The Long-term Benefit of Surgery on Health-Related Quality of Life in Patients With Inflammatory Bowel Disease

Richard C. Thirlby, MD; Marco A. Sobrino, MD; James B. Randall, RN

**Hypotheses:** Health-related quality of life (HRQL) has been shown to improve dramatically shortly after surgery in patients with inflammatory bowel disease (IBD). Our hypotheses were that (1) improved HRQL would be maintained long term in patients after surgery for ulcerative colitis and (2) the improved HRQL in patients with Crohn disease would decline with long-term follow-up.

**Design:** Consecutive series of patients undergoing surgery for IBD between June 1994 and January 2000 prospectively investigated as a cohort outcomes study.

**Patients:** Data were obtained in 139 patients. The diagnoses were Crohn disease (n=56) and ulcerative colitis (n=83).

**Intervention:** Patients with Crohn disease underwent resections with or without stricturoplasties; all but 5 patients with ulcerative colitis underwent ileal pouch-anal anastomoses.

**Main Outcome Measure:** Health status was measured using the Health Status Questionnaire (HSQ) preoperatively and then every 3 months postoperatively.

**Results:** Preoperative HSQ scores were very low in all 8 scales of the HSQ. Postoperatively, HRQL measures improved significantly (P<.05) both in patients with Crohn disease and ulcerative colitis, with scores equal to or better than published scores in the general population. In patients with Crohn disease, the scores improved significantly after surgical resection and steadily increased despite disease recurrence and reoperations. The HRQL at last follow-up was equivalent to the general population. The improvements were statistically significant in patients followed up for more than 1 year in 7 of 8 scales of the HSQ.

**Conclusions:** These results confirm that HRQL is poor in patients with IBD referred for possible operation. Surgical resection resulted in significant improvement in HRQL. More important, the results were durable. With follow-up up to 6 years, the HRQL in this cohort was equal to or better than norms for the general population both in patients with ulcerative colitis and with Crohn disease. We believe these data justify aggressive surgical intervention in many patients with IBD and support the prospective study of HRQL by surgeons treating patients with chronic diseases.

**Arch Surg. 2001;136:521-527**
PATIENTS, MATERIALS, AND METHODS

STUDY DESIGN

Subjects eligible for this study were all patients with IBD seen by a single surgeon (R.C.T.) between June 1994 and January 2000. All patients were considered candidates for bowel resection. Three questionnaires were completed preoperatively for each patient: (1) a patient-completed disease-specific questionnaire, (2) a physician-completed disease-specific symptom assessment form, and (3) a patient-completed Health Status Questionnaire (HSQ).

An intraoperative form was also completed for each patient that described disease activity and operative procedures. We attempted to obtain HSQs 3, 6, and 12 months postoperatively and then at each follow-up visit subsequently. In addition, HSQs were mailed to the patients.

HEALTH STATUS QUESTIONNAIRE

RAND developed the 36-Item Short-Form Health Survey (SF-36) to capture behavioral dysfunction caused by health problems.17,18 The HSQ is a modification of the SF-36, adding 3 questions (total of 39) to assess better emotional or mental health.16 The HSQ has been demonstrated to be easy to understand by patients, takes only minutes to complete, and has excellent consistency and reliability between subgroups of patients with different chronic diseases.16 The 39 questions are collapsed into 8 dimensions or scales of health status. The scores for each scale ranged from 0 to 100, with higher scores for better health status.

The 8 scales are general health perception (6 questions), physical functioning (10 questions), physical role limitations (the interference with work by physical health) (4 questions), emotional role limitations (interference with work by emotional problems) (3 questions), social functioning (2 questions), mental health (8 questions), bodily pain (2 questions), and energy/vitality (4 questions).

PATIENT POPULATION

One hundred thirty-nine patients were evaluated. Twenty-five percent of patients were self-referred, and 75% were physician referred. Twenty-five did not undergo operation; therefore, the study population was 104 patients.

The clinical diagnosis was Crohn disease in 56 patients and ulcerative colitis (UC) in 83 patients. Patients seen only for complications of perianal Crohn disease were excluded from the analysis. The average age of the patients with Crohn disease was 41 years (range, 21-77 years), with 24 men and 32 women. The average age of patients with UC was 44 years (range, 15-75 years), with 39 men and 24 women. Table 1 summarizes the current medications and associated comorbidities of the study cohort.

STATISTICAL ANALYSIS

Mean scores for all HSQ scores were calculated. The changes in HSQ results in individual patients with both preoperative and postoperative data were calculated as gain scores. Because the HSQ gain scores exhibited a nonnormal distribution, the Wilcoxon matched pairs signed rank test was used, with P<.05 considered significant.

RESULTS

SURGICAL INTERVENTION

All but 5 patients with UC underwent ileal J-pouch-anal anastomoses. All patients with Crohn disease underwent resections with or without strictureplasties. The distributions of small bowel, colon, and small bowel plus colon disease were 57% (n=22), 34% (n=13), and 9% (n=4), respectively. In patients with small bowel disease, 53% of patients had more than 30 cm of bowel resected. Fifty-nine percent had single small bowel site involvement, and 41% had multiple small bowel site involvement. Patients with Crohn disease who required only incision and drainage of perirectal abscesses were not evaluated in the present study. Nearly all patients with Crohn disease were treated postoperatively with prophylactic mesalamine, azathioprine, or methotrexate. Five patients with Crohn disease required additional abdominal operations.

HSQ SCORES

Figure 1 shows the average HSQ scores for all forms recorded preoperatively and postoperatively (ie, multiple postoperative HSQs on most patients). These data, therefore, represent the average HRQL for all patients seen by a surgeon for preoperative consultation and the average HRQL for all patients subsequent to surgical resections, regardless of duration of follow-up. For comparison, Figure 1 includes previously published data on the general population and on patients with diabetes mellitus.16 As shown, the preoperative HRQL in our patients was poor, with particularly low scores in the scales of general health (health perception), role limitations due to physical health, social functioning, and energy level. These scores are lower than those reported for patients with diabetes in 7 of 8 scales. Postoperative scores, on the other hand, are virtually equal to the scores for the general population.

Preoperative and postoperative HSQs were obtained in 104 patients. Changes between scores in each of the 8 scales were then individually compared with cal-
ulate gain scores. The gain scores in all 104 patients and in those followed up for more than 12 months are summarized in Table 2. For the entire study cohort and in patients with UC, the improvements were significant in all 8 scales of the HSQ. Most of the improvements seen in the patients with Crohn disease, although large, were not statistically significant.

CROHN DISEASE

The results in patients with Crohn disease are illustrated in Figure 2 and Figure 3. Figure 2 shows the mean preoperative and postoperative HSQ scores in the 39 patients who completed both forms, regardless of duration of follow-up. The scores for role limitations due to physical health, bodily pain, and energy level are low preoperatively. With an average follow-up of about 16 months (range, 3-76 months), scores increased, approaching normal levels in 7 of 8 domains. Figure 3 shows the preoperative and postoperative scores in the patients with Crohn disease followed up for more than 12 months (mean, 26 months; range, 12-76 months). The mean HSQ scores in all 8 domains were actually higher in these patients than those in the entire group. Mean scores were higher than previously published norms for the general population in all 8 scales.

ULCERATIVE COLITIS

Figure 4 and Figure 5 show the results in the patients with UC. Figure 4 shows the mean preoperative and postoperative HSQ scores in the 65 patients who completed both forms, regardless of duration of follow-up. Preoperatively, the lowest scores are in the scales of role limitations due to physical health, social functioning, and energy level. Postoperatively, with an average follow-up of only 5 months (range, 3-70 months), scores equaled or exceeded published norms in 7 of 8 domains, remaining low in role limitations due to physical health. As shown in Figure 5, the results in patients with UC followed up for more than 1 year (mean, 32 months; range, 12-70 months) were similar. However, the average scores were lower that those reported for the general population in all 8 scales. However, preoperative scores were very low, especially in the scales of role limitations due to physical health and energy level. The improvements in the ability to function at home and work were dramatic: 49 points in patients followed up for more that 12 months. Similarly, the improvements in social functioning, bodily pain, and energy levels were also significant ($P<.001$).

EFFECT OF LONG FOLLOW-UP

Figure 6 shows the average gain scores for the patients followed up for 12 months or less and those followed up for more than 12 months. Gain scores were higher in the long follow-up group (mean follow-up, 29 months; range, 12-76 months). The improvements were statistically significant in all 8 scales in all patients with IBD and in the patients with UC but only 1 of 8 scales in patients with Crohn disease.

EFFECTS OF AGE AND SEX

Our previous study$^{15}$ suggested that the beneficial effects of surgery were greater in younger patients. The gain
scores (and average postoperative raw HSQ scores) in patients less than or greater than our mean age are listed in Table 3. In contrast to our previous study, the present findings suggest that older patients benefit as much as or more than younger patients. Positive gain scores were nearly identical in 6 or 8 scales and greater in 2 scales: role limitations due to physical health and bodily pain. Also shown in Table 2 are scores in men and women. There was no apparent difference, except for a greater improvement in bodily pain in women.

**COMMENT**

We believe that patient-derived HRQL data are the most important clinical outcomes that one can measure. The present study confirms that surgical interventions positively affect HRQL in patients with IBD. Patients with IBD have markedly impaired HRQL, with most studies suggesting that patients with Crohn disease have poorer HRQL than patients with UC. Emphasizing the importance of the patient’s perspective, family members and physicians consistently underestimate dysfunction or HRQL. Assessment of HRQL, or health status from the patient’s perspective, is a powerful tool to assess and quantify disease outcomes. It is important to note, however, factors that are disease independent can affect HRQL. Factors that affect quality of life (eg, socioeconomic status and marital status) influence HRQL scores. For example, employed persons with high annual incomes score consistently higher on HRQL testing than an unemployed cohort with similar chronic disease status. Women with disease activities similar to men report lower HRQL scores. In other words, patients with identical disease-specific symptoms, such as stool frequency, might have divergent social or emotional reactions to their disease and, as a result, divergent HRQL. Because our study does not control for or evaluate disease-independent factors, it is important to acknowledge that our results demonstrate an association between surgical intervention and improved HRQL in patients with IBD.
What is the best measure of HRQL in patients with IBD? We chose a generic test, the HSQ. The advantage of the HSQ is that it has been tested in diverse population groups and has been established to have high interval consistency and validity.16 The disadvantage is that the HSQ could be insensitive to small but clinically significant disease-specific factors. Many have reported the use of an IBD-specific questionnaire, the Inflammatory Bowel Disease Questionnaire (IBDQ).19,20 The IBDQ has been shown to be both reliable and valid and responsive to changes in outpatients being treated medically for Crohn disease. Many experts recommend the use of a disease-specific measure such as the IBDQ since it may have better responsiveness or sensitivity to treatment and will relate to areas routinely analyzed by clinicians.1,3,19,20 However, in our opinion, it can improperly associate disease-specific symptoms with improved HRQL. For example, stool frequency (a measure in the IBDQ) in patients after ileal reservoir procedures may be higher than before surgery, but the overall sense of well-being is consistently improved. In addition, the IBDQ has not been thoroughly assessed for evaluating surgical patients.3 In summary, we continue to use the generic HSQ as our assessment tool in preference to the disease-specific IBDQ recognizing advantages and disadvantages. The fact that the improvements in HRQL as assessed by the HSQ were very large in our study suggests that it is sensitive enough to be used in patients with IBD.

The present study is not a controlled trial comparing HRQL in patients treated medically with those treated surgically. Most patients were referred by gastroenterologists and were refractory to maximal medical measures. Thus, this cohort is a select group of patients with more severe disease than those in most medical series and cannot be compared with a cohort of patients treated medically. We report HSQ scores after surgery equal to the general population, which is superior to the results in most medical series. Studies7,8,11,15 of patients with IBD managed medically, although showing improvement in HRQL, have found that scores in patients with IBD in remission are significantly worse than those in healthy individuals, suggesting that HRQL is permanently impaired in patients with IBD. We and others7,8,11,15 have reported normalization of HRQL in patients with UC. Our previous study, with short follow-up, concluded that patients with Crohn disease also had near normal HRQL after surgery.15 After surgical resections of Crohn disease, endoscopic, clinical, and surgical recurrences occur in approximately 95%, 70%, and 50% of patients, respectively. Our expectation, therefore, was that the longer follow-up in the present study would result in frequent disease recurrence and resultant decline in HRQL. Despite many clinical recurrences and additional resections in a few patients, HRQL was actually better in the patients with Crohn disease with longer follow-up. However, we acknowledge that most patients undergoing operation have positive expectations regarding clinical outcomes, and our significant effects on HRQL may reflect a significant placebo effect. Regardless, we believe our results demonstrating postoperative HRQL equivalent to that of healthy controls support the conclusion that surgical intervention is effective in patients with IBD.3,12

In conclusion, the present study confirms the finding of our preliminary report: the HRQL of patients with IBD referred for surgery is poor and improves markedly after surgery. With larger numbers and longer follow-up, we have shown that these findings are durable in patients with UC and in those with Crohn disease. As expected, in patients with Crohn disease, frequent recurrence negatively affected HRQL. However, HRQL at last follow-up was equivalent to the general population. Given the established effect of HRQL on patient satisfaction, health care expenditures, and utilization, we believe that surgical intervention is underused in the care of patients with IBD. Ideally, future studies in these patients should be collaborative and multidisciplinary, incorporating prospective assessment of the effects of both medical and surgical interventions on the HRQL of these patients.

Presented at the 108th Scientific Session of the Western Surgical Association, Dana Point, Calif, November 14, 2000.

Corresponding author and reprints: Richard C. Thirlby, MD, Section of General, Thoracic, and Vascular Surgery, Virginia Mason Medical Center, 1100 Ninth Ave, PO Box 900 (C6-GSUR), Seattle, WA 98111 (e-mail: Richard.Thirlby@vmmc.org).

Table 3. Mean Gain Scores After Operation*

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Age £49 y</th>
<th>Age &gt;49 y</th>
<th>Men (n = 53)</th>
<th>Women (n = 53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health perception</td>
<td>20 (74)</td>
<td>16 (62)</td>
<td>17 (75)</td>
<td>11 (72)</td>
</tr>
<tr>
<td>Physical functioning</td>
<td>17 (89)</td>
<td>10 (79)</td>
<td>3 (81)</td>
<td>7 (79)</td>
</tr>
<tr>
<td>Role physical</td>
<td>67 (89)</td>
<td>41 (68)</td>
<td>22 (59)</td>
<td>26 (62)</td>
</tr>
<tr>
<td>Role emotional</td>
<td>19 (93)</td>
<td>19 (74)</td>
<td>14 (84)</td>
<td>19 (85)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>29 (92)</td>
<td>25 (67)</td>
<td>19 (76)</td>
<td>21 (72)</td>
</tr>
<tr>
<td>Mental health</td>
<td>15 (79)</td>
<td>12 (73)</td>
<td>6 (78)</td>
<td>5 (27)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>36 (83)</td>
<td>23 (66)</td>
<td>9 (70)</td>
<td>20 (70)</td>
</tr>
<tr>
<td>Energy/vitality</td>
<td>32 (70)</td>
<td>27 (50)</td>
<td>17 (74)</td>
<td>22 (34)</td>
</tr>
</tbody>
</table>

*Values in parentheses are mean postoperative Health Status Questionnaire scores.
REFERENCES


DISCUSSION

Merrill T. Dayton, MD, Salt Lake City, Utah: (Discussion read by Dr Fabrizio Michelassi in Dr Dayton's absence.) I would like first to convey to you Dr Dayton's apologies for not being here as President Rothhammer announced. He had to leave this morning very early because of the death of a very close friend.

Dr Thirlby and colleagues have conducted an interesting study evaluating "health-related quality of life" or HRQL. Unlike "survival" or "recurrence," which have traditional clinical endpoints, the authors assess an end point based on the patient's perspective. The simple question asked is this: does surgical intervention improve the quality of life in patients with Crohn disease and UC?

The authors attempt to answer the question by administering a health quality questionnaire to 104 patients preoperatively, postoperatively, and every 3 months for the first year plus at each subsequent visit. The questionnaire utilized in this study has been validated and found to have excellent consistency and reliability.

The results of their study revealed that preoperative HRQL is very low compared to the general population; stated more simply, these are sick patients who have miserable lives because of their disease. However, the postoperative HRQL demonstrates that the quality of life improves considerably after surgical intervention. Importantly, this improvement in quality of life holds up over time.

I have a few questions I would like the authors to address at this time. (1) It is unclear from the manuscript what percentage of patients are available for at least 1 year of follow-up, and what was the mean length of follow-up? (2) While most patients with UC that is "moderately severe" or "severe" will have improved quality of life after surgical intervention, the subset of patients with "mild" or "asymptomatic" disease who undergo the operation for dysplasia or cancer could conceivably have a worse quality of life because of high stool frequency associated with the procedure and incontinence. Did you do a subset analysis of this group to see if their quality of life was worsened after surgery? (3) The authors state in the manuscript that "most of the improvements seen in patients with Crohn disease, although large, were not statistically significant." Can the authors explain why patients with Crohn disease seem to benefit less from surgical intervention? Could it be explained by recurrence alone? (4) Because the HRQL results are less impressive for Crohn disease than UC, does this mean that surgery should play a lesser role in the management of Crohn disease and that we should rely more on pharmacological strategies?

In conclusion, I would like to commend the authors for studying patients' quality of life. It is important to know that we really do do these people to enjoy their lives and have as high a quality of life as possible.

Edward T. Peter, MD, Red Bluff, Calif: I think this is a very important study pointing out that we are beginning to pay attention to the quality of life and not just surgical results. I brought along on this trip a bestseller that I picked up called “The Human Side of Cancer” by Holland and Louis. It pays attention to quality-of-life factors in patients who have cancer. I have been interested in the quality of life in breast cancer patients for a number of years, paying attention to whether the patients have had reconstruction or no reconstruction. I think it is important for all of us to pay attention to quality-of-life factors. I would like to ask the authors whether they have expanded this study to patients with malignancies as well.

Erwin R. Thal, MD, Dallas, Tex: I noticed that most of the patients in the UC group had the anal pouch or the pull-through operation. I was just curious if you had a subset and looked at the quality of life of those patients who just had illectomies.

Richard A. Prinz, MD, Chicago, Ill: Most surgeons do a global assessment of the outcome of their operations and include quality of life in that. Have you checked the surgeon assessment of outcome with the health survey questionnaire results to see if there was correlation? If so, it may not be necessary to do this because it does take a substantial amount of effort.

Basil A. Pruitt, Jr, MD, San Antonio, Tex: How do you avoid a “learning” or placebo effect? With repeated administration of the same instrument, it seems quite possible that the patients will learn or want to give a more optimistic response than is actually the case. Second, how many patients didn't return the questionnaire, and are they the ones in whom improvement was very limited or actually nil? Third, how do you explain your finding that HRQL did not deteriorate even in those patients with UC who had surgery? (2) The authors state in the manuscript that “most of the improvements seen in patients with Crohn disease, although large, were not statistically significant.” Can the authors explain why patients with Crohn disease seem to benefit less from surgical intervention? Could it be explained by recurrence alone? (4) Because the HRQL results are less impressive for Crohn disease than UC, does this mean that surgery should play a lesser role in the management of Crohn disease and that we should rely more on pharmacological strategies?

I would like to ask the authors whether they have expanded this study to patients with malignancies as well.

Fabrizio Michelassi, MD, Chicago: I agree completely with you that we need more and more often to look at quality of life, especially in benign diseases and IBD. My question is whether...
in the patients with Crohn disease you have done subset analysis based on primary disease vs recurrent disease, location of disease (ie, small bowel vs colon or anorectal), and indication to surgery (such as failure of medical treatment, obstructive or septic complications).

Dr Thirlby: We have been advocates of quality-of-life measurements for years. However, we need to acknowledge some limitations of these studies. For example, quality of life is not the same as HRQL. Quality of life is determined by socioeconomic factors such as marital status, financial status, and gender. People who are employed and in a happy marriage score much better on HRQL surveys than those who are single and unemployed. As surgeons, we have no control of these factors. Any HRQL study must be suspect if you have not controlled for nonmedical factors in quality of life, and obviously I have done none of that in this study.

There is also clearly going to be a placebo effect in this study, which Dr Pruitt alluded to. These patients have invested their heart and soul in an operation and they want to do better, and I don't think that there is any question that that does affect their ability to fill out this form honestly. We all must acknowledge that placebo effect.

The other thing that we need to realize is that not every disease is amenable to HRQL studies. I dare say that not too many vascular surgeons make their patients with abdominal aortic aneurysms feel better with their procedure. It is very hard to make a woman with asymptomatic breast cancer feel better with surgical intervention. So there are obviously limitations.

Having said that, I really do believe I have demonstrated today very impressive effects of surgery on quality of life. In previously published medical series, patients with Crohn disease in medical remission have poorer quality of life than the general population. We have shown quite convincingly that our quality-of-life results are equivalent to the general population after surgical resection.

Dr Dayton asked several questions, one about the mean follow-up. The mean follow-up in this study is just less than 2 years, which is long enough to make this study very valid. Our previous study had a mean follow-up of only about 6 months, and in most of the studies in the medical and surgical literature, follow-up is less than 1 year, so I do believe this is a long follow-up study.

Dr Pruitt's questions were about a subset analysis in patients with dysplasia. Many of these patients come to our office with 1 stool a day, they feel well, and they have dysplasia in the setting of longstanding, almost asymptomatic, UC. I have not done a specific subset analysis in those patients. The numbers are too small. Anecdotally, all of us who operate on these patients routinely here note at 1-year follow-up that these patients feel much better than they did before surgery and that they didn't realize how sick they had been prior to this operation. I will say, however, that I tell patients with familial adenomatous polyposis, or FAP, that I am going to adversely affect their quality of life.

Dr Pruitt, I acknowledged the placebo effect was significant. What about patients lost to follow-up? I have lost several patients to follow-up. At the Virginia Mason Medical Center, virtually all of the patients who go through our Gastrointestinal Department with recurrences do come back to my office. I personally think patients lost to follow-up are likely doing better than the average, which improves my results rather than makes them worse. I think the patients who experience recurrence come back and find us.

Our subset analyses show that patients with colonic Crohn disease do better than those with small bowel Crohn disease. I have not been able to demonstrate any difference in length of resection or numbers of resections. In fact, again, without statistical significance, our impression is the patients with worse disease have greater improvements after surgical resections.

I would like to close with one last slide. I recently received this letter from an ileal pouch patient. I think it demonstrates the effect of surgical intervention on patients with IBD. She states, “I count my blessings every day. I think of my scars as beauty marks that remind me of how much I have to be thankful for. I wear them proudly. I haven't felt this good since I was diagnosed with UC 10 years ago.”

We have all received letters like this. We smile; we show them to our residents, show them to our nurse, and then we regretfully put them in the trash. What I have done today is objectively measure the sentiments expressed on this slide. I would contend that measurements that objectively quantify the sentiments expressed on this slide are more important than any other measure we can select, whether it be cost of care, length of stay, stool frequency, or disease recurrence. I urge all of you to do the same.