

Effect of order of presentation of a generic and a specific health-related quality of life instrument in knee and hip osteoarthritis: a randomized study¹

A. C. Rat M.D., Ph.D.^{†‡*}, C. Baumann Ph.D.[†], S. Klein B.Cs., Research Nurse[†],
D. Loeuille M.D., Ph.D.[‡] and F. Guillemin M.D., Ph.D.[†]

[†] Centre d'Epidémiologie Clinique EA 4003, CIC-EC-INSERM CIE6,
Service D'épidémiologie et évaluation Cliniques, CHU de Nancy, France

[‡] Service de Rhumatologie, CHU Nancy, France

Summary

Objective: Current guidelines recommend using both a generic and a specific instrument to measure quality of life (QoL) among people with chronic diseases. However, the two questionnaires may not be independent, which raises the issue of whether the order in which they are completed influences their value. We aimed to assess the effect of order of presentation of a generic (SF36) and a specific (Osteoarthritis Knee and Hip QoL [OAKHQOL]) QoL instrument administered to patients with knee and hip osteoarthritis (OA).

Methods: We recruited 341 outpatients from rheumatology and orthopaedic surgery clinics. Demographic and clinical data and responses to the SF36 and OAKHQOL self-administered questionnaires were collected at baseline and 10 days after inclusion; surgical patients were also assessed 6 months post-surgery. The order of presentation of the instruments was randomized at inclusion.

Results: The order of instrument presentation had no significant effect on response rate, number of dimensions with a floor or a ceiling effect, or questionnaire scores. In one of 13 dimensions (social support dimension of the OAKHQOL), test–retest reliability was slightly better when the generic SF36 questionnaire was presented first (intra-class correlation coefficient (ICC) 0.84 vs 0.55). The analysis of sensitivity to change and discriminant ability did not favour one group over the other.

Conclusion: The order of presentation of a generic and a specific QoL questionnaire to patients with OA had an only marginal effect on the quality of responses and the QoL scores obtained.

© 2007 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Key words: Methodology, Osteoarthritis, Quality of life, Psychometry.

Introduction

Current guidelines recommend using both a generic and a specific instrument to measure quality of life (QoL) among people with chronic diseases¹. In health research studies, generic questions covering broader issues are generally addressed first, followed by the more disease-specific or focused ones. The designers of the SF36 (a generic instrument) recommend presenting their tool first to be consistent with the standard followed when normative data were gathered from a general population and to allow for comparisons between data for different diseases². However, some experts recommend administering the instruments in a randomized order³.

Although the importance of the order of items within questionnaires has been documented⁴, not much is known about the influence of the order of presentation of composite instruments when several scales are included within the same package. The literature to date is inconclusive^{3,5–15}.

There are several ways in which the order of presentation might be expected to influence responses to questionnaires^{4,16–18}. For example, patient fatigue can lead to increased levels of missing data and careless answers towards the end of a package. Completing the first questionnaire may prepare respondents for the second questionnaire, thus giving them a better understanding of the fields covered. There may also be a repetition effect, that is, respondents who have answered a specific question about a particular issue may feel that they are repeating themselves if they address the same issue in response to a similar but more general question. If the specific question is asked second, it may be seen as a request for a deeper response⁴. The framing effect refers to the phenomenon whereby the meaning of a particular question varies according to the context in which it is presented^{4,16–18}. The type of questionnaire presented first may influence how the overall concept of QoL is perceived. Starting with a condition-specific questionnaire emphasizes dimensions particularly affected by the disease. The halo effect refers to the influence of a particular aspect

¹This study was supported by the Clinical Epidemiology Center, INSERM CIE6 (Institut National de la Santé et de la Recherche Médicale), Health Ministry, Nancy University Hospital, France.

*Address correspondence and reprint requests to: Dr Anne-Christine Rat, Centre d'Epidémiologie Clinique, CIC-EC-INSERM CIE6, Service D'épidémiologie et Évaluation Cliniques, Hôpital Marin-CHU de Nancy, 2 Avenue du Maréchal de Lattre de Tassigny, C.O. N°34, 54035 Nancy Cédex, France. Tel: 33-383852163; Fax: 33-383851205; E-mail: ac.rat@chu-nancy.fr

Received 20 March 2007; revision accepted 29 July 2007.

of a situation on other aspects, which suggests that categories are not perceived as distinct¹⁶.

Because QoL instruments are widely used in a number of different fields, their administration should be standardized so that results can be compared. The objective of the present study was to assess the influence of the order of presentation of a generic and a specific QoL instrument to patients with knee and hip osteoarthritis (OA) to clarify the effect of order of presentation on the quality of the information gathered and on the QoL scores obtained.

Materials and methods

SAMPLE

Patients with knee or hip OA were recruited from March 2003 to June 2005 from orthopaedic surgery clinics if they were scheduled to receive prosthetic replacement surgery within 3 months and from rheumatology outpatient clinics. Among the inclusion criteria was a diagnosis of OA according to American College of Rheumatology criteria^{19,20} and no other disabling disorder.

MEASURES

Sociodemographic and clinical data were collected at baseline. All patients were asked to complete the SF36 (generic questionnaire) and the OAKHQOL (OsteoArthritis Knee and Hip QoL: specific questionnaire) at baseline (T0) and at 10 days (T1) after inclusion; surgical patients were also assessed 6 months (T2) after intervention.

The SF36 is the most popular generic QoL instrument used for patients with OA²¹, particularly in the context of hip and knee prosthetic surgery. Its validity has been widely documented, and it has been extensively applied and adapted in numerous countries²². It contains 36 items in the following eight dimensions: physical functioning, physical role, mental health, emotional role, bodily pain, social functioning, vitality, and general health. Two established summary component scores (physical and mental) are also used here.

The OAKHQOL (2.3) is the first specific QoL questionnaire for patients with knee and hip OA. It is self-administered and comprises 43 items in five dimensions: physical activity, mental health, pain, social support, and social activities plus three independent items²³. Dimension scores are standardized from 0 (worst QoL) to 100 (best QoL).

RANDOMIZATION OF ORDER OF PRESENTATION OF INSTRUMENTS

Patients were randomized into two groups, with stratification on the joint concerned and the recruitment setting (medical or surgical clinics): the

generic/specific (G/S) group (174 patients) received the SF36 first, and the specific/generic (S/G) group (167 patients) received the OAKHQOL first. At T1, half of the patients crossed over to the other order of presentation and in half the order was unchanged. The crossover design allowed the order effect to be investigated for the same subject. At T2 (surgical patients), the order of presentation was the same as at baseline (Table I).

STATISTICAL ANALYSIS

The effect of the order of presentation was assessed by comparing the G/S and S/G responses with regard to number of missing items, number of dimensions showing a floor or ceiling effect, QoL scores, and reliability and responsiveness. Construct validity and discriminative ability were also compared.

Characteristics

Differences in characteristics between the groups were tested using the chi-square or Fisher exact test for categorical variables and Student's *t* test or Wilcoxon rank sum test for continuous variables.

Quality of the responses

The quality of responses was assessed at T0 (baseline) by comparing the number of missing items per patient in the two groups by the Wilcoxon rank sum test for each questionnaire and each dimension.

A mean shift in score potentially attributable to the order of presentation could result in a higher or lower floor or ceiling effect, which would translate into more or less responsiveness in longitudinal studies. To assess these effects, the groups were compared for the number of dimensions showing a floor (0) or ceiling (100) effect per patient on Wilcoxon rank sum testing at T0.

Within-patient order effects (e.g., differences in SF36 scores when administered first vs second for the same person) were investigated using generalized estimating equation (GEE) with T0 and T1 data. Only patients who crossed over between T0 and T1 participated in this analysis. The effect of presentation order was tested in terms of the number of missing responses and number of dimensions showing a ceiling or floor effect.

QoL scores

Differences between the G/S and S/G groups in mean SF36 and OAKHQOL scores were compared at baseline by the Wilcoxon rank sum test and by applying GEE to longitudinal data.

Test-retest validity

To determine the order of presentation that resulted in the least measurement error and to compare the reliability of the different dimensions by group,

Table I
Sociodemographic data and clinical characteristics by order of presentation group

		Mean (SD)		P	
		Group G/S (N = 174)	Group S/G (N = 167)		
Age (years)		66.2 (10.4)	65.4 (9.7)	0.31	
Charlson score	(0–34)	0.7 (0.8)	0.6 (0.7)	0.40	
Pain intensity	VAS: 0–100	57.5 (23.3)	59.6 (21.3)	0.51	
Walking distance (m)		798.8 (855.9)	1187.4 (1752.1)	0.19	
		N (%)			
Sex	Male	62 (36.3)	72 (43.4)	0.18	
Body mass index (kg/m ²)	>30	59 (34.1)	46 (28.2)	0.24	
Recruitment	Rheumatology	100	95	0.91	
	Surgery	74	72		
Osteoarthritic joint (most symptomatic)	Hip	78	78	0.62	
	Knee	85	77		
	Hip and knee	11	12		
Kellgren score	3 and 4	69 (76.7)	66 (71.0)	0.38	
T0	Baseline	174	167		
T1	10 days	G/S then G/S	G/S then S/G	S/G then S/G	S/G then G/S
		65	65	68	53
T2	6 months	28	30		

Group G/S: generic questionnaire (SF36) presented first; group S/G: specific questionnaire (OAKHQOL) presented first; SD: standard deviation; VAS=visual analog scale; T0 = baseline for all patients; T1 = 10 days after inclusion for all patients; T2 = survey of surgical patients 6 months after surgery.

ICCs between T0 and T1 were compared after Fisher's z transformations. Only data for patients tested in the same order at T0 and T1 were analyzed.

Construct validity

The type of questionnaire presented first can influence how the broad concept of QoL is perceived and may affect correlations between QoL scores and clinical characteristics. For example, when the first tool used to measure QoL is condition-specific, dimensions linked to the impact of the disease are likely to carry more weight. To test whether the correlations between OAKHQOL scores and clinical characteristics (visual analogue scale [VAS] scores for pain and walking distance) and between OAKHQOL scores and SF36 scores were higher when the specific (OAKHQOL) instrument was presented first rather than second, Pearson correlation coefficients were compared after Fisher's z transformations.

Discriminative ability

The abilities of the two orders of presentation to discriminate between rheumatology and surgery patients were compared at baseline by regression analysis.

Sensitivity to change

To determine which order of presentation was most likely associated with a treatment effect, the responsiveness of the different dimensions between T0 and T2 was analyzed by standardized response means (SRMs). The distribution of the SRM of each dimension was assessed with a bootstrap procedure (100 iterations), and data for the two groups were compared by Student's t test.

Statistical analyses involved use of SAS, version 9.1 (SAS Institute Inc., 2002). A P value of less than 0.05 was considered significant.

Results

CHARACTERISTICS

A total of 341 patients were included in the study. Their mean age was 66 years, and 60% were women; 51% of the joints affected were knees; 43% of patients were seen just before surgery, and the remaining 57% were seen in

rheumatology clinics. Table I summarizes the demographic and clinical characteristics of the two groups of patients; none of the differences were significant.

The second questionnaire was sent by mail at T1 (10 days after inclusion), and 26% of the patients did not return this questionnaire. Significantly more patients recruited in surgery did not return the second questionnaire (33% surgical vs 22% nonsurgical patients). Only scores for social support (OAKHQOL) and role physical (SF36) dimensions differed between responders and nonresponders after adjusting for surgical/nonsurgical status, with no trend observed.

QUALITY OF THE RESPONSES

No significant differences were observed in number of missing items and the number of dimensions showing a floor (0) or ceiling (100) effect at baseline (Table II). Table II also shows the results of the GEE analyses of within-patient effects. The regression coefficients describe the relation between the order of presentation and the number of missing items or the number of dimensions with a ceiling or floor effect on the basis of the within-person analysis. No association was found other than in the number of missing items. However, this model showed a significant carryover effect (the interaction between order of presentation and sequence [G/S then S/G or S/G then G/S]) that rendered the GEE analysis uninterpretable. If an interaction is significant, the difference in the number of missing items between the groups differed according to the sequence, and only the baseline analysis is interpretable.

QoL SCORES

QoL dimension scores did not differ significantly between the two groups at baseline (Table III). In addition to statistical results, we show a difference between the two order

Table II
Comparison of number of missing items and number of dimensions with floor or ceiling effect by order of presentation group for each instrument

	OAKHQOL (43 items – five dimensions)			SF36 (36 items – eight dimensions)		
	Med	Mean (SD)	P	Med	Mean (SD)	P
Baseline						
<i>Number of missing items</i>						
Group G/S (N = 174)	0	1.3 (4.0)	0.32	0	0.9 (2.3)	0.19
Group S/G (N = 167)	0	1.1 (2.2)		0	0.8 (3.5)	
<i>Number of dimensions with ceiling effect</i>						
Group G/S	0	0.1 (0.3)	0.91	0	1.0 (1.0)	0.54
Group S/G	0	0.1 (0.3)		0	1.0 (1.0)	
<i>Number of dimensions with floor effect</i>						
Group G/S	0	0.4 (0.7)	0.94	1	0.6 (1.0)	0.92
Group S/G	0	0.3 (0.6)		1	0.6 (0.9)	
	β	95% CI	P**	β	95% CI	P**
<i>Change (T0 – T1)*</i>						
G/S then S/G (N = 55) vs S/G then G/S (N = 67)						
Number of missing items	-0.4	-1.0-0.2	0.22	-1.6	-3.1-0.3	0.02†
Number of dimensions with ceiling effect	0.2	-0.1-0.5	0.21	0.1	-0.2-0.4	0.51
Number of dimensions with floor effect	-0.2	-0.7-0.3	0.40	0.0	-0.2-0.2	0.92

Med = median.

*Regression analysis of an order effect within-patients by GEE; β the difference in number of missing items or dimensions with a ceiling or floor effect between T0 and T1 based on within-patient differences controlling for the sequence (G/S then S/G or S/G then G/S).

†Carryover effect (interaction between order of presentation and sequence) P = 0.0014; all the interactions for the other models were nonsignificant.

**P results of the test for the order effect.

Table III
Comparison of mean QoL dimension scores by order of presentation

	Baseline difference (G/S group S/G group)			Change (T0 – T1)*		
	Means	95% confidence intervals	P	β	Order effect, P	Carryover effect [†] , P
OAKHQOL						
Physical activities	-1.4	-6.4–3.6	0.58	-0.01	0.66	0.19
Mental health	1.5	-3.7–6.7	0.68	-0.01	0.67	0.92
Pain	-2.0	-7.6–3.6	0.49	-0.04	0.29	0.67
Social support	2.4	-2.8–7.5	0.42	-0.06	0.05	0.02
Social activities	-0.2	-5.8–5.5	0.92	-0.08	0.19	0.01
SF36						
Physical functioning	3.5	-1.5–8.5	0.17	0.05	0.22	0.60
Physical role	2.9	-5.3–11.1	0.48	0.12	0.18	0.08
Mental health	0.6	-3.8–4.9	0.76	-0.06	0.04	0.03
Role emotional	2.8	-6.5–12.1	0.46	-0.03	0.76	0.08
Bodily pain	-2.4	-6.7–1.9	0.17	-0.04	0.27	0.82
Social functioning	-3.2	-8.3–1.9	0.36	-0.02	0.51	0.18
Vitality	-0.3	-4.3–3.7	0.96	-0.00	0.95	0.34
General health perception	0.1	-3.9–4.1	0.75	-0.04	0.06	0.77
Physical component summary	0.3	-1.4–2.1	0.69	0.01	0.52	0.81
Mental component summary	-0.5	-2.9–2.0	0.78	-0.06	0.01	0.02

The scores of QoL questionnaires range from 0 (worst) to 100 (best QoL).

*Regression analysis of an order effect within-patients by GEE, β difference in QoL scores between T0 and T1 based on within-patient differences controlling for the sequence (G/S then S/G or S/G then G/S).

†Carryover effect: interaction between order of presentation and sequence.

groups of less than five points, which is considered by the developers of the SF36 to be under the threshold of a meaningful difference²².

For the SF36, with the exception of the mental component summary, QoL scores did not differ when the questionnaire was administered first vs second to the same person. However, the OAKHQOL social support and social activity scores and SF36 mental health and mental component summary scores showed a carryover effect. In other words, QoL

differences between the order groups varied according to the sequence (G/S then S/G or S/G then G/S), and for these dimensions, only the baseline analysis is interpretable.

TEST-RETEST VALIDITY

With the exception of the OAKHQOL social support dimension, the reliability of which was better when the generic SF36 questionnaire was presented first, the two order

Table IV
Comparison of test–retest reliability and responsiveness after surgery according to the order of presentation group

	Reproducibility (T0 – T1)*		Responsiveness after surgery [†]			
	G/S group then G/S group (N = 65)	S/G group then S/G group (N = 68)	Group G/S (N = 28)		Group S/G (N = 30)	
	ICC	ICC	SRM	(95% CI)	SRM	(95% CI)
OAKHQOL						
Physical activities	0.87	0.83	1.09	1.05–1.13	1.08	1.04–1.12
Mental health	0.86	0.84	0.79	0.75–0.82	0.78	0.75–0.82
Pain	0.85	0.87	1.11	1.06–1.16	1.25	1.21–1.29**
Social support	0.84	0.55*	0.09	(0.05–0.13)	0.07	0.03–0.11
Social activities	0.73	0.76	0.31	(0.26–0.35)	0.33	0.29–0.37
SF36						
Physical functioning	0.79	0.73	0.65	0.60–0.79	0.81	0.76–0.85**
Physical role	0.72	0.61	0.64	0.61–0.68	0.57	0.54–0.60**
Mental health	0.77	0.71	0.62	0.57–0.66	0.60	0.56–0.46
Role emotional	0.70	0.69	0.45	0.42–0.49	0.42	0.38–0.46
Bodily pain	0.78	0.82	1.11	1.07–1.14	1.20	1.16–1.23**
Social functioning	0.66	0.68	0.76	0.72–0.80	0.33	0.30–0.37**
Vitality	0.73	0.77	1.20	1.15–1.26	0.85	0.81–0.88**
General health perception	0.81	0.83	0.53	0.48–0.57	0.54	0.50–0.56
Physical component summary	0.80	0.85	0.91	0.86–0.95	0.95	0.90–0.99
Mental component summary	0.78	0.73	0.60	0.56–0.64	0.40	0.35–0.44**

*Test–retest reliability was analyzed for nonsurgical and surgical patients at T0 and T1 and compared with the z test after Fisher's z test transformations of the ICC.

†Responsiveness after surgery compared between groups by Student's t test after assessment of the distribution of the SRM with a bootstrap procedure. *P < 0.05; **P < 0.0001.

Table V
Comparison of construct validity by order of presentation group at baseline

	Group G/S (N = 174), R	Group S/G (N = 167), R
Physical activities (OAKHQOL/SF36)	0.68	0.63
Physical activities (OAKHQOL)/walking distance	0.42	0.35
Mental health (OAKHQOL/SF36)	0.77	0.79
Pain (OAKHQOL)/VAS pain	0.41	0.29
Pain (OAKHQOL/SF36)	0.68	0.67
Social activities (OAKHQOL/SF36)	0.40	0.38

R = correlation coefficients between QoL scores and between QoL scores and clinical indices. Correlation coefficients were compared with the z test after Fisher's z test transformations of the coefficients. None of the differences between correlation coefficients were significant.

groups did not differ in reproducibility of the dimensions between T0 and T1 (Table IV).

CONSTRUCT VALIDITY

Correlations between OAKHQOL scores and clinical variables (VAS pain and walking distance scores) or SF36 scores were not significantly influenced by the order of instrument presentation (Table V).

DISCRIMINATIVE ABILITY

The order of instrument presentation had no effect on discriminative ability of the instruments (Table VI).

SENSITIVITY TO CHANGE

SRMs for the OAKHQOL and the SF36 pain scores and the SF36 physical functioning score were significantly higher in the S/G than G/S group and those for the SF36 physical role, social functioning, vitality and mental component summary scores were higher in the G/S than S/G group (Table IV). However, the differences between the SRMs were small, with no significant differences observed in any other dimension.

INFLUENCE OF LEVEL OF EDUCATION OR STAGE OF OA (SURGICAL OR NONSURGICAL) BY THE ORDER OF INSTRUMENT PRESENTATION

An effect of order of instrument presentation was noted only for patients with a primary or university level of education for the SF36 role emotional dimension in baseline analysis and for physical activities and general health dimensions in the within-patients analysis. The test-retest reliability between the two order groups for the OAKHQOL social support dimension differed only for patients with a primary level of education.

At baseline, the stage of OA had no influence on the order effect. Only nonsurgical patients showed a carryover effect for the OAKHQOL social dimension scores and SF36 mental health dimension score. As well, only nonsurgical patients showed an order effect for the within-patients analysis in the SF36 bodily pain dimension. Whatever the stage of OA, the test-retest reliability differed between the two order groups for the OAKHQOL social support dimension.

Discussion

This randomized study of the order of presentation of a generic and a specific questionnaire in hip and knee OA

Table VI
Comparison of the discriminant abilities of the two order groups between nonsurgical and surgical patients

	Group G/S (N = 174)		Group S/G (N = 167)		P [†]
	Mean R - S*	SD	Mean R - S*	SD	
OAKHQOL					
Physical activities	18.8	21.9	16.5	21.9	0.55
Mental health	4.0	23.5	3.3	24.8	0.57
Pain	12.6	24.5	9.0	27.1	0.46
Social support	-9.1	23.1	-4.5	24.7	0.36
Social activities	5.9	26.3	6.3	26.1	0.95
SF36					
Physical functioning	14.0	22.7	12.7	22.1	0.17
Physical role	21.5	37.5	21.3	36.7	0.48
Mental health	2.9	19.9	6.9	20.5	0.79
Role emotional	16.7	42.1	25.6	41.9	0.55
Bodily pain	9.2	19.8	10.7	19.3	0.25
Social functioning	0.0	24.9	3.2	23.2	0.22
Vitality	5.3	18.0	8.3	18.8	0.89
General health perception	-5.5	18.2	1.0	18.8	0.96
Physical component summary	4.0	8.0	3.7	7.6	0.64
Mental component summary	1.2	11.5	4.3	10.7	0.73

*Mean R - S is the difference in QoL scores between nonsurgical and surgical patients. The QoL scores range from 0 (worst) to 100 (best QoL).

†Discriminant abilities of the two groups were compared by regression analysis.

revealed essentially no effect of instrument order on QoL measurement. The order of presentation influenced neither the item response rate nor the number of dimensions with a floor or ceiling effect. In only one dimension, social support was reliably better in the group given the generic questionnaire first. Sensitivity to change, construct validity and discriminant ability was unaffected.

Although the order of items within a questionnaire has been consistently shown to be important⁴, not much is known about the impact on results of variations in the presentation of composite measurement scales within the same package. Moreover, the results of different studies are inconsistent. Investigations in the field of psychology have demonstrated that changing the order of measurement scales within the same package can alter the pattern of responses^{5,6,12}. In a study comparing patients' responses to questions about their disability with and without specification of the affected area, the order of presentation influenced the results⁷. In another randomized study, the scores on a general-assessment VAS were influenced by its presentation relative to the presentation of the SF36¹¹.

Studies of QoL instruments have yet to show any effect of order of their presentation. In a study of HIV patients, administration of the EuroQol (generic questionnaire) before the medical outcomes study HIV health survey resulted in higher mean scores than when the order was reversed, but no differences were significant⁸. In another randomized study, the effect of order was assessed when two generic instruments and three specific QoL instruments for asthma and angina were administered. Neither response rate nor response speed was affected. A few dimension scores of the SF36 (social functioning, mental health, general health perception and vitality) were affected by the order of presentation (scores were lower when the generic questionnaire was presented first), but the results were inconsistent across conditions and scales and the differences were small. The internal consistency of the scales remained stable, and the correlation between the generic and specific scores showed a slight trend towards greater strength when the condition-specific measure was presented first⁹. In a study of 190 patients with cancer, two specific QoL instruments appeared to perform equally well, whatever their sequence of presentation within a questionnaire package¹⁰.

However, in a study of patients with hypertension, the SF36 mean emotional role and mental component summary scores were higher when the specific instrument was completed first rather than second¹⁴. A randomized study looking at the SF36, the Hospital Anxiety and Depression Scale (HADS) and vision-targeted instruments revealed that changing the order of administration resulted in significant variations in scores for the SF36 mental health dimension and the HADS depression and anxiety subscales: all scores indicated poorer mental health status when generic instruments followed vision-targeted instruments; however, the differences were small and not clinically relevant¹⁵. Finally, a study of four questionnaires presented at different times relative to the period to which they referred found no effect on answers, but response rates were higher and questionnaires were returned more quickly¹³. Thus, in our study as in other studies, in terms of consistency and clinical relevance, no QoL score was influenced by the order of presentation of questionnaires.

Since responses to generic and specific QoL questionnaires do not appear to be strongly affected by their order of presentation, self-completed instruments may allow respondents to review the entire package before responding, with no pressure to present a consistent image to an

interviewer^{9,13}. Provision of clear instructions helps patients stay concentrated and motivates them to complete items as accurately as possible. Some factors assumed to be affected by order (framing effect, halo effect, repetition effect, fatigue) have opposite influences and cancel out each other. Patients ready for surgery may be so highly focused on their joint problems that the order of presentation does not matter. The same may be true for rheumatology patients, because OA of the lower limbs has an important impact on QoL. Because the frequency of comorbidities was low in our study (mean Charlson index of 0.6), the general QoL of respondents probably related mainly to their disease.

One limitation of the present investigation is that its results are difficult to generalize fully. Order of instrument presentation had no effect on QoL of our patients with OA but may be important when longer or different questionnaires are involved and other pathologies are assessed. For example, psychiatric patients tend to be susceptible to loss of concentration. Therefore the effect of presentation order should be confirmed in any given situation.

As well, questionnaires must undergo a rigorous development process and a thorough analysis of their validity before considerations such as order of presentation are addressed.

Furthermore, patients recruited for our study were representative of the knee and hip OA population consulting in rheumatology or in orthopaedic surgery but were not of the general population. Therefore, our findings cannot be fully generalizable to a population-based sample of the disease but are valid in most situations involving the two questionnaires. Finally, a larger sample size might have resulted in a small effect of order, if any exists, but would have been of low clinical significance.

Conclusions

The order of presentation of a generic and a specific QoL questionnaire to patients with OA had only a marginal effect on the quality of responses and the QoL scores obtained. Because a large number of studies have involved the interchangeable use of generic and specific instruments, these results are reassuring by showing that their order of presentation does not affect the results. But when the aim is to compare patients across diseases or against reference populations, starting with the generic questionnaire would follow how normative data are gathered and minimize differences in a framing effect seen in results across different diseases.

Acknowledgements

The authors would like to acknowledge M Alfonsi, F Arnould, X Ayrat, P Bouillot, I Cahitte, I Chary-Valckenaere, P Chaspoux, P Fener, H Gaudin, MF Gérard, F Goupy, Le Maitre, J P Gros, D Hannouche, E Hiraux, N Jeanson, JM Lardry, D Nebout, H Ouakil, S Perrot, S Poir-audeau, J Pourcel, P Prost, P Rabany, M Rousseau, P Thomas, F Touzard, G Vançon, L Vastel, and JP Voilquin.

References

1. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989;27:S217–32.

2. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998;51:903–12.
3. Barry MJ, Walker-Corkery E, Chang Y, Tyll LT, Cherkin DC, Fowler FJ. Measurement of overall and disease-specific health status: does the order of questionnaires make a difference? *J Health Serv Res Policy* 1996;1:20–7.
4. Sudman S, Bradburn N. *Asking Questions. A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass 1982.
5. Jensen PS, Watanabe HK, Richters JE. Who's up first? Testing for order effects in structured interviews using a counterbalanced experimental design. *J Abnorm Child Psychol* 1999;27:439–45.
6. Lucas CP. The order effect: reflections on the validity of multiple test presentations. *Psychol Med* 1992;22:197–202.
7. Marx RG, Hogg-Johnson S, Hudak P, Beaton D, Shields S, Bombardier C, *et al*. A comparison of patients' responses about their disability with and without attribution to their affected area. *J Clin Epidemiol* 2001;54:580–6.
8. Wu AW, Jacobson DL, Berzon RA, Revicki DA, van der Horst C, Fichtenbaum CJ, *et al*. The effect of mode of administration on medical outcomes study health ratings and EuroQol scores in AIDS. *Qual Life Res* 1997;6:3–10.
9. McColl E, Eccles MP, Rousseau NS, Steen IN, Parkin DW, Grimshaw JM. From the generic to the condition-specific?: Instrument order effects in Quality of Life Assessment. *Med Care* 2003;41:777–90.
10. Cheung YB, Wong LC, Tay MH, Toh CK, Koo WH, Epstein R, *et al*. Order effects in the assessment of quality of life in cancer patients. *Qual Life Res* 2004;13:1217–23.
11. Briancon S, Empereur F, Guillemin F, Hercberg S. Effect of VAS orientation and ordering on general perception of health within a quality of life questionnaire. *Qual Life Res* 1999;8:650.
12. Laberge M, Fournier S, Freeston MH, Ladouceur R, Provencher MD. Structured and free-recall measures of worry themes: effect of order of presentation on worry report. *J Anxiety Disord* 2000;14:429–36.
13. Dunn KM, Jordan K, Croft PR. Does questionnaire structure influence response in postal surveys? *J Clin Epidemiol* 2003;56:10–6.
14. Leidy NK, Legro M, Schmier JK, Zyczynski T, McHorney CA, Coyne K. Does order of administration influence subject response on generic and condition-specific measures of health related quality of life? *Qual Life Res* 2000;9:337.
15. Childs AL. Effect of order of administration of health-related quality of life interview instruments on responses. *Qual Life Res* 2005;14:493–500.
16. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to their Development and Use*. 2nd edn. New York: Oxford University Press; 1995.
17. Guyatt GH, Cook DJ, King D, Norman GR, Kane SL, van Ineveld C. Effect of the framing of questionnaire items regarding satisfaction with training on residents' responses. *Acad Med* 1999;74:192–4.
18. Armstrong K, Schwartz JS, Fitzgerald G, Putt M, Ubel PA. Effect of framing as gain versus loss on understanding and hypothetical treatment choices: survival and mortality curves. *Med Decis Making* 2002;22:76–83.
19. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, *et al*. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;34:505–14.
20. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, *et al*. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;29:1039–49.
21. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. *J Bone Joint Surg Am* 2004;86:963–74.
22. Ware J, Kosinski M, Keller S. *SF-36 Physical and Mental Health Summary Scale: A User's Manual*. Boston, MA: The Health Institute; 1994.
23. Rat AC, Coste J, Pouchot J, Baumann M, Spitz E, Retel-Rude N, *et al*. OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis. *J Clin Epidemiol* 2005;58:47–55.

Appendix

Investigators

Osnowycz M.D.	Service de Rhumatologie, CHU Nancy, France
D. Mainard M.D.	Service de Chirurgie Orthopédique et Traumatologique, CHU Nancy, France
J. P. Delagoutte M.D.	Service de Chirurgie Orthopédique et Traumatologique, CHU Nancy, France
D. D. Fisher M.D.	Service de Rhumatologie, Hôpital Avicenne, Bobigny, France
C. Cuny M.D.	Service de Chirurgie Orthopédique et Traumatologique, CHR Metz-Thionville, France
R. Nizard M.D.	Service D'orthopédie, Hôpital Lariboisière, Paris, France
J. S. Le Quintrec M.D.	Service D'orthopédie, Hôpital Cochin, Paris, France
T. Bégué M.D.	Service D'orthopédie Hôpital Avicenne, Bobigny, France