotic treatment for secondary prevention after acute myocardial infarction. *Eur Heart J* 1998; **19:** 74–79.

[11] Griffiths C, Sturdy P, Naish J, Omar R, Dolan S, Feder G. Hospital admissions for asthma in east London: associations with characteristics of local general practices, prescribing, and population. *Br Med J* 1997; **314:** 482–486.

[12] Neville RG, Hoskins G, Smith B, Clark RA. Observations on the structure, process and clinical outcomes of asthma care in general practice. *Br J Gen Pract* 1996; **46:** 583–587.

[13] Field MJ, Lohr KN (eds). *Guidelines for Clinical Practice: From Development to Use*. Washington DC: National Academy Press, 1992.

[14] Grimshaw J, Eccles M, Russell I. Developing clinically valid practice guidelines. *J Eval Clin Pract* 1995; **1:** 37–48.
[15] North of England Stable Angina Guidelines Development Group. Evidence based guideline for the primary care management of stable angina: summary version. *Br Med J* 1996; **312:** 827–832.
[16] North of England Asthma Guideline Development Group. North of England evidence based guidelines development project: summary version of evidence based guideline for the primary care management of asthma in adults. *Br Med J* 1996; **312:** 762–766.
[17] Eccles MP, Clapp Z, Grimshaw JM *et al*. North of England evidence based guidelines development project: methods of guideline development. *Br Med J* 1996; **312**: 760–761.
[18] Baker R. General practice in Gloucestershire, Avon and Somerset: explaining variations in standards. *Br J Gen Pract* 1992; **42:** 415–418.
[19] Elovainio M, Eccles M. Attitudes towards guidelines and a scale for measuring them. In Thorsen T, Makela M (eds). *Changing Professional Practice*. Copenhagen: DSI: Danish Institute for Health Services Research and Development, 1999: 153–167.
[20] West MA. Innovation amongst health care professionals. *Social Behav* 1989; **4:** 173–184.
[21] Campbell M, Steen N, Grimshaw J, Eccles M, Mollison J, Lombard C. Design and statistical issues in implementation research. In Thorsen T, Makela M (eds). *Changing Professional Practice*. Copenhagen: DSI: Danish Institute for Health Services Research and Development, 1999: 57–76.
[22] O'Neil M *et al*. Read Codes Version 3: a user led terminology. *Methods Inf Med* 1995; **34:** 187–192.
[23] Cannon PJ, Connell PA, Stockley IH, Garner ST, Hampton JR. Revalence of angina as assessed by a survey of prescriptions for nitrates. *Lancet* 1988; **1(8592):** 979–981.
[24] Purves IN, Edwards C. Comparison of prescribing unit with index including both age and sex in assessing general practice prescribing costs. *Br Med J* 1993; **306:** 496–498.
[25] Campbell SM, Roland MO, Shekelle PG, Cantrill JA, Buetow SA, Cragg DK. Development of review criteria for assessing the quality of management of stable angina, adult asthma, and non-insulin dependent diabetes mellitus in general practice. *Qual Health Care* 1999; **8:** 6–15.
[26] Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; **27 (3 Suppl):** S217–S232
[27] Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30:** 473–483.
[28] Jenkinson C, Coulter A, Wright L. Short form 36 (SF36) health survey questionnaire: normative data for adults fo working age. *Br Med J* 1993; **306:** 1437–1440.
[29] Brooks RwtEG. Euroqol—the current state of play. *Health Policy* 1996; **37:** 53–72.
[30] Spertus JA, Winder JA, Dewhurst TA, Deyo RA, Fihn SD. Monitoring the quality of life in patients with coronary artery disease. *Am J Cardiol* 1994; **74:** 1240–1244.
[31] Spertus JA, Winder JA, Dewhurst TA *et al*. Development and evaluation of the Seattle Angina Questionnaire: a new functional status measure for coronary artery disease. *J Am Coll Cardiol* 1999; **25:** 333–341.
[32] Steen IN, Hutchinson A, McColl E *et al*. Development of a symptom-based outcome measures for asthma. *Br Med J* 1994; **309:** 1065–1069.
[33] Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschke R, Hiller TK. Evaluation of impairment of health related quality of life in asthma: development of a questionnaire for use in clinical trials. *Thorax* 1992; **47:** 76–83.
[34] Juniper EF, Guyatt GH, Ferrie PJ, Griffith LE. Measuring quality of life in asthma. *Am Rev Respir Dis* 1993; **147:** 832–838.
[35] Anonymous. *Cost-effectiveness in Health and Medicine*. New York: Oxford University Press, 1996.
[36] Drummond M, O'Brien B, Stoddart G, Torrance G. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 1997.
[37] Rousseau N, McColl E, Eccles M, Hall L. Qualitative methods in implementation research. In Thorsen T, Makela M (eds). *Changing Professional Practice*. Copenhagen: DSI: Danish Institute for Health Services Research and Development, 1999: 99–116.
[38] Yin RK. *Case Study Research: Design and Methods*. Beverly Hills, California: Sage Publications Inc., 1984.