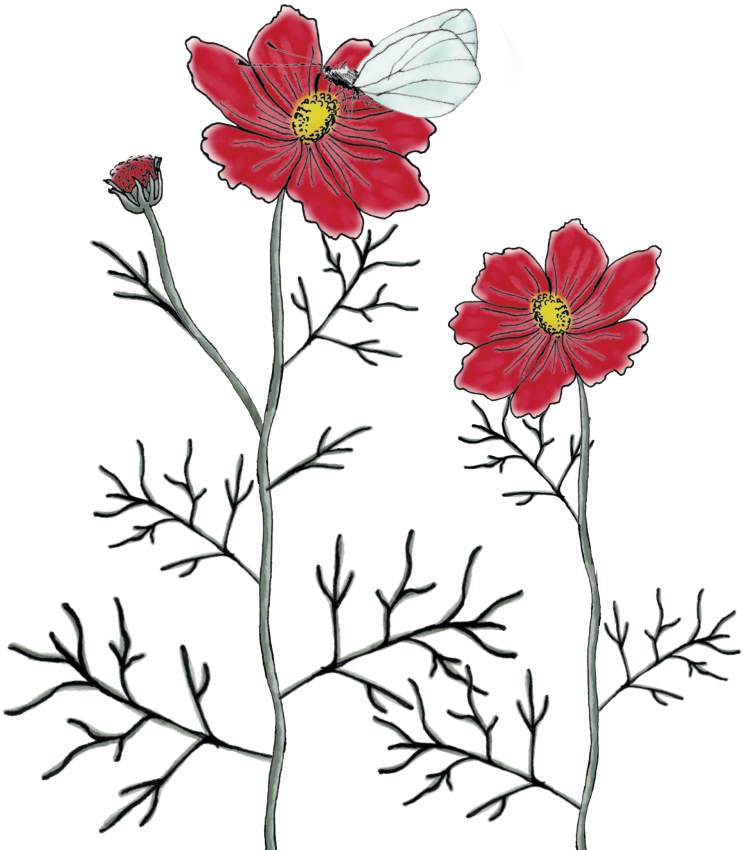
Measuring shared decision making in oncology: an informed approach

Bomhof-Roordink, H.



6

PATIENT AND PHYSICIAN SHARED DECISION-MAKING BEHAVIOURS IN ONCOLOGY: EVIDENCE ON ADEQUATE MEASUREMENT PROPERTIES OF THE iSHARE QUESTIONNAIRES

Hanna Bomhof-Roordink
Anne M. Stiggelbout
Fania R. Gärtner
Johanneke E.A. Portielje
Cor D. de Kroon
Koen C.M.J. Peeters
Karen J. Neelis
Jan Willem T. Dekker
Trudy van der Weijden
Arwen H. Pieterse
iSHARE study group

iSHARE study group: Stephanie M. de Boer, Liesbeth J. Boersma, Daan ten Bokkel Huinink, Jeroen Buijsen, Marissa Cloos-van Balen, Hans Gelderblom, Lobke van Leeuwen-Snoeks, Irene M. Lips, S. Rajen S. Ramai, Hossain Roshani, Marije Slingerland, Ben G.L. Vanneste, Lisette M. Wiltink.

Patient Education and Counseling. 2021. Online ahead of print

ABSTRACT

Objectives

We have developed two Dutch questionnaires to assess the shared decision-making (SDM) process in oncology; the iSHAREpatient and iSHAREphysician. In this study, we aimed to determine: scores, construct validity, test-retest agreement (iSHAREpatient), and inter-rater (iSHAREpatient-iSHAREphysician) agreement.

Methods

Physicians from seven Dutch hospitals recruited cancer patients, and completed the iSHAREphysician and SDM-Questionnaire-physician version. Their patients completed the: iSHAREpatient, nine-item SDM-Questionnaire, Decisional Conflict Scale, Combined Outcome Measure for Risk communication And treatment Decision-making Effectiveness, and five-item Perceived Efficacy in Patient-Physician Interactions. We formulated, respectively, one (iSHAREphysician) and 10 (iSHAREpatient) a priori hypotheses regarding correlations between the iSHARE questionnaires and questionnaires assessing related constructs. To assess test-retest agreement patients completed the iSHAREpatient again 1-2 weeks later.

Results

In total, 151 treatment decision-making processes with unique patients were rated. Dimension and total iSHARE scores were high both in patients and physicians. The hypothesis on the iSHAREphysician and 9/10 hypotheses on the iSHAREpatient were confirmed. Test-retest and inter-rater agreement were >.60 for most items.

Conclusions

The iSHARE questionnaires show high scores, have good construct validity, substantial test-retest agreement, and moderate inter-rater agreement.

Practice implications

Results from the iSHARE questionnaires can inform both physician- and patient-directed efforts to improve SDM in clinical practice.

1. INTRODUCTION

Those who have not experienced the intricacies of clinical practice demand measures that are easy, precise, and complete—as if a sack of potatoes was being weighed. True, some elements in the quality of care are easy to define and measure, but there are also profundities that still elude us. We must not allow anyone to belittle or ignore them; they are the secret and glory of our art.

Avedis Donabedian[1]

Measurement of shared decision making (SDM) remains a challenge.[2-4] The SDM process in which patients, their loved ones and healthcare professionals together arrive at treatment decisions incorporating patients' values and preferences is not easy to capture in a measurement instrument. SDM happens both during and outside consultations,[5] involves both observable (e.g., information-giving) and covert (e.g., thinking about the options) behaviours, and includes behaviours of both patients and healthcare professionals.[6, 7] Current SDM measurement instruments do not cover all of these aspects, and substantially differ in which SDM elements are assessed.[8, 9] Many often-used measurement instruments assess only healthcare professionals' behaviour (e.g., OPTION,[10] CollaboRATE)[11] or do not assess patient behaviour independently of physician behaviour (e.g., nine-item SDM-Questionnaire (SDM-Q-9),[12] SDM-Questionnaire-physician version (SDM-Q-Doc),[13] impeding the assessment of patients' role.

We developed the Dutch iSHARE questionnaires to assess SDM in oncology, from both a patient (iSHAREpatient) and physician (iSHAREphysician) viewpoint.[14] We chose the oncology setting since cancer patients often face preference-sensitive decisions.[15, 16] The SDM construct was informed by an SDM model in oncology based on stakeholders' views, and by a review of SDM models across healthcare settings published until June 2016. The iSHARE questionnaires include both patient and physician behaviours. Cancer patients and physicians were extensively involved during the development process, in line with quality criteria for the development of health-related measurement instruments.[17]

We aimed to a) describe scores obtained by the iSHARE questionnaires in an oncology setting, and determine b) construct validity of the iSHARE questionnaires, c) test-retest agreement of the iSHAREpatient, and d) agreement between scores on the iSHAREpatient and iSHAREphysician.

2. METHODS

2.1 Study design

In this multicentre study, we asked physicians from seven Dutch hospitals to complete a questionnaire after each consultation with a unique eligible patient, between June 2018 and December 2019. Participating patients were asked to complete a questionnaire after the consultation, and again 1-2 weeks later. We aimed for 50 physicians, each including at least two patients, based on the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist.[18-20] The Medical Ethical Committee of the Leiden University Medical Centre (LUMC) approved the study (NL50551.058.14, P14.207), which was conducted according to the Dutch Medical Research Involving Human Subjects Act.

2.2 Participant recruitment

We approached physicians treating cancer patients for participation, and asked consenting physicians to recruit consecutive unique eligible patients. Patients were eligible if they had been diagnosed with cancer, were ≥18 years old, able to speak and write Dutch, had a consultation in which a decision to start, stop, change or forgo treatment with curative or palliative intent was discussed, and had a life expectancy of over three months. We aimed to assess the measurement properties of the iSHARE questionnaires in a sample representing the heterogeneity of cancer treatment decisions, and therefore asked physicians from a range of cancer specialties to approach patients.

The physicians provided patients with an information letter, an informed consent form, and a post-consultation questionnaire, and asked them if they agreed to being called by the researchers. If so, we contacted them to ask if they had questions and if they were willing to participate. Consenting patients sent us their signed informed consent form and the completed questionnaire. We only used the physician's questionnaire if the patient had provided informed consent.

2.3 Data collection

Physicians reported their birth year, gender, year of start of specialization, working place, and specialty. They completed the iSHAREphysician[14] and the SDM-Q-Doc[13] post-consultation on paper or online. They also reported the patient's primary tumour type and curative/palliative intent of the treatment discussed. Patients completed the: iSHAREpatient,[14] SDM-Q-9,[12] Decisional Conflict Scale (DCS),[21] Combined Outcome Measure for Risk communication And treatment Decision-making Effectiveness (COMRADE),[22] five-item Perceived Efficacy in Patient-Physician Interactions (PEPPI-5),[23] and birth date, gender, education, month and year of most recent cancer diagnosis, and number of consultations they had in mind while completing the questionnaire, on paper or online. We sent consenting patients the iSHAREpatient again on paper or via email, whichever they preferred, within a few days after we had received the initial questionnaire. To match patients and physicians, the paper version of the questionnaire included a study code that was unique for each unique decision-making process. In case patients or physicians completed the questionnaires online, they used a link to the online database questionnaire system Qualtrics, and entered the study code. We entered the data from the paper questionnaires in Qualtrics.

2.4 iSHAREpatient and iSHAREphysician

The iSHAREpatient (Box 1) and iSHAREphysician (Box 2) have the same, but mirrored 15 items,[14] with a six-point unbalanced scale, ranging from 'not at all' (0) to 'completely' (5).[24] They encompass the same construct, consisting of six dimensions (i.e., Choice awareness, Medical information, Preferences, Deliberation, Time for deliberation, Decision). The items relate to these six dimensions, which we do not assume to be necessarily correlated,[2, 25, 26] leading us to adopt a formative measurement model (i.e., the items form the construct).[14] The dimensions aim to assess the complete SDM process both during and outside consultations, and include both patient and physician behaviours. Depending on whether a decision has already been made or not, either the score on item 15 or item 16 is relevant to compute the

score on dimension six.[14] If a patient or physician had indicated that a decision had been made, or if the response to that item was missing, we report the score on item 15; otherwise, we report the score on item 16.

We calculated dimension scores (range, 0-5) and a total score (the sum of the dimension scores; range, 0-30) for both iSHARE questionnaires. We applied a linear transformation to obtain a 0 to 100 total score ((score/30)*100). Higher dimension and total scores indicate higher levels of SDM. We only report dimension and total scores if all the respective items had been completed; the formative nature of the construct makes imputation of missing values inappropriate.

2.5 Construct validity of the iSHAREpatient and iSHAREphysician

We determined construct validity by testing hypotheses about correlations between the iSHARE questionnaires and questionnaires measuring related constructs. We formulated a priori hypotheses based on the content of the respective scales, subscales and items and/or on the construct they aim to assess. For example, we expected the COMRADE subscale 'satisfaction with communication' to correlate positively with the iSHAREpatient, based on the content of the items. We tested hypotheses on total score level for both iSHARE questionnaires and on dimension level for the iSHAREpatient (Table 5). We further expected the three iSHAREpatient items on patient-initiated behaviour (items 7, 13, 14) each to correlate with the PEPPI-5. We expected a correlation of >.30 or <-.30 for each hypothesis. We did not formulate hypotheses at the dimension level for the iSHAREphysician or the iSHAREpatient dimensions Choice Awareness, Deliberation, and Time for Deliberation, since we could not find questionnaires measuring related constructs from the same viewpoint.

2.5.1 SDM-Q-9 and SDM-Q-Doc

The SDM-Q-9[12] and SDM-Q-Doc[13] assess SDM from respectively patient and physician perspective. They each include nine items that are scored on a six-point scale from 'completely disagree' (0) to 'completely agree' (5). The raw score ranges from 0 to 45 and is multiplied by 20/9, resulting in a score from 0 to 100. Higher scores indicate higher levels of SDM.[12, 13] Both questionnaires have been validated in the oncology setting,[27-29] and have been translated and validated in Dutch.[30] Cronbach's α's were .90 (SDM-Q-9) and .85 (SDM-Q-Doc).

2.5.2 COMRADE

The COMRADE aims to measure effectiveness of risk communication and treatment decision making in consultations, and consists of two subscales: satisfaction with communication (10 items) and confidence in decision (10 items). The response scale ranges from 'strongly disagree' (1) to 'strongly agree' (5).[22] We calculated subscale scores based on the original factor analysis that was provided by the developer. Both subscale scores range from 0 to 100, with higher scores indicating more satisfaction or confidence, respectively. The COMRADE has been translated in Dutch.[31] Cronbach's α's were .91 (satisfaction with communication) and .90 (confidence in decision).

2.5.3 DCS

The DCS is a 16-item questionnaire assessing the level of decisional conflict; the five-point scale items range from 'strongly agree' (0) to 'strongly disagree' (4).[21] The scale consists of five subscales: feeling uncertain (3 items), feeling uninformed (3 items), feeling unclear about values (3 items), feeling unsupported (3 items), and ineffective decision making (4 items).[32] To calculate the subscale scores, item scores are summed, divided by the number of items in the subscales and multiplied by 25, with scores ranging from 0 to 100. The total score ranges from 0 to 64, is multiplied by 25/16, resulting in a standardized score from 0 to 100. Higher scores indicate higher decisional conflict. The DCS has been translated and validated in Dutch, in an oncology setting.[33] Cronbach's α's were .69 (feeling uncertain), .73 (feeling uninformed), .58 (feeling unclear about values), .32 (feeling unsupported) and .82 (ineffective decision making).

2.5.4 PEPPI-5

The PEPPI-5 aims to measure patients' perceived self-efficacy in obtaining medical information and attention to their medical concerns from physicians. The response scale ranges from 'not at all confident' (1) to 'very confident' (5) and the total score ranges from 5 to 25, with higher scores representing higher perceived self-efficacy in patient-physician interactions.[23] The PEPPI-5 has been translated and validated in Dutch, in patients with osteoarthritis.[34] Cronbach's α was .91.

2.6 Test-retest agreement of the iSHAREpatient

We assessed test-retest agreement of the iSHAREpatient, that is, the extent to which item scores for patients with a stable perception of the SDM process were the same for repeated measurements over time.[35] The COSMIN study design checklist[20] requires participants to be stable during the chosen interval, and the interval to be long enough to avoid them recalling their scores at first administration; we expected a time window of 1-2 weeks to be appropriate between test and retest. We excluded patients who answered affirmatively to one or both of the following questions at retest: 'Please think back to the time you filled in the questionnaire for the first time. Do you have different thoughts regarding the decision-making process now, compared to the thoughts you had back then?' and 'Have you had another conversation with the physician in the meantime?'.

We did not consider it feasible to assess test-retest agreement for the iSHAREphysician. We did not expect physicians to be able to recall the treatment decision-making process for a particular patient well enough over a period of 1-2 weeks to complete the iSHAREphysician again for that patient.

2.7 Inter-rater agreement between the iSHAREpatient and iSHAREphysician

In accordance with the COSMIN study design checklist[20] we determined agreement (not correlation) between the scores on the iSHAREpatient and iSHAREphysician.

Box 1. iSHAREpatient†[14]

iSHARE: Deciding together on the treatment of cancer

When completing this questionnaire, please think of the last time you spoke to your doctor in the hospital about the treatment options. This may have been in one or multiple conversations. When you are completing the questionnaire, please think about all these conversations.

The statements are about the doctor and about yourself. Some statements may look similar, but ask about something different.

For each statement, tick the answer that fits best. There are no right or wrong answers, it is your opinion that matters. Your answers will remain anonymous, so the doctor will not see them.

This questionnaire is not about how satisfied you are with your doctor. It is about what your doctor said or did during the conversation.

Do you find the information mentioned above clear?

☐ Yes

☐ No. Please state what is not clear to you: .....

1. The doctor explained what the advantages of the treatment options are

not at all	hardly	a little	for a large part	almost completely	completely
☐	☐	☐	☐	☐	☐

2. The doctor explained what the disadvantages of the treatment options are
3. The doctor explained the advantages and disadvantages of each treatment option equally well
4. The doctor checked whether I understood the advantages of the treatment options
5. The doctor checked whether I understood the disadvantages of the treatment options
6. The doctor told me how the treatment options differ from each other
7. I asked questions about the treatment options
8. At the beginning of the conversation, the doctor said that there was a choice with regard to my treatment*
9. The doctor said that it matters what I think is important*
10. The doctor checked whether he/she understood what was important to me
11. The doctor helped me to weigh up the advantages and disadvantages of the treatment options
12. The doctor gave me time to weigh up the advantages and disadvantages of the treatment options (during or after the conversation)
13. I told the doctor what was important to me
14. I weighed up the advantages and disadvantages of the treatment options (before, during or after the conversation)

Has a decision about treatment been made?

☐ *Yes, the decision has been made* → *please fill in question 15 below*

☐ *No, the decision has not been made* → *please fill in question 16 below*

15. The decision takes into account what I consider to be important
16. The doctor has discussed with me what I need in order to weigh up the advantages and disadvantages of the treatment options

† This is an English translation of the original Dutch iSHAREpatient. A translation agency translated the iSHAREpatient using a forward-backward approach.

*Items 8 and 9 of the iSHARE questionnaires assess whether the physician created choice awareness. Theoretically, we consider this as the first step of the SDM process. We decided against putting the items at the start of the questionnaires because patients did not seem to critically reflect on what was asked if they were presented first. We recommend future users to adopt the same approach.

Box 2. iSHAREphysician†[14]

iSHARE: Deciding together on the treatment of cancer

When completing this questionnaire, please think about the consultation in which you discussed the decision about the treatment with the patient. You may have had several consultations with the patient about this decision. When you are completing the questionnaire, please think about all these consultations.

The statements are about the patient and about yourself. There are no right or wrong answers.

1. I explained what the advantages of the treatment options are

not at all	hardly	a little	for a large part	almost completely	completely
☐	☐	☐	☐	☐	☐

2. I explained what the disadvantages of the treatment options are
3. I explained the advantages and disadvantages of each treatment option equally well
4. I checked whether the patient understood the advantages of the treatment options
5. I checked whether the patient understood the disadvantages of the treatment options
6. I told the patient how the treatment options differ from each other
7. The patient asked questions about the treatment options
8. At the beginning of the conversation, I said that there was a choice with regard to the patient's treatment*
9. I said that it matters what the patient thinks is important*
10. I checked whether I understood what was important to the patient
11. I helped the patient to weigh up the advantages and disadvantages of the treatment options
12. I gave the patient time to weigh up the advantages and disadvantages of the treatment options (during or after the conversation)
13. The patient told me what was important to him/her
14. The patient weighed up the advantages and disadvantages of the treatment options (before, during or after the conversation)

Has a decision about treatment been made?

☐ *Yes, the decision has been made* → *please fill in question 15 below*

☐ *No, the decision has not been made* → *please fill in question 16 below*

15. The decision takes into account what the patient considers to be important
16. I discussed with the patient what he/she needs in order to weigh up the advantages and disadvantages of the treatment options

†This is an English translation of the original Dutch iSHAREphysician. The translation is based on the translation of the iSHAREpatient.

*Items 8 and 9 of the iSHARE questionnaires assess whether the physician created choice awareness. Theoretically, we consider this as the first step of the SDM process. We decided against putting the items at the start of the questionnaires because patients did not seem to critically reflect on what was asked if they were presented first. We recommend future users to adopt the same approach.

2.8 Statistical analyses

2.8.1 Selection and missing values

We excluded test and/or retest patient questionnaires if they had been completed >30 days post-consultation, and physician questionnaires if they had been completed >7 days post-consultation (Figure 1). We assumed that a longer period would be detrimental to participants' recollection of the decision-making process.

We handled missing values according to authors' recommendations, if provided in the original or Dutch validation paper (see section 2.5).[12, 13, 34] For the other questionnaires and the iSHARE questionnaires (see section 2.4), we only report scores when all respective items had been completed. We report sample sizes per analysis, since these may differ due to missing values.

2.8.2 Analyses

Descriptive statistics were used to report scores on all questionnaires. Hypotheses were tested by calculating Spearman correlation coefficients between the scores on the iSHARE questionnaires and the respective comparison questionnaires, as the data were non-normally distributed on all scales. We determined test-retest agreement and inter-rater agreement by calculating agreement and the corresponding 95% confidence intervals (CIs).[36, 37] Due to the non-normally distributed data it was not possible to calculate weighted kappa's. For test-retest agreement we defined agreement as the same item score obtained both at test and retest: (X00+X11+X22+X33+X44+X55)/(X01+X02+X03+X04 +X05+X10+X12+…+X54), where e.g., X33 means that for both test and retest the item score was 3. For inter-rater agreement, we allowed the item scores to differ one point, since we considered it acceptable if scores from the respective viewpoints somewhat differed. To illustrate, a score of 5 on an iSHAREpatient item and a score of 4 on the same iSHAREphysician item (i.e., X54), was considered as agreement. Consequently, proportion agreement (P) was defined as: (X00+X01+X10+X11+X12+X21+X22+X23+X32+X33+X34+X43+X44+X45+X54+X55) /(X02+X03+X04+X05+X13+X14+…+X53). The corresponding CIs were calculated as follows:

$$P_{low} = P - c_{\frac{\alpha}{2}}\sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{\mathrm{n}} - \frac{1}{2n}} \qquad P_{high} = P + c_{\frac{\alpha}{2}}\sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{\mathrm{n}} + \frac{1}{2n}}$$

When agreement was close to 0 or 1 (i.e. ≤.3 or ≥.7), we applied the Fleiss correction to the corresponding CIs. These CIs were calculated as follows:[36]

$$P_{low} = \left(2nP + {c_{\frac{\alpha}{2}}}^2 - 1\right) - c_{\frac{\alpha}{2}}\sqrt{\frac{{c_{\frac{\alpha}{2}}}^2 - \left(2 + \frac{1}{\mathrm{n}}\right) + 4\mathrm{P}(\mathrm{n}(1-\mathrm{P}) + 1)}{2(\mathrm{n} + {c_{\frac{\alpha}{2}}}^2)}}$$

$$P_{high} = \left(2nP + {c_{\frac{\alpha}{2}}}^2 - 1\right) + c_{\frac{\alpha}{2}}\sqrt{\frac{{c_{\frac{\alpha}{2}}}^2 + \left(2 + \frac{1}{\mathrm{n}}\right) + 4\mathrm{P}(\mathrm{n}(1-\mathrm{P}) - 1)}{2(\mathrm{n} + {c_{\frac{\alpha}{2}}}^2)}}$$

where n is the sample size and $c_{\frac{\alpha}{2}}$ the percentile cut-off for the standard normal distribution (i.e., 1.96 for the 95% CI). CIs for agreement were calculated in Excel version 2010. We used SPSS version 25 to perform all other analyses. A p-value <.05 was considered statistically significant.

3. RESULTS

3.1 Participants

In total, 156 patients and 51 physicians participated in the study (Table 1). Fifty-seven eligible patients who had been approached for participation by their treating physician and took the study information home, did not provide consent. We do not know how many eligible patients have been approached and declined immediately. In total, 151 treatment decision-making processes were rated by both patients and physicians, with a range of one to seven per physician. Five decision processes were only rated by patients and eleven only by physicians (Figure 1). Patients completed the initial questionnaire 6.0+6.0 (range, 0-29) days post-consultation and physicians 0.2+0.8 (range, 0-7) days post-consultation. Eighty-five patients thought about more than one consultation while completing the questionnaire.

3.2 Responses on the iSHAREpatient and iSHAREphysician

Both the iSHAREpatient and iSHAREphysician showed few missing values (Table 2). The iSHAREpatient and iSHAREphysician dimension scores showed a distribution skewed toward higher scores (Figure 2). Median total scores (interquartile range (IQR)) were 95.0 (77.1-99.5) (iSHAREpatient) and 75.0 (61.1-90.7) (iSHAREphysician) (Table 3). In total, 35 (23%) patients and for 15 (10%) treatment decision-making processes physicians gave the highest possible total score (100).

3.3 Construct validity of the iSHARE questionnaires

Table 3 displays the median total and subscale scores on the comparison questionnaires used for hypotheses testing. The hypothesis formulated for the iSHAREphysician was confirmed. Nine out of ten hypotheses formulated for the iSHAREpatient were also confirmed (Table 5).

3.4 Test-retest agreement iSHAREpatient

In total, 112 patients completed the iSHAREpatient for the second time within 30 days post-consultation, of which 45 were excluded for various reasons (Figure 1). Mean time between test and retest was 11.1+3.7 (range, 4-24) days. Agreement at item level ranged from .55 (item 11) to .84 (item 15) (Table 4). Three patients had reported that no decision had been made at both test and retest and completed item 16 twice; agreement was .00. A post-hoc analysis in which we allowed item scores to differ one point, showed agreement ranging from .79 (item 7) to .97 (item 15) (Table 4).

3.5 Inter-rater agreement between the iSHAREpatient and iSHAREphysician

Inter-rater agreement between the iSHARE questionnaires ranged from .55 (item 12) to .79 (item 1 and 15). Seven patients and physicians both had reported that no decision had been made and completed item 16; agreement was .43 (Table 2).

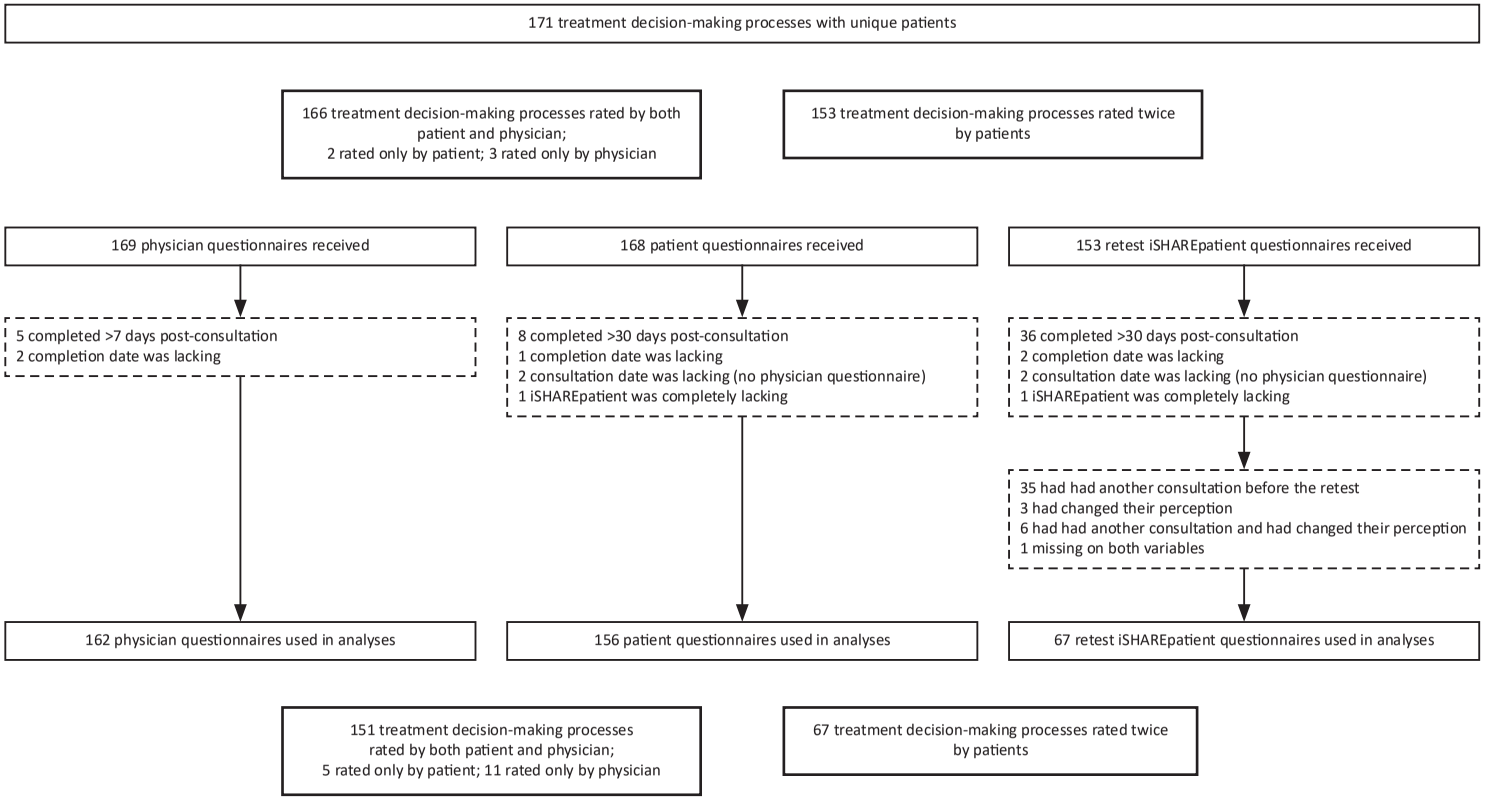


Figure 1. Flow diagram of participants

Table 1. Patient (n=156) and physician (n=51, who rated 162 treatment decision-making processes) socio-demographic, and disease- or work-related characteristics

	N*	Percentage or mean ± SD
Patients		
Sex, female	67	43%
Age, years	156	67.5 ± 12.5
Education level	153	
Low	46	30%
Intermediate	43	28%
High	64	42%
Primary tumour type	156	
Gastro-intestinal	42	27%
Urological	36	23%
Breast	22	14%
Lung	17	11%
Haematological	13	8%
Gynaecological	10	6%
Other	16	11%
Treatment intent	154	
Curative	90	58%
Palliative	59	38%
Other	5	3%
Months since most recent cancer diagnosis	143	
0-3	66	46%
4-12	34	24%
>12	43	30%
Physicians		
Sex, female	24	47%
Age, years	51	44.4 ± 9.6
Years since start specialist training	51	15.8 ± 8.4
Hospital	52	
Academic (n=2)	33	65%
Non-academic (n=5)	18	35%
Specialty	51	
Radiotherapy	17	33%
Medical Oncology	11	22%
Urology	6	12%
Surgery	4	8%
Gynaecology	3	6%
Pulmonology	4	8%
Other	6	12%

*Numbers do not always add up to the total sample size, due to missing values.
SD=standard deviation

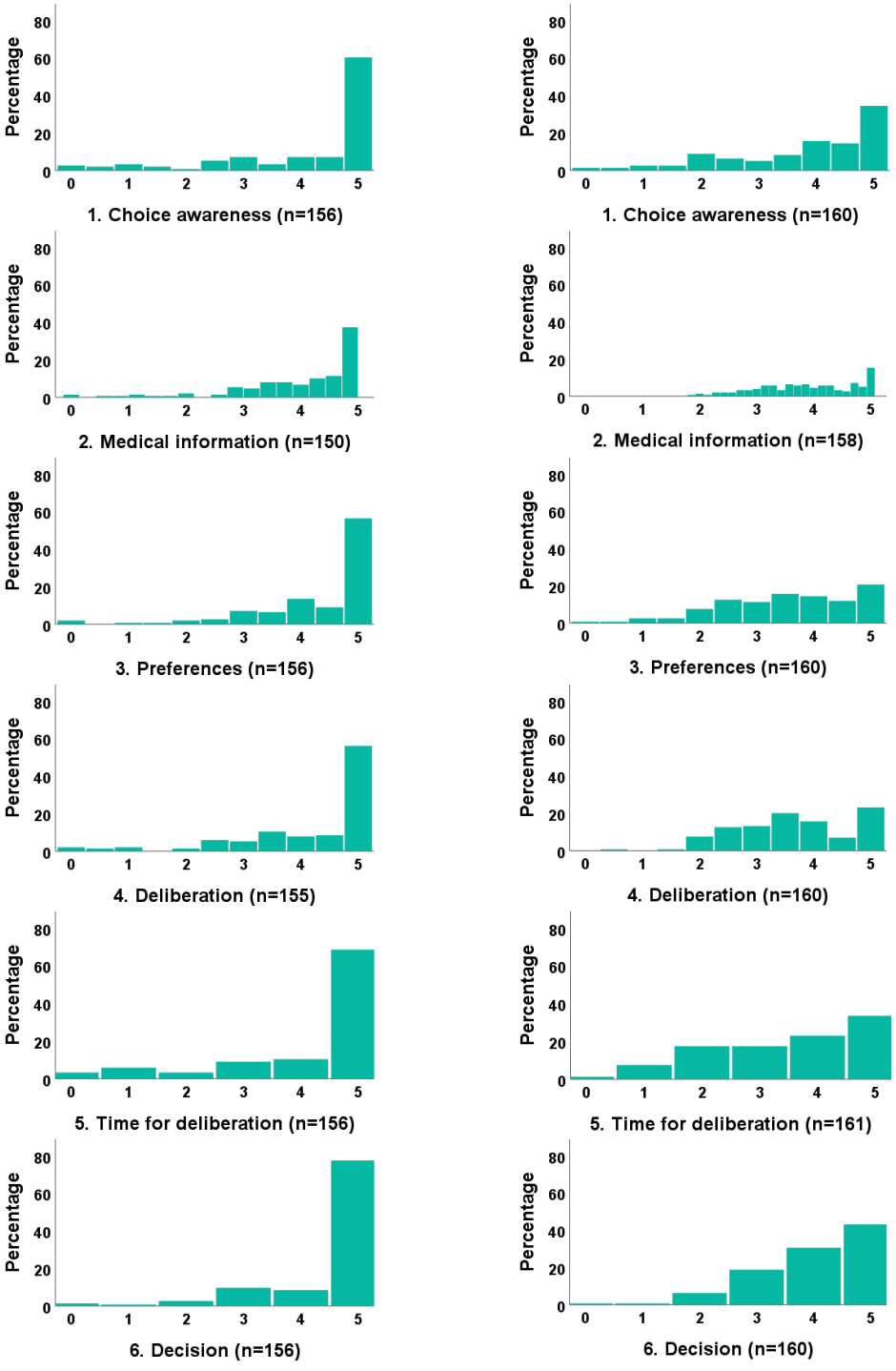


Figure 2. Dimension scores on the iSHARE questionnaires

Table 2. Median, interquartile range, range, and agreement on item level for the iSHAREpatient (n=156 treatment decision-making processes) and iSHAREphysician (n=162 treatment decision-making processes)

	iSHAREpatient		**iSHAREphysician**		**Agreement¯**	
	N	Median (IQR), range	N	Median (IQR), range	N	P (95%CI)
Item score (0-5)						
1. Physician explained advantages of treatment options	155	5.0 (4.0-5.0), 0-5	162	4.0 (4.0-5.0), 0-5	150	.79 (.71* - .84*)
2. Physician explained disadvantages of treatment options	156	5.0 (3.0-5.0), 0-5	162	4.0 (4.0-5.0), 2-5	151	.72 (.64* - .78*)
3. Physician explained (dis)advantages equally well	154	5.0 (3.0-5.0), 0-5	161	4.0 (3.0-5.0), 0-5	148	.68 (.60 - .76)
4. Physician checked patient's understanding of advantages	155	5.0 (4.0-5.0), 0-5	162	4.0 (3.0-5.0), 1-5	150	.67 (.59 - .75)
5. Physician checked patient's understanding of disadvantages	154	5.0 (4.0-5.0), 0-5	162	4.0 (3.0-5.0), 1-5	149	.65 (.57 - .73)
6. Physician told how treatment options differ	155	5.0 (3.0-5.0), 0-5	158	4.0 (3.0-5.0), 0-5	146	.63 (.55 - .71)
7. Patient asked for clarification	155	5.0 (3.0-5.0), 0-5	161	4.0 (2.0-5.0), 0-5	149	.60 (.52 - .68)
8. Physician said there is a choice	156	5.0 (3.0-5.0), 0-5	160	4.0 (3.0-5.0), 0-5	149	.65 (.57 - .73)
9. Physician said patient's opinion is important	156	5.0 (4.0-5.0), 0-5	161	4.0 (3.0-5.0), 0-5	150	.67 (.59 - .75)
10. Physician checked if he/she understood what is important for patient	156	5.0 (4.0-5.0), 0-5	160	4.0 (3.0-5.0), 0-5	149	.61 (.53 - .69)
11. Physician helped patient weighing (dis)advantages	156	5.0 (3.0-5.0), 0-5	161	4.0 (3.0-5.0), 0-5	150	.63 (.55 - .71)
12. Physician gave patient time for weighing (dis)advantages	156	5.0 (4.0-5.0), 0-5	161	4.0 (2.0-5.0), 0-5	150	.55 (.47 - .64)
13. Patient told physician what is important to him/her	156	5.0 (4.0-5.0), 0-5	161	4.0 (3.0-5.0), 0-5	150	.59 (.50 - .67)
14. Patient weighed (dis)advantages[#]	155	5.0 (4.0-5.0), 0-5	160	4.0 (3.0-5.0), 0-5	148	.59 (.51 - .68)
15. Decision takes into account what is important for patient[^]	140	5.0 (5.0-5.0), 0-5	141	4.0 (4.0-5.0), 0-5	124[¯]	.79 (.71* - .85*)
16. Physician discussed what patient needs for weighing options[^]	16	4.0 (3.0-5.0), 0-5	19	4.0 (2.0-4.0), 1-5	7[¯]	.43 (-.01 - .87)

[¯]Agreement was defined as the same item score or one point difference obtained by the iSHAREpatient and iSHAREphysician, since we considered it acceptable if scores from the respective viewpoints differed somewhat.

[#]In our development article[14] we recommend researchers to split this item if they would like to determine whether weighing the options happened during or outside the consultation. In this study we presented these items at the end of the patient questionnaire: I weighed up the advantages and disadvantages of the treatment options during the conversation. Median (IQR): 4.0 (3.0-5.0) I weighed up the advantages and disadvantages of the treatment options before or after the conversation. Median (IQR): 4.0 (2.0-5.0).

[¯]We only report agreement if both patient and physician agreed on whether a decision had been made; a treatment decision was made according to both patient and physician for 124 decision-making processes, while no treatment decision was made according to both patient and physician for seven decision-making processes. For 10 decision-making processes the patient indicated that a decision had been made, while no decision had been made according to the physician; it was the other way around for eight decision-making processes.

^We report item 15 if a patient/physician had reported that a decision had been made or if the response to that item was missing; we report item 16 if a patient/physician had reported that no decision had been made.
*Fleiss correction applied
CI = Confidence interval; IQR = interquartile range; P = proportion agreement

Table 3. Median and interquartile range for dimension and total scale scores of the iSHAREpatient (n=156 treatment decision-making processes) and iSHAREphysician (n=162 treatment decision-making processes), and for total and subscale scores of the comparison questionnaires

		Patient		**Physician**	
	Item	N*	Median (IQR)	N*	Median (IQR)
iSHARE dimension scores			**iSHAREpatient**		**iSHAREphysician**
1. Choice awareness (0-5)	8,9	156	5.0 (3.5-5.0)	160	4.0 (3.0-5.0)
2. Medical information (0-5)	1-7	150	4.4 (3.6-5.0)	158	3.9 (3.3-4.7)
3. Preferences (0-5)	10,13	156	5.0 (4.0-5.0)	160	3.5 (2.5-4.5)
4. Deliberation (0-5)	11,14	155	5.0 (3.5-5.0)	160	3.5 (3.0-4.5)
5. Time for deliberation (0-5)	12	156	5.0 (4.0-5.0)	161	4.0 (2.0-5.0)
6. Decision (0-5)	15 or 16	156	5.0 (5.0-5.0)	160	4.0 (3.0-5.0)
iSHARE total score (0-100)		149	95.0 (77.1-99.5)	155	75.0 (61.1-90.7)
			SDM-Q-9		**SDM-Q-Doc**
SDM-Q (0-100)		151	88.9 (71.1-97.8)	161	77.8 (69.4-88.9)
COMRADE					
Satisfaction with communication (0-100)		130	72.0 (63.1-78.2)		
Confidence in decision# (0-100)		130	78.7 (71.0-79.3)		
DCS# (0-100)		149	15.6 (5.5-25.8)		
Feeling uncertain# (0-100)		153	16.7 (0.0-41.7)		
Feeling uninformed (0-100)		152	16.7 (0.0-25.0)		
Feeling unclear about values (0-100)		151	25.0 (0.0-33.3)		
Feeling unsupported (0-100)		152	8.3 (0.0-25.0)		
Ineffective decision making (0-100)		153	0.0 (0.0-12.5)		
PEPPI-5 (5-25)		155	24.0 (20.0-25.0)		

*Numbers do not always add up to the total sample size, due to missing values.

#No a priori hypothesis was formulated regarding the correlation between this total or subscale score and either of the iSHARE questionnaires (Table 5); scores are reported for sake of information.

COMRADE = Combined Outcome Measure for Risk communication And treatment Decision-making Effectiveness; DCS = Decisional Conflict Scale; IQR = interquartile range; PEPPI-5 = five-item Perceived Efficacy in Patient-Physician Interactions; SDM-Q-9 = nine-item SDM-Questionnaire; SDM-Q-Doc = SDM-Questionnaire-physician version

Table 4. Test-retest agreement on item level for the iSHAREpatient (n=67 treatment decision-making processes)

	Agreement¯		**Agreement¯**
	N	P (95%CI)	P (95%CI)
1. Physician explained advantages of treatment options	67	.64 (.52 - .76)	.85 (.74* - .91*)
2. Physician explained disadvantages of treatment options	67	.60 (.47 - .72)	.88 (.77* - .93*)
3. Physician explained (dis)advantages equally well	67	.64 (.52 - .76)	.84 (.72* - .90*)
4. Physician checked patient's understanding of advantages	67	.70 (.58* - .79*)	.91 (.81* - .95*)
5. Physician checked patient's understanding of disadvantages	66	.61 (.48 - .73)	.85 (.73* - .91*)
6. Physician told how treatment options differ	67	.64 (.52 - .76)	.84 (.72* - .90*)
7. Patient asked for clarification	67	.60 (.47 - .72)	.79 (.67* - .86*)
8. Physician said there is a choice	67	.72 (.59* - .80*)	.87 (.76* - .92*)
9. Physician said patient's opinion is important	67	.79 (.67* - .86*)	.93 (.83* - .96*)
10. Physician checked if he/she understood what is important for patient	67	.69 (.57 - .81)	.91 (.81* - .95*)
11. Physician helped patient weighing (dis)advantages	67	.55 (.43 - .68)	.85 (.74* - .91*)
12. Physician gave patient time for weighing (dis)advantages	67	.64 (.52 - .76)	.91 (.81* - .95*)
13. Patient told physician what is important to him/her	67	.76 (.64* - .84*)	.94 (.85* - .97*)
14. Patient weighed (dis)advantages	67	.70 (.58* - .79*)	.91 (.81* - .95*)
15. Decision takes into account what is important for patient^	61	.84 (.71* - .90*)	.97 (.88* - .98*)
16. Physician discussed what patient needs for weighing options^	3	.00 (.03* - .56*)	.33 (-.37 - 1.03)

¯Agreement was defined as the same item score obtained both at test and retest.

¯Agreement was defined as the same item score obtained both at test and retest, or one point difference as post-hoc analysis.

^We report item 15 if a patient had reported that a decision had been made or if the response to that item was missing; we report item 16 if a patient had reported that no decision had been made.

*Fleiss correction applied

CI = confidence interval; P = proportion agreement

Table 5. Correlations between the iSHARE and other questionnaires

iSHARE questionnaire	Comparison scale - subscale		
		N	Spearman Rho*
iSHAREphysician	SDM-Q-Doc	155	.84⁻
iSHAREpatient	SDM-Q-9	144	.77⁻
	COMRADE – Satisfaction with communication	125	.68⁻
iSHAREpatient dimension (item)			
2. Medical information (1-7)	DCS – Feeling uninformed	146	-.44⁻
2. Medical information (7)[a]	PEPPI-5	154	.31⁻
3. Preferences (10,13)	DCS – Feeling unclear about values	151	-.43⁻
3. Preferences (13)[a]	PEPPI-5	155	.40⁻
4. Deliberation (14)[a]	PEPPI-5	154	.27
6. Decision (15)[^]	DCS – Ineffective decision making	138	-.46⁻
6. Decision (16)[^]	DCS – Feeling unsupported	15	-.66⁻

Note. The expected correlation was >.30 for the SDM-Q-9, SDM-Q-Doc, COMRADE and PEPPI-5, and <-.30 for the DCS.
[^]We report item 15 if a patient had reported that a decision had been made or if the response to that item was missing; we report item 16 if a patient had reported that no decision had been made.
*p<.01
[a] Items measuring patient behaviour
⁻Hypothesis was confirmed
COMRADE = Combined Outcome Measure for Risk communication And treatment Decision-making Effectiveness; DCS = Decisional Conflict Scale; IQR = interquartile range; PEPPI-5 = five-item Perceived Efficacy in Patient-Physician Interactions; SDM-Q-9 = nine-item SDM-Questionnaire; SDM-Q-Doc = SDM-Questionnaire – physician version

4. DISCUSSION AND CONCLUSION

4.1. Discussion

In this study, we determined the measurement properties of the iSHAREpatient and the iSHAREphysician designed to assess SDM in oncology. As opposed to many existing questionnaires, the iSHARE questionnaires are based on a clear definition of the construct, provide a comprehensive assessment of the SDM process in- and outside consultations, and allow the assessment of both patient and physician behaviours.[2, 14] We have conducted a large-scale study, including patients and physicians from academic and non-academic hospitals, physicians from different specialties, patients with a variety of cancer diagnoses, and with treatment intents being either curative or palliative. The current analyses have shown high dimension and total scores on both iSHARE questionnaires, and good construct validity of the iSHARE questionnaires. The iSHAREpatient showed substantial test-retest agreement. Further, the iSHARE questionnaires show moderate inter-rater agreement.

The iSHARE questionnaires, and especially the iSHAREpatient, showed high scores. More than 15% of the patients reported the highest possible score, which may be considered as a moderate ceiling effect.[38] Patient SDM questionnaires are known for ceiling effects.

These may be caused by the so-called halo effect, leading people to unconsciously alter their judgment of others' attributes based on their judgment of unrelated attributes.[39] To illustrate, if physicians are perceived to be friendly, the halo effect leads patients to evaluate their information-giving behaviours favourably instead of critically assessing them. Methods to reduce these effects, such as reflecting (stop-and-think) before rating the SDM process, have not been shown successful in patients.[3] We aimed to avoid ceiling effects by using an unbalanced response scale, that is, using a scale with more positively-labelled than negatively-labelled response options, thereby enabling more differentiation.[24] We further explicitly stated in the introduction of the iSHAREpatient that the questionnaire is not about satisfaction with the physician (Box 1).[14] However, these precautions do not seem to have adequately addressed the problem. The high scores may have resulted from recruiting physicians from our network (i.e., researcher selection bias), some of whom had been trained in SDM and whose patients may actually have experienced high levels of SDM. Moreover, physicians may have, consciously or unconsciously, selectively approached patients with whom the decision-making process was, or was expected to be, shared (i.e., physician selection bias). In addition, patients who declined participation may have been less involved in decision making (i.e., patient selection bias). A clear indication that our sample suffered from selection bias were the remarkably high scores on the other questionnaires too. Two recent studies in Dutch cancer patients[40, 41] showed substantially lower SDM-Q-9 scores and higher decisional conflict scores. In addition, two recent studies in Dutch cancer patients[42] and Dutch cancer survivors[43] showed somewhat lower patients' perceived self-efficacy compared to our sample. It is therefore important to await the scores in other samples before drawing definitive conclusions about the high scores. Of note, treatment decision making is often distributed across consultations and time[44] and half of the patients indeed thought about more than one consultation while completing the questionnaire.

The iSHARE questionnaires showed only very small numbers of missing values and no specific patterns, implicating acceptability of the items for both patients and physicians, and no systematic bias. Regardless, more research is needed on how to deal with missing values for instruments assessing formative constructs.

Our results demonstrated good construct validity (i.e., >75% of the results confirm our hypotheses)[45] of the iSHARE questionnaires. The iSHAREpatient and iSHAREphysician correlated highly (>.50) with the SDM-Q-9 and SDM-Q-Doc, indicating that the questionnaires measure the same construct.[46] The iSHARE questionnaires offer a more valid assessment of the SDM process since they cover both patient and physician behaviours. Hypotheses with regard to correlations with the COMRADE and DCS subscales were confirmed, adding to the proof for construct validity. Internal consistency of the DCS subscales seemed sub-optimal, a problem identified previously.[47] Further, two of three hypotheses regarding the PEPPI-5 were confirmed. To our knowledge no appropriate questionnaires were available at the time of designing the study for construct validity testing of any of the iSHAREphysician dimensions, nor for the Choice Awareness, Deliberation and Time for Deliberation dimensions of the iSHAREpatient. We recommend hypotheses testing for the other iSHARE dimensions once appropriate measurement instruments become available.

We determined test-retest agreement for the iSHAREpatient. This is a strength of the study, as this has not frequently been established for patient SDM questionnaires.[2] While several guidelines are available for kappa and intraclass correlations,[45, 48] we are not aware of any criteria to label the proportion agreement. Using the labels proposed for the kappa,[49] we propose that a proportion agreement of ≤.30 is 'slight'; >.30 'fair'; >.50 'moderate'; >.70 'substantial', and >.90 'almost perfect'. This results in substantial agreement for four, moderate for eleven, and slight for one of the iSHAREpatient items. Higher agreement may be found if the period between the two assessments is even shorter. The time period should be long enough, so that participants will not remember their previous answers; yet patients risk forgetting about their and their physician's behaviours if the period is too long. In addition, test-retest agreement of a questionnaire evaluating a decision-making process may be different from one evaluating, e.g., a state such as quality of life, or an attitude. Consequently, we did a post-hoc analysis in which we allowed the item scores to differ one point; agreement was almost perfect for seven items, substantial for eight items and fair for one item. All in all, the results demonstrate substantial test-retest agreement.

We applied the same criteria to the agreement between the iSHAREpatient and iSHAREphysician scores, allowing one point difference; agreement was substantial for three, moderate for 12 and fair for one item, demonstrating moderate inter-rater agreement overall. As noted, some physicians had been trained in SDM and may have reflected more critically on the decision process than their patients. Patients' and physicians' ratings of communication, including SDM in oncology[27, 28] are known to correlate poorly, but it should be noted that correlations are not the appropriate measure for agreement.[50, 51] Only few studies calculated the kappa and proportion agreement.[50] To the best of our knowledge, we are the first to have calculated proportion agreement for patient and physician SDM scores in oncology, which makes it hard to compare results. We aimed to achieve good inter-rater agreement by using the same underlying construct for both questionnaires, using the same items and most importantly, extensively involving both patients and physicians throughout the development process of the questionnaires.[14] We recommend future users of the iSHARE questionnaires to consider which perspective is most feasible to determine or to use both, bearing in mind that they represent different perspectives.

The iSHARE questionnaires contain two versions of the last item; for the majority of decision-making processes a decision had been made, so item 15 (The decision takes into account what is important for the patient) was reported. As a consequence there were not enough data to determine agreement for item 16 (The physician discussed what the patient needs to weigh the options). The iSHARE questionnaires may be applicable to healthcare settings outside of oncology, but we advise content validity testing first. We also recommend to determine cross-cultural validity when using the iSHARE questionnaires in languages other than Dutch. Finally, the findings should be considered in light of several limitations. As discussed, different forms of selection bias might have been present. Further, we aimed to include a broad range of patients, including in terms of education. Forty percent were highly educated, which may limit the representativeness of the sample for the patient population.

4.2. Conclusion

The iSHAREpatient and iSHAREphysician demonstrate good construct validity, substantial test-retest agreement (iSHAREpatient), and moderate inter-rater agreement. The dimension and total scores were high, which may have largely been caused by selection bias.

4.3 Practice Implications

Results obtained using the iSHARE questionnaires provide information about the entire SDM process, about both patient and physician behaviours, from the perspective of patient and/or physician, and may be administered before or after the final decision has been made. The results may inform both physician- and patient-directed efforts to improve SDM in clinical practice, and dimension scores can be used to determine the impact of interventions or training on specific aspects of the SDM process.

REFERENCES

1. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743–8.
2. Gärtner FR, Bomhof-Roordink H, Smith IP, Scholl I, Stiggelbout AM, Pieterse AH. The quality of instruments to assess the process of shared decision making: A systematic review. PLoS One 2018;13:e0191747.
3. Kunneman M, LaVecchia CM, Singh Ospina N, Abu Dabrh AM, Behnken EM, Wilson P, et al. Reflecting on shared decision making: A reflection-quantification study. Health Expect 2019;22:1165–72.
4. Blumenthal-Barby J, Opel DJ, Dickert NW, Kramer DB, Tucker Edmonds B, Ladin K, et al. Potential Unintended Consequences Of Recent Shared Decision Making Policy Initiatives. Health Aff (Millwood) 2019;38:1876–81.
5. Clayman ML, Gulbrandsen P, Morris MA. A patient in the clinic; a person in the world. Why shared decision making needs to center on the person rather than the medical encounter. Patient Educ Couns 2017;100:600–4.
6. Bomhof-Roordink H, Gärtner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. BMJ Open 2019;9:e031763.
7. Salzburg Global Seminar. Salzburg statement on shared decision making. BMJ 2011;342:d1745.
8. Norful AA, Dillon J, Baik D, George M, Ye S, Poghosyan L. Instruments to measure shared decision making in outpatient chronic care: a systematic review and appraisal. J Clin Epidemiol 2020;121:15–9.
9. Bouniols N, Leclere B, Moret L. Evaluating the quality of shared decision making during the patient-carer encounter: a systematic review of tools. BMC Res Notes 2016;9:382.
10. Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Shared decision making: developing the OPTION scale for measuring patient involvement. Qual Saf Health Care 2003;12:93–9 .
11. Elwyn G, Barr PJ, Grande SW, Thompson R, Walsh T, Ozanne EM. Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters. Patient Educ Couns 2013;93:102–7.
12. Kriston L, Scholl I, Holzel L, Simon D, Loh A, Harter M. The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. Patient Educ Couns 2010;80:94–9.
13. Scholl I, Kriston L, Dirmaier J, Buchholz A, Harter M. Development and psychometric properties of the Shared Decision Making Questionnaire–physician version (SDM-Q-Doc). Patient Educ Couns 2012;88:284–90.
14. Bomhof-Roordink H, Gärtner FR, van Duijn-Bakker N, van der Weijden T, Stiggelbout AM, Pieterse AH. Measuring shared decision making in oncology: Development and first testing of the iSHAREpatient and iSHAREphysician questionnaires. Health Expect 2020;23:496–508.
15. Kane HL, Halpern MT, Squiers LB, Treiman KA, McCormack LA. Implementing and evaluating shared decision making in oncology practice. CA Cancer J Clin 2014;64:377–88.
16. Politi MC, Studts JL, Hayslip JW. Shared decision making in oncology practice: what do oncologists need to know? Oncologist 2012;17:91–100.
17. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge University Press. Cambridge. 2011.
18. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539–49.
19. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651–7.

20. Mokkink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, de Vet HCW., et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Version July 2019. Available from: *https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf*. Date last accessed: 09-01-2021.

21. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995;15:25–30.

22. Edwards A, Elwyn G, Hood K, Robling M, Atwell C, Holmes-Rovner M, et al. The development of COMRADE–a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. Patient Educ Couns 2003;50:311–22.

23. Maly RC, Frank JC, Marshall GN, DiMatteo MR, Reuben DB. Perceived efficacy in patient-physician interactions (PEPPI): validation of an instrument in older persons. J Am Geriatr Soc 1998;46:889–94.

24. Bomhof-Roordink H, Gärtner FR, Stiggelbout AM, Pieterse AH. Measuring shared decision making: Choice of response scale matters. Abstract presented at the International Shared Decision Making Conference, Quebec City, Canada (2019). Available from: *https://fourwaves-sots.s3.amazonaws.com/static/media/uploads/2019/06/28/isdm2019-oralsessionsbooklet-2019-06-28.pdf*.

25. Dowie J. Shared decision making is a Preference-sensitive Formative Construct: the Implications. Eur J Pers Cent Healthc 2019;7:506–17.

26. Wollschlager D. Short communication: Where is SDM at home? putting theoretical constraints on the way shared decision making is measured. Z Evid Fortbild Qual Gesundhwes 2012;106:272–4.

27. Calderon C, Jimenez-Fonseca P, Ferrando PJ, Jara C, Lorenzo-Seva U, Beato C, et al. Psychometric properties of the Shared Decision-Making Questionnaire (SDM-Q- 9) in oncology practice. Int J Clin Health Psychol 2018;18:143–51.

28. Nejati B, Lin CC, Imani V, Browall M, Lin CY, Brostrom A, et al. Validating patient and physician versions of the shared decision making questionnaire in oncology setting. Health Promot Perspect 2019;9:105–14.

29. Calderon C, Ferrando PJ, Carmona-Bayonas A, Lorenzo-Seva U, Jara C, Beato C, et al. Validation of SDM-Q-Doc Questionnaire to measure shared decision- making physician's perspective in oncology practice. Clin Transl Oncol 2017;19:1312–9.

30. Rodenburg-Vandenbussche S, Pieterse AH, Kroonenberg PM, Scholl I, van der Weijden T, Luyten GP, et al. Dutch Translation and Psychometric Testing of the 9- Item Shared Decision Making Questionnaire (SDM-Q-9) and Shared Decision Making Questionnaire-Physician Version (SDM-Q-Doc) in Primary and Secondary Care. PLoS One 2015;10:e0132158.

31. van der Krieke L, Emerencia AC, Boonstra N, Wunderink L, de Jonge P, Sytema S. A web-based tool to support shared decision making for people with a psychotic disorder: randomized controlled trial and process evaluation. J Med Internet Res 2013;15:e216.

32. O'Connor A. User Manual—Decisional Conflict Scale. 1993. Updated 2010. Available from: *https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Decisional_Conflict.pdf*. Date last accessed: 09-01-2021.

33. Koedoot N, Molenaar S, Oosterveld P, Bakker P, de Graeff A, Nooy M, et al. The decisional conflict scale: further validation in two samples of Dutch oncology patients. Patient Educ Couns 2001;45:187–93.

34. ten Klooster PM, Oostveen JC, Zandbelt LC, Taal E, Drossaert CH, Harmsen EJ, et al. Further validation of the 5-item Perceived Efficacy in Patient-Physician Interactions (PEPPI-5) scale in patients with osteoarthritis. Patient Educ Couns 2012;87:125–30.

35. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737–45.

36. de Vet HCW, Dikmans RE, Eekhout I. Specific agreement on dichotomous outcomes can be calculated for more than two raters. J Clin Epidemiol 2017;83:85–9.

37. de Vet HCW, Mullender MG, Eekhout I. Specific agreement on ordinal and multiple nominal outcomes can be calculated for more than two raters. J Clin Epidemiol 2018;96:47–53.

38. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293–307.

39. Nisbett RE, DeCamp Wilson T. The Halo Effect: Evidence for Unconscious Alteration of Judgments. J Pers Soc Psychol 1977;35:250–6.

40. Henselmans I, van Laarhoven HWM, van Maarschalkerweerd P, de Haes H, Dijkgraaf MGW, Sommeijer DW, et al. Effect of a Skills Training for Oncologists and a Patient Communication Aid on Shared Decision Making About Palliative Systemic Treatment: A Randomized Clinical Trial. Oncologist 2020;25:e578–88.

41. Wieldraaijer T, de Meij M, Zwaard S, van Weert H, Wind J. Introducing a time out consultation with the general practitioner between diagnosis and start of colorectal cancer treatment: Patient-reported outcomes. Eur J Cancer Care (Engl) 2019;28:e13141.

42. Nguyen MH, Smets EM, Bol N, Loos EF, van Laarhoven HW, Geijsen D, et al. Tailored Web-Based Information for Younger and Older Patients with Cancer: Randomized Controlled Trial of a Preparatory Educational Intervention on Patient Outcomes. J Med Internet Res 2019;21:e14407.

43. van der Hout A, van Uden-Kraan CF, Holtmaat K, Jansen F, Lissenberg-Witte BI, Nieuwenhuijzen GAP, et al. Role of eHealth application Oncokompas in supporting self-management of symptoms and health-related quality of life in cancer survivors: a randomised, controlled trial. Lancet Oncol 2020;21:80–94.

44. Williams D, Edwards A, Wood F, Lloyd A, Brain K, Thomas N, et al. Ability of observer and self-report measures to capture shared decision-making in clinical practice in the UK: a mixed-methods study. BMJ Open 2019;9:e029485.

45. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147–57.

46. Schellingerhout JM, Heymans MW, Verhagen AP, de Vet HC, Koes BW, Terwee CB. Measurement properties of translated versions of neck-specific questionnaires: a systematic review. BMC Med Res Methodol 2011;11:87.

47. Lam WW, Kwok M, Liao Q, Chan M, Or A, Kwong A, et al. Psychometric assessment of the Chinese version of the decisional conflict scale in Chinese women making decision for breast cancer surgery. Health Expect 2015;18:210–20.

48. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011;64:96–106.

49. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

50. Rottele N, Schopf-Lazzarino AC, Becker S, Korner M, Boeker M, Wirtz MA. Agreement of physician and patient ratings of communication in medical encounters: A systematic review and meta-analysis of interrater agreement. Patient Educ Couns 2020;103:1873–82.

51. Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. BMC Med Res Methodol 2020;20:293.