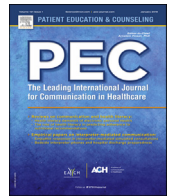




Contents lists available at ScienceDirect

Patient Education and Counseling

journal homepage: www.elsevier.com/locate/pateducou



Text-only and picture conversation aids both supported shared decision making for breast cancer surgery: Analysis from a cluster randomized trial

Renata W. Yen^a, Marie-Anne Durand^{a,b}, Camille Harris^a, Sarah Cohen^c, Abigail Ward^c, A. James O'Malley^a, Danielle Schubbe^a, Catherine H. Saunders^{a,d}, Glynn Elwyn^{a,*}

^a The Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, Lebanon, NH, USA

^b Université Toulouse III Paul Sabatier, Toulouse, France

^c Dartmouth College, Hanover, NH, USA

^d Centers for Health and Aging, Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA

ARTICLE INFO

Article history:

Received 14 February 2020

Received in revised form 30 June 2020

Accepted 18 July 2020

Keywords:

Shared decision making

Breast cancer

Encounter decision aids

Measuring shared decision making

Patient-centered care

Decision aids

Encounter conversation aids

Conversation aids

ABSTRACT

Objectives: To determine if two encounter conversation aids for early-stage breast cancer surgery increased observed and patient-reported shared decision making (SDM) compared with usual care and if observed and patient-reported SDM were associated.

Methods: Surgeons in a cluster randomized trial at four cancer centers were randomized to use an Option Grid, Picture Option Grid, or usual care. We used bivariate statistics, linear regression, and multilevel models to evaluate the influence of trial arm, patient socioeconomic status and health literacy on observed SDM (via OPTION-5) and patient-reported SDM (via **collaboRATE**).

Results: From 311 recordings, OPTION-5 scores were 73/100 for Option Grid ($n = 40$), 56.3/100 for Picture Option Grid ($n = 144$), and 41.0/100 for usual care ($n = 127$; $p < 0.0001$). Top **collaboRATE** scores were 81.6 % for Option Grid, 80.0 % for Picture Option Grid, and 56.4 % for usual care ($p < 0.001$). Top **collaboRATE** scores correlated with an 8.60 point (95 %CI 0.66, 13.7) higher OPTION-5 score ($p = 0.008$) with no correlation in the multilevel analysis. Patients of lower socioeconomic status had lower OPTION-5 scores before accounting for clustering.

Conclusions: Both conversation aids led to meaningfully higher observed and patient-reported SDM. Observed and patient-reported SDM were not strongly correlated.

Practice implications: Healthcare providers could implement these conversation aids in real-world settings.

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Women diagnosed with early-stage breast cancer often have a choice between surgery to remove the entire breast (mastectomy) or surgery to remove part of the breast (breast-conserving surgery) [1,2]. Survival is the same for both surgical options and both are typically offered [1,3,4]. However, they each have unique risks and benefits that women value differently [1]. This preference-sensitive decision is well-suited to shared decision making (SDM), where patients and clinicians use evidence-based

information and consider patients' needs and personal goals to make a treatment choice [5,6].

Encounter conversation aids are tools designed to be used during visits to engage patients and clinicians in SDM. They have been shown to have a positive impact on patient-clinician collaboration [7,8]. The Option Grid and Picture Option Grid for early-stage breast cancer surgery are two paper-based encounter conversation aids developed to help women with early-stage breast cancer and their clinicians decide between mastectomy and breast-conserving surgery plus radiation [9,10]. The Option Grid and Picture Option Grid have been demonstrated to be acceptable and feasible in a clinic setting with women of both lower and higher socioeconomic status [10]. In a mixed-methods study, they both were perceived to facilitate SDM, particularly for those of lower health literacy [10]. Option Grids have been developed over many years and tested in many clinical settings [7,11,12].

* Corresponding author at: The Dartmouth Institute for Health Policy & Clinical Practice, Dartmouth College, One Medical Center Drive, Williamson Building, Level 5, Lebanon, NH, 03756, USA.

E-mail address: glynnelwyn@gmail.com (G. Elwyn).

In the context of disseminating these tools, it is important to determine whether they facilitate SDM [13]. There is more than one way to assess whether SDM is achieved, including from the patient perspective, clinician perspective, or an observer's perspective [14]. Previous researchers have not found consistent or significant correlation between patient-reported and observed SDM. Several studies found that patient-reported and observed SDM were not significantly associated [15–17]. One found high levels of agreement [18]. To our knowledge, no study has assessed this relationship for encounter conversation aids.

SDM can also be influenced by patient and clinician characteristics. Patients that are disadvantaged due to their socioeconomic status, education, literacy, or even geographic location might benefit the most from SDM interventions [19,20]. It is therefore important to understand how interventions such as the Option Grid, Picture Option Grid and other similar tools might affect the degree of SDM achieved for disadvantaged patients.

We had three aims for this analysis: Aim 1: determine if Option Grid or Picture Option Grid encounter conversation aids increased observed SDM compared with usual care when used in surgical visits for early-stage breast cancer; Aim 2: determine if observed SDM varied by patient characteristics indicative of being disadvantaged, and; Aim 3: determine if observed SDM and patient-reported SDM overall correlate, while controlling for intervention and patient characteristics.

2. Methods

2.1. Design

We conducted a secondary analysis of data collected in the *What Matters Most* trial (ClinicalTrials.gov: NCT03136367) [21]. The parent trial was conducted at four NCI-designated cancer centers in the U.S. In the parent trial, we recruited adult women (female at birth) with a confirmed diagnosis of stage I–IIIA breast cancer. The women saw one of 16 surgeons participating in the trial. We randomized surgeons to one of three arms: (1) a text-only conversation aid (Option Grid), (2) a picture-enhanced conversation aid (Picture Option Grid), or (3) usual care. We used balanced block randomization to account for the range in number of surgeons at each site (range: 2–5). Surgeons in intervention arms (Option Grid and Picture Option Grid) received training on the use of the tool.

The Option Grid and Picture Option Grid are encounter patient conversation aids developed by several of the study authors. They are designed to support SDM when women are deciding between breast-conserving surgery plus radiation versus mastectomy for treating their breast cancer [9,10]. Both tools display Frequently Asked Questions to provide evidence-based information on the two surgical options for breast cancer. Picture Option Grid includes

simpler text and images to help support understanding among patients of lower health literacy and lower socioeconomic status [21]. Table 1 includes a comparison of the two interventions based on the Template for Intervention Description and Replication (TIDieR) checklist [22]. Both interventions are available in the supplemental file.

We screened women for inclusion via the electronic medical record and confirmed eligibility with their surgeon. After consenting, women in the trial completed a baseline questionnaire. Women then met with their surgeon to discuss breast cancer treatment options. Women that provided additional consent had their encounters audio recorded. Some women who were unaware of their diagnosis pre-encounter consented to the trial during the appointment and therefore were not able to consent to recording. Depending on their surgeon's arm assignment, women received one of the interventions or usual care during the encounter. Women completed an additional survey immediately after the encounter. More information on the trial's procedures, surgeon training, and interventions can be found in the trial protocol [21].

2.2. Outcome measures

2.2.1. Demographics

We collected patient demographics in the baseline questionnaire, including education, race/ethnicity, annual household income, household size, preferred language, and health literacy. We measured health literacy using Chew's single item measure asking patients on a 5-point Likert scale, "How confident are you filling out medical forms by yourself?" [23]. We dichotomized this measure for analysis by considering patients who selected the highest option ("Extremely") as having adequate health literacy versus all others having inadequate health literacy. This measure has been previously validated and demonstrated to have high sensitivity for detecting inadequate health literacy at this cut off [23–25]. We extracted patient age and insurance details from the electronic medical record.

2.2.2. Observed shared decision making

We assessed observed SDM using the collected audio recordings and the validated Observer OPTION-5 instrument [26]. OPTION-5 has been tested and shown to have acceptable psychometric qualities [27]. The five items in the measure ask raters to score the consultation on how much the clinician: 1) confirms that alternatives exist, 2) reassures that they will support the patient to become informed, 3) gives information or checks understanding about the options, 4) makes an effort to elicit the patient's preferences, and 5) integrates the patient's elicited preferences. Each of the five items is scored from zero to four for a summary score ranging from zero to 20 and a scaled score ranging from 0 to 100.

Table 1
Key attributes of the study interventions.*

	Option Grid	Picture Option Grid
Summary	Paper-based tool that includes nine frequently asked questions about breast cancer surgery in one table.	Paper-based tool that includes nine frequently asked questions about breast cancer with one to five images (pictures or icon arrays) per question and page dedicated to writing additional comments and questions.
Number of pages	1 page	4 pages
Number of images	0	18
Flesh-Kincaid grade level	6.6	6.5
Tested with disadvantaged women	No	Yes
Was the intervention individually tailored	No	No

* based on the Template for Intervention Description and Replication (TIDieR) checklist.

Two raters (SC and AW) received standardized OPTION-5 training which included reading the OPTION-5 manual as well as completing an interactive, online training module with the support of a trained facilitator. Both raters were external to the study team before their training and had no previous experience in assessing SDM. After completing training, the two raters independently assessed five recordings. The raters met with other study team members (including other trained raters) to discuss disagreements in item-level scoring and overall scores. After each rater had each assessed 100 recordings, the team met again to review disagreements, after which raters rescored all recordings with ratings that differed by at least two points on the raw scale. Once this calibration was complete, the raters independently assessed all remaining recordings. We used the average of the two raters scaled scores in our analyses and assessed overall agreement levels.

2.2.3. Patient-reported shared decision making

We assessed patient-reported SDM using **collaboRATE**, a 3-item validated measure [28]. **collaboRATE** has been shown to have good psychometric properties [29,30]. Patients completed **collaboRATE** immediately after their surgical consultation. **collaboRATE** has three questions: 1) how much effort was made to help you understand your health issues?, 2) how much effort was made to listen to the things that matter most to you about your health issues?, and 3) how much effort was made to include what matters most to you in choosing what to do next? Each item is rated on a scale from 0 (“No effort was made”) to 9 (“Every effort was made”). We used the top score approach for analysis, grouping patients scoring 9 on all three items versus all others. This approach maximizes the ability to detect variance and mitigate against ceiling effects [29,31].

2.3. Analysis

We used Stata 15 for these secondary data analyses [32]. Significance was assumed at a p-value of less than 0.05.

2.3.1. Determining socioeconomic status

We used patient insurance, education, and annual household income to determine SES. In order to use insurance coverage as a proxy for socioeconomic status, we dichotomized insurance status as underinsured (including Medicaid, Medicare without supplemental, or uninsured) versus adequately insured (including employer-based and Medicare with supplemental insurance). We dichotomized educational attainment as less than a bachelor's degree versus a bachelor's degree or more to distinguish between higher and lower levels of education. We dichotomized annual household income and considered patients to be in the higher income category if they were above 138 % of the federal poverty level the year they were recruited based on income and household size [33]. Patients were considered lower SES if they met at least two of the three criteria for these three dichotomized variables (underinsured, lower educational attainment, and lower income).

2.3.2. Interrater reliability of the OPTION-5 data

We determined interrater reliability of the two raters for each of the five items on OPTION-5 using a weighted Cohen's kappa [27,34].

2.3.3. Analyses corresponding to aim 1

We used descriptive and bivariate statistics to evaluate OPTION-5 scores across arms and surgeons. We used a mixed-effects multilevel linear regression model to analyze OPTION-5 scores and assess the effect of three levels of clustering in our results: 14 out of 16 trial surgeons at seven clinics across four cancer centers. Two of

the 16 surgeons didn't have analyzable recordings due to low patient volume or patients consenting during the encounter. We used a second mixed-effects multilevel linear regression model to assess whether the length of time the surgeon had been in the trial and the number of times the surgeon had used the tool (“surgeon learning effects”) had an effect on the results. For this analysis, two predictors were added to the model: (1) the time elapsed from the start of recruitment to each consultation (“time”), and (2) the order patients were seen in by each surgeon, measured by the date of the consultation, which served to account for the number of times the surgeon had used the tool (“patient order”). We considered surgeons, clinics, and sites to be random effects and trial arm, time, and patient order to be fixed effects.

2.3.4. Analyses corresponding to aim 2

We used two-sample t-tests to examine the bivariate relationship between OPTION-5 scores and patient SES and health literacy. In order to assess if OPTION-5 scores varied by selected patient characteristics indicative of being disadvantaged, we added patient SES (higher versus lower) and health literacy level (adequate versus inadequate) as fixed effects into the second multilevel model used for aim 1 and assessed for any changes to effect size.

2.3.5. Analyses corresponding to aim 3

We used bivariate statistics to test the correlation between **collaboRATE** top scores and OPTION-5 scores overall, and by arm, and calculated a point-biserial correlation coefficient (r_{pb}), which was used due to the comparison of a continuous and a binary random variable. We used simple and multiple linear regression to control for patient SES and health literacy in this relationship with OPTION-5 as the dependent variable and **collaboRATE** top scores, SES, and health literacy as independent variables. We used a mixed-effects linear regression to assess this relationship while accounting for arm as well as site-, clinic-, and surgeon-level clustering.

3. Results

3.1. Enrollment and patient characteristics

In the parent trial, 622/812 women consented to participate and 440 consented to having their consultations recorded. We analyzed 311 usable recordings across 14 surgeons. We excluded data from patients who became ineligible for the trial due to not having a choice between both surgical options or if other issues prevented us from collecting the recording. In the included sample, 40 were in the Option Grid arm, 144 in the Picture Option Grid arm, and 127 in the usual care arm. See Fig. 1 for the flow of patients in the study and reasons for exclusion from analysis.

The majority of the patients in the analyzed sample ($n = 311$) were non-Hispanic White (73.6 %), English-speaking (91.0 %), higher SES (66.6 %), and had private insurance (71.7 %). Only about half (54 %) had adequate health literacy and 35.7 % had a bachelor's degree. There were statistically significant differences in race/ethnicity, preferred language, SES, and insurance type when we looked at the sample of women who consented to recording ($n = 440$) against those who did not consent to recording ($n = 178$). Women who consented to recording were more likely to be non-Hispanic White, speak English as their primary language, have private insurance, and be of higher SES. Table 2 includes patient characteristics as well as tests for differences between women who did and did not consent to recording.

3.2. Interrater reliability on OPTION-5

The per item weighted kappa on OPTION-5 indicated either fair (items 1 and 2) or moderate agreement (items 3–5). The overall

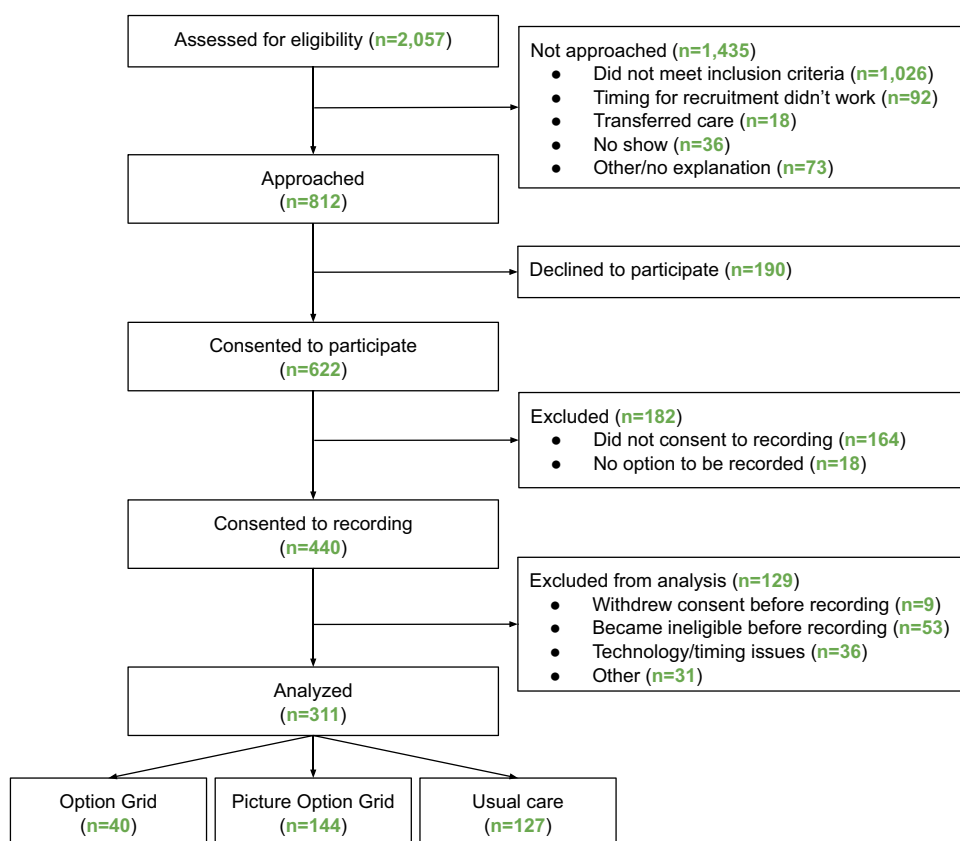


Fig. 1. Patient flow and recording consent.

weighted kappa indicated moderate agreement. See Table 3 for the overall and item-level weighted kappa values.

3.3. Observed SDM (Aim 1)

The mean OPTION-5 score was 73.0 (SD = 14.2) in the Option Grid arm, 56.3 (SD = 21.9) in the Picture Option Grid arm, and 41.0 in the usual care arm (SD = 27.5). In an ANOVA, the difference between these scores was statistically significant ($F = 32.0$, $p < 0.0001$). In a two-sample t -test the difference between Option Grid and Picture Option Grid was also statistically significant ($t = 4.6$, $p < .0001$). Mean surgeon scores ranged from 12.8 (SD = 15.7) to 79.1 (SD = 12.8). See Fig. 2 for mean surgeon scores stratified by arm. There were observed outliers in the usual care arm since surgeons who were well practiced in SDM prior to the trial were not asked to change their approach. There was also a surgeon in the Picture Option Grid arm who was an outlier which could have been due to not using the tool as intended or at all in some encounters.

In the mixed-effects multilevel regression model adjusted for surgeon, clinic, and site clusters, using Option Grid was associated with a 32.6 (95 % CI 14.8, 50.4) point increase in OPTION-5 score when compared to usual care ($p < 0.001$). Using Picture Option Grid was associated with a 23.7 (95 % CI 7.4, 40.0) point increase in OPTION-5 score ($p = 0.004$) when compared to usual care.

In the mixed-effects multilevel model adjusted for clustering and surgeon learning effects (time and patient order), OPTION-5 scores did not change significantly. Using Option Grid was associated with a 32.1 (95 % CI 14.2, 50.1) point increase in OPTION-5 score when compared to usual care ($p < 0.001$). Using Picture Option Grid was associated with a 24.0 (95 % CI 7.5, 40.5) point increase in OPTION-5 score ($p = 0.004$). See Fig. 3 for the crude and adjusted OPTION-5 scores.

3.4. Adjusting observed SDM for patient characteristics (Aim 2)

Crude mean OPTION-5 scores were 43.1 for patients of lower SES and 56.8 for patients of higher SES ($t = 4.6$, $p < 0.0001$). They were 50.9 for patients of inadequate health literacy and 52.9 for patients of adequate health literacy ($t = 0.67$, $p = 0.50$). There were significant differences when stratifying by arm for SES but not for health literacy (Supplemental Table 1). Adding patient SES and health literacy into the multilevel regression model adjusting for clustering and surgeon learning effects, did not meaningfully change the estimated effect size. Compared to usual care, using the Option Grid was associated with a 32.2 (95 % CI 15.0, 49.5) point improvement in OPTION-5 score and Picture Option Grid was associated with a 23.9 (95 % CI 8.1, 39.6) point improvement in OPTION-5 score, both compared to usual care. See Fig. 3 for the OPTION-5 scores adjusted for patient characteristics.

3.5. Comparing observed and patient-reported SDM (Aim 3)

The percent of patients with a top score on **collaborATE** was highest in the Option Grid arm (81.6 %), followed by Picture Option Grid (80.0 %), then usual care (56.4 %) ($X^2 = 19.6$, $p < 0.001$). The difference in scores between the two intervention arms was not statistically significant ($p > 0.05$). There were some differences by patient SES and health literacy: more patients of lower SES had top scores in the usual care arm and more patients with adequate health literacy had top scores in the Picture Option Grid arm (see Supplemental Table 1). Sixteen patients were missing **collaborATE** scores due to not completing some or all of the **collaborATE** measure or withdrawing from the study before **collaborATE** was completed.

Table 2

Participant characteristics by recording consent.*

	Analyzed (n = 311)	Consented (n = 440)	Did not consent (n = 178)	p-value*
Age, mean (SD)	60.5 (12.2)	60.0 (12.1)	59.0 (13.4)	0.34
Race/ethnicity, n (%)				<0.001
Black, non-Hispanic	35 (11.3)	58 (13.2)	38 (21.3)	
Hispanic	30 (9.6)	38 (8.6)	40 (22.5)	
Asian	4 (1.3)	5 (1.1)	14 (7.9)	
White, non-Hispanic	229 (73.6)	315 (71.6)	78 (43.8)	
Other	7 (2.3)	11 (2.5)	2 (1.1)	
Prefer not to say/missing	6 (1.9)	13 (3.0)	6 (3.4)	
Preferred language, n (%)				<0.001
English	283 (91.0)	396 (90.0)	138 (77.5)	
Spanish	24 (7.7)	33 (7.5)	28 (15.7)	
Mandarin	2 (0.6)	2 (0.5)	3 (1.7)	
Other	2 (0.6)	2 (0.5)	6 (3.4)	
Prefer not to say/missing	0 (0.0)	7 (1.6)	3 (1.7)	
Education, n (%)				0.39
Less than a college degree	201 (64.6)	258 (58.6)	110 (61.8)	
4-year college degree or higher	110 (35.4)	176 (40.0)	64 (36.0)	
Prefer not to say/missing	0 (0.0)	6 (1.4)	4 (2.2)	
Insurance, n (%)				0.007
Public or uninsured	91 (29.3)	125 (28.4)	70 (39.3)	
Private	220 (70.7)	315 (71.6)	107 (60.1)	
Missing	0 (0.0)	0 (0.0)	1 (0.6)	
Income[~], n (%)				0.06
Below 138 % FPL	76 (24.4)	100 (22.7)	48 (27.0)	
Above 138 % FPL	172 (55.3)	249 (56.6)	79 (44.4)	
Prefer not to say/missing	63 (20.3)	91 (20.7)	51 (28.7)	
Socioeconomic status[^], n (%)				0.02
Low	104 (33.4)	132 (30.0)	71 (39.9)	
High	207 (66.6)	308 (70.0)	107 (60.1)	
Health literacy[†], n (%)				0.39
Inadequate	141 (45.3)	192 (43.6)	85 (47.8)	
Adequate	168 (54.0)	240 (54.5)	91 (51.1)	
Prefer not to say/missing	2 (0.6)	8 (1.8)	2 (1.1)	
Surgical choice, n (%)				0.49
Breast conserving surgery	211 (67.8)	257 (58.4)	90 (50.6)	
Mastectomy	57 (18.3)	67 (15.2)	28 (15.7)	
Missing	43 (13.8)	116 (26.4)	60 (33.7)	

* Comparing those who consented v did not consent to recording, 2-sample t-tests used for continuous variables, chi-square tests used for dichotomous and categorical variables.

[~] Income = dichotomized as above or below 138 % federal poverty level (FPL) in the year of recruitment and based on household size.

[^] Low socioeconomic status defined meeting at least two criteria of the following: inadequately insurance, lower income, and lower education.

[†] Measured using Chew's single item health literacy screener and dichotomized using top-box scoring to indicate adequate health literacy.

In the bivariate analysis, patients with a top **collaboRATE** score had a mean OPTION-5 score 7.18 (95 %CI 0.66, 13.7) points higher than those without a top score ($t = 2.17$, $p = 0.03$). The point-biserial correlation coefficient (r_{pb}) was 0.13 (95 %CI 0.01, 0.23), indicating a weak but positive correlation between **collaboRATE** and OPTION-5 scores. When accounting for patient SES and health literacy in the multiple linear regression, patients with a top **collaboRATE** score had a mean OPTION-5 score 8.6 (95 %CI 2.3, 14.9) points higher than those without a top score ($t = 2.67$, $p = 0.008$). When adjusting for arm as well as surgeon-, clinic-, and site-level clusters, top **collaboRATE** scores were not correlated with higher OPTION-5 scores. Being in the lower SES group was associated with lower

OPTION-5 scores when adjusting for patient demographics but not when adjusting for trial arm and surgeon-, clinic-, and site-level clusters. See [Table 4](#) for the crude and adjusted relationships between OPTION-5 and **collaboRATE** scores.

4. Discussion and conclusion

4.1. Discussion

Our analysis included 311 patients and 14 surgeons across seven clinics. For aim 1, we found that both the Picture Option Grid and Option Grid led to higher levels of observed SDM with Option Grid having the highest scores overall. These differences in scores remained when accounting for clustering, surgeon learning effects, and patient characteristics. For aim 2, we found that both tools also led to higher levels of patient-reported SDM. For aim 3, we found that across arms, observed and patient-reported SDM were weakly correlated when measured by OPTION-5 and **collaboRATE**, which did not change when adjusting for arm, patient SES, and patient health literacy. We also found that when not adjusting for clustering, OPTION-5 scores were lower among patients of lower SES.

The mean observed SDM scores in both intervention arms were higher than other studies that have used the measure or previous

Table 3

Weighted Cohen's kappa by OPTION-5 item.

	Agreement	Expected Agreement	Kappa	SE	Z	p-value
Overall score	85.9 %	68.6 %	0.55	0.03	16.0	<0.0001
Item 1	81.9 %	70.3 %	0.39	0.04	10.9	<0.0001
Item 2	78.0 %	63.9 %	0.39	0.04	10.4	<0.0001
Item 3	85.3 %	63.6 %	0.60	0.04	15.4	<0.0001
Item 4	83.80 %	63.10 %	0.56	0.04	14.5	<0.0001
Item 5	82.20 %	61.60 %	0.54	0.04	14.1	<0.0001

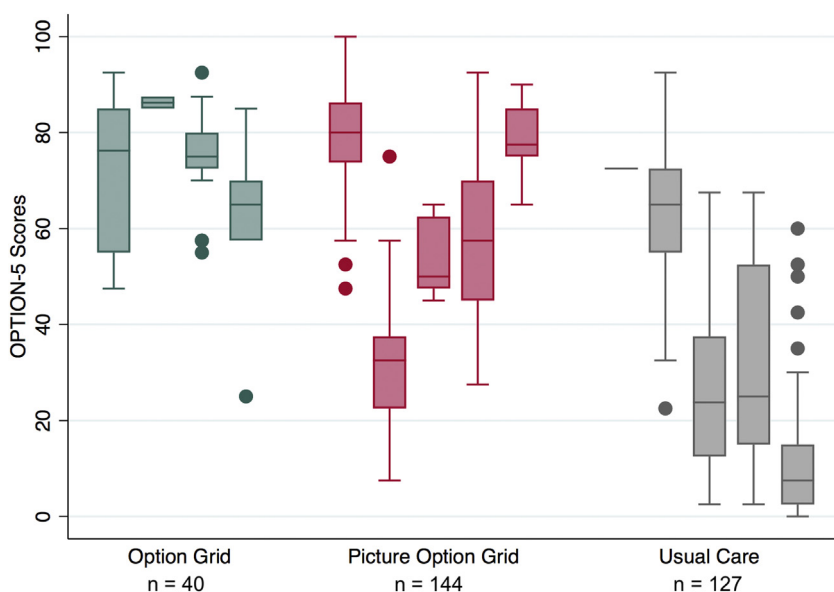
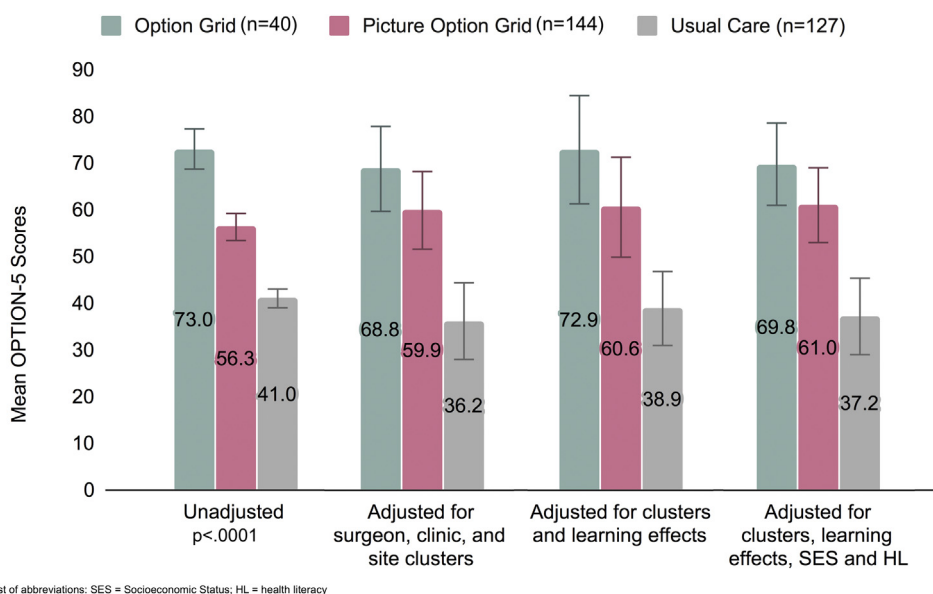


Fig. 2. Box plots for individual OPTION-5 surgeon scores, stratified by arm (surgeon n = 14, patient n = 311).



* List of abbreviations: SES = Socioeconomic Status; HL = health literacy

Fig. 3. Crude and Adjusted mean OPTION-5 scores by arm.

*List of abbreviations: SES = Socioeconomic Status; HL = health literacy.

versions of the measure [27,35,36]. The mean scores in both intervention arms were above 40 which corresponds to at least baseline ability in SDM (raw score of two or above in all five measure items) [27]. A score of 40 indicates there were meaningful SDM conversations occurring in the consultation. This indicates that both the Option Grid and Picture Option Grid interventions meaningfully improved the extent to which SDM occurred in the consultation. There were also statistical and clinically meaningful differences in observed SDM between the two intervention arms. The intervention is dependent on the surgeon using it appropriately to engage in SDM and there were noted outliers in the Picture Option Grid arm, likely leading to a lower score as compared to Option Grid. The Picture Option Grid was also designed for patients of lower socioeconomic status and only one-third of our sample was of lower SES, thus potentially limiting the use of the

intervention during the encounter and the observed effect of the intervention.

Our findings that observed and patient-reported SDM are not strongly correlated when controlling for arm assignment align with prior findings from Wunderlich and colleagues, Geessink and colleagues, and Scholl and colleagues [15–17]. Interrater agreement for observed SDM was generally higher in these three studies compared to ours. These results contrast with the findings of Pass and colleagues that showed good agreement between patient-reported and observed SDM. The assessments by Pass and colleagues used a different approach and may not be comparable [18].

Lack of agreement between observed and patient-reported SDM reflects the complexity of measuring SDM [5]. It is a reminder that while researchers and clinicians strive to implement and

Table 4

Crude and adjusted coefficients comparing collaboRATE and OPTION-5.*.

	Crude OPTION-5 score	Crude		Adjusted for SES and HL		Adjusted for SES, HL, Arm and Clustering [^]	
		Coefficient (95 %CI)	p-value	Coefficient (95 %CI)	p-value	Coefficient (95 %CI)	p-value
CollaboRATE							
Not top score	47.2	<i>referent</i>	–	–	–	<i>referent</i>	–
Top score	54.4	7.18 (0.66, 13.69)	0.031	8.60 (2.27, 14.93)	0.008	–1.33 (–5.15, 2.48)	0.49
Socioeconomic status							
Higher SES	56.8	<i>referent</i>	–	<i>referent</i>	–	<i>referent</i>	–
Lower SES	43.1	–15.32 (–21.39, –9.26)	<0.001	–15.36 (–21.57, –9.14)	<0.001	–3.68 (–7.62, 0.27)	0.07
Health literacy							
Adequate HL	52.9	<i>referent</i>	–	<i>referent</i>	–	<i>referent</i>	–
Inadequate HL	50.9	–1.62 (–7.58, 4.34)	0.59	1.43 (–4.44, 7.31)	0.63	0.082 (–3.36, 3.52)	0.96
Arm							
Usual Care	41.0	<i>referent</i>	–	–	–	<i>referent</i>	–
Option Grid	56.3	33.69 (24.9, 42.5)	<0.001	–	–	25.16 (9.49, 40.84)	0.002
Picture Option Grid	73.0	16.55 (10.60, 22.51)	<0.001	–	–	34.89 (17.80, 53.37)	<0.001

* Abbreviations: SES = socioeconomic status, HL = health literacy.

[^] Site-, clinic-, and surgeon-level clustering.

measure SDM in routine clinical settings, we must continue to study how patients perceive SDM [15]. Patients may not have experienced high levels of SDM in previous encounters so their scores might be more likely to be related to general appraisals such as satisfaction. Patients may assess SDM by factors such as "agreement" with their clinician [37], and by social desirability bias [38]. Because of these factors, we must determine the best way to ensure patients feel they took part in SDM even if observed and patient-reported SDM agreement was not reached within the consultation.

Being of lower SES was associated with lower OPTION-5 scores when not accounting for surgeon-, clinic-, and site-level clusters. This was significant across all arms and therefore the interventions had no effect on this relationship. This was unexpected given prior research showing that disadvantaged populations might benefit the most from SDM interventions and therefore have higher observed and patient-reported SDM scores [19]. However it is in-line with the large body of research that breast cancer patients of lower SES are more likely to have poorer communication with their clinician [39–41]. However, this significance was lost when accounting for the multiple levels of clustering in our analysis: four sites, seven clinics and 14 surgeons. Clustering played a large role in all of our results as successful use of the tool varied across surgeons and sociodemographic variables varied across sites. Further research into this relationship would be appropriate to determine if this relationship was truly significant given how large of an effect clustering had on this outcome.

The strengths of this analysis include the large number of recordings and paired data from both observations and patient surveys. Because of this, we were able to compare different perspectives on SDM for 311 patients across 14 surgeons. These data were also part of a large randomized controlled trial that recruited patients in three geographically diverse regions of the United States.

The analysis has limitations: we did not collect baseline measurements for the clinicians and therefore were not able to assess their pre-trial SDM scores. We did, however, conduct standardized training including role play and subsequent feedback for all surgeons in intervention arms in an effort to achieve uniform use of the conversation aids. This seems to have occurred as the OPTION-5 scores didn't increase as the surgeons used the tools during the trial. In the trial, we did not account for patient volume at the level of each surgeon when randomizing. This led to an imbalance in the number of patients per arm, with much fewer patients in the Option Grid arm. Randomization at the surgeon levels was necessary to avoid contamination, but the differing

number of patients per arm affected our analyses. The raters in this analysis could not be blinded to study arms because they were listening to the audio recordings of the clinical encounters, in which the interventions were referenced. The variation in scores between surgeons was higher than expected leading to wide ranges in OPTION-5 and collaboRATE scores within arms. Finally, there were statistical differences between patients who agreed to having their visits recorded and those who did not, particularly related to race/ethnicity, SES, and insurance status. These factors limit the generalizability of our findings.

4.2. Conclusion

Both the Picture Option Grid and Option Grid, two paper-based encounter conversation aids, led to meaningfully improved observed and patient-reported SDM across a wide range of socioeconomic strata. From this analysis, we cannot determine if either tool works better with different types of patients. Our findings support the use of both types of tools in clinical settings for patients with early-stage breast cancer.

4.3. Practice implications

The results of this study support findings from a range of studies and indicate that encounter-based conversation aids increase SDM and can be used by clinicians in clinical settings. Further studies need to establish how these types of tools can be implemented at scale.

Funding

Research reported in this manuscript is funded through a Patient-Centered Outcomes Research Institute (PCORI) Award (1511-32875). The statements presented in this manuscript are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee. The study funder has no role in any aspects of the set up or execution of the study.

Ethical approval

Ethics approval for this study was provided by the Dartmouth Committee for the Protection of Human Subjects on June 8, 2017. Montefiore Medical Center provided authorization agreement to rely on Dartmouth's approval on September 8, 2017. Ethics

approval was provided by The Washington University in St. Louis Institutional Review Board on May 9, 2017. NYU School of Medicine provided ethics approval on August 29, 2017.

Data statement

A full copy of this data set is available upon request by emailing glynelwyn@gmail.com.

CRediT authorship contribution statement

Renata W. Yen: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing - original draft, Writing - review & editing, Supervision. **Marie-Anne Durand:** Conceptualization, Investigation, Methodology, Project administration, Data curation, Visualization, Writing - review & editing, Funding acquisition, Supervision. **Camille Harris:** Data curation, Formal analysis, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Sarah Cohen:** Investigation, Writing - review & editing. **Abigail Ward:** Investigation, Writing - review & editing. **A. James O'Malley:** Funding acquisition, Methodology, Supervision, Validation, Writing - review & editing. **Danielle Schubbe:** Investigation, Writing - review & editing. **Catherine H. Saunders:** Investigation, Writing - review & editing. **Glyn Elwyn:** Conceptualization, Funding acquisition, Supervision, Writing - review & editing.

Declaration of Competing Interest

GE has edited and published books that provide royalties on sales by the publishers: the books include *Shared Decision Making* (Oxford University Press) and *Groups* (Radcliffe Press). He owns copyright in measures of shared decision making and care integration, namely **collaboRATE**, **integrate** (measure of care integration), **considereRATE** (patient experience of care in serious illness), **coopeRATE** (measure of goal setting), **toleRATE** (clinician attitude to shared decision making, Observer OPTION-5 and Observer OPTION-12 (observer measures of shared decision making).

He has in the past provided consultancy for organizations, including: 1) Emmi Solutions LLC who developed patient decision support tools; 2) National Quality Forum on the certification of decision support tools; 3) Washington State Health Department on the certification of decision support tools; 4) SciMentum LLC, Amsterdam (workshops for shared decision making).

He is the Founder and Director of &think LLC which owns the registered trademark for Option Grids™ patient decision aids; Founder and Director of SHARPNETWORK LLC, a provider of training for shared decision making. He provides advice in the domain of shared decision making and patient decision aids to: 1) Access Community Health Network, Chicago (Adviser to Federally Qualified Medical Centers); 2) EBSCO Health for Option Grids™ patient decision aids (Consultant); 3) Bind On Demand Health Insurance (Consultant), 4) PatientWisdom Inc (Adviser); 5) abridge AI Inc (Chief Clinical Research Scientist).

M-A D helped develop the Option Grid patient decision aids, which are licensed to EBSCO Health. She receives consulting income from EBSCO Health, and may receive royalties in the future. M-A D is a consultant for ACCESS Community Health Network.

CHS holds copyright in The **considereRATE** Suite of tools.

No other competing interests declared.

Acknowledgements

Thank you to the patients who participated in the What Matters Most study. Thank you to the clinicians who used the decision aids

and allowed us to record their conversations. Thank you to the clinic staff for their help with the coordination of our study at each site. Thank you to the Coproduction Laboratory for their thoughtful comments and suggestions. Thank you to Robin Paradis Montibello for helping with Observer OPTION-5 training.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.pec.2020.07.015>.

References

- [1] Breast Cancer Treatment (Adult) (PDQ®)–Patient Version, National Cancer Institute, 2019. (Accessed 10 January 2020) <https://www.cancer.gov/types/breast/patient/breast-treatment-pdq>.
- [2] Treatment of early-stage breast cancer, JAMA 265 (1991) 391–395.
- [3] I. Jatoi, M.A. Proschian, Randomized trials of breast-conserving therapy versus mastectomy for primary breast cancer: a pooled analysis of updated results, Am. J. Clin. Oncol. 28 (2005) 289–294.
- [4] M.M. Poggi, D.N. Danforth, L.C. Sciuto, S.L. Smith, S.M. Steinberg, D.J. Liewehr, C. Menard, M.E. Lippman, A.S. Lichter, R.M. Altman, Eighteen-year results in the treatment of early breast carcinoma with mastectomy versus breast conservation therapy: the National Cancer Institute Randomized Trial, Cancer 98 (2003) 697–702.
- [5] G. Elwyn, M.A. Durand, J. Song, J. Aarts, P.J. Barr, Z. Berger, N. Cochran, D. Frosch, D. Galasiński, P. Gulbrandsen, P.K.J. Han, M. Härter, P. Kinnarsley, A. Lloyd, M. Mishra, L. Perestelo-Perez, I. Scholl, K. Tomori, L. Trevena, H.O. Witteman, T. Van der Weijden, A three-talk model for shared decision making: multistage consultation process, BMJ 359 (2017) j4891.
- [6] E.D. Collins, C.P. Moore, K.F. Clay, S.A. Kearing, A.M. O'Connor, H.A. Llewellyn-Thomas, R.J. Barth Jr., K.R. Sepucha, Can women with early-stage breast cancer make an informed decision for mastectomy? J. Clin. Oncol. 27 (2009) 519–525.
- [7] G. Elwyn, A. Lloyd, N. Joseph-Williams, E. Cording, R. Thomson, M.-A. Durand, A. Edwards, Option Grids: shared decision making made easier, Patient Educ. Couns. 90 (2013) 207–212.
- [8] P. Scalia, M.-A. Durand, J.L. Berkowitz, N.P. Ramesh, M.J. Faber, J.A.M. Kremer, G. Elwyn, The impact and utility of encounter patient decision aids: systematic review, meta-analysis and narrative synthesis, Patient Educ. Couns. 102 (2019) 817–841.
- [9] M.-A. Durand, S. Alam, S.W. Grande, G. Elwyn, Much clearer with pictures": using community-based participatory research to design and test a Picture Option Grid for underserved patients with breast cancer, BMJ Open 6 (2016) e010008.
- [10] S. Alam, G. Elwyn, S. Percac-Lima, S. Grande, M.-A. Durand, Assessing the acceptability and feasibility of encounter decision aids for early stage breast cancer targeted at underserved patients, BMC Med. Inform. Decis. Mak. 16 (2016) 147.
- [11] G. Elwyn, T. Pickles, A. Edwards, K. Kinsey, K. Brain, R.G. Newcombe, J. Firth, K. Marrin, A. Nye, F. Wood, Supporting shared decision making using an Option Grid for osteoarthritis of the knee in an interface musculoskeletal clinic: a stepped wedge trial, Patient Educ. Couns. 99 (2016) 571–577.
- [12] M. Fay, S.W. Grande, K. Donnelly, G. Elwyn, Using Option Grids: steps toward shared decision-making for neonatal circumcision, Patient Educ. Couns. 99 (2016) 236–242.
- [13] V.M. Montori, M. Kunneman, J.P. Brito, Shared decision making and improving health care: the answer is not in, JAMA 318 (2017) 617–618.
- [14] F.R. Gärtner, H. Bomhof-Roordink, I.P. Smith, I. Scholl, A.M. Stiggelbout, A.H. Pieterse, The quality of instruments to assess the process of shared decision making: a systematic review, PLoS One 13 (2018) e0191747.
- [15] T. Wunderlich, G. Cooper, G. Divine, S. Flocke, N. Oja-Tebbe, K. Stange, J.E. Lafata, Inconsistencies in patient perceptions and observer ratings of shared decision making: the case of colorectal cancer screening, Patient Educ. Couns. 80 (2010) 358–363.
- [16] N.H. Geessink, E.H. Ofstad, M.G.M. Olde Rikkert, H. van Goor, J. Kasper, Y. Schoon, Shared decision-making in older patients with colorectal or pancreatic cancer: determinants of patients' and observers' perceptions, Patient Educ. Couns. 101 (2018) 1767–1774.
- [17] I. Scholl, L. Kriston, J. Dirmmaier, M. Härter, Comparing the nine-item Shared Decision-Making Questionnaire to the OPTION Scale - an attempt to establish convergent validity, Health Expect. 18 (2015) 137–150.
- [18] M. Pass, J. Belkora, D. Moore, S. Volz, K. Sepucha, Patient and observer ratings of physician shared decision making behaviors in breast cancer consultations, Patient Educ. Couns. 88 (2012) 93–99.
- [19] M.-A. Durand, L. Carpenter, H. Dolan, P. Bravo, M. Mann, F. Bunn, G. Elwyn, Do interventions designed to support shared decision-making reduce health inequalities? A systematic review and meta-analysis, PLoS One 9 (2014) e94670.
- [20] R.C. Forcino, M. Thygeson, A.J. O'Malley, M.J. Meinders, G.P. Westert, G. Elwyn, Do collaboRATE scores reflect differences in perceived shared decision-making across diverse patient populations? Evidence from a large-scale patient experience survey in the United States, J. Patient Exp. (2019) 2374373519891039.

- [21] M.-A. Durand, R.W. Yen, A.J. O'Malley, M.C. Politi, S. Dhage, K. Rosenkranz, K. Weichman, J. Margenthaler, A.N.A. Tosteson, E. Crayton, S. Jackson, A. Bradley, R.J. Volk, K. Sepucha, E. Ozanne, S. Percac-Lima, J. Song, J. Acosta, N. Mir, G. Elwyn, What matters most: protocol for a randomized controlled trial of breast cancer surgery encounter decision aids across socioeconomic strata, *BMC Public Health* 18 (2018) 241.
- [22] T.C. Hoffmann, P.P. Glasziou, I. Boutron, R. Milne, R. Perera, D. Moher, D.G. Altman, V. Barbour, H. Macdonald, M. Johnston, S.E. Lamb Kadoorie, M. Dixon-Woods, P. McCulloch, J.C. Wyatt, A.W.C. Phelan, S. Michie, Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide, *BMJ* 348 (2014) g1687.
- [23] L.D. Chew, J.M. Griffin, M.R. Partin, S. Noorbaloochi, J.P. Grill, A. Snyder, K.A. Bradley, S.M. Nugent, A.D. Baines, M. Vanryse, Validation of screening questions for limited health literacy in a large VA outpatient population, *J. Gen. Intern. Med.* 23 (2008) 561–566.
- [24] L.S. Wallace, E.S. Rogers, S.E. Roskos, D.B. Holiday, B.D. Weiss, Brief report: screening items to identify patients with limited health literacy skills, *J. Gen. Intern. Med.* 21 (2006) 874–877.
- [25] L.D. Chew, K.A. Bradley, E.J. Boyko, Brief questions to identify patients with inadequate health literacy, *Fam. Med.* 36 (2004) 588–594.
- [26] G. Elwyn, M. Tsulukidze, A. Edwards, F. Légaré, R. Newcombe, Using a “talk” model of shared decision making to propose an observation-based measure: observer OPTION 5 Item, *Patient Educ. Couns.* 93 (2013) 265–271.
- [27] P.J. Barr, A.J. O'Malley, M. Tsulukidze, M.R. Gionfriddo, V. Montori, G. Elwyn, The psychometric properties of Observer OPTION5, an observer measure of shared decision making, *Patient Educ. Couns.* 98 (2015) 970–976.
- [28] G. Elwyn, P.J. Barr, S.W. Grande, R. Thompson, T. Walsh, E.M. Ozanne, Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters, *Patient Educ. Couns.* 93 (2013) 102–107.
- [29] P.J. Barr, R. Thompson, T. Walsh, S.W. Grande, E.M. Ozanne, G. Elwyn, The psychometric properties of CollaboRATE: a fast and frugal patient-reported measure of the shared decision-making process, *J. Med. Internet Res.* 16 (2014) e2.
- [30] R.C. Forcino, M. Thygeson, A.J. O'Malley, M.J. Meinders, G.P. Westert, G. Elwyn, Measuring patient-reported shared decision-making to promote performance transparency and value-based payment: assessment of collaboRATE's group-level reliability, *J. Patient Exp.* (2019) 2374373519884835.
- [31] R.C. Forcino, P.J. Barr, A.J. O'Malley, R. Arend, M.G. Castaldo, E.M. Ozanne, S. Percac-Lima, C.D. Stults, M. Tai-Seale, R. Thompson, G. Elwyn, Using CollaboRATE, a brief patient-reported measure of shared decision making: results from three clinical settings in the United States, *Health Expect.* 21 (2018) 82–89.
- [32] StataCorp, Stata Statistical Software: Release 15, StataCorp LLC, n.d., College Station, TX, 2017.
- [33] U.S. Department of Health and Human Services, Prior HHS Poverty Guidelines and Federal Register References, Office of the Assistant Secretary for Planning and Evaluation, 2019. (Accessed 3 January 2020) <https://aspe.hhs.gov/prior-hhs-poverty-guidelines-and-federal-register-references>.
- [34] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, *Biometrics* 33 (1977) 159–174.
- [35] M.A. Vortel, S. Adam, A.V. Port-Thompson, J.M. Friedman, S.W. Grande, P.H. Birch, Comparing the ability of OPTION12 and OPTION5 to assess shared decision-making in genetic counselling, *Patient Educ. Couns.* 99 (2016) 1717–1723.
- [36] N. Couët, S. Desroches, H. Robitaille, H. Vaillancourt, A. Leblanc, S. Turcotte, G. Elwyn, F. Légaré, Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument, *Health Expect.* 18 (2015) 542–561.
- [37] L.A. Shay, J.E. Lafata, Understanding patient perceptions of shared decision making, *Patient Educ. Couns.* 96 (2014) 295–301.
- [38] D.L. Phillips, K.J. Clancy, Some effects of “Social desirability” in survey studies, *Am. J. Sociol.* 77 (1972) 921–940.
- [39] M.B. Mac Bride, L. Neal, C.A. Dilaveri, N.P. Sandhu, T.J. Hieken, K. Ghosh, D.L. Wahner-Roedler, Factors associated with surgical decision making in women with early-stage breast cancer: a literature review, *J. Womens Health* 22 (2013) 236–242.
- [40] L.A. Siminoff, G.C. Graham, N.H. Gordon, Cancer communication patterns and the influence of patient characteristics: disparities in information-giving and affective behaviors, *Patient Educ. Couns.* 62 (2006) 355–360.
- [41] S. Willems, S. De Maesschalck, M. Deveugele, A. Derese, J. De Maeseneer, Socio-economic status of the patient and doctor-patient communication: does it make a difference? *Patient Educ. Couns.* 56 (2005) 139–146.