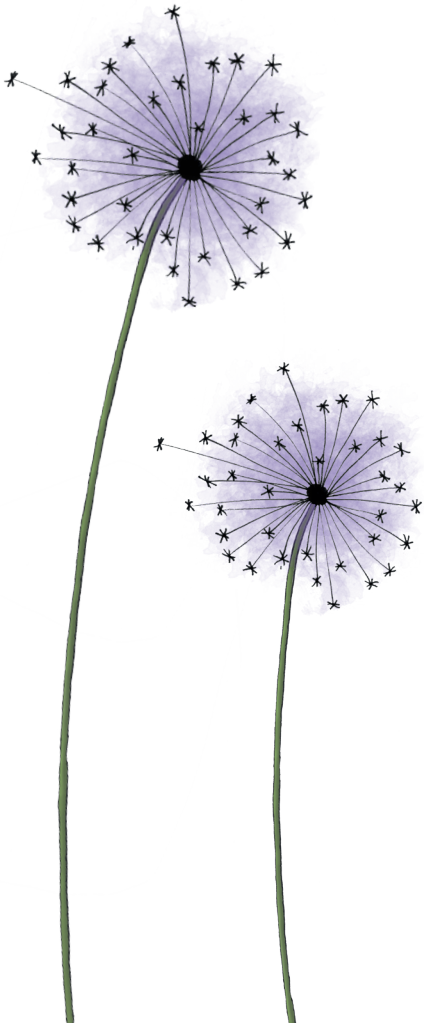
Measuring shared decision making in oncology: an informed approach

Bomhof-Roordink, H.



7

GENERAL DISCUSSION

Extensive research into shared decision making (SDM) between patient and healthcare professionals started in the 21th century, after the first models had been published in 1997[1] and 1999.[2] Research activities in the SDM field may be categorized in four different domains; *'(i) definition of SDM and development of frameworks, (ii) development and psychometric testing of measurement scales, (iii) development and evaluation of SDM interventions and (iv) implementation [of SDM] in routine practice.'*[3] While SDM interventions have been developed and campaigns to foster SDM implementation have been started, we noted that there were still measurement difficulties, and that SDM measurement was often limited to physician behaviour, while patient responsibilities were described in the first SDM models. Therefore, we set out to address the first two domains: we aimed to answer the following two fundamental questions in this thesis; *What is SDM and How can the SDM process be measured, specifically for oncology?*

We critically appraised the then available SDM measurement instruments, thereby informing researchers on their strengths and weaknesses. There was an overall lack of evidence on the quality of measurement instruments (Chapter 2). We showed both the variety and the consensus in components present in forty SDM models, what roles healthcare professionals and patients are described to have according to these models, and how the presence of certain components has varied over time. We presented an SDM map displaying the components seen as key per healthcare setting Chapter 3). In a qualitative study, we showed that SDM in oncology extends to time outside consultations and includes both patient and physician behaviours according to the relevant stakeholders (Chapter 4).

Informed by all these data and by using the original COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist,[4, 5] we developed the iSHARE questionnaires. The iSHAREpatient and ISHAREphysician assess the SDM process from different viewpoints, include both patient and physician behaviours, are aimed to assess the full SDM process, and may also be administered before a final decision has been made (Chapter 5). Finally, we validated the iSHARE questionnaires in a sample of patients and physicians, including patients with different cancer diagnoses and treated with varying intents, and physicians from differently specialties. The iSHARE questionnaires have adequate measurement properties and are fit to assess specific parts of the SDM process (Chapter 6).

In this chapter, we will discuss our main findings, reflect on the methods we have used and propose implications for research and clinical practice.

Actors in SDM

Early SDM models already identified patient behaviour to be part of this decision-making process, next to healthcare professional behaviour.[1, 2] In our review of SDM models, we found that patient and healthcare professionals are identified as actors in the majority of the models published until September 2019 (Chapter 3). Still, it might feel uncomfortable for healthcare professionals to give patients some responsibility, or to at least expect something from them, to achieve SDM. We know from qualitative research and from studies using the Control Preferences Scale (CPS) that patients want to participate in decision making and see

specific roles for themselves, both in- and outside oncology.[6-12] Indeed, in our qualitative study about SDM in oncology (Chapter 4) it became clear that patients have specific roles in SDM. Patients and members of the general population emphasized the importance of patients communicating openly, as was found before.[6, 13] This may help physicians in getting to know their patients, and patients need to feel 'known' to actively participate in SDM.[14] Healthcare professionals should see patients as a person, not as just a patient.[14, 15]

Healthcare professionals and developers of SDM interventions should be aware of the essence of time for patients' role in SDM. Time in which the patient can reflect on the decision, sleep on it, and discuss it with significant others if they wish to do so (Chapter 4). Also in an oncology setting there is often time to take a few days before making the final decision. This might implicate a different organization of healthcare, where time would be routinely scheduled between the consultation in which a diagnosis, or a change in the status of the disease, is communicated and possible treatment options are discussed, and the consultation in which treatment decisions are made. First having time to process and accept the diagnosis and thereafter considering the treatment options is expected to facilitate patient involvement.[16, 17]

Next to the patient and physician, also others may have a role in the SDM process, such as nurses,[18, 19] general practitioners,[20] caregivers,[21] family members, and significant others.[22] Roles of others have only recently been described and may receive more attention in future studies on defining and measuring SDM, but also when designing SDM interventions. A few existing examples are a scale measuring role competency of oncology nurses[23] and SDM interventions aiming at nurses, both in oncology[24] and in primary care.[25] Specifically for the oncology setting, we found that treatment options are often discussed with relatives or the general practitioner outside of the consultation with the specialist (Chapter 4).

While patient behaviour is part of the SDM process, existing SDM measurement instruments most often only assess healthcare professional behaviour (e.g., OPTION,[26] CollaboRATE[27]), or include both patient and physician behaviour in one item (e.g., nine-item SDM-Questionnaire (SDM-Q-9)[28] and SDM-Questionnaire-physician-version (SDM-Q-Doc)[29]). We urge researchers who develop SDM measurement instruments to include items dedicated to patient behaviours. In the iSHARE questionnaires, we included items on either physician or patient roles. The iSHARE questionnaires therefore can also be used to assess the effects of campaigns and interventions aiming at patient behaviour. Note that we deliberately chose to include only patient and physician behaviours. Others may be involved, but if they are not, SDM is no less.

Development and validation of the iSHARE questionnaires

We used the original version of the COSMIN checklist[4, 5] to rate the development and validation studies included in our systematic review (Chapter 2). We further used the checklist as a guideline during the development and validation process of the iSHARE questionnaires, although it had been developed to rate the methodological quality of studies on measurement properties. In recent years, many different COSMIN checklists and tools have been published: the risk of bias checklist[30] (which substitutes the original checklist)

and guideline,[31] both for systematic reviews of patient-reported outcome measures (PROM); the study design checklist;[32] the methodology for evaluating the content validity of PROMS;[33] the risk of bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments;[34] and the reporting guideline for studies on measurement properties.[35] We recommend researchers to use the appropriate COSMIN tools and to report on this, and especially to use the reporting guideline,[35] to make sure all necessary information is reported for future users of the measurement instrument.
To date, the guidelines available are relevant to the development and the validation of measurement instruments based on a reflective measurement model, and to a lesser extent when a formative measurement model is assumed. The former is more common. However, to us, such guidelines would have been useful, given that we assumed a formative measurement model for the iSHARE questionnaires. The COSMIN group might develop a guideline on how to deal with item selection and validation of a measurement instrument when a formative measurement model is assumed.

During all the phases of the development and the validation of the iSHARE questionnaires, we asked patients and physicians for their opinions and incorporated their feedback. Their involvement contributed to the design of questionnaires that turned out to be feasible during the cognitive interviews (Chapter 5). We also aimed to reach acceptable agreement between the scores on the iSHAREpatient and iSHAREphysician, by using this approach. The low number of missing values in the validation study (Chapter 6) suggested acceptability of both the iSHAREpatient and the iSHAREphysician.

We tested content validity quantitatively in a sample of 12 cancer patients and 11 physicians. The then available original COSMIN checklist did not make a difference between quantitative and qualitative data collection for content validity testing, and ≥10 was considered to be an adequate sample size.[4, 5] The more recent COSMIN study design checklist[32] (and the COSMIN risk of bias checklist[30, 36]) now recommends content validity testing using a sample size of >50 for quantitative studies. We performed content validity testing of the iSHAREphysician at the level of the domains (note, the 13 iSHARE domains are clustered in six dimensions), while the original COSMIN checklist (and the more recent COSMIN risk of bias checklist[30, 36] and COSMIN study design checklist[32]) states it should be done on the item level.[4, 5] We indeed asked patients to assess each item but decided against asking this from physicians, due to the time investment it would require. Somehow, it is strange though that we were hesitant to ask additional time investment of healthcare professionals to test the items while at the same time asking so much more time and effort of patients. We, as researchers, should consider the time we ask from both patient and healthcare professionals in the development and validation of measurement instruments. A measurement instrument is not directly beneficial to either patients or healthcare professionals, other than for example, a decision aid. For future development and validation studies it might be smart to reduce the individual burden by only presenting half of the items that should be assessed to one group and the other half to another group, thereby doubling the required sample.

We tested hypotheses about the correlation between the scores on the iSHAREpatient (dimensions) and the scores on several measurement instruments from the patient viewpoint

(the SDM-Q-9, the Decisional Conflict Scale (DCS), the Combined Outcome Measure for Risk communication And treatment Decision-making Effectiveness, and the five-item Perceived Efficacy in Patient-Physician Interactions (PEPPI-5)) to determine construct validity. We were only able to determine construct validity of the iSHAREphysician by formulating a hypothesis about the correlation between the scores on the iSHAREphysician and the scores on the SDM-Q-Doc. At the start of the validation study, no other measurement instrument was available to assess the SDM process from the viewpoint of the physician that had been validated in Dutch. This is still true today. Also, no questionnaires measuring related constructs from the physician viewpoint are available, on which to base hypotheses for validation. When these will become available, it will be valuable to formulate hypotheses on the dimension level to further validate the iSHAREphysician.

We decided to not compare the iSHARE questionnaires to a measurement instrument assessing the SDM process from an observer viewpoint (e.g., to the OPTION-12 or the OPTION-5) for two reasons. First and foremost, poor correlations have been found between scores on self- versus observer-based SDM measurement instruments.[3, 37, 38] The different perspectives on the SDM process may be due to differences in the construct underlying the respective measurement instruments, as well as to an inherently different view on the extent to which SDM occurred from these different perspectives. If the correlation between the scores on the iSHARE questionnaires and an observer-based coding scheme was low, we should doubt whether this informs us about the validity of the iSHARE questionnaires. Second, we would have faced logistic challenges which we could not solve in a pragmatically enough manner to be able to collect the necessary data.

In the field, Pearson[39, 40] and Spearman[38] correlations have been reported as parameters of test-retest, inter- and intra-rater reliability (Chapter 2). However, the COSMIN group recommends to calculate intraclass correlation coefficients for continuous scores and (weighted) Kappa's for dichotomous, nominal, and ordinal scores as reliability parameters.[32, 34] We were not able to calculate weighted Kappa's because the data were not distributed normally (Chapter 6) and therefore calculated agreement, which is considered to be a parameter of measurement error for dichotomous, nominal, and ordinal scores. To determine measurement error for continuous scores, calculation of the Standard Error of Measurement, Smallest Detectable Change, Limits of Agreement or Coefficient of Variation is recommended.[32, 34] We call upon the field to calculate and report both on reliability and measurement error, if possible.

Response scales

In a side-study, we tested four different response scales to determine the most appropriate one for the iSHARE questionnaires (Chapter 5). We were in particular interested in possible ceiling effects, as SDM measurement instruments from the patient viewpoint have shown to be prone to them.[41-43] We compared the scores obtained with the different response scales, using the draft version of the iSHAREpatient questionnaire. We had anticipated that a five-point scale ranging from 'not done at all' to 'done completely' would show fewer ceiling effects compared to a five-point scale ranging from 'totally disagree' to 'totally agree', due to the focus on actual behaviour in the response options. Further, respondents tend to agree

with questionnaire items regardless of their content, referred to as acquiescence bias.[44] We also tested a five-point positively unbalanced scale, ranging from 'not done at all' to 'done completely', i.e., with more labels on the positive end of the scale. The scale provided more choice and detail if someone would like to rate the item positively. This unbalanced format is known to reduce ceiling effects, when compared to a balanced scale in patient satisfaction measurement.[45] Finally, we included a visual analogue scale (VAS) with the ends labelled as 'not done at all' and 'done completely', as the VAS is known to show fewer ceiling effects compared to a Likert scale.[46]

The results obtained by the VAS scale showed fine results for the mean, standard deviation, range, and the coefficient of variation. Upon further inspection, the histogram showed a bimodal distribution (i.e., two distant peaks in the distribution), which can indicate inconsistent use of response options.[47] This underscores the importance of visual inspection of collected data. During the cognitive interviews we also saw problems with the completion of the VAS, and we decided against the use of the VAS scale based on these observations.

The five-point unbalanced 'done' scale showed favourable results as well. Informed by these findings, and in combination with the results of the cognitive interviews, we decided to use a six-point scale with two negative response options ('not at all' [helemaal niet gedaan] and 'hardly' [bijna niet gedaan]), and four positive response options ('a little' [een beetje gedaan], 'for a large part' [voor een groot deel gedaan], 'almost completely' [bijna helemaal gedaan], and 'completely' [helemaal gedaan]). Despite our attempts to limit ceiling effects, they were undeniably present in our validation study and scores were even higher (the SDM-Q-9[48, 49] and the PEPPI-5[48, 50, 51]) and lower (the DCS[48, 49]) than in other Dutch oncology samples. We have discussed the possibility of researcher, physician, and patient bias (Chapter 6). Another explanation might be that patients and physicians did not closely read the labels and may have in fact used it as a 'balanced' scale. A different approach, in which the response options are presented as words that should be circled instead of boxes that should be checked[52] might result in patients and physician closely reading the labels.

We encourage researchers in the SDM field to test different response scales when developing a measurement instrument. This enables them to choose the one with the largest range, the most variation, and fewest ceiling or floor effects. Next to these quantitative parameters, the focus of research on response scales should also be on their interpretability and feasibility for respondents.

Using the iSHARE questionnaires

The iSHARE questionnaires were developed and tested in Dutch, implying exclusion of cancer patients who do not speak or read Dutch. We involved patients throughout the development process and made efforts to formulate clear items, but the samples in which we tested feasibility and acceptability were highly educated (Chapters 4, 5 and 6). We therefore do not know whether low literate patients may experience difficulties in understanding the items. We did not explicitly test the iSHAREpatient in low literate patients and recommend developers of SDM instruments to do so. Users of the iSHARE questionnaire and other measurement instruments may perform additional testing in patients with various levels of education and

health literacy, especially when translating existing measurement instruments.

The iSHARE questionnaires may be used to establish baseline levels of SDM in a particular setting, or a change in the SDM level due to an intervention or training. We expect that the iSHARE questionnaires show fewer ceiling effects in a new sample compared to the effects demonstrated in our validation study (Chapter 6), allowing the detection of improvement. Of course, responsiveness of the iSHARE questionnaires should be assessed first. Responsiveness is a measurement property that is seldom assessed; it was only done for the CollaboRATE (Chapter 2). Responsiveness needs to be assessed in a longitudinal design in which a measurement instrument is administered twice and a change should be expected between the two assessments, e.g., as a result of an intervention. A priori hypotheses on the change scores need to be formulated. A feasible option to assess responsiveness, is to include the questionnaire under study (e.g., the iSHAREpatient) next to another questionnaire (e.g., the SDM-Q-9) which is used to evaluate the effect of an intervention to foster SDM. The construct approach may then be used,[32, 53] in which hypotheses are formulated about the expected direction and magnitude of correlations between change scores on the SDM-Q-9 and the iSHAREpatient. Note that in this case, the data collected for the iSHAREpatient should only be used to determine responsiveness of the iSHAREpatient and not to draw conclusions about the effect of the intervention. Apart from the possibility of the iSHARE questionnaires and other questionnaires to measure change over time, it might be valuable to discuss what change is clinically relevant in what context, compared to being statistically significant. We call upon the field to determine the minimal important change values in future research, and anchor-based methods may be used to that end.[54, 55] Future users of the iSHARE questionnaires who aim to use the questionnaires to assess the effect of an intervention, may critically review its dimensions and individual items, to determine in advance on which of them their intervention might have an effect. If only relevant dimensions or items are included in a questionnaire, study load for both patients and healthcare professional decreases. For example, items 9 and 10 (i.e., 9. The doctor said that it matters what I think is important and 10. The doctor checked whether he/she understood what was important to me) may show higher scores when a patient has completed a PROM and patient and healthcare professional have discussed it during the consultation. Although PROMs are increasingly linked to SDM,[56, 57] their impact on SDM has not yet been assessed. In the future, specific dimensions of the iSHARE questionnaires might be used to do so. We also encourage future users to critically assess whose viewpoint should be measured. Since the agreement between the iSHARE questionnaires is moderate, it might be enough to assess only one viewpoint, thereby reducing study burden.

Our validation study (Chapter 6) indicates that for the majority of decision-making processes (131/149), patients and physicians agree on whether a decision has been made. This is higher than was found previously in an oncology setting.[58] We cannot be sure whether this agreement was specific for our sample, and therefore ask future users of both iSHARE questionnaires to report their findings.

We described how to calculate a total score on both iSHARE questionnaires (Chapter 5) and reported this in our validation study (Chapter 6) to inform future users. We urge them to

carefully consider the use of total versus dimensions scores, since dimension scores are more informative for a formative construct. We do not explicitly describe weighing of the dimensions to arrive at a total score (Chapter 5), but we did so by combining dimension scores into a total score, instead of computing a total score based on individual items. Combining dimension scores results in proportionally giving more weight to dimensions with fewer items. If we would have computed total scores based on individual items, the total score would largely be determined by the seven of 15 items that refer to asking for and providing medical information. The information component makes up a large part of SDM measurement instruments e.g.,[59] and therefore has a major impact on the obtained SDM scores. We consider all dimensions to be of equal importance and call on future users to follow this approach.

Measuring SDM in future research

In 2017, we identified 40 SDM measurement instruments, including 21 original versions, four revised versions, and 15 translated versions (Chapter 2). In subsequent years, adapted (e.g., CollaboRATE^{pediatric60}) and translated versions (e.g., the Japanese versions of the SDM-Q-9 and SDM-Q-Doc[61, 62] and an Arabic version of the SDM-Q-9[63]) have been published, as well as papers on the development and validation of a patient questionnaire; the SDM Process Scale.[64-66] The SureScore[67] and the Alberta Shared decision-maKing Measurement Instrument (ASK-MI)[68] questionnaires have recently been developed to assess SDM, with both a patient and a clinician version. Also, a new observer measurement instrument has been developed, the 4SDM.[49] Both the 4SDM and the OPTION[12] were able to detect change as a result of an SDM training for oncologists, and we therefore recommend to further validate the 4SDM (e.g., content validity, construct validity, intra-rater and inter-rater agreement) and to publish about its development and validation process.

We recommend further validation and reliability testing of existing instruments, including the iSHARE questionnaires. In addition to further validation, we call authors to always describe their construct when developing or validating a measurement instrument, or to explicitly refer to a source in which the construct is described. It will help future users to determine whether the construct matches their SDM model, and thus whether the measurement instrument is useful to assess SDM in their study.

If one assumes a somewhat different underlying SDM construct than the ones underlying any of the already existing measurement instruments, one may edit an existing measurement instrument to meet one's purposes. However, authors should not simply edit items or remove or add items and refer to the original questionnaire without mentioning the changes. The known measurement properties do not longer apply to an adapted version. One solution is to explicitly describe the changes and the reason for them in the Methods section of an article, or to include the adapted version in an Appendix. Another approach is to present it as an officially adapted version of an existing measurement instrument and to give it a new name. An example of adapting an already existing and validated measurement instrument to match another SDM model, is the adaptation of the OPTION-5 into OPTION^{MCC}.[69] The same may apply if SDM needs to be measured from another viewpoint. CollaboRATE, a patient questionnaire, was adapted into the CollaboRATE^{pediatric} in German to assess SDM from the

patient, parent, and parent-proxy viewpoint.[60] The SDM-Q-9 and SDM-Q-doc have been adapted into the Care SDM-Questionnaire for care receivers (SDM-C-patient) and the Care SDM-Questionnaire for care providers (SDM-C-provider), to measure SDM between patient and healthcare professionals other than physicians.[70] We strongly recommend to determine the measurement properties of these instruments. Adaptations of existing measurement instruments may also include adding an assessment of behaviour of others, next to that of the patient and the physician.

It might be valuable to set up an international item bank for SDM, given the many measurement instruments available (Chapter 2). Especially if a formative measurement model is assumed and given that many different SDM models exist, researchers might benefit from the opportunity to create their own SDM measurement instrument based on the SDM model that fits their setting best. Researchers then do not have to formulate new, unvalidated items, and can compose a fitting combination of validated dimensions or items, which matches their specific construct. For this reason, we need more data on measurement properties on both the dimension and item levels, over and above information on a total scale level. We provide evidence on agreement on the item level for the iSHARE questionnaires, and for some iSHAREpatient items hypothesis testing was done on the item level (Chapter 6). Item scores for the SDM-Q-9 have also been published.[42] To the extent that such evidence is available for other SDM measurement instruments, it may be included in the item bank to inform future users. The dimensions and items should have a clear description and all information available on measurement properties in different settings should be reported. Validated versions of the items in other languages may also be included. Note that item response theory (IRT) only applies to reflective, and not formative measurement models. If one assumes a reflective measurement model, IRT is normally used to determine the item characteristic curves (i.e., a plot that shows the association between a patient's underlying ability and the probability of a particular response to the item) of the items. Item difficulty and patient ability are linked to each other in an IRT model,[71] and may inform researchers on which items to select from the item bank, but this approach is not applicable to an SDM item bank. SDM items do not by definition differ in difficulty, other than for example items on walking ability (a unidimensional construct). Someone who cannot walk will by definition not be able to run either and will answer items accordingly. Researchers should instead select items from the item bank based on their content.

We studied the SDM process, which includes time during and outside consultations. These consultations were face-to-face consultations. This thesis started with a comparison between the complex behaviours needed for this decision-making process, in comparison to the more easy-to-implement behaviours recommended to slow the spread of COVID-19. COVID-19 has changed much in the main SDM playing field: the consultation room. Patients have been requested to come alone, not to shake hands, to wear a mask, and to keep physical distance whenever possible. They have increasingly been invited to digital appointments instead of in-person ones. Some of these and other effects of the pandemic on healthcare delivery most probably will stay in the coming years. We do not yet know how this may affect communication, and more specifically SDM between patients and healthcare professionals. Attention should be paid to what different behaviours it may potentially require from the

participants. The literature on remote SDM focusses on technological features and less on how conducting consultations remotely affects collaboration between patients and healthcare professionals.[72] A study with simulated consultations showed that perceived SDM did not differ significantly between face-to-face and screen-to-screen consultation,[73] which is promising. However, to draw firm conclusions on how remote consultations impacts the interaction between patients and healthcare professionals and SDM, more research in real-life clinical practice is needed.

SDM as the norm

SDM between patients and healthcare professionals should become the norm, especially when the patient wishes to be involved. Indeed, even in a more acute situation SDM is the preferred approach to be used when there is clinical equipoise or uncertainty about the best approach for this patient, when the patient can be involved, and when there is sufficient time to do so.[74] If a patient prefers to be involved but is not yet ready to be, due to e.g., patient-related characteristics, healthcare professionals' directed efforts to help make the patient ready may enable participatione.[75] The given that SDM slowly becomes the norm is reflected in its establishment by Dutch law[76] and its promotion by healthcare professional organisations, patient organisations, and the government.[77] As a consequence, there might be the tendency to formulate norms regarding the required level of SDM in a certain setting. In Italy for example, the level of SDM was proposed as a quality indicator for breast cancer care.[78] However, it is questionable if making SDM a requirement for organizational accreditation would benefit SDM implementation.[79] In fact, we call insurers and policy makers to refrain from benchmarking for SDM. The SDM field has still work to do to guarantee valid and meaningful measurement instruments for different contexts, to ensure that those in charge of SDM implementation and reimbursement know wat they should measure.[80] When the level of SDM is used as a quality indicator or is used for accreditation, one should be aware that the way it is measured may have a major impact on what behaviour is actually promoted. It may result in rewarding easy to measure healthcare professional behaviour, such as the use of decision aids, instead of promoting the awareness of all that is needed to truly involve patients in making decisions about their care.[81, 82]

Conclusion

Healthcare professionals and researchers should be aware of the role that patients have in SDM and enable their participation. To that end, healthcare professionals should get to know their patients and see them as a person with a condition, and not only as a patient. Providing time to consider options outside of the consultation may be an important facilitator of patient involvement. Intervention developers should design them in such a way that the interventions support patients in their unique role, which may differ depending on the setting.

The iSHARE questionnaires assess both patient and physician behaviours and cover the entire SDM process. Both patients and physicians were involved in all steps of the development and validation and they show adequate measurement properties. We recommend use of the iSHARE questionnaires in an oncology setting.

The variety of existing SDM models and measurement instruments is not necessarily a problem and may be a natural result of the formative nature of the SDM construct. The SDM map may be used to determine which SDM components are relevant in specific setting. Both SDM models and SDM measurement instruments benefit from a clear description, as this enables future users to apply them in appropriate ways. For SDM measurement instruments, including the iSHARE questionnaires, further high-quality validation studies are needed, and especially responsiveness should be assessed. We recommend using existing measurement instruments, by adapting and renaming them if needed, or by building an item bank, enabling researchers to compose a fitting combination of items or dimensions. Instrument developers should consider the assessment of patients' role and the formative measurement model. Finally, we recommend to always involve patients and healthcare professionals in the development and validation of SDM measurement instruments. While this work continues, more knowledge on SDM measurement will become available. This will help finding answers to challenges still present in the field.

REFERENCES

1. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med.1997;44:681-92.
2. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med. 1999;49:651-61.
3. Scholl I, Kriston L, Dirmaier J, Harter M. Comparing the nine-item Shared Decision-Making Questionnaire to the OPTION Scale - an attempt to establish convergent validity. Health Expect. 2015;18:137-50.
4. 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.
5. 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.
6. Shay LA, Lafata JE. Understanding patient perceptions of shared decision making. Patient Educ Couns. 2014;96:295-301.
7. Holland WC, Hunold KM, Mangipudi SA, Rittenberg AM, Yosipovitch N, Platts-Mills TF. A Prospective Evaluation of Shared Decision-making Regarding Analgesics Selection for Older Emergency Department Patients With Acute Musculoskeletal Pain. Acad Emerg Med. 2016;23:306-14.
8. Hahlweg P, Kriston L, Scholl I, Brahler E, Faller H, Schulz H, et al. Cancer patients' preferred and perceived level of involvement in treatment decision-making: an epidemiological study. Acta Oncol. 2020:1-8.
9. Josfeld L, Keinki C, Pammer C, Zomorodbakhsch B, Hubner J. Cancer patients' perspective on shared decision-making and decision aids in oncology. J Cancer Res Clin Oncol. 2021;147:1725-32.
10. Moth E, McLachlan SA, Veillard AS, Muljadi N, Hudson M, Stockler MR, et al. Patients' preferred and perceived roles in making decisions about adjuvant chemotherapy for non-small-cell lung cancer. Lung Cancer. 2016;95:8-14.
11. Brom L, Hopmans W, Pasman HR, Timmermans DR, Widdershoven GA, Onwuteaka-Philipsen BD. Congruence between patients' preferred and perceived participation in medical decision-making: a review of the literature. BMC Med Inform Decis Mak. 2014;14:25.
12. Noteboom EA, May AM, van der Wall E, de Wit NJ, Helsper CW. Patients' preferred and perceived level of involvement in decision making for cancer treatment: A systematic review. Psychooncology. 2021;30:1663-79.
13. Hamann J, Kohl S, McCabe R, Buhner M, Mendel R, Albus M, et al. What can patients do to facilitate shared decision making? A qualitative study of patients with depression or schizophrenia and psychiatrists. Soc Psychiatry Psychiatr Epidemiol. 2016;51:617-25.
14. Thorne S, Oliffe JL, Stajduhar KI. Communicating shared decision-making: cancer patient perspectives. Patient Educ Couns. 2013;90:291-6.
15. 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.
16. Joseph-Williams N, Elwyn G, Edwards A. Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns. 2014;94:291-309.
17. Keij SM, van Duijn-Bakker N, Stiggelbout AM, Pieterse AH. What makes a patient ready for Shared Decision Making? A qualitative study. Patient Educ Couns. 2021;104:571-7.

18. Bos-van den Hoek DW, Thode M, Jongerden IP, Van Laarhoven HWM, Smets EMA, Tange D, et al. The role of hospital nurses in shared decision-making about life-prolonging treatment: A qualitative interview study. J Adv Nurs. 2021;77:296-307.

19. Savelberg W, Boersma LJ, Smidt M, Goossens MFJ, Hermanns R, van der Weijden T. Does lack of deeper understanding of shared decision making explains the suboptimal performance on crucial parts of it? An example from breast cancer care. Eur J Oncol Nurs. 2019;38:92-7.

20. Dobler CC, Midthun DE, Montori VM. Quality of Shared Decision Making in Lung Cancer Screening: The Right Process, With the Right Partners, at the Right Time and Place. Mayo Clin Proc. 2017;92:1612-6.

21. Schuster F, Holzhuter F, Heres S, Hamann J. 'Triadic' shared decision making in mental health: Experiences and expectations of service users, caregivers and clinicians in Germany. Health Expect. 2021;24:507-15.

22. Legare F, Stacey D, Gagnon S, Dunn S, Pluye P, Frosch D, et al. Validating a conceptual model for an inter-professional approach to shared decision making: a mixed methods study. J Eval Clin Pract. 2011;17:554-64.

23. Tariman JD, Katz P, Bishop-Royse J, Hartle L, Szubski KL, Enecio T, et al. Role competency scale on shared decision-making nurses: Development and psychometric properties. SAGE Open Med. 2018;6:2050312118807614.

24. Berger-Hoger B, Liethmann K, Muhlhauser I, Steckelberg A. Implementation of shared decision-making in oncology: development and pilot study of a nurse-led decision-coaching programme for women with ductal carcinoma in situ. BMC Med Inform Decis Mak. 2017;17:160.

25. Lenzen SA, Daniels R, van Bokhoven MA, van der Weijden T, Beurskens A. Development of a conversation approach for practice nurses aimed at making shared decisions on goals and action plans with primary care patients. BMC Health Serv Res. 2018;18:891.

26. 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.

27. 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.

28. 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.

29. 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.

30. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Qual Life Res. 2018;27:1171-9.

31. 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.

32. 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.

33. Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27:1159-70.

34. 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.

35. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res. 2021;30:2197-218.

36. Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist. Version July 2018. Available from: *https://www.cosmin.nl/wp-content/uploads/COSMIN-RoB-checklist-V2-0-v17_rev3.pdf*. Date last accessed: 09-02-2022.

37. Geessink NH, Ofstad EH, Olde Rikkert MGM, van Goor H, Kasper J, Schoon Y. Shared decision-making in older patients with colorectal or pancreatic cancer: Determinants of patients' and observers' perceptions. Patient Educ Couns. 2018;101:1767-74.

38. Kasper J, Heesen C, Kopke S, Fulcher G, Geiger F. Patients' and observers' perceptions of involvement differ. Validation study on inter-relating measures for shared decision making. PLoS One. 2011;6:e26255.

39. Barr PJ, O'Malley AJ, Tsulukidze M, Gionfriddo MR, Montori V, Elwyn G. The psychometric properties of Observer OPTION(5), an observer measure of shared decision making. Patient Educ Couns. 2015;98:970-6.

40. Martin LR, DiMatteo MR, Lepper HS. Facilitation of patient involvement in care: development and validation of a scale. Behav Med. 2001;27:111-20.

41. 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.

42. Rencz F, Tamasi B, Brodszky V, Gulacsi L, Weszl M, Pentek M. Validity and reliability of the 9-item Shared Decision Making Questionnaire (SDM-Q-9) in a national survey in Hungary. Eur J Health Econ. 2019;20:43-55.

43. 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.

44. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol. 2003;88:879-903.

45. Moret L, Nguyen JM, Pillet N, Falissard B, Lombrail P, Gasquet I. Improvement of psychometric properties of a scale measuring inpatient satisfaction with care: a better response rate and a reduction of the ceiling effect. BMC Health Serv Res. 2007;7:197.

46. Voutilainen A, Pitkaaho T, Kvist T, Vehvilainen-Julkunen K. How to ask about patient satisfaction? The visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric Likert scale. J Adv Nurs. 2016;72:946-57.

47. Huebner M, Vach W, le Cessie S. A systematic approach to initial data analysis is good research practice. J Thorac Cardiovasc Surg. 2016;151:25-7.

48. 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.

49. 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-e88.

50. 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.

51. 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.

52. Hendriks AA, Vrielink MR, Smets EM, van Es SQ, De Haes JC. Improving the assessment of (in) patients' satisfaction with hospital care. Med Care. 2001;39:270-83.

53. Mokkink L, Terwee C, de Vet H. Key concepts in clinical epidemiology: Responsiveness, the longitudinal aspect of validity. J Clin Epidemiol. 2021;140:159-62.

54. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010;63:524-34.

55. Terwee CB, Peipert JD, Chapman R, Lai JS, Terluin B, Cella D, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. Qual Life Res. 2021;30:2729-54.

56. Damman OC, Jani A, de Jong BA, Becker A, Metz MJ, de Bruijne MC, et al. The use of PROMs and shared decision-making in medical encounters with patients: An opportunity to deliver value-based health care to patients. J Eval Clin Pract. 2020;26:524-40.

57. Vilans. Samen beslissen met de vragenlijst TOPICS-SF. Available from: *https://www.zorgvoorbeter.nl/zorgvoorbeter/media/documents/thema/persoonsgerichte-zorg/vragenlijst-topics-sf-samen-beslissen.pdf*. Date last accessed: 09-02-2022.

58. Leppin AL, Humeniuk KM, Fernandez C, Montori VM, Yost K, Kumbamu A, et al. Was a decision made? An assessment of patient-clinician discordance in medical oncology encounters. Health Expect. 2015;18:3374-81.

59. Kunneman M, Henselmans I, Gärtner FR, Bomhof-Roordink H, Pieterse AH. Do Shared Decision-Making Measures Reflect Key Elements of Shared Decision Making? A Content Review of Coding Schemes. Med Decis Making. 2019;39:886-93.

60. Herrler A, Gorig T, Georg S, De Bock F, Ullrich C, Eichinger M. Assessment of shared decision making in pediatrics: Developing German scales for patients aged 7-18 years, parents and parent-proxy reports (CollaboRATE(pediatric)). Patient Educ Couns. 2021;104:634-41.

61. Goto Y, Miura H, Son D, Arai H, Kriston L, Scholl I, et al. Psychometric Evaluation of the Japanese 9-Item Shared Decision-Making Questionnaire and Its Association with Decision Conflict and Patient Factors in Japanese Primary Care. JMA J. 2020;3:208-15.

62. Goto Y, Miura H, Son D, Scholl I, Kriston L, Harter M, et al. Association between physicians' and patients' perspectives of shared decision making in primary care settings in Japan: The impact of environmental factors. PLoS One. 2021;16:e0246518.

63. Alzubaidi H, Hussein A, Mc Namara K, Scholl I. Psychometric properties of the Arabic version of the 9-item Shared Decision-Making Questionnaire: the entire process from translation to validation. BMJ Open. 2019;9:e026672.

64. Valentine KD, Mancini B, Vo H, Brodney S, Cosenza C, Barry MJ, et al. Using Standardized Videos to Examine the Validity of the Shared Decision Making Process Scale: Results of a Randomized Online Experiment. Med Decis Making. 2021:272989X211029267.

65. Valentine KD, Vo H, Fowler FJ, Jr., Brodney S, Barry MJ, Sepucha KR. Development and Evaluation of the Shared Decision Making Process Scale: A Short Patient-Reported Measure. Med Decis Making. 2021;41:108-19.

66. Fowler FJ, Jr., Sepucha KR, Stringfellow V, Valentine KD. Validation of the SDM Process Scale to Evaluate Shared Decision-Making at Clinical Sites. J Patient Exp. 2021;8:23743735211060811.

67. 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.

68. Manhas KP, Olson K, Churchill K, Faris P, Vohra S, Wasylak T. Measuring shared decision-making and collaborative goal setting in community rehabilitation: a focused ethnography using cross-sectional surveys in Canada. BMJ Open. 2020;10:e034745.

69. Pel-Littel RE, Buurman BM, van de Pol MH, Yilmaz NG, Tulner LR, Minkman MM, et al. Measuring triadic decision making in older patients with multiple chronic conditions: Observer OPTION(MCC). Patient Educ Couns. 2019;102:1969-76.

70. Goto Y, Yamaguchi Y, Onishi J, Arai H, Harter M, Scholl I, et al. Adapting the patient and physician versions of the 9-item shared decision making questionnaire for other healthcare providers in Japan. BMC Med Inform Decis Mak. 2021;21:314.

71. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge University Press. Cambridge. 2011.

72. 72. Hartasanchez SA, Heen AF, Kunneman M, Garcia-Bautista A, Hargraves IG, Prokop LJ, et al. Remote shared decision making through telemedicine: A systematic review of the literature. Patient Educ Couns. 2022;105:356-65.

73. Tates K, Antheunis ML, Kanters S, Nieboer TE, Gerritse MB. The Effect of Screen-to-Screen Versus Face-to-Face Consultation on Doctor-Patient Communication: An Experimental Study with Simulated Patients. J Med Internet Res. 2017;19:e421.

74. Probst MA, Kanzaria HK, Schoenfeld EM, Menchine MD, Breslin M, Walsh C, et al. Shared Decisionmaking in the Emergency Department: A Guiding Framework for Clinicians. Ann Emerg Med. 2017;70:688-95.

75. Fisher KA, Tan ASL, Matlock DD, Saver B, Mazor KM, Pieterse AH. Keeping the patient in the center: Common challenges in the practice of shared decision making. Patient Educ Couns. 2018;101:2195-201.

76. Ubbink DT GP, Gosens T , Brand PLP. Meer 'samen beslissen' nodig door aangescherpte Wgbo. Ned Tijdschr Geneeskd 2021;165:D5775.

77. Partners van campagne 'Begin een goed gesprek'. Available from: *https://begineengoedgesprek.nl/hallo-zorgverlener/partners/*. Date last accessed: 18-12-2021.

78. Maes-Carballo M, Gomez-Fandino Y, Reinoso-Hermida A, Estrada-Lopez CR, Martin-Diaz M, Khan KS, et al. Quality indicators for breast cancer care: A systematic review. Breast. 2021;59:221-31.

79. Scholl I, Kobrin S, Elwyn G. "All about the money?" A qualitative interview study examining organizational- and system-level characteristics that promote or hinder shared decision-making in cancer care in the United States. Implement Sci. 2020;15:81.

80. 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.

81. Kunneman M, Montori VM, Shah ND. Measurement with a wink. BMJ Qual Saf. 2017;26:849-51.

82. Durand MA, Barr PJ, Walsh T, Elwyn G. Incentivizing shared decision making in the USA--where are we now? Healthc (Amst). 2015;3:97-101.