

Shared Decision Making

Measurement using the OPTION instrument

Authors:

Glyn Elwyn, Adrian Edwards, Michel Wensing, Richard Grol

Collaborators:

Michel Cauchon, Silvia Fontanesi, Claudia Goss, Martin Härter, Michel Labrecque, Wolf Langewitz, France Légaré, Andreas Loh, Elizabeth Martínez-Gibson, Monica Paccaloni, Lidia Del Piccolo, Charlene Pope, Michaela Rimondini, Jason Robertson, Silvia Rodriguez-Sabater, Graham Thornicroft, Myrra Vernooij, Trudy van der Weijden, Hub Wollersheim, Christa Zimmermann

The work was funded by a collaboration between the Department of General Practice, School of Medicine, Cardiff University, Wales and the Centre for Quality of Care Research, Radboud University Nijmegen Medical Centre, Netherlands.

A licence to use the OPTION instrument, free of charge for research purposes, is available from the author: email elwyng@cardiff.ac.uk

Copies of the book may be ordered from: elwyng@cardiff.ac.uk

The studies presented in this book were performed with assistance from many colleagues. We would like to thank them for their hard work and support. The work was supported by the following organisations:

- Department of General Practice, School of Medicine, Cardiff University
- Centre for Quality of Care Research, Radboud University Nijmegen Medical Centre, Netherlands
- Primary Care Group, School of Medicine, Swansea University
- Wales Office of Research and Development for Health and Social Care, Wales Assembly Government
- Department of Health, United Kingdom, Health in Partnership Initiative

We are grateful to the following journals for the permission to publish the studies included:

- Health Expectations
- Family Practice
- Quality and Safety in Health Care

First published in Wales in 2005 by Cardiff University.

Copyright:	Glyn Elwyn
Print:	Ponsen and Looijen BV, Wageningen
Cover illustration:	Sian Koppel
ISBN:	0-9550975-0-9

Contents

	page
1. Introduction	5
2. The use of OPTION - research and improvement	7
3. Measuring the involvement of patients in shared decision making: a systematic review of instruments	13
4. Option publications	37
4.1 Shared decision making: developing the OPTION scale for measuring patient involvement	37
4.2 The OPTION scale: measuring the extent that clinicians involve patients in deciding making tasks	51
4.3 Achieving involvement: process outcomes from a cluster randomized trial of shared decision making skill development and use of risk communication aids in general practice	61
5. OPTION Manual	77
6. OPTION Scale	91
6.1 English version	93
6.2 Dutch version	95
6.3 French version	97
6.4 German version	99
6.5 Spanish version	101
6.6 Italian version	105
7. OPTION Training Pack	107

1. Introduction

Richard Grol

There is an increasing awareness that patients can and should play an important role in deciding on their care, in defining optimal care, and in improving health care delivery. Popular concepts such as patient centred care, patient empowerment, shared decision making illustrate this emancipation of patients. This development has different backgrounds [1,2]. There is an ethical and legal pressure on professionals to share information with patients and to involve patients in decisions regarding their health and care provision. Many patients also wish to be more involved and their needs should, where possible, be met. There is increasing evidence that involvement of patients may result in better processes and outcomes of care. It further helps professionals to reflect on their patient's needs and preferences, it can result in the implementation of evidence based care, and it can also result in better self-management and more patient satisfaction. Patients have to be seen as co-producers of their own health, as their decisions and behaviour influence health outcomes to a large extent. So, it is logical to share responsibilities with patients. Finally, patient involvement may contribute to more choice and more competition between care providers, and as a result hopefully increased quality. It will be clear that there is a case for empowering patients. The question, however, is how to achieve such an ideal, since studies and observations in clinical practice show that patients are often hardly involved in decisions regarding their health care situation.

There are various ways to involve patients in health care and health care improvement [1,2]. A global distinction has been made between:

- influencing patient's expectations: preparing and educating patients before or during their decisions whether to seek care and what the best place for seeking care would be. This can be done by education on health problems through mass media and through the Internet, or by providing public reports on the quality of care by different care providers.
- involvement in care processes: involving patient's needs and preferences during care provision.
- For instance, needs of patients can be explored by specific instruments, education can be given tailored to the needs of patients, specific self-management tools can be introduced giving patients more responsibility for their own care, or patients can be involved in decisions by shared decision making processes, the use of specific decision aids or the introduction of risk communication tools and tables.
- patient feedback on care given: this includes the use of satisfaction surveys, but also complaint procedures, idea boxes and patient participation groups.
- involving patients in policy making: finally. patients may be involved at a higher level in health policies, for instance in priority setting and policy making at a national, regional or local level, or by involving patients in guideline or indicator setting.

Many tools for involving patients are thus available but the research evidence on the value and effectiveness of such approaches is still scarce; theory and anecdotal experiences dominate the debate [3]. Nevertheless we are dealing with an extremely interesting and challenging area, which may change the

landscape of health care in the next decades. Within this landscape of more autonomy and empowerment of patients the developments regarding “shared decision making” are among the most interesting. This has become a term used to describe the process whereby patients participate in decision-making processes about health care issues, typically in consultations [4]. There is wide agreement that information and decision making with patients should be “shared”, but exactly how to do that most effectively in what type of patients is yet unclear. It is not an easy concept, as de Haes [5] formulated it: “The ethical principle of autonomy is not necessarily beneficial and may conflict with the principle of beneficence”. It is also obvious that patients differ and may have different wishes with regard to involvement in care decisions. Older and younger patients, and patients with simple acute or serious chronic conditions will differ with respect to the manner and the extent of involvement. Research in this area, for instance on the appropriate way of using decision aids, is still inconclusive [6]. In short, a challenging field for research and development.

The OPTION-instrument, developed by Elwyn and colleagues [7,8], and presented in this book and manual fits into this developmental process. It is the first instrument specifically developed and validated for measuring the extent and quality of shared decision making by clinical professionals.

It has been developed in a step by step approach, starting with exploring and defining the concept and identifying other instruments related to this concept. This led to the conclusion that a specific instrument to measure this new concept was needed. After that a rigorous process of development, testing, revising, testing, etc. has been followed resulting in the current 12-item instrument. The OPTION-instrument was next tested in a variety of settings, running from primary care physicians to oncologists; many experts and clinicians around the world gave their opinions and experiences. It proved to be a feasible instrument for both research and practical purposes.

This book presents the process of development of OPTION as well as the final version translated in different languages. It is meant for research on shared decision making as well as a practical tool for education, evaluation and quality improvement on the communication between patients and their care providers.

References

1. Wensing M, Grol R. The patients' role in improving quality. In: Jones R (ed). Oxford textbook of Primary Medical Care. Oxford: Oxford University Press, 2004.
2. Elwyn G, Rhydderch M, Edwards A. Shared decision making. In: Silagy C, Britten N (eds). Oxford Textbook of Primary Medical Care. Oxford: Oxford University Press, 2003.
3. Grol R. Evidence based patient choice: Foreword. In: Edwards A, Elwyn G (eds). Evidence based patient choice. Oxford: Oxford University Press, 2001.
4. Elwyn G. Shared decision making in clinical practice (PhD Thesis). Nijmegen: Centre for Quality Care Research, Nijmegen University, 2001.
5. de Haes HC, Molenaar S. Patient participation and decision control: are patient autonomy and well-being associated? *Med Decis Making* 1997; 17: 353-4.
6. O'Connor AM, Stacey D, Rovner D, Holmes-Rovner M, Tetroe J, Llewellyn-Thomas H, Entwistle V, Rostom A, Fiset V, Barry M, Jones J. Decision aids for people facing health treatment or screening decisions (Cochrane Review) CD001431. 2001, *Cochrane Database Syst Rev*. p. 3.
7. Elwyn G, Edwards A, Wensing M, Hood K, Robling M, Atwell C, Grol R. Shared decision making: developing the OPTION scale for measuring patient involvement. *Qual Saf Health Care* 2003; 12: 93-9.
8. Elwyn G, Hutchings H, Edwards A, Rapport F, Wensing M, Cheung WY, Grol R. The OPTION scale: measuring the extent that clinicians involve patients in decision making tasks. *Health Expect* 2005; 8: 34-42.

2. The use of OPTION - research and improvement

Adrian Edwards, Michel Wensing

Introduction

OPTION is a new and validated instrument to assess the extent that shared decision making is happening in a health care consultation [1]. It is in the public domain and free to access. As with many other instruments, it can be used for different aims, including research, improvement, and accountability [2]. Each aim has specific purposes, audiences, and requirements for the measurements (see Table 1). In research, the validity of the measure is vitally important and the emphasis is usually on evaluating the effects of an intervention on a sample of clinicians. OPTION assesses whether there are demonstrable changes in the pattern of consultations.

Table 1. Characteristics of measurement for improvement, accountability, and research (adapted from Solberg et al, 1997)

	Improvement	Accountability	Research
Purposes	Motivation and focus for improvement; baseline and evaluation of changes	Basis for choice by customers and/or accreditation and recertification	Accumulation of knowledge and understanding
Audiences	Professionals, health care managers	Purchasers, patients	Scientists, public, users (clinicians)
Requirements for measurement	Few, feasible and responsive measures	Very few, highly precise and valid measures with complex procedures to account for confounding	Many, complex data collection, very precise and valid, large sample size

In improvement, measures are primarily used to evaluate and refine interventions so as to target their delivery and impact more accurately. To do this it is important that a measure is sensitive and responsive at the level of individual clinicians. For accountability, measures must essentially be able to compare between individual providers, and confounding from patient case mixes must be avoided. This chapter argues that the OPTION instrument has been developed primarily for research [1]. It can be used for improvement, but it has not been validated for accountability.

Research

The objective of research is to contribute to the accumulation of knowledge and understanding. In the case of the OPTION instrument this refers to knowledge about communication between health professionals and patients, a construct which the instrument has been shown to measure validly and consistently (with reliability). This knowledge can be assimilated into our understanding of the processes of professional-patient communication, gaining deeper insights into what appears to be effective as training interventions for professionals and what improves the outcomes that matter most to patients [3]. The development of the OPTION instrument was based on a theoretical analysis of the concept of shared decision making [4] and empirical work with experienced general practice educationalists [5]. These analyses led to the specification of specific competencies for the health professional regarding shared decision making (chapter 1). A

systematic literature review showed that very few validated instruments addressed aspects of shared decision making adequately, while none covered all the relevant competencies [6]. Therefore, the OPTION instrument was developed, and in further psychometric development and assessment, was shown to have high levels of validity and reliability [1] [7].

No instrument is 'proven valid' because validation is a never ending process. The OPTION instrument was mainly tested in patients and general practitioners from the UK, so use of the (translated) instrument for research in other populations is undergoing further validation. The validation procedures were based on a methodology that was mainly developed in the social sciences [8-10] and which has been used widely in health care [11]. Other perspectives on validation may also be valuable (e.g. from clinometric, econometrics). Nevertheless, validity and reliability of the instrument have received substantial psychometric assessment and from this we conclude that the instrument is ready for use in many research spheres [1,7].

Principally we conclude that the instrument has shown validity in the observational setting [1,7]. It also shows discriminant ability in the context of a randomised controlled trial [12]. It remains to be seen whether it shows responsiveness in the research setting. In practice this means that researchers can use OPTION to assess the extent to which clinicians implement shared decision making in practice from cross-sectional data, such as a series of audio- or video-tapes of consultations from one clinician or a group of clinicians. Researchers can also be confident to use the instrument to assess the effects of interventions (for example training interventions or the provision of decision aids) to groups of clinicians in the context of a randomised trial. OPTION is able to detect both clinically and statistically important differences after an intervention to groups of clinicians or their patients [12]. Responsiveness, however – the ability to show changes at individual clinician level, such as before to after an intervention – has not yet been demonstrated. This latter point will be returned to in considering the role of OPTION in educational and training settings.

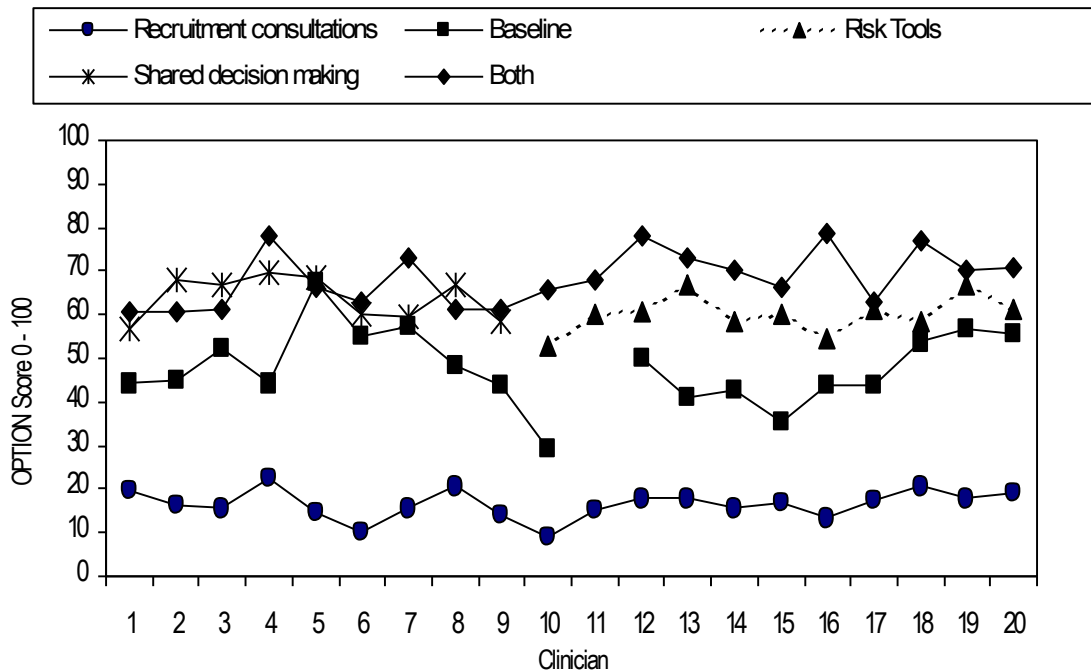
Improvement

Improvement refers to activities to improve professional practice or organisation of care, such as continued professional education, programmes to enhance guideline implementation, or formative evaluations. These projects use measures mainly for tailoring the interventions and for evaluating their impact. Compared to research, there is less emphasis on validation and more attention to the feasibility of conducting the measurements. Sensitivity to differences in performance and responsiveness to change are crucial features in evaluations of improvement initiatives. For instance, to be useful for improvement purposes, the OPTION instrument should be able to detect improvements (if occurring) after a training session to develop clinician skills in shared decision making. Improvement is often measured at the level of health care institutions rather than individual health professionals. Therefore reliability and responsiveness need to be assessed not only at the lowest level of patients or consultations (as they are in validation for research purposes), but also at higher aggregation levels.

As indicated above, OPTION's responsiveness to change has not yet been formally evaluated at the level of individual clinicians. However evidence from a trial that used a modified 'interrupted time series' design,

showed changes in scores both at individual clinician level and at aggregated level for clinicians [12]. This is illustrated by the figure below, which indicates how scores for 20 clinicians changed (increased) as they progressed through the study.

Figure 1. OPTION score changes during a trial with modified interrupted time series design [12].



In the figure responsiveness is indicated by scores increasing for most clinicians from pre study (circles) to study baseline (squares) to first active study phase (asterisks and triangles) to second active study phase (diamonds). These findings give preliminary indication of responsiveness. We hope to develop and test educational written feedback on the basis of OPTION scores, and to analyse the instrument's responsiveness formally in the near future.

Regarding feasibility, it is clear that using the OPTION instrument requires a substantial investment. Consultations have to be sampled and taped; observers have to be recruited to rate the taped consultations, trained (and paid); data have to be processed and analysed. Users should consider whether this is acceptable for their purposes. Currently, there is no 'simple' version of OPTION that is more feasible than the current version.

Accountability

In accountability, the performance of health care providers is compared and reported to professional or legal bodies (for accreditation and certification) or the general public (public reports). As the measurements are used for selection of providers, through consumer choice or summative assessments, it is crucial that they are highly valid, reliable and not confounded. It should be clear what the optimal level of shared decision making is and the measurement procedures should be resistant to manipulation by interested parties.

On the former of these, it is not clear what the optimal level of OPTION scores should be. This is because higher scores are not necessarily better, according to the context. The appropriate level of involvement, as reflected in the OPTION score, is determined by the situation (the nature of the health condition – acute or chronic health conditions – and setting) but also and importantly by the patient's desire for involvement in decision making. Certain situations may not be characterized by clinician equipoise, such as in considering the prescription of antibiotics for upper respiratory tract infections. These will require specific approaches to balance involvement and choice with the need for evidence-based practices motivated by public health goals [13]. The skill is for professionals to match their communication approaches – and level of involvement achieved – to the contextual aspects and patient preferences. Because OPTION explores this variable context, it is applicable for educational and reflective interventions such as described under 'Quality Improvement' above, but it has not been developed and is not applicable for accountability purposes.

While the validity has received much attention, it is less clear how the patient mix influences the scores per provider so that comparisons across professionals may be confounded. For instance, can a clinician with a younger patient population, perhaps with higher education attainment levels – and more likely to engender higher OPTION scores – be compared to a clinician with an older patient population, perhaps with lower educational attainment levels – and in whom the expectations for involvement and actual involvement levels may be lower. Many variables have complex influences, including context and patient case mix, and further research is needed before OPTION may be ready for use in accountability.

Conclusion

OPTION has shown good validity and reliability in the research setting. Raters need to be trained and standardised in their assessments, making it a labour intensive method, but it yields important findings about the process of consultations. It has the ability to provide valuable data for researchers and participants in studies alike, to show the effect of interventions and the way consultation behaviours have changed.

OPTION holds promise for use in the observational setting, such as in educational or other quality improvement activities. It should not be used for accountability purposes. Validation is a continuous process and in particular the non-English language versions of the instrument need testing. To facilitate this work we request that users of the instrument contribute data to a shared database for these further analyses.

References

1. Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Fleeting glimpses: measuring shared decision making in primary care using the OPTION instrument. *Qual Saf Health Care* 2003; 12: 93-9.
2. Solberg L, Mosser G, McDonald S. The three faces of performance measurement: improvement, accountability, and research. *J Qual Improve* 1997; 23: 135-47.
3. Edwards A, Elwyn G. How should 'effectiveness' of risk communication to aid patients' decisions be judged? A review of the literature. *Med Decis Making* 1999; 19: 428-34.
4. Elwyn GJ, Edwards A, Kinnersley P. Shared decision making in primary care: the neglected second half of the consultation. *Br J Gen Pract* 1999; 49: 477-82.
5. Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision-making and the concept of equipoise: defining the 'competences' of involving patients in health care choices. *Br J Gen Pract* 2000; 50: 892-9.

6. Elwyn G, Edwards A, Mowle S, Wensing M, Wilkinson C, Kinnersley P, et al. Measuring the involvement of patients in shared decision making: a systematic review of instruments. *Pat Educ Couns* 2001; 43: 5-22.
7. Elwyn G, Hutchings H, Edwards A, Rapport F, Wensing M, Cheung I, Grol R. The OPTION scale: measuring the extent that clinicians involve patients in decision making tasks. *Health Expect* 2005; 8: 34-42.
8. Streiner DL, Norman GR. Ch 3: Devising the items. In: *Health measurement scales: a practical guide to their development and use*. 2nd ed. Oxford: Oxford University Press, 1995: 15-27.
9. Jenkinson C, McGee H. Generic health status profiles. In: *Health status measurement: a brief but critical introduction*. 1st edn. Abingdon, Oxfordshire: Radcliffe Medical Press, 2000: 27-52.
10. McDowell I, Jenkinson C. Development standards for health measures. *Journal of Health Services Research and Policy* 1996, 1: 238-45.
11. Garratt A, Ruta D, Russell I, Macleod K, Hunt P, McKinlay A, et al. Developing a condition specific measure of health for patients with dyspepsia and ulcer related symptoms. *J Clin Epidemiol* 1996; 49: 565-71.
12. Elwyn G, Edwards A, Hood K, Robling M, Atwell C, Russell I, Wensing M, Grol R. Achieving involvement: process outcomes from a cluster randomised controlled trial of shared decision making skill development and use of risk communication aids in general practice. *Fam Pract* 2004; 21: 335-44.
13. Elwyn GJ, Gwyn R, Edwards A. Is 'shared decision-making' feasible in a consultation about patient expectations for antibiotics in a viral upper respiratory tract infection? *Health Expect* 1999; 2: 105-17.

3. Measuring the involvement of patients in shared decision making: a systematic review of instruments

Glyn Elwyn, Adrian Edwards, Steve Mowle, Michel Wensing, Clare Wilkinson, Paul Kinnersley, Richard Grol. *Patient Education and Counselling* 2001; 43(1): 5 - 22.

Permission for re-publication granted by Elsevier 10/12/04.

Abstract

We wanted to determine whether a research instruments exist which focus on measuring to what extent health professionals involve patients in treatment and management decisions. A systematic search and appraisal of the relevant literature was conducted by electronic searching techniques, snowball sampling and correspondence with field specialists. The instruments had to concentrate on assessing patient involvement in decision-making by observation techniques (either direct or using audio or videotaped data) and contain assessments of the core aspects of 'involvement', namely: evidence of patients being involved (explicitly or implicitly) in decision-making processes, a portrayal of options and a decision-making or deferring stage. Eight instruments met the inclusion criteria. But we did not find any instruments that had been specifically designed to measure the concept of 'involving patients' in decisions. The results reveal that that little attention has been given to a detailed assessment of the processes of patient involvement in decision-making. The existing instrumentation only includes these concepts as sub-units within broader assessments, and does not allow the construct of patient involvement to be measured accurately. Instruments developed to measure 'patient-centredness' are unable to provide enough focus on 'involvement' because of their attempt to cover so many dimensions. The concept of patient involvement (shared decision-making; informed collaborative choice) is emerging in the literature and requires an accurate method of assessment.

Introduction

Although there is increasing interest in the outcomes of involving patients in aspects of healthcare decisions, albeit with a recognition that a flexible approach is needed in practice [1], there is no agreed construct to describe 'involvement' [2]. 'Patient-centredness' is proving to be too ill-defined, [3] a method that in reality contains many constructs, and a recent comparison of instruments designed to measure it revealed the difficulty of achieving reliable tools [3,4]. Although involving patients is an important element of patient-centred practice, patient participation in decision-making has not been defined in sufficient detail to allow rigorous evaluation. Research into the roles patients prefer within decision-making processes has been mostly based on hypothetical scenarios [5,6] and reveals a spectrum of views. Hypothetical determinations may not equate with the views of patients who have experienced actual involvement in decision-making. There is evidence from studies on screening that the wishes of patients who are initially uninformed change after they have become aware of the harms and benefits of different treatment options [7]. This is likely to be especially true if the clinician is skilled at providing information and is sensitive to anxieties that may be generated by the potential responsibility of decision-making. It is also important to conceptualise patient involvement as a process that will inevitably vary from one consultation to another.

We were unaware of a method to measure 'involvement', and therefore undertook a systematic search of the literature with the aim of appraising the instruments identified.

Patient involvement can be viewed as occurring along a spectrum, from paternalism at one end to complete autonomy at the other [8]. 'Shared decision-making' involves both the patient and the clinician being explicit about their values and treatment preferences [9]. The approach involves arriving at an agreed decision, to which both parties have contributed their views. The stages and skills of 'shared decision-making' are being investigated by firstly using qualitative methods to investigate how practitioners and patients conceptualise 'involvement', and secondly by an empirical study which analyses consultations that aim to 'share decisions' [10-12].

Two assumptions underpin this review. Firstly, that involvement in decision-making is a negotiated event that occurs between a clinician and patient, either explicitly, or as is more common, implicitly.

The second assumption is that choices legitimately exist in most clinical situations, and that it is acceptable - vital according to those who place autonomy first amongst ethical principles- to portray options to patients, at least to some level of detail (excepting extremis, intellectual impairment, unconsciousness and psychiatric risk). Any attempt to measure involvement in decision-making should therefore consider to what degree (if any) a health professional portrays choices and invites patients to participate in the decisions, along with other processes that may be associated (such as an exploration of views, concerns, and fears). Involvement is not considered as a rhetorical gesture. Successful 'involvement' starts from the position of respecting a patient's right to autonomy and self-determination, even when a fully informed patient, aware of a contrary professional viewpoint, decides a divergent treatment or management plan. The ethical stance assumed here is one of optional autonomy rather than mandatory autonomy (where patient involvement in decision-making is a requirement) [13].

Decision-making in a clinical setting involves many factors, including prior experience, existing knowledge, trust and confidence in the clinician, personality traits, exposure and access to information, satisfaction with the consultation process, and the influence of family and others [13]. Despite this complex context, we consider that patient involvement in the decision-making process within the consultation is an important construct to measure accurately, for many reasons. It is necessary if we are to gauge how involvement contributes to determining adherence to treatment choices, and whether involvement *per se* contributes in other ways to potential health gain.

Objectives

Having first appraised the literature on how professionals should most appropriately involve patients in decision making processes [14], and completed a qualitative study on the 'competences' required [11], we undertook a systematic search for instruments that focused on an evaluation of the extent professionals involve patients in decision-making (and the quality) as observed by a third party. This is not to dismiss the literature that has focused on perceived involvement (as viewed by clinician and patient) and which has an important predictive effect on patient outcome [15,16]. We consciously excluded such tools because of our focus on actual behaviour within the consultation. This is justified by an argument that each perspective,

(i.e. observed participation and the views of physicians and the patients), needs robust measures so that valid comparisons can be made, and conclusions drawn about the most effective 'participatory' behaviours. Exploratory qualitative work provided a framework for our conceptualisation of patient involvement in decision-making - see Box 1 [11,17]. We wanted to establish whether existing instruments were capable of providing valid and reliable measurements of 'patient involvement' to a level that is satisfactory for quantitative research purposes.

Box 1. Stages and competencies of involving patients in healthcare decisions [11,17]

1	Implicit or explicit involvement of patients in decision-making process
2	Explore patient ideas, fears and expectations regarding possible treatments
3	Portrayal of equipoise and options
4	Identify preferred format and provide tailor-made information
5	Checking process: understanding of information and reactions (e.g. ideas, fears, and expectations of possible options)
6	Acceptance of process and decision-making role preference
7	Make, discuss or defer decisions
8	Arrange follow-up

Methods

The methods of systematic reviewing have been developed primarily to summarise research that investigates the effectiveness of interventions [18]. This review applies the concept of a systematic and explicit method of assessment to the area of instrumentation. There are agreed methods for both developing and confirming the validity and reliability of health measurement instruments, which will be used as the basis for assessing the quality of instruments in this review [19].

Search Strategy

We sought to identify studies that reported the development or use of instruments that aimed to evaluate clinical interactions. Identified instruments were then assessed to see if they had the ability to measure whether, and to what extent, clinicians were, in a broad sense, 'involving' patients in health care decisions.

We searched the following databases: Medline (1986-98) CinAHL (1986-1999) Psychlit (1986-1998), Embase (1986-1998), ASSIA (1986-1998). The search strategy for Medline required articles to match against (i) one or more MeSH or textword terms relating to *decision-making* or *patient involvement*, and (ii) MeSH or textword terms describing *methods of assessing the consultation*. The MeSH terms were correspondingly modified for use in different databases. Full details of the search strategy are available.

This subject area is not well indexed. We therefore used a strategy designed to achieve high recall/sensitivity rather than precision/specificity. A large number of titles and abstracts were generated from these searches. Two authors independently assessed this output and retrieved relevant articles for further assessment. Forward searches for citations of papers reporting the development and first use of relevant instruments were conducted on the ISI Science and Social Science Citation Indices. We checked the reference list of identified papers and corresponded with 60 experts in this research area, determined by the authors as experts in the field of health communication research (list available).

Criteria for considering studies

Two criteria had to be met for studies to be included in this review, and were based on the widely agreed premise that patients can only contribute to the decision-making process if choices are explored within the communication process. The measures had to:

Involve assessments made by direct or indirect observation of the consultation (i.e. by audio or videotape recording). Include assessment of the core aspects of 'involving' patients 'in the process of decision-making', and therefore contain items that covered at least one of the following stages (Box 1):

- a) involvement of patients in decision-making processes
- b) a portrayal of options
- c) a decision-making or decision-deferring stage.

The inclusion criteria were applied in two stages. The first stage involved the evaluation of all identified instruments that measure patient-clinician interaction (actual or simulated practice). It could then be ascertained whether aspects of the second criterion were met. Decisions regarding inclusion/exclusion were checked by a second author (AE). Instruments that met both criteria were appraised in depth against an agreed checklist by two assessors (GE and AE), and by correspondence with the original authors when it was necessary to obtain further details.

Data extraction

Data extraction was carried out by GE and checked by AE. Authors were contacted with requests for copies of their instruments if details or questions were missing from published reports. Data were extracted in order to examine two broad aspects of the instruments. Firstly, descriptive features for each instrument were collected (Table 1): the stated aim, the theoretical or conceptual basis (or the theoretical or conceptual framework of the paper, methods of assessment, reports of instrument development and/or first use); the scenario(s) or aspects of the concept to be considered, the setting in which it was first used; and the apparent scope of its use. Included in these descriptive categorisations is information about the means of data collection and the existence of a guidance manual. Instruments that met the inclusion criteria were compared against a conceptual framework which describes the competencies which professionals consider to be key features of patient-participation in decision-making (see Box 1).

Secondly, there are the methodological issues that determine the quality of instruments and these are covered in Table 2. They concern the development of the scale (and its items) and to what extent validity and reliability have been assessed (see footnotes to Table 2) [19].

Results

The searching strategy identified a total of 4929 abstracts from the following databases: combined listing from Medline, Psychlit and Embase, 2460; CinAHL, 2395; ASSIA, 74. After dual and independent assessments, a total of 107 articles were retrieved for detailed appraisal. Information and articles were received from 29 of the 60 authors contacted (see acknowledgements); 52 consultation assessment instruments that met only the first inclusion criterion of this review are listed in Table 3.

Reporting of results

Eight instruments were found to include items that measured patient involvement in decision-making as defined by the inclusion criteria. Descriptive details and an analysis of their development, validity and reliability data are provided in Tables 1 and 2 respectively. It will be noted that there are four scales that fulfilled the 'measure involvement' criteria (see table 3) that are not appraised. Pendleton's Consultation Rating Scale [20], the Royal Australian College of General Practitioners [21], the Royal College of General Practitioners (UK) examination criteria [22] either had items which only mentioned the concept of involvement or in the case of the examination criteria were checklists that have not been subjected to any validation exercises in a research setting. The Leeds Rating Scale [23] was not included as the concept of involvement was mentioned only as a broad approach within the interaction. Using these strict criteria we would have also excluded the Calgary-Cambridge Observation Guides (formative assessment tools) but we felt the detailed items included deserved comparison with other existing instruments. Both tables are arranged alphabetically by instrument name.

Synthesis of results

An appraisal of these instruments reveals that there has been an initial interest in the 'second half' of the consultation [14] but that very little attention has been given to a detailed assessment of the processes of participation [2]. It is clear that involving patients in decision-making, either implicitly or explicitly, providing treatment options, information about choices and then engaging in a decision-making stage are 'constructs' that have not been considered to any significant depth in clinical interaction analysis. To date, the existing instrumentation only includes these concepts as sub-units within broader comprehensive assessments.

Constructs which are apparent in the literature (*italics* indicate phrases or items within original instruments).

Six constructs (Table 1) underpin the instruments that meet the inclusion criteria. Four of these are more focused in nature, and are concerned respectively with *problem-solving* [24], *exploration of patient concerns* [25], assessment of patient *reliance* [26] and *informed decision-making* [27]. The other constructs have broader scopes: two aim to assess communication skills in a global sense [28,29], and two aim to assess components of *patient-centredness* [4,30].

Problem-Solving Instrument (Pridham, 1980) [24]

Pridham's work considered *problem-solving* and the *construction of self-management plans* based on the analysis of five consultations [24]. The method however was to assign codes to each 10-second intervals and calculate counts of categorisations, namely 'scanning, formulating, appraising, willingness to solve, planning, implementing (all applied to problems within the consultation. The instrument was not based on worded items. No further work has been published on this instrument.

Levels of Physician Involvement (Marvel, 1994) [25]

Marvel's [25] adaptation of Doherty's levels of physician involvement with families [31] is included but the instrument does not fully address patient involvement in decision-making. The prime aim of the measure is to evaluate the range of skills that physicians use to address the psychosocial concerns of patients (and their families). For example, at the second level (of 5 ascending stages) the rater is asked to consider:

Level 2: Individual focus

Orientating question: What information should be exchanged with the patient to make the correct diagnosis and to design and agree upon a treatment plan?

Although options and decision-making are not specified, the design and agreement could be understood as a measure of patient involvement. The primary aim is not to consider patient involvement in decision-making, as items at other stages indicate.

Decision-Making Checklist (Makoul, 1992) [26]

The aim of this instrument is to assess how the consultation influences patients' perceptions of control and correlates with observed behaviours. Makoul's work was undertaken as part of a research thesis using a large sample of general practice consultations in the United Kingdom. The Decision-Making Checklist [26] (binary responses) concentrate on information provision. Three items (from a total of 24) focus on decision-making:

Item 17 Did the MD (doctor) ever seem to give the PT (patient) any responsibility for deciding how to deal with the health problem?

Item 18 Did the MD ever explain possible treatments to the PT? (over and above naming the treatment)

Item 19 Did the MD ever involve the patient in choosing a treatment for the health problem? (e.g. "which alternative")

Informed decision-making (Braddock, 1997) [27]

The aim of this instrument is to characterise the consent and decision-making process in consultations. Braddock's coding of consultations using an informed *decision-making* model [27] is an approach which, although it requires validation, has the benefit of having a firm theoretical stance and mirrors sequences that professionals suggest are needed in order to involve patients in decision-making [17].

MAAS-Global (van Thiel, 1991) [29]

The aim of this instrument is to determine the degree of proficiency of medical interviewing skills. Van Thiel's adaptation of earlier MAAS scoring lists has resulted in the current MAAS-Global instrument [32]. This scoring list is now designed for use in general practice consultations. The sixth phase (named 'management') contains the following four items:

- *shared decision-making, discussing alternatives, risks and benefits*
- *discussing feasibility and adherence*
- *determining who will do what and where*
- *asking for patient response.*

Scores are given to each phase (scored '0' for absent, 6 for excellent): The manual (in translation) indicates that the criteria for 'excellent' require that 'the physician discusses the treatment plan and provides the

patient [sic] the opportunity to share his or her views, that the advantages and disadvantages of the treatments are described, and that depending on the condition, it may be necessary to discuss alternatives. The criteria continue by addressing the need to be sensitive to patient preferences and to make adequate review arrangements.

Calgary-Cambridge Observation Guides (Kurtz, 1996) [28]

The aim of the Calgary-Cambridge Observation Guides is to act as a basis for formative assessment. The guides provide the most extensive list of items but they are not designed to be research instruments [28]. Nevertheless, the second guide which covers the ‘explanation and planning’ stage of consultations provides converging confirmation of the ‘construct’ of patient involvement in decision-making, as depicted by the items within the section on negotiating a ‘mutual plan of action’:

- 27 *Discusses options, e.g. no action, investigation, medication or surgery, non-drug treatments*
- 28 *Provides information on action or treatment offered*
 - a) *name*
 - b) *steps involved, how it works*
 - c) *benefits and advantages*
 - d) *possible side-effects*
- 29 *Elicits patient’s understanding, reactions, and concerns about plans and treatments, including acceptability*
- 30 *Obtains patient’s views of need for action, perceived benefits, barriers and motivation; accepts and advocates alternative viewpoint as necessary*
- 31 *Takes patient’s lifestyle, beliefs, cultural background and abilities into consideration*
- 32 *Encourages patient to be involved in implementing plans, to take responsibility and to be self-reliant*
- 33 *Asks about patient support systems, discusses other support available*

Component 3 of the patient-centredness (Stewart, 1995) [33]

Component 3 of the *patient-centredness* scoring instrument (which covers the concept of ‘finding common ground’) provides spaces to list which *problems the doctor has clearly defined* and whether *opportunities for questions* were provided.

Raters are also asked to assess whether patients have been involved in *mutual discussions* and in agreeing *goals for treatment and management*. Binary (yes/no) responses are possible. Total scores provide an overall index of *common ground*. Although the instrument can assess whether ‘discussion’ occurs, it cannot distinguish whether choice is provided and to what extent patients are involved in the decision-making process.

Euro Communication (Mead, 1999) [4]

This instrument was developed specifically for use in a current study and measures a doctor’s patient-centred behaviour across five dimensions. Preliminary validation work comparing it to two other instruments reveals that three of the dimensions cover aspects of patient involvement in decision-making: problem definition, decision-making, patient ambivalence.

Validity and Reliability Testing of Instruments

The development of instruments to evaluate professional communication skills has taken place mostly in a generalist clinical setting; the eight instruments in Table 1 exemplify that trend. The quality of the instruments that met the reviews' inclusion criteria, compared to the rigorous psychometric standards of validity and reliability testing (item development based on qualitative techniques, followed by quantitative refinement and selection, and determination of sensitivity and responsiveness) is generally low (Table 2). Evolution of the MAAS instrument for instance has moved from the assessment of basic communication skills of medical students to the formative development of doctors training in general practice. It is a global index of ability across many different facets of communication skills. The most cited instrument aims to assess 'patient-centredness' but this is increasingly recognised to be a multi-dimensional construct. Braddock's tool was developed from ethical principles [27], and Makoul's instrument based on the construct of 'reliance' [26] but the path taken from theoretical concept to item formulation, refinement and selection is not described. Many of the identified instruments have not been validated and the results of concurrent validity of Stewart's instrument when conducted outside the original development setting point to the need for further refinements [4]. Braddock and Marvel report inter-rater agreements without adjusting for agreement by chance. The use of generalisability theory [19] as a means of providing reliability coefficients based the number of raters and the number of consultations is limited to studies conducted on the MAAS-Global instrument.

Clustering of existing instrument items

It is possible however to match the items identified within these eight instruments against a suggested chronological staging of 'patient involvement', which we have based on the competencies identified in Box 1 and on existing literature in the field [1,9,11,34] (Box 2). This matching process illustrates how the identified instruments vary in the extent to which they contain items that cover the broad sequences described. This comparison of items has the potential, if combined with further inductive work, to guide the construction of a patient involvement instrument.

Box 2. Clustering of existing items into identifiable ‘stages’ of patient involvement in decision-making (arranged alphabetically by instrument)

Stages of patient involvement	Items found in existing instruments
Involvement in decision-making process (i.e. agreeing the problem and the need for a decision)	<ul style="list-style-type: none"> • <i>involving patients in problem definition</i> (Euro Communication) [4] • <i>the clinical issue and nature of decision</i> (Informed decision-making) [27] • <i>problems the doctor has clearly defined</i> (Patient-centredness) [30, 35]
Exploring ideas, fears and expectations	<ul style="list-style-type: none"> • <i>takes patient’s lifestyle, beliefs, cultural background and abilities into consideration</i> (Calgary-Cambridge Observation Guides) [28] • <i>exploring issues of patient ambivalence</i> (Euro Communication) [4] • <i>discussion of uncertainties</i> (Informed decision-making) [27] • <i>discussing feasibility and adherence</i> (MAAS-Global) [32] • <i>opportunities for questions</i> (Patient-centredness) [33] • <i>mutual discussions</i> (Patient-centredness) [30, 35] • <i>goals for treatment and management</i> (Patient-centredness) [30, 35]
Option portrayal	<ul style="list-style-type: none"> • <i>discusses options, e.g. no action, investigation, medication or surgery, non-drug treatments</i> (Calgary-Cambridge Observation Guides) [28] • <i>discussion of alternatives</i> (Informed decision-making) [27] • <i>What information should be exchanged with the patient to ... design and agree a treatment plan?</i> (Levels of involvement) [25] • <i>shared decision-making, discussing alternatives, risks and benefits</i> (MAAS-Global) [32]
Provide information (risk communication)	<ul style="list-style-type: none"> • <i>Provides information on action or treatment offered</i> (Calgary-Cambridge Observation Guides) [28] • <i>Did the MD ever explain possible treatments to the PT? (over and above naming the treatment)</i> (Communication and decision-making) [26] • <i>Discussion of pros and cons of alternatives</i> (Informed decision-making) [27] • <i>What information should be exchanged with the patient to ... design and agree a treatment plan?</i> (Levels of involvement) [25] • <i>shared decision-making, discussing alternatives, risks and benefits</i> (MAAS-Global) [32]
Checking process: understanding of information and reactions	<ul style="list-style-type: none"> • <i>elicits patient’s understanding, reactions, and concerns about plans and treatments, including acceptability</i> (Calgary-Cambridge Observation Guides) [28] • <i>Obtains patient’s views of need for action, perceived benefits, barriers and motivation; accepts and advocates alternative viewpoint as necessary</i> (Calgary-Cambridge Observation Guides) [28] • <i>assessment of patient understanding</i> (Informed decision-making) [27]
Acceptance of process and decision-making role preference/making decisions	<ul style="list-style-type: none"> • <i>encourages patient to be involved in implementing plans, to take responsibility and to be self-reliant. Asks about patient support systems. Discusses other support available</i> (Calgary-Cambridge Observation Guides) [28] • <i>did the MD (doctor) ever seem to give the PT (patient) any responsibility for deciding how to deal with the health problem?</i> (Communication and decision-making) [26] • <i>did the MD ever involve the patient in choosing a treatment for the health problem? (e.g. “which alternative”)</i> (Communication and decision-making) [26] • <i>involving patient in decision-making regarding management</i> (Euro Communication) [4] • <i>asking patient to express a preference</i> (Informed decision-making) [27] • <i>determining who will do what and where</i> (MAAS-Global) [32]
Opportunity to review decision-making	<ul style="list-style-type: none"> • <i>asking for patient response</i> (MAAS-Global) [32]

Discussion

Principal findings

Existing instrumentation in the field of professional-patient interaction research and evaluation does not enable the construct of patient involvement to be measured comprehensively. Although an important finding, it is not a surprising one. None of the instruments we found (and included) were designed specifically to measure 'patient involvement'. The study of interactive communication within clinical consultations was pioneered in the 1960s, and many instruments have been developed since to evaluate the clinician-patient interaction. Nevertheless, the majority of existing observational tools have been situated within a paternalistic paradigm of interpersonal communication. The instruments that met our criteria are generic tools, capable of considering all types of clinical decision-making scenarios but they vary extensively in the detail to which they measure 'involvement'. The MAAS-Global and Decision-Making [26] checklists for example do not cover the issue of *mutual plan of action* in as much detail as the Calgary-Cambridge Observation Guides. Those tools that had some items on 'involvement' lacked a clear conceptual and theoretical underpinning.

Instruments developed to measure 'patient-centredness' are unable to provide enough focus on 'involvement', and their quality has been questioned, mainly because of their attempt to cover so many dimensions within consultations [3,4,36]. The concept of patient involvement (shared decision-making [9,37]; informed collaborative choice [38]) is emerging in the literature and demands an accurate method of assessment.

It is important to recognize that how a construct is defined and understood will determine efforts to design measurements. The principles of 'shared decision making' (where professional and patient values are integrated to arrive at a final decision differ from those of the 'informed choice' model, where patients are regarded as fully autonomous, and expected to make their own decisions [9]. It is clear for that an active patient involvement in the decision making process was not part of the patient-centred consultation method, at least within early conceptualisations. Measurements will either need to state which model they are assessing (state underlying assumptions) or be capable of taking neutral observational stances, whilst having items that determine empirically which model the clinician is following. It also seems clear that some of the stages and competencies (see box 1) will be easier to operationalise into items than others, and this is exemplified by the frequency of items within the instruments that were included in our detailed appraisal (see Table 1). Assessing 'implicit' involvement may be impossible to observe reliably, as would any aspects that depend on the assessment of patient perceptions. To attain reliability, instruments in this area will have to narrow their focus on behaviours that can be directly observed (e.g. providing options, data about harms and benefits, checking understanding and so forth) and to attain validity, be based on competencies that are at least feasible in actual practice [11,39].

Strengths and weaknesses

This systematic review of instruments in the field of professional-patient interaction examines for the first time the extent of psychometric development and testing which has underpinned existing instrumentation in this area. It provides an assessment of the degree to which validity and reliability issues have been

considered when measurement tools are developed and provides a comparison of items within existing instruments. Although we made extensive use of the technique of snowball sampling and contacted over 60 cited authors we found that publications in this area are spread over many journals which are either not, or poorly, indexed and we would welcome information about any instruments that have escaped our attention. Although we were able to obtain translations of work done in the Netherlands and contacted colleagues in Germany, we may have omitted other work not published in English. Some extensively used instruments were excluded (e.g. the Roter Interaction Analysis System [40]). Although this instrument included items that code patient question-asking and information provision, its dependence on the summing of coding categories precluded it as a tool capable of identifying an involvement process [41].

Implications for researchers and policymakers

Existing instruments have not been specifically developed to measure 'patient involvement' in clinical interactions: the tools were developed for different purposes. Those that have items relevant to this construct are not well developed or validated. It remains to be decided whether the instruments described in this review should guide the design of a measure of patient involvement. Valid instrumentation should be derived from a well-defined construct with item selection based on qualitative inquiry, and then rigorously developed according to psychometric principles.

To what extent the development of such an instrument should be guided by patient (consumer) or professional perspectives is a moot issue? The communication steps in Box 1 are derived from qualitative work on both patient and professional viewpoints and provide a firm basis for conceptualising how clinicians should approach this task, and could guide instrument development. As no 'gold standard' exists, construct validity should be determined by means of hypothesis testing (using extreme groups, convergent and discriminant validity testing methods) [19]. The list of items evident in these eight instruments (Box 2) provides at least a starting point for discussion with professionals and consumers [11].

Although there is work to suggest that patient perceptions of involvement are an important component of any 'effect' that increasing the participatory nature of the consultation might have, there is a parallel need to be able to 'identify' the communication skills that result in differing perceptions. Correlating empirical practice against high perceptions of 'involvement' may well be one method of identifying 'good practice'. But there is also a need to determine the construct of 'involvement', determine the contributory competencies, and develop an acceptable instrument to determine the levels of proficiency attained. This study allows us to move closer to that possibility. Proposals to involve existing research groups who have an interest in this area in the development of an 'involvement' instrument would strengthen the work and avoid the duplication of under-used evaluation methods in the field of health communication research [42].

Table 1. Descriptive data for instruments that consider involvement in decision-making

Instrument, first author, Index Publication, Country	Conceptual or construct framework	Method of assessment	Aspects of decision-making considered (Numbers correspond to skills and stages described in Box 1)	Types of decisions considered Context of first use	Manual Availability Citation total (SCI/SSCI) of Index Publication
Calgary-Cambridge Observation Guides (Kurtz, 1996) [28] Canada	Communication skills derived by expert consensus.	Checklist of defined behaviours and stages.	Within the explanation and planning phase a section exists, which is called 'shared decision-making' which lists key stages of offering choices, checking views and negotiating acceptable management plan. (1,2,3,4,5,6,7,8).	All types. Developed within the undergraduate communication course, University of Calgary.	Published Observation Guides available. Citations: 2
Communication & Decision-Making Checklist (Makoul, 1992) [26] United States	Investigation of 'reliance': the degree to which patients rely on clinicians for decision-making.	Checklist of defined behaviours and stages (binary responses).	Is information provided about medication and involvement in decision-making? Discussion about medication. Equality in consultation. Number of options mentioned. Involvement in decision-making. (1,2,3,7).	All types. Consultations in UK-based general practice.	No manual available. Citations: 6
Elements of informed decision making (Braddock, 1997) [27] United States	Informed consent.	Binary scoring of defined elements of informed decision-making.	Nature of decision, alternatives, pros/cons, uncertainties, patient understanding and preferences. (1,2,3,4,5,6,7)	All types. Family practice in United States.	No manual available. Citations: 11
Euro-communication Scale (Mead, 1999) [4] United Kingdom	Patient-centred consulting style.	Rating scale applied to 5 defined behavioural dimensions.	Patient involvement in problem definition and management decision-making, self-efficacy and clinician responsiveness. (1, 2,6,)	All types. Consultations in UK-based general practice.	No manual available. Citations: 0
Levels of Physician Involvement (Marvel, 1994) [25] United States	Exploration and management of patient and family concerns.	Levels of involvement coded and quantified.	Level 2 describes the competency of collaborative information exchange, i.e. 'what information should be exchanged to diagnose, design and agree a treatment plan'. (1,3)	All types. Family practice in United States.	No manual available. Citations: 11

Measuring patient involvement

Instrument, first author, Index Publication, Country	Conceptual or construct framework	Method of assessment	Aspects of decision-making considered (Numbers correspond to skills and stages described in Box 1)	Types of decisions considered Context of first use	Manual Availability Citation total (SCI/SSCI) of Index Publication
MAAS-Global, (van Thiel, 1991) [29] Netherlands	Communication skills derived and defined by expert consensus.	Rating scales applied to defined behaviours and stages.	Discussing alternatives; discussing risks/benefits; checking processes. (1,2,3,5,6,8).	All types. Communication skills of medical undergraduates.	Yes. Dutch manual available. Citations: 11
Patient-centredness: Component 3: Finding common ground. (Stewart, 1995) [30], Canada	Patient-centred consulting style.	Checklist of defined behaviours and stages, with binary scoring. Total score expressed.	'Mutual' discussion about goals for treatment and management. (1,2,3,5).	All types. Family medicine in Canada.	Instrument and guidance available from authors. Citations: 0
Problem-Solving Method (Pridham, 1980) [24] United States	Interpersonal problem-solving skills.	Intervals (10-second duration) are coded according to itemised process sheet.	Problem-solving; guiding further action; self-management plan development; evaluation of problem-solving process. (1,2,3).	All types. Family practice in United States.	No manual available. Citations: 0

Citation data obtained from Science Citation Index (SCI); Social Science Citations Index (SSCI), BIDS ISI Service, 16/9/99.

Table 2. The development, validity and reliability testing of instruments that met inclusion criteria †

Instrument (First author)	How was the instrument developed?	Reported Validity Assessments	Reported Reliability Assessments
Calgary-Cambridge Observation Guides (Kurtz, 1996) [28]	The guides were developed and refined over 20 years within the undergraduate communication course of the University of Calgary, and have been adapted by reference to the cumulative literature on doctor-patient communication [43].	Content validity confirmed by authors. Other validity aspects not measured systematically as the guides are formative, not 'research' measures.	No published data.
Communication & Decision-Making Checklist (Makoul, 1992) [26]	No details are provided about how items were developed or selected for inclusion in the checklist [26].	Content validity confirmed by authors. The thesis and publications to date do not provide further data apart from hypotheses testing within the thesis which support the validity of measuring 'reliance' (on self or on physician).	Inter-rater reliability coefficient (K) = 0.97 [26]
Elements of informed decision making (Braddock, 1997) [27]	The authors 'synthesised' the ethical models of informed consent as presented in the bioethical literature and devised a 6-item list : 'elements of informed decision-making'.	Content validity confirmed by authors. The publication does not provide data regarding further validation or construct hypothesis testing.	Inter-rater 'agreement' 77%.
Euro-communication Scale (Mead, 1999) [4]	The scale was devised specifically for use in the Euro-communication study. No data available regarding its development but the authors of the index publication admit that it has been limited.	Poor concurrent validity with two other measures of patient-centredness. Significant positive associations with: GP acquaintance with patient, GP age, consultation length, proportion of eye contact and importance placed on psychological factors by GP [4].	Internal consistency (Cronbach's alpha) = 0.90. Inter-rater agreement: intraclass correlation coefficient = 0.34.
Levels of Physician Involvement (LPI) (Marvel, 1993) [25]	The LPI was developed from Doherty's 'levels of physician involvement with families' [31], but no details are provided regarding the adaptation of the group measure to an instrument designed for a dyad interaction.	Content validity confirmed by authors. The publications to date do not provide data regarding further validation or construct hypothesis testing.	Inter-rater 'agreement' 79%.
MAAS-Global, (van Thiel, 1992) [32]	The instrument has been extensively developed from an original checklist of history-taking and advice giving in a medical student training context (1987). Development took place within a series of iterative assessments of communication skills. The revised version (MAAS-R, 1989) was modified by van Thiel in 1992 and is now known as MAAS-Global.	Content validity confirmed by authors. The instrument is used throughout the Netherlands for communication skill assessments in general practice (1999). van Thiel confirms that publications to date do not provide data regarding further validation or construct hypothesis testing (personal communication, 1999).	Use of generalizability coefficients. Inter-rater reliability MAAS-Global (intraclass correlations) = 0.78

Measuring patient involvement

Instrument (First author)	How was the instrument developed?	Reported Validity Assessments	Reported Reliability Assessments
Patient-centredness: Component 3: Finding common ground (Stewart, 1995) [30].	The existing measurements of patient-centredness were developed over the last 20 years by a research group in Ontario, Canada, and based on the conceptualisations of Levenstein, Henbest and McWhinney. Development of the instruments took place within the studies into patient-centredness conducted mostly within the context of family medicine in Canada by the Ontario group.	Good concurrent validity with 'global scores of experienced communication researchers' ($r=0.85$). Associations found with patients' subjective perceptions of 'finding common ground' but not with perceptions that the 'doctor explored the illness experience'. Construct validity: not systematically tested [30].	Inter-rater reliability coefficient = 0.83; intra-rater: $r= 0.73$ [30].
<i>Problem-Solving Observation Method (Pridham, 1980) [24]</i>	<i>This was developed by the principal investigator to assign codes to each 10-second interval which differentiate whether participants in the clinical interaction were 'organising', 'formulating', 'orientating', 'guiding', and 'planning' within an overarching construct of 'problem-solving'. No evidence exists that this instrument has been used subsequently.</i>	<i>Content validity confirmed by authors. The publication does not provide any further validity data or construct hypothesis testing.</i>	<i>Only 5 consultations analysed and inappropriate statistical analysis performed.</i>

† Footnote to Table 2 **Assessing issues of validity and reliability** (it is important to emphasise that we are not assessing the ability of the instruments to measure 'involvement in decision making' (they were not developed to undertake that task), but reporting published data.)

Validity: Face validity indicates whether an instrument 'appears' to either the users or designers to be assessing the correct qualities. It is essentially a subjective judgement. Content validity is similarly a judgement by one or more 'experts' as to whether the instrument samples the relevant or important 'content' or 'domains' within the concept to be measured. An explicit statement by an expert panel should be a minimum requirement for any instrument. However, to ensure that the instrument is measuring what is intended, methods that go beyond peer judgements are usually required. For this study, the instrument should reflect the understanding given to patient involvement in decision-making: agreement that a defined problem needing a management decision exists; that valid options are available; and that both information and opinions contribute to the process of decision-making.

If similar instruments already exist it is possible to consider criterion validity and construct validity. *Criterion validity* is usually defined as the correlation of a scale with some other measure of the trait of disorder under study (ideally a 'gold standard' in the field). *Construct validity* refers to the ability of the instrument to measure the 'hypothetical construct' which is at the heart of what is being measured. (For example, in this review an ideal instrument should be capable of measuring the level of patient involvement in decision-making achieved within the consultation.) If however, no other similar measure exists it is not possible to compare against another scale. For example, it emerges that a 'gold standard' for measuring patient involvement in decision-making is not available. Construct validity is then determined by designing experiments which explore the ability of the instrument to 'measure' the construct in question. This is often done by applying the scale to different populations, which are known to have differing amounts of the property to be assessed. By conducting a series of converging studies the construct validity of the new instrument can be determined. High correlation with aspects of 'patient-centredness', global measures of communication skills or patient perceptions of 'having their views' considered could be postulated, and

investigated for example. An additional method would be to measure 'patient involvement' within a sample of consultations and to test hypotheses within that population e.g. that elderly patients, or patients from low educational or social class are involved to lesser extents than other groupings.

Reliability: Internal consistency: this assumes that the instrument is assessing one dimension or concept and that scores in individual items would be correlated with scores in all other items. These correlations are usually calculated by comparing items (Cronbach's alpha, Kuder-Richardson, split halves). Instruments which assess 'the consultation' rarely focus on one concept and it is not usually possible to assess internal consistency (although different elements of 'good' consulting could be expected to correlate). Stability: this is an assessment of the ability of the instrument to produce similar results when used by different observers (inter-rater reliability) or by the same observer on different occasions (intra-rater reliability). Does the instrument produce the same results if used on the same sample on two separate occasions (test-retest reliability)? The production of reliability coefficients by using generalisability theory is advocated where measurements are undertaken in complex interactions by multiple raters [44].

Table 3. Clinical interaction measures: a list of instruments identified and considered

Instrument name, date, first author	Description of instrument	Data collection	Addresses 'involvement'	Portrays options	Considers decision-making
Type A: Instruments that measure concepts, stages or defined tasks within consultations					
Arizona Clinical Interview Rating Scale, (Stillman, 1977) [45, 46]	Assesses 16 interviewing skills using 5-point scale, under 6 headings. Organisation, Timeline, Transitional Statements, Questioning Skills, Rapport and Documentation of Data.	Direct or recorded data analysis	(-)	(-)	(-)
Assessment of videotapes (Cox, 1993) [47]	This 37-item rating scale aims to distinguish between 'good' and 'bad' consultations.	Videotape analysis	(-)	(-)	(-)
Barrett-Lennard Relationship Inventory [48]	A 64-item inventory divided across four variables: empathy, level of regard, unconditionality of regard and congruence.	Direct or recorded data analysis	(-)	(-)	(-)
Bensing's General Consultation Judgement (Bensing, 1991) [49]	Measures the attention given by a practitioner to the 'psychosocial care' provided within the consultation. A general judgement is made (on a scale of 1 to 10) against a set of 5 items that describe psychosocial care qualities.	Videotape analysis	(-)	(-)	(-)
Brown University Interpersonal Skill Evaluation (Burchard, 1990) [50]	Assesses the interpersonal skills of surgeons using a 40-item list divided into four sections: 'establishing rapport', 'demonstrating skills and procedures', 'testing for feedback' and 'providing appropriate closing'.	Direct or recorded data analysis.	(-)	(-)	(-)
Calgary-Cambridge Observation Guides (Kurtz, 1996) [28]	The aim of the guide is to act as a basis for formative assessment. Communication skills derived by expert consensus. Checklist of defined behaviours and stages.	<i>Direct or recorded data analysis.</i>	(+)	(+)	(+)
Category Observation Scheme (Mazucca, 1983) [51]	Eleven behaviours are categorised. Although there is an explicit category named 'sharing medical data', the focus is on data transfer and patient understanding.	Videotape analysis	(-)	(-)	(-)

Measuring patient involvement

Instrument name, date, first author	Description of instrument	Data collection	Addresses 'involvement'	Portrays options	Considers decision-making
<i>Communication & Decision-Making Checklist</i> (Makoul, 1992) [26]	<i>This checklist has items that cover whether information was provided about medication and whether patients were involved in decision-making within general practice consultations.</i>	Audio or videotape analysis.	(+)	(+)	(+)
Consultation Rating Scale (Hays, 1990) [52]	This consultation tasks rating scale uses evaluations such as 'explanations were adequate' or 'trainee listened attentively'.	Videotape analysis.	(-)	(-)	(-)
Communication Rating System CRS (Hulsman, 1998) [53]	Based on the Utrecht Consultation Assessment Method the CRS assesses 7 behavioural categories. Although information 'effectiveness' is itemised, no evaluation of involvement in decision-making occurs.	Audiotape analysis.	(-)	(-)	(-)
Daily Rating Form of Student Clinical Performance (White, 1991) [54]	This 6 section 17-item rating scale was developed to provide feedback to medical students on their interviewing skills.	Direct analysis.	(-)	(-)	(-)
<i>Elements of Informed Decision Making</i> (Braddock, 1997) [27]	This 6-item list covers the elements of 'informed consent'.	Videotape analysis.	(+)	(+)	(+)
Euro Communication Scale (Mead, 1999) [4]	A 5-item (dimensions) rating scale to assess patient-centredness.	Videotape analysis.	(+)	(+)	(+)
General Practice Interview Rating Scale (Verby, 1979) [55]	A 17-item 4 point rating scale of interviewing skills.	Audiotape analysis.	(-)	(-)	(-)
Interpersonal and Communication Skills Checklist (Cohen, 1976) [56]	A 17-item checklist developed for use by simulated patients after consultations.	Observation by simulated patients.	(-)	(-)	(-)
Interpersonal Skills Rating Form (Schnabl, 1995) [57]	A 13-item graded checklist developed to be used by standardized patients to assess the interpersonal skills of 4 th year medical students.	Observation by simulated patients.	(-)	(-)	(-)
Lehmann-Cote Checklist (Lehmann, 1990) [58]	A 41-item checklist assessed the 'presence' or 'absence' of tasks in chronological order within a consultation.	Videotape analysis.	(-)	(-)	(-)
Levels of Involvement (Marvel, 1994) [25]	This tool assesses the degree to which physicians explore patient psychosocial concerns.	Videotape analysis.	(+)	(+)	(+)
Lovett's Techniques of Interviewing Peer-Assessment Form (Lovett, 1990) [59]	This is a peer-review checklist covering interviewing skills developed within a communication course in psychiatry.	Direct analysis.	(-)	(-)	(-)
<i>MAAS-Global</i> , (van Thiel, 1991) [29]	<i>Communication skills derived by expert consensus. Measurement of medical interviewing skills (student assessment originally, but now adapted for general practice.)</i>	<i>Direct or recorded data analysis.</i>	(+)	(+)	(+)
<i>Patient-centredness: Component 3: Finding common ground</i> (Stewart, 1995) [30, 35]	<i>Items assess the degree of 'common ground' achieved within consultations. This is the third section of a 3 component instrument designed to measure patient-centredness.</i>	<i>Audio or videotape analysis.</i>	(+)	(-)	(+)

Using the OPTION instrument

Instrument name, date, first author	Description of instrument	Data collection	Addresses 'involvement'	Portrays options	Considers decision-making
Pendleton's Consultation Rating Scale (Pendleton, 1984) [20]	A 14-item consultation rating scale. Paired opposing statements are scored for agreement on a linear analogue scale. One item asks if the 'patient is involved in management adequately and appropriately' but there is no further elaboration.	Direct or recorded data analysis.	(+)	(-)	(-)
Physician Behaviour Checklist (PBCL) (Blanchard, 1986) [60]	A checklist developed to assess the behaviours of oncologists during ward rounds. Some items cover the discussion of tests and future treatment, but none that identify patient involvement in decision-making process.	Direct analysis.	(-)	(-)	(-)
Royal Australian College of General Practitioners Evaluation (Nyman, 1997) [21]	A checklist developed to assess the consulting skills of trainee general practitioners in Australia. One item asks if the patient was 'involved' in decision-making.	Direct observation.	(+)	(-)	(-)
Royal College of General Practitioners Examination Criteria (RCGP, 1998) [22].	Membership of the RCGP is by examination or by assessment. The criteria for consulting skills include one item about sharing 'management options' with the patient.	Videotape analysis.	(-)	(+)	(-)
Standard Index of Communication and Discrimination (SIC/SID): Levels of Response Scale (LRS) (Carkuff, 1969) [61, 62]	This index aims to measure the concepts of 'empathy, respect, concreteness, genuineness and confrontation' in communication processes.	Audiotape analysis.	(-)	(-)	(-)
Summative Assessment of General Practitioners (Campbell, 1996) [63, 64]	Based on the Pendleton consultation 'tasks', the instrument uses a 6-point scale. It is designed as a summative assessment of registrars in general practice.	Videotape analysis.	(-)	(-)	(-)
Interactional Styles Taylor (1989) [65]	Coding system devised to analyse interactional styles, including paternalism, maternalism, shared decision-making, and mixed styles.	Videotape analysis.	(-)	(-)	(-)
Telephone Assessment of TALK (Kosower, 1995) [66]	TALK is an acronym for 'trust, assert, listen and KISS (know, inquire, solve and stroke). This instrument categorises 24 generic behaviours into a conceptual framework and items are scored on a 5-point Likert scale.	Videotape analysis.	(-)	(-)	(-)
Teaching Communication Behaviour Scale (Clark, 1997) [67]	<i>10-item scale that lists behaviours associated with achieving compliance with long-term medication (e.g. asthma treatment).</i>	Videotape analysis.	(-)	(-)	(-)
University of Leeds Consultation Rating Scale (Stanley, 1985) [23]	<i>Communication skills derived by expert consensus. The aim of the guide is to act as a basis for formative assessment. Rating scales applied to defined behaviours and stages. UK-based general practice type consultations.</i>	Videotape analysis.	(-)	(-)	(+)
Utrecht Consultation Assessment Method [68]	UCAM is a checklist (incorporating a rating scale) which is divided into two categories: 'patient-centred approach' and 'doctor-patient interaction'. No further development work is being conducted on this instrument (personal communication).	Videotape analysis.	(-)	(-)	(-)

Measuring patient involvement

Instrument name, date, first author	Description of instrument	Data collection	Addresses 'involvement'	Portrays options	Considers decision-making
Type B: Interaction analysis measures - instruments which assess interactions at the 'micro' level (coded utterances or segments)					
Byrne and Long's Checklist of Behaviours (Byrne, 1992) [69]	This method subdivided the consultations into 85 'units of sense' and categorised them into three sets (doctor-centered, patient-centered and negative behaviour). The units were counted and a total score calculated for the consultation.	Audiotape analysis.	(-)	(-)	(-)
Cancer Specific Interaction Analysis System (CN-LOGIT) (Butow, 1991) [70]	This is an interaction analysis which was developed to assess the relationship between satisfaction with the consultation and the process and contents of consultations with an oncologist.	Audiotape analysis.	(-)	(-)	(-)
Davis Observation Coding 1991 (Callahan, 1991) [71]	Assessment of 20 behaviours (e.g. chatting, structuring interaction and counselling). 15-second intervals are coded. It is noteworthy that the operational definition for 'structuring interaction', a behaviour in which the patient's preferred and actual role in decision-making could be considered, specifically 'excluded planning treatment'.	Direct or recorded data analysis.	(-)	(-)	(-)
Faulkner's Communication Rating Scale [72]	Assessment of psychological concern by analysis of individual 'utterances'.	Transcript analysis.	(-)	(-)	(-)
Interaction System for Interview Evaluation (ISIE-81)[73]	Coding system developed by National Board of Medical Examiners for 2-second intervals or behaviour change (whichever comes first).	Videotape analysis.	(-)	(-)	(-)
Multi-dimensional Interaction Analysis System [74]	An interactional analysis method that lists 36 content areas and scores 'questioning, informing and supportiveness'.	Direct or recorded data analysis.	(-)	(-)	(-)
Measurement of Physician-Patient Communication (Kaplan, 1989) [75]	This coding scheme (a modification of Bales' interaction analysis) aimed to assess the attempts by patients to 'control' the interaction and judged the pattern of the consultation by quantifying utterances by both doctor and patient.	Audiotape analysis.	(-)	(-)	(-)
Medical Communication Behaviour System (MBCS) (Wolraich, 1986) [76]	13 provider behaviours and 10 patient behaviours are itemised and quantified. Physician behaviours are divided into 3 categories: Content, Affective and Negative Behaviours. The instrument is situated in the paternalistic paradigm. For instance, the item 'advice/ suggestion' is explained as, 'statements providing advice or suggestion on what the patient <i>should</i> do'. (Their italicisation).	Audio or videotape analysis.	(-)	(-)	(-)
Method for the Interactional Analysis of Doctor / Patient Consultation (Butler, 1992) [77]	This instrument codes the consultation by 'floorholding units' that are defined in terms of the content and form of communication categories e.g. physical agenda, emotional agenda and social agenda.	Videotape analysis.	(-)	(-)	(-)
McGee's Coding Method (McGee, 1998) [78]	Coding for patient utterances according to type of questions and by category (illness, treatment regimen, medical procedure, non-medical). Emphasis on information elicitation and verification.	Videotape analysis.	(-)	(-)	(-)

Using the OPTION instrument

Instrument name, date, first author	Description of instrument	Data collection	Addresses 'involvement'	Portrays options	Considers decision-making
Ockene's Counselling Assessment (Ockene, 1988) [79]	A three-item rating scale measuring the elicitation of feelings and information, and the provision of information.	Videotape analysis.	(-)	(-)	(-)
Patient-Doctor Communication Instrument (Waitzkin, 1985) [80]	This instrument gauges the 'amount' (in terms of time) of information transmittal that occurs between physicians and patients, and collects information about situational and sociodemographic variables that affect this area.	Audiotape analysis.	(-)	(-)	(-)
Physician-Patient Interaction Coding System (Makoul, 1992) [26]	This is an intricate coding scheme which assigns codes to 'events of interest' within speech turns (both patient and doctor utterances). Two phases are defined as 'problems' and 'solutions'. Within these phases, subordinate codes are described.	Audiotape analysis.	(-)	(-)	(-)
<i>Problem-Solving Observation Method</i> (Pridham, 1980) [24]	<i>Assess the interactive problem-solving behaviour of clients and clinicians. Interpersonal problem-solving skills.</i>	<i>Audio or videotape analysis.</i>	(+)	(+)	(+)
University of Leeds Consultation Rating Scale (Stanley, 1985) [23]	A 10-item rating scale designed for use within an undergraduate communication skills course.	Videotape analysis.	(+)	(+)	(-)
Relational Communication Control Coding Scheme [81]	Adaptation of Rogers [82] scheme which determines whether statements are assertions, questions, statements, 'talk-overs' or other categories and determines overall 'control' within interactions.	Audio or videotape analysis.	(-)	(-)	(-)
Roter's Interaction Analysis System (RIAS) (Roter, 1989) [40, 83]	The RIAS is derived from Bales' work assessing small group processes [84] but applies to the dyad scenario (i.e. doctor and patient) and consists of means of categorising each verbal utterance (distinguished in task-related behaviour and socio-emotional behaviour) and a set of global affect-ratings. Examples of the utterance codes include: Agree (shows agreement); [?] Med (closed medical question); Gives-Med (gives information-medical condition). The instrument's perspective is revealed by the following question: 'Did the physician summarize his/her recommendations near the end of the visit?'	Audio or videotape analysis.	(-)	(-)	(-)
Street's Coding Structure [85]	Utterances are coded into 9 categories.	Videotape analysis.	(-)	(-)	(-)
Verbal Response Mode (VRM) (Stiles, 1979) [86-88]	This system is based on work in psychotherapy that had developed (by observation) a framework of 'response' modes: Question, Advisement, Silence, Interpretation, Reflection, Edification, Acknowledgement, Confirmation and Unscorable.	Transcript analysis.	(-)	(-)	(-)
Verhaak [89]	This study used a coding system designed to observe the detection of psychological symptoms in primary care consultations. One item covered patient-centred behaviour during the prescribing phase.	Videotape analysis.	(-)	(-)	(-)

References

1. Charles C, Gafni A, Whelan T. Decision making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Soc Sci Med* 1999; 49: 651-61.
2. Entwistle V. Towards constructive innovation and rigorous evaluation: a new series on methods for promoting and evaluating participation. *Health Expect* 1999; 2: 75-7.
3. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000; 51: 1087-110.
4. Mead N, Bower P. Measuring patient-centredness: a comparison of three observation-based instruments. *Pat Educ Couns* 2000; 39: 71-80.
5. Guadagnoli E, Ward P. Patient participation in decision-making. *Soc Sci Med* 1998; 47(3): 329-39.
6. Benbassat J, Pilpel D, Tidhar M. Patients' preferences for participation in clinical decision-making: a review of published surveys. *Behav Med* 1998; 24: 81-88.
7. Wolf AMD, Nasser JF, Wolf AM, Schorling JB. The impact of informed consent on patient interest in prostate-specific antigen screening. *Arch Intern Med* 1996; 156: 1333-6.
8. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. *JAMA* 1992; 267: 2221.
9. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). *Soc Sci Med* 1997; 44: 681-92.
10. Towle A, Godolphin W. Framework for teaching and learning informed shared decision making. *BMJ* 1999; 319(7212): 766-71.
11. Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision-making and the concept of equipoise: defining the competences of involving patients in healthcare choices. *Br J Gen Pract* 2000; 50: 892-900.
12. Elwyn G, Edwards A, Wensing M, Hibbs R, Wilkinson C, Grol R. Shared decision-making observed: visual displays of communication sequence and patterns. *J Eval Clin Pract* 2001; 7(2): 211-21.
13. Schneider CE. *The practice of autonomy: patients, doctors, and medical decisions*. New York: Oxford University Press, 1998.
14. Elwyn G, Edwards A, Kinnersley P. Shared decision-making: the neglected second half of the consultation. *Br J Gen Pract* 1999; 49: 477-82.
15. Brown J, Stewart M, Tessier S. *Assessing communication between patients and doctors: a manual for scoring patient-centred communication*. 1995, Centre for Studies in Family Medicine, University of Western Ontario, Canada.
16. Kaplan SH, Gandek B, Greenfield S, Rogers W, Ware JE. Patient and visit characteristics related to physicians' participatory decision-making style. *Med Care* 1995; 33: 1176-87.
17. Elwyn G, Edwards A, Gwyn R, Grol R. Towards a feasible model for shared decision-making: a focus group study with general practice registrars. *BMJ* 1999; 319: 753-7.
18. NHS Centre for Reviews and Dissemination. *Undertaking systematic reviews of research or effectiveness*. York: University of York: NHS CRD, 1996.
19. Streiner N, Norman GR. *Health Measurement Scales: A practical guide to their development and use*. 2nd ed. Oxford: Oxford University Press, 1995.
20. Pendleton D, Schofield T, Tate P, Havelock P. *The consultation: an approach to learning and teaching*. Oxford: Oxford University Press, 1984.
21. Nyman KC, Sheridan B. Evaluation of consulting skills of trainee general practitioners. *Aust Fam Physician* 1997; 26(Suppl. 1): S28-35.
22. RCGP. *Assessment of consulting skills: MRCGP Examination criteria*. London: Royal College of General Practitioners, 1998.
23. Stanley IM, Webster CA, Webster J. Comparative rating of consultation performance: a preliminary study and proposal for collaborative research. *J Roy Coll Gen Pract* 1985; 35: 375-80.
24. Pridham KF, Hansen MF. An observation methodology for the study of interactive clinical problem-solving behaviour in primary care settings. *Med Care* 1980; 18: 360-75.
25. Marvel MK, Schilling R. Levels of physician involvement with patients and their families: a model for teaching and research. *J Fam Pract* 1994; 39: 535-44.
26. Makoul G. *Perpetuating passivity: a study of physician-patient communication and decision-making*. Evanston, IL: Northwestern University, 1992.
27. Braddock CH, Fihn SD, Levinson W, Jonsen AR, Rearlman RA. How doctors and patients discuss routine clinical decisions: informed decision making in the outpatient setting. *J Gen Intern Med* 1997; 12: 339-45.
28. Kurtz SM, Silverman JD. *The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organising the teaching in communication training programmes*. *Med Educ* 1996; 30: 83-9.
29. van Thiel J, Kraan HF, van der Vleuten CPM. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht history-taking and advice checklist. *Med Educ* 1991; 25: 224-9.

30. Stewart M, Brown JJ, Donner A, McWhinney IR, Oates J, Weston W. The impact of patient-centred care on patient outcomes in family practice. London, Ontario: Centre for Studies in Family Medicine (Final report), 1995.
31. Doherty WJ, Baird MA. Developmental levels in family-centred medical care. *Fam Med* 1986; 18: 153-6.
32. van Thiel J, van der Vleuten C, Kraan H. Assessment of medical interviewing skills: generalisability of scores using successive MAAS-versions. in *Approaches to the assessment of Clinical Competence: Proceedings of the Fifth Ottawa Conference*. Dundee, Scotland, 1992.
33. Stewart M, Brown JB, Weston WW, McWinney IR, McWilliam CL, Freeman TR. *Patient Centred Medicine: Transforming the Clinical Method*. Thousand Oaks, CA: Sage Publications, 1995.
34. Towle A. *Physician and Patient Communication Skills: Competencies for Informed Shared Decision-Making*. Informed Shared Decision-Making Project: Internal Report. Vancouver, Canada: University of British Columbia, 1997.
35. Henbest R, Stewart M. Patient-centredness in the consultation. 1: a method for assessment. *Fam Pract* 1990; 7: 249-54.
36. Kinnersley P. The patient-centredness of consultations and the relationship to outcomes in primary care. Unpublished MD thesis, University of Bristol, 1997.
37. Gafni A, Charles C, Whelan T. The physician-patient encounter: the physician as a perfect agent for the patient versus the informed decision-making model. *Soc Sci Med* 1998; 47: 347-54.
38. DiMatteo MR, Reiter RC, Gambone JC. Enhancing medication adherence through communication and informed collaborative choice. *Health Communication* 1994; 6: 253-65.
39. Towle A, Godolphin W. Framework for teaching and learning informed shared decision making. *BMJ* 1999; 319: 766-9.
40. Roter DL. *The Roter method of interaction process analysis*. Baltimore: The John Hopkins University, Department of Health Policy and Management, 1991.
41. Inui TS, Carter WB. Problems and prospects for health service research on provider-patient communication. *Med Care* 1985; 23: 521-38.
42. Boon H, Stewart M. Patient-physician communication assessment instruments: 1986 to 1996 in review. *Pat Educ Couns* 1998; 35: 161-76.
43. Kurtz S, Silverman J, Draper J. *Teaching and Learning Communication Skills in Medicine*. Abingdon: Radcliffe Medical Press, 1998.
44. Cronbach LJ, Gleser GC, Nanda H, Rajaratnam N. *The dependability of behavioural measurements: theory of generalizability for scores and profiles*. New York: Wiley, 1972.
45. Stillman P. Construct validation of the Arizona Clinical Interview Rating Scale. *Educational and Psychological Measurement* 1977; 37(4): 1031-8.
46. Stillman PL, Sabers DL, Redfield DL. Use of para-professionals to teach and evaluate interviewing skills in medical students. *Paediatrics* 1976; 60: 165-9.
47. Cox J, Mullholland H. An instrument for assessment of videotapes of general practitioners' performance. *BMJ* 1993; 306: 1043-6.
48. Simmons J, Roberge L, Kendrick Fr SB, Richards B. The interpersonal relationship in clinical practice: the Barrett-Lennard relationship inventory as an assessment instrument. *Evaluation and the Health Professions* 1995; 18: 112.
49. Bensing JM. Doctor-patient communication and the quality of care. *Soc Sci Med* 1991; 32: 1301.
50. Burchard KW, Rowland-Morin PA. A new method of assessing the interpersonal skills of surgeons. *Acad Med* 1990; 65: 274.
51. Mazzuca SA, Weiberger M, Kurpius DJ, Froehle TC, Heister M. Clinician communication associated with diabetic patients' comprehension of their therapeutic regimen. *Diabetes Care* 1983; 6: 347-351.
52. Hays RB. Assessment of general practice consultations: content validity of a rating scale. *Med Educ* 1990; 1990(24): 1110-6.
53. Hulsman RL. *Communication skills of medical specialists in oncology*. Utrecht: NIVEL Netherlands Institute of Primary Health Care, 1998.
54. White DG, Tiberius R, Talbot Y, Schiralli V, Richett M. Improving feedback for medical students in a family medicine clerkship. *Can Fam Phys* 1991; 37: 64-70.
55. Verby JE, Holden P, Davis RH. Peer review of consultations in primary care: the use of audiovisual recordings. *BMJ* 1979; 1: 1686-8.
56. Cohen DS, Colliver JA, Marcy MS, Fried ED, Swartz MH. Psychometric properties of a standardized-patient checklist and rating-scale form used to assess interpersonal and communication skills. *Acad Med* 1996; 71(S1): S87-89.
57. Schnabl GK, Hassard TH, Kopelow ML. The assessment of interpersonal skills using standardised patients. *Acad Med* 1995; 66 (Suppl. 9): S534-6.
58. Lehmann F, Cote L, Bourque A, Fontaine D. Physician-patient interaction: a reliable and valid checklist of quality. *Can Fam Phys* 1990; 36: 1711-6.
59. Lovett LM, Cox A, Abou-Saleh M. Teaching psychiatric interview skills to medical students. *Med Educ* 1990; 24: 243.
60. Blanchard CG, Ruckdeschel JC, Fletcher BA, Blanchard EB. The impact of oncologists' behaviors on patient satisfaction with morning rounds. *Cancer* 1986; 58: 387-393.

61. Carkuff RR. The prediction of the effects of teacher-counsellor education: the development of communication and discrimination selection indexes. *Couns Educ Supervision* 1969; 8: 265-72.
62. Engler CM, Saltzman GA, Walker ML, Wolf FM. Medical student acquisition and retention of communication and interviewing skills. *J Med Educ* 1981; 56: 572-9.
63. Campbell LM, Howie JGR, Murray TS. Use of videotaped consultations in summative assessment in general practice. *Br J Gen Pract* 1993; 45: 137-41.
64. Campbell LM, Murray TS. Summative Assessment of vocational trainees: results of a 3-year study. *Br J Gen Pract* 1996; 46: 411-4.
65. Taylor SG, Pickens JM, Geden EA. Interactional styles of nurse practitioners and physicians regarding patient decision making. *Nursing Research* 1989; 38: 50-5.
66. Kosower E. An assessment and teaching tool for interpersonal communication skills. *Acad Med* 1995; 70(5): 452-3.
67. Clark NM, Gong M, Schork M, Maiman LA, Evans D, Hurwitz ME. A scale for assessing health care providers' teaching and communication behaviour regarding asthma. *Health Education and Behaviour* 1997; 24: 245-256.
68. Pieters HM, Touw-Otten FW, De Melker RA. Simulated patients in assessing consultation skills of trainees in general practice vocational training: a validity study. *Med Educ* 1994; 28: 226-33.
69. Byrne PS, Long BEL. *Doctors talking to patients*. London: HMSO, 1976.
70. Butow PN, Dunn SM, Tattersall MHN, Jones QJ. Computer-based interaction analysis of the cancer consultation. *Br J Cancer* 1995; 71: 115-21.
71. Callahan EJ, Bertakis KD. Development and validation of the Davis Observation Code. *Fam Med* 1991; 5(3): 19-24.
72. Argent J, Faulkner A, Jones A, O'Keefe C. Communication skills in palliative care: development and modification of a rating scale. *Med Educ* 1994; 28: 559-65.
73. Badger L, deGruy F, Hartman J, Plant MA, Leeper J, Ficken R, Maxwell A, Rand E, Anderson R, Templeton B. Psychosocial interest, medical interviews and the recognition of depression. *Arch Fam Med* 1994; 3: 899-907.
74. Charon R, Greene MG, Adelman RD. Multi-dimensional interaction analysis: a collaborative approach to the study of medical discourse. *Soc Sci Med* 1994; 39: 955-65.
75. Kaplan SH, Greenfield S, Ware JE. Assessing the effects of physician-patient interactions on the outcomes of chronic disease. *Med Care* 1989; 27(Suppl): S 110-27.
76. Wolraich ML, Albanese M, Stone G, et al. Medical communication behaviour system: an interaction analysis system for medical interactions. *Med Care* 1986; 24: 891-903.
77. Butler NM, Campion PD, Cox AD. Exploration of doctor and patient agendas in general practice consultations. *Soc Sci Med* 1992; 35: 1145-55.
78. McGee DS, Cegala D. Patient communication skills training for improved communication competence in the primary care medical consultation. *Journal of Applied Communication Research* 1998; 26: 412-30.
79. Ockene J, Quirk M, Goldberg R, Kristeller J, Donnelly G, Kalan K, Gould B, Greene H, Harrison-Atlas R, Pease J, Pickens S, Williams J. A residents' training program for the development of smoking intervention skills. *Arch Intern Med* 1988; 148: 1039-45.
80. Waitzkin H. Information giving in medical care. *J Health Soc Behaviour* 1985; 26: 81.
81. Cecil DW. Relational control patterns in physician-patient clinical encounters: continuing the conversation. *Health Communication* 1998; 10: 125-49.
82. Rogers LE, Farace RV. Analysis of relational communication in dyads: new measurement procedures. *Human Communication Research* 1975; 1: 222-39.
83. Roter DL. The Roter method of interaction process analysis (unpublished manuscript). Baltimore, MD: The John Hopkins University, 1989.
84. Bales R. *Interaction Process Analysis*. Cambridge, Massachusetts: Addison-Wesley, 1950.
85. Street RLS. Analyzing communication in medical consultations: do behavioural measures correspond to patients' perceptions? *Med Care* 1992; 30: 976-88.
86. Stiles WB. *Manual for a taxonomy of verbal response modes*. Chapel Hill: Institute for Research in Social Science, University of North Carolina, 1978.
87. Stiles WB, Putnam SM, Wolf MH, James SA. Interaction exchange structure and patient satisfaction with medical interviews. *Med Care* 1979; 17: 667-81.
88. Stiles WB, Putnam SM, James SA, Wolf MH. Dimensions of patient and physician roles in medical screening interviews. *Soc Sci Med* 1979; 13A: 335-41.
89. Verhaak P. Detection of psychologic complaints by general practitioners. *Med Care* 1988; 26: 1009-20.

4. OPTION publications

4.1. Shared decision making: developing the OPTION scale for measuring patient involvement

Glyn Elwyn, Adrian Edwards, Michel Wensing, Kerry Hood, Christine Atwell, Richard Grol.

Quality and Safety in Health Care 2003; 12: 93-99.

Permission for re-publication granted by BMJ Journals 23/12/04.

Background: A systematic review has shown that no measures of the extent to which healthcare professionals involve patients in decisions within clinical consultations exist, despite the increasing interest in the benefits or otherwise of patient participation in these decisions.

Aims: To describe the development of a new instrument designed to assess the extent to which practitioners involve patients in decision making processes.

Design: The OPTION (observing patient involvement) scale was developed and used by two independent raters to assess primary care consultations in order to evaluate its psychometric qualities, validity, and reliability.

Study sample: 186 audio-taped consultations collected from the routine clinics of 21 general practitioners in the UK.

Method: Item response rates, Cronbach's alpha, and summed and scaled OPTION scores were calculated. Inter-item and item-total correlations were calculated and inter-rater agreements were calculated using Cohen's kappa. Classical inter-rater intra-class correlation coefficients and generalisability theory statistics were used to calculate inter-rater reliability coefficients. Basing the tool development on literature reviews, qualitative studies and consultations with practitioner and patients ensured content validity. Construct validity hypothesis testing was conducted by assessing score variation with respect to patient age, clinical topic "equipoise", sex of practitioner, and success of practitioners at a professional examination.

Results: The OPTION scale provided reliable scores for detecting differences between groups of consultations in the extent to which patients are involved in decision making processes in consultations. The results justify the use of the scale in further empirical studies. The inter-rater intraclass correlation coefficient (0.62), kappa scores for inter-rater agreement (0.71), and Cronbach's alpha (0.79) were all above acceptable thresholds. Based on a balanced design of five consultations per clinician, the inter-rater reliability generalisability coefficient was 0.68 (two raters) and the intra-rater reliability generalisability coefficient was 0.66. On average, mean practitioner scores were very similar (and low on the overall scale of possible involvement); some practitioner scores had more variation around the mean, indicating that they varied their communication styles to a greater extent than others.

Conclusions: Involvement in decision making is a key facet of patient participation in health care and the OPTION scale provides a validated outcome measure for future empirical studies.

Introduction

The involvement of patients in shared decision making has been the subject of debate [12], with some claiming that it should be mandatory while others point out the problems [3], but it remains an area where few empirical studies have been conducted [4]. A systematic review has shown that there is no existing measure of the extent to which healthcare professionals involve patients in decisions within clinical consultations [5]. Although some instruments include some components of patient involvement [6–11], they were found to be insufficiently developed to measure accurately this facet of communication in patient-clinician interactions. The underlying ethical principles of patient autonomy and veracity underpin this development and, coupled with the interest of consumers, professionals and policy makers, drive a research need to ascertain whether achieving greater involvement in decision making is associated with improved patient outcomes.

The area is complex and the concept is not easy to measure. It is reported that, typically, less than 50% of patients wish to be involved in the decision making processes [1,12,13] despite the possibility that “involvement” could have a positive effect on health outcomes [7,14,15]. Recent qualitative research conducted with a wide range of consumer and patient groups revealed only minor reservations about participation in decision making processes, provided the process was sensitive to individual preferences at any given time points [16,17].

Patients stated that professionals should definitely provide information about treatment options, but should respect the extent to which patients wish to take on decision making responsibilities in clinical settings. The underlying principles of the shared decision making method have been described elsewhere [18–20] and, following a literature review [5,21] and a series of qualitative and quantitative studies [5,21–24], a skills framework has been proposed [25]. This framework is composed of a set of competences that include the following steps:

- problem definition (and agreement);
- explaining that legitimate choices exist in many clinical situations, a concept defined as professional “equipoise” [25];
- portraying options and communicating risk about a wide range of issues - for example, entry to screening programmes or the acceptance of investigative procedures or treatment choices); and
- conducting the decision process or its deferment.

These are all aspects of consultations that need to be considered by an instrument designed to assess whether clinicians engage patients in decisions [25]. It is the accomplishment of these competences that forms the conceptual basis for the OPTION scale.

OPTION (acronym for “observing patient involvement”) is an item based instrument completed by raters who assess recordings of consultations (audio or video). It has been developed to evaluate shared decision making specifically in the context of general practice, but it is intended to be generic enough for use in all types of consultations in clinical practice. The OPTION scale is designed to assess the overall shared decision making process. In summary, it examines whether problems are well defined, whether options are formulated, information provided, patient understanding and role preference evaluated, and decisions examined from both the professional and patient perspectives.

Some suggest that clinical practice should be categorised by a taxonomy of policies - that is, whether the screening, testing, or treatment under consideration is a “standard”, a “guideline”, or an “option”—and that clinicians should vary the degree of patient involvement on this basis. “Standards” theoretically provide strong evidence of effectiveness and strong agreement about best treatment. “Guidelines” are less prescriptive and, where there are “options”, the evidence regarding effectiveness or otherwise is unclear. It is then proposed that patient involvement be reserved for situations where clear “options” exist. This scale was designed, however, from the standpoint that there are opportunities for patients to be involved in decisions across the spectrum of evidence for effectiveness or professional agreement about best practice. Firstly, there are few situations where interventions are free from harm, and so it is almost always appropriate to raise awareness about such outcomes. Secondly, patients have legitimate perspectives on many social and psychological aspects of decisions whereas the evidence base almost certainly restricts itself to providing data about the biomedical aspects of decision making. The instrument developed was therefore a generic tool capable of assessing the extent to which clinicians involve patients in decisions across a range of situations, excluding emergencies or other compromised circumstances. The aim of the study was to enable accurate assessments of the levels of involvement in shared decision making achieved within consultations in order to provide research data for empirical studies in this area. This paper describes the development of the instrument and assesses its ability to discriminate involvement levels and the decision making methods used in consultations within and between differing practitioners by reporting key aspects of the tool’s validity and reliability using a sample of consultations recorded in a general practice setting.

Methods

The psychometric characteristics of the OPTION scale were applied to a sample of audio-taped consultations collected from the routine clinics of 21 GPs and rated by two observers. Validity issues were considered at both theoretical (construct emergence) and item formulation and design stages; construct validity was also investigated. The reliability of the scale was calculated by assessing response rates, inter-item and item-total correlations, inter-rater agreement (kappa), and inter- and intra-rater reliability coefficients using both classical and generalisability theory statistical methods. Approval to conduct the work was obtained from the Gwent local research ethics committee.

Overall design features

The content validity of the instrument was developed by appraising existing research and undertaking qualitative studies to define the clinical competences of patient involvement in shared decision making in clinical consultations [5,18,19, 25].

Content validity and concept mapping

The development process followed established guidelines [26]. The systematic review [5] allowed existing scales - especially measures of related concepts such as “patient centred-ness” and “informed decision making” [7,27] - to be considered critically. Qualitative studies using key informants to clarify and expand the competences revealed that clinicians have specific perceptions about what constitutes “involvement in decision making” which are matched in part, but not entirely, by patient views [25] and emphasised the

importance of checking patient role preference (item 10, table 2). The use of design and piloting iterations involving both patient and clinician groups ensured content validity and formulated items. In addition, a sample of consultations in which clinicians were intent and experienced at involving patients in discussions and sharing decisions were purposively chosen and examined [23]. Thus, the theoretical construct was refined by an assessment of clinical practice [22]. The synthesis of this body of work enabled the development of a theoretical framework for patient involvement in decision making and informed the design of the OPTION instrument.

Instrument and scale development

An 18-item pilot instrument was used by five GP key informants [25] and one non-clinical rater to assess six simulated audio-taped consultations; item refinement and scale development involved three iterative cycles over a 12 month period. These simulated consultations had been modelled to contain differing levels of patient involvement and decision making methods. This process reduced item ambiguity, removed value laden wordings, and resulted in short and (where possible) positively worded items [26]. A 5-point scale, anchored at both ends with the words “strongly agree” and “strongly disagree”, was used to avoid the loss of scoring efficiency in dichotomised measures [26]. Revisions included removing two duplicative items, increasing the focus on observable “clinician behaviour” rather than attempting to assess patient perceptions of the consultation, and modifying item sequence. This version was subjected to further piloting using a second calibration audiotape containing modelled consultations (two “paternalistic” consultations, three “shared decision making” and two “informed choice” examples). These consultations were rated by two non-clinical raters using the OPTION scale and two other scales—namely, the determination of “common ground” developed by Stewart et al in Ontario [7] and Braddock’s measure of “informed decision making” [27] - which were selected as the most comparable scales identified [5]. The raters provided written feedback and regarded the pilot 16-item OPTION instrument as a more acceptable and feasible tool. For the assessment of the simulated tapes the OPTION scale achieved an inter-rater reliability correlation coefficient of 0.96 compared with a score of 0.76 for the Braddock scale and 0.4 for the Stewart “common ground” scale. These initial results were therefore promising and a stable version of the instrument (June 2000) was described in a manual for raters. By participating in item revision and the development of the manual drafting, the raters were integrated into a calibration process before applying the instrument to a series of naturally occurring consultations.

Data collection: practitioner and patient samples

To test the instrument, recordings of consultations were taken from the recruitment phase of a proposed trial of shared decision making and risk communication [28]. As part of the recruitment process to the study, GPs in Gwent, South Wales were asked to audiotape consecutive consultations during a routine consulting session in general practice. To be eligible for possible recruitment into the trial the GPs had to have been principals in a general practice for at least 1 year and less than 10 years. The potential sample pool of 104 GPs in 49 practices (mean age 41 years, 62% men) was initially approached by letter (followed by telephone contact) and asked to participate in a research trial. As far as we are aware, these volunteer practitioners were naïve to the concepts that we were measuring and had not been exposed to any training or educational interventions that could have influenced their proficiency in this area. Patients attending on

the specified recording dates gave their consent using standard procedures, and their age and sex were recorded. Apart from these consent procedures, no other stipulations were imposed and the data collected contained recordings covering the range of conditions typically seen in routine general practice sessions. Each consultation recording (Spring 2000) was rated in the autumn using the OPTION instrument by two calibrated raters who were non-clinical academics in social sciences and who remained independent of the main research team. Tapes are available for re-assessment. A random sample of 21 consultations (one per clinician) was selected for test-retest analysis and repeated ratings conducted by the two raters.

Data analysis

The data were analysed by taking the response to each item and calculating a summed OPTION score which was then scaled to lie between 0 (least involved) and 100 (most involved). Inter-item and item-total correlations were calculated and inter-rater agreements were calculated using Cohen's kappa.

As well as assessing a classical inter-rater intraclass correlation coefficient, the inter-rater and intra-rater reliability coefficients of the instrument were calculated using the statistical techniques described in generalisability theory [29, 30]. This theory uses modified analysis of variance techniques to generate "generalisability coefficients" [26]. The methods enable multiple sources of error variance to be calculated and subsequent generalisations to be made about the degree to which these sources are contributing to the overall variability. This allows decisions to be made about the effect of changing the characteristics of the measurement process -for example, number of raters or number of consultations per practitioner [26]- in order to assess the instrument's reliability. We also estimated whether consultation scores clustered within practitioners by calculating an intraclass correlation coefficient [31] and the homogeneity of the scale by calculating Cronbach's alpha [32]. Using the mean scores of the two raters, the Kaiser-Meyer-Olkin measure of sampling adequacy was assessed, inter-item correlations and item-total correlation were calculated, and confirmatory factor analysis performed to determine whether the scale could be legitimately considered as a measure of a single construct.

Assessment of the construct validity of the OPTION instrument was conducted by examining four hypothetical constructs —namely, that the OPTION score level would be influenced by patient age (negative), sex of clinician (positive in favour of female), qualification of clinician (positive), and whether the clinical topic was one where clinical equipoise existed (positive). The existence of equipoise was determined by a clinical assessment of the audiotape sample content (GE). Studies have also examined the effect of sex of the physician on communication within consultations. Although an area of debate [33], Hall et al [34] found that female physicians made more partnership statements than male physician and Coates' review [35] reported a broad consensus that female language is generally more cooperative. Although there is no consistent evidence, we examined this by comparing the mean OPTION scores for the eight female clinicians with those of their 13 male colleagues (t test). In 1995 the examination for membership of the Royal College of General Practitioners, UK (MRCGP) introduced a video assessment and listed shared decision making as a merit criterion. Although there exists evidence that GPs in training do not involve patients in decision making [36], it was conjectured that success in the examination (at any time, before 1995, or after 1995) might be associated with higher scores (t test), although we did not expect strong

correlations. It has been established in cross sectional studies that increasing patient age leads to less patient preference for involvement [12,13], and we assessed the correlation (Pearson) between OPTION scores and patient age. It was also hypothesised from previous qualitative work that decisions were more likely to be shared in consultations that contained clinical problems characterised by professional equipoise such as hormone replacement therapy [25]. The consultations were differentiated (by GE) according to this characteristic and any significant differences between the mean OPTION scores were determined (weighted t test). No attempt was made to establish criterion (specifically concurrent) validity.

Table 1. Demographic and clinical characteristics of the recorded consultations (n=186)

Male/female	60 (32%)/126 (68%)
Age (years)	Mean 43.3, SD 20.6, range 4 months–83 years
Duration of consultation (min)	Mean 8.2, SD 4.0, median 7.3, range 22.5
<i>Clinical problems</i>	
Respiratory	28
Musculoskeletal	27
Dermatological	21
Psychological	13
Cardiovascular	12
Hypertension	11
HRT	11
Other	63

Results

Sample characteristics

Of the potential sample pool of 104 practitioners, 21 GPs in separate practices who showed interest in being recruited into the trial provided a tape of a routine clinic before receiving any detailed information about the proposed research. These GPs represented a slightly younger group than the sampling frame (mean age 38 years), identical M:F ratio (38% female), and 16 (76%) had been successful in the membership examination of the Royal College of General Practitioners compared with an overall membership level of 54% in the sampling frame. Of the 242 consecutive patients approached in all practices, 12 (5%) declined to have the consultation recorded (the maximum refusal in any one practice was three patients in a series of 15). The remaining 230 consultations were assessed and, after removing consultations where there were technical recording problems, 186 consultations were available for analysis (average of 8.8 consultations per practitioner). There was no age and sex difference between the consultations excluded because of poor recordings and those included for analysis. One practitioner recorded five consultations but most recorded eight or more. There were twice as many consultations with women in the sample and 66% of the patients seen were aged between 30 and 70 years. The demographic and clinical characteristics of the recorded consultations are summarised in table 1.

Scale refinement

The performance of the 16-item scale was analysed in detail. Four of the items had been formulated to try and discriminate between styles of clinician decision methods to distinguish between paternalism, on the one hand, and the transfer of decisional responsibility to the patient on the other. The other 12 items had

been constructed to determine performance within a construct of a defined set of steps and skills. The reliability of items that attempted to differentiate between decision making styles was poor, and a decision was made to focus on a scale that was composed of the items that specifically evaluated the agreed competence framework. It is the reliability and construct validity of this 12-item scale that is reported.

Table 2. Option item response, missing value rates (%), and Cohen's kappa

OPTION scale item	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Missing values (%)	Kappa score*
The clinician identifies a problem(s) needing a decision making process	49.5	33.1	11.0	4.3	1.3	0.8	0.61 (0.31)
The clinician states that there is more than one way to deal with an identified problem ("equipoise")	6.2	3.2	5.4	13.4	71.0	0.8	0.82(0.50)
The clinician lists "options" including the choice of "no action" if feasible	6.7	4.0	7.0	9.7	71.8	0.8	0.75(0.51)
The clinician explains the pros and cons of options to the patient (taking "no action" is an option)	3.5	3.2	9.4	11.6	71.5	0.8	0.68(0.43)
The clinician checks the patient's preferred information format (words/numbers/visual display)	0	0	0.3	0.5	98.4	0.8	0.98(0.98)
The clinician explores the patient's expectations (or ideas) about how the problem(s) are to be managed	0.5	1.1	8.6	18.8	69.9	1.1	0.75(0.34)
The clinician explores the patient's concerns (fears) about how problem(s) are to be managed	1.3	4.6	12.1	22.0	59.1	0.8	0.53(0.42)
The clinician checks that the patient has understood the information	0.8	1.1	35.2	26.9	34.9	1.1	0.38(0.10)
The clinician provides opportunities for the patient to ask questions	1.9	3.2	40.1	17.2	36.0	1.6	0.20(-0.08)
The clinician asks for the patient's preferred level of involvement in decision making	0.8	1.3	4.0	8.1	84.9	0.8	0.86(0.66)
An opportunity for deferring a decision is provided	1.1	2.4	4.8	7.5	83.3	0.8	0.83(0.66)
Arrangements are made to review the decision (or the deferment)	19.4	7.8	35.2	5.4	30.9	0.8	0.58(0.44)

*Kappa scores are for agreement across sum of "agree, neutral and disagree" scale points; scores in parentheses are kappa scores for 5-point scale agreement.

Response rates to OPTION items

Items 1, 2, 3, 4, and 6 had a range of responses across the 5-point scale but with a predominance of low scores (see table 2 for summary of responses to items). Oversights in item completion led to an average of 0.9% missing values that were distributed evenly across all items (see table 2). The results indicate that the clinicians generally did not portray equipoise (71% strongly disagree); they did not usually list options (71.8% strongly disagree); they did not often explain the pros and cons of options (71.5% strongly disagree); and they did not explore patients' expectations about how the problems were to be managed (69.9% strongly disagree). Responses to items 7, 8, and 9 revealed most variation across scale points. Item

7 asked whether the clinician explored the patients' concerns (fears) about how the problem(s) were to be managed: the response was 81.1% disagreement and 12.1% neutral. A similar pattern of disagreement with the assertion that the clinician "checks patient understanding" and provides "opportunities for questions" (items 8 and 9) was obtained but with higher scores for the neutral scale point (35.2% and 40.1%, respectively). Clinicians were infrequently observed to "ask patients about their preferred level of involvement in decision making" (84.9% strongly disagree).

Opportunities for deferring decisions were rarely observed (item 11, 3.5% agreement) but an arrangement to review problems in the consultation was made in over a quarter of the consultations (item 12, 27.2% agreement). To summarise, the responses obtained indicate that the consultations recorded during these routine surgeries are characterised by low levels of patient involvement in decision making and a largely paternalistic approach by the GPs. This is confirmed by the fact that the items that assess equipoise, option listing, and information provision (items 2, 3 and 4) achieved a mean agreement response rate of 8.6%.

Reliability of the OPTION score (summed and scaled scores)

For all 12 items the mean Cohen kappa score was 0.66, indicating acceptable inter-rater agreement for this type of instrument after correcting for chance 0.37. Exclusion of item 9 (which requires further attention because of its low kappa score) increased the mean kappa score to 0.71. For the kappa scores the scale was aggregated to three points (agree, neutral, disagree; see table 2). Five point kappa scores are shown in parentheses. Coefficient alpha (Cronbach's alpha) was 0.79, indicating little redundancy in the scale (using the mean of the two rater scores). The inter-rater intraclass correlation coefficient for the OPTION score was 0.62. Based on a balanced design of the first five consultations on each practitioner's audiotape, the inter-rater reliability generalisability coefficient was 0.68 (two raters) and, using the test-retest data, the intra-rater reliability generalisability coefficient was 0.66. The corrected item-total correlations lay between 0.35 and 0.66 except for items 1 and 5 which had correlations of 0.05 and 0.07, respectively. Kaiser-Meyer-Olkin measure of sampling adequacy was 0.82, indicating a very compact pattern of item correlation and justifying the use of factor analysis.

Confirmatory factor analysis using principal components revealed that variable loading scores in a forced single factor solution resulted in scores that were above 0.36 (the recommended thresholds for sample sizes of approximately 200) for all except items 1 and 5 (-0.10 and 0.09). Item 1 asked whether a "problem" is identified by the clinician and perhaps should be regarded as a gateway item to the scale—that is, if a problem is not identified then it is difficult to see how the other items can be scored effectively. Item 5 had a low endorsement rate which was anticipated given current practice. Items 2–4 and 6–12 had a mean factor loading of 0.64. A total of 35.2% of the variance was explained by one latent component. Of a total of 66 possible inter-item correlations, 49 were above 0.25 (mean $r = 0.40$).

Given these reliability indicators, the overall mean (SD) OPTION score for all clinicians on a scale of 0–100, averaged across both rater scores, was 16.9 (7.7), 95% confidence interval 15.8 to 18.0, with a minimum score of 3.3 and a maximum of 44.2 across the sample. The scores are skewed towards low values (see fig 1). At the individual clinician level the mean OPTION scores lay between 8.8 and 23.8 with an intraclass correlation coefficient of 0.22 (across individual means), indicating significant clustering of consultation

scores within clinicians. These scores and the quartiles for each practitioner are shown in fig 2. Note that some clinicians have a much wider range of involvement score, indicating a more variable consulting style. The results show that the general level of patient involvement achieved in these consultations was low.

Figure 1. Distribution of OPTION scores.

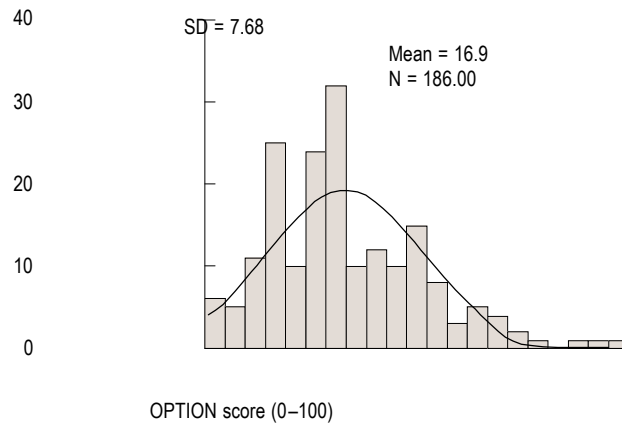
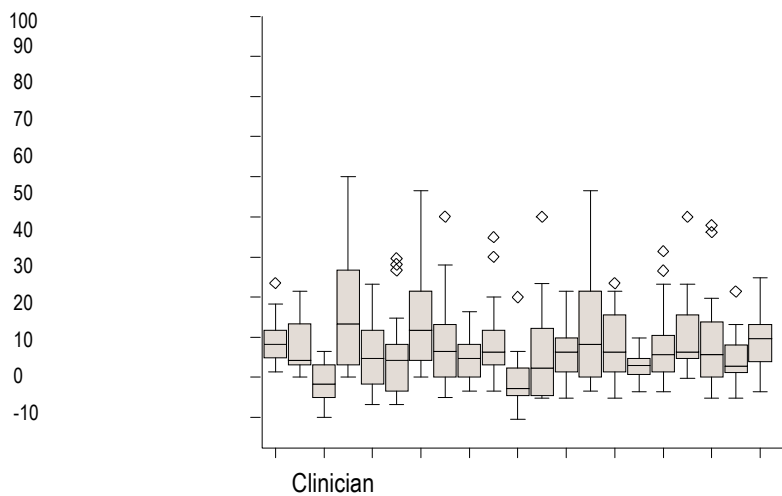


Figure 2. Mean OPTION scores for clinicians (box plots)



Construct validity

Two constructs were found to be correlated with levels of involvement in decision making—namely, patient age and the existence of a clinical topic where professional equipoise could be expected. The correlation coefficient between the mean OPTION score and patient age (adult age range) was -0.144 ($p < 0.01$) and confirmed the hypothesis that involvement levels reduced as patient age increased. Although this was a small sample, it was found that consultations that contained clinical problems characterised by having a greater likelihood of professionals exhibiting equipoise about treatment choice ($n = 15$ consultations, 8.1%), such as discussion of HRT or depression, had a mean OPTION score of 21.6 which was significantly higher than the mean scores achieved in consultations where equipoise topics did not occur (16.4, $p < 0.01$,

weighted t test), confirming the hypothesis that involvement increases where this characteristic exists. Sex of the clinician and success or otherwise in the MRCGP examination were not associated with differences in OPTION scores.

Discussion

Principal findings

The results of this study show that the OPTION scale provides a method of scoring the extent to which clinicians involve patients in the decision making process at the consultation level.

Based on the psychometric characteristics reported, we were satisfied that the scale could be used to provide a score for the competence framework we had defined as “shared decision making”. Although there is little overall variance between practitioners, there is considerable variability within practitioners, as shown by the differing quartile ranges around their mean scores (fig 2). Some clinicians have a narrower range of scores than others. This may indicate that these clinicians are able to modify their involvement levels across different consultations and to adapt it to the preferred roles of patients in these interactions. This is, however, a conjecture that needs further investigation.

The content validity of the instrument was based on formulating the items from the existing literature, using the results of a series of studies designed to understand how patient involvement can best be achieved in professional practice, followed by subsequent development using an iterative design and assessment cycle. The results with the instrument in this sample of consultations indicate that low levels of involvement in shared decision making are achieved by GPs and that paternalism is the typical “modus operandi” in routine consultations. These practitioners volunteered to take part in a research study on communication skills, and represent those with a high level of confidence in their skills who were aware that we were recording their consultations. Results from other practitioners are likely to be at least on a par or, most likely, lower.

The results indicate that the OPTION instrument achieves acceptable levels of measurement reliability for use in research settings. By focusing on a specific dimension this scale seems to have acceptable levels of reliability compared with similar measures [38,39]. Construct validity was supported by a correlation between involvement scores and patient age and the existence of clinical equipoise in the consultation (although the sample was limited); both hypotheses are supported by previous findings. The lack of correlation between involvement scores and sex of the practitioner or success at the MRCGP was not unexpected, given the weak evidence for these hypotheses.

Strengths and weaknesses of the study

The strength of this study lies in the method of instrument development and a rigorous application of scale development procedures [40]. Some weaknesses were however noted during the study. Most consultations in general practice contain more than one problem solving issue and it is impractical to apply the OPTION instrument to every single presenting problem. Raters are therefore required to agree an index problem. Guidance is given for this issue in a revised manual. In summary, the problem is chosen for which the prime attention is given during the consultation or for which the clinician achieves the greatest involvement score, as the aim is to score demonstrated ability not to calculate involvement across all possible decisions.

Secondly, parent and child consultations required additional guidelines (advising that the interaction between the clinician and the adult was assessed), and the raters had to judge which was the main patient participant where teenagers were being consulted.

It was not possible to estimate concurrent validity (correlation of the measure with some other scale of the concept or trait to be assessed) as there was neither a “gold standard” nor a comparable instrument available. Correlation with patient opinions about their preferred and achieved involvement levels will be reported in further studies from trials conducted in parallel with this validation study [41].

Psychometric assessment also revealed areas where further instrument refinement is necessary. Item 1 may need to be conceptualised as a “gateway” item in which the assessment of involvement in decision making cannot continue if no agreed problem can be identified. Although item 5 has a relatively high kappa score, the response rate was skewed and the factor loading is low. The item is retained, however, as it asks about a feature (use of risk communications tools) that is known not to occur in current service settings. As interventions to change this situation are being introduced, however, the results are likely to change with time as decision aids are introduced into clinical settings [42]. Item 9 questions whether clinicians “provide opportunities for the patient to ask questions” but it has low kappa scores and a factor loading score below 0.2. This item needs modification and further testing to overcome the variation in scoring judgement. There is also a need to consider changing the scale from one that measures magnitude rather than attitude.

Implications for research and formative skill development

OPTION scores for these routine consultations taken from general practice in a UK setting are low. For some items almost no responses were registered, for example, there was 99.7% disagreement with item 5 which asked if the clinician “checks the patient’s preferred information format”. Further research work in this area will involve presenting information in different formats and it is known that, when practitioners develop the skills of involving patients, there is a tendency for a pendulum effect. Retaining these items and others that reveal skewed or “floor” scores should enhance the ability of the instrument to register change.

The OPTION scale can therefore be used to determine the extent to which clinicians involve patients in clinical decisions. It should be noted that the results show that some practitioners have a wider scatter of scores than others. This result is congruent with the theoretical stance that practitioners should be flexible in their consulting style and adapt to the nature of the problem and the patient preference for participation in clinical decisions, although we cannot be certain that this has occurred. It is noteworthy, however, that these OPTION scores are low and it is anticipated that higher scores will be evident after periods of skill development. The instrument should be used to determine scores at a group level (mean scores) or at consultation levels and not to provide a definitive OPTION score that is taken to be characteristic of that practitioner’s ability, unless attention is given to case mix, sample size, and confidence interval estimation. The responsiveness of the instrument to change (increased levels of patient involvement in decision making after skill development) will be validated in further evaluations. It should be emphasised that this tool is designed as an evaluation of a consultation process. It does not measure patient’s preferred role, their contribution to the consultation interaction (also important), or their perceived levels of involvement or

satisfaction. Without this measure of communication process we believe that a vital piece of the presumed linkage between patient involvement and improved outcomes in health care is missing.

Implications for practice

In the face of the widespread acceptance that patient centred-ness is a fundamental goal in clinical practice [43] and that sharing decisions is one of the key components of this approach, the result of this study confirms that the practice of GPs, as represented by this sample (who are an “above average” sample in terms of MRCGP membership and willingness to participate in this type of research), lies far away from espoused models in books and communication skills courses [44, 45] and, indeed, the wishes of certain patients.⁴⁶ Do data from service contexts challenge these espoused models? Are the ideals of patient centred-ness and involvement in decision making completely unrealistic for day to day service contexts? Given that clinicians are consistently positive about the principles of patient centred-ness and patient participation in decision making processes, perhaps the issue of skill development is only a small obstacle and the structural constraints, particularly the lack of time and readily accessible and relevant information about the harms and benefits of health-care interventions, are the true limiting factors. These practitioners volunteered to have their consultations studied but, even so, the results reveal a very limited degree of patient participation. This study, among many others [36, 47–49], provides additional evidence for the assertion that successful patient participation demands more time than is currently allocated. Perhaps these results also lend support to others for the need to harness technologies such as decision aids [42] so that consultations have firmer foundations for partnerships.

References

- Schneider CE. *The practice of autonomy: patients, doctors, and medical decisions*. New York: Oxford University Press, 1998.
- Gwyn R, Elwyn G. When is a shared decision not (quite) a shared decision? Negotiating preferences in a general practice encounter. *Soc Sci Med* 1999; 49: 437–47.
- Edwards A, Elwyn G, eds. *Evidence based patient choice: inevitable or impossible*. Oxford: Oxford University Press, 2001.
- Bekker H, Thornton JG, Airey CM, et al. *Informed decision making: an annotated bibliography and systematic review*. *Health Technology Assess* 1999;3(1).
- Elwyn G, Edwards A, Mowle S, et al. Measuring the involvement of patients in shared decision making: a systematic review of instruments. *Pat Educ Couns* 2001; 43: 5–22.
- Makoul G. *Perpetuating passivity: a study of physician-patient communication and decision-making*. Evanston, IL: Northwestern University, 1992.
- Stewart M, Brown JJ, Donner A, et al. *The impact of patient-centred care on patient outcomes in family practice*. Final report. London, Ontario: Centre for Studies in Family Medicine, 1995.
- van Thiel J, Kraan HF, van der Vleuten CPM. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht history-taking and advice checklist. *Med Educ* 1991; 25: 224–9.
- Marvel MK, Schilling R. Levels of physician involvement with patients and their families: a model for teaching and research. *J Fam Pract* 1994;39:535–44.
- Kurtz SM, Silverman JD. *The Calgary-Cambridge Referenced Observation Guides: an aid to defining the curriculum and organising the teaching in communication training programmes*. *Med Educ* 1996; 30: 83–9.
- Roter DL. *The Roter method of interaction process analysis*. Baltimore: The John Hopkins University, Department of Health Policy and Management, 1991.
- Benbassat J, Pilpel D, Tidhar M. Patients' preferences for participation in clinical decision-making: a review of published surveys. *Behav Med* 1998; 24: 81–8.
- Guadagnoli E, Ward P. Patient participation in decision-making. *Soc Sci Med* 1998; 47: 329–39.
- Coulter A. Partnerships with patients: the pros and cons of shared clinical decision-making. *J Health Serv Res Policy* 1997; 2: 112–21.

15. Stewart M. Studies of health outcomes and patient-centred communication. In: Stewart M, Brown JB, Weston WW, et al, eds. *Patient-centred medicine*. Sage, CA: Thousand Oaks, 1995.
16. Jenkins L, Britten N, Barber N, et al. Improving doctor-patient communication about drugs. Final Report to the Department of Health. London: GKT Concordance Unit, 2001.
17. Edwards A, Elwyn G, Smith C. Consumers' views of quality: identifying the consultation outcomes of importance to consumers, and their relevance to 'shared decision-making' approaches. *Health Expect* 2001; 4: 151–61.
18. Charles C, Gafni A, Whelan T. Decision making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Soc Sci Med* 1999; 49: 651–61.
19. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). *Soc Sci Med* 1997; 44: 681–92.
20. Charles C, Redko C, Whelan T, et al. Doing nothing is no choice: lay constructions of treatment decision-making among women with early-stage breast cancer. *Sociology of Health and Illness* 1998; 20: 71–95.
21. Elwyn G, Edwards A, Kinnorsley P. Shared decision making: the neglected second half of the consultation. *Br J Gen Pract* 1999; 49: 477–82.
22. Elwyn G, Edwards A, Gwyn R, et al. Towards a feasible model for shared decision-making: a focus group study with general practice registrars. *BMJ* 1999; 319: 753–7.
23. Elwyn G, Edwards A, Wensing M, et al. Shared decision-making observed: visual displays of communication sequence and patterns. *J Eval Clin Pract* 2001; 7: 211–21.
24. Elwyn G, Gwyn R, Edwards AGK, et al. Is a 'shared decision' feasible in a consultation for a viral upper respiratory tract infection: assessing the influence of patient expectations for antibiotics using discourse analysis. *Health Expect* 1999; 2: 105–17.
25. Elwyn G, Edwards A, Kinnorsley P, et al. Shared decision making and the concept of equipoise: defining the competences of involving patients in healthcare choices. *Br J Gen Pract* 2000; 50: 892–9.
26. Streiner N, Norman GR. *Health measurement scales: a practical guide to their development and use*. 2nd ed. Oxford: Oxford University Press, 1995.
27. Braddock CH, Fihn SD, Levinson W, et al. How doctors and patients discuss routine clinical decisions: informed decision making in the outpatient setting. *J Gen Intern Med* 1997; 12: 339–45.
28. Edwards A, Elwyn G. The use of shared decision making skills and risk communication tools: a primary care trial. London: Department of Health (Health in Partnership Programme), 1999.
29. Shavelson RJ, Webb NM. *Generalisability theory: a primer*. London: Sage, 1991.
30. Cronbach LJ, Gleser GC, Nanda H, et al. The dependability of behavioural measurements: theory of generalisability for scores and profiles. New York: Wiley, 1972.
31. Kerry SM, Bland JM. Statistics notes: the intraclass correlation coefficient in cluster randomisation. *BMJ* 1998; 316: 1455–60.
32. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
33. Skelton JR, Hobbs FDR. Descriptive study of cooperative language in primary care consultations by male and female doctors. *BMJ* 1999; 318: 576–9.
34. Hall JA, Irish JT, Roter DL, et al. Gender in medical encounters: an analysis of physician and patient communication in a primary care setting. *Health Psychol* 1994;13: 384–91.
35. Coates J. *Women, men and language*. London: Longman, 1993.
36. Champion P, Tate P, Foulkes J, et al. Patient-centredness is rare among candidates in the MRCGP video examination: analysis of large cohort. *BMJ* 2002; 325: 691-2.
37. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall, 1991.
38. Mead N, Bower P. Measuring patient-centredness: a comparison of three observation-based instruments. *Patient Educ Couns* 2000; 39: 71–80.
39. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000; 51: 1087–110.
40. Boon H, Stewart M. Patient-physician communication assessment instruments: 1986 to 1996 in review. *Patient Educ Couns* 1998; 35: 161–76.
41. Edwards A, Elwyn G, Hood K, et al. The development of COMRADE: a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. *Pat Educ Couns* 2003; 50: 311-22.
42. O'Connor AM, Rostom A, Fiset V, et al. Decision aids for patients facing health treatment or screening decisions: systematic review. *BMJ* 1999; 319: 731–4.
43. Stewart M. Towards a global definition of patient centred care. *BMJ* 2001; 322: 444–5.
44. Stewart M, Brown JB, Weston WW, et al. *Patient centred medicine: transforming the clinical method*. Thousand Oaks, CA: Sage Publications, 1995.
45. Kurtz S, Silverman J, Draper J. *Teaching and learning communication skills in medicine*. Abingdon: Radcliffe Medical Press, 1998.

46. Little P, Everitt H, Williamson I, et al. Preferences of patients for patient centred approach to consultation in primary care: observational study. *BMJ* 2001; 322: 468–72.
47. Howie JGR, Heaney DJ, Maxwell M. Measuring quality in general practice: a pilot study of a needs, process and outcome measure. London: Royal College of General Practitioners, 1997.
48. Barry CA, Bradley CP, Britten N, et al. Patients' unvoiced agendas in general practice consultations: qualitative study. *BMJ* 2000; 320: 1246–50.
49. Stevenson FA, Barry CA, Britten N, et al. Doctor-patient communication about drugs: the evidence for shared decision making. *Soc Sci Med* 2000; 50: 829–40.

4.2. The OPTION scale: measuring the extent that clinicians involve patients in decision making tasks

Glyn Elwyn, Hayley Hutchings, Adrian Edwards, Frances Rapport, Michel Wensing, Wai Yee Cheung, Richard Grol. *Health Expectations* 2005; 8: 34-42.

Permission for re-publication granted by Blackwell Publishing 22/12/04.

Introduction

Although there is an increasing call for clinicians to involve patients in decision making about healthcare interventions, it is by no means clear how this communication strategy should be achieved and measured. A review of the literature revealed that there was a lack of validated tools for this purpose [1]. The overall intent of the communication strategy, often known as 'shared decision making' [2], is for the patient to be made aware that there are important decisions to be considered, and that these decisions cannot be taken by the clinician alone. Patients will need to be encouraged and assisted to take on the task of understanding the relevant information and to share their values and views with clinicians. The principles of shared decision making have been described and reviewed [3]. The skills (competences) have been elucidated and discussed [4].

It should be acknowledged that a debate exists as to whether 'shared decision making' can or should be undertaken in all clinical interactions. One school of thought argues that shared decision making should only be implemented where there is a 'genuine' choice operating, and refers to a classification of clinical situations into those that should follow a *standard* of care, a *guideline* or to situations where *options* may be legitimately discussed with patients. This view argues that the measurement of shared decision making could only take place where the clinical situation warrants the provision of options. In contrast to this view, the OPTION scale takes a different conceptual stance and accepts that it is difficult, if not impossible, to judge where and when patients will want to partake in decisions. Some will not want to learn of options even when there is evidence of definite uncertainty. What is important conceptually is that the possibility of shared decision making is enhanced, and for this to happen, clinicians have to *involve* patients in the process of understanding the nature of the problem, understanding that there are uncertainties and different likelihoods of harms and benefits and finally that the patient can, if they wish, influence the decision itself. The OPTION scale regards these steps as constituting the process of involving patients in decision making. It does not purport to measure shared decision making, although the authors are convinced that a shared decisions making process could take place in consultations with low OPTION scores.

A previous publication described how a scale that measured the extent to which clinicians involved patients in decision making was developed and validated (Observing patient involvement, abbreviated to be known as the OPTION scale) [5]. This work described a theoretical construct, provided details of scale design stages and item formulation [6]. It was recognised however that the instrument had aspects that required further attention [5]. In particular, difficulties had been encountered with the scaling characteristics

and with the phraseology and order of some items. An attitudinal scale had been used and this had led to an overuse of a mid point that represented uncertainty about the evaluation. The aim of this paper is to address the aspects that required attention and to report an improved, definitive instrument that could be used to measure the extent to which clinicians involve patients in decision making processes — a tool that is available to researchers and educators for use in research and skill development.

Method

The study examined the psychometric characteristics of a revised OPTION scale applied to a sample of audio taped consultations, collected from the routine clinics of 21 general practitioners. Approval to conduct the work was obtained from the Gwent Local Research Ethics Committee. To conduct the validation study, the revised scale was used by two non-clinical lecturers in social sciences who remained independent of each other and of the research team, and who were trained in its use. These two raters had used the previous OPTION scale version to assess the same set of consultation recordings [5]. The recordings were taken from the recruitment phase of a research study, specifically a trial of shared decision making and risk communication. As part of the recruitment process to the study, general practitioners in Gwent, South Wales, were asked to audiotape consecutive consultations during a routine consulting session in general practice. To be eligible, the practitioners had to have worked in general practice for at least one year and less than ten years.

The potential sample of 104 practitioners in 49 practices were invited to participate. As far as we are aware, these volunteer practitioners were naïve to the concepts that we were measuring and had not been exposed to any training or educational interventions that could have influenced their proficiency in this area. Details of the recruitment process have been published elsewhere [5]. In order to test inter-rater reliability the two raters rated all the consultations independently and a random sample of 21 consultations (1 per clinician) was selected for repeated ratings by the two raters in order to examine intra-rater reliability.

The revised scale

The main difference between the previously published scale and the scale used in this study was the manner in which the observable competences were rated. Raters were previously asked to consider whether they agreed or disagreed (on a 5 point scale ranging from strongly agree to strongly disagree) whether certain skills (such as ways to deal with a problem, preferred patient approach to receiving information, etc.) were observed during the consultations. It was reported that this led to difficulties, whereby raters used the score of '3' to indicate indecision regarding the competences. It was apparent that this uncertainty had multiple origins; there was uncertainty whether the competence was present or not, and also uncertainty as to whether the activity was undertaken with a high degree of skill or not. We suspect that this tendency to use the mid point for both types of uncertainty inflated the OPTION scores as it is a recognised problem of attitudinal scales [7]. In order to address this problem, the revised scale was designed to measure the *magnitude* of skill rather than the *attitude* towards the described competences. The score '0' was allocated to the situation where the competency described was not observed, other

scores (1 to 4) were allocated to increasing levels of achievement for the described competence (see Figure 1). Minor alterations to wording and some re-sequencing of items was also performed

In order to address this problem, the revised scale was designed to measure the extent of skill rather than agreement about observable behaviours towards the described competences, a shift from an *attitudinal* to a *magnitude* based scale. A detailed manual was developed that defined the levels of observed behaviour that should be scored. The score '0' was allocated to the situation where the competency described was not observed, other scores (1 to 4) were allocated to increasing levels of achievement for the described competence (see Figure 1). Minor alterations to wording and some re-sequencing of items were also performed.

Figure 1. The revised scale scoring guidance (excerpt)

Scale Score	Definition
0	The behaviour is not observed
1	A minimal attempt is made to exhibit the behaviour
2	The clinician asks the patient about their preferred way of receiving information to assist decision
3	The behaviour is exhibited to a good standard
4	The behaviour is observed and executed to a high standard

Analysis

The data were analysed by studying the responses to each item, the scale reliability was assessed with inter-item and item-total correlations, summarised by Cronbach's alpha. Rater agreement was assessed by using Cohen's kappa (on a five point scale and a two point binary scale) and inter- and intra-rater agreement was assessed using intra-class correlation coefficient (ICC). ICC scores above 0.40, 0.60 and 0.80 were interpreted as fair, moderate and substantial agreement respectively [8]. Exploratory factor analysis (oblique rotation taking eigen values of 1.1) was used to determine factor loadings [9]. Further assessment using a forced one-factor analysis (oblique rotation) was also performed.

Results

Sample

Out of the potential sample pool of 104 practitioners, 21 general practitioners (from separate practices) provided a tape of a routine clinic prior to receiving detailed information about the proposed research trial [10]. These practitioners represented a slightly younger group than the sampling frame: average age 38 years, the male to female ratio was identical to the sampling frame (38% female); 76% (16/21) of the general practitioners recruited had been successful in the membership examination of the Royal College of General Practitioners, compared with an overall membership level of 54% in the sampling frame, and could have been expected perhaps to show skills associated with a more recently trained and motivated group. Of the 242 consecutive patients approached in all practices, 12 (5%) declined to have the consultation recorded (the maximum refusal in any one practice was 3 patients in a series of 15 patients).

The remaining 230 consultations were assessed and after removing consultations where there were technical recording problems, 186 consultations were available for analysis (average of 8.8 consultations per practitioner). There was no age and sex difference between the consultations excluded because of

poor recordings compared to those included for analysis. One practitioner recorded 5 consultations but the majority recorded 8 or more consultations. Consultations with women were twice as frequent in the sample and 66% of the patients seen were between 30 and 70 years of age. The demographic and clinical characteristics of the recorded consultations have been reported previously [5]; in summary, there were 126 female and 60 male patients; the patient ages were between 4 months to 83 years, the mean duration of consultations was of 8.2 minutes and the majority of consultations dealt with respiratory, musculoskeletal, dermatological and psychological problems, the typical spectrum of general practice.

Rating patterns

All items with the exception of items 8 and 9 showed a predominance of zero scores (see Table 1 for baseline scores). Items 8 and 9 showed the most variation across the scale, although results were still confined to the lower scores. None of the items had a score that exceed '2' for any consultation. There were no missing scores. Compared to previously published results [5], these results show a greater level of scoring consistency and a lower level of missing values.

Table 1. Option item response (%) for two observers, Cohen's kappa and intra-class correlation (ICC)

OPTION Item		0	1	2	3	4	Kappa*	ICC
1. The clinician draws attention to an identified problem as one that requires a decision making process	Observer 1	96.4	3.6	0	0	0	0.53 (0.52)	0.33
	Observer 2	93.6	4.5	1.8	0	0		
2. The clinician states that there is more than one way to deal with the identified problem ('equipoise')	Observer 1	91.8	6.4	1.8	0	0	0.88 (0.88)	0.93
	Observer 2	91.8	6.4	1.8	0	0		
3. The clinician assesses the patient's preferred approach to receiving information to assist decision making	Observer 1	99.1	0.9	0	0	0	0.98 (0.98)	0.98
	Observer 2	100	0	0	0	0		
4. The clinician lists 'options', which can include the choice of 'no action'	Observer 1	90	9.1	0.9	0	0	0.64 (0.76)	0.77
	Observer 2	84.5	11.8	3.6	0	0		
5. The clinician explains the pros and cons of options to the patient (taking 'no action' is an option)	Observer 1	95.5	4.5	0	0	0	0.70 (0.70)	0.70
	Observer 2	91.8	8.2	0	0	0		
6. The clinician explores the patient's expectations (or ideas) about how the problem(s) are to be managed	Observer 1	95.5	4.5	0	0	0	0.56 (0.56)	0.56
	Observer 2	98.2	1.8	0	0	0		
7. The clinician explores the patient's concerns (fears) about how problem(s) are to be managed	Observer 1	95.5	3.6	0.9	0	0	0.51 (0.59)	0.61
	Observer 2	92.7	7.3	0	0	0		
8. The clinician checks that the patient has understood the information	Observer 1	91.8	8.2	0	0	0	0.08 (0.10)	0.11
	Observer 2	36.4	61.8	1.8	0	0		
9. The clinician offers the patient explicit opportunities to ask questions during the decision making process	Observer 1	56.4	43.6	0	0	0	0.45 (0.48)	0.48
	Observer 2	60	38.2	1.8	0	0		
10. The clinician elicits the patient's preferred level of involvement in decision-making	Observer 1	99.1	0.9	0	0	0	0.98 (0.98)	0.98
	Observer 2	100	0	0	0	0		
11. The clinician indicates the need for a decision making (or deferring) stage	Observer 1	100	0	0	0	0	0.84 (0.84)	0.84
	Observer 2	91.8	8.2	0	0	0		
12. The clinician indicates the need to review the decision (or deferment)	Observer 1	79.1	20	0.9	0	0	0.67 (0.67)	0.61
	Observer 2	79.1	20.9	0	0	0		

*Kappa: scores are for five-point scale agreement. Scores in brackets are for agreement across binary scale points (no involvement / involvement)

Factor analysis

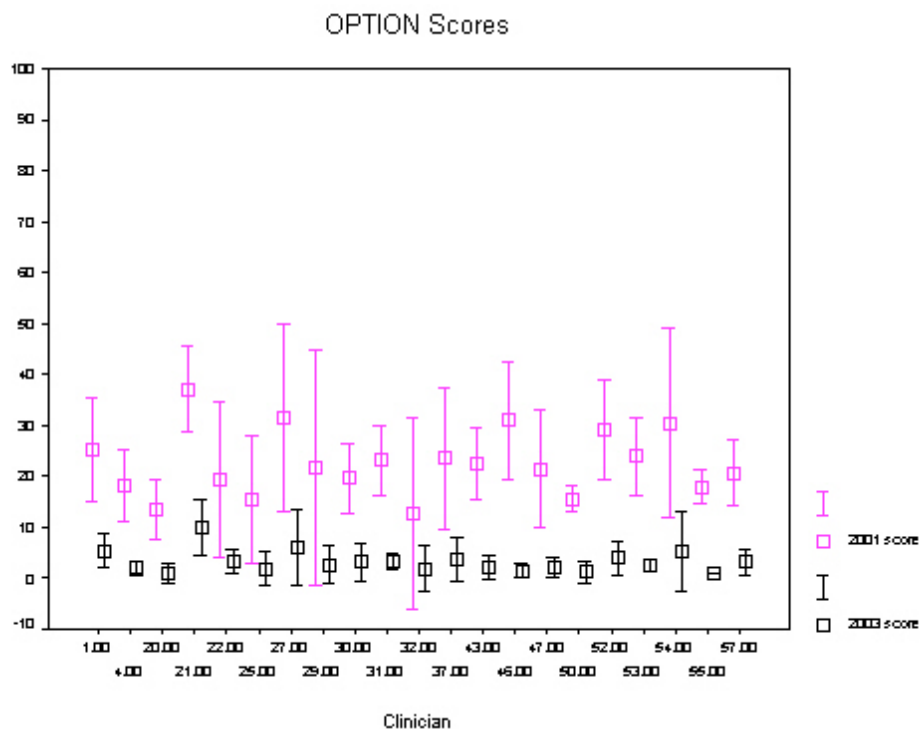
Exploratory factor analysis (oblique rotation taking eigen values of 1.1) was used to determine factor loadings [9].

The scree plot showed the presence of two factors, the distribution of questions to each factor revealed a Cronbach's alpha of 0.80 for the first factor and a Cronbach's alpha of 0.44 for the second factor. Cronbach's alpha based on all 12 items was 0.68. The 12-item single factor solution explained 28% of the variability. Items 1 and 8 had low inter-rater reliability. The removal of these items produced a small improvement in scale performance (32% of variability explained). We could not determine a pattern to item loadings on the two factors and because the scale was developed to match an agreed set of competences [4], a judgement was made that it was best to use a single factor to retain the 12-item scale in its entirety.

Inter-rater reliability

With the exception of item 8, the five point scale Cohen's kappa scores ranged from 0.45 to 0.98, indicating acceptable inter-rater agreement after correcting for chance [8]. Aggregating the rating scores to produce a two-point binary scale showed similar kappa values (see Table 1). The inter-rater intra-class correlation coefficient (ICC) for the total OPTION score was 0.77, with values ranging from 0.11 to 0.98 for the individual items, which again showed good levels of agreement, with the exception of item 9. They compare well with inter-rater reliability from the 2001 data, but the overall scores for the OPTION scale are lower than previously as a result of changing from the attitude to magnitude measures (see Table 2 and Figure 2 for results and graphical representation). For all 12 items, the mean Cohen kappa score was 0.66, indicating acceptable inter-rater agreement for this type of instrument, after correcting for chance [11]. Item 8 had the lowest kappa and ICC scores. Item 9 also showed a difference in inter-rater reliability. Items 8 and 9, therefore, need attention to definition in a revised manual and a focused calibration of raters. No changes have been made to item wording however. Compared to previously published results [5], these scores indicate a marginal improvement in reliability, identical kappa scores and an increase in inter-rater ICC from 0.62 to 0.77.

Figure 2. OPTION Scores (2001 and 2003 Scale)



Test-retest reliability

The test-retest data were based on a reduced sample of one consultation per practitioner (n=21) where the raters scored the consultations for a second occasion. The re-test results confirmed a predominance of low scores. The inter-rater ICC for the total OPTION score was 0.53.

This ICC was slightly lower than the score of 0.77 achieved for the ratings achieved for the full set of consultations using this scale. At individual item level, Cohen's kappa measured on a five-point scale ranged from -0.05 to 1 indicating good agreement for some items, but poor agreement for others. ICC scores at individual item for intra-rater reliability ranged from -0.05 to 0.66 for observer 1, and 0 to 0.66 for observer 2. Despite having only weak ICCs for individual items, the ICC for the total OPTION scores showed a good level of agreement for both observer 1 (0.82) and observer 2 (0.65). As with the initial ratings, test-retest data confirm that the OPTION instrument cannot be regarded as reliable at the individual item level (see ICC scores in Table X but when summed, OPTION ICC scores indicate substantial agreement according to suggested interpretations [8]).

Table 2. Mean transformed OPTION 2003 and 2001 scores (0=min, 100=max) for each practitioner

GP number	Mean OPTION 2003 score (95% CI)	Mean OPTION 2001 score (95% CI)	Number of consultations
1	5.42 (2.07, 8.76)	25.42 (15.25, 35.58)	5
4	2.08 (0.73, 3.44)	18.23 (11.26, 25.20)	4
20	1.04 (-0.87, 2.96)	13.54 (7.64, 19.44)	5
21	10.00 (4.47, 15.53)	37.08 (28.67, 45.49)	5
22	3.33 (0.85, 5.82)	19.32 (3.91, 34.73)	5
25	1.88 (-1.35, 5.10)	15.42 (2.93, 27.90)	5
27	6.04 (-1.41, 13.49)	31.36 (13.01, 49.71)	5
29	2.71 (-1.02, 6.43)	21.67 (-1.35, 44.68)	5
30	3.13 (-0.42, 6.67)	19.58 (12.59, 26.57)	5
31	3.33 (1.65, 5.02)	23.15 (16.14, 30.11)	5
32	1.88 (-2.64, 6.39)	12.71 (-6.22, 31.64)	5
37	3.75 (-0.50, 8.00)	23.54 (9.53, 37.55)	5
42	2.29 (0.61, 3.98)	Not rated	5
43	2.29 (-0.02, 4.61)	22.50 (15.50, 29.50)	5
46	1.46 (-0.02, 2.93)	31.00 (19.41, 42.60)	5
47	2.08 (0.04, 4.13)	21.50 (9.98, 33.01)	5
50	1.25 (-0.88, 3.38)	15.63 (13.21, 18.04)	5
52	3.96 (0.62, 7.31)	29.17 (19.23, 39.10)	5
53	2.50 (1.34, 3.66)	23.96 (16.36, 31.56)	5
54	5.21 (-2.61, 13.02)	30.42 (11.73, 49.10)	5
55	1.04 (1.04, 1.04)	17.92 (14.70, 21.14)	5
57	3.13 (0.71, 5.54)	20.63 (14.13, 27.12)	5

Discussion

Principal Findings

When compared to the first version of the OPTION scale, the revised scale, when applied to the same data set, has resulted in a small improvement in the scale's reliability and to lower scores for the levels of involvement achieved by the practitioners. Factor analysis confirms that it is acceptable to regard the scale as a single construct. We conclude therefore that OPTION is sufficiently reliable to be used for formal

assessment at the level of the whole instrument (all 12 items) but that the scale can be used with more flexibility in professional education settings.

Although there is moderate variability when raters are assessed on an item by item basis, the agreements between raters at the level of the overall OPTION score is high (ICC scores for total OPTION score was 0.77), a level that is acceptable for the evaluation of a set of consultations per practitioner (e.g. between 5 and 10 consultations), where aggregate scores would be used for determining overall competence. It should be noted that the scale is only scored at, or close to, the 'floor' level. The vast majority of scores given were '0' or '1'. Although it could be argued that this is a weakness of the scale in that it does not display sensitivity to existing practice, we prefer to argue, based on parallel work in discourse analysis, that the scale reflects the reality of current routine practice. Clinical encounters do not typically contain examples of practitioners displaying the skills of shared decision making. We therefore consider the scale both reliable and valid to use in research contexts. We have demonstrated the scale's ability to show an increase in skill levels, as exhibited in a clinical controlled trial reported elsewhere [10]. We briefly considered the characteristics of a binary scale, where a 'yes' or 'no' answer could be given to observable competences for each item (Table 1). A simplified binary scale could be used, for instance, in educational contexts, where scores could be given out of a maximum of 12, indicating success or otherwise at addressing each specific competence, with the possibility of rapid feedback in a formative learning context.

Strengths and weaknesses

The major strength of this study is that for the first time a scale has been developed which can be used as a valid and reliable measure for shared decision making in clinical encounters. It builds on a rigorous development path, addresses the weaknesses of the previous instrument and replicates the assessment of a set of consultations taken from day to day practice using the same calibrated raters. The double rating is recommended if the instrument is to be used for research at the level of individual consultations.

If however the aim is to achieve an overall 'involvement' score at the practitioner level, provided there are at least 5 consultations available per clinician, we consider the scale reliable enough if single ratings are undertaken. Changing the scale from an assessment of attitude to an assessment of magnitude (observable skills) has added to our confidence in the assessment of skill attainment at the practitioner level.

The major weakness is the recognised clustering of low scores: scores of '0' or '1' predominate. It could be argued that the scale has been poorly designed, that it has no ability to discriminate between practitioners who are working in routine contexts to the best of their ability. It could also be argued that, on the basis of this data, there could be no confidence that the scale could be sensitive to skills at a higher level or to increases in skill attainment. However, data from a controlled trial (using the previously published scale) demonstrates that the scale is capable of detecting changing skill levels [10]. Although in ideal circumstances, we would re-establish these finding with a larger sample and in different settings, we conclude that these results are based on the accurate use of a valid and reliable scale in routine clinical

contexts - that practitioners with no previous training in shared decision making achieve very low levels of patient involvement in decision making.

Findings in context

A systematic review of instruments that aimed to measure shared decision making did not reveal the presence of reliable valid instruments in this field [1]. It is therefore difficult to undertake comparative studies in order to establish concurrent and criterion validity. It may be possible in the future to undertake comparative studies with instruments that have been recently published that aim to undertake similar evaluations [12]. It will be important to continue to study the characteristics of the OPTION tool in different clinical contexts, as it is known that different clinical specialities have a different cultural ethos. If the OPTION tool was used to evaluate the consultations of clinical geneticists, for example, it may be that their scores would be different, given that consultations in genetics are of significantly greater duration and that there is an increased awareness of the need to involve patients in the generic assessment process [13].

It is important to bear in mind that the results were observed in general practitioners working in day-to-day settings, having had no special exposure to the concepts of shared decision making and that the duration of the consultation were, on average, 8 minutes. It is becoming widely recognised that achieving greater levels of patient involvement requires additional time [10]. In other words, current practice militates against involving patients in decisions, and even with the most highly skilled communicators in primary care, we would be surprised if substantially higher levels of patient involvement could be achieved without at least a 50% increase in consultations duration (i.e. 10-12 minutes per consultations). It is important, however, to concede that time, although necessary, is not sufficient on its own [14]. There is evidence from the membership examination of the Royal College of General Practitioners that confirms this view. Given the opportunity to provide their 'best' consultations for assessment and where it is known that shared decision making will be among the most valued criteria, doctors fail to demonstrate these competences [15]. Such findings are confirmed when more qualitative methods are used to analyse medical practice [16-18].

Implications

The implications of this work can be summarised as follows. This instrument provides a means of assessing to what degree clinicians involve patients in decision making. To meet the increasing call to inform patients about the harms as well as the benefits of interventions, the lack of the necessary communication skills and the barriers to their development and implementation need to be addressed at policy levels. How best to develop these skills is a matter for educationalists at both undergraduate and postgraduate levels, whilst the uncertainties about the outcomes of involving patients will require further investigation [19]. As part of this work, the OPTION tool provides a means of assessing progress.

References

1. Elwyn G, Edwards A, Mowle S, Wensing M, Wilkinson C, Kinnersley P, Grol R. Measuring the involvement of patients in shared decision making: a systematic review of instruments. *Pat Educ Couns* 2001; 43: 5-22.
2. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). *Soc Sci Med* 1997; 44: 681-92.
3. Charles C, Gafni A, Whelan T. Decision making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Soc Sci Med* 1999; 49: 651-61.
4. Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision making and the concept of equipoise: defining the competences of involving patients in healthcare choices. *Br J Gen Pract* 2000; 50: 892-9.
5. Elwyn G, Edwards A, Wensing M, Hood K, Robling M, Atwell C, Grol R. Shared decision making: developing the OPTION scale for measuring patient involvement. *Qual Saf Health Care* 2003; 12: 93-99.
6. Streiner N, Norman GR. *Health Measurement Scales: A practical guide to their development and use*. 2nd ed. Oxford: Oxford University Press, 1995.
7. Galtung J. *Theory and methods of social research*. Oslo: Universitetsforlaget, 1967.
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74.
9. Jolliffe IT, Morgan BJT. Principal component analysis and exploratory factor analysis. *Stat Methods Med Res* 1992; 1: 69-75.
10. Elwyn G, Edwards A, Hood K, Robling M, Atwell C, Russell I, Wensing M, Grol R. Achieving involvement: process outcomes from a cluster randomised controlled trial of shared decision making skill development and use of risk communication aids in general practice. *Fam Pract* 2004; 21: 337-46.
11. Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall, 1991.
12. Braddock CH, Edwards KA, Hasenberg MH, Laidley TL, Levinson W. Informed decision making in outpatient setting: time to get back to basics. *JAMA* 1999; 282: 2313-20.
13. Elwyn G, Gray J, Clarke A. Shared decision making and non-directiveness in genetic counselling. *Journal of Medical Genetics* 2000; 37: 135-8.
14. Freeman GK, Horder JP, Howie JGR, Hungin AP, Hill AP, Shah NC, Wilson A. Evolving general practice consultation in Britain: issues of length and context. *BMJ* 2002; 324: 880-2.
15. Champion P, Foulkes J, Neighbour R, Tate P. Patient centredness in the MRCGP video examination: analysis of large cohort. *BMJ* 2002; 325: 691-2.
16. Gwyn R, Elwyn G. When is a shared decision not (quite) a shared decision? Negotiating preferences in a general practice encounter. *Soc Sci Med* 1999; 49: 437-447.
17. Gwyn R, Elwyn G, Edwards A, Mooney A. The problematic of decision-sharing: deconstructing 'cholesterol' in a clinical encounter. *Health Expect* 2003; 6: 242-54.
18. Elwyn G, Gwyn R, Edwards AGK, Grol R. Is a 'shared decision' feasible in a consultation for a viral upper respiratory tract infection: assessing the influence of patient expectations for antibiotics using discourse analysis. *Health Expect* 1999; 2: 105-17.
19. Edwards A, Elwyn G, Atwell C, Hood K, Houston H, Kinnersley P. Patient-based outcome results from a cluster randomised trial of shared decision making skill development and use of risk communication aids in general practice. *Fam Pract* 2004; 21: 347-54.

4.3. Achieving involvement: process outcomes from a cluster randomized trial of shared decision making skill development and use of risk communication aids in general practice

Glyn Elwyn, Adrian Edwards, Kerry Hood, Mike Robling, Christine Atwell, Ian Russell, Michel Wensing, Richard Grol and the Study Steering Group. *Family Practice* 2004; 21(4): 337-46.

Permission for re-publication granted by Oxford University Press 9/12/04.

Abstract

Background: A consulting method known as 'shared decision making' (SDM) has been described and operationalised in terms of several 'competences'. One of these competences concerns the discussion of the risks and benefits of treatment or care options – 'risk communication'. Few data exist on clinicians' ability to acquire skills and implement the competences of shared decision making or risk communication in consultations with patients.

Aims: To evaluate the effects of skill development workshops for shared decision making and the use of risk communication aids on the process of consultations.

Design: Cluster randomised trial with crossover.

Participants: 20 recently qualified general practitioners; 747 patients with known atrial fibrillation, prostatism, menorrhagia or menopausal symptoms, invited to a consultation to review their condition or treatments. Half the consultations were randomly selected for audio-taping, of which 352 patients attended and were successfully audiotaped.

Setting: Urban and rural general practices in Gwent, South Wales.

Interventions: After baseline, participating doctors were randomised to receive training in (1) shared decision making skills or (2) the use of simple risk communication aids, using simulated patients. The alternative training was then provided for the final study phase. Patients were randomly allocated to a consultation during baseline or Intervention 1 (SDM or risk communication aids) or Intervention 2 phases. A randomly selected half of the consultations were audio taped from each phase.

Outcomes: raters (independent, trained and blinded to study phase) assessed the audiotapes using a validated scale to assess levels of patient involvement (OPTION: observing patient involvement), and to analyse the nature of risk information discussed. Clinicians completed questionnaires after each consultation assessing perceived clinician-patient agreement and level of patient involvement in decisions.

Principal analysis: multi-level modelling with the OPTION score as the dependent variable, and rater, consultation and clinician levels of data, standardised by rater within clinician.

Results: following each of the interventions, the clinicians significantly increased their involvement of patients in decision making (OPTION score increased by 10.6 following risk communication training (95% CI 7.9-13.3; $p < 0.001$) and by 12.9 after shared decision making skill development (95% CI 10-15.8, $p < 0.001$), a moderate effect size. The level of involvement achieved by the risk communication aids was significantly increased by the subsequent introduction of the skill development workshops (7.7 increase in OPTION score, 95% CI 3.4 - 12; $p < 0.001$). The alternative sequence (skills followed by risk communication aids) did not achieve this effect. The use of most risk information formats increased after the provision of specific risk communication aids ($p < 0.001$). Clinicians using the risk communication tools

perceived significantly higher patient and clinician agreement on treatment ($p < 0.001$), patient satisfaction with information ($p < 0.01$), clinician satisfaction with decision ($p < 0.01$) and general overall satisfaction with the consultation ($p < 0.001$) than those who were exposed to SDM skill development workshops.

Conclusions: These clinicians were able to acquire the skills to implement shared decision making competences and to use risk communication aids. Each intervention provided independent effects. Further progress towards greater patient involvement in health care decision making is possible and skill development in this area should be incorporated into postgraduate professional development programmes.

Introduction

Shared decision making (SDM), a short hand term used for the process of involving patients in clinical decisions, has been the subject of debate in the recent literature on interpersonal communication in health care [1,2]. Although the principles of the method are described [3], the competences outlined [4,5] and a measure proposed [6], there is uncertainty about the proposal [7,8] and some doubt that the concept can be applied in clinical settings [9,10]. Although there are feasibility studies reported [11] there is little agreement, or evidence, about how to implement SDM. In essence, the ethos of shared decision making is one where professionals should work to define problems with sufficient clarity and openness so that patients can comprehend the options and uncertainties that surround most decisions in health care and therefore appreciate that choices have to be made between competing options. A part of this process includes the discussion of the harms and benefits relating to these options⁴, a stage of 'risk communication'. 'Shared decision making' is only one stage of the consultation, which must be viewed holistically, but it is a stage that has until recently received relatively attention, particularly in professional skill development [12-14]. Such consideration should also not be divorced from evidence about what patients value most from consultations, including doctors' ability to listen, empathise and support, be interested in the effect of problems on patients' lives and to maintain a positive approach [15,16].

In particular, few studies have investigated shared decision making and risk communication to great depth in actual clinical settings [12-14]. Hulsman and Bensing's review noted the inadequacy of research designs reported to date [17]. Some evidence is emerging about the benefits of shared decision making in practice [14], but most reviews have been more general in their scope, addressing the effects of 'patient-centred approaches' [18-20]. The potential problems of taking a 'shared approach' to decision making have also been highlighted [9 21]. As yet for 'shared decision making' the unanswered questions are twofold: firstly, it is not clear whether clinicians working in everyday settings *can* improve their skills to involve patients in decisions. Secondly, we need to know whether the model for developing shared decision making should be based on information provision or interpersonal communication skill development.

This study sought to operationalise *shared decision making* and *risk communication aids* as specific and comparable interventions to clinicians. These were not viewed as mutually exclusive interventions. The aim was to investigate whether each intervention, first separately and then combined with the other, would increase clinicians' ability to involve patients in decision making processes. If this was found to be the case, the secondary aim was to evaluate whether acquiring skills in SDM should happen before or after exposure to the risk communication aids.

Methods

Setting and subjects

The study took place in Gwent, South Wales, which includes urban, suburban and rural areas that cross a range of socio-economic levels [22], and had Local Research Ethics Committee approval. General practitioners were approached for consent to participate in the study. Eligibility for the trial was restricted to those who had been principals in general practice for between one and ten years at the start. Participants also had to be tape-recorded for a surgery session (approximately 10 patients). Both this, and the former criterion (through exposure to recent training methods) were adopted to achieve a group of participants who were likely to be familiar with the data collection methods intended. Only one practitioner per practice would be recruited. Three practitioners were excluded from the potential sample pool of 104 practitioners because of prior exposure to the training content during developmental work conducted earlier in other areas. In all, 101 practitioners from 49 practices were approached, initially by letter (followed by telephone contact) and asked to participate.

Patients were approached by the practices for consent to participate in the study if they were known to have one of the four following conditions: non-valvular atrial fibrillation; prostatism; menorrhagia or menopausal symptoms. These conditions were selected because they are characterised by having more than one treatment option available, and about which the professionals as a group are likely to have no clear preference —‘equipoise’ [23]. This is not the only context for shared decision making in practice but it is a context in which the opportunity for shared decision making is greater than if equipoise is not present [4]. Although patients with these problems can be found in all practices, the incidence of patients presenting with these problems *de novo* is low. The trial therefore proactively identified previous attenders, avoiding the problems associated with clinician-based patient recruitment. These patients were identified from READ codes on electronic practice databases by staff from the practices using a standard protocol, assisted by a research officer (CA). If READ code searches were insufficient, then specific treatments were identified (such as digoxin for atrial fibrillation). For this stage, patient records were then checked to ensure eligibility. Inclusion and exclusion criteria are listed in Table 1. Patient recruitment and allocation are described in more detail in the accompanying paper [24].

Table 1. Patient inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
General	Consultation within previous 3 years. Identification by computer codes or repeat medications.	> 75 years of age.
Specific	Non-valvular atrial fibrillation	Valvular heart disease
	Lower urinary tract obstruction: diagnosis of ‘prostatism’ (range of synonyms).	History of prostate cancer. Previous prostate surgery. Raised prostatic specific antigen level.
	Diagnosis of ‘menorrhagia’ (range of synonyms)	Hysterectomy
	Diagnosis of ‘menopause’ related problems (range of synonyms). Ages 45 – 55 current or previous users of HRT.	

The interventions

Separate interventions to enhance clinician skills in either shared decision making or the use of risk communication aids were devised (see Box 1 & related figures) and piloted [11,25]; they were provided to the clinicians before each active trial phase (see below).

Box 1. Trial interventions

The theoretical basis underlying both these skill developments was the extended model of interpersonal interaction outlined by Hargie [31] which proposes that as skill 'perceptions' are translated into motor responses (speech and actions), a sequence of feedback loops ensure that performance is modified. Repeated cycles lead to fluent skill acquisition. This is a widely accepted basis for communication skill development in clinical contexts [5,32]. The simulated patients involved in each intervention were non-medical people with considerable experience as simulators from under-graduate medical education.

Shared decision making

Practitioners randomised to this intervention attended two workshops where a standardised and previously piloted skill development process was undertaken using presentations, discussions and participation in consultations with simulated patients (facilitated by GE). The background literature on shared decision making [3,33,34] was outlined and participants asked to debate its relevance to clinical practice. Building on the work of Boyatzis [35] the skills (competences) of shared decision making [4] were described and demonstrated using simulated consultations. This provided opportunities for all the participants to comment on the method, using an observational competence checklist. Simulated patients were also encouraged to comment. Participants were asked to consult with the simulated patients using pre-prepared scenarios involving the study conditions. At the second workshop, participants were asked to consider the competences in more depth. By the end of the workshop, all participants had conducted and received feedback from at least one consultation with a simulated patient. No data relating to treatments for the study conditions were provided to participating clinicians receiving this training before risk communication training.

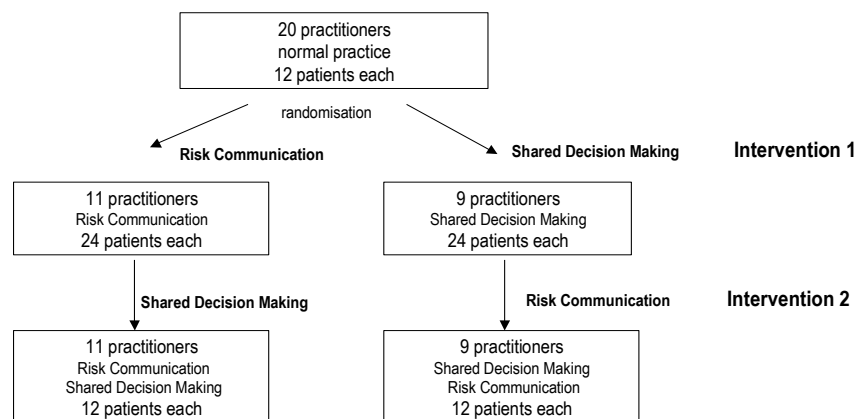
Risk communication aids

Similar workshops addressing risk communication were facilitated by AE. The risk communication aids consisted of tabulated data and visual displays of risk estimates (histograms and bar charts) for the four study conditions. The risk data were based on systematic reviews [36-40] (and other publications, see below) and presented as the best evidence available at the time of the trial. *Risk communication* was defined in order to distinguish it from other terms in common use, such as *risk management* and *risk analysis*. Recent research in this area was summarised, including evidence of effectiveness of interventions [13] and the expressed needs of clinicians [41] and patients [42]. The participants were provided with treatment outcome information for the study conditions in the following range of formats: summary statements, bar charts, numerical statistical information and source publications. Examples are shown in the following figures. Participants were told about the derivation of the risk aids, advised on how they might be incorporated in consultation discussions and then asked to use them in simulated patient consultations. Participants were directed first to conduct a simulated consultation using only numerical data, followed by a consultation in which they would use only graphical data displays. They then conducted further consultations in which they chose the most appropriate format to use with the individual patients – the recommended strategy to apply in the trial setting itself. The consultations were conducted in pairs, where colleagues alternated between clinician and observer roles. This was repeated until each participant had received feedback after conducting two or three consultations using the risk communication aids across a range of conditions. A plenary group discussion, which included the patient simulators, allowed the group to share learning points and consider the application of the materials in clinical practice.

Design

A cluster randomised design with crossover was chosen (Figure 1). The 'cluster' was each participating clinician and the patients who consulted them. This method offered the greatest potential to gain understanding about the effects of the training interventions alone and in combination, and whether the sequence of skill acquisition was important. The latter evidence would be important in applying the results of this trial to continuing professional development initiatives.

Figure 1. Randomised trial design with crossover



- NB Numbers of practitioners in each arm show actual numbers randomised; numbers of patients show maximum numbers intended in each phase of trial design;
- In each phase, random allocation to normal surgery time or protected 'research clinic' for audio-taping;
- Patient-based outcomes assessed at exit and at 1 month.

All patients consulted the participating clinician only once for the study. This 'review type' consultation was randomly allocated to one of the three trial phases (baseline, risk communication or shared decision making only, or combined interventions). In each of these phases, the consultations were also randomly allocated to occur alongside routine surgery consultations or in a 'research clinic'. This clinic was also held at the participating clinician's own health centre but was protected so that other interruptions and time constraints could be minimised and the required assessments (audio-taping and patient interviews) could be achieved. If required, locum payments were provided. Although the clinicians were advised to conduct consultations in their usual way, the encounters in the research clinics had the potential to take approximately 50% more time than usual. By evaluating the effect of the interventions, and the added element of protected time, the trial sought to be explanatory in nature [26]. Patients unable to attend an allocated appointment were offered an alternative appointment (or excluded from the study and replaced by other patients if alternatives were inconvenient or if they did not attend).

All randomisations were undertaken by random number generation and allocations by the trial statistician (KH) were concealed from those implementing the interventions or assessments. Both clinicians and patients were informed that the trial was investigating 'communication skills' but were otherwise 'blinded' to the decision making or risk communication focus of the interventions.

Outcome measurement

This paper reports the evaluation of process in the consultations, and the results of questionnaires completed by clinicians after each study consultation. Patient outcomes are reported elsewhere [24]. The principal process measure was the OPTION (observing patient involvement) scale [6]. Audiotapes were recorded on the randomly selected half of study consultations that were scheduled for the 'research clinics'.

Two post-doctoral social science researchers were trained to score audiotapes tapes independently using the OPTION scale and manual. This scale requires raters to listen to audiotapes of consultations, and make judgements regarding the demonstration of 12 behaviours that cover the competences of shared decision making [4]. All consultation recordings were intended to be rated by two raters and ratings were undertaken blind to study group allocation of clinicians or patients. Inter-rater differences for OPTION scores for consultations (clusters in clinicians) were assessed. An example of the first items in the OPTION scale is shown in Box 2.

Box 2. The first items in the OPTION scale to assess the extent of patient involvement in decision making.

- 1 The clinician *draws attention to* an identified problem as one that requires a decision making process.
- 2 The clinician *states* that there is more than one way to deal with the identified problem ('equipoise').
- 3 The clinician *assesses* the patient's preferred approach to receiving information to assist decision making (e.g. discussion, reading printed material, assessing graphical data, using videotapes or other media).
- 4 The clinician *lists* 'options', which can include the choice of 'no action'.
- 5 The clinician *explains* the pros and cons of options to the patient (taking 'no action' is an option).
- 6 The clinician *explores* the patient's *expectations* (or ideas) about how the problem(s) are to be managed.

(the rater scores each item on a five point scale from 'strongly agree' to 'strongly disagree' according to a manual, available from the authors)

In addition observations were made about the nature of risk information used in these research clinic consultations. The raters recorded whether or not the discourse included (at any time in the consultation) general risk statements, descriptives (such as 'likely', 'rare'), comparisons to everyday risks, numerical information, absolute risk information, relative risk information, numbers-needed-to-treat formats or visual depiction of risks. These different methods of communicating risk were all covered in the intervention workshops, with emphasis on clinicians using what they felt was most appropriate with individual patients.

The participating clinicians also completed questionnaires after every study consultation ('research clinic' and routine surgery study consultations). These addressed clinician perceptions of the level of clinician-patient agreement, the patient's satisfaction with information provided and the clinician's satisfaction with the decision and overall consultation.

Sample size

Sample size calculations were based on patient outcomes and on providing 80% power (5% significance levels) to detect a change of 20 percentage points in either direction from a baseline of 50% for binary variables, and full details are in the accompanying paper [24]. The sample size requirements were for 240 in each Phase, 960 for the whole trial. Each doctor would consult with 48 patients (see Figure 1). However, only half the consultations were to be audiotaped (6 patients per clinician at baseline, 12 at the RC or SDM phase and 6 at the final phase). Complete patient attendance would therefore have resulted in 480 consultations being available for tape-recording and OPTION ratings. This gave 80% power to detect (at 5% significance levels) a change of 6.6 points on the OPTION scale (from earlier work this amounts to a 'moderate' effect size [6,27]).

Data processing and analysis

Data from the questionnaires were entered into SPSS files (error rate of items <0.5% in a 10% random sample). Further data cleaning included correction of out-of-range values. Scores were calculated from the mean of other valid items for up to 2 missing item ratings out of the 12.

The primary outcome (OPTION) was assessed with multilevel modelling with MLwiN software [28,29]. Explanatory variables were entered as fixed effects in a regression model with the OPTION score as the dependent variable. The improvement of fit from allowing the effect to be random was also assessed. A three level model was fitted with rater at level 1, consultation at level 2 and clinician at level 3 to the data standardised by rater within clinician (i.e. adjusting for consistent scoring differences between raters). The model assessed the extent to which variability in OPTION could be explained by the rater, the medical condition (3 dummy variables) and the intervention (risk communication, shared decision making as single interventions, the effect of combining the interventions and the effect of the order in which they were received). The extent to which the interventions had a random or fixed effect was also assessed. The model also assessed the extent to which variability in OPTION could be explained by practitioner variables (age, gender and membership of the Royal College of General Practitioners), patient variables (age, condition). The *sequence* effect was entered as the last explanatory variable. Improvements in the model were assessed using reductions in the log likelihood. The mean OPTION scores (and 95% confidence intervals) for each general practitioner during the phases of the trial were calculated.

Secondary outcomes were assessed at the cluster level using t-tests weighted for cluster size. These outcomes included the types of risk information used in consultations, and clinician perceptions of doctor-patient agreement, patient's satisfaction with information and clinician satisfaction ratings.

Results

Recruitment and participant flow

Twenty-one out of 49 practices (42.8%) had a doctor who agreed to participate, have a surgery session audio taped, and had sufficient practice computerisation to enable identification of a patient sample. One doctor dropped out after baseline phase. The remaining practitioners, 12 men and 8 women, had an average age of 38. These characteristics did not differ from the eligible sample frame (101 practitioners with average age of 41, 62 % male). 80% of the participating doctors had Membership of the Royal College of General Practitioners, compared with an 54% in the overall sample approached.

A total of 2585 patients was approached and 1135 (43.9%) consented to take part in the trial. The full flow chart for patients in the study, and analysis of bias from non-consent or participation is given in the accompanying paper [24]. Patients were randomly selected from those consenting to be invited to attend study appointments with their doctor. The mean age of patients recruited was as follows: those with prostatic symptoms 63, atrial fibrillation 65, menorrhagia 45 and hormone replacement therapy 56 years. Audiotape recording was scheduled for 480 of the study consultations. Some patients could not attend and were replaced by other patients. In all, 566 patients were invited; 391 attended. Audiotaping was successful for 352 consultations (92%; see Figure 2 and Table 2 on web-version). There was no difference

in the mean consultation lengths at baseline, phase 1 and phase 2 (overall consultation mean duration was 12.5 minutes).

Figure 2. Sequencing of interventions and participation of practitioners and patients in trial of shared decision making skill development

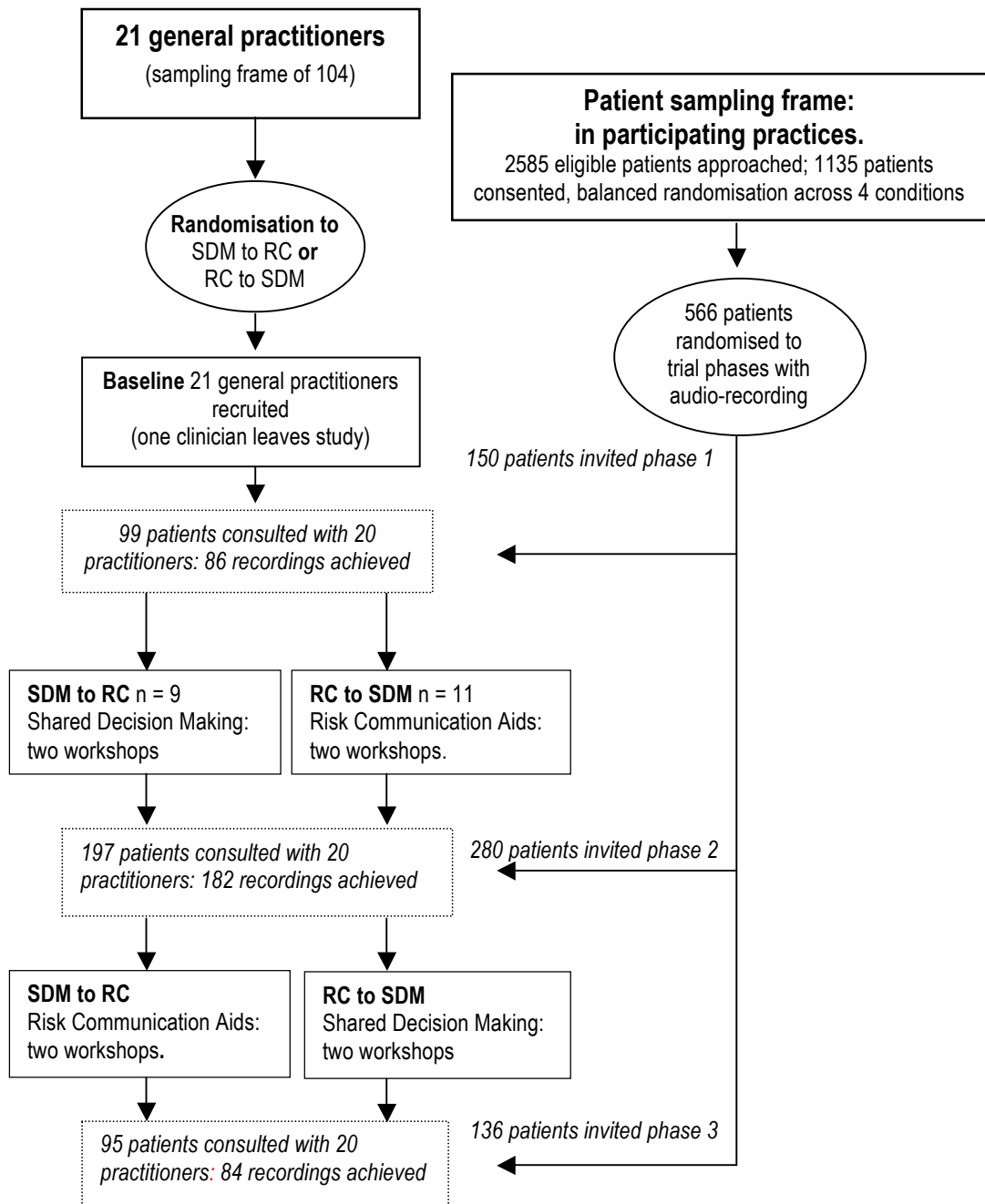


Table 2. OPTION Ratings available for analysis in each trial phase

Trial phase	Ratings	HRT	Prostate	Menorrhagia	Atrial fibrillation
Baseline	85 pairs, 1 single	32	22	14 (1) *	17
SDM to RC	80 pairs, 3 single	28	22 (2)	14 (1)	16
RC to SDM	99 pairs	37	28	22	12
Combined	82 pairs, 2 single	27 (1)	27	13	15 (1)
Totals	698	124 (1)	99 (2)	63 (2)	60 (1)

* Numbers in brackets indicate the number of consultations that had single ratings.

OPTION (patient involvement in decision making) scores

A total of 698 completed OPTION ratings entered the multilevel analysis. 643 ratings had corresponding complete data on patient age and condition, but these showed no association with OPTION score, so the full sample was used. No clinician or ‘condition’ variables resulted in a significant improvement in the model, so these were not adjusted for in the model. No explanatory variables were shown to be better represented by a random effect. The intra-cluster correlation coefficient (ICC) was 0.18 in the baseline phase. The final model is given in Table 3. Mean scores for each participant are shown in Figure 3 and Table 4. Consistent inter-rater differences for OPTION scores were identified. Analyses therefore used standardised scores within clinicians and the results remained significant after this adjustment.

Table 3. Effects of risk communication (RC) or shared decision making (SDM) training or both on OPTION scores*.

Intervention	Coefficient	95% C.I	Significance
RC training	10.6	(7.9, 13.3)	<0.001
SDM training	12.9	(10.0, 15.8)	<0.001
Both	-10.6 ^a	(-15.1, -6.1)	<0.001
Order of training	7.7 ^b	(3.4, 12.0)	<0.001

*Regression coefficients from 3-level model.

^a indicates that OPTION scores in the combined intervention group were 10.6 less than the sum of the individual effects

^b indicates that doctors receiving RC before SDM had OPTION scores 7.7 higher than the reverse order

Chi-square statistic = 152.2; d.f. = 4; p<0.001; Final -2 log likelihood = 5200.8

Figure 3. Mean OPTION Scores (standardised) across trial phases for the participating doctors, numbered 1-20

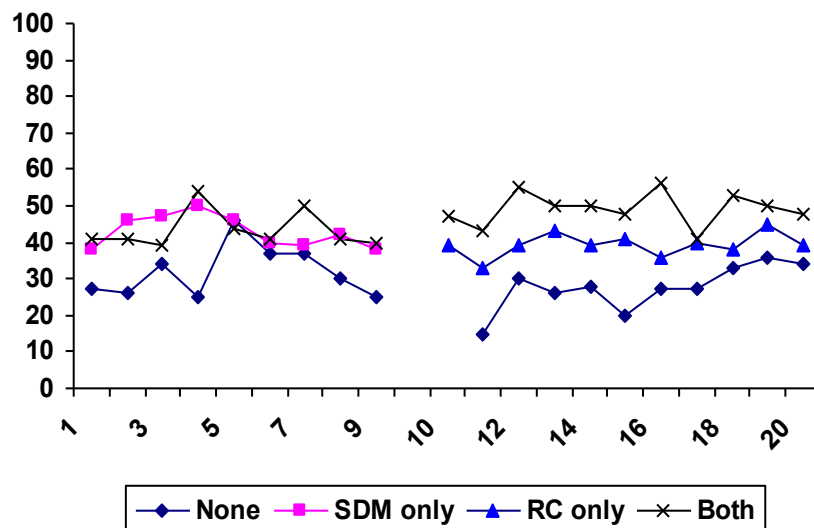


Table 4. OPTION Scores* for clinicians across study phases
For inclusion on web-version, not paper journal publication

Clinician Groups	Baseline	Shared decision making	Risk Tools	After both interventions
SDM then RC Group Mean	32 (13.8)	43 (13.6)		43 (11.0)
RC then SDM Group Mean	27 (14.7)		39 (11.8)	50 (11.8)
Overall Mean	29 (14.5)			47 (12.1)

(*OPTION scores are standardised by rater; values are mean with SD inflated for clustering).

Table 3 shows that there were significant increases in patient involvement as a result of both the risk communication (OPTION score increased by 10.6, 95% CI 7.9-13.3; $p < 0.001$) and shared decision making skill development workshops (12.9 increase, 95% CI 10-15.8, $p < 0.001$). A significant addition in patient involvement as a result of receiving both interventions was only seen in those who received risk communication intervention first then shared decision making skill development second ('RC then SDM': 7.7 increase, 95% CI 3.4 - 12; $p < 0.001$).

Use of risk information in consultations

The use of risk information across the phases of the study is shown in Table 5. The content of risk information discussed in consultations changed dramatically in association with the risk communication intervention, including the provision of the packs with information in different formats. The changes were noted across all categories (formats) of risk information, but were most evident in the use of visual formats (charts).

Taking the group randomised to receive risk communication first, most categories of risk information increased between baseline and Phase 1 with generally little further change into Phase 2 ($p < 0.001$). The exceptions ('comparisons drawn' and using individualised risk estimates) did not increase, offering validation of the rating exercise, as these were not made available to doctors as part of the training.

In the doctors receiving shared decision making training first, few changes were seen between baseline and phase 1. After receiving the risk communication intervention, large changes in the use of risk information were seen, mirroring those in the other group after their risk communication intervention. Again, the 'comparisons drawn' category showed no change, and the individualised risk estimates remained at low levels.

In summary, clinicians increased the proportion of consultations in which they used several categories of risk information after the risk communication training intervention. This was not shown when the groups received shared decision making training. Statistically significant changes were shown for all of the numerical information items, and were largest for the visual formats of risk information.

Table 5. Consultations where types of risk information were used between clinician groups

	Risk (n = 11 doctors)			SDM (n = 9 doctors)		
	Baseline (%)	Phase 1 ^a (%)	Phase 2 ^a (%)	Baseline (%)	Phase 1 ^b (%)	Phase 2 ^b (%)
General statements**	40	65	68	53	48	71
Likely / rare *	15	35	62	24	25	65
Comparisons drawn	0	1	0	3	1	0
Individualised risk	0	4	2	11	2	18
Numerical information c**	4	78	78	21	20	85
Absolute risk**	2	34	26	8	7	32
Relative risk**	2	32	18	5	7	26
Number needed to treat**	0	12	6	3	2	3
Visual format**	6	77	72	5	2	88

(results are weighted mean percentage of consultations in which the types of information were used and t-tests between clinician groups (weighted for cluster size, 18 degrees of freedom, critical t = 2.1) *p<0.01 **p<0.001

^a Phase 1 for these clinicians comprised only training for risk communication aids; Phase 2 comprised shared decision making skills in addition to the previous risk communication aids.

^b Phase 1 for these clinicians comprised only training for shared decision making skills; Phase 2 comprised risk communication aids in addition to the previous risk communication aids.

^c Numerical information includes absolute and relative risk and number needed to treat formats.

Clinicians' views on the consultations

Clinicians showed significant differences between the RC and SDM arms (see Table 6 on web-version). Doctors receiving the risk communication tools and training first perceived significantly higher doctor-patient agreement on treatment ($p<0.001$), patient satisfaction with information ($p<0.01$), doctor satisfaction with decision ($p<0.01$) and general overall satisfaction ($p<0.001$) with the consultation than those who were exposed to shared decision making training. The latter group of doctors showed lower scores after the interventions. The differences were largely maintained in the second intervention phase. That is, even when provided with the risk communication training and tools, the group of doctors who had received SDM training first still reported lower levels of satisfaction, agreement etc. *Vice versa*, doctors who had received risk communication training first maintained their higher levels of satisfactions and agreement, even when later given the shared decision making training which appeared less beneficial (to doctors) in the first phase.

Table 6. Doctors' assessment of consultations across the phases of the study

	Risk (n = 11 doctors)			SDM (n = 9 doctors)		
	Baseline (%)	Phase 1 ^a (%)	Phase 2 ^a (%)	Baseline (%)	Phase 1 ^b (%)	Phase 2 ^b (%)
Patient and clinician agreed on treatment	71	77**	79	66	58**	54
Patient satisfied with information	68	66*	80	70	54*	63
Clinician satisfied with decision	80	79*	83	73	67*	66
Overall satisfaction	74	63**	79	73	55**	52
Mean number of consultations per GP	9.5	17.5	7.5	9.1	16.0	7.9
Total consultations	105	193	82	82	144	71

(percentages are weighted for cluster size) *p<0.01 weighted t-test between groups (Phase 1); **p<0.001

^a Phase 1 for these clinicians comprised only training for risk communication aids; Phase 2 comprised shared decision making skills in addition to the previous risk communication aids.

^b Phase 1 for these clinicians comprised only training for shared decision making skills; Phase 2 comprised risk communication aids in addition to the previous risk communication aids.

Discussion

Principal findings

The clinicians in this study demonstrated greater involvement of patients in treatment decision making after skill development workshops. They also integrated the risk communication aids by using the graphical illustrations in scheduled review consultations with real patients after the training intervention. Both interventions independently increased patient involvement levels. It appears that the most effective way to increase clinicians' abilities to involve patients is to familiarise them with detailed information before discussing skill development techniques. Patient involvement in decisions did not vary significantly between the four clinical conditions or for patient age differences. The clinicians also *perceived* higher levels of patient involvement, improved clinician-patient agreement and increased patient satisfaction with the consultations. These findings should be taken in conjunction with those of the accompanying paper on patient-based outcomes [24].

Strengths and weaknesses of the study

The strength of this study is that it operationalised 'shared decision making' by using a rigorously developed competences framework [4,5] and that it used a specific and validated scale (OPTION) as its principal outcome measure [6]. This study focuses on practitioners in settings that were as near as possible to normal service conditions given the data collection requirements. At patient level the comparison was randomised between all comparison groups. At clinician level the cross-over design allowed assessment of the differential effects of acquiring the skills and competences for shared decision making and risk communication in different sequences.

Weaknesses of the study include a degree of bias from participant representativeness and data captured. It is likely that the general practitioners were motivated clinicians with higher than average confidence and ability regarding interpersonal communication skills. A pure control group was not included in the design, due to the issues of control-group disengagement from a complex study protocol occurring over a number of months [30]. Thus some of the changes demonstrated may represent Hawthorne effects. Patient non-attendance was more likely among younger and female patients. Nearly 30% fewer consultations were audiotaped than intended. Thus bias or reduced power in the study to detect intervention effects were possible.

The 'review' nature of the study consultations may have hindered discussions about treatment choices although the trial maximised its chances of showing effects by using conditions characterised by clinician 'equipose' [4,31]. The assessment of risk information used in the audio-taped consultations was restricted to frequency of using information and did not address skills with which the information was discussed.

The multi-level modelling could not test a period (referred to in this design as 'Phase') effect separately. Period and carry-over effects are confounded here so we have tested for carry-over (interaction between the two types of training), attributing changes between period ('Phases') 1 and 2 as the combined effect of the training. This is justified on the grounds that the randomisation of patients to period minimised case-

mix differences in the periods. The potential reduction in power from assessing both sequence effects of training and the individual effects via the cross-over design is acknowledged.

Interpretation in context of setting and intervention

There are very few data examining the effect of providing interventions such as these with real patients in UK general practice. The results demonstrate that the interventions led to significant changes in the process of consultations, as detected by the OPTION scale and analysis of the types of risk information discussed, and to changes in clinicians' perceptions of the consultation. These findings must be taken in the context of a *lack* of change in patient-based outcomes [24].

The increased involvement levels after both interventions, with an additive effect when shared decision making skills are provided *after* the introduction of risk communication aids, indicates that skill development and information provision can lead to changes in the clinical interaction. One explanation for the sequential enhancement of OPTION scores is that clinicians may have used the risk communication aids to reinforce professional decisions after the risk communication workshops [25]. Doing this in the consultation led clinicians also to demonstrate a number of the competences of SDM -and hence increased scores on OPTION. However, some of the (perhaps finer) competences were only covered by the SDM intervention- such as portraying equipoise and inviting patient choice. However the converse was not found: the doctors who received the SDM skills intervention first then showed no further enhancement in OPTION scores after risk communication training. This statistically significant difference is likely to be important in practice. This finding informs how continuing professional development initiatives can take forward efforts to enhance patient involvement in health care decisions. Risk communication training and provision of (decision) aids should precede skill development in SDM. This is not only because the levels of involvement were highest among doctors following this sequence, but also because doctors' confidence and satisfaction with the process and decisions appeared greater with this sequence.

Further work is needed in this area. In particular this should assess the sustainability of shared decision making skills and evaluate whether clinicians can apply these skills to 'new' decisions over an increased range of conditions. This trial was also explanatory in nature [26]. It used an idealised setting and subjects to evaluate the efficacy of the interventions. The participating clinicians reported that the interventions were highly acceptable. Having demonstrated that it is possible to achieve change in consultation process, a pragmatic trial is desirable to assess the generalisability to routine service settings. Only in the light of such further evidence would it be valid to suggest that all general practitioners should receive specific training on risk communication and shared decision making. We recognise that in reality education and training should be part of a process rather than the single interventions evaluated here, and might be expected to achieve greater or longer-lasting effects as a result. In the meantime the efficacy of these interventions to affect consultation processes indicate the potential for improvements in clinical practice and the value of further postgraduate communication skills development for those keen to progress in this area. The efficacy of these interventions to date also indicates the value of developing decision support materials and technologies that offer easily utilised information for both clinicians and patients. Further

research should address practitioner skills for using such decision aids and the discussion of the information in ways that are flexible to meet the needs of individual patients.

Conclusion

The skills and steps of shared decision making (including the need for risk communication) are valid competences that can be operationalised. Experienced and motivated practitioners are willing to develop skills in this area. As a result of the trial interventions, the clinicians significantly increased their involvement of patients in decision making, as measured by the OPTION scale, and the frequency of discussing risk information with patients. Greatest effects were evident among the doctors receiving shared decision making training after risk communication training and provision of decision aids.

This could be the basis for continuing professional development initiatives in this area.

References

1. Charles C, Gafni A, Whelan T. Decision making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Soc Sci Med* 1999; 49: 651-661.
2. Elwyn G, Charles C. Shared Decision Making: inevitable or impossible? In: Edwards A, Elwyn G, editors. *Evidence Based Patient Choice*. Oxford: Oxford University Press, 2001.
3. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). *Soc Sci Med* 1997; 44: 681-92.
4. Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision making and the concept of equipoise: defining the competences of involving patients in healthcare choices. *Br J Gen Pract* 2000; 50: 892-899.
5. Towle A, Godolphin W. Framework for teaching and learning informed shared decision making. *BMJ* 1999; 319: 766-71.
6. Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Fleeting glimpses: measuring shared decision making in primary care using the OPTION instrument. *Qual Saf Health Care* 2003;12:93-99.
7. McKinstry B. Do patients wish to be involved in decision making in the consultation? A cross sectional survey with video vignettes. *BMJ* 2000; 321: 867-871.
8. Stevenson FA, Barry CA, Britten N, Barber N, Bradley CP. Doctor-patient communication about drugs: the evidence for shared decision making. *Soc Sci Med* 2000; 50: 829-40.
9. Auerbach SM. Should patients have control over their own health care?: empirical evidence and research issues. *Ann Behav Med* 2000; 22: 246-59.
10. Gwyn R, Elwyn G. When is a shared decision not (quite) a shared decision? Negotiating preferences in a general practice encounter. *Soc Sci Med* 1999; 49: 437-447.
11. Elwyn G, Edwards A, Gwyn R, Grol R. Towards a feasible model for shared decision-making: a focus group study with general practice registrars. *BMJ* 1999; 319: 753-7.
12. Elwyn G, Edwards A, Kinnersley P. Shared decision making: the neglected second half of the consultation. *BJGP* 1999; 49: 477-82.
13. Edwards AGK, Hood K, Matthews EJ, Russell D, Barker J, et al. The effectiveness of one-to-one risk communication interventions in health care: a systematic review. *Med Decis Making* 2000; 20: 290-7.
14. Bekker H, Thornton JG, Airey CM, Connelly JB, Hewison J, Robinson MB, et al. Informed decision making: an annotated bibliography and systematic review. *Health Technol Assess* 1999;3(1).
15. Little P, Everitt H, Williamson I, Warner G, Moore M, Gould C, et al. Observational study of effect of patient centredness and positive approach on outcomes of general practice consultations. *BMJ* 2001; 323: 908-11.
16. Edwards A, Elwyn G, Smith C, Williams S, Thornton H. Consumers' views of quality in the consultation and their relevance to "shared decision making" approaches. *Health Expect* 2001; 4: 151-62.
17. Hulsman RL, Ros WJ, Winnubst JA, Bensing JM. Teaching clinically experienced physicians communication skills. A review of evaluation studies. *Med Educ* 1999; 33: 655-8.
18. Griffin S, Gillard S, Veltman M, Grant J, Stewart M, Kinmonth AL. A systematic review of trials of interventions to improve the interactions between patients and practitioners. The 29th Annual Scientific Meeting of the Association of University Departments of General Practice. Bournemouth: CBC Oxford Ltd on behalf of the AUDGP, 2000.
19. Lewin SA, Skea ZC, Entwistle V, Zwarenstein M, Dick J. Interventions for providers to promote a patient-centred approach in clinical consultations. *The Cochrane Library*. Oxford: Update Software., 2002.
20. Stewart M. Effective Physician Patient Communication and Health Outcomes. *CMAJ* 1995; 152: 1423-33.

21. Schneider CE. The practice of autonomy: patients, doctors, and medical decisions. New York: Oxford University Press, 1998.
22. Public Health Department Gwent Health Authority. Townsend indices of socio-economic status for post code enumeration districts in Gwent, South Wales. Newport: Public Health Department Gwent Health Authority, 1998.
23. Weijer C, Shapiro S, Glass K, Enkin M. Clinical equipoise and not the uncertainty principle is the moral underpinning of the randomised controlled trial. *BMJ* 2000; 321: 756-8.
24. Edwards A, Elwyn G, Atwell C, Hood K, Houston H, Kinnersley P, et al. Patient-based outcome results from a cluster randomised trial of shared decision making skill development and use of risk communication aids in general practice. *Family Practice* 2004; 21: 345-52.
25. Edwards A, Elwyn G, Gwyn R. General practice registrar responses to the use of different risk communication tools: problems and opportunities. *BMJ* 1999; 319: 749-52.
26. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *Journal of Chronic Diseases* 1967;20:637-648.
27. Kazis LE, Anderson JJ, Meenan R. Effect sizes for interpreting changes in health status. *Med Care* 1999; 27: S178-S189.
28. Goldstein H, Rasbash J, Plewis I, Draper D, Browne W, Yang M, et al. A user's guide to MLwiN. London: Institute of Education, 1998.
29. MLwiN [program]. London: Institute of Education, University of London <http://www.ioe.ac.uk/multilevel/index.html>, 2001.
30. Moore H, Summerbell C, Vail A, Greenwood D, Adamson A. The design features and practicalities of conducting a pragmatic cluster randomised trial of obesity management in primary care. *Statistics in Medicine* 2001; 20: 331-40.
31. Edwards A, Elwyn GJ. How should 'effectiveness' of risk communication to aid patients' decisions be judged? A review of the literature. *Med Decis Making* 2000;19: 428-34.
32. Hargie ODW, editor. *The Handbook of Communication Skills*. London: Routledge, 1997.
33. Stewart M, Brown JB, Weston WW, McWinney IR, McWilliam CL, Freeman TR. *Patient Centred Medicine: Transforming the Clinical Method*. Thousand Oaks, CA: Sage Publications, 1995.
34. Guadagnoli E, Ward P. Patient participation in decision-making. *Soc Sci Med* 1998; 47: 329-39.
35. Coulter A. Partnerships with patients: the pros and cons of shared clinical decision making. *Journal of Health Services Research and Policy* 1997; 2: 112-21.
36. Boyatzis RE. *The competent manager. A model for effective performance*. New York: John Wiley and Sons, 1982.
37. Nuffield Institute for Health UoL. Benign Prostatic Hyperplasia. *Effective Health Care Bulletin* 1995;2(2):1-16.
38. Coulter A, Kelland J, Peto V, Rees M. Treating menorrhagia in primary care. *Int J Technol Assess Health Care* 1999; 11: 456-71.
39. Grady D, Rubin SM, Pettiti DB, Fox CS, Black D, Ettinger B, et al. Hormone therapy to prevent disease and prolong life in postmenopausal women. *Ann Intern Med* 1992; 117: 1016-37.
40. Benavente O, Hart R, Koudstaal P, Laupacis A, McBride R. Antiplatelet therapy for preventing stroke in patients with non-valvular atrial fibrillation and no previous history of stroke or transient ischemic attacks. *Update Software: Cochrane Library* 1999; (2):1-12.
41. Benavente O, Hart R, Koudstaal P, Laupacis A, McBride R. Oral anticoagulants for preventing stroke in patients with non-valvular atrial fibrillation and no previous history of stroke or transient ischemic attacks. *Update Software: Cochrane Library* 1999; (2):1-16.
42. Edwards A, Matthews E, Pill RM, Bloor M. Communication about risk: diversity among primary care professionals. *Fam Pract* 1998; 15: 296-300.
43. Edwards A, Elwyn G, Smith C. Consumers' views of quality: identifying the consultation outcomes of importance to consumers, and their relevance to 'shared decision-making' approaches. *Health Expect* 2001; 4: 151-61.

5. OPTION Manual

OPTION

Rater Manual

**Observing patient involvement
Evaluating the extent that clinicians involve patients in decisions**

Glyn Elwyn, Adrian Edwards, Michel Wensing and Richard Grol

2005 version

In collaboration with:

Jill Bourne, Wales
Richard Byng, England
Evelyn Chan, USA
Wai Yee Cheung, Wales
Elizabeth Cox, USA
John Dowell, Scotland
Tina Eriksson, Denmark
Ellen Fischer, USA
Guido Goelen, Brussels
Claudia Goss, Verona
Wolfgang Himmel, Germany
Margaret Holmes Rovner, USA
Kerry Hood, Wales
Amanda Howe, England
Hayley Hutchings, Wales
Christoph Kindler, Switzerland
Alex Krist MD, USA
Wolf Langewitz, Switzerland
Tara Lawn, England
France Légaré, Canada
Cathy Lises, Wales
Young Mi Kim, USA
Matthias Nuebling, Germany
Charlene Pope, USA
Frances Rapport, Wales
Mike Robling, Wales
Angela Towle, Canada
Trudy van der Weijden, the Netherlands
Myrra Vernooij, the Netherlands
Heidi Weber, Switzerland
Marjorie Weiss, England
Clare Wilkinson, Wales
Hub Wollersheim, the Netherlands

Contents

1. General Issues
2. Scale design and item definitions
3. OPTION scale (Research Version)
4. OPTION scale (Educational feedback version)

1. General issues

This scale is designed to measure the extent to which clinicians (medical, nursing or other relevant professional) involves patients in decisions within consultations. The authors recognise that these interactions are complex and involve contributions from parties, patients and clinicians. A consultation where there is clear evidence of participation in the decision will as likely have contributions from patients in equal measure to the contributions of the professional. Nevertheless, when developing the scale we recognised the difficulty of designing a scale that assessed the contribution of both agents simultaneously and decided to concentrate the OPTION scale on assessing the skills exhibited by the clinician. This approach is also justified by the differential power relationship that exists in patient-professional interactions, and that if professionals do not offer opportunities, and promote patient involvement in decision making, then the process is highly unlikely to be observed. Nevertheless, we concede that the patient contribution to the process of participation in decision needs to be evaluated in parallel and are making efforts to support the development of such as scale.

Research education etc

The scale is designed so that it can be applied to all types of consultations, it is therefore intended to give an involvement score for all 'types' of consultations, whether the encounter is one that involves a first consultation about a problem or a review of a previously discussed problem, reassurance, or any of a range of possible categories of encounters. It is recognised however, that patient involvement in decisions is going to be dependent on the *type* of consultation, so it is important to record the overall consultation *type* and to record the *index* clinical condition that the rater uses as the basis for the assessment of involvement in the decision making process. These variables are added to the rating dataset so that the analysis can assess the degree of involvement achieved in differing types of consultations and with respect to the topic or conditions discussed (see Table 1). Often there is more than one problem in the consultation. A practical decision should be taken to score the process for an *index problem*. An index problem is the problem where the highest degree of involvement occurs within the overall consultation, as the aim is to identify the practitioner's *ability* to involve patients. Where there is more than one OPTION rater, they will need to agree the *index problem* for the rating.

Raters should therefore use the total duration of the consultation to score the OPTION scale, recognising that involvement in decision making can be at the level of involvement in problem solving decisions (e.g. there are two possible diagnoses here, here are the options and possible ways of investigating them...), involvement in treatments (there are a range of possible treatments for this problem, let me explain a bit more about them...), or involvement in further management (there are a number of ways in which this problem could be managed, lets consider them in more detail...). If the clinician involves the patient in any of these types of decision, then the OPTION score would be used to assess the proficiency of the skills exhibited. It is recognised that often, some of the skills may have taken place in previous encounters, but this data will be correlated with the consultation type (new, review or composite).

Table 1. Consultation data: index problem and consultation categories

Index problem	In new and review consultations, the index problem will be relatively easy to identify, unless there are multiple unrelated new problems presented for diagnosis (and / or review). Where multiple problems are presented, the index problem should be chosen on the basis of the highest degree of patient involvement in decision making achieved by the clinician. For composite type consultations, the index problem is the one where the clinician provides most evidence of involving the patient in decision making. A broad description of the <i>index problem</i> should be provided, e.g. headache, indigestion, heavy period, depression, chest pain. The terms used should describe the main issue(s) presented or discussed and use terms used in the actual consultations. Diagnostic accuracy is not required unless there is evidence that this has been achieved within the interaction.
Consultation type	
New	The main focus is on the presentation of a new problem for diagnosis or advice. The focus in this type of consultation is on taking the patient's story and taking steps to arrive at a diagnosis, i.e. requesting investigations or arranging follow up.
Review	Main focus is follow-up of a recently presented new problem, e.g. results of investigations are investigated and decisions are made regarding further management or treatment. Patients returning to report ineffective treatment would be categorised as a review consultation, unless they also had new problems to present.
Composite	New problem(s) and action or review(s) of previous problems. e.g. a respiratory infection plus request for repeat medication for high blood pressure.

In addition to data about the rater identity, consultation type and condition considered, information about the patient and the practitioner is collected (age, gender and information about additional postgraduate qualification, such as Membership of a professional college. These variables are part of the OPTION dataset and are used to evaluate construct validity.

Consultations often involve more than two individuals and when obtaining consent a record should be made of the ages and gender of the individuals who are involved in the consultation. A parent presents a child for instance, or two people consult about a mutual concern (husband, wife, mother, daughter and so on). These consultations are often of a complex nature, and the interaction involves many conversations about problems and decisions. In most adult-child consultations, it will be clear that the decision making discourse will occur between the adult and the practitioner, and it is the age and gender of the person engaged in the consultation process that should be recorded on the OPTION scale, although a note of the age of the child can also be made under the index problem. Where an adult accompanies a teenager, the age of the person to engage in the consultation process should be recorded (i.e. a rater judgement). In consultations where more than one adult is present, the rater should indicate which individual takes the primary role in the consultation process and the clinician behaviour should be judged in relation to this interaction.

2. Scale design and item definitions

A five-point scale is used to assess the existence of a communication behaviour (competence). The first point on the scale, namely 0, is used when the behaviour is not observed in the consultation. Details about how each scale point should be given to differing skill levels of behaviours observed are provided in this manual. In general terms, the five levels (0-4) will correspond to the following general outline:

Score	Description
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The behaviour is observed and a minimum skill level achieved.
3	The behaviour is exhibited to a good standard.
4	The behaviour is exhibited to a very high standard.

Raters should use the scale points when a behaviour observed corresponds to the descriptions provided in this manual. A set of calibration audiofiles is available from the OPTION Group for those who want to become OPTION raters.

The Scale Items

Item 1	The clinician <i>draws attention to an identified problem as one that requires a decision making process.</i>
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician draws attention to a problem that requires a decision making process.
3	The behaviour is observed and the clinician puts emphasis on the decision making process required.
4	The behaviour is observed and executed to a high standard.

To embark on a decision making process, there has to be clarity about a specific problem or problems. In order to involve the patient in a decision, it should be clear that a decision making process is taking place. The skill to be observed therefore is the ability to identify, emphasise, draw attention to a problem (e.g. high blood pressure, menopausal symptoms, atrial fibrillation etc), as one where a decision exists about further action, and that it needs to be considered by both clinician and patient. In other words, the patient's attention is focused on the fact that the consultation is one where a decision making process is being considered and that the clinician is going to involve the patient, if they so wish, in considering the problem.

For this behaviour to occur there has to be a degree of agreement about the nature of the problem. The problem need not necessarily be a diagnosis where there are choices between treatments or form of management. It is also possible to share a decision about whether or not to take a test, order an investigation or send off a referral. The item therefore assess the clarity with which the clinician draws the patients attention to the 'problem' that needs a decision making process.

This item does not attempt to cover the issue of diagnosis as such – for example a patient with a headache may want to be reassured that this symptom is not due to a tumour (we expect that such tasks have been completed before a discussion about what to do (problem management) can occur. So in the instance of a patient with a headache where the clinician is not unduly concerned about the possibility of serious pathology, the clinician could proceed by saying, “so we agree that you have a headache, and that it is unlikely to be due to a serious problem. There are a number of ways in which we could proceed, and I will explain these to you so that you can let me know your views about what would suit you best”. This type of statement, where the agreed problem is 'headache symptom' could then proceed to the behaviour of drawing attention to the making process, and would be given a score of 4. No attempt to draw attention to a need for a decision making process should be scored as 0. Attempts to draw attention to the need to

embark on a decision making process, should thereafter be scored on the degree of skill exhibited. A score of 1 should be given if the attempts is very brief or perfunctory; a score of 2 if the clinician draws attention to a problem that requires a decision making process (baseline skill level); a score of 3 should be given when the clinician puts emphasis on the decision making process required; score of 4 given when the skill is exhibited to a high standard, e.g. supplementary explanations and evidence of patient recognising the need to engage in the process of decision making.

Often there is no clarity about problems, or at least no clarity about the decisions to be taken about the problem or problems identified. If this is the case, this item is given a score of 0. In other words, the skill of drawing attention to the need for a decision making process is not observed. Despite a score of 0 for item 1, the rest of the scale should be completed for the consultation.

Item 2	The clinician states that there is more than one way to deal with the identified problem ('equipoise').
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician conveys the sense that the options are valid and need to be considered in more depth.
3	The clinician explains 'equipoise' in more detail, that options have pros and cons that need to be considered
4	The behaviour is observed and executed to a high standard.

More than one way of managing problems exist in many (if not most) clinical situations, and there is always the choice between providing an intervention and not doing so, i.e. acting conservatively or making a conscious decision to review the need to intervene at a further consultation. Where there is a perfunctory attempt to convey the existence of more than one option then a score of 1 should be give. A score of 2 should be given when the clinician conveys the sense that the options are valid and need to be considered in more depth (baseline skill level). A score of 3 should be given when the clinician explains 'equipoise' in more detail and that options have pros and cons that need to be considered. Where the clinician also explains 'why' choices are available e.g. there is genuine professional uncertainty as to the 'best' way of managing the problem (clinical equipoise) the behaviour will have been executed to a high standard and a score of 4 is given.

Item 3	The clinician assesses the patient's preferred approach to receiving information to assist decision making (e.g. discussion in consultations, read printed material, assess graphical data, use videotapes or other media).
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician asks the patient about their preferred way of receiving information to assist decision.
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Although it is entirely feasible for a clinician to exhibit all the behaviours outlined in this framework in one consultation, it is also recognised that the level of information exchange required to prepare patients for participation in decision making is time consuming and requires reflection about the implications. There are therefore many different approaches that can be used to achieve this purpose. In many instances, clinicians and patients wish to discuss the options and their characteristic in the relevant consultation so that decisions can be made and action taken. But there are also many other possible approaches, and the

arrival of decisions aids that range from paper-based data to digital interactive methods are transforming the degree to which the process of information exchange, and therefore, decision making, is undertaken. Practitioners are becoming aware of these tools, and as they become more readily available to patients, clinicians will need to assess if patients wish to engage with these methods in order to participate more fully in decisions. A score of 2 (baseline skill level) would be given to the clinician who asks about the patient's preferred method of receiving information. A score of 3 would be given for doing this behaviour well e.g. the clinician who states that there are many ways in which information can be conveyed, many of which need the patient to read outside the consultation, and who then assesses the patient's preferred method. A score of 4 would be given for giving many examples of the types of information formats and media available for the patient, and then providing an opportunity for the patient to select their preferred method or methods.

Item 4	The clinician <i>lists</i> 'options', which can include the choice of 'no action'.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician lists options.
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Listing options is different to providing details about each option. The listing of options allows the patient to get an overview of the decision structure, i.e. "This problem has three possible solutions, A or B or C. Let's now consider these options in more detail". This item should receive a score of 1 if a minimal or perfunctory attempt is made to exhibit the behaviour. This item should receive a scores of 2 if the clinician *lists* the options as distinct possibilities that are available, e.g. using 'either / or' phrasing to describe the existence of options (baseline skill level). Careful listing of all possible options, including the choice of taking no action, or deferring the decision should be given a score of 3. If the clinician exhibited this behaviour to a high standard, a score of 4 should be given.

Item 5	The clinician <i>explains</i> the pros and cons of options to the patient (taking 'no action' is an option).
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The behaviour is observed and information about pros and cons are provided.
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 5 should be scored according to the extent that each option is described. A score of 1 should be given if the clinician fails to provide information about more than one option. A score of 2 should be given if the clinician provides details about the pros and cons of the options (baseline skill level). A score of 3 requires the behaviour is exhibited to a good standard. Scores of 4 are given to a clinician who does this task to a high standard.

Item 6	The clinician explores the patient's <i>expectations</i> (or ideas) about how the problem(s) are to be managed.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician explicitly asks the patient what they expected (thought) about the actions required to manage the problem(s).
3	The behaviour is performed to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 6 judges the proficiency of the clinician at exploring the patient's perspective on how the problem was going to be managed. The clinician needs to assess what did the patient think was going to happen (expect) and what ideas might they have had about the *management* of the problem. These ideas are typically difficult to assess. Patients are known to find it difficult to respond to these concepts and are often reluctant to share their views about these issues, for a range of reasons. However, skilled clinicians are able to find ways to explore these views by using open ended questions, suggesting a range of common fears, using pauses, being alert to verbal and physical cues and so on. It is skills such as these that should be assessed in item 6. For example, a patient who thought that she had a menopausal problem might have *expected* further tests or to be prescribed some medication. To assess this expectation, clinicians have to ask patients about their prior conceptions about problem management. Skilled practitioners use phrases such as, 'Patient commonly expect to be sent for an operation or to have tablets provided for this kind of problem, I wonder what you thought might be offered to you?' If the clinician does not attempt to ascertain the patient's views about their *expectations* then this item should receive a score of 0. Unskilled or perfunctory attempts to uncover a patient's ideas or expectations about management should be given a score of 1. Doing this task in a skilled way (using some of the outlined techniques) should be given a score of 2 (baseline skill level). If this behaviour is exhibited and leads to supplementary questions to clarify expectations or ideas, a score of 3 should be given (i.e. exploration of expectations takes place). A score of 4 should be given where this behaviour is achieved to high standards and where the patient's views are discussed and addressed.

Item 7	The clinician explores the patient's <i>concerns</i> (fears) about how problem(s) are to be managed.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician explicitly asks the patient to voice their fears or concerns about the possible actions required to manage the problem(s).
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 7 judges the proficiency of the clinician at exploring the patient's *concerns* and *fears* regarding how the problem was going to be managed. Asking about *concerns* and *fears* requires clinicians to ascertain worries or anxieties that patients may have held. For example, a man who has 'prostatism' (a condition where prostate gland enlargement causes urinary flow obstruction) might have discussed the problem with peers and is worried about the future risk of surgery. These fears are often difficult to assess unless they are explicitly sought. Patients find question about fears and concerns difficult to address, and are often reluctant to share their views about these issues. Skilled clinicians are able to explore these fears and ideas (using open ended questions, suggesting a range of common fears, using pauses, being alert to

verbal and physical cues and so on) and it is these skills that should be assessed in item 7. For example, a clinician might ask, 'Many patients have concerns and fears about treatments or tests, side-effects and so on. Do you have concerns of this sort?' If the clinician does not attempt to ascertain the patient's views about their fears or concerns then this item should receive a score of 0. Unskilled or perfunctory attempts to uncover a patient's fears or concerns about management should be given a score of 1. Exhibiting this behaviour should be given a score of 2 (baseline skill level). If this behaviour is exhibited and leads to supplementary questions to clarify concerns, a score of 3 should be given (i.e. exploration of fears takes place). A score of 4 should be given where this behaviour is achieved to high standards and where the patient's fears or concerns are discussed and addressed.

Item 8	The clinician checks that the patient has <i>understood</i> the information.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician checks the patient's understanding of information provided in the consultation (or assimilated by the patient from other potential sources).
3	The behaviour is observed to a good standard (e.g. exploration of understanding)
4	The behaviour is observed and executed to a high standard.

Item 8 judges whether the clinician actively explores the patient's understanding of the information that has been provided by the clinician (or assimilated by the patient from other potential sources). A perfunctory attempt to check that the patient has understood the relevant information should be given a score 1. To score 2 on this item, an explicit question would have to be posed to the patient asking whether they had understood the information provided or obtained from other sources (baseline skill level). A score of 3 should be given for the clinician that explores the nature of the patients understanding by using statements like: 'I'd like to check that you have understood the information about the possible options. Would you like to let me know what you now understand about this issue?' A score of 4 should be given if the behaviour is executed to a high standard.

Item 9	The clinician offers the patient explicit <i>opportunities</i> to ask questions during the decision making process.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician explicitly asks the patient to voice a question, using phrases such as: 'Do you have any questions?'
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 9 judges whether the clinician provides opportunities for the patient to ask questions and to clarify issues in the consultation that may be broader than that of understanding information. If the clinician provides pauses or other opportunities for queries to be raised, such as appropriate pace within the discourse for example, a score of 1 should be given. If the clinician explicitly asks the patient to voice a question, using phrases such as, 'Do you have any questions?', then a score of 2 should be given (baseline skill level). If the clinician is more specific and asks the patient whether they have questions about the options and the management of the identified problem(s) then a score of 3 should be given. The clinician who achieves this task to a high standard will allow time for the patient to respond and will check if there are any other or supplementary questions and should be given a score of 4. Patients often ask

questions in consultations and these questions cannot be attributed to a behaviour or statement by the clinician. Evidence in the consultation of patient asking questions should be taken into consideration when scoring this item and attention given to the climate and pace of the interview. If the rater judge that the clinicians has created the climate for patient queries, then the behaviour is met and high scores should be given.

Item 10	The clinician elicits the patient's <i>preferred level of involvement</i> in decision making.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician asks the patient about their preferred role in the decision making process.
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 10 judges whether the clinician actively explores the patient's wishes about the role they want to play in making the decisions in the consultation, i.e. a communication about the type of communication. It may be that patients want to be actively involved but are denied the opportunity. Perhaps other patients do not wish to take any part in the decision making process but the clinician assumes that they prefer involvement and acts against their preferred role. It is difficult to assess a patient's preferred role and the rater is not asked to make judgements about this but to assess if an attempt is made to clarify this issue.

A score of 1 should be given if the attempt to elicit the patients preferred role (active or passive) in decision making is perfunctory or rushed. If the clinician explicitly asks the patient about their preferred role then a score of 2 should be given (baseline skill level). Patients often do not understand this question and if the clinician provided further explanation and continues to assess their role preference a score of 3 should be given. A score of 4 should be given to the clinician who asks this question in a way that is easy for patients to understand and which signals that the clinician is sensitive to the decisional responsibility that is being expected of the patient.

Item 11	The clinician indicates the need for a <i>decision making (or deferring)</i> stage.
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The clinician indicates that a time has come where a decision (or deferment) is required.
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 11 judges whether the clinician indicates the need to make, or defer, a decision. This stage can involve a short summary of the options and perhaps an exchange of views, i.e. both clinician and patient views are made explicit and summary statements are elicited, e.g. 'Have you come to a view about this issue?'. Note that the issue of decisional agency (i.e. whether paternalism or not is the *modus operandi*, how the decision is made between the participants and who takes 'control' is not evaluated in this scale). It is also possible the decision point involves deferment, an acceptance that time is required for further information to be obtained, further views explored or for further reflection to occur. If the clinician does not clearly indicate that a time has come where a decision (or deferment) is required, then a score of 0 should be given. A perfunctory or unclear attempt to indicate the need for a decision making state should be given a score of 1. A clear statement such as, 'Perhaps its time now to make a decision about what should be

done', should be given a score of 2 (baseline skill level). Exhibiting this behaviour to good standard should be given a score of 3. A clinician that achieves this task to a high standard will have signalled the transition from consideration of information and views to one of deliberation and closure should be given a score of 4. Note that it is assumed that the decisions considered are not 'urgent' or 'life-threatening' and that it is reasonable in clinical terms to allow sufficient time for a decision making process.

Item 12	The clinician indicates the need to review the decision (or <i>deferment</i>).
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The behaviour is observed.
3	The behaviour is exhibited to a good standard.
4	The behaviour is observed and executed to a high standard.

Item 12 judges whether the clinician provides an opportunity to review the decision or deferment, i.e. to allow time for a decision to be reconsidered and if necessary revised or altered. A perfunctory (e.g. that the patient should be seen again) or rushed attempt should be given a score of 1. If the clinician indicates that the patient should be seen again to re-consider the decision, then a score of 2 should be given (baseline skill level). If this behaviour is performed to a good standard a score of 3 should be given. A clinician that exhibits this behaviour to a high standard, e.g. makes it very explicit and encourages this approach should be given a score of 4.

3. OPTION Observing patient involvement (Research Version)



© June 2004 elwyng@cardiff.ac.uk

Rater Name		Clinician Code		Date of rating	DD	MM	YY
				Consultation number			
				Consultation duration (m, s)			
				Practitioner (M = 1, F = 2)	Age	Sex	
				Patient (M = 1, F = 2)	Age	Sex	
				New Consultation	1		
				Review Consultation	2		
				Composite Consultation	3		
Description of index problem							

1	The clinician <i>draws attention to</i> an identified problem as one that requires a decision making process.	0	1	2	3	4
2	The clinician <i>states</i> that there is more than one way to deal with the identified problem ('equipoise').	0	1	2	3	4
3	The clinician <i>assesses</i> the patient's preferred approach to receiving information to assist decision making (e.g. discussion, reading printed material, assessing graphical data, using videotapes or other media).	0	1	2	3	4
4	The clinician <i>lists</i> 'options', which can include the choice of 'no action'.	0	1	2	3	4
5	The clinician <i>explains</i> the pros and cons of options to the patient (taking 'no action' is an option).	0	1	2	3	4
6	The clinician explores the patient's <i>expectations</i> (or ideas) about how the problem(s) are to be managed.	0	1	2	3	4
7	The clinician explores the patient's <i>concerns</i> (fears) about how problem(s) are to be managed.	0	1	2	3	4
8	The clinician checks that the patient has <i>understood</i> the information.	0	1	2	3	4
9	The clinician offers the patient explicit <i>opportunities</i> to ask questions during the decision making process.	0	1	2	3	4
10	The clinician elicits the patient's <i>preferred level of involvement</i> in decision-making.	0	1	2	3	4
11	The clinician indicates the need for a <i>decision making</i> (or <i>deferring</i>) stage.	0	1	2	3	4
12	The clinician indicates the need to review the decision (or <i>deferral</i>).	0	1	2	3	4

Score	Description
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The behaviour is observed and a minimum skill level achieved.
3	The behaviour is exhibited to a good standard.
4	The behaviour is exhibited to a very high standard.

4. OPTION Observing patient involvement (Educational Feedback Version)

© June 2004 elwyng@cardiff.ac.uk

Rater Name		Clinician Code		Date of rating	DD	MM	YY
				Consultation number			
				Consultation duration (m, s)			
				Practitioner (M = 1, F = 2)	Age	Sex	
				Patient (M = 1, F = 2)	Age	Sex	
				New Consultation	1		
				Review Consultation	2		
				Composite Consultation	3		
Description of index problem							

1	The clinician <i>draws attention</i> to an identified problem as one that requires a decision making process.	Yes 1	No 0
2	The clinician <i>states</i> that there is more than one way to deal with the identified problem ('equipoise').	Yes 1	No 0
3	The clinician <i>assesses</i> the patient's preferred approach to receiving information to assist decision making (e.g. discussion, reading printed material, assessing graphical data, using videotapes or other media).	Yes 1	No 0
4	The clinician <i>lists</i> 'options', which can include the choice of 'no action'.	Yes 1	No 0
5	The clinician <i>explains</i> the pros and cons of options to the patient (taking 'no action' is an option).	Yes 1	No 0
6	The clinician explores the patient's <i>expectations</i> (or ideas) about how the problem(s) are to be managed.	Yes 1	No 0
7	The clinician explores the patient's <i>concerns</i> (fears) about how problem(s) are to be managed.	Yes 1	No 0
8	The clinician checks that the patient has <i>understood</i> the information.	Yes 1	No 0
9	The clinician offers the patient explicit <i>opportunities</i> to ask questions during the decision making process.	Yes 1	No 0
10	The clinician elicits the patient's <i>preferred level of involvement</i> in decision-making.	Yes 1	No 0
11	The clinician indicates the need for a <i>decision making</i> (or <i>deferring</i>) stage.	Yes 1	No 0
12	The clinician indicates the need to review the decision (or <i>deferment</i>).	Yes 1	No 0

6. OPTION Scale

6.1	English version	93
6.2	Dutch version and translation process	95
6.3	French version and translation process	97
6.4	German version and translation process	99
6.5	Spanish version and translation process	101
6.6	Italian version and translation process	105

6.1 English Version



OPTION Observing patient involvement

© June 2004 elwyng@cardiff.ac.uk

Rater Name		Clinician Code		Date of rating	DD	MM	YY
				Consultation number			
				Consultation duration (m, s)			
				Practitioner (M = 1, F = 2)			
				Patient (M = 1, F = 2)			
				New Consultation			
				Review Consultation			
				Composite Consultation			
Description of index problem							

1	The clinician <i>draws attention to</i> an identified problem as one that requires a decision making process.	0	1	2	3	4
2	The clinician <i>states</i> that there is more than one way to deal with the identified problem ('equipoise').	0	1	2	3	4
3	The clinician assesses the patient's preferred approach to receiving information to assist decision making (e.g. discussion, reading printed material, assessing graphical data, using videotapes or other media).	0	1	2	3	4
4	The clinician <i>lists</i> 'options', which can include the choice of 'no action'.	0	1	2	3	4
5	The clinician <i>explains</i> the pros and cons of options to the patient (taking 'no action' is an option).	0	1	2	3	4
6	The clinician explores the patient's <i>expectations</i> (or ideas) about how the problem(s) are to be managed.	0	1	2	3	4
7	The clinician explores the patient's <i>concerns</i> (fears) about how problem(s) are to be managed.	0	1	2	3	4
8	The clinician checks that the patient has <i>understood</i> the information.	0	1	2	3	4
9	The clinician offers the patient explicit <i>opportunities</i> to ask questions during the decision making process.	0	1	2	3	4
10	The clinician elicits the patient's <i>preferred level of involvement</i> in decision-making.	0	1	2	3	4
11	The clinician indicates the need for a <i>decision making</i> (or <i>deferring</i>) stage.	0	1	2	3	4
12	The clinician indicates the need to review the decision (or <i>deferment</i>).	0	1	2	3	4

Score	Description
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The behaviour is observed and a minimum skill level achieved.
3	The behaviour is exhibited to a good standard.
4	The behaviour is exhibited to a very high standard.

6.2 Dutch Version



OPTION Observatie patient inbreng

© Juni 2004 elwyng@cardiff.ac.uk

Rater Naam	<input type="text"/>	Arts Code	<input type="text"/>	Datum van rating	Dag	Maand	Jaar
				Consult nummer			
				Consult duur (m, s)			
				Arts (M = 1, V = 2)	LFT		GSL
				Patient (M = 1, V = 2)	LFT		GSL
				Nieuw Consult		1	
				Follow-up Consult		2	
				Samengesteld Consult		3	
Omschrijving van index probleem							

1	De arts vestigt de aandacht op een specifiek probleem als een probleem waarover een beslissing moet worden genomen.	0	1	2	3	4
2	De arts geeft aan dat er meer dan één manier is om met het betreffende probleem om te gaan ('gelijkwaardigheid').	0	1	2	3	4
3	De arts stelt vast op welke wijze de patiënt informatie zou willen krijgen ter ondersteuning van het besluitvormingsproces (bijv. overleg, voorlichtingsfolder, grafieken bekijken, videobanden of andere media).	0	1	2	3	4
4	De arts geeft een opsomming van de keuzemogelijkheden, inclusief de eventuele keuze om 'niets te doen'.	0	1	2	3	4
5	De arts geeft uitleg over voors en tegens van de keuzemogelijkheden aan de patiënt (waarbij 'niets doen' ook een keuzemogelijkheid is).	0	1	2	3	4
6	De arts gaat na welke verwachtingen (of ideeën) de patiënt heeft over de aanpak van het probleem.	0	1	2	3	4
7	De arts gaat na welke zorgen (angsten) de patiënt heeft over de aanpak van het probleem.	0	1	2	3	4
8	De arts controleert of de patiënt de informatie heeft begrepen.	0	1	2	3	4
9	De arts geeft de patiënt expliciet de mogelijkheid om vragen te stellen tijdens het besluitvormingsproces.	0	1	2	3	4
10	De arts gaat stelt vast in welke mate de patiënt betrokken wil worden in de besluitvorming.	0	1	2	3	4
11	De arts geeft aan dat het noodzakelijk is om over te gaan tot het nemen van een beslissing (of om deze uit te stellen).	0	1	2	3	4
12	De arts geeft aan dat het mogelijk is om op de beslissing terug te komen (of het uitstel van de beslissing).	0	1	2	3	4

Score	Omschrijving
0	Dit gedrag wordt niet waargenomen.
1	Een minimale poging is gedaan om dit gedrag te vertonen.
2	De arts vestigt de aandacht op een specifiek probleem als een probleem waarover een beslissing moet worden genomen.
3	Het gedrag wordt waargenomen en de arts legt nadruk op het besluitvormingsproces dat nodig is.
4	Dit gedrag wordt waargenomen en voldoet aan zeer hoge maatstaven.

Dutch translation process

Stage One:

The English instrument (items and answering scale) has been translated into Dutch by two independent researchers. A meeting was held and consensus was reached about this forward translation.

Stage Two:

Next, the Dutch version was translated back into English by two other independently working researchers.

Stage Three:

The result was compared with the original English version, which led to a few additional changes in the Dutch version.

Stage Four:

The explanation to the 12 items was translated by the four researchers; each of them translated the explanation of 3 items.

6.3 French version



OPTION Observer l'implication du patient

© Juin 2004 elwyng@cardiff.ac.uk

Nom de l'évaluateur/trice		Code clinicien/ne		Date de l'évaluation	JJ	MM	AA
Code de la consultation							
Durée de la consultation (m, s)							
Professionnel/le (M = 1, F = 2)				Age	Sexe		
Patient/e (M = 1, F = 2)				Age	Sexe		
Consultation Nouvelle				1			
Consultation Subséquente				2			
Consultation Mixte				3			
Description du problème principal							

1	Le clinicien <i>attire l'attention</i> sur un problème identifié nécessitant un processus de prise de décision.	0	1	2	3	4
2	Le clinicien <i>mentionne</i> qu'il y a plus qu'une option pour faire face au problème identifié ('contre-poids').	0	1	2	3	4
3	Le clinicien <i>évalue</i> comment le patient préfère recevoir l'information afin d'aider celui-ci à prendre une décision (i.e. discussion, lecture de texte imprimé, documentation écrite, illustration graphique, dessins, vidéos ou autres moyens).	0	1	2	3	4
4	Le clinicien <i>présente</i> les différentes options possibles incluant le choix de « ne rien faire ».	0	1	2	3	4
5	Le clinicien <i>explique</i> les avantages (le pour) et les inconvénients (le contre) associés aux différentes options (incluant celle de « ne rien faire »).	0	1	2	3	4
6	Le clinicien explore <i>les attentes</i> (ou les idées) du patient concernant les manières avec lesquelles le problème peut être géré.	0	1	2	3	4
7	Le clinicien explore <i>les inquiétudes</i> (craintes) du patient concernant les manières avec lesquelles le problème peut être géré.	0	1	2	3	4
8	Le clinicien <i>vérifie</i> que le patient a <i>compris</i> l'information.	0	1	2	3	4
9	Le clinicien donne <i>l'opportunité</i> au patient de poser des questions tout au long du processus de prise de décision.	0	1	2	3	4
10	Le clinicien précise <i>le rôle que le patient préfère</i> jouer dans la prise de décision.	0	1	2	3	4
11	Le clinicien indique <i>le besoin de prendre une décision</i> (ou de la reporter).	0	1	2	3	4
12	Le clinicien indique le besoin de revoir la décision (ou de la reporter).	0	1	2	3	4

Échelle de réponses	
0	Le comportement n'est pas observé.
1	Le comportement est à peine amorcé.
2	Le comportement est observé et est démontré minimalement.
3	Le comportement est observé et est démontré avec compétence.
4	Le comportement est observé et est démontré avec un très haut niveau de compétence.

French translation process

Stage one:

France Légaré and Michel Cauchon independently translated the OPTION scale in French. They sent their respective version to Michel Labrecque.

Stage two:

All three (Michel Labrecque, Michel Cauchon and France Légaré) met and reached a consensus on the French translation. Meanwhile, France Légaré received comments from Franck Zenasni in Paris about the French used in the first French version.

Stage three:

The French version was sent to Glyn Elwyn who sent it to three back translators. Glyn received the three English back translations for comparison.

Stage four:

Glyn Elwyn and France Légaré met and reached a consensus about a second French version that integrates comments from stages one, two and three. A final French version of the scale was agreed.

6.4 German version



OPTION Beobachtung der Einbeziehung von Patienten

© Juni 2004 elwyng@cardiff.ac.uk

Name des Bewerbers		Arzt-Code		Datum der Bewertung	Tag	Monat	Jahr
				Konsultation Nr.			
				Dauer der Konsultation (Minuten, Sekunden)			
				Arzt (männlich = 1, weiblich = 2)	Alter	Geschlecht	
				Patient (männlich = 1, weiblich = 2)	Alter	Geschlecht	
				Erste Konsultation	1		
				Folge- Konsultation	2		
				Gemischte- Konsultation	3		
Beschreibung des Indexproblems							

1	Der Arzt <i>lenkt die Aufmerksamkeit</i> auf ein bestimmtes Problem, das einer Entscheidung bedarf.	0	1	2	3	4
2	Der Arzt <i>teilt mit</i> , dass es mehr als einen Weg gibt, um mit dem identifizierten Problem umzugehen.	0	1	2	3	4
3	Der Arzt <i>klärt</i> , in welcher Art und Weise der Patient Informationen erhalten möchte, um eine Entscheidung treffen zu können (z. B. im Gespräch, durch das Lesen von Informationsmaterial, durch die Präsentation grafisch aufbereiteter Daten, durch Videos oder andere Medien).	0	1	2	3	4
4	Der Arzt <i>zählt</i> Optionen auf, worunter auch die Option fallen kann, nichts zu tun.	0	1	2	3	4
5	Der Arzt <i>erläutert</i> dem Patienten die Vor- und Nachteile der verschiedenen Optionen (<i>nichts tun</i> ist ebenfalls eine Option).	0	1	2	3	4
6	Der Arzt exploriert <i>die Erwartungen</i> (oder Ideen) eines Patienten, wie mit dem Problem (den Problemen) umgegangen werden soll.	0	1	2	3	4
7	Der Arzt exploriert die <i>Sorgen</i> (Befürchtungen) des Patienten, wie mit dem Problem (den Problemen) umgegangen werden soll.	0	1	2	3	4
8	Der Arzt vergewissert sich, dass der Patient die Informationen <i>verstanden</i> hat.	0	1	2	3	4
9	Der Arzt bietet dem Patienten explizit <i>Möglichkeiten</i> an, während des Entscheidungsprozesses Fragen zu stellen.	0	1	2	3	4
10	Der Arzt findet heraus, in welchem <i>Ausmass</i> der Patient bei der Entscheidungsfindung beteiligt werden möchte.	0	1	2	3	4
11	Der Arzt weist darauf hin, dass es notwendig ist, eine <i>Entscheidung jetzt zu treffen</i> (oder <i>aufzuschieben</i>).	0	1	2	3	4
12	Der Arzt weist darauf hin, dass es notwendig ist, noch einmal auf die Entscheidung zurückzukommen (oder auf das Aufschieben der Entscheidung).	0	1	2	3	4

Score	Beschreibung
0	Die Kompetenz wird nicht beobachtet .
1	Ein minimaler oder oberflächlicher Versuch wird unternommen, die Kompetenz zu zeigen.
2	Die Kompetenz wird beobachtet und erreicht ein minimales Niveau.
3	Die Kompetenz entspricht einem guten Standard.
4	Die Kompetenz entspricht einem sehr hohen Standard.

German translation process

Stage one:

The starting point for the translation of the OPTION scale into German was the English version authorised by Glyn Elwyn and colleagues who developed the scale. Two research groups were involved in that translation process. They have broad experiences in the field of medical communication and health services research and include the following: Professor Dr Wolf Langewitz, Head of Dept., Psychosomatics and Internal Medicine, University Hospital of Basel, Switzerland, and Prof. Dr. Martin Haerter, Head of Dept. of Psychiatry and Psychotherapy, Clinical Epidemiology and Health Services Research University Hospital of Freiburg, Germany.

Stage two:

Both groups used native speakers, Wolf Langewitz and Andreas Loh, to perform independent translations into German.

Stage three:

In an expert panel consisting of Martin Haerter, Wolf Langewitz and Andreas Loh, the differences between these translations were identified and the linguistic and content elements were discussed and optimised. A final German version was agreed.

Stage four:

The final German version was then back translated into English by Rigo Brueck and Anthony Tyndall who are two independent native speakers. Both English back translations were sent to Glyn Elwyn, who authorised the German translation of the OPTION-scale.

6.5 Spanish Version



OPTION Observación de participación de pacientes

© Junio 2004 elwyng@cardiff.ac.uk

Nombre del/la evaluador/a		Código del/la clínico/a		Fecha de la evaluación	Día	Mes	Año
Número de consulta							
Duración de la consulta (min, seg)							
Personal médico (M = 1, F = 2)				Edad		Sexo	
Paciente (M = 1, F = 2)				Edad		Sexo	
Consulta nueva				1			
Consulta subsecuente				2			
Consulta en colaboración				3			
Descripción del problema principal							

1	El/la clínico/a <i>llama la atención</i> a un problema ya identificado que requiere un proceso de toma de decisión.	0	1	2	3	4
2	El/la clínico/a <i>menciona</i> que hay más de una manera de tratar el problema ('equilibrio').	0	1	2	3	4
3	El/la clínico/a <i>evalúa</i> la manera en que el/la paciente prefiere recibir la información para ayudar con el proceso de la toma de decisión (ej: conversación, lectura de material impreso, evaluación de datos gráficos, uso de vídeos u otros medios).	0	1	2	3	4
4	El/la clínico/a <i>da una lista</i> de 'opciones', incluyendo la posibilidad de 'no hacer nada.'	0	1	2	3	4
5	El/la clínico/a <i>explica</i> las ventajas y desventajas de las opciones al/la paciente (incluyendo 'no hacer nada' como una opción).	0	1	2	3	4
6	El/la clínico/a <i>explora</i> las <i>expectativas</i> (o ideas) del/la paciente sobre cómo se puede tratar el/los problema/s.	0	1	2	3	4
7	El/la clínico/a <i>explora</i> las inquietudes (miedos) del/la paciente sobre cómo se puede tratar el/los problema/s.	0	1	2	3	4
8	El/la clínico/a se asegura que el/la paciente haya <i>entendido</i> la información.	0	1	2	3	4
9	El/la clínico/a da la <i>oportunidad</i> de hacer preguntas concretas durante el proceso de toma de decisión.	0	1	2	3	4
10	El/la clínico/a obtiene el <i>nivel preferido de participación</i> del/la paciente en la toma de decisión.	0	1	2	3	4
11	El/la clínico/a indica que hay que dar el paso de <i>tomar una decisión</i> (o <i>aplazarla</i>).	0	1	2	3	4
12	El/la clínico/a indica que hay que <i>revisar</i> la decisión (o <i>aplazarla</i>).	0	1	2	3	4

	Niveles de observación
0	No se observa el comportamiento.
1	Se hace el mínimo intento de mostrar el comportamiento.
2	Se observa el comportamiento y se llega a un nivel mínimo de destreza.
3	Se muestra el comportamiento a un buen nivel.
4	Se muestra el comportamiento a un nivel muy alto.

Spanish translation process

Stage one:

Authorized by Professor Glyn Elwyn, Dr. Silvia Rodríguez-Sabater, an assistant professor of Spanish and Linguistics and a native speaker of Spanish and Catalan, made an independent Spanish translation of the OPTION decision making scale. She consulted with Dr. Jason Roberson, an assistant professor of Spanish and Linguistics with experience in medical interpretation who is a native English speaker fluent in Spanish, prior to submission for a formal back translation into English. They raised issues concerning nuances in Spanish concepts involved in the decision making process. Further discussions involved the use of the subjunctive in discussing differences in expressing uncertainty in the two languages. These differences were held for later discussion by the entire translation team. The process used is based on the model provided in the "Guidelines for the Process of Cross Cultural Adaptation of Self-Report Measures" (Beaton, Bombardier, Guillemin, and Ferraz, 2000), later substantiated by Kovacs et al (2002).

Stage two:

Dr. Rodríguez-Sabater gave the Spanish translation to Dr. Elizabeth Martínez-Gibson, an associate professor of Spanish and Linguistics who is both a native Spanish and English speaker and who also speaks French. Dr. Martínez-Gibson made an independent back translation from the Spanish version into English for comparison with the English original of the OPTION scale.

Stage three:

As a first level of cognitive debriefing (Gandek, 2003), the native Spanish speakers (Rodríguez-Sabater and Martínez-Gibson) then reviewed variations between the three versions. They discussed British English versus American English variations in particular medical concepts (for example, 'review consultation' versus 'follow-up visit'). They consulted the French version of OPTION to clarify linguistic and cultural concepts that share a Latin perspective. The concept of 'equipoise' was discussed for an equivalent acceptable in Spanish, choosing 'equilibrio.'

Stage four:

As a final level of cognitive debriefing, the primary translation team, Dr. Rodríguez-Sabater and Dr. Martínez-Gibson met with Dr. Charlene Pope, a native Romanian and English speaker with some knowledge of French and Spanish who is a health service researcher and originally a nurse-midwife. All are trained as sociolinguists, with Dr. Pope furnishing a clinician perspective for the content area of the language translation. The team compared the original version of the OPTION scale with the Spanish version, the back translation, and the French version for synthesis and final translation.

The team decided that some shifts in language reflect both syntactic and semantic differences as cognitive and affective dimensions of the pragmatics of decision making as well as word equivalence. For example, neither 'involvement' in the Title and Item 10 in the context of clinician-patient interaction or 'deal with' in Item 2 in the context of decision making are either words or concepts common in Spanish. In the Title, the replacement of 'involvement' with the Spanish word for participation reflects the closest expectations of acceptable behavior as well as a more fitting word equivalent. The plural for patients in the Title reflects Spain's current language policy on fairness in genderless identification. In Item 2, the concept 'deal with' required consideration of the variations in the decision making process that would be expected culturally in clinician-patient interaction. Looking at the spectrum (from Examine - to Manage - to Treat - to Solve), 'tratar' or treat appears more in keeping with more action-oriented assumptions about problems for Spanish speakers. In Item 5, the choice of 'no action' did not make as much sense to either native Spanish or second language Spanish speakers as 'no hacer nada' or 'doing nothing,' considered a more understandable cultural construct. In Item 7, the team felt that 'concern' in a patient context captured the disquiet of 'inquietudes' rather than the fears and worries inherent in the concept of preoccupation, the alternative choice. In Item 8, the verb 'se asegura' was softened with the subjunctive form 'haya' later in the sentence to reflect more desired uncertainty, considered part of a cultural fit (Hendrickson, 2003). Finally, Item 11 prompted native Spanish speakers to reflect that the concept 'need' feels artificial or culturally inappropriate, as a condition of individual deficit rather than in the context of a desired action. The translation 'take a step towards a decision' reflects the more impersonal metaphor of motion rather than personal incapacity which the English 'need' suggests. These suggestions were submitted to Dr. Elwyn, the developer of the OPTION scale for final approval and synthesis.

Stage 5:

Given the variations within the Spanish-speaking world, the translation team for the OPTION Spanish version chose to use a standard variation of the language. Although the translators are native speakers of Spanish from Spain, they are linguists who have worked with all varieties of Spanish in Latin America and the United States.

References:

Beaton D, Bombardier C, Guillemin F, Ferraz M. Guidelines for the process of cross-cultural adaptation for self-report measures. *Spine* 2000; 25: 3186-3191.

Gandek B, Alcoque A, Andrew-Hobs M, Davis K. Translating the short-form Headache Impact Test (HIT) in 27 countries: Methodological and conceptual issues. *Quality of Life Research* 2003; 12: 975-979.

Hendrickson SG. Beyond translation...cultural fit. *Western Journal of Nursing Research* 2003; 25: 593-608.

Kovacs F. et al. Validation of the Spanish version of the Roland-Morris Questionnaire. *Spine* 2002; 27: 538-542.

6.6 Italian Version



OPTION Osservare il coinvolgimento del paziente

© Giungio 2004 elwyng@cardiff.ac.uk

Nome del valutatore		Codice del medico		Data di valutazione	Giorni	Mese	Anno
Numero della consulenza Durata della consulenza (m,s) Médico (M = 1, F = 2) Paziente (M = 1, F = 2) Nuova consultazione Consultazione successiva Consultazione mista							
				età		Sesso	
				età		Sesso	
					1		
					2		
					3		
Descrizione del problema indice:							

1	Il medico focalizza l'attenzione su un problema specifico che richiede un processo decisionale	0	1	2	3	4
2	Il medico dice che c'è più di un modo di affrontare il problema identificato ("equità")	0	1	2	3	4
3	Il medico valuta in che modo il paziente preferisca ricevere le informazioni che lo possano aiutare nel processo decisionale (e.g. parlando con il medico, leggendo materiale informativo, attraverso la presentazione di grafici, l'uso di videocassette o altri mezzi)	0	1	2	3	4
4	Il medico elenca le "opzioni", in cui può essere inclusa anche la scelta di non fare niente	0	1	2	3	4
5	Il medico spiega i pro e i contro delle opzioni al paziente (Il non fare niente è un'opzione)	0	1	2	3	4
6	Il medico esplora le aspettative e/o le idee del paziente rispetto al modo di gestire il problema (o problemi)	0	1	2	3	4
7	Il medico esplora le preoccupazioni e/o le paure del paziente rispetto al modo di gestire il problema (o problemi)	0	1	2	3	4
8	Il medico verifica che il paziente abbia capito l'informazione	0	1	2	3	4
9	Il medico offre al paziente in modo esplicito delle opportunità di far domande durante il processo decisionale.	0	1	2	3	4
10	Il medico valuta in che misura il paziente vorrebbe essere coinvolto nel processo decisionale.	0	1	2	3	4
11	Il medico indica la necessità di prendere una decisione ora o di rimandarla.	0	1	2	3	4
12	Il medico segnala la necessità di rivedere e/o rimandare la decisione.	0	1	2	3	4

Livello	Descrizione
0	Il comportamento non si è osservato
1	Si è fatto un tentativo minimo di mostrare il comportamento
2	Il comportamento si è osservato e si è raggiunto un minimo livello di abilità
3	Il comportamento si è mostrato ad un buon livello
4	Il comportamento si è mostrato ad un livello eccellente

Italian translation process

Stage one

Authorized by Professor Glyn Elwyn the English instrument and manual has been translated into Italian by two independent Italian native speakers (Claudia Goss & Silvia Fontanesi). After a discussion an agreement on a first version was found.

Stage two

This version was checked for the language fluency and consistence by a school teacher of Italian and then it was back translated into English by Professor Christa Zimmermann.

Stage three

This English version was compared with the original English version by an English native speaker (Professor Graham Thornicroft).

Stage four

In an expert panel consisting of Claudia Goss, Christa Zimmermann, Lidia Del Piccolo, Michela Rimondini & Monica Paccaloni, some language differences were discussed and a final version of the instrument was agreed upon. The main differences with the original English version are:

- The term “review consultation” (consulazione di revisione) has been translated into “subsequent consultation” (consultazione successiva) as the term “review” in Italian is not so frequently used for medical consultations.
- the term “clinician”(clinico) has been translated with “doctor” (medico) as in Italy it is rare that health professionals other the doctor deliver information.
- In item 2. “equipoise” (equilibrio) has been translated with “equity” (equità) because this term expresses more clearly that there are more than one the way to deal to the identified problem and there are no preferences *a priori*.
- Item 11. the adverb “now” (ora) has been added and the word “stage” (stadio) has been deleted to better clarify that there is a specific moment of the consultation in which the doctor express the need to decide what to do.

7. OPTION Training Pack

Contents

1. Introduction	107
2. How to use the OPTION instrument	107
3. Using the test consultations	107
4. Checking your scores	108

A more detailed training manual and CD are available from the author.
Please contact prof. dr. Glyn Elwyn, e-mail: elwyng@cardiff.ac.uk

1. Introduction

A systematic review has shown that no measures of the extent to which healthcare professionals involve patients in decisions within clinical consultations exist, despite the increasing interest in the benefits or otherwise of patient participation in these decisions.

The OPTION scale was developed and used by two independent raters to assess primary care consultations in order to evaluate its psychometric qualities, validity and reliability. The following background papers are available in section five of this pack:

- Elwyn G, Edwards A, Kinnersley P. Shared decision-making in primary care: the neglected second half of the consultation. *British Journal of General Practice* 1999; 49: 477-482.
- Elwyn G, Edwards A, Wensing M, Hibbs R, Wilkinson C, Grol R. Shared decision making observed in clinical practice: visual displays of communication sequence and patterns. *Journal of Evaluation in Clinical Practice* 2001; 7: 211-221.
- Elwyn G, Edwards A, Mowle S, Wensing M, Wilkinson C, Kinnersley P, Grol R. Measuring the involvement of patients in shared decision-making: a systematic review of instruments. *Patient Education and Counselling* 2001; 43: 5-22.
- Elwyn G, Edwards A, Wensing M, Atwell C, Hood K, Grol R. Fleeting glimpses: measuring shared decision making in primary care using the OPTION instrument. *Quality & Safety in Health Care* 2003; 12: 93-9.
- Elwyn G, Edwards A, Kinnersley P, Grol R. Shared decision making and the concept of equipoise: the competencies of involving patients in healthcare choices. *British Journal of General Practice* 2000; 50(460): 892-899.
- Elwyn G, Edwards A, Hood K, Robling M, Atwell C, Russell I, Wensing M, Grol R and the Study Steering Group. Achieving involvement: process outcomes from a cluster randomised trial of shared decision making skill development and use of risk communication aids in general practice. *Family Practice* 2004; 21(4): 337-346.
- Edwards A, Elwyn G, Hood K, Atwell C, Robling M, Houston H, Kinnersley P, Russell I and the Study Steering Group. Patient-based outcome results from a cluster randomised trial of shared decision making skill development and use of risk communication aids in general practice. *Family Practice* 2004; 21 (4): 347-354.

2. How to use the OPTION instrument

Please find attached a copy of the OPTION manual and scale.

The OPTION scale consists of twelve items which are used to score each consultation.

Once you have looked at the scale, the manual will help you to learn about and gain a better understanding of the OPTION scoring process. The manual provides a description of each of the twelve items in the scale and offers suggestions as to which score to select.

There are useful examples of what to look for in consultations which guide you to select the most accurate score for each item.

Most studies have used two raters to score each consultation, but it is also possible to use only one rater.

3. Using the test consultations

Once you have read the manual and scale, please listen to the attached CD containing seven test consultations.

These consultations involve a range of different patients who are visiting their GP. Please score the consultations using the OPTION scale.

4. Checking your scores

The attached table can be used to assess your scores for the test consultations.

By adding up the score for each consultation you will arrive at a number between 0 and 48. We recommend that if you are using two raters that you use the mean of those two scores, for example:

Score one: $\frac{12}{48}$ + score two: $\frac{16}{48}$ = mean score of $\frac{14}{48}$

It is also recommended that this raw score should be transformed to a scale of 0 to 100 by using the following calculation:

$$(\frac{14}{48} \times 100) = 29$$

This allows the OPTION score to be represented by a score that lies between 0 and 100.

Suggested scores for the seven consultations are as follows:

Consultation	Mean Score from two raters	Translated to score out of 100
One	$\frac{7}{48}$	15
Two	$\frac{30}{48}$	63
Three	$\frac{31}{48}$	65
Four	$\frac{9}{48}$	19
Five	$\frac{22}{48}$	46
Six	$\frac{32}{48}$	67
Seven	$\frac{30}{48}$	63



SAMPLE SCORING SHEET OPTION Observing patient involvement

© June 2004 elwyng@cardiff.ac.uk

Rater Name	Rater 1	Clinician Code	001	Date of rating	DD 2	MM 6	YY 2004
Consultation number				1			
Consultation duration (m, s)				15 mins			
Practitioner (M = 1, F = 2)				Age	49	Sex	1
Patient (M = 1, F = 2)				Age	45	Sex	2
New Consultation				1			
Review Consultation				2			
Composite Consultation				3			
Description of index problem							
Atrial Fibrillation							

1	The clinician <i>draws attention to</i> an identified problem as one that requires a decision making process.	0	1	2	3	4
2	The clinician <i>states</i> that there is more than one way to deal with the identified problem ('equipoise').	0	1	2	3	4
3	The clinician <i>assesses</i> the patient's preferred approach to receiving information to assist decision making (e.g. discussion, reading printed material, assessing graphical data, using videotapes or other media).	0	1	2	3	4
4	The clinician <i>lists</i> 'options', which can include the choice of 'no action'.	0	1	2	3	4
5	The clinician <i>explains</i> the pros and cons of options to the patient (taking 'no action' is an option).	0	1	2	3	4
6	The clinician explores the patient's <i>expectations</i> (or ideas) about how the problem(s) are to be managed.	0	1	2	3	4
7	The clinician explores the patient's <i>concerns</i> (fears) about how problem(s) are to be managed.	0	1	2	3	4
8	The clinician checks that the patient has <i>understood</i> the information.	0	1	2	3	4
9	The clinician offers the patient explicit <i>opportunities</i> to ask questions during the decision making process.	0	1	2	3	4
10	The clinician elicits the patient's <i>preferred level of involvement</i> in decision-making.	0	1	2	3	4
11	The clinician indicates the need for a <i>decision making</i> (or <i>deferring</i>) stage.	0	1	2	3	4
12	The clinician indicates the need to review the decision (or <i>deferment</i>).	0	1	2	3	4

Score	Description
0	The behaviour is not observed.
1	A minimal attempt is made to exhibit the behaviour.
2	The behaviour is observed and a minimum skill level achieved.
3	The behaviour is exhibited to a good standard.
4	The behaviour is exhibited to a very high standard.

Total score: 8