Knudsen *et al. Research Involvement and Engagement* (2026) 12:90
https://doi.org/10.1186/s40900-026-00922-4

Research Involvement
and Engagement

RESEARCH **Open Access**

Involving patient partners as appraisers of shared decision making: an explanatory sequential mixed-methods study

Bettina Mølri Knudsen[1,2*], Mette Byg Josefsen[1], Mona Muusmann Petersen[3], Poul Risom[4], Henriette Witte Nielsen[2,5], Frede Donskov[2,6], Stine Rauff Søndergaard[7], Karina Dahl Steffensen[2,8], Dawn Stacey[2,9,10] and Lea Lund[1,2]

Abstract

Background Patient and public involvement (PPI) is increasingly recognised as a valuable component of health research, yet patients are rarely included in data analysis within shared decision making (SDM) studies. We aimed to assess the feasibility and value of involving patient partner appraisers of SDM using the OPTION^5 instrument. Specific objectives were to examine interrater reliability between patient partners and researchers, and to explore patient partners' perspectives on observed differences in interrater reliability.

Methods An explanatory sequential mixed-methods design was used. Four patient partners and three researchers used the validated Danish version of OPTION^5 to independently score 102 audio-recorded oncology consultations from two randomised trials evaluating patient decision aids delivered before versus during consultations. Each consultation was scored by one trained patient partner and two researchers. Scores were converted to a 0–100 scale. Agreement was assessed using two-way mixed-effects, absolute-agreement intraclass correlation coefficients (ICC) for single and average measures. Item-level analyses included unadjusted and Bonferroni-adjusted p-values. Three patient partners participated in a semi-structured focus group interview to discuss differences in interrater reliability and to reflect on their assessments based on lived experience. Interview data were analysed using a primarily deductive thematic approach informed by hermeneutic–phenomenological principles.

Results The four patient partners appraised 25–27 consultations each. Total OPTION^5 scores were median 60.0 for patient partners and 52.5 for researchers. Overall single-rater agreement was moderate (ICC = 0.60, 95% CI 0.46–0.71), and good when averaging raters (ICC = 0.75, 95% CI 0.63–0.83). Item-level differences were small. Although item 2 showed a nominal difference before correction for multiple testing, no item-level differences remained statistically significant after Bonferroni correction. Qualitative analysis revealed three overarching themes: emotion-driven scoring (atmosphere, tone, empathy and recognition), navigating the OPTION^5 instrument (growing confidence and calibration over time), and evolving roles as co-researchers.

---

*Correspondence:
Bettina Mølri Knudsen
Bettina.Knudsen@rsyd.dk

Full list of author information is available at the end of the article

Conclusions Involving patient partners as appraisers of SDM using OPTION^5 was feasible, and their appraisal ratings showed moderate reliability with researcher ratings. Findings suggest that, while using the same structured OPTION^5 scoring framework, patient partners may draw on broader relational cues when interpreting observable SDM behaviours in clinical consultations.

Trial registration Ethical approval was obtained from the Research Ethics Committee at the University of Southern Denmark (approval ID 23/555) for the involvement of patient partners as co-researchers.

Plain English summary

We wanted to find out whether patients can help assess how well doctors involve patients in decisions about their care during cancer consultations. In this study, four patient partners and three researchers listened to recordings of 102 cancer consultations from two studies of patient decision aids, which are tools designed to help patients take part in decisions. The patient partners and researchers each rated how well the cancer patients in the consultations were involved. Overall, the patient partners' ratings were fairly close to the researchers' ratings. This suggests that it is possible for patients to take part in this kind of research assessment. When talking to the patient partners about their experience with the assessment task, they explained that they often noticed things that went beyond the checklist of shared decision making used in the study. For example, they paid close attention to the atmosphere of the conversation, the doctor's tone of voice, empathy, and whether the patient seemed recognised and understood. They also said that they became more confident in using the rating tool over time. These findings show that involving patient partners in analysing research data is both possible and valuable. Patients may notice important parts of shared decision making that researchers or doctors do not always capture, especially the human and emotional side of the consultation. This suggests that patient involvement in research can add a broader and more real-world understanding of what good shared decision making looks like in clinical practice.

Keywords Patient partner, PPI, Patient voice, Shared decision making, SDM, Oncology consultations, OPTION^5

Background

Shared decision making (SDM) has gained significant momentum in cancer care, where treatment decisions often involve complex options and difficult trade-offs. SDM is an approach where clinicians and patients collaborate to select treatments based on clinical evidence as well as patients' informed preferences [1]. However, SDM is a behaviour that is challenging to measure because different measurement approaches get different results and capture different perspectives, including patient-reported experiences, clinician-reported behaviour, and observer-rated patient-clinician communication. Patient decision aids (PDA) are often used to support SDM by helping structure clinical counselling and facilitating patient involvement in decisions [2].

Patient and public involvement (PPI) in health research is defined as an active partnership in which people with lived experiences take on agreed roles throughout a study. In this approach, research is conducted *with* or *by* patients and members of the public rather than *to*, *about* or *for* them [3]. There are clear ethical, societal and scientific reasons for PPI. A review of frameworks highlights three key arguments: (1) patients have a right to contribute to research on their condition, (2) their lived experiences can enhance the quality, efficiency, and relevance of research, and (3) PPI increases accountability and transparency [4].

Over recent decades, PPI has moved from an aspiration to an expectation in many settings. Across Europe, Canada and the US, enthusiasm for PPI is strong and well-established in cancer care [5]. From a practical perspective, PPI has shown particular value in clinical trials, for example in shaping trial design, selecting outcomes and improving recruitment [6, 7]. However, despite this momentum, practice remains inconsistent. The depth and stage of involvement vary widely [5], and patients are still rarely included in data analyses phases of research [8]. Little is known about the involvement of patient partners in measuring SDM behaviour. In SDM measurement specifically, barriers to involving patient partners may mirror those reported in broader PPI and collaborative data analysis literature, including uncertainty about appropriate methods and levels of involvement, the need for adequate preparation and training, and the challenge of integrating patient perspectives systematically into analysis [9].

To contribute to the evidence on how PPI can be implemented, this study examines the involvement of patient partners in data analysis. The study is based on two randomised controlled trials with breast and colorectal cancer patients, investigating the use of pre-consult digital versus in-consult paper-based PDAs during consultations with clinicians [10]. Outcomes in the two trials were measured from the perspectives of patients, clinicians, and observers [10]. Observer-rated instruments, such as the OPTION^5 instrument (acronym for "observing patient involvement" [11]), are one approach to measuring observable SDM behaviours in consultations, including identifying the need for a decision, supporting deliberation, providing information about options, eliciting patient preferences, and integrating preferences into the decision [12]. In this study, patient partners participated as members of the research team and used the OPTION^5 instrument to evaluate the quality of real-life oncology consultations. The overall aim was to assess the feasibility and value of involving patient partners as appraisers by examining the level of agreement and potential differences in OPTION^5 scores between patient partners and researchers, and to qualitatively explore patient partners' perspectives on these differences and how they interpreted their roles as co-researchers.

Methods

Study design

We used an explanatory mixed-methods design [13], in which quantitative and qualitative methods were combined to address different parts of the research question. First, the quantitative part examined the level of agreement between patient partners and researchers in appraising the quality of SDM in consultations using OPTION^5. This was followed by a qualitative exploration of patient partners' experiences to help explain how they made their assessments and how these perspectives may have contributed to similarities and differences in ratings. The study involved four steps: establishing the team of researchers and patient partners, training in appraising quality of consultations using the OPTION^5, analysing audiotaped consultations, and exploring patient partners´ experiences (Fig. 1). The study design and reporting followed the GRIPP2 (Guidelines for Reporting of Patient and Public Involvement in Research) framework [14]. The GRIPP2 checklist is provided in Additional file 1.

Step 1: Establishing the research team

The research team consisted of three researchers (two female, one male) and four patient partners (one female, three male), all of whom were cancer patients. The patient partners were recruited through the Patient and Relatives' Council at two hospitals in the Region of Southern Denmark (Lillebaelt Hospital and Esbjerg Hospital). The principal investigator (PI) was invited to attend council meetings at both sites with the aim of recruiting interested patient partners. The patient partners were not selected for previous research experience, but for their lived experience as cancer patients and their interest in contributing to the study. As members of patient and relatives' councils, they had experience contributing patient perspectives in healthcare development or advisory contexts. The first four volunteers were included (two from each site). The PI had previous experience with SDM research, PDA evaluation, and use of the OPTION instrument, including previous work on translation and cultural adaptation of the Danish version [15].

In addition to the PI, the research team included two researchers affiliated with the Department of Oncology at Esbjerg Hospital, both of whom were healthcare professionals with experience in consultations with cancer patients. Apart from the PI, neither the two additional researchers nor the four patient partners had prior experience using OPTION^5 or similar observer-based rating instruments, although all had a basic understanding of SDM. Patient partners were not financially compensated for their participation, but travel expenses were reimbursed by the PI.

Step 2: Training in appraising quality of consultations using OPTION^5

Prior to formal scoring, the rating team participated in a training session led by the PI to ensure consistency in applying the five OPTION^5 items. The session included an introduction to the purpose of OPTION^5, a review of the five items and scoring categories, and discussion of how SDM behaviours could be identified in audio-recorded consultations. Training used three non-oncology consultations (patients with gallstones or hernia) to introduce the patient partners to the scoring process without the emotional strain specific to cancer diagnoses. After each training consultation, all raters first completed their scores independently, followed by a group discussion of differences in ratings and interpretation of the scoring manual. The discussion focused on clarifying how the OPTION^5 criteria should be applied, rather than on reaching consensus about future scoring. The session aimed to achieve initial calibration and reduce interrater variation. Early differences were noted as patient partners tended to rate consultations more positively, and researchers more critically. After three training cases,

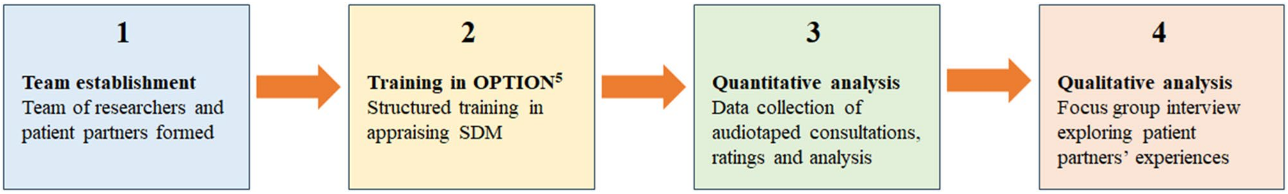

Fig. 1 The four steps comprising the explanatory mixed-methods design

alignment was achieved. Training was conducted in a single session in June 2024.

To address potential emotional strain, the four patient partners and PI met with a psychologist from the oncology department both before training and midway through scoring to reflect on their experiences and emotional wellbeing.

Step 3: Analysing audiotaped consultations

Following training, all 102 audio-recorded oncology consultations (32 with colorectal cancer patients and 70 with breast cancer patients) were independently scored by one patient partner, the PI and one of the two researchers using the validated Danish version of OPTION^5 [15]. OPTION^5 is a five-item scale [16] measuring patient involvement in clinical decision making across five different items, each rated on a 5-point Likert scale ranging from 0 (zero effort observed) to 4 (exemplary effort observed).

To ensure data security and provide emotional support, patient partners conducted their scoring sessions together with the PI either in person (in a meeting room at the hospital) or online, with the audio played from the PI's computer. This setup prevented direct access to recordings and allowed for debriefing after each consultation. During these debriefings, patient partners could discuss emotional reactions to the recordings and raise concerns, for example related to feelings of inadequacy in the research process, exposure to sensitive topics, willingness to be honest in their assessments, or difficulties separating the consultation content from their own experiences as patients. These reflections were documented in an impact log [17]. The scoring itself was completed independently by the patient partners and was not discussed with the PI during or after the scoring sessions. The two researchers had full access to the recordings and independently scored half of the consultations within a defined timeframe. Consequently, all consultations were independently scored by one patient partner, the PI and one additional researcher, ensuring that each recording received three separate OPTION^5 assessments. For the quantitative comparisons, the PI and additional researcher ratings were combined into a single reference value by calculating their mean (hereafter referred to as the researcher-mean score). This mean provides a single summary estimate of the researchers' assessment for each consultation and allows direct comparison with the patient partner rating. Because the analysis involved paired observations (patient partner score vs. researcher-mean score for the same consultation) and normality of score, patient partner and researcher-mean scores were compared using two-sided Wilcoxon signed-rank tests. All analyses were performed in Stata 19.5. Audio recordings were securely stored on a dedicated server at OPEN (Open Patient data Explorative Network, Odense University Hospital, Region of Southern Denmark).

All OPTION^5 ratings were completed using printed paper scoring forms. The PI subsequently entered all item-level ratings into an Excel dataset. To ensure data accuracy, a research assistant independently verified the entered data against the original paper scoring forms before statistical analysis. The statistical analyses were conducted by the PI and subsequently validated by a researcher with statistical expertise who was not involved in the OPTION^5 scoring.

OPTION^5 ratings were converted to a 0–100 scale (sum of five items, each scored 0–4, rescaled). Total OPTION^5 scores for patient partners and researchers are presented descriptively as medians. Paired differences are presented as median paired differences. Paired patient partner and researcher-mean scores were compared using two-sided Wilcoxon signed-rank tests. Item-level comparisons were considered exploratory; for these five comparisons, both unadjusted exact p-values and Bonferroni-adjusted p-values were calculated, with interpretation based on the adjusted p-values.

Agreement between patient partner and researcher-mean total scores was analysed using two-way mixed-effects, absolute-agreement intraclass correlation coefficients (ICC) for single and average measures, with 95% confidence intervals. We report both single-measure reliability (ICC(A,1)) and average-measure reliability (ICC(A,2)). Single-measure reliability (ICC(A,1)) reflects the expected reliability of a single rating from an individual rater. Average-measure reliability (ICC(A,2)) reflects the reliability of the mean of the two researcher ratings compared with the patient partner rating for each consultation. Following established guidelines, ICC values < 0.50 were interpreted as poor reliability, values between 0.50 and 0.75 as moderate reliability, values between 0.75 and 0.90 as good reliability, and values > 0.90 as excellent reliability [18]. Score distributions were visualised using a boxplot.

Step 4: Exploring patient partners' experiences appraising consultations

The four patient partners were invited to participate in a focus group interview using a semi-structured interview guide (Additional file 2). Open-ended questions addressed the patient partners' first-hand reflections on the observed scoring patterns, including differences in OPTION^5 ratings and level of agreement between themselves and the researchers, what they considered meaningful when assigning OPTION scores, and how they perceived their role as co-researchers in the project. The session lasted 120 min. At the start of the session, patient partners were presented with an overview of the OPTION^5 scoring calculation and ICC calculations, as

Table 1 Characteristics of analysed consultations by raters

Raters	Consultations rated			Duration of audio-recordings (minutes)	Participated in focus group
	Total (*n*)	Breast (*n*)	Colorectal (*n*)		
Patient partner 1	25	17	8	1,030	Yes
Patient partner 2	27	20	7	1,153	Yes
Patient partner 3	25	19	6	939	Yes
Patient partner 4	25	14	11	987	No
PI	102	70	32	4,109	n/a
Researcher 2	51	19	32	2,077	n/a
Researcher 3	51	51	0	2,032	n/a

n/a = not applicable

well as the boxplot illustrating group-level medians and the corresponding ICC values. This was followed by a focus group discussion facilitated by the PI, who ensured that all three patient partners were actively involved in the conversation. The interview was audio-recorded and subsequently transcribed verbatim for analysis.

To analyse the focus group transcript, we used thematic analysis [19] informed by a hermeneutic–phenomenological approach [20]. The analysis was iterative and interpretive, involving repeated cycles of reading, reflection, and re-reading of the transcript to engage closely with the patient partners' accounts. The initial analysis was guided by deductive themes related to the study aims and the interview guide, while remaining open to new meanings emerging from the data. The focus group transcript was therefore coded inductively to capture nuances in patient partners' experiences. Initial coding was conducted by the PI and subsequently reviewed by the last author. To enhance methodological integrity and reflexivity, preliminary codes and themes were discussed between the PI and the last author, with attention to how the PI's dual role as trainer, facilitator, and analyst could influence interpretation.

All analyses were performed using NVivo software (release 14.23.2, QSR International Pty Ltd, Victoria, Australia).

Results

A total of 102 audio-recorded oncology consultations were scored between July and December 2024 (see Table 1). Of the four patient partners, three participated in the focus group as one patient partner passed away after completing the scoring period and before the focus group took place.

The OPTION^5 instrument was median 60.0 out of 100 for patient partners and 52.5 for researchers ($p = 0.072$). Patient partners scored higher in 50 consultations, lower in 45, and the same for 7. The mean paired difference was +4.8 points (patient partners higher) with higher central tendency and greater variability among patient partners (Fig. 2).

Patient partners researcher-mean

Overall agreement between patient partners and the researcher-mean across all consultations was moderate for single-rater agreement (ICC(A,1) = ICC 0.60, 95% CI 0.46–0.71) and good for the average of both raters

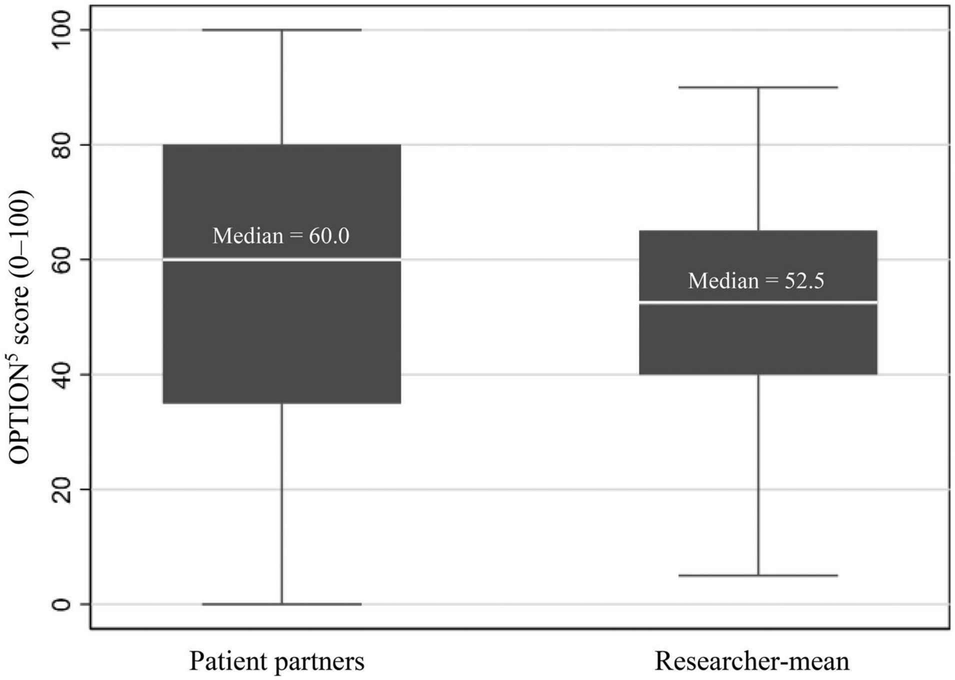

Fig. 2 Distribution of OPTION^5 total scores (0–100, rescaled) for patient partners and researcher-mean. Median values are shown for each group

Table 2 Item-level differences between patient partners and researcher-mean on OPTION^5 (0–4)

OPTION^5 item	**Mean difference (pat–res)**	***p*-value**	**Bonferroni-adjusted *p*-value**
Item 1 For the health issue being discussed, the clinician draws attention to or confirms that alternate treatment or management options exist or that the need for a decision exists. If the patient rather than the clinician draws attention to the availability of options, the clinician responds by agreeing that the options need deliberation.	+0.08	0.697	1.000
Item 2 The clinician reassures the patient or re-affirms that the clinician will support the patient to become informed or deliberate about the options. If the patient states that they have sought or obtained information prior to the encounter, the clinician supports such a deliberation process.	+0.27	0.034	0.170
Item 3 The clinician gives information or checks understanding about the options that are considered reasonable (this can include taking no action), to support the patient in comparing alternatives. If the patient requests clarification, the clinician supports the process.	+0.19	0.068	0.340
Item 4 The clinician makes an effort to elicit the patient's preferences in response to the options that have been described. If the patient declares their preference(s), the clinician is supportive.	+0.20	0.146	0.730
Item 5 The clinician makes an effort to integrate the patient's elicited preferences as decisions are made. If the patient indicates how best to integrate their preferences as decisions are made, the clinician makes an effort to do so.	+0.22	0.113	0.565

Mean difference (pat–res) indicates patient partner score minus researcher-mean score (0–4 scale). Bonferroni-adjusted p-values were calculated for the five item-level comparisons and for inferential interpretation. Adjusted p-values above 1.000 were truncated at 1.000

Table 3 Agreement between patient partner and researchers on OPTION^5 total scores (0–100)

Cancer type	***n* (pairs)**	**Patient partner median**	**Researcher median**	**Paired difference, median**	**ICC (A,1)**	**ICC (A,2)**
Overall	102	60.00	52.50	+4.8	0.60	0.75
Colorectal (first dataset)	32	85.00	58.75	+13.75	0.43	0.60
Breast (second dataset)	70	42.50	47.50	−2.50	0.56	0.72

ICC(A,1) reflects single-measure agreement between patient partner and researcher-mean scores, while ICC(A,2) reflects the reliability of the average of these two ratings

(ICC(A,2) = 0.75, 95% CI 0.63–0.83). Although the point estimates suggest moderate and good reliability, respectively, interpretation based on the lower bounds of the confidence intervals indicates poor-to-moderate reliability for single ratings and moderate reliability for average ratings [18]. In the OPTION^5 per-item analyses, patient partners scored higher than the researcher-mean across all five items. Item 2 showed the largest difference, with patient partners scoring + 0.27 points higher than researchers on the 0–4 item scale. This difference was nominally significant before correction for multiple testing but did not remain statistically significant after Bonferroni correction. No item-level differences were statistically significant after correction for multiple testing (Table 2).

In consultations involving patients with colorectal cancer, patients partner scored substantially higher than researchers (medians 85.0 vs. 58.8; Wilcoxon $p < 0.001$). Agreement between raters was ICC(A,1) = 0.43 (95% CI 0.00–0.71). Although the point estimate suggests moderate reliability, the lower confidence bound indicates poor reliability [18]. In the subsequent breast cancer consultations, researchers scored higher than patient partners, although the difference was small (median 42.5 vs. 47.5; Wilcoxon $p = 0.63$). Agreement was ICC(A,1) = 0.56 (95% CI 0.37–0.70), corresponding to moderate reliability based on the point estimate but poor-to-moderate reliability when interpreted according to the lower confidence bound (Table 3) [18].

Results from focus group interview

The focus group was performed September 2025 (see Table 1 for participants). During the analysis, three overarching themes were identified: (A) emotion-driven scoring system, (B) navigating the OPTION^5 instrument in practice, and (C) learning and evolving roles as co-researchers (see Fig. 3).

Theme A: Emotion-driven scoring system

Patient partners described how their scoring decisions were shaped by emotional responses to the consultations. They emphasised that their assessments of SDM were influenced not only by observable behaviours but also by how the interaction felt from a patient perspective. They

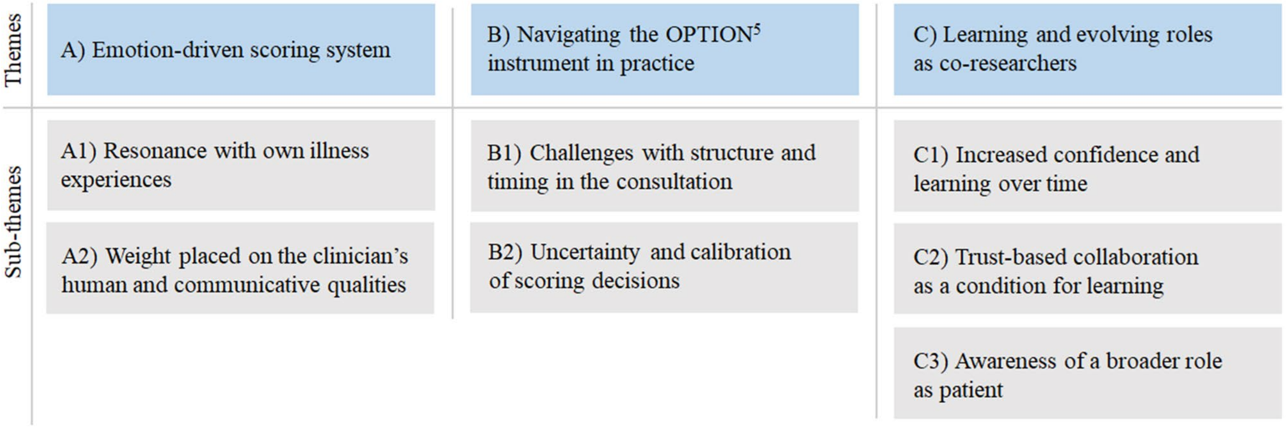

Fig. 3 Patient partners' experiences appraising consultations, and roles as co-researchers

ascribe this to having a different starting point: "*…the researchers, they might look for the holes in the cheese. We look for the cheese itself when we score… because we have two different starting points*" (Patient partner 3).

Subtheme A1: Resonance with own illness experiences

Patient partners explained that their assessments of SDM were strongly shaped by emotional responses rooted in their own patient stories. They described how recognition, identification, and memories of past encounters with clinicians influenced what they paid attention to and how they interpreted the recorded consultations: "*It's the atmosphere. How you're received. I'm extremely sensitive to that. I've always thought, when I sat waiting for results, and the healthcare professional who fetched me from the waiting room: how do the person look at me, does the person smile? You're hyper-aware. So I also think about whether they make sure the patient feels okay when they start the consultation. I think that's what I listen for first*" (Patient partner 2).

This emotional resonance meant that scoring SDM was intertwined with personal experiences of vulnerability, uncertainty, or relief, making the assessment process both reflective and affective: "*I think, we draw on some experiences we've had in the healthcare system. And when I hear a conversation where I can sense empathy, understanding, and openness from the clinician conducting the conversation with the patient, then I think I score it more positively*" (Patient partner 1). When a consultation evoked familiar situations or feelings, patient partners were often inclined to score it more positively, particularly when they sensed improvement compared with previous negative experiences in healthcare: "*We know very well how poor a consultation can be. Here, we see that it's going the right way. So we might get a bit more cheerful and crank the score up a bit*" (Patient partner 3).

Subtheme A2: Weight placed on the clinician's human and communicative qualities

Patient partners placed considerable emphasis on the clinician's relational behaviour when evaluating SDM. Empathy, warmth, authenticity, and the ability to meet the patient at eye level were highlighted as central cues that shaped their scoring. They also viewed the clinician's ability to adapt language, tone, and explanations to the patient's context as an important indicator of good SDM: "*So for me it's the atmosphere that stands out, and whether the doctor is able to choose words in a way that the person sitting opposite can understand what is being said… and understand the message. That's one of the first things I notice… Are they on the same level? Can the doctor get down to the level where the patient is…?*" (Patient partner 3). These human and communicative competencies acted as emotional anchors in their decision making: if the clinician created a trusting atmosphere, chose accessible words, and demonstrated genuine attentiveness, patient partners were more inclined to perceive the consultation as embodying SDM, even when the formal behaviours measured by OPTION^5 were only partially present.

Theme B: Navigating the OPTION^5 instrument in practice

Patient partners explained that using OPTION^5 required constant movement between listening to the consultation and interpreting the structure of the tool. They described this balancing act as demanding but increasingly manageable as they became more familiar with the instrument: "*It took a little while before I realised that this wasn't the script running in full time. I initially felt this was the script the doctor also followed, and it took a few rounds to realise: no, they jump back and forth*" (Patient partner 3).

Subtheme B1: Challenges with structure and timing in the conversation

Patient partners found it challenging to align the structure of the OPTION^5 instrument with the dynamic flow of real-life consultations. They experienced the conversations as non-linear, with frequent shifts between decision making, information sharing, and broader biographical or emotional topics, whereas the instrument followed a fixed item sequence: "*At least at first it was hard, but honestly also later, because the conversations went in all sorts of directions. And much of it wasn't directly about shared decision making. Other things came in, and luckily so - like the patient's life and how she was doing. And then you have to keep refocusing. Is it shared decision making when the doctor asks about these things, or is it just a doctor–patient conversation?*" (Patient partner 1). This mismatch made it difficult to maintain an overview of where they were in the scoring form and to decide which parts of the conversation should be linked to which item. They described repeatedly moving their attention between the recording, the headings, and the response options: "*I also think it was difficult, because I couldn't remember the headings, and I think especially at the beginning, I got carried away listening to the conversation instead of listening for SDM. I had to keep going back to the heading, and I didn't find that easy*" (Patient partner 2). This constant switching between following the interaction and orienting themselves in the tool was experienced as cognitively taxing, and some found the layout dense and demanding.

Subtheme B2: Uncertainty and calibration of scoring decisions

Patient partners reported substantial uncertainty about how to choose between neighbouring score levels, particularly early in the project. They described hesitating, going back to revise earlier scores, and sometimes feeling unsure about what should "count" as evidence for a higher or lower rating; "*I would say that I have sometimes sat there hesitating, thinking, is this a* [score of] *3 or a 4? Because it is definitely not a 2. But is it a 3 or a 4? And then I might think, okay, if I'm weighing it like that, it must be because I think it's good, so I believe that in this case it's a 4*" (Patient partner 1). Over time, repeated use of the instrument helped them develop a sense of what different score levels represented, which reduced but did not eliminate doubt. All patient partners described becoming "sharper" and more confident in their decisions as they gained familiarity with the logic of OPTION^5, while still recognising that their scoring remained influenced by comparison across consultations and by their own expectations of what good SDM should look like. Patient partner 3 said: *"Whether routine gives something? We should preferably get sharper and sharper"*, which aligned with patient partner 2's perception: *"Yes, we really should. I truly hope we improved along the way. I hope we learned".*

Theme C: Learning and evolving roles as co-researchers

Patient partners highlighted a gradual sense of growth throughout their involvement in the project. They described becoming more confident, more reflective, and more aware of their contribution as patient representatives in a wider research context: *"We bring ourselves and our experiences… It doesn't just come by itself, you could say. It takes some groundwork, and then it worked out"* (Patient partner 2).

Subtheme C1: Increased confidence and learning over time

Patient partners experienced their involvement in the project as an intensive learning process. They described how repeated exposure to recorded consultations and systematic work with the OPTION^5 instrument gave them a deeper understanding of SDM as a practice, beyond their prior experiences as patients. Over time, they felt more secure in their role, less anxious about "doing it wrong," and more able to articulate why they had scored in a particular way: "*At the beginning… I thought, this is really difficult. The others* [patient partners] *are probably much better than me. I'm probably a fool when it comes to figuring this out…So I got a bit nervous until I thought, no, I'll score the way I believe it should be, based on my patient perspective. And then it really doesn't matter if I mess up and the rest of you score more precisely…*" (Patient partner 1). The same patient partner also framed the process as a form of education in SDM: *"I also told my wife, when I got home, that I am actually in the middle of an education. A real education in shared decision making…"* (Patient partner 1). This growing confidence was experienced as a sense of competence and legitimacy.

Subtheme C2: Trust-based collaboration as a condition for learning

Patient partners highlighted trust, respect, and a sense of equality as key features of the collaboration, noting that they did not feel positioned as "lay people" in opposition to "experts" but as partners whose perspectives were taken seriously: *"It has been good, the way you* [the PI] *approached us. It wasn't researcher–patient or researcher–layperson. There has been eye level all the way through. And the things that could have been difficult were addressed along the way. I think the introduction to everything was excellent"* (Patient partner 3). This trust-based environment made it possible to share personal experiences, express doubt, and engage with emotionally charged material without fear of judgement: *"I have felt completely safe throughout the process. Even though emotions rise now and then, I think it has been incredibly exciting and educational"* (Patient partner 2).

Subtheme C3: Awareness of a broader role as patient representatives

Over the course of the project, patient partners became increasingly aware that their contributions extended beyond the immediate scoring task. They began to see themselves as bearers of patient knowledge who could influence research agendas and professional discussions about SDM more broadly: *"We receive so much knowledge, and we carry it outside these walls. We are all outgoing people, and this knowledge needs to be spread... The more people who go through this process, the more of us there are to represent these perspectives and talk about it everywhere"* (Patient partner 2). Several described taking insights from the project into conferences, networks, and informal conversations with other patients and healthcare professionals: *"A Danish network is holding a conference in a few months... And I'm invited... Because I took part last time too, and I said some things that I think startled them a bit... So I'm curious to see how it will go"* (Patient partner 1). This fostered a sense of empowerment: they were not only providing data but also helping to shape how SDM is understood, implemented, and discussed in oncology care and in the wider health system.

Discussion

In line with the explanatory sequential mixed-methods design, the quantitative results first identified patterns of agreement and variation between patient partner and researcher ratings, while the qualitative focus group subsequently explored patient partners' reflections on these scoring patterns and their experiences of using OPTION^5. Across the two components, the findings were connected primarily through complementarity and explanation: the quantitative results showed moderate overall agreement, early divergence in ratings, and later convergence, while the qualitative findings described how patient partners navigated the instrument, became more confident over time, and drew on relational and emotional aspects of the consultations when scoring. Taken together, the findings demonstrated the feasibility and value of involving patient partners as appraisers of SDM in oncology consultations using OPTION^5. Inter-rater reliability between patient partners and researchers reached a moderate level overall and improved when scores were averaged and over time as they gained experience. Although patient partners tended to give slightly higher scores, especially in the early colorectal consultation dataset, this pattern may reflect increasing familiarity with the instrument, differences between the clinical contexts, or both, as timing and cancer type were not separable in the present study. Themes from the qualitative findings further showed patient partners based their assessments on emotional resonance, attention to relational aspects of the consultation, and their own lived experiences as cancer patients. Together, these insights suggest that patient partners offer interpretive value that complements, rather than replaces, scoring by researchers.

The observed differences between colorectal and breast cancer consultations may have several explanations. First, the colorectal consultations were scored earlier in the process, whereas the breast cancer consultations were scored later. The qualitative findings suggest that patient partners became more confident and familiar with OPTION^5 over time, which may partly explain why ratings became more similar in the later breast cancer dataset. Second, the two datasets differed in clinical context and decision situation, which may have influenced the SDM behaviours present in the consultations. However, the purpose of the present study was not to compare SDM across cancer types or intervention contexts, but to examine agreement between patient partner and researcher ratings of the same audio-recorded consultations. Nevertheless, the subgroup findings should be interpreted cautiously and primarily as possible indications of rater learning, contextual variation, or both.

In our study, patient partners tended to rate consultations more positively than researchers, with the largest descriptive item-level difference observed for OPTION^5 item 2, capturing information exchange and support for deliberation. Similar discrepancies between patient-reported and OPTION^5 observer-rated SDM have been documented previously, with patient measures often yielding high scores while OPTION^5 ratings remain modest and only weakly correlated with patient reports [21]. This pattern is consistent with known challenges in SDM measurement, including ceiling and halo effects in patient-reported experience measures and conceptual issues in how SDM constructs are operationalised [22–24]. However, the mixed-methods design of the present study allowed us to move beyond a purely quantitative interpretation of higher scores as a ceiling effect. The qualitative findings showed that patient partners' scoring was shaped by nuanced reasoning related to empathy, communication quality, and perceptions of "being met," suggesting that higher scores were not simply lenient but grounded in patient-based evaluative criteria.

Integration of patient perspectives in observer-based scoring is rarely reported in SDM research, where researchers typically conduct ratings. Evidence from other fields shows that service-user involvement in appraising can introduce new insights, reveal interpretive blind spots, and broaden conceptualisations of quality [25]. The qualitative findings from this study suggest that patient partner perspectives may add value to the interpretation of observer-based SDM assessments, but they also highlight an important methodological tension. Patient partners' emphasis on atmosphere, tone, recognition and the relational quality of the consultation aligns with broader SDM frameworks that position relational quality as a foundational component of shared deliberation rather than a separate domain [26]. However, these relational and emotional impressions are not explicit OPTION^5 scoring criteria. OPTION^5 is designed as a structured observer-based instrument assessing observable SDM behaviours, and this structured focus is one of its methodological strengths. Patient partner involvement may therefore add value not by changing the OPTION^5 scoring criteria, but by helping researchers interpret what structured SDM measures do and do not capture. This raises an important methodological question about whether relational and value-based aspects of SDM should be assessed separately from OPTION^5. Patient partners' emphasis on empathy, recognition, safety, and being met as a person aligns with Oueslati et al.'s [27] description of SDM as underpinned by values such as shared control, a safe and supportive environment, and decisions tailored to patients. Rather than suggesting that OPTION^5 should be expanded, our findings indicate that patient partner involvement may help identify complementary dimensions of SDM that could be explored through separate qualitative appraisal or additional measures alongside observer-based instruments.

One notable finding is the misalignment between OPTION^5's structured linear process and the dynamic, non-linear nature of real consultations. Patient partners explicitly struggled with this during early scoring, but they also used the challenge to articulate what they believed mattered most from their patient perspective. Their emotionally grounded scoring led to higher ratings early on, but converged toward researcher-mean scores over time, indicating that with training and experience they can apply the instrument consistently while maintaining patient-centred interpretive insights. The study further demonstrates that involving patient partners in SDM rating teams is feasible, produces acceptable reliability, and enriches interpretation of the data. This has practical implications for future SDM research: patient partners can be meaningfully integrated into data analysis, provided adequate support and when opportunities for debriefing are offered. Importantly, while raters received structured training before data analysis, the observed improvement in reliability over time suggests that future studies may benefit from longer and more iterative training sessions. Future studies could also explore mixed rating teams in which patient partners and researchers are involved together during training, calibration, and interpretation. This aligns with previous implementation work in breast cancer care, where patients contributed to integrating the patient perspective when feedback on consultations was provided to clinical teams [28].

Beyond these practical and methodological implications, the findings also raise more fundamental questions about how SDM is conceptualised and measured. Patient partners consistently attended to relational and emotional aspects of the consultation, as well as the overall interactional climate, which are not fully captured by the OPTION^5 items. This suggests that patients may hold a broader, more holistic understanding of SDM than that reflected in clinician- or researcher-led observer instruments. As such, this study not only informs how patient partners can be involved in SDM measurement, but also contributes to the wider measurement literature by highlighting potential conceptual gaps between existing SDM instruments and patient perceptions of what constitutes meaningful SDM in clinical encounters.

One strength of our study is the contribution of empirical evidence about involving patient partners in structured observer-based SDM assessments. Its mixed-methods design enabled a richer understanding of both scoring patterns and the interpretive processes underpinning them. The use of real clinical consultations and a validated instrument further strengthened the study. Limitations include greater divergence between ratings in the early block of colorectal cancer consultations, suggesting a more intensive training effect that should be considered when designing similar studies. Also, because breast cancer consultations accounted for most of the sample, the overall agreement estimates were more strongly influenced by this dataset. At the same time, the colorectal consultations showed greater divergence between patient partner and researcher ratings, possibly reflecting both dataset-specific differences and the fact that these consultations were scored earlier in the process. Another limitation is that ratings of both patient partners and researchers were based on audio recordings rather than transcripts. Although audio recordings capture tone, pauses and other paralinguistic features relevant to SDM, the absence of transcripts may have limited the ability to review specific wording or revisit segments during scoring. One additional limitation is the small number of patient partner raters. Although this was consistent with the feasibility-oriented nature of the study, it may limit the transferability of the findings to other patient partner groups or settings. The small number of patient partners also means that individual scoring tendencies may have influenced the overall agreement estimates. Future studies should explore whether similar patterns of agreement and patient partner contributions are observed with larger and more diverse rater groups. Sex may also have influenced scoring patterns or interpretations. Given the small number of raters, we were not able to examine whether rater or researcher sex influenced OPTION^5 scoring. Finally, although patient partner scoring was completed independently and scores were not discussed with the PI, the PI's presence during scoring sessions and debriefings may have influenced patient partners' confidence, reflections, or interpretation of the task. To reduce this risk, debriefings focused on emotional reactions and practical concerns rather than on OPTION^5 scores. Further methodological work is also needed to establish best practices for training and supporting patient partners in emotionally demanding rating tasks. The loss of one patient partner before the focus group may also have reduced the diversity of perspectives represented in the qualitative findings.

Conclusion

This exploratory study suggests that involving patient partners as appraisers in SDM research using OPTION^5 was feasible and valuable. Their assessments provide complementary perspectives grounded in lived experience, relational sensitivity, and emotional insight - dimensions that are central to how patients experience SDM, but not fully captured by existing observer-based SDM instruments. With structured support and training, patient partners achieve acceptable inter-rater reliability with researchers while enhancing the interpretive depth of SDM evaluation. These findings underscore the importance of including patient voices not only in the design of research but also in the analytic processes that shape our understanding of clinical communication and SDM. Through involvement, patient partners developed growing confidence, methodological awareness, and a sense of legitimacy in their roles as co-researchers. They reported that contributing to the research expanded their understanding of SDM and let to personal and professional growth, making them meaningful contributors both inside and outside the study context.

Abbreviations

GRIPP	Guidelines for Reporting of Patient and Public Involvement in Research
ICC	Intraclass correlation coefficients
OPEN	Open patient data explorative network
PI	Principal investigator
PPI	Patient and public involvement
PDA	Patient decision aid
SDM	Shared decision making

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40900-026-00922-4.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We acknowledge OPEN, the Open Patient data Explorative Network at Odense University Hospital, Region of Southern Denmark, for hosting the qualitative study data, which were managed using NVivo software (release 14.23.2, QSR International Pty Ltd, Victoria, Australia). The authors would like to thank psychologist Marianne Andersen from the Department of Oncology for her support of patient partners, and to participating patients and healthcare professionals of the Department of Oncology of at Vejle and Hospital South West of Jutland for their contributions to the study.Two patient partners are co-authoring this paper, and another patient partner, Kai Baier, is gratefully acknowledged. A special and heartfelt thanks is extended to the fourth patient partner, who sadly passed away in February 2025. He was deeply grateful and enthusiastic about his involvement in the study. This paper is dedicated to him.

Author contributions

Conception – BK, DS, KS and LL Funding acquisition – BK and KS Methodology and supervision – DS, KS and LL Data analysis – BK, MMP, PR, HWN, FD and LL Validation of statistical analysis – MBJ Drafting of the article – BK All authors critically revised the article and approved the final manuscript.

Funding

Open access funding provided by University of Southern Denmark. The work was funded by the Region of Southern Denmark, the Cancer Foundation and the Research Fund at Lillebaelt Hospital.

Data availability

The datasets used and analysed during the current study are available from the corresponding authors on reasonable request.

Declarations

Ethics approval and consent to participate

Both breast and colorectal cancer patients and healthcare staff provided written consent to audio recording, including agreement that the recordings would be listened to and rated by both fellow patients and researchers. Data collection and processing were approved by the Region of Southern Denmark (approval ID 22/9278). Ethical approval was obtained from the Research Ethics Committee at the University of Southern Denmark (approval ID 23/555) for the involvement of patient partners as co-researchers, and all patient partners have signed a confidentiality agreement.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Center for Shared Decision Making, Lillebaelt Hospital - University Hospital of Southern Denmark, Beriderbakken 4, Vejle DK-7100, Denmark
[2]Department of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark
[3]Patient Partner, Haderslev, Denmark
[4]Patient Partner, Esbjerg, Denmark
[5]Department of Public Health, University Hospital of Southern Denmark, Esbjerg, Denmark
[6]Department of Oncology, University Hospital of Southern Denmark, Esbjerg, Denmark
[7]Department of Oncology, Lillebaelt Hospital - University Hospital of Southern Denmark, Vejle, Denmark
[8]Faculty of Health Sciences, Aalborg University, Aalborg, Denmark
[9]School of Nursing, University of Ottawa, Ottawa, Canada
[10]Centre for Implementation Research, The Ottawa Hospital Research Institute, Ottawa, Canada

Received: 29 March 2026 / Accepted: 10 June 2026

References

1. Coulter A, Collins A. Making shared decision-making a reality. London: King's Fund; 2011. p. 621.
2. Stacey D, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Reviews. 2024(1).

3. What is public involvement in research? Accessed 3 Oct 2025]; Available from: https://www.hra.nhs.uk/planning-and-improving-research/best-practice/public-involvement/
4. Greenhalgh T, et al. Frameworks for supporting patient and public involvement in research: systematic review and co-design pilot. Health Expect. 2019;22(4):785–801.
5. Colomer-Lahiguera S, et al. Patient and public involvement in cancer research: A scoping review. Cancer Med. 2023;12(14):15530–43.
6. Boote J, Baird W, Sutton A. Public involvement in the design and conduct of clinical trials: a review. Int J Interdisciplinary Social Sci. 2011;5(11).
7. Crocker JC, et al. Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. BMJ. 2018;363.
8. Sweeney A, et al. Hearing the voices of service user researchers in collaborative qualitative data analysis: the case for multiple coding. Health Expect. 2013;16(4):e89–99.
9. Stocker R, et al. Patient and public involvement in care home research: Reflections on the how and why of involving patient and public involvement partners in qualitative data analysis and interpretation. Health Expect. 2021;24(4):1349–56.
10. Knudsen BM, et al. Impact of timing and format of patient decision aids for breast cancer patients on their involvement in and preparedness for decision making-the IMPACTT randomised controlled trial protocol. BMC Cancer. 2024;24(1):336.
11. Elwyn G, et al. Shared decision making: developing the OPTION scale for measuring patient involvement. Qual Saf health care. 2003;12(2):93–9.
12. Elwyn G, Grande SW, Barr P. Observer OPTION 5 manual. Darthmouth: The Dartmouth Institute for health policy and clinical practice; 2016.
13. Creswell JW, Clark VLP. Designing and conducting mixed methods research. Sage; 2017.
14. Staniszewska S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ. 2017;358.
15. Knudsen BM, Olling K, Dahl K, Steffensen. Translation and cultural adaption of Observer OPTION5 and OPTION12 in a Danish context. Health Lit Communication Open. 2024;2(1):2436891.
16. Elwyn G, et al. A three-talk model for shared decision making: multistage consultation process. BMJ. 2017;359.
17. Kok M. Guidance document: evaluating public involvement in research. 2018.
18. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
19. Clarke V, Braun V. Thematic analysis. J Posit Psychol. 2017;12(3):297–8.
20. Giorgi A. Sketch of a psychological phenomenological method. Phenomenology and psychological research. 1985:8–22.
21. Diendéré G, et al. Observer ratings of shared decision making do not match patient reports: an observational study in 5 family medicine practices. Med Decis Making. 2021;41(1):51–9.
22. van der Krieke L, et al. 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(10):e216.
23. Doherr H, et al. Use of the 9-item Shared Decision Making Questionnaire (SDM-Q-9 and SDM-Q-Doc) in intervention studies—a systematic review. PLoS ONE. 2017;12(3):e0173904.
24. Kunneman M, et al. Reflecting on shared decision making: A reflection-quantification study. Health Expect. 2019;22(5):1165–72.
25. Brett J, et al. A systematic review of the impact of patient and public involvement on service users, researchers and communities. Patient-Patient-Centered Outcomes Res. 2014;7(4):387–95.
26. Elwyn G, et al. Shared decision making: a model for clinical practice. J Gen Intern Med. 2012;27(10):1361–7.
27. Oueslati R, et al. What value structure underlies shared decision making? A qualitative synthesis of models of shared decision making. Patient Educ Couns. 2024;124:108284.
28. van Veenendaal H, et al. Effects and working mechanisms of a multilevel implementation program for applying shared decision-making while discussing systemic treatment in breast cancer. Curr Oncol. 2022;30(1):236–49.

Publisher's note