

Measuring patient-centredness, the neglected outcome in fertility care: a random multicentre validation study

Inge W.H. van Empel^{1,*}, Johanna W.M. Aarts¹, Ben J. Cohlen², Dana A. Huppelschoten³, Joop S.E. Laven⁴, Willianne L.D.M. Nelen¹, and Jan A.M. Kremer¹

¹Department of Obstetrics and Gynaecology, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands ²Fertility clinic, Isala Clinics, Zwolle, The Netherlands ³Department of Obstetrics and Gynaecology, Jeroen Bosch Ziekenhuis, 's-Hertogenbosch, The Netherlands ⁴Division of Reproductive Medicine, Department of Obstetrics and Gynaecology, Erasmus MC, The Netherlands

*Correspondence address: E-mail: i.vanempel@obgyn.umcn.nl

Submitted on May 11, 2010; resubmitted on July 19, 2010; accepted on July 22, 2010

BACKGROUND: High-quality fertility care should be effective and safe, but also patient-centred. However, a suitable instrument for measuring patient-centredness is lacking. This study aims to develop and validate an instrument that can reliably measure patient-centredness in fertility care: patient-centredness questionnaire-infertility (PCQ-infertility).

METHODS: The PCQ's content, addressing 53 care aspects, was generated by seven focus groups with 54 infertile patients. Besides background questions, the questionnaire included one 'experience item' and one 'importance item' for each care aspect. Thirty Dutch fertility clinics were invited to participate in the validation study. The questionnaire was sent at random to 1200 infertile couples. Psychometric tests included inter-item and reliability analyses. Importance scores were calculated. The discriminative power was determined using multilevel analysis.

RESULTS: The questionnaire was completed by 888 infertile couples (net response 75%) from 29 clinics. The ultimate PCQ-infertility, comprising 46 items and seven subscales, appeared reliable and valid for measuring patient-centredness in fertility care. Of the seven subscales, 'communication' received the best ratings and 'continuity' the worst. 'Honesty and clearness on what to expect from fertility care' appeared most important to patients. Significant differences between clinics were found, even after case-mix adjustment.

CONCLUSION: This study resulted in a valid, reliable and strongly discriminating instrument for measuring patient-centredness in fertility care. The PCQ-infertility can identify shortcomings on patient-centredness and can be adopted for quality improvement. Therefore, fertility care can now be monitored and benchmarked on patient-centredness, as well as on live birth and complication rates.

Key words: patient-centredness / benchmarking / fertility care / measurement instrument / surveys

Introduction

Integrating all elements of high-quality care into daily care is one of the challenges health-care providers face today. Core elements, such as (cost-)effectiveness and safety, but also patient-centredness should be integrated to accomplish the best possible emotional and physical health in each patient (Institute of Medicine, 2001; Bengoa *et al.*, 2006; Keirns and Goold, 2009). Patient-centred care, which is guided by patients' values and is responsive to individual patients' needs, will bring patients many benefits (Institute of Medicine, 2001). It enables them to be heard and their ideas, concerns and expectations to be addressed (Clark, 2008) eventually leading to positive care

experiences. Patient-centred care could also contribute to better cooperation between patients and care providers, which would reduce misunderstandings, complaints and litigation, and make the health-care system more cost-effective (Patwardhan and Patwardhan, 2009).

In reproductive medicine, quality measures mainly concentrate on effectiveness (e.g. pregnancy rates) and safety (e.g. frequency of multiples), while patient-centredness is neglected (Gunby *et al.*, 2008; van Empel *et al.*, 2008; Nyboe Andersen *et al.*, 2009). Although infertile couples experience many weaknesses and needs in their care (van Empel *et al.*, 2010a), patient-centredness is increasingly recognized as important for the quality of reproductive medicine

(Dancet *et al.*, 2010). Given the high dropout rates together with substantial physical and emotional burden of fertility treatments, infertile couples would particularly benefit from care tailored to their individual needs (Gerteis *et al.*, 1993; Verhaak *et al.*, 2007; Verberg *et al.*, 2008).

Patient-centredness is ideally monitored by surveys measuring patients' specific experiences, rather than by surveys measuring global satisfaction (Cleary, 1999; Wensing and Elwyn, 2003; Patwardhan and Patwardhan, 2009). For reliably monitoring and benchmarking patient-centredness in fertility care, a validated measurement instrument is needed which is appropriate for patients with all kinds of medically assisted reproduction (MAR) and applicable to all sorts of fertility clinics (Dancet *et al.*, 2010; van Empel *et al.*, 2010a). To date, however, such an instrument does not exist.

Therefore, this study aims at developing a valid and widely usable instrument [patient-centredness questionnaire-infertility (PCQ-infertility)] that can (1) reliably measure patient-centredness in fertility care, and (2) discriminate between the extent of patient-centredness between fertility clinics.

Materials and Methods

For the development of the PCQ-infertility, qualitative methods (focus groups) and quantitative methods (validation survey) were used, both supported by a literature study.

Focus groups

Patients' preferences are best elicited by focus groups (Mullen, 1999). We organized focus groups with infertile patients to conceptualize patient-centredness within the infertility context and to generate questionnaire items. This strongly contributes to the new measurement instrument's content validity. For obtaining a varied, representative focus group sample, both childless couples and couples with offspring were invited. A total of 24 couples and six additional women were recruited, originating from 13 fertility clinics situated in three Dutch regions (East, West and North). Patients were subdivided into seven focus group discussions, which were conducted by three researchers (I.W.H.v.E., D.A.H. and W.L.D.M.N.) in autumn 2008. All participants were undergoing or had completed MAR. Focus groups were moderated using the Picker Institute's established general model of patient-centredness (www.pickerinstitute.org) comprising eight domains: accessibility; information, communication and education; involvement of family and friends; respect for patients' values; coordination and integration; continuity and transition; physical support; and emotional support. To elicit care aspects important to patients and discover what 'patient-centred fertility care' implies, patients' positive and negative care experiences were discussed using open-ended questions. Patients were also asked to complete a short questionnaire on demographics (e.g. age and obstetric history).

Focus groups discussions lasted 2½ h on average. All were recorded and transcribed verbatim. Transcripts were thematically analysed by two researchers (I.W.H.v.E. and D.A.H.) independently and discussed among them to increase coding reliability. A third researcher (W.L.D.M.N.) reviewed the identified care aspects to ensure consistency with the original data. Differences in interpretation between researchers were small and consensus was mostly promptly achieved. Finally, 729 relevant quotes were extracted from the transcripts. Quotes were grouped into 81 care aspects that together constituted the concept 'patient-centred fertility care'.

Questionnaire development

From the 81 care aspects, 53 were selected for the pilot version of the PCQ-infertility, based on their frequency and intensity in the focus groups (Krueger, 1998). Before the remaining care aspects were converted into questionnaire items, the structure of several questionnaire families had been studied (Hays *et al.*, 1999; Jenkinson *et al.*, 2002; Stubbe *et al.*, 2007a,b; Singer *et al.*, 2010). Then, two researchers (I.W.H.v.E. and J.W.M.A.) independently formulated one 'experience item' and one 'importance item' for each remaining care aspect. Discussion between three researchers (I.W.H.v.E., J.W.M.A. and W.L.D.M.N.) led to consensus on the best items formulations. Since the aim was to develop a manageable questionnaire that is easy to complete for most fertility patients and that does not include 'skip items', we chose to tailor the questionnaire to couples instead of to women and men separately. To facilitate patients in answering the questions, the best-fitting answer category per item was chosen. For the 53 experience items four answering formats were selected: (a) no, yes (nine items); (b) never, sometimes, usually, always (19 items); (c) definitely no, somewhat no, somewhat yes, definitely yes (eight items); and (d) no, yes but insufficiently, yes definitely (11 items). Six items received answer categories tailored to that specific question. All importance items had the same format ('*how important did you find it having ...?*') and same answer categories (not important, fairly important, important and extremely important). For the questionnaire's order of items, the patient's care pathway was followed. Items on diagnostics thus came before items on treatment. For describing the study population and examining case-mix differences, 20 questions on patients' background were added to the questionnaire, such as age, ethnic background and treatment type.

The draft PCQ-infertility was pretested among 15 infertile couples and five care professionals (gynaecologists, fertility nurses, psychologist) and consequently some last alterations were made. The pilot version of the PCQ-infertility consisted of 127 items: 53 items on patient's experiences regarding patient-centred care aspects; 53 items about the importance patients assigned to the questioned care aspects; 20 background questions; and one satisfaction mark (range 0–10) to express patients' global satisfaction with care. The questionnaire's final page was reserved for written comments about patients' personal experiences with the clinic and for suggestions to improve the questionnaire.

Data collection

A total of 30 fertility clinics in the northern, eastern and western parts of the Netherlands were invited by three regional coordinating gynaecologists (B.J.C., J.A.M.K. and J.S.E.L.) to participate in the validation study. After gaining participation approval, clinics were asked to extract the address files of all patients who underwent MAR in their clinic between April and June 2009 from their diagnosis treatment combination coding system. Patient data were entered in an excel database and duplicates removed. From the database including 3061 individual patient couples, a random sample of 1200 couples was taken. The number of sampled couples per clinic depended on the size of their infertility outpatient clinic, ranging from 25 couples for smaller clinics to 75 for the largest IVF-centres. The 1200 couples were sent the pilot PCQ-infertility between July and September 2009. Since 11 questionnaire packages were returned unopened, probably because of wrong addresses, 1189 couples received a questionnaire package. The questionnaire was accompanied by instructions, a refusal form and a postage-paid return envelope. Couples were asked to complete the questionnaire together. Participation in the survey was voluntary and anonymity was guaranteed. In the Netherlands, institutional ethics committee approval was not required for this survey. All couples were sent a reminder card 3 weeks following the initial mailing. Subsequently, 2 weeks later non-responders received a reminder

with a copy of the questionnaire. Data of incoming questionnaires were entered into SPSS (Statistical Package for the Social Sciences: version 16.0 for Windows[®], SPSS Inc., Chicago, IL, USA).

Analyses

The aim was obtaining a measurement instrument that: (1) is feasible, reliable and valid; (2) can identify fertility care's most important weaknesses according to patients; and (3) can discriminate between patient-centredness in different fertility clinics. Hence, respectively (1) the PCQ's psychometric properties, (2) quality improvement scores, and (3) the PCQ's discriminative power were determined.

Psychometric properties

The PCQ's feasibility, reliability and validity we assessed by testing the (a) appropriateness of items; (b) internal consistency; and (c) construct validity.

Appropriateness of items

First, negatively posed items (Q6, Q7, Q32, Q47, Q48, Q49 and Q52) were mirrored. For each care aspect the experience score (0 = most negative, 3 = most positive), importance score (0 = not important, 3 = extremely important) and proportion negative experiences (percentage of respondents with an experience score of 0 or 1) was calculated. Subsequently, patients' written comments were analysed. When many comments were made regarding a certain item, rephrasing or exclusion of the item was considered. Furthermore, items selected for omission were (1) extremely skewed items (>90% in extreme answer category); (2) items with a high non-response (>5% missing values); (3) relatively unimportant items (importance score <1.5); and (4) redundant items (Pearson's ρ between two items >0.80).

Internal consistency

Then, guided by the Picker model of patient-centredness, the internal consistency of the total scale and subscales was assessed by computing Cronbach's alpha coefficients and item-total correlations (ITCs). Alphas from 0.70 and higher were aspired; scales with alphas lower than 0.60 were considered unacceptable. Items not contributing to subscale reliability (ITC > 0.20) were omitted (Nunnally, 1978; Kline, 1986). Furthermore, we confirmed whether each item was in the right subscale by correlating items with the subscale means. Items that correlated more highly on subscales other than the one to which it was assigned were displaced if plausible, and otherwise eliminated (Streiner and Norman, 2008). Then, subscales with their items have been established. For patient-centredness (total scale) and each reliable subscale, a mean score was calculated (range 0–3) by summing up the responses to the individual items and dividing these sum scores by the number of items filled in. Patients who filled out half or less of the items within a subscale were excluded from further analyses of that subscale.

Construct validity

To assess the questionnaire's construct validity within infertile couples, the following hypotheses were tested, based on previous studies within the fertility care context (Hammarberg et al., 2001; Mourad et al., 2009; van Empel et al., 2010a,b): (1) patients who experience more patient-centredness are more satisfied with their care; (2) each instrument's subscale aims at measuring a part of the same construct (patient-centredness) and is therefore positively and significantly correlated with other subscales; (3) patients who had (a) access to their medical records, (b) a lead physician, (c) received written information and (d) scheduled treatment evaluations are more positive regarding the patient-centredness of their care than patients without these conditions; (4) patients who achieved pregnancy have more positive experiences regarding patient-centred care;

and (5) patients receiving assisted reproductive technology (ART) are more positive regarding the patient-centredness perceived than patients receiving non-ART treatments like intrauterine insemination.

Finally, the ultimate PCQ-infertility was reciprocally converted from Dutch into English by a bilingual translator.

Quality improvement scores

To identify aspects of patient-centred care that have priority for improvement, quality improvement scores (QI scores) were calculated. This score represents the maximum mean score of 3 – the perceived mean experience on a care aspect, multiplied by the importance score of the same care aspect (range 0–3). Consequently, QI scores could vary from 0 to 9; the higher the score, the more need there is for improvement.

Discriminative power

An elaborate multivariate multilevel regression analysis was performed with two purposes in mind: (1) to assess the PCQ's ability to measure differences in patient-centredness between fertility clinics (benchmark capability) and (2) to evaluate if case-mix adjustment is necessary when measuring patient-centredness. First, correlation analyses were performed to evaluate co-linearity between patients' background characteristics using a non-parametric correlation coefficient (Spearman's ρ). In case of two strongly correlating variables ($\rho > 0.40$), the clinically most relevant characteristic was kept. Secondly, univariate multilevel regression analyses were performed with remaining variables on patient characteristics and (sub)scale mean scores. Characteristics with $P < 0.20$ in the univariate analysis were allowed in the multivariate regression model. Subsequently, a multivariate multilevel analysis with manual backward elimination was performed using the remaining patient characteristics. Two nested models were fitted to the data. The first model was a random-intercept model without explanatory variables (0-model). Characteristics were entered and fixed in the final model; P -values of <0.05 were considered statistically significant. Separate multilevel analyses were performed for the total scale and its reliable subscales. To assess how much variance in each 0-model is attributable to differences in patient characteristics (case-mix), the proportional change in variance (PCV) was calculated according to Merlo et al. (2005).

Per clinic, case-mix adjusted mean dimension scores were calculated using a general linear model (univariate). To determine any between-clinic differences on patient-centredness, one-way ANOVA analyses were performed on uncorrected and case-mix adjusted mean scores.

Finally, the PCQ-infertility's benchmark capability was determined by calculating intra-cluster correlation coefficients (ICCs). The ICC accounts for the relatedness of clustered data (e.g. patients clustered in fertility clinics) by comparing the variance within clusters with the variance between clusters (Killip et al., 2004). Thus the ICC provides an estimate of the total variance in experienced patient-centredness attributable to differences between fertility clinics. For each reliable subscale, an ICC was calculated in both the 0- and final model, with random intercept at the clinic level.

Analyses were performed using SPSS (version 16.0 for Windows[®], SPSS Inc., Chicago, IL, USA).

Each participating clinic was sent a detailed feedback report of their performance regarding patient-centredness, including a personalized list of quality improvement scores and their subscale mean scores compared with the national scores.

Results

Respondents

Detailed information on the focus group participants is given in Table 1 (left column).

Table I Demographic characteristics of focus group and survey participants.

Characteristic	Focus groups (24 couples and 6 women)	Survey (n = 888 couples)
Median age (years, range)		
Women	33 (24–41)	33 (20–45)
Partner	36 (26–44)	35 (21–61)
Ethnic background ^a (%)		
Dutch / Western / non-Western		
Women	100 / 0 / 0	85 / 5 / 10
Partner	96 / 0 / 4	87 / 3 / 9
Level of education ^b (%)		
Low–medium / high		
Women	57 / 43	58 / 42
Partner	46 / 54	62 / 38
Lesbian couples (%)	3.3	1
Median duration of infertility (months, range)	n.r. ^c	34 (2–174)
Median experience in fertility care (months, range)	n.r. ^c	20 (1–164)
Childless couples (%)	67	71
Diagnosis (%)		
Male factor ^d / female factor ^e / both ^f / unexplained	n.r. ^c	27 / 26 / 10 / 37
Treatment type (%)		
ART ^g / non-ART ^h	50 / 50	51 / 49
Pregnant at time of the study (%)	7	19
Self-reported health (%)		
Bad / not good, not bad / (very) good	n.r. ^c	1 / 10 / 89

^aFor ethnic background the 'Statistics Netherlands' classification was used. This Dutch governmental institution classifies ethnicity according to citizens' country of birth and to that of their parents. Immigrants include both those who are foreign-born (first generation) and those who have at least one foreign-born parent (second generation). Categories were: (1) Dutch, (2) Western (Europe, USA, Canada, Australia, New Zealand, Japan and Israel), (3) Non-Western (immigrants from remaining countries, including Morocco, Surinam and Turkey).

^bLow = primary or lower vocational education; middle = secondary or intermediate vocational education; high = higher professional education or university.

^cn.r. = not registered.

^dLow semen quality.

^eIrregular ovulation, polycystic ovary syndrome, tubal factor, endometriosis, mucus hostility.

^fBoth male and female infertility diagnosis found.

^gAssisted reproductive technology (ART), encompassed IVF, ICSI, cryopreservation and testicular sperm extraction.

^hNon-ART included ovulation induction and intrauterine insemination with or without controlled ovarian stimulation.

In the validation study, 29 of the 30 invited clinics participated. In total, 888 respondents (75%) filled out the PCQ-infertility. Sixty-three per cent of the respondents filled out the questionnaire together with their partner. Respondents' characteristics are presented in the right-hand column of Table I. Sixty-two couples returned a refusal form. Various reasons were given for non-participation, for example, having language problems, being too emotional or having too little experience with the fertility clinic. There was no difference in age between responders and non-responders ($P = 0.56$). No differences in responses were found between the responding couples and women who filled out the questionnaire alone. Respectively, 15% and 12% of the women and partners had an ethnic background other than the Dutch. At the time of the study, 19% of the women were pregnant.

Analyses

Psychometric analyses

Appropriateness of items. The seven omitted experience items that did not meet the psychometric criteria are presented in Table II together

with their reason for exclusion. For instance, item Q53 was excluded because patients commented that transition problems could be caused by both their previous and current clinics.

Internal consistency. Internal consistency analyses determined there were seven domains in which patient-centredness could be reliably measured: accessibility; information; communication; patient involvement; respect for patients' values; continuity and transition; and competence. After correlating all items with the subscale means, two items had to be displaced (Q4 from *patient involvement* to *respect* and Q6 from *communication* to *competence*). Mean scores and Cronbach's alphas of these subscales were adapted. Table III provides the final items per subscale, together with the subscale mean score and alpha. On average, 'communication' was best rated by patients; 'continuity and transition' was rated worst. The ITCs and proportion of negative experiences per item are also presented in Table III. Item responses diverged considerably among patients, even when items came in succession. For instance, 52% of the respondents reported

Table II Omitted items with reason of omission.

Omitted items (<i>n</i> = 7)		Reason for omission
Q1	Staff handed useful websites with reliable information on infertility and ART	Relatively unimportant
Q8	Staff handed useful websites for having contact with fellow patients	Relatively unimportant
Q16	It was clear what to do each day during the treatment period	Extremely skewed
Q33	Serious investigation or treatment results reported at unexpected moment	Extremely skewed
Q37	Having offered several options when making a new appointment	Not contributing to scale reliability
Q38	Treatment was also possible on weekend days	Not contributing to scale reliability
Q53	Smooth transition of medical records from previous clinic	Many negative comments

Table III The final PCQ-infertility item description and psychometric properties.

Item	Dimension scales with accompanying items	Mean score (SD)	% nE ^a	ITC ^b	α^c
Accessibility (2 items; FQ1–2; <i>n</i> = 886 ^d)		2.13 (0.78)			0.70
Q35	Telephonic access of the hospital		22	0.55	
Q36	Accessibility of the team for questions (by email or phone)		30	0.55	
Information (11 items; FQ3–13; <i>n</i> = 885)		2.03 (0.63)			0.71
Q2	Receiving written information		22	0.35	
Q5	Contact numbers for urgent problems at nights or weekends		42	0.32	
Q7*	Treatment situations when instructions by a nurse were missed		25	0.33	
Q9	Information on how and where to get psychosocial support		63	0.38	
Q10	Comprehensiveness of information on investigations		9	0.41	
Q11	Receiving an overview of treatment plan with time schedule		50	0.38	
Q13	Several treatment options were discussed		28	0.41	
Q14	Comprehensiveness of information on treatment		6	0.51	
Q17	Clear explanation on possible side- effects medication		52	0.40	
Q18	Sound instructions on how to inject hormones		4	0.32	
Q50	Periodical evaluations to overlook treatment period		54	0.35	
Communication (7 items; FQ14–20; <i>n</i> = 887)		2.53 (0.50)			0.81
Q3	Honesty and clarity on what to expect of the fertility services		15	0.50	
Q12	Physician discussed the results of investigations with you		19	0.42	
Q19	Physician listened carefully		8	0.60	
Q21	Physician took you seriously		5	0.64	
Q23	Physician took enough time		11	0.70	
Q32*	Staff were talking about you instead of talking to you		3	0.43	
Q34	Staff's willingness to talk about errors or incidents		24	0.54	
Patient involvement (3 items; FQ21–23; <i>n</i> = 881)		2.38 (0.64)			0.72
Q15	If preferred, decision-making was shared with you		21	0.49	
Q22	Physician was open to your opinion and ideas about treatment		13	0.64	
Q24	Opportunity to ask physician questions		9	0.55	
Respect for patient's values (7 items; FQ24–30; <i>n</i> = 885)		1.98 (0.76)			0.83
Q4	Having access to own medical records		67	0.38	
Q20	Physician had empathy with your emotions and actual situation		13	0.66	
Q25	Physician took interest in you as a person		32	0.67	
Q28	Staff involved your partner in your treatment		24	0.65	
Q29	Staff paid attention to the emotional impact of infertility		43	0.71	

Continued

Table III Continued

Item	Dimension scales with accompanying items	Mean score (SD)	% nE ^a	ITC ^b	α^c
Q30	Personal attention and support of nurses		52	0.59	
Q31	Nurses showed understanding for your situation		20	0.61	
Continuity and transition (7 items; FQ31–37; $n = 886$)		1.95 (0.56)			0.64
Q43	No more than 4 different physicians involved in your treatment		26	0.35	
Q44	Regularity in seeing the same physician		43	0.52	
Q45	Having a lead physician for evaluation and decision-making		34	0.44	
Q46	One caregiver as central point for problems or questions		66	0.32	
Q47*	Having received contradictory information or advice		5	0.31	
Q48*	Need to repeat the same story to different physicians		9	0.38	
Q49*	Contradictory policy adhered by different caregivers		4	0.35	
Competence (6 items; FQ38–43; $n = 888$)		2.45 (0.39)			0.71
Q6	Staff used difficult words without explaining them		2	0.33	
Q26	Physician was well prepared for your appointments		16	0.54	
Q27	Professional skills physician(s)		3	0.52	
Q42	Seen within 15 min of appointment time		71	0.34	
Q51	Fertility outpatient department well organized		8	0.50	
Q52*	Staff worked disorderly		2	0.44	
Care organization^e (single items; FQ44–46)				0.46	
Q39	Being seen within 3 weeks after physician's appointment was made		11	0.29	–
Q40	Waiting time between first visit and receiving treatment plan		27	0.30	
Q41	'Unnecessary' waiting time between two treatments		18	0.29	–
Overall patient-centredness (46 items; $n = 887$)		2.19 (0.43)			0.92

*In the original questionnaire, these items were negatively posed. For analyses, these items were mirrored.

^anE = the proportion of negative experiences with that aspect, in %.

^bCorrected ITC for each item within a domain are shown.

^cCronbach's alpha of whole domains (α) are shown. The calculated alpha's of accessibility, information, communication, patient involvement, respect for patient's values, continuity and transition, competence and care organization are based on, respectively, 747, 649, 312, 854, 518, 867, 863 and 725 patients.

FQ = the item number(s) in the final questionnaire.

^d n = the number of patients who were calculated a subscale's mean score.

^eCare organization was not a reliable dimension. Therefore, Q39, Q40 and Q41 need to be interpreted as single items.

receiving no or insufficient information on possible side effects of medication (Q17), whereas only 4% responded negatively regarding the hormone injection instructions (Q18). Bias caused by the halo-effect (answering patterns) is therefore less likely (Rubin, 1969), which contributes to the PCQ's validity.

The domain 'care organization', comprising three items, had an unacceptable low alpha of 0.46. Therefore, no mean score for this domain could be calculated. For its sufficient ITCs and importance, items Q39, Q40 and Q41 were kept in the final questionnaire, but need to be considered as single items. This altogether makes the final PCQ-infertility being a reliable scale (alpha 0.92) composed of 46 experience items.

Construct validity. All hypotheses could be accepted, which confirms the PCQ's construct validity. Patients who experienced more patient-centredness in their care were more satisfied ($\rho = 0.73$, $P = 0.01$). All PCQ's subscales were positively and significantly ($P = 0.01$) correlated with each other ($\rho = 0.18$ – 0.76). Patients with access to their medical records experienced more patient-centredness in their care than patients without this access ($P < 0.001$). The same applied to patients

who had a lead physician ($P < 0.001$), received written information ($P < 0.001$), and had scheduled treatment evaluations ($P < 0.001$). Furthermore, pregnant patients and ART-patients experienced a higher level of patient-centredness than patients who were not pregnant ($P = 0.034$) and received non-ART treatments ($P < 0.001$). In view of respondents' written comments, four of the 46 questions were slightly adapted. One answer category had been added to Q4 ('I don't know'), Q7 ('Around the pregnancy test') and Q45 ('Yes, but I saw him/her sporadically'). Additionally, items Q45 and Q46 were rephrased to improve clarity. The English version of the PCQ-infertility and a manual how to use it are available as supplementary data.

Quality improvement scores

The twelve care aspects with the highest QI scores are presented in Table IV. Given its QI score of 4.15, 'assigning each patient one contact person (e.g. a nurse) for questions' should have the highest priority for improving patient-centredness. This care aspect also received the highest mean negative experience score. As can be seen in Table IV, Q11 (supplying patients with an overview of the treatment plan and a time schedule) received a high QI score too

Table IV Twelve highest quality improvement scores.

Item	Quality aspect	I ^a	nExp ^b	QI ^c
Q46	Assign each patient one staff member (e.g. a nurse) for questions/problems	2.08	1.99	4.14
Q11	Supply patients with an overview of the treatment plan and a time schedule	2.31	1.50	3.47
Q4	Make each patient get access to own medical records during treatment	1.80	1.91	3.44
Q17	Provide information on possible side effects of prescribed medication	2.34	1.36	3.18
Q43	Assure no more than 4 different physicians are involved in patient's treatment	2.01	1.51	3.04
Q50	Schedule periodical evaluations with physician to overlook treatment period	2.05	1.45	2.97
Q44	Guarantee patients regularity in seeing the same physician	2.06	1.38	2.84
Q9	Provide information on how and where to get psychosocial support	1.54	1.83	2.82
Q5	Provide contact numbers for urgent problems at nights or weekends	2.08	1.26	2.62
Q30	Personal attention and support of nurses	1.79	1.45	2.60
Q45	Make each couple has a lead physician for evaluations and decisions	2.38	1.03	2.45
Q29	Pay attention to any emotional impact of fertility problems	2.29	1.02	2.34

^aI = importance score, with possible range from 0 to 3. The higher I, the more important the care aspect was to patients.

^bnExp = mean negative experience score = the maximum mean score of 3 – the perceived mean experience on the care aspect. The nExp has a possible range from 0 to 3. The higher the nExp, the more negative experiences patients had.

^cQI = I × nExp. QIs have a possible range from 0 to 9. The higher the QI, the higher is the improvement potential.

(3.46), since it was scored as highly important yet insufficiently met. Of all 46 care aspects, the most important was Q3 ('honesty and clarity on what to expect of the fertility services'). This item got an importance score (I) of 2.8 out of 3. 'Comprehensiveness of information on treatment' (Q14, I = 2.76) was the second most important care aspect.

Discriminative power

Table V demonstrates the results of the multilevel analyses. The intercepts in both models represent patients' mean scores on overall patient-centredness and the seven subscales (possible range 0–3). High scores represent positive experiences with care. For all mean scores, variation on the patient's level significantly differs from zero in both the 0- and final model (seventh column Table V). Significant variation at clinic level was found for overall patient-centredness and for the subscales information, communication, respect, continuity and competence. For patient involvement, significant variation was found only in the 0-model. Regression coefficients (column 3–6) show that patient characteristics 'type of treatment', 'women's level of education', 'partner's gender' and 'achieved pregnancy' are significantly associated with the outcome variables. For instance, undergoing ART is associated with experiences that are more positive regarding patient-centredness in terms of information, patient involvement, respect and overall patient-centredness. Conversely, being highly educated results in lower scores on patient-centredness and several subscales. The PCV ranged from 0 to 18.6% (Table V, column 9). This means the aforementioned patient characteristics explain only a small part of the total variance detected in the 0-models, except for the information subscale. Other characteristics did not explain any variation in perceived patient-centredness.

Case-mix adjusted mean scores for overall patient-centredness ranged from 2.53 (SE 0.10) for the best scoring clinic to 1.66 (SE 0.13) for the worst. Per dimension, clinics' case-mix adjusted mean scores ranged from 2.63 (SE 0.23) to 1.65 (SE 0.21) for 'accessibility';

from 2.45 (SE 0.15) to 1.09 (SE 0.23) for 'information'; from 2.82 (SE 0.14) to 1.88 (SE 0.15) for 'communication'; from 2.82 (SE 0.24) to 1.74 (SE 0.24) for 'patient involvement'; from 2.62 (SE 0.28) to 1.21 (SE 0.31) for 'respect'; from 2.63 (SE 0.09) to 1.44 (SE 0.12) for 'continuity'; and from 2.74 (SE 0.06) to 1.97 (SE 0.10) for 'competence'. For each scale, significant differences in both uncorrected and adjusted mean scores between clinics were found ($P \geq 0.001$). Since our total patient sample included only eight lesbian couples, mean scores were not adjusted for partner's gender.

In the final model, differences between participating fertility clinics appeared to be responsible for 11–21% of the variance in domains of patient-centredness (ICCs, last column).

Discussion

This multicentre study resulted in the first validated instrument for measuring patient-centredness in fertility care. By using the PCQ-infertility, patients' experiences with patient-centred fertility care can be reliably surveyed and benchmarked.

Over the past decades, several questionnaire studies have been conducted to evaluate patients' perspective of fertility care (Bromham et al., 1988; Sabourin et al., 1991; Sundby et al., 1994; Souter et al., 1998; Malin et al., 2001; Schmidt et al., 2003; Redshaw et al., 2007; Haagen et al., 2008). Studies with the best quality are those by Souter et al. (1998) and Haagen et al. (2008) (Dancet et al., 2010). Both were multicentric, with questionnaires based on both qualitative research and literature review. However, the questionnaire of Haagen et al. (2008) is tailored to intrauterine insemination patients, concentrates only on a part of the patient-centredness concept, and is not fully validated. The questionnaire of Souter et al. (1998) encompasses the entire concept of patient-centredness, but is not validated at all; its psychometric properties are unknown.

The PCQ measures patients' specific experiences rather than their global satisfaction, and can accordingly be adopted for improving the

Table V Model fitting results multilevel analysis for the domains of patient-centred fertility care.

	Intercept	Treatment type ^a	Education women ^a	Gender partner ^a	Achieved pregnancy ^a	Var patient ^b	Var clinic ^c	PCV ^d	ICC ^e
Accessibility									
0-model	2.14 (2.06–2.22)					0.590*	0.021	Reference	n.c. ^f
Final model	2.23 (2.15–2.32)		–0.21 (–0.32; –0.11)			0.583*	0.017	1.8%	n.c.
Information									
0-model	1.92 (1.78–2.05)					0.297*	0.118*	Reference	0.28
Final model	2.52 (2.07–2.97)	0.37 (0.28; –0.46)	–0.09 (–0.16; –0.02)	–0.73 (–1.16; –0.29)	0.11 (0.02; 0.20)	0.274*	0.064*	18.6%	0.19
Communication									
0-model	2.50 (2.43–2.57)					0.226*	0.029*	Reference	0.11
Final model	3.03 (2.63–3.42)			–0.55 (–0.93; –0.16)	0.11 (0.03; 0.19)	0.222*	0.028*	1.9%	0.11
Patient involvement									
0-model	2.36 (2.29–2.43)					0.384*	0.023*	Reference	0.06
Final model	2.89 (2.38–3.41)	0.15 (0.05; –0.24)		–0.06 (–1.12; –0.12)	0.15 (0.04; 0.25)	0.380*	0.017*	2.3%	n.c.
Respect									
0-model	1.91 (1.78–2.04)					0.492*	0.094*	Reference	0.16
Final model	1.83 (1.69–1.96)	0.24 (0.13; –0.36)	–0.11 (–0.21; –0.02)		0.14 (0.01; 0.26)	0.485*	0.071*	5.1%	0.13
Continuity									
0-model	1.95 (1.85–2.05)					0.249*	0.067*	Reference	0.21
Final model	2.40 (2.01–2.80)			–0.45 (–0.83; –0.07)		0.249*	0.066*	0.0%	0.21
Competence									
0-model	2.41 (2.34–2.48)					0.129*	0.028*	Reference	0.18
Final model	2.80 (2.51–3.08)		–0.05 (–0.10; –0.00)	–0.36 (–0.63; –0.09)		0.127*	0.028*	1.3%	0.18
Patient-centredness									
0-model	2.15 (2.07–2.22)					0.157*	0.031*	Reference	0.16
Final model	2.62 (2.29–2.95)	0.15 (0.08; –0.21)	–0.08 (–0.13; –0.02)	–0.51 (–0.84; –0.19)	0.09 (0.03; 0.16)	0.152*	0.023*	7.5%	0.13

* $P < 0.05$.

^aReference groups are for treatment type 'patients with a non-ART treatment'; for education women 'low–medium education'; for gender partner 'male'; and for 'achieved pregnancy' = 'no pregnancy achieved'.

^bVar patient = variance at the patient level.

^cVar clinic = variance at the hospital level.

^dPCV = Proportional change in variance = (Total Var 0-model – Total Var Final model)/Total Var 0-model. E.g. for information, the PCV is $((0.297 + 0.118) - (0.274 + 0.064)) / (0.297 + 0.118) = 0.186$

^eICC = Var hospital/(Var patients + Var hospital).

^fn.c. = not calculated. The ICC is not calculated since the variance at the hospital level (Var clinic) was not significant.

quality of fertility care (Cleary, 1999). First, tailored information on fertility clinics' performance gives fertility care professionals an insight into the clinic's weaknesses through their patients' eyes (Mourad et al., 2009; van Empel et al., 2010a). Despite some professionals' scepticism (Patwardhan and Patwardhan, 2009), unsatisfactory results from 'internal feedback' appear to be an incentive for quality improvement (Cleary, 1999; Berwick et al., 2003; Coulter and Ellins, 2007). Second, since the PCQ can distinguish 'weak' from 'strong' performing fertility clinics, it can be adopted for benchmark purposes on patient-centredness. The public image threat means that benchmark information can stimulate quality improvement as well, especially when a clinic scores significantly lower than others (Fung et al., 2008; Faber et al., 2009; Riiskjaer et al., 2010). Another use of public performance data on patient-centredness is patients' opportunity to compare fertility clinics on accessibility, information, competence and so on. This way, patients can make an informed choice of fertility clinic, which will strengthen their position (Coulter and Ellins, 2007).

Particularly continuity of care, respect for patient's values and information could be markedly improved in the clinics studied. Furthermore, two-thirds of the participants had a negative experience with the information provision about how and where to get psychosocial support (Q9). A possible explanation for this regrettable finding is that psychosocial care is not always an integral part of fertility care in the Netherlands, especially not in smaller non-ART clinics. Quality improvement scores can help health professionals in prioritizing which aspects to pay attention to first, to improve care more accurately. Quality improvement scores have been presented before in a similar study for Breast Care (Damman et al., 2009a,b), but their priority list for quality improvement showed completely different items than those in the current study. This illustrates the significance of surveys customized per care type (Patwardhan and Patwardhan, 2009).

A strength of the PCQ-infertility is its thoroughly developmental and validation process using both qualitative and quantitative methods (National Science Foundation, 1997). For instance, focus groups analysis and questionnaire's item formulation were carried out by two researchers independently, which increases validity and reliability (Converse and Presser, 1986; Miles and Huberman, 1994). Validity was carefully tested by many hypotheses and was not disturbed through bias by the halo-effect (Rubin, 1969). To further establish construct validity in future research, it would be interesting to test whether patients who have experienced repeated treatment failure have also more negative perceptions of fertility care. Furthermore, the PCQ's discriminative power can be considered as strength, given the high ICCs compared with similar instruments that intend benchmarking on patients' experiences (Sjetne et al., 2007; Stubbe et al., 2007a,b). One-way ANOVA confirmed significant clinic differences in patient-centredness. These differences are illustrated by the large differences found in the mean scores between clinics. For example, mean scores for information ranged from 1.20 (SD 0.63) to 2.50 (SD 0.40) on a scale from 0 to 3. Some mean scores, though, have quite high standard deviations, presumably caused by the small number of respondents per clinic (15–20 for smallest clinics). A fourth strong point is the large patient sample of the validation study ($n = 888$), which was random, and diverse. Together with the satisfying response rate (75%), this careful sampling ensures representativeness for the entire Dutch fertility population and

contributes to the PCQ's general applicability. Since the PCQ's items are not specific for the Dutch care setting only, the instrument is probably easily applied in other countries, although applicability should be assessed before using it outside the Netherlands.

However, some limitations of our study and questionnaire need to be addressed. First, the PCQ includes only items on care delivered by gynaecologists, fertility physicians and fertility nurses. Therefore, the PCQ cannot be adopted for evaluating fertility care delivered by other professionals of patients' fertility care network, like andrologists, psychologists and embryologists. However, thanks to the focus on 'mainstream fertility care', the questionnaire is of convenient length, has an extremely low non-response per item (on average 1%), and fits most fertility care settings. Second, albeit widely recommended (Silow-Corroll et al., 2006; Elwyn et al., 2007; Dancet et al., 2010), standardizing patient-centredness measurement remains a contradiction in terms to some extent. The PCQ evaluates care aspects relevant to mainstream infertile patients, whereas needs, expectations and priorities can differ somewhat among patients (Schmidt et al., 2003; Delgado et al., 2008). Accordingly, tailoring care to the individual patient is still required. A third limitation is the reliability of the dimension 'continuity of care', which is acceptable ($\alpha = 0.64$), but should be improved in future versions. This relatively low reliability may be explained by the dimension's diverse answering categories and its two dichotomous items (Q45 and Q46). Although Cronbach's alpha is the most widely used index to estimate scale reliability (Sun et al., 2007), it underestimates the true reliability when scales include dichotomous items or items that are not strictly parallel (Brogden, 1946). In the PCQ's final version, however, item Q45 has three answering categories instead of two, and the item description of Q46 has been improved. Therefore, a higher reliability of 'continuity of care' can be expected in future surveys.

Benchmark data on patient-centredness should reflect the actual performance of a specific clinic, and not its different composition of patient profiles. Therefore, we performed case-mix adjustment for three of the four determinants found significant in the multilevel regression analysis. Before 'adjusting' for lesbian couples as standard procedure, more research is deemed necessary to establish the impact of the partner's gender. Multilevel analysis is currently the best available tool for case-mix adjustment (Damman et al., 2009a,b; Hekkert et al., 2009). Interestingly, after adjustment for treatment type, level of education and achieved pregnancy, differences in mean scores between clinics were even larger than before adjustment. However, case-mix adjusters can unintentionally adjust for systematic differences in care delivery to different patient groups, but is not able to adjust for bias caused by differences in patients' expectations of care (Bago d'Uva et al., 2008). For the 'calibration' of responses, the use of anchoring vignettes can be investigated as alternative for case-mix adjustment (King et al., 2004).

In conclusion, this study provides a valid, reliable and strongly discriminating instrument to measure patient-centredness in fertility care: the PCQ-infertility. It can offer clinics detailed insight in their performance according to patients, and allows tailored quality improvement and benchmarking. From now on, the quality of fertility care cannot only be monitored and benchmarked on live birth and complication rates, but also on patient-centredness. Future cross-national research should establish the PCQ's value for infertile populations beyond the Netherlands.

Supplementary data

Supplementary data are available at <http://humrep.oxfordjournals.org/>.

Acknowledgements

We thank all patient couples for their participation and the 29 participating clinics for their cooperation in data-sampling. Jacky Boivin is acknowledged for her help with the English translation of the PCQ-infertility.

Funding

This work was funded by the Department of Obstetrics and Gynaecology of the Radboud University Nijmegen Medical Centre, The Netherlands.

References

- Bago d'Uva T, Van Doorslaer E, Lindeboom M, O'Donnell O. Does reporting heterogeneity bias the measurement of health disparities? *Health Econ* 2008;**17**:351–375.
- Bengoa R, Kawar R, Key P, Leatherman S, Massoud R, Saturno P. *Quality of Care: A Process for Making Strategic Choices in Health Systems*. Geneva: WHO Press, 2006.
- Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. *Med Care* 2003;**41**:130–138.
- Brogden HE. Variation in test validity with variation in the distribution of item difficulties, number of items, and degree of their intercorrelation. *Psychometrika* 1946;**11**:197–214.
- Bromham D, Blamer B, Clay R, Hamer R. Disenchantment with infertility services: a survey of patients in Yorkshire. *Br J Fam Plann* 1988;**14**:3–8.
- Clark J. The narrative in patient-centred care. *Br J Gen Pract* 2008;**58**:896.
- Cleary PD. The increasing importance of patient surveys. Now that sound methods exist, patient surveys can facilitate improvement. *BMJ* 1999;**319**:720–721.
- Converse J, Presser S. *Survey Questions. Handcrafting the Standardized Questionnaire*. Beverly Hill: Sage, 1986.
- Coulter A, Ellins J. Effectiveness of strategies for informing, educating, and involving patients. *BMJ* 2007;**335**:24–27.
- Damman OC, Hendriks M, Sixma HJ. Towards more patient centred healthcare: a new consumer quality index instrument to assess patients' experiences with breast care. *Eur J Cancer* 2009a;**45**:1569–1577.
- Damman OC, Stubbe JH, Hendriks M, Arah OA, Spreeuwenberg P, Delnoij DM, Groenewegen PP. Using multilevel modeling to assess case-mix adjusters in consumer experience surveys in health care. *Med Care* 2009b;**47**:496–503.
- Dancet EA, Nelen WL, Sermeus W, De LL, Kremer JA, D'Hooghe TM. The patients' perspective on fertility care: a systematic review. *Hum Reprod Update* 2010; Epub ahead of print.
- Delgado A, Andrés Lopéz-Fernández L, de Dios Luna J, Gil N, Jiménez M, Puga A. Patient expectations are not always the same. *J Epidemiol Community Health* 2008;**62**:427–434.
- Elwyn G, Buetow S, Hibbard J, Wensing M. Measuring quality through performance. Respecting the subjective: quality measurement from the patient's perspective. *BMJ* 2007;**335**:1021–1022.
- Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. *Med Care* 2009;**47**:1–8.
- Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med* 2008;**148**:111–123.
- Gerteis M, Edgman-Levitan S, Daley J, Delbanco TL. *Through the patient's eyes*. San Francisco: Jossey-Bass Publishers, 1993. USA.
- Gunby J, Bissonnette F, Librach C, Cowan L. Assisted reproductive technologies (ART) in Canada: 2004 results from the Canadian ART register. *Fertil Steril* 2008;**89**:1123–1132.
- Haagen EC, Hermens RP, Nelen WL, Braat DD, Kremer JA, Grol RP. Subfertile couples' negative experiences with intrauterine insemination care. *Fertil Steril* 2008;**89**:809–816.
- Hammarberg K, Astbury J, Baker H. Women's experience of IVF: a follow-up study. *Hum Reprod* 2001;**16**:374–383.
- Hays RD, Shaul JA, Williams VS, Lubalin JS, Harris-Kojetin LD, Sweeny SF, Cleary PD. Psychometric properties of the CAHPS 1.0 survey measures. Consumer assessment of health plans study. *Med Care* 1999;**37**:MS22–MS31.
- Hekkert KD, Cihangir S, Kleefstra SM, van den BB, Kool RB. Patient satisfaction revisited: a multilevel approach. *Soc Sci Med* 2009;**69**:68–75.
- Institute of Medicine. *Crossing the Quality Chasm. A New Health System for the 21st Century*. Washington, DC: National Academy Press, 2001, USA.
- Jenkinson C, Coulter A, Bruster S. The picker patient experience questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health Care* 2002;**14**:353–358.
- Keirns CC, Goold SD. Patient-centered care and preference-sensitive decision making. *JAMA* 2009;**302**:1805–1806.
- Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med* 2004;**2**:204–208.
- King G, Murray CJL, Salomon JA, Tandon A. Enhancing the validity and cross-cultural comparability of measurement in survey research. *Am Polit Sci Rev* 2004;**98**:191–207.
- Kline P. *Handbook of Test Construction*. London: Methuen, 1986.
- Krueger RA. Analyzing and reporting focus group results. In: Morgan DL, Krueger RA (eds). *The Focus Group Kit*. London: Sage, 1998, 36–37.
- Malin M, Hemmink E, Raikonen O, Sihvo S, Perala ML. What do women want? Women's experiences of infertility treatment. *Soc Sci Med* 2001;**53**:123–133.
- Merlo J, Yang M, Chaix B, Lynch J, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people. *J Epidemiol Community Health* 2005;**59**:729–736.
- Miles MB, Huberman AM. *Qualitative Data Analysis. An Expanded Sourcebook*, 2nd edn. Thousand Oaks: Sage, 1994.
- Mourad SM, Nelen WLDM, Akkermans RP, Vollebergh JH, Grol RP, Hermens RP, Kremer JA. Determinants of patients' experiences and satisfaction with fertility care. *Fertil Steril* 2009; Epub ahead of print.
- Mullen PM. Public involvement in health care priority setting: an overview of methods for eliciting values. *Health Expect* 1999;**2**:222–234.
- National Science Foundation. *User-Friendly Handbook for Mixed Method Evaluations*. Virginia: National Science Foundation, 1997
- Nunnally J. *Psychometric Theory*. New York: McGraw-Hill, 1978.
- Nyboe Andersen A, Goossens V, Bhattacharya S, Ferraretti AP, Kupka MS, de MJ, Nygren KG. Assisted reproductive technology and intrauterine inseminations in Europe, 2005: results generated from European registers by ESHRE: ESHRE. The European IVF Monitoring Programme (EIM), for the European Society of Human Reproduction and Embryology (ESHRE). *Hum Reprod* 2009;**24**:1267–1287.

- Patwardhan A, Patwardhan P. Are consumer surveys valuable as a service improvement tool in health services? A critical appraisal. *Int J Health Care Qual Assur* 2009;**22**:670–685.
- Redshaw M, Hockley C, Davidson LL. A qualitative study of the experience of treatment for infertility among women who successfully became pregnant. *Hum Reprod* 2007;**22**:295–304.
- Riiskjaer E, Ammentorp J, Nielsen JF, Kofoed PE. Patient surveys—a key to organizational change? *Patient Educ Couns* 2010;**78**:394–401.
- Rubin D. Halo effect in individual response to an attitude questionnaire. *Percept Mot Skills* 1969;**28**:258.
- Sabourin S, Wright J, Duchesne C, Belisle S. Are consumers of modern fertility treatments satisfied? *Fertil Steril* 1991;**56**:1084–1090.
- Schmidt L, Holstein BE, Boivin J, Sangren H, Tjornhoj-Thomsen T, Blaabjerg J, Hald F, Andersen AN, Rasmussen PE. Patients' attitudes to medical and psychosocial aspects of care in fertility clinics: findings from the Copenhagen multi-centre psychosocial infertility (COMPI) research programme. *Hum Reprod* 2003;**18**:628–637.
- Silow-Corroll S, Alteras T, Stepnick L. *Patient-Centered Care for Underserved Populations: Definition and Best Practices*. Washington, DC: Economic and Social Research Institute, 2006, 1–43.
- Singer S, Kuhnt S, Momenghalibaf A, Stuhr C, Dimmel-Hennersdorf U, Köhler U, Einenkel J. Patients' acceptance and psychometric properties of the EORTC QLQ-CX24 after surgery. *Gynecol Oncol* 2010;**116**:82–87.
- Sjetne IS, Veenstra M, Stavem K. The effect of hospital size and teaching status on patient experiences with hospital care: a multilevel analysis. *Med Care* 2007;**45**:252–258.
- Souter VL, Penney G, Hopton JL, Templeton AA. Patient satisfaction with the management of infertility. *Hum Reprod* 1998;**13**:1831–1836.
- Streiner DL, Norman GR. Selecting the items. In: Streiner DL, Norman GR (eds). *Health Measurement Scales. A Practical Guide to Their Development and Use*. 4th edn. Oxford: Oxford University Press, 2008, 77–102.
- Stubbe JH, Brouwer W, Delnoij DM. Patients' experiences with quality of hospital care: the consumer quality index cataract questionnaire. *BMC Ophthalmol* 2007a;**7**:14.
- Stubbe JH, Gelsema T, Delnoij DM. The consumer quality index hip knee questionnaire measuring patients' experiences with quality of care after a total hip or knee arthroplasty. *BMC Health Serv Res* 2007b;**7**:60.
- Sun W, Chou CP, Stacy AW, Ma H, Unger J, Gallaher P. SAS and SPSS macros to calculate standardized Cronbach's alpha using the upper bound of the phi coefficient for dichotomous items. *Behav Res Methods* 2007;**39**:71–81.
- Sundby J, Olsen A, Schei B. Quality of care for infertility patients. An evaluation of a plan for a hospital investigation. *Scand J Soc Med* 1994;**22**:139–144.
- van Empel IWH, Nelen WLDM, Hermens RPMG, Kremer JAM. Coming soon to your clinic: high-quality ART. *Hum Reprod* 2008;**23**:1242–1245.
- van Empel IWH, Nelen WLDM, Tepe ET, van Laarhoven EAP, Verhaak CM, Kremer JAM. Weaknesses, strengths and needs in fertility care according to patients. *Hum Reprod* 2010a;**25**:142–149.
- van Empel IWH, Hermens RPMG, Akkermans RP, Hollander KW, Nelen WLDM, Kremer JAM. Organizational determinants for patient-centered fertility care: a multilevel analysis. *Fertil Steril* 2010b; in press.
- Verberg MF, Eijkemans MJ, Heijnen EM, Broekmans FJ, De KC, Fauser BC, Macklon NS. Why do couples drop-out from IVF treatment? A prospective cohort study. *Hum Reprod* 2008;**23**:2050–2055.
- Verhaak CM, Smeenk JM, Evers AW, Kremer JA, Kraaimaat FW, Braat DD. Women's emotional adjustment to IVF: a systematic review of 25 years of research. *Hum Reprod Update* 2007;**13**:27–36.
- Wensing M, Elwyn G. Methods for incorporating patients' views in health care. *BMJ* 2003;**326**:877–879.