Use of the SF-36 questionnaire to predict quality-of-life improvement after laparoscopic colorectal resection for endometriosis

G. Dubernard¹, R. Rouzier¹, E. David-Montefiore¹, M. Bazot² and E. Darai¹,³

¹Department of Obstetrics and Gynecology, Hôpital Tenon, AP-HP, Université Pierre et Marie Curie Paris VI, 4 rue de la Chine, 75020 Paris, France; ²Department of Radiology, Hôpital Tenon, AP-HP, Université Pierre et Marie Curie Paris VI, 4 rue de la Chine, 75020 Paris, France

³Correspondence address. Tel: +33-1-56-01-73-18; Fax: +33-1-56-01-73-17; E-mail: emile.darai@tnn.ap-hop-paris.fr

BACKGROUND: Laparoscopic colorectal resection for endometriosis can improve quality of life (QOL), but the results vary widely from one woman to another. The aim of this study was to determine whether the preoperative results on the Physical Component Summary (PCS) and Mental Component Summary (MCS) subscales of the SF-36 questionnaire could predict the improvement in QOL after surgery. METHODS: The predictive value of the subscales was first evaluated on a training set of 57 patients. A mathematical model, quantified with respect to discrimination and calibration was then applied to the validation set of 36 patients. RESULTS: Women with preoperative PCS and MCS scores below 37.5 and 44.5, respectively, had 80.7% and 84.2% probabilities of seeing their scores improve after surgery, whereas women with preoperative scores above 46.5 and 47.5, respectively, had probabilities of 0% and 10.7% to improve their scores. CONCLUSIONS: With our mathematical model, the postoperative improvement in QOL can be reliably predicted. This model should help to identify those women who are most likely to benefit from this major surgery.

Keywords: colorectal endometriosis; laparoscopy; SF-36 health status; quality of life

Introduction

Deep infiltrating endometriosis (DIE) is defined as infiltration of anatomical structures and organs by endometriotic tissue composed of endometrial glands and stroma (Koninckx et al., 1991; Nisolle and Donnez, 1997; Anaf et al., 2000). The exact incidence of DIE in the general population is not known, but DIE is estimated to affect 20% of women with endometriosis (Chapron et al., 1999). Colorectal endometriosis represents almost 90% of all forms of intestinal endometriosis (Coronado et al., 1990; Bailey et al., 1994; Tran et al., 1996; Jerby et al., 1999) and is one of the most incapacitating forms, significantly altering quality of life (QOL) (Garry et al., 2000; Redwine and Wright, 2001).

The fibrotic component represents about 80% of DIE lesions, and medical therapies are ineffective (Anaf et al., 2002; Thomassin et al., 2004; Darai et al., 2005). As a result, colorectal endometriosis must often be surgically removed. Previous studies have demonstrated the feasibility of laparoscopic segmental colorectal resection for endometriosis, and also its efficacy on both digestive and gynaecological symptoms (Redwine et al., 1996; Urbach et al., 1998; Redwine and Wright, 2001; Thomassin et al., 2004; Darai et al., 2005; Dubernard et al., 2006; Darai et al., 2007; Seracchioli et al., 2007). Moreover, an improvement in QOL has been reported on the basis of visual analogue scales (Jones and Sutton, 2003; Darai et al., 2005; Dubernard et al., 2006). These results were subsequently confirmed by the use of validated questionnaires (Abbott et al., 2003; Jones et al., 2004). Recently, using the short-form (SF)-36 questionnaire, Dubernard et al. (2006) observed a significant improvement in QOL among women undergoing laparoscopic colorectal resection for endometriosis. However, although some women had a major improvement in the SF-36 score after surgery, others had only a slight improvement. It would therefore be useful to be able to distinguish between these subgroups of women preoperatively in order to select the best candidates for surgery.

In women with Douglas pouch obliteration by endometriosis, Redwine and Wright found that the presence of a palpable nodule reproducing on physical examination the spontaneous pain reported by the patient was predictive of a good response to surgery (Redwine and Wright, 2001). However, this sign appears to depend on practitioner experience and is not applicable to limited or non-palpable colorectal endometriosis. No
other markers potentially identifying good candidates for surgery have been tested in this setting.

In this study, we first examined on a training set of patients whether preoperative results of the SF-36 questionnaire were predictive of the change in QOL after laparoscopic colorectal resection for endometriosis. We then used a validation set of patients to test the utility of the SF-36 questionnaire for predicting the improvement in QOL after this major operation.

Materials and Methods

Patients

Between April 2001 and December 2006, we performed a prospective study including 93 women with colorectal endometriosis referred to the gynaecology department of Tenon Hospital, Paris. Before surgery, all the women underwent both magnetic resonance imaging and rectal endoscopic sonography. In order to avoid a possible bias linked to the type of surgery, only women undergoing segmental laparoscopic colorectal resection were selected. All the patients gave their informed consent to participate in the study and were informed of the possible need for laparconversion.

Methods

Laparoscopic colorectal resection was performed as previously described (Darai et al., 2005, 2007; Dubernard et al., 2006).

All the women completed a validated French translation of the SF-36 Health Status questionnaire (Perneger et al., 1995; Lepege et al., 1998). The SF-36 is a multi-purpose health survey with only 36 questions. It yields an eight-scale profile of functional health and well-being scores, as well as psychometrically based physical and mental health summary measures (standardized). The eight scales are hypothesized to form two distinct higher-ordered clusters due to the physical and mental health variance that they have in common. Among the eight scales, three [physical functioning (PF), role physical (RP), bodily pain (BP)] correlate most strongly with the physical component and contribute most to the Physical Component Summary (PCS) score. The mental component correlates best with the mental health (MH), role emotional (RE) and social functioning (SF) scores, which also contribute most to the Mental Component Summary (MCS) score. Two of the scales [vitality (VT) and general health (GH)] have noteworthy correlations with both components.

Mathematical model and statistical analysis

The study population was composed of two groups of women: a training set used to build the model and an independent validation set to test the model. All the women completed the SF-36 questionnaire both before and after the operation. The training set consisted of women who underwent laparoscopy for colorectal endometriosis and whose QOL results have already been published (Dubernard et al., 2006).

Recursive partitioning (RP) was used to determine SF-36 cut-offs predicting an improvement in QOL. RP is a technique which can be applied to mine large data sets in order to uncover hidden patterns within the data and to elucidate statistically significant subgroupings within the data. RP is non-parametric in nature, imposing no \textit{a priori} restrictions on the distributional forms of the predictor variables. The central result is a RP algorithm simple and intuitive. At each step, the RP program determines for each variable (in this case, score results) cut-points that optimally separate patients into homogeneous groups. The RPART program, implemented in R, was used to generate decision trees depicting the classification rules generated by RP.

For each scale, we considered that an improvement occurred if the postoperative score was better than the preoperative score (except for BP: improvement if postoperative score > preoperative score + 10). For the standardized PCS and MCS measures, we considered that there was an improvement if the postoperative score was improved by 1 SD (based on the preoperative score of the training set) compared with the preoperative score. Of note, the standard deviations were not significantly different between the data sets and the preoperative or postoperative measures.

The model performance was quantified with respect to calibration on the training set. Calibration corresponds to the agreement between observed outcome frequencies and predicted probabilities. It was studied with graphical representations of the relationship between the observed outcome frequencies and the predicted probabilities (calibration curves) for groups of patients defined by the final leaves in RP model. A calibration curve can be approximated by a regression line with intercept \( \alpha \) and slope \( \beta \). These parameters can be estimated in a logistic regression model with the event as outcome and the linear predictor as the only covariate. Well-calibrated models have \( \alpha = 0 \) and \( \beta = 1 \) (ideal predictions are plotted on a 45° line that cross the origin). Therefore, a sensible measure of calibration is a likelihood ratio statistic testing the null hypothesis that \( \alpha = 0 \) and \( \beta = 1 \). The statistic has a \( \chi^2 \) distribution with 2 df (unreliability [\( U \)]-statistic) (Cox, 1958). Of note, the unreliability [\( U \)]-statistic is satisfactory if the \( P \)-value is not significant. All analyses were performed by RR, using the R package with Rpart, Design, Hmisc and Lexis libraries (http://lib.stat.cmu.edu/R/CRAN/).

To compare epidemiological and clinical characteristics, parametric and non-parametric continuous variables were compared with Student’s \( t \)-test and the Mann–Whitney test, and categorical variables were compared with the \( \chi^2 \) test or Fisher’s exact test, as appropriate. Correlations were identified with ANOVA and Spearman’s test. \( P \)-values of <0.05 were considered statistically significant.

Results

Comparison of epidemiological and clinical characteristics

No difference in epidemiological characteristics was found between the two sets of patients (Table I). The only difference in the preoperative SF-36 scores was a higher BP score in the training set than in the validation set (\( P < 0.0001 \)) (Table II).

After laparoscopic segmental colorectal resection, an improvement was observed on the eight scales of the SF-36 questionnaire in both the training set and the validation set (Fig. 1). Strong improvements were noted in BP, RP, SF and

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Training set ((n = 57))</th>
<th>Validation set ((n = 36))</th>
<th>(P)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age (years)</td>
<td>31 (21–49)</td>
<td>34 (25–52)</td>
<td>NS</td>
</tr>
<tr>
<td>Median body mass index (range)</td>
<td>19 (18–30)</td>
<td>25 (19–29)</td>
<td>NS</td>
</tr>
<tr>
<td>Median parity (range)</td>
<td>0.5 (0–3)</td>
<td>0.4 (0–2)</td>
<td>NS</td>
</tr>
<tr>
<td>Median gestation (range)</td>
<td>0.5 (0–4)</td>
<td>0.4 (0–2)</td>
<td>NS</td>
</tr>
<tr>
<td>Percentage of women with previous surgery for endometriosis</td>
<td>50%</td>
<td>60%</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS, not significant.
RE (Fig. 1). The median preoperative PCS in the training set and validation set were, respectively, 44 (SD: 8.3) and 49 (SD: 9.9), whereas the median postoperative PCS values were, respectively, 53 (SD: 6.1) and 50 (SD: 9.0). The median MCS in the training set and validation set was, respectively, 47 (SD: 5.2) and 46 (SD: 4.9) preoperatively, and, respectively, 50 (SD: 5.3) and 49 (SD: 4.1) postoperatively.

Model development

Table III shows the results of the model based on the training set of 57 women with colorectal endometriosis. For each item of the SF-36 questionnaire, the model determined cut-points separating the patients into three subgroups (except for RE: 2 subgroups) with different probabilities of improvement. All the calibration curves were correct, with no significant differences between predicted and observed probabilities of improvement in the validation set.

We also build a RP model and created trees for the PCS and MCS that used different preoperative SF-36 values. The trees are represented in Figs 2 and 3. Women with a preoperative PCS score below 37.5 had an 80.7% probability of having an improved PCS score after surgery. Women with a preoperative PCS score between 46.5 and 37.5 had a 33.3% probability of a score improvement after surgery, whereas no postoperative improvement in PCS could be expected by women with a preoperative score over 46.5 (Fig. 2).

Women with a preoperative MCS score below 44.5 had an 84.2% probability of a score improvement after surgery. Women with a preoperative MCS score between 44.5 and 47.5 had a 30% probability of improving their score after surgery, whereas women with a preoperative MCS score above 47.5 had only a 10.7% probability (Fig. 3).

Validation

Postoperatively, in the validation set, the results of the mathematical model on the physical composite score and mental composite score were, respectively, 50 and 49. We tested the calibration of the models on the validation set and found no significant difference between the predicted and the observed

---

**Table III.** Results of the mathematical model based on each item of the SF-36 questionnaire in the training set of 57 women with colorectal endometriosis.

<table>
<thead>
<tr>
<th>Item of the SF-36 questionnaire</th>
<th>Initial score, Initial score,</th>
<th>Probability of improvement, Probability of improvement,</th>
<th>P for validation set*</th>
</tr>
</thead>
<tbody>
<tr>
<td>General health</td>
<td>&lt;36</td>
<td>36–63, &gt;63</td>
<td>0.06</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>&lt;53</td>
<td>53–97, &gt;97</td>
<td></td>
</tr>
<tr>
<td>Role physical</td>
<td>&lt;37</td>
<td>37–62, &gt;62</td>
<td>0.16</td>
</tr>
<tr>
<td>Role emotional</td>
<td>&lt;17</td>
<td>0%</td>
<td>0.99</td>
</tr>
<tr>
<td>Social functioning</td>
<td>&lt;19</td>
<td>19–56, &gt;56</td>
<td>0.68</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>&lt;31</td>
<td>32–56, &gt;56</td>
<td>0.07</td>
</tr>
<tr>
<td>Vitality</td>
<td>&lt;38</td>
<td>38–52, &gt;52</td>
<td>0.37</td>
</tr>
<tr>
<td>Mental health</td>
<td>&lt;54</td>
<td>54–65, &gt;65</td>
<td>0.09</td>
</tr>
</tbody>
</table>

*Unreliability [U]-statistic: satisfactory if the P-value is not significant.
probabilities (Table III). The calibration of the model was also good for the PCS and MCS, with no significant difference between predicted and observed probabilities ($P = 0.57$ and $P = 0.22$, respectively, for PCS and MCS). This confirmed that the model was sufficiently robust to predict the improvement in QOL after segmental colorectal resection for endometriosis (Fig. 4). In this latter figure, the model performance was quantified with respect to calibration on the training set: the agreement between the predicted probabilities ($x$-axis) and the observed outcome frequencies ($y$-axis) was good as points were plotted close of the 45° line that crosses the origin (perfect prediction). This tool can thus be used to inform patients about the probability of improvement in each item of the SF36 questionnaire, and in the PCS and MCS.

Discussion
This study shows that preoperative assessment of QOL with the SF-36 questionnaire can predict the QOL improvement after laparoscopic segmental colorectal resection for endometriosis. Like Redwine and Wright (2001), we found that women with higher preoperative symptom scores had greater improvement of postoperative symptom scores. But to predict the postoperative improvement of symptoms, Redwine and Wright use the ability to reproduce the patient’s pain during the preoperative examination. According to the severity and the variability of symptoms described by patients and to the postoperative morbidity, this sign appears limited. As we previously demonstrated, using the MOS SF-36 questionnaire, we encompassed not only pain but also impact on the others components of the QOL (Dubernard et al., 2006). These QOL questionnaires are more pertinent for selecting good candidates for radical surgical resection of deep pelvic endometriosis.

It is a major challenge for physicians treating women with colorectal endometriosis to select the best candidates for colorectal resection. Indeed, this is a major operation exposing women to a risk of potentially severe complications such as rectovaginal fistulae and de-novo dysuria, with no guarantee of an improvement in QOL. One way of predicting the response to anticancer treatment is to use mathematical models to identify predictive factors. In RP, a divisive clustering technique, data are divided on the basis of specified attributes. RP is a non-parametric type of analysis that repeatedly subdivides data into smaller and smaller subgroups based on characteristics that predict the desired end-point. The goal is to construct subgroups that, ideally, consist entirely of subjects in one end-point category or another. In this study, we used a RP model to test the capacity of the SF-36 questionnaire to predict the change in QOL of women undergoing colorectal resection for endometriosis. Validation of the mathematical model on a training set of women undergoing segmental colorectal resection for symptomatic endometriosis demonstrated that SF-36 score cut-points could be used to predict the QOL improvement after surgery.

The mathematical model based on the training set robustly predicted the change in QOL when tested on a validation set. Indeed, calibration of the RP model showed a strong correlation between the predicted and the observed improvements in SF-36 scores in the independent validation set, even though the preoperative BP score was lower in the validation set than in the training set.

PCS and MCS cut-points were defined to predict the improvement in QOL. A preoperative PCS score below 37.5 was associated with an 80.7% probability of QOL improvement for this component after colorectal resection, whereas a preoperative MCS score below 44.5 was associated with an
84.2% probability of QOL improvement for this component. For patients with a preoperative PCS and/or MCS predictive score, respectively, higher than 46.5 or 47.5, surgery cannot be recommended. The risk of surgical complications is too high compared with the QOL improvement expected. For these patients, medical treatments (hormonal or non-hormonal) should be preferred.

This study confirms that laparoscopic segmental colorectal resection for endometriosis can improve QOL (Fig. 1). The improvement was similar to that observed in other studies using validated questionnaires (Redwine et al., 1996; Urbach et al., 1998; Garry et al., 2000; Redwine and Wright, 2001; Abbott et al., 2003; Thomassin et al., 2004; Darai et al., 2005; Dubernard et al., 2006; Darai et al., 2007; Seracchioli et al., 2007). Abbott et al. (2003) demonstrated that concomitant hysterectomy and bilateral salpingo-oophorectomy during removal of endometriotic lesions was associated with a larger QOL improvement than fertility-sparing surgery. Few of our patients underwent radical surgery, probably because the study population included a large proportion of young, often nulliparous women wishing to preserve their childbearing potential. However, conservative surgery carries a high risk of recurrence. In particular, young women have a higher rate of pain recurrence (Fedele et al., 2004; Vignali et al., 2005). Moreover, the absence of bowel resection in women with DIE and intestinal endometriosis is the factor most strongly associated with the actuarial recurrence rate (Fedele et al., 2004).

Certain limitations of this study must be underlined. First, our results only concern women with bowel endometriosis, the most severe form of pelvic endometriosis carrying a risk of severe complications after surgical excision. Second, the study took place in a single gynaecology department, and our algorithm thus requires validation by other teams. Third, our algorithm only concerns women undergoing laparoscopy, and therefore may not be relevant to women undergoing laparotomy. However, we have previously observed a significant improvement in SF-36 scores after laparoconversion, with no difference relative to women who underwent purely laparoscopic procedures (Dubernard et al., 2006).

Although all items of the SF-36 questionnaire were improved in our patients, some were improved more than others, suggesting that a simplified questionnaire could be used in clinical practice. An SF-12 questionnaire already exists. This questionnaire reproduces the eight-scale profile with fewer levels than SF-36 scales and yields less precise scores, as would be expected for single-item and two-item scales. For large group studies, these differences are not as important, because confidence intervals for group averages in health scores are largely determined by sample size. But our sample is quite short (93 patients) and we wanted to have the highest discriminative index to demonstrate the benefits of the SF-36 questionnaire to select our patients for this complex surgery. This is the reason why we used the original SF-36 questionnaire. Instead of using a simpler set of questions, we plan to implement a web site to help the practitioner using our model.

We explored all the scores and component summaries as proof of principle that our approach is robust, although it should be noted that a patient with very specific symptoms due to endometriosis might see little improvement in her QOL after surgery.

In conclusion, our results confirm that laparoscopic segmental colorectal resection for endometriosis is associated with a significant improvement in QOL, as evaluated with the SF-36 questionnaire. We also show that preoperative SF-36 scores can be used to predict the degree of change in QOL after this major operation. This validated questionnaire, available in several languages, offers a simple tool to help practitioners select and inform women who might benefit from laparoscopic segmental colorectal resection for endometriosis.

References


Submitted on October 8, 2007; resubmitted on November 28, 2007; accepted on December 11, 2007.