

# The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks

Glyn Elwyn MSc PhD FRCGP,\* Hayley Hutchings PhD,† Adrian Edwards MRCP MRCGP MPhil PhD,‡ Frances Rapport PhD,§ Michel Wensing PhD,¶ Wai-Yee Cheung PhD\*\* and Richard Grol PhD††

\*Professor, Primary Care Group, School of Medicine, University of Wales Swansea, Swansea, UK, †Lecturer, Primary Care Group, School of Medicine, University of Wales Swansea, Swansea, UK, ‡Reader, Primary Care Group, School of Medicine, University of Wales Swansea, Swansea, UK, §Julian Tudor Hart Research Fellow, Primary Care Group, School of Medicine, University of Wales Swansea, Swansea, UK, ¶Senior Lecturer, Centre for Quality of Care Research (WOK), Radboud University Medical Centre, the Netherlands, \*\*Senior Lecturer, Primary Care Group, School of Medicine, University of Wales Swansea, Swansea, and ††Professor, Centre for Quality of Care Research (WOK), Radboud University Medical Centre, Nijmegen, the Netherlands

---

## Correspondence

Glyn Elwyn  
Professor, Primary Care Group  
School of Medicine  
University of Wales Swansea  
Grove Building  
Singleton Park  
Swansea SA2 8PP  
UK  
E-mail: g.elwyn@swansea.ac.uk

## Accepted for publication

5 November 2004

**Keywords:** communication skills, patient involvement, psychometric characteristics, scale development, shared decision making

---

## Abstract

**Objective** To examine the psychometric properties of a revised scale, named ‘observing patient involvement in decision making’ (OPTION), by analysing its reapplication to a sample of routine primary care consultations. The OPTION instrument assesses to what degree clinicians involve patients in decision making.

**Design** Cross-sectional assessment of medical interaction by two calibrated raters.

**Setting** Primary care.

**Participants** Twenty-one general practitioners provided 186 consultations for assessment.

**Measurements** Observational score using the OPTION instrument.

**Results** Compared with the first version of the OPTION scale, the revised scale that uses a *magnitude* instead of an *attitude* scale, when applied to the same data set, resulted in 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. 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 (the intraclass correlation coefficient 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 performance.

**Conclusions** We conclude that OPTION is sufficiently reliable to be used for formal assessment at the level of the whole instrument (all 12 items).

## Introduction

Although there is an increasing call for clinicians to involve patients in decision making about health-care 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.<sup>1</sup> The overall intent of the communication strategy, often known as ‘shared decision making’,<sup>2</sup> 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.<sup>3</sup> The skills (competences) have been elucidated and discussed.<sup>4</sup>

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 refer 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 ‘observing patient involvement in decision making’ (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, influ-

ence 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 (OPTION scale).<sup>5</sup> This work described a theoretical construct, provided details of scale design stages and item formulation.<sup>6</sup> However, it was recognized that the instrument had aspects that required further attention.<sup>5</sup> 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 midpoint 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.

## Methods

The study examined the psychometric characteristics of a revised OPTION scale applied to a sample of audiotaped 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.<sup>5</sup> 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 1 year and <10 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.<sup>5</sup> In order to test inter-rater reliability, the two raters rated all the consultations independently and a random sample of 21 consultations (one 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 five-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 midpoint for both types of uncertainty inflated the OPTION scores as it is a recognized problem of attitudinal scales.<sup>7</sup> 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 Table 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 Table 1). Minor alterations to wording and some re-sequencing of items was also performed.

#### Analysis

The data were analysed by studying the responses to each item, the scale reliability was assessed with inter-item and item-total correlations, summarized 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 intraclass correlation coefficient (ICC). ICC scores above 0.40, 0.60 and 0.80 were interpreted as fair, moderate and substantial agreement respectively.<sup>8</sup> Exploratory factor analysis (oblique rotation taking eigenvalues of 1.1) was used to determine factor loadings.<sup>9</sup> Further assessment using a forced one-factor analysis (oblique rotation) was also performed.

**Table 1** The revised scale scoring guidance

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

## Results

### Sample

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.<sup>10</sup> 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 three 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 with those included for analysis. One practitioner recorded five consultations but the majority recorded eight 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;<sup>5</sup> 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 min 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 2 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 with previously published results,<sup>5</sup> these results show a greater level of scoring consistency and a lower level of missing values.

### Factor analysis

Exploratory factor analysis (oblique rotation taking eigenvalues of 1.1) was used to determine factor loadings.<sup>9</sup> 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,<sup>4</sup> 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.<sup>8</sup> Aggregating the rating scores to produce a two-point binary scale showed similar kappa values (see Table 2). The inter-rater 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.

**Table 2** Option item response (%) for two observers, Cohen's kappa and intraclass 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).

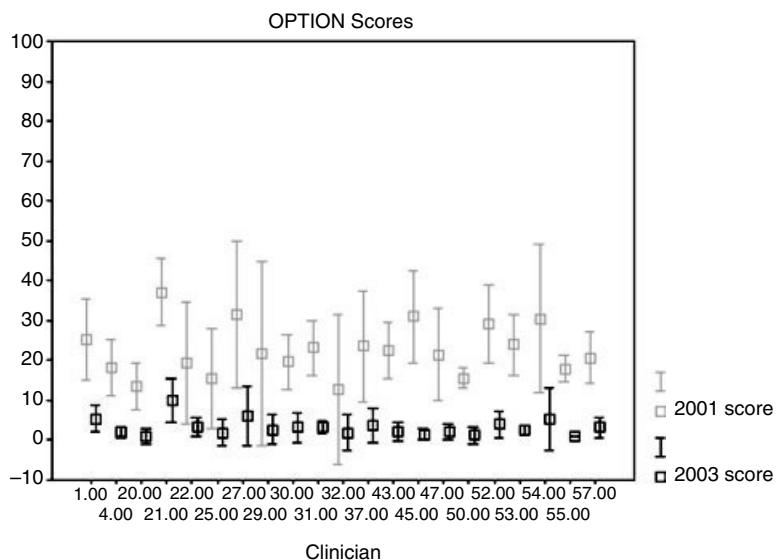
They compare well with inter-rater reliability from the 2001 data, but the overall scores for the practitioners are lower than rated previously as a result of changing from the attitude to magnitude measures (see Table 3 and Figure 1 for results and graphical representation). For all 12 items, the mean Cohen  $\kappa$  score was 0.66, indicating acceptable inter-rater agreement for this type of instrument, after correcting for chance.<sup>11</sup> 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 with previously published results,<sup>5</sup> these scores indicate a marginal improvement in reliability, identical kappa scores and an increase in inter-rater ICC from 0.62 to 0.77.

#### 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 retest 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).<sup>8</sup>

**Table 3** 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



**Figure 1** OPTION scores (2001 and 2003 scale).

**Discussion**

**Principal findings**

When compared with 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 performance. 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.<sup>10</sup> 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 2). 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 five 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 recognized 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.<sup>10</sup> Although in ideal circumstances, we would re-establish these findings 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.<sup>1</sup> 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.<sup>12</sup> 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.<sup>13</sup>

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 min. It is becoming widely recognized that achieving greater levels of patient involvement requires additional time.<sup>10</sup> 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 consultation duration (i.e. 10–12 min per consultations). It is important, however, to concede that time, although necessary, is not sufficient on its own.<sup>14</sup> 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.<sup>15</sup> Such findings are confirmed when more qualitative methods are used to analyse medical practice.<sup>16–18</sup>

### Implications

The implications of this work can be summarized 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, while the uncertainties about the outcomes of involving patients will require further investigation.<sup>19</sup> As part of this work, the OPTION tool provides a means of assessing progress.

### References

- 1 Elwyn G, Edwards A, Mowle S *et al.* Measuring the involvement of patients in shared decision making: a systematic review of instruments. *Patient Education and Counseling*, 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). *Social Science and Medicine*, 1997; **44**: 681–692.
- 3 Charles C, Gafni A, Whelan T. Decision making in the physician-patient encounter: revisiting the shared treatment decision-making model. *Social Science and Medicine*, 1999; **49**: 651–661.
- 4 Elwyn G, Edwards A, Kinnorsley P, Grol R. Shared decision making and the concept of equipoise: defining the competences of involving patients in healthcare choices. *British Journal of General Practice*, 2000; **50**: 892–899.
- 5 Elwyn G, Edwards A, Wensing M *et al.* Shared decision making: developing the OPTION scale for measuring patient involvement. *Quality and Safety in Health Care*, 2003; **12**: 93–99.
- 6 Streiner N, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*, 2nd edn. 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**: 159–174.
- 9 Jolliffe IT, Morgan BJT. Principal component analysis and exploratory factor analysis. *Statistical Methods in Medical Research* 1992; **1**: 69–75.
- 10 Elwyn G, Edwards A, Hood K *et al.* 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. *Family Practice*, 2004; **21**: 337–346.
- 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. *Journal of the American Medical Association*, 1999; **282**: 2313–2320.
- 13 Elwyn G, Gray J, Clarke A. Shared decision making and non-directiveness in genetic counselling. *Journal of Medical Genetics*, 2000; **37**: 135–138.
- 14 Freeman GK, Horder JP, Howie JGR *et al.* Evolving general practice consultation in Britain: issues of length and context. *British Medical Journal*, 2002; **324**: 880–882.
- 15 Champion P, Foulkes J, Neighbour R, Tate P. Patient centredness in the MRCGP video examination: analysis of large cohort. *British Medical Journal*, 2002; **325**: 691–692.
- 16 Gwyn R, Elwyn G. When is a shared decision not (quite) a shared decision? Negotiating preferences in a general practice encounter. *Social Science and Medicine*, 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 Expectations*, 2003; **6**: 242–254.
- 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 Expectations*, 1999; **2**: 105–117.
- 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. *Family Practice*, 2004; **21**: 347–354.