Are you SURE?

Assessing patient decisional conflict with a 4-item screening test

France Légaré MD PhD CCMF FCMF  Stephen Kearing MSc  Kate Clay MA RN  Susie Gagnon MA  Denis D’Amours MD  Michel Rousseau PhD  Annette O’Connor RN PhD

ABSTRACT

OBJECTIVE To assess the reliability and validity of the 4-item SURE (Sure of myself; Understand information; Risk-benefit ratio; Encouragement) screening test for decisional conflict in patients.

DESIGN Cross-sectional study.

SETTING Four family medicine groups in Quebec and 1 rural academic medical centre in New Hampshire.

PARTICIPANTS One hundred twenty-three French-speaking pregnant women considering prenatal screening for Down syndrome and 1474 English-speaking patients referred to watch condition-specific video decision aids.

MAIN OUTCOME MEASURES Cronbach α was used to assess the reliability of SURE. A factorial analysis was performed to assess its unidimensionality. The Pearson correlation coefficient was computed between SURE and the Decisional Conflict Scale to assess concurrent validation. A t test procedure comparing the SURE scores of patients who had made decisions with the scores of those who had not was used to assess construct validation.

RESULTS Among the 123 French-speaking pregnant women, 105 (85%) scored 4 out of 4 (no decisional conflict); 10 (8%) scored 3 (≤3 indicates decisional conflict); 7 (6%) scored 2; and 1 (1%) scored 1. Among the 1474 English-speaking treatment-option patients, 981 (67%) scored 4 out of 4; 272 (18%) scored 3; 147 (10%) scored 2; 54 (4%) scored 1; and 20 (1%) scored 0. The reliability of SURE was moderate (Cronbach α of 0.54 in French-speaking pregnant women and 0.65 in treatment-option patients). In the group of pregnant women, 2 factors accounted for 72% of the variance. In the treatment-option group, 1 factor accounted for 49% of the variance. In the group of pregnant women, SURE correlated negatively with the Decisional Conflict Scale (r = -0.46; P<.0001); and in the group of treatment-option patients, it discriminated between those who had made a choice for a treatment and those who had not (P<.0001).

CONCLUSION The SURE screening test shows promise for screening for decisional conflict in both French- and English-speaking patients; however, future studies should assess its performance in a broader group of patients.

EDITOR’S KEY POINTS

• Decisional conflict refers to a patient’s uncertainty about the course of action to take when the choices involve risk, loss, regret, or a challenge to personal life values.

• The 4-item SURE (Sure of myself; Understand information; Risk-benefit ratio; Encouragement) screening test was developed to help health professionals identify patients with clinically significant decisional conflict as quickly as possible.

• Results of this study indicate that the SURE screening test has acceptable psychometric properties and is suitable for screening for decisional conflict in patients facing clinical decisions in primary care. As such, the tool can improve how decisions are made in family medicine, benefiting patient outcomes as a result.

This article has been peer reviewed.  
Can Fam Physician 2010;56:e308-14
Évaluer le conflit décisionnel chez les patients à l’aide d’un test de dépistage en 4 volets

France Légaré MD PhD CCMF FCMF  Stephen Kearing MSc  Kate Clay MA RN  Susie Gagnon MA  Denis D’Amours MD  Michel Rousseau PhD  Annette O’Connor RN PhD

**RÉSUMÉ**

**OBJECTIF** Évaluer la fiabilité et la validité du test SURE servant à dépister le conflit décisionnel chez les patients (SURE est l’acronyme en anglais pour : sûr de soi; comprendre l’information; rapport risques-avantages; encouragement).

**TYPE D’ÉTUDE** Étude transversale.

**CONTEXTE** Quatre groupes de médecine familiale au Québec et un centre médical universitaire rural au New Hampshire.

**PARTICIPANTS** Groupe de 123 femmes enceintes francophones envisageant le dépistage du syndrome de Down et 1 474 patients anglophones référés pour visionner une vidéo sur l’aide à la décision liée à leur maladie.

**PRINCIPAUX PARAMÈTRES ÉTUDIÉS** On a utilisé le coefficient α de Cronbach pour évaluer la fiabilité de SURE. On a effectué une analyse factorielle pour mesurer son unidimensionnalité. On a calculé le coefficient de corrélation de Pearson entre SURE et l’Échelle du conflit décisionnel pour évaluer la validation concurrente. Pour évaluer la validation du construct, on a effectué un test t comparant les résultats SURE des patients qui avaient pris leur décision avec ceux des personnes qui ne l’avaient pas prise.

**RÉSULTATS** Parmi les 123 femmes enceintes francophones, 105 (85%) ont eu des résultats de 4 sur 4 (aucun conflit décisionnel); 10 (8%) ont eu un résultat de 3 (≤3 indique un conflit décisionnel); 7 (6%) ont eu 2; et 1 (1%) a eu un résultat de 1. Parmi les 1 474 patients anglophones du groupe envisageant les options de traitement, 981 (67%) ont eu des résultats de 4 sur 4; 272 (18%) ont eu 3; 147 (10%) ont eu 2; 54 (4%) ont eu 1; et 20 (1%) ont eu 0. La fiabilité de SURE était modérée (α de Cronbach de 0,54 chez les femmes enceintes francophones et de 0,65 chez les patients du groupe des options de traitement). Dans le groupe des femmes enceintes, une part de 72% de la variation était attribuable à 2 facteurs. Dans le groupe des options de traitement, 1 facteur expliquait 49% de la variation. Dans le groupe des femmes enceintes, il y avait une corrélation négative entre SURE et l’Échelle du conflit décisionnel (r = -0,46; P < 0,0001); dans le groupe des options de traitement, le test faisait une distinction entre ceux qui avaient fait un choix de traitement et ceux qui n’avaient pas décidé (P < 0,0001).

**CONCLUSION** Le test de dépistage SURE est prometteur pour évaluer le conflit décisionnel chez les patientes francophones et les patients anglophones; par contre, les études futures devraient mesurer son efficacité dans un groupe plus large de patients.

Cet article a fait l’objet d’une révision par des pairs. Can Fam Physician 2010;56:308-14

**POINTS DE REPÈRE DU RÉDACTEUR**

- Le conflit décisionnel désigne l’incertitude des patients entourant la marche à suivre quand leurs choix comportent des risques, une perte, un regret éventuel ou un conflit avec des valeurs de vie personnelles.
- Le test de dépistage en 4 volets SURE (sûr de soi; comprendre l’information; rapport risques-avantages; encouragement) a été conçu pour aider les professionnels de la santé à identifier aussitôt que possible les patients qui vivent un conflit décisionnel important sur le plan clinique.
- Les résultats de cette étude indiquent que les propriétés psychométriques du test de dépistage SURE sont acceptables et que ce test convient dans le dépistage du conflit décisionnel chez les patients aux prises avec des décisions cliniques en soins primaires. Comme tel, l’outil peut améliorer la façon dont les décisions sont prises en médecine familiale et, par conséquent, les résultats chez les patients.
most primary health care decisions occur in contexts of uncertainty. Because the probability of risks and benefits in a population cannot be directly attributed at the individual level, uncertainty is inevitable when making decisions in family medicine. In a study of shared decision making in primary care, 54% of 924 patients who had made decisions with their doctors stated that they were uncertain about which option was best. It is clear from this that both family physicians and their patients must manage uncertainty when making clinical decisions.

Decisional conflict refers to an individual’s perception of uncertainty about the course of action to take when the choices involve risk, loss, regret, or a challenge to personal life values. In lay terms, it indicates an individual’s level of comfort with a decision. The Decisional Conflict Scale (DCS) is a unique research-oriented tool that assesses decisional conflict in patients. Unlike the Physician Reaction to Uncertainty scale and the Ambiguity Aversion Medical scale, which were developed to assess the actors’ predisposition to uncertainty (a trait), the DCS assesses individual perceptions of uncertainty about which course of action to take (a state). The DCS counts 16 items and has been translated and validated in Dutch, French, and Spanish. The DCS scores correlate with scores for patient knowledge, treatment and screening intentions, and decisional regret, as well as the patient’s blaming of his or her doctor for a negative outcome. Physicians can use the DCS to reduce the downstream effects of unresolved decisional conflict by evaluating decisional conflict in their patients and providing appropriate support.

However, the time required to administer the DCS discourages its use. For that reason, we sought to develop a tool that would help health professionals identify patients with clinically significant decisional conflict as quickly as possible. We accordingly developed the 4-item SURE (Sure of myself; Understand information; Risk-benefit ratio; Encouragement) screening test for decisional conflict in patients and used a 2-step process to assess its reliability and validity.

**METHODS**

Creating the 4-item SURE test
We based our selection of the 4 items on core concepts of the Ottawa Decision Support Framework, which are relevant at all stages of decision making: feeling uncertain, feeling informed, feeling clear about values, and feeling supported in decision making. A fifth core concept (ie, the perceived effectiveness of the choice made) was not applicable to all stages of decision making. The wording of the French and the English questions was developed concurrently and framed in a positive manner to match the acronym SURE. The resulting 4-item test was field-tested with experts and graduate students taking clinical courses in decision support. Table 1 shows the French and English versions of the 4-item SURE screening test.

**Table 1. The SURE test: A) English and B) French versions. A response of yes scores 1 and a response of no scores 0; a score of <4 is a positive result for decisional conflict.**

<table>
<thead>
<tr>
<th>A) SURE ACRONYM IN ENGLISH</th>
<th>ENGLISH VERSION OF TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sure of myself</td>
<td>Do you feel SURE about the best choice for you?</td>
</tr>
<tr>
<td>Understand information</td>
<td>Do you know the benefits and risks of each option?</td>
</tr>
<tr>
<td>Risk-benefit ratio</td>
<td>Are you clear about which benefits and risks matter most to you?</td>
</tr>
<tr>
<td>Encouragement</td>
<td>Do you have enough support and advice to make a choice?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B) SURE ACRONYM IN FRENCH</th>
<th>FRENCH VERSION OF TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sûr</td>
<td>Êtes-vous certain de ce qui constitue le meilleur choix pour vous?</td>
</tr>
<tr>
<td>Utilité de l’information</td>
<td>Est-ce que vous connaissez les bénéfices et risques de chacune des options?</td>
</tr>
<tr>
<td>Risques-bénéfices à balancer</td>
<td>Avez-vous le sentiment de savoir ce qui est le plus important pour vous à l’égard des risques et bénéfices?</td>
</tr>
<tr>
<td>Encouragement</td>
<td>Avez-vous suffisamment de soutien afin de faire votre choix?</td>
</tr>
</tbody>
</table>

Copyright O’Connor and Légaré, 2008.

**Clinical setting, study participants, and data collection**
Between April 2007 and December 2008, 2 distinct data-collection processes were performed in a stepwise approach, using convenience samples. The first group of consecutive patients consisted of French-speaking pregnant women who were considering prenatal screening for Down syndrome; they were recruited from 4 family medicine groups in Quebec city. Women were eligible if they were between the ages of 18 and 34 years, had no family history of genetic disorders, and were between 8 and 12 weeks pregnant. Women whose pregnancies were at risk were excluded. Patients were told that the study aimed at describing shared decision making in the context of prenatal screening. After their first routine prenatal screening test,
consultations, the women provided sociodemographic information and completed a self-administered 16-item DCS. Each item was measured on a 5-point Likert scale (1 = strongly agree and 5 = strongly disagree). The women also completed the self-administered 4-item SURE test (Table 1) with 2 response categories: yes (score = 1) and no (score = 0).

The second group consisted of consecutive English-speaking patients at a rural academic medical institution (ie, the Dartmouth Hitchcock Medical Center in New Hampshire) who were referred to watch condition-specific video decision aids as part of their standard process of care. The conditions addressed by the videos were chronic low back pain, spinal stenosis, herniated disk, hip osteoarthritis, knee osteoarthritis, prostate cancer treatments, early-stage breast cancer surgery, and breast reconstruction after cancer. The video for each condition provided information about treatment options, discussed the potential benefits and risks of each option, invited the viewer to consider the values he or she associated with each option, and reviewed the importance of patient involvement in decision making. After watching the video decision aids, study participants completed a self-administered questionnaire, which included the 4-item SURE questions. They did not complete the 16-item DCS.

Data processing and analysis
In both groups of patients, we performed descriptive data analyses and assessed the internal reliability of the SURE test by computing Cronbach α. The tetrachoric correlation coefficient was used to assess item-to-item correlations and the Pearson correlation coefficient was used to assess item-to-total correlations. Because decisional conflict is a state rather than a trait, it was not appropriate to assess intrarater reliability over time. Also, as the SURE test is a self-administered instrument, interrater reliability was similarly irrelevant. Factor analysis was performed using the principal components analysis method for factor extraction, with varimax orthogonal rotation. The number of factors retained was based on the minimum eigenvalue of 1 criterion.

In the group of French-speaking pregnant women, we computed the mean DCS score by adding the values for all items, dividing the sum by 16, and multiplying the product by 25. The DCS scores ranged from 0 (no decisional conflict) to 100 (high decisional conflict). Previous research shows that women whose scores exceed 37.5 experience clinically significant decisional conflict. Totals of the SURE test were computed by adding the response scores of the 4 questions. We then assessed the criterion validity of the SURE test with the DCS by using a Pearson correlation coefficient—the hypothesis being that SURE scores would correlate negatively with DCS scores. (A perfect score on the SURE test indicates no decisional conflict; while a high score on the DCS indicates high decisional conflict.) In the group of treatment-option patients, treatment intentions were dichotomized (ie, those who made a choice about treatment versus those who were unsure about treatment), and SURE items were summarized as the frequency (percentage) of endorsed responses. Construct validity by extreme groups was assessed using a t test procedure—the hypothesis being that SURE would discriminate between patients who made choices of treatment and patients who did not. All calculations were performed using Statistical Analysis System version 9.1. Patients were not compensated financially. The study was approved by the institutional review boards of the institutions where data collection took place.

**RESULTS**

**Participant characteristics**
We approached 180 French-speaking pregnant women registered at family medicine clinics in Quebec and requested their participation. Of these women, 21 were ineligible and 11 declined to participate. Of the 148 women recruited (response rate of 82%), 141 completed the DCS and 123 completed the SURE test. Table 2 summarizes participants’ sociodemographic characteristics. Of the 141 participants who completed the DCS, 7 presented clinically significant decisional conflict. Of the 123 participants who completed the SURE test, 108 (85%) scored 4 out of 4, 10 (8%) scored 3, 7 (6%) scored 2, and 1 (1%) scored 1.

| Table 2. Patient characteristics: Mean (SD) age was 28.6 (3.51) years for French-speaking pregnant patients in Quebec and 59.3 (13.2) years for English-speaking patients facing treatment decisions in New Hampshire. |
|---------------------------------|-------------------------------|-------------------------------|
| CHARACTERISTIC                  | FRENCH-SPEAKING PATENTS       | ENGLISH-SPEAKING PATENTS       |
|                                 | N = 123, N (%)                | N = 1474, N (%)                |
| Female                          | 123 (100)                     | 765 (52)                      |
| Education*                      |                               |                               |
| • University                    | 62 (50)                       | 671 (46)                      |
| • Some college education or a high school diploma | 56 (46) | 689 (47) |
| • Less than a high school diploma | 5 (4)                         | 96 (7)                        |
| *Education category for English-speaking patients does not add to 100 owing to missing data. |

The English-speaking patients in the treatment-option group were systematically distributed video decision aids and SURE questionnaires. A total of 1474 patients (34%) completed and returned the questionnaire. Of these 1474 patients, 981 (67%) patients scored 4 out of 4; 272 (18%) scored 3; 147 (10%) scored 2; 54 (4%)
scored 1, and 20 (1%) scored 0 (Figure 1). Table 3 summarizes patients’ yes responses to the 4-item SURE questions.

Reliability of SURE
The internal reliability of SURE was moderate (Cronbach $\alpha$ was 0.54 in French-speaking pregnant women and 0.65 in English-speaking treatment-option patients). In the group of pregnant women, removing 1 item (ie, support) produced a higher value (Cronbach $\alpha=0.61$). In the group of treatment-option patients, all item-to-item correlations were positive and ranged from 0.46 to 0.71. Item-to-total correlation results are presented in Table 4. In pregnant women, 1 item (ie, support) was negatively correlated with 2 items (ie, knowledge and values) and was poorly correlated with the total score. This item showed no variance in this group of respondents.
Concurrent and construct validation

As expected, in the group of French-speaking pregnant women, the SURE score correlated negatively with the DCS score ($r = -0.46; P < .0001$). Also as expected, in the English-speaking treatment-option group, patients who had not made choices about treatment ($n = 225$) had lower mean (SD) SURE scores than those who had ($n = 1249$) (2.6 [1.0] vs 3.6 [0.8], respectively; $P < .0001$). More specifically, among the treatment-option patients who scored 4 on the SURE test, only 4% had not made treatment choices compared with 34% of patients who had scored 3, 41% of patients who had scored 2, 46% of patients who had scored 1, and 35% of patients who had scored 0 (Figure 1).

A factorial analysis of the SURE test in the group of pregnant women indicated that 2 factors accounted for 72% of the variance. Three items (ie, knowledge, value, and certainty) loaded under 1 factor. The other item (ie, support) loaded under the second factor. As expected, in the treatment-option group, 1 factor accounted for 49% of the variance.

Table 4. Item-to-total Pearson correlation coefficient results for the 2 participant groups

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>SURE OF MYSELF</th>
<th>UNDERSTAND INFORMATION</th>
<th>RISK-BENEFIT RATIO</th>
<th>ENCOURAGEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>French-speaking pregnant women, N = 123</td>
<td>0.47</td>
<td>0.32</td>
<td>0.59</td>
<td>0.07</td>
</tr>
<tr>
<td>English-speaking treatment-option patients, N = 1474</td>
<td>0.46</td>
<td>0.33</td>
<td>0.45</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Our findings suggest that the SURE test has acceptable psychometric properties and is suitable for screening for decisional conflict in French- and English-speaking patients facing clinical decisions in primary care. As such, the tool can improve how decisions are made in family medicine, benefiting patient outcomes as a result. We expect this phenomenon to occur in a 2-step process.

First, to the best of our knowledge, SURE is the only clinically oriented instrument that helps physicians identify patients experiencing clinically significant decisional conflict. In developing the 4-item test, our team drew inspiration from the 4-item CAGE questionnaire (Have you ever felt that you should cut down on your drinking? Have people annoyed you by criticizing your drinking? Have you ever felt bad or guilty about your drinking? Have you ever had a drink first thing in the morning to steady your nerves or to get rid of a hangover [eye-opener?] used for screening for alcohol abuse in clinical practice. The literature reports that CAGE’s short format makes it more attractive to busy clinicians than the longer version and thereby increases its use. It is not unlikely that the 4-item SURE test will act similarly, making clinicians more willing to seek to identify patients experiencing substantial decisional conflict.

Second, the downstream effects of unresolved clinically significant decisional conflict in patients suggest that the use of SURE in clinical practice has the potential to improve patient outcomes. In the group of French-speaking pregnant women who had consulted with their family physicians, testing revealed that 7 (6%) of them were experiencing clinically significant decisional conflict regarding prenatal screening for Down syndrome. This relatively high incidence suggests real benefits to coupling the use of the SURE tool with a decision support system so that positive test results (a score of 3 or less) trigger physicians to help patients make decisions or refer patients to appropriate resources.

Limitations

Notwithstanding the interest of our findings, our study has several limitations. First, we reported on data collections that occurred independently. The drawbacks of this procedure are offset by the fact that it produced a larger sample size with more descriptive data and better evidence of the tool’s acceptability (2 languages and 2 countries) than would otherwise be possible. It also allowed us to assess the tool’s relevance in different clinical situations. Second, in the group of French-speaking pregnant women, we observed a less-than-optimal value of Cronbach α and could not confirm SURE’s unidimensionality. This suggests that there might be a need to modify 1 item (ie, support). However, it is possible that the lack of variance in the SURE scores (most individuals had perfect scores) might have contributed to this observation, as the magnitude of a reliability coefficient is highly dependent on the homogeneity of subjects in a group. Thus, it is possible that we underestimated the true reliability of the SURE test. For this reason, it is important that future studies assess the performance of SURE in groups with a larger proportion of patients experiencing clinically significant decisional conflict.

Conclusion

Results from this study suggest that the SURE test is suitable for screening for decisional conflict in French- and English-speaking patients with a variety of health conditions. Future research should assess the performance of the SURE test with a more diverse group of patients.
Dr Légaré is a Full Professor in the Department of Family Medicine at Laval University in Quebec city, Que, and Canada and Tier 2 Canada Research Chair in Implementation of Shared Decision Making in Primary Care. Mr Kearing is a research consultant and Mrs Clay is a decision support coach at the Dartmouth Institute for Health Policy and Clinical Practice in New Hampshire. Mrs Gagnon is a research assistant at the CHUQ Research Centre at the Hôpital St-François d’Assise in Quebec city. Drs D’Amours and Rousseau are professors in the Department of Family Medicine at Laval University. Dr O’Connor is a Distinguished Professor in the School of Nursing at the University of Ottawa in Ontario.

Contributors
All authors contributed to concept and design of the study; data gathering, analysis, and interpretation; and preparing the manuscript for submission.

Competing interests
None declared

Correspondence
Dr France Légaré, CHUQ Research Centre, Hôpital St-François d’Assise, 10 rue de l’Espinay, Quebec city, QC G1L 3L5; telephone 418 525-4351; e-mail france.legare@mfa.ulaval.ca

References