

Effect of Intervention in Inflammatory Bowel Disease on Health-Related Quality of Life: A Critical Review

Robert G. Maunder, M.D., F.R.C.P.(C),* Zane Cohen, M.D., F.R.C.S.(C),†
Robin S. McLeod, M.D., F.R.C.S.(C),† Gordon R. Greenberg, M.D., F.R.C.P.(C)‡

*From the Departments of *Psychiatry, †Surgery, and ‡Medicine, University of Toronto and the Inflammatory Bowel Disease Centre, Mount Sinai Hospital, Toronto, Ontario, Canada*

Health-related quality of life (HRQOL) is a quantitative measurement of subjective perception of health state, including emotional and social aspects. It can be reliably measured with several valid instruments. Previous reviews of the literature suggested inadequate attention to HRQOL in studies of interventions in inflammatory bowel disease. **PURPOSE:** This study was undertaken to assess the current status of the quality of measurement of HRQOL in studies of inflammatory bowel disease and to review the clinical conclusions warranted by the literature. **METHOD:** Medline was searched for articles relating to ulcerative colitis, Crohn's disease, or inflammatory bowel disease and quality of life since 1981. The articles found were reviewed for citations of further articles. The adequacy of HRQOL measure was assessed and graded, and the study design was categorized to assess the strength of the literature on the whole. **RESULTS:** A trend was found toward higher quality of HRQOL measurement in the period 1988 to 1994 compared with 1981 to 1987. Most of the improvement was because of increased use of standardized and multidomain but unvalidated and unpublished questionnaires for measurement. **CONCLUSIONS:** Confidence in the following clinical conclusions in studies of surgical interventions in inflammatory bowel disease is limited by study design: that pelvic pouch is not inferior to ileostomy, that specific domains of HRQOL are differentially affected by different surgical procedures, and (with less confidence) that surgery is helpful in Crohn's disease. Medical studies have demonstrated that high quality HRQOL measures can be integrated into randomized, prospective trials. Clinically equivalent treatments have shown differential effects on HRQOL: 9 mg daily of budesonide is superior to 15 mg, and hydrocortisone foam enemas are superior to prednisolone. Home parenteral nutrition has received modest support, limited again by study design. It is recommended that standard tests of HRQOL be used to increase comparability of studies and to increase the quality of this literature in general. In particular the Inflammatory Bowel Disease Questionnaire, Rating Form of Inflammatory Bowel Disease Concerns, and Direct Questioning of Objectives are recommended. [Key words: Quality of life; Ulcerative colitis; Crohn's disease; Inflammatory bowel disease; Pelvic pouch; Ileostomy]

Maunder RG, Cohen Z, McLeod RS, Greenberg GR. Effect of intervention in inflammatory bowel disease on health-related quality of life: a critical review. *Dis Colon Rectum* 1995;38:1147-1161.

Over the past 15 to 20 years, health-related quality of life (HRQOL) has developed from an abstract concept to a reliably measured index of subjective health status. It has been advocated as a useful measure in evaluating interventions in many chronic diseases. Measurement of HRQOL facilitates health care planning^{1,2} and is useful in evaluating clinical interventions.³

Inflammatory bowel disease (IBD) has a significant impact on HRQOL. Whereas qualitative and semi-quantitative reviews of HRQOL in IBD are now abundant,⁴⁻⁶ studies using reliable and valid HRQOL measures have been less common. In a 1987 review⁷ only 2 of 53 surgical trials in all illnesses measuring quality of life were randomized, controlled, and prospective. This review examines studies that use HRQOL as an outcome measure of medical and surgical interventions in IBD. It has two purposes: to determine the quality of HRQOL measurement and to summarize the conclusions that this literature allows regarding treatment of IBD. A brief review of the relevant issues in measurement of HRQOL will clarify the criteria by which the literature was reviewed.

Health-related quality of life has been defined as "a global measure of the patient's perceptions, illness experience, and functional status that incorporates social, cultural, psychologic, and disease-related factors."⁸ It is essentially an attempt to quantify the patient's subjective evaluation of these aspects of experience. Guidelines for the development and use of HRQOL measures are available.^{3,9,10} It is generally measured in several domains probing physical, emotional, and social function. Typically HRQOL domains have included sexual activity, social activity, ability to work or attend school, sports and recreation, and body image. The instrument should generate at least one summary score and possibly several subscores. Measures of HRQOL may be applicable to many chronic diseases (general instruments) or may be dis-

Address reprint requests to Dr. Maunder: Department of Psychiatry, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5.

ease-specific. Several general measures have been available since the early 1980s, including those commonly used in the validation of IBD-specific instruments—the sickness impact profile (SIP),¹¹ the RAND questionnaire,¹² and the quality-of-life index.¹³ IBD-specific measures have included questionnaires that are clinician-administered^{14–16} or self-administered.^{17, 18} Questionnaires that generate only qualitative information should be distinguished from instruments that allow statistical testing. In this review “measure” is used to describe any qualitative or quantitative probe of HRQOL, whereas “instrument” is reserved for a measure whose results can be tested statistically.

Each method of measuring HRQOL has advantages and limitations. General instruments provide comparison between diagnostic groups and may be useful in health service planning. General instruments also provide a standard against which the validity of new disease-specific instruments can be tested. Disease-specific instruments provide more information about particular concerns that are most relevant to a given population. They may be more responsive to change in HRQOL (evaluative instruments), or they may be more sensitive to differences between groups (discriminative instruments).¹⁰ Measures chosen in a study will depend on the purposes for measurement.

In reviewing the literature, the following criteria were considered to be important for optimal measurement of HRQOL.^{10, 19, 20} 1) Reproducibility: studies using self-developed, unpublished instruments to measure HRQOL do not allow for replication of findings or certainty in interpreting their results. 2) Reliability: measures of general and disease-specific HRQOL should show test-retest reliability in stable populations and inter-rater reliability if they are not self-administered. 3) Validity: there is no “gold standard” of HRQOL. Standard measures of HRQOL should at least show face and consensual validity. Better, construct validity should be demonstrated by correlation with other measures of function according to prediction.¹⁴ Comparison with independently validated general measures of HRQOL also bolsters confidence in validity.¹⁸ 4) Ease of use: instruments should be affordable and not too time-consuming. The SIP, for example, takes approximately 90 minutes to complete, whereas some disease-specific instruments take as little as 5 to 10 minutes. 5) Responsiveness to change: to be useful in intervention studies, evaluative measures of HRQOL must be as sensitive as possible to changes in subjective state. Reliable but relatively insensitive measures will not detect changes

in state before and after interventions or will require larger sample sizes.¹⁰ 6) Meaningfulness of results: the results of HRQOL measurement should be readily interpretable. Semiquantitative categories (good, fair, poor) provide very little interpretive information. Multidomain instruments allow scrutiny of the components of HRQOL, which may be differentially affected by disease and treatment. 7) Sampling of patient’s perspective: QOL categorization based on clinician’s impression measures a different phenomenon. Health-related quality of life is subjective by definition. Correlations between physician global ratings and validated HRQOL measures are positive but vary in their strength.^{14, 16, 18}

Two IBD-specific HRQOL instruments have now been demonstrated to have reasonable reliability, responsiveness, and validity for evaluative use: the IBD Questionnaire (IBDQ)¹⁶ and the Rating Form of IBD Patient Concerns (RFIPC).¹⁸ A third, the Quality of Life Questionnaire developed by Farmer *et al.*,¹⁵ has been validated as an IBD-specific discriminative instrument.

The IBDQ is a 32-item, clinician-administered questionnaire that was designed as an evaluative instrument for clinical trials. It was refined from a pool of 150 qualitative patient concerns generated by careful questioning of 77 IBD patients and a multidisciplinary group of clinicians in daily contact with IBD patients. The number of items was reduced by determining the items chosen most frequently and rated most important by 97 IBD patients.¹⁴ Serial testing to improve wording and consultation with experienced clinicians resulted in the 32-item IBDQ. These items fell into four domains: gastrointestinal symptoms, systemic symptoms, emotional dysfunction, and social dysfunction. The final version of the IBDQ was tested for reproducibility, responsiveness, and validity.¹⁰ Reproducibility in stable patients was better for emotional and social function than for the symptom domains, and there was a trend toward improvement in all scores, suggesting a practice effect. The IBDQ was more responsive to changes in clinical function than the RAND questionnaire, and changes in scores were greater in ulcerative colitis (UC) than Crohn’s disease (CD) patients. The IBDQ was found to be responsive enough to be suitable for small *n* clinical trials (20 patients per group in parallel group designs). Validity was tested by comparison of actual to predicted correlations to physician and patient global ratings, the RAND questionnaire, and a disease activity index. The validated IBDQ has more recently been modified to a

self-administered 36-item form¹⁷ and has been further validated in a CD population.²¹

The RFIPC is a 25-item, self-administered questionnaire developed by Drossman and colleagues.²² Twelve items were drawn from patient concerns and gathered by an experienced clinician, with other items generated by video-taped patient interviews designed to encourage disclosure of subjective concerns. In the early 21-item version,²³ 150 patients were asked to rate each concern on a visual analog scale, with results reported *via* an average "sum-score." Correlation of the RFIPC to the SIP was 0.46 for UC and 0.48 for CD patients. Significant correlation to physician's global assessment ($r = 0.48$) was found in UC patients only. Physician's global rating is the only disease activity measurement in this study, but it supports the conclusion that the RFIPC is more responsive to changes in disease activity than the SIP. Test-retest reliability over a two-week to ten-week period for the sumscore was 0.87. Information about changes in disease activity for this period is not reported (an ideally responsive test would show similar score only in those patients whose condition is unchanged). A subsequent mail survey of 320 UC and 671 CD patients²² provides more evidence of validity. Factor analysis yields four factors: impact of disease, sexual intimacy, complications of disease, and body stigma. Factors were used to determine four similarly named indices from among the questionnaire items. All four indices associate strongly to disease and demographic variables, which were then factored out for subsequent analyses. Comparisons of the four RFIPC factor domains to subscales of the SIP and to the symptom check list (SCL-90, an indicator of psychologic distress) and global ratings of well-being yield significant correlations for the impact of disease factor. This factor is significantly associated with poor sense of health and well-being on self-report, increased psychologic distress on SCL-90, and poorer function on all SIP subscales. The sexual intimacy factor of the RFIPC is significantly associated with the psychologic distress scale of the SIP. A significant advantage of multi-item questionnaires such as the RFIPC is the information they provide about the relative importance of different patient concerns, which allowed Drossman *et al.*^{22, 23} to rank patient's disease-related concerns. This could allow for a finer evaluation of the precise impact of clinical intervention on specific concerns, rather than a single measure of whether HRQOL gets better or worse.

A clinician-administered HRQOL questionnaire has

also been developed by Farmer and colleagues¹⁵ to provide brief and practical HRQOL measurements. This questionnaire uses 18 questions in four domains (functional/economic, social/recreational, affect/life, and general) with responses on the Likert 1 to 5 scale (strongly agree to strongly disagree). These questions were selected from a 45-item questionnaire as the items that most significantly distinguished between four groups of IBD patients (CD surgical, CD nonsurgical, UC surgical, UC nonsurgical). Test-retest reliability over a two-week period was good (Spearman correlation coefficient, 0.75–0.95), and inter-interviewer kappa values were reported as "high." "The SIP and IBD quality-of-life questionnaires had statistically significant correlations between similar components. Both instruments were equal in discriminating the [most severely ill] Crohn's surgical patients," but unfortunately comparison data are not provided. Responsiveness to change in clinical status is not tested. This questionnaire shows an ability to distinguish cross-sectionally between groups but has not been designed or tested as an evaluative instrument for clinical trials.

Other HRQOL instruments have been adopted for use in IBD trials. The Time Trade-Off Technique (TTOT)^{24, 25} is a modification of the von Neumann standard gamble developed for cost-effectiveness analysis of different treatment programs competing for health care funding. In the TTOT patients are asked to compare two imaginary "alternatives of certainty," health state 1 (the state being tested such as "on TPN" (total parenteral nutrition) or "with ileostomy") for a normal life span and health state 2 (good health) for time x (less than a normal life span). Time x is varied to determine the point of indifference between alternatives. The resulting time "traded off" is converted into an index between 0 (death) and 1 (health). Negative values for states "worse than death" are allowed. The TTOT is a relatively insensitive discriminator for states close to 1.²⁴ Direct Questioning of Objectives (DQO)²⁵ was developed to test HRQOL in patients on TPN. Patients generate a list objectives in life that they are then weighted according to the extent to which they are important to the patient. Objectives are scored (0–1.0) according to the patients ability to achieve them in each of the health states being compared. DQO thus provides for highly individualized domains of HRQOL. McLeod and colleagues²⁶ provided validation data for the TTOT and DQO in 93 patients with ulcerative colitis, finding correlations in the expected direction to scores on the

Beck depression inventory (TTOT, $r = 0.50$ /DQO, $r = 0.61$), SIP (0.49/0.59), and global assessments by physician (0.44/0.52), patient (0.52/0.79), and significant other (0.43/0.60). The DQO was a better discriminative instrument, because it, but not the TTOT, distinguished between mild, moderate, and severely affected patients. The TTOT was also found by Mitchell *et al.*¹⁴ to lack discriminative power in a UC population.

METHODS

A literature search of the Medline database since 1966 for articles concerning UC, CD, IBD, or colitis and quality of life revealed 150 citations. Consistent with the assumption that qualitative studies provide little reliable information about the effects of intervention on HRQOL and the intent to evaluate the quality of the HRQOL literature since introduction of valid measures, citations before introduction of reliable general measures of HRQOL (designated as 1981) were deemed to be qualitative and were not reviewed (17 articles). Review articles, case reports, editorials, cross-sectional studies, and naturalistic follow-up studies not directly concerned with clinical interventions were scanned for citations of relevant studies but not themselves included in this analysis. Eleven new articles were added to the review. Articles were also excluded if they made no direct measure of HRQOL, were not written in English, concerned primarily children, did not relate to IBD, and were not in peer-reviewed journals. A total of 45 articles were thus included for review.

Studies were sorted in two ways: according to the quality of HRQOL measurement and study design and then according to their clinical content. In the first review studies were divided into four groups: prospective and controlled, prospective and uncontrolled, retrospective and controlled, retrospective and uncontrolled. Then, to indicate the quality of HRQOL measurements, three standards were defined to categorize the studies. Group I includes studies that use HRQOL measures composed of several items in multiple appropriate domains *and* for which there is evidence of reliability and validity. Any study using the IBDQ, RFIPC, or other validated evaluative instruments would be in this class. Group II includes studies that are either multidomain *or* have demonstrated reliability and validity but not both. Group III includes the remainder of the studies that used measures of HRQOL ranging from semiquantitative to qualitative

(*e.g.*, a single question as to whether one is satisfied with an operation in retrospect). Studies that did not adequately communicate the content of self-made instruments were also included in Group III. In the case of borderline determinations, articles were classified with higher quality studies.

In the second sort, studies were grouped according to their clinical content so results could be reviewed separately for 1) different surgical interventions in UC, 2) surgical treatment of CD, 3) medical treatment of IBD, and 4) home parenteral nutrition. These categories exhausted the topics addressed by the reviewed articles. In each section, reports of higher quality according to the first review were given greater weight in drawing overall conclusions.

RESULTS AND DISCUSSION

Results are presented first by study design and then by the clinical questions that have been addressed with HRQOL as an outcome measure.

Study Design

The 45 intervention studies meeting inclusion criteria for review were grouped according to the quality of their HRQOL instrument. One study²⁶ included both prospective and retrospective wings and was included in both categories, giving a total of 46 studies. The resulting distribution of studies is given in Table 1. Twenty-two of forty-six studies (48 percent) use Group III HRQOL measures. Twenty-nine (63 percent) are retrospective and uncontrolled. Nineteen (41 percent) are retrospective, uncontrolled, and in the lowest class of HRQOL measurement.

More recent studies have included more controlled trials and have shifted toward the use of multidomain HRQOL instruments (Table 1). Most of this improvement is accounted for by a large increase in the number of retrospective studies using Group II instruments. In the period from 1988 to 1994, 13 of 27 studies used Group II measures (7 controlled, 6 uncontrolled), 9 used Group III, and 5 used Group I. There are only five prospective, controlled trials in this literature. Two of these use a Group I instrument to measure HRQOL. The quality of studies examining HRQOL in IBD is improving. Nonetheless, the preponderance of clinical trials has been retrospective, uncontrolled studies with the lowest standard of HRQOL measurement. Retrospective and uncontrolled design is the easiest and least expensive to implement. These studies have sometimes examined

Table 1.
Distribution of Studies by Study Design, Quality of HRQOL Instrument, and Year

	1981-1987		1988-1994		Total	
	No.	(%)	No.	(%)	No.	(%)
Prospective and controlled						
Quality I	0	(0)	2	(7.1)	2	(4.3)
Quality II	1	(5.3)	1	(3.6)	2	(4.3)
Quality III	1	(5.3)	0	(0)	1	(2.1)
Prospective and uncontrolled						
Quality I	0	(0)	1	(3.6)	1	(2.1)
Quality II	0	(0)	1	(3.6)	1	(2.1)
Quality III	1	(5.3)	0	(0)	1	(2.1)
Retrospective and controlled						
Quality I	1	(5.3)	1	(3.6)	2	(4.3)
Quality II	1	(5.3)	5	(17.9)	6	(12.8)
Quality III	0	(0)	1	(3.6)	1	(2.1)
Retrospective and uncontrolled						
Quality I	1	(5.3)	1	(3.6)	2	(4.3)
Quality II	2	(10.5)	6	(21.4)	8	(17.0)
Quality III	11	(57.9)	8	(29.6)	19	(41.3)
Total	19	(100)	27	(100)	46	(100)

McLeod *et al.* (1991)²⁶ has retrospective and prospective wings that are counted separately.

series of patients whose interventions predate the availability of more rigorous tests of HRQOL, and so they have provided information not otherwise available about possible effects of clinical interventions on the quality of life in IBD. Nonetheless, the information they provide is quite limited. The several studies that have only surveyed retrospective satisfaction with a single surgical procedure are prone to considerable bias, because one is likely to favor a major decision previously made. Significantly, these studies each support the value of the intervention studied. Several studies²⁷⁻³¹ show that groups that differ in HRQOL by a quantitative measure report similarly high levels of global satisfaction. This should warn against accepting patient global satisfaction as a usefully discriminative test.

The improvement in study design in recent years has consisted for the most part of an increase in the number of standardized self-designed questionnaires, most of which remain poorly validated. This is an advance on qualitative reports of outcome, but it results, unfortunately, in a great variety of studies that are difficult to compare with one another. Unpublished questionnaires are hazardous to interpret. Furthermore, the ability to statistically test HRQOL results for significance is of little value if the instrument is of undetermined validity.

There are well-validated and reliable instruments available for evaluative use in clinical trials in IBD: the

IBDQ¹⁶ and the RFIPC.²² Also available are discriminative instruments for cross-sectional studies: Farmer's¹⁵ questionnaire and DQO.²⁵ The next phase in HRQOL measurement in IBD populations should include the adoption of standard instruments and a greater emphasis on randomized, controlled prospective trials.

Clinical Questions

Surgical Approaches to UC. Four surgical approaches to the treatment or cure of ulcerative colitis have been advocated over the past 15 years: proctocolectomy with conventional ileostomy, proctocolectomy with Kock continent ileostomy (Kock pouch), colectomy with ileorectal anastomosis, and proctocolectomy with ileoanal anastomosis (pelvic pouch). Advocates of the pelvic pouch differ on technical details such as optimal size of the pouch, level of anastomosis, the way in which the anastomosis is performed, whether ileostomy is performed, and the way in which functional outcomes are measured. Ileorectal anastomosis is generally discouraged because of continuing inflammatory disease and risk of cancer in the unresected rectum. Because mortality is low but bowel function is inevitably impaired to some degree after surgery and because other forms of morbidity, especially pouchitis and repeat surgery, are common,³² HRQOL could be a useful outcome mea-

sure to evaluate these procedures. This clinical controversy has thus generated a large number of clinical trials measuring HRQOL.

HRQOL has been measured preoperatively and postoperatively to demonstrate the value of surgery in ulcerative colitis.²⁶ This study measured HRQOL with the TTOT and DQO in 20 patients who had various types of surgery. A significant and robust improvement in HRQOL by both measures was found approximately one year after surgery. No comparison between surgical methods is possible with this small sample. Of their two disease-specific measures, the DQO was found to correlate inversely with disease severity preoperatively, whereas the TTOT did not. This is consistent with the finding of Mitchell *et al.*¹⁴ that the TTOT lacks discriminative power but does not diminish the finding of a robust improvement after surgery.

Comparison of HRQOL between types of surgery has been attempted. A controlled and randomized prospective study³³ has been done to address some technical disputes but did not include quality-of-life measures. A prospective but not randomized comparison of triplicated *vs.* duplicated pelvic ileal reservoirs in 31 patients receiving the pelvic pouch³⁴ did address HRQOL. Unfortunately the measures of HRQOL were restricted to patients' impressions that their social lives were restricted by bowel function (triplicated, 3/12; duplicated, 11/19), successful return to work (10/10, 10/15), and what can be concluded from a detailed survey of bowel history. Trends toward the superiority of the triplicated pouch were generally not statistically significant, although the sample is small. The only other prospective study addressing HRQOL³⁵ is an uncontrolled examination of 100 patients with proctocolectomy and ileoanal anastomosis (J-pouch or S-pouch). This study found relatively high levels of pouchitis (30 percent at 2 years), frequent evacuation (10 percent ≥ 7 /day), night evacuation (40 percent more than once weekly), and mucous soiling (30 percent) at one year. Despite this, patient satisfaction was "generally excellent." Sexual function was found to be impaired postoperatively much more frequently in women (27 percent improved/22 percent deteriorated) than men (30 percent/4 percent).

Retrospective studies allow more limited confidence in conclusions. Nonetheless, controlled studies with more rigorous HRQOL instruments provide the best information available about the effects of surgery on quality of life in ulcerative colitis.

Quality of life after pelvic pouch has been com-

pared in retrospective series with Brooke ileostomy,^{26, 27, 36} Kock pouch,^{26, 27} cholecystectomy,³⁷ and medical treatment.²⁸ Results of the procedure have also been compared between patients with ulcerative colitis and familial adenomatous polyposis.^{29, 38}

Among these studies, the retrospective wing of the article by McLeod *et al.*²⁶ has the most rigorously validated instruments for measuring HRQOL. By using the TTOT and DQO in 93 patients one year following surgery, they found no significant difference between patients with conventional ileostomy, Kock pouch, and pelvic pouch. Each group had scores similar to the postoperative group in the prospective wing of the study and comparably low Beck, SIP, and global assessment scores.

On the other hand, Köhler and Pemberton and colleagues^{27, 36, 37} at the Mayo Clinic published a series of articles that suggest that the pelvic pouch does result in HRQOL superior to that following conventional and continent ileostomy. Pemberton *et al.*³⁶ compared 298 patients with pelvic pouch (89 percent UC/11 percent FAP) with 406 patients with Brooke ileostomy (95 percent UC). Pelvic pouch patients showed significantly less restriction in all seven domains (social activity, sports, housework, recreation, family relations, sexual activity, travel) such that decreased restriction predicted surgical group. In a subsequent article²⁷ these results were combined with a retrospective survey of activity in 313 patients with Kock (continent) ostomy, ostensibly to determine if stoma or continence is the key factor in determining HRQOL. The pelvic pouch group was superior to Kock pouch on two activity domains (sports and sex), whereas the Kock ileostomy was superior to Brooke ileostomy with respect to sports and sex and inferior with respect to travel. At face value these results suggest that quality of life following pelvic pouch is clearly superior to conventional ileostomy, with more ambiguous differentiation between the Kock pouch and either procedure (five of seven activity areas being undifferentiated between the continent ileostomy and pelvic pouch).

The superiority of pelvic pouch to Kock pouch with respect to sexual activity is not broken down by gender, an important consideration in light of the data suggesting greater difficulty with sexual activity in women than in men following pelvic pouch^{28, 35} and inconsistent reports of sexual difficulties in both genders with ostomy.³⁹

There is conflict between the Mayo Clinic conclusions and those of the Toronto group. Possible expla-

nations include insensitive HRQOL measures in the Toronto study (*i.e.*, a false-negative test missing an actual difference between groups), an invalid HRQOL instrument in the Mayo Clinic series (*i.e.*, a false-positive result finding a spurious difference), differing patient populations, differing surgical technique or expertise, or instruments that sample different aspects of HRQOL that vary independently. As has previously been noted, there is in fact some concern that the TTOT may lack discriminative power. The DQO was able to discriminate disease severity in the prospective study and correlates well to other measures relevant to HRQOL. Discriminative power, however, is not necessarily equivalent to evaluative responsiveness. In addition, the DQO concerns problem areas identified by individual patients, which increases the presumed relevance of the identified "objectives" but may well lead it to address different concerns than those chosen in the Mayo Clinic questionnaire. It is unfortunate that in their impressive series of studies Köhler and Pemberton and colleagues^{27, 36, 37} have not yet demonstrated that their instrument is a valid and reliable measure of HRQOL, which does not allow us to rule out the possibility of a false-positive discrimination of procedures. This concern may be amplified by the remarkable finding in the third study in this series³⁷ that one to eight years postoperatively pelvic pouch patients were not significantly different than a matched cohort of cholecystectomy patients (matched except for an older postcholecystectomy population), which suggests better than expected function in the UC group, worse than expected function in the cholecystectomy group (perhaps based in part on more advanced age), or problems with validity or discriminative ability in their HRQOL instrument. Clearly, this is a question that could benefit from further investigation, using a demonstrably responsive and valid instrument as the outcome measure.

Sagar *et al.*²⁸ compared surgical with medical management of UC, a comparison that is important to many patients. This study also differentiates surgical techniques (pelvic pouch above or at the dentate line; duplicate, triplicate, or quadruplicate pouch design), but few differences are reported here, possibly because of fewer patients in these cells. Conclusions regarding the differences found are limited by the inevitable lack of randomization. Among the significant findings, the surgical group showed more nocturnal bowel movements than the medical group and greater loss of work time over the one year that

included their surgery but less urgency to defecate and less social restriction. Differences in sexual function were greatly different between genders, with men in the medical group and women in the surgical group reporting more difficulty. Both groups reported high levels of satisfaction and similar depression and anxiety scores on the Hospital Anxiety and Depression questionnaire.

Of the two articles that have addressed differences in outcome between patients with ulcerative colitis and those with familial adenomatous polyposis (FAP) receiving the same operation,^{29, 38} Fujita *et al.*²⁹ found similar bowel function in the two groups but greater satisfaction in the UC group. Among symptomatic patients, satisfaction was correlated with a measure of "ego strength." Their HRQOL instrument was a 95-item, multidomain mail questionnaire. Inter-item reliability for question pairs dealing with similar content was calculated but not the reliability of the test to generate similar scores in stable patients over time. No validity data were provided except for the face validity concluded from its development by a stoma therapist and revision based on patient input. Another group³⁸ also found similar bowel function in these two populations but did not find differences in HRQOL either. This study used a four-item, four-domain, self-designed questionnaire.

Uncontrolled, retrospective studies form the remainder of the literature in this area. Methodologic problems include using the patient's memory of the period with a temporary ileostomy as a "control" or comparison against which to measure the HRQOL with an ileoanal or ileorectal anastomosis.^{30, 40-44} It is likely that psychologic adaptation to a temporary ileostomy is different than to an ileostomy that is expected to be permanent. Similarly the question of whether one is satisfied with the surgery as a sole measure of postoperative HRQOL^{45, 46} or believes that surgery has led to improved general quality of life^{40, 43} is very problematic. Any retrospective evaluation of an essentially irreversible decision is open to a strong bias in the direction of validating the decision. Furthermore the studies previously reviewed, which include both multidomain HRQOL instruments and global satisfaction reports, have shown poor discriminative power in the satisfaction rating.

Several uncontrolled, retrospective studies have attempted more comprehensive evaluation of HRQOL. Self-designed questionnaires offer more detailed information about the components of HRQOL but lack testing of validity, reliability, and responsiveness.

None of the studies reported in the remainder of this section have used validated HRQOL instruments.

Anseline⁴² designed a questionnaire based on Spitzer and colleague's¹³ quality-of-life index. He used several items in four domains to evaluate 18 UC patients with proctocolectomy and ileoanal anastomosis. He found postoperative improvement in ability to work, travel, diet, sporting activity, and social life in those patients with preoperative problems in those areas. However, he also found postoperative deterioration in ability to travel in three of seven patients not previously restricted and a more complicated mixture of improvement and deterioration in sexual function. Sexual activity was reduced in 14 of 18 patients preoperatively. With temporary ileostomy no one reported improvement and 6 of 18 reported abstinence. Following pelvic pouch six remained abstinent, and one more was "severely restricted" because of pain, loss of interest and "divorce precipitated by the presence of the ileostomy." Sexual function improved in some patients previously restricted and deteriorated significantly in some previously unaffected. The data are unfortunately not broken down by gender. No statistical analysis of this data is supplied.

Skarsgard *et al.*³⁰ reported that 75 UC and FAP patients preferred pelvic pouch to temporary ileostomy in several areas of social interaction and self-confidence. Patients were apparently asked to imagine life with a permanent ileostomy based on their experience with a temporary one, although details of the questionnaire are not published. Greater percentages of patients suggested no preference regarding the ability to do their job (47 percent), interest in sexual activity (42 percent), and partner's sexual interest (72 percent). Eleven percent could not return to their previous employment because of the pouch, 16 percent required "excessive sick leave" postoperatively, and 30 percent of women reported postoperative sexual dysfunction. Despite these complications, 92 percent reported they would choose the pelvic pouch again.

Perrault *et al.*⁴¹ followed up 66 patients six months to several years after proctocolectomy and ileoanal anastomosis. They used a J-pouch in 22 patients and straight ileoanal pull-through in 44 patients. Patients had a median age of 16 years, but no differences were detected between child and adult respondents. Their four-domain questionnaire addressed HRQOL only through general satisfaction (85 percent) and missing time at work or school (in both groups 23–25 percent missed one to three days of school). Using these

relatively insensitive probes, no difference in HRQOL could be detected between these procedures.

Wexner and colleagues⁴⁴ reported continued improvement more than five years postoperatively in 114 patients with pelvic pouch. Work and school attendance was similar (approximately 90 percent) before and after surgery. With respect to sexual activity 86 percent were active preoperatively. Postoperatively 50 percent reported no change, 23 percent improved, and 27 percent deteriorated (although most of these report mild deterioration and do not attribute it to the pouch). Sexual activity data are not broken down by gender. Sixty-three percent felt that improvements in function continued for several years.

Although there are no definitive studies that recommend any one treatment in ulcerative colitis, one can have more confidence in recommendations that come from several sources using different methodology. Although the retrospective comparison studies involving the pelvic pouch have been inconsistent, the pelvic pouch operation has not been found inferior to ileostomy (conventional or continent) in any series, and some have suggested superiority. Further evidence from prospective (unrandomized for ethical reasons) comparison of pelvic pouch and ileostomy with validated HRQOL evaluative instruments is needed.

With more reliable studies one may be able to look to particular domains of HRQOL to guide treatment decisions. One could suggest, for example, that among UC patients coming to surgery pelvic pouch would be the superior choice for those particularly concerned preoperatively with body image problems if these are predictably exacerbated by ileostomy. Women may be less likely than men to find the pelvic pouch satisfactory because of the disproportionate incidence of sexual problems caused by dyspareunia (women might even choose medical more than surgical treatment if the risk of increased sexual difficulties is considered unacceptable). The desirability of easy travel or of participation in sports could guide the choice of procedure. HRQOL measures may allow more flexible and rationale planning of treatment in this case.

Fujita and colleagues²⁹ suggested that the greater satisfaction among UC patients with similar functional outcomes to FAP patients after pelvic pouch is because of higher "ego strength." Unfortunately, psychoanalytic concepts such as ego strength have been even more resistant to quantification than quality of life.⁴⁷ Beyond the individual case study, one must

interpret such measures with caution. Greater satisfaction suggests a smaller discrepancy between expectation and result, which may be caused by greater adaptability to imposed reality (roughly equivalent to ego strength), to lower expectations, or to an attitude of accepting resignation (the "given up" attitude of depression). Nonetheless there is a discrepancy between functional results and subjective satisfaction, which at least emphasizes the value of measuring HRQOL and leads to further study. The only other study comparing these groups³⁸ found similar satisfaction regardless of diagnosis but is limited by a four-item HRQOL instrument that may have lacked the required discriminative power.

Quality of Life Following Ileostomy. Only one study of HRQOL following ileostomy has used a validated HRQOL instrument. Drossman *et al.*⁴⁸ surveyed 990 UC and CD patients by mail through the National Foundation for Ileitis and Colitis using the SIP, the SCL-90, and (one presumes from the other publications emerging from this survey) the RFIPC. They examined four study groups: UC colectomy, UC non-colectomy, CD ostomy, and CD nonostomy. HRQOL was better in the postoperative groups. Physical and psychosocial function on SIP were similar across groups. CD ostomates had more trouble at work. UC colectomy had less distress on SCL-90. Both surgical groups had fewer disease-related concerns (including concern about having an ostomy and about body image) than nonsurgical groups. The retrospective design and grouping together of several surgical approaches in the UC colectomy group limit any conclusions beyond the better function and adjustment of the UC surgical group and a more limited but generally good adjustment of the CD ostomy group.

Patients who have been converted from conventional Brooke ileostomy to a Kock pouch provide a group who can be "used as their own controls" in comparing one arrangement to the other. Gerber *et al.*⁴⁹ interviewed 80 such patients about changes in lifestyle and found extremely positive results, with 87 to 97 percent reporting improvements in social activities, athletics, and appearance (ability to wear street and beach attire) and 76 percent reporting an improvement in sexual relations. Sjødahl and colleagues⁵⁰ asked patients to compare their function with the Kock pouch with function before colostomy and with that of Brooke ileostomy. Compared with Brooke ileostomy improvements were reported for sport (95 percent), travel (78 percent), employment and education (76 percent), and family life (62 per-

cent). For sexual activity, 74 percent noted improvement since conversion to Kock pouch, but only 64 percent noted improvement since before colectomy. Both studies found that satisfaction was equivalent among patients who required a revision to restore continence and those who did not. It should be noted that the degree of superiority of the continent ileostomy reported here is much greater than was found in the controlled comparison of Köhler *et al.*,²⁷ and there is inconsistency among the domains (travel being better among conventional ileostomy patients in Köhler and colleague's study).

Five studies used multidomain questionnaires or structured interview to assess quality of life following conventional ileostomy. McLeod and colleagues⁵¹ found that most patients considered their physical health improved postoperatively. Work status was constant in 64 percent, with 22 percent working less. Rates of medical disability were constant preoperatively and postoperatively. There were only insignificant trends toward better quality of life with better ostomy function in each domain. Regarding restrictions of lifestyle, results were similar in each category (except clothing selection), with 33 to 46 percent noting no change, 37 to 44 percent noting improvement, and 15 to 21 percent noting worsening. Only three percent indicated regretting the surgery. Martinsson *et al.*⁵² found some degree of disability in most stoma patients but few complaints of significant restrictions. Awad *et al.*³¹ also presents modest support for ileostomy. Improvements in housework and gardening were noted in 15 percent (51 percent unchanged), holidays 21 percent (42 percent unchanged), and sex life 7 percent (42 percent unchanged, 47 percent deteriorated). No statistical analysis is provided. Küchenhoff and colleagues⁵³ compared four groups: cancer and IBD patients with bowel resection and end-to-end anastomosis or ostomy. Stoma groups (cancer and colostomy, UC and ileostomy) showed greater preoperative and postoperative depression and greater social inhibition. Kennedy⁵⁴ examined 39 UC patients after ileostomy with a comprehensive battery of tests. The HRQOL items included reports of sexual activity (psychologic difficulties in 23 percent of men and 93 percent of women) and social activity (restricted in 21 percent).

Studies of the effects of ileostomy on HRQOL provide only modest support for the procedure. There is no strong evidence of an enhanced ability to work, and there is a suggestion of deterioration in sexual function postostomy. Each of these trials lacks comparability with the pelvic pouch studies described above because of the inclusion of patients with

Crohn's disease, which presumably weights results toward poorer long-term function. McLeod *et al.*⁵¹ found only insignificant trends toward greater HRQOL with better ostomy function, which suggests a lack of discriminative power in the instrument used.

The problems with studies of conversions that use an earlier Brooke ileostomy as a retrospective "control" are similar to those described above for comparing pelvic pouch to temporary ileostomy. A bias toward favoring the more recent procedure can be expected.

Surgical Approaches to Crohn's Disease. Evaluation of surgery in Crohn's disease may also benefit from measurement of HRQOL. As opposed to ulcerative colitis, for which surgery is considered curative, patients with Crohn's disease expect a course of exacerbation and remission that is likely to continue postoperatively. The cost and benefit of any treatment in terms of quality of life should be a key factor in decision-making for patients and professionals.

Early studies comparing radical resection of bowel with wide margins of excision to more conservative procedures⁵⁵⁻⁵⁹ use semiquantitative measures of HRQOL, assigning patients to one of three categories based on the clinician's impression of state of health and ability to work and have normal leisure-time activities. The unresponsiveness of this measure is suggested by the results of Lindhagen *et al.*⁵⁶ who report 85 percent of patients in QOL I (the highest possible rating) following initial bowel resection, despite recurrence in 37 percent and up to three resections of bowel (five of seven patients with recurrence after their third resection still find themselves in QOL I).

Meyers⁶⁰ interviewed 51 patients with CD five to ten years after elective surgery. He used a 39-item questionnaire addressing five areas of psychosocial functioning (personal relations, school and job performance, recreation, sexual function, and body image). Patients were asked to retrospectively assess these areas six months preoperatively, one year postoperatively, and at the time of interview. Overall, patients reported significantly less severe symptoms at interview than preoperatively as expected. Differences between these measures are large (100 percent symptomatic at surgery, approximately one-half that at follow-up), which may support an effect of the intervention itself. Ninety-two percent felt the surgery had been helpful. All components of HRQOL surveyed improved, less so in the subgroup with an ileostomy (14), who had poor body image and poorer

overall assessment. A subgroup with recurrent illness (27) had less overall improvement after surgery, which was significant only for the overall measure of dysfunction and for symptoms (which demonstrates a tautology—that those with recurrence experience more symptoms).

Cooper and colleagues⁶¹ reported 42 patients with CD sparing the rectum who underwent colectomy and ileorectal anastomosis. They experienced seven percent operative mortality because of anastomotic breakdown. Of the 14 people questioned selected from the 22 people who continued to have a functioning anastomosis 1 to 14 years postoperatively, 85 percent reported unrestricted social activity and regular employment on the five-domain HRQOL instrument used.

Halme⁶² followed 98 patients treated for CD for a mean period of ten years. By using a three-grade QOL designation (good, reduced, poor), he found a substantial difference between patients with large bowel involvement treated with (procto)colectomy and ileostomy and all other treatment groups. The former showed 90 percent good QOL, whereas other groups (ileocecal resection, colectomy with ileorectostomy, colostomy with proctectomy) had approximately one-third of patients in the reduced group. Considering the apparent lack of responsiveness of the three-category method of measuring QOL in earlier studies, this is a striking finding. Ability to work was greatest in the nonsurgical group and the ileostomy group.

The study by Küchenhoff *et al.*⁵³ noted above used multivariate analysis to compare stoma with non-stoma groups and cancer with ileitis/colitis groups but not CD resection with cancer resection, the only comparison that would be readily interpretable in this discussion. Similarly Kennedy⁵⁴ included 12 CD patients in his study of the effects of ileostomy, but the group is mixed with the UC patients in the analysis.

In a more easily quantified variation on the retrospective question "Do you prefer the surgery you had?" Scott and Hughes⁶³ asked 80 CD patients after ileocolonic resection if they would have preferred their surgery earlier, later, or at the same time it was done. Of the 70 patients who replied, 74 percent indicated they would have preferred it sooner, and none indicated they would have preferred it later. When quantifying how many months earlier they would prefer, these patients gave a median time of 12 months. Reasons for preferring earlier operation included severity of symptoms (97 percent), ability to eat normally (86 percent), feeling after resection (62

percent), and abolishing drugs (43 percent). Although this innovative measure lacks the information available in a multidomain HRQOL instrument, it is a significant step forward from the retrospective questioning of satisfaction in earlier studies.

All studies addressing surgical treatment of CD are uncontrolled, and none used a validated HRQOL instrument. The dangers of placing too much confidence in uncontrolled studies are even greater in this case. Any chronic illness that waxes and wanes will yield large differences in measurements of severity, which depend largely on the time of sampling. Thus a comparison of current function to preoperative function will predictably show better function currently because the time when a person chooses palliative surgery presumably marks a nadir in satisfaction and function, as opposed to the more arbitrarily determined date of follow-up when some proportion of patients will be in remission by chance.

Assessment of the surgical management of Crohn's disease is greatly hampered by the use of physician-rated scales that assign patients to one of three or four levels of HRQOL. Physician ratings cannot truly measure HRQOL, which is by definition the patient's subjective opinion. The insensitivity of these scales further reduces confidence in the conclusions of the studies using them that have generally supported "radical" resection of bowel in CD.

Meyers⁶⁰ early contribution used a much better HRQOL questionnaire, but as noted the lack of controls does not allow us to distinguish natural remission from treatment effect among the improved population. Comparison with cross-sectional survey data in the Crohn's population could clarify this, but results from other studies are inconsistent. Farmer and Michener⁶⁴ showed similar quality of life to Meyers's surgical group (approximately 50 percent reporting some impairment) in both operated and nonoperated patients, whereas other surveys suggested a much lower incidence of impairment in an unselected group of Crohn's patients.⁶⁵⁻⁶⁷ No studies show levels of HRQOL significantly worse than Meyers's patients, which is what would be required to have confidence that his results are better than a chance finding. In general, the available studies of surgical interventions for Crohn's disease provide quite limited support for the value of surgery and no convincing argument for the superiority of any particular surgical approach.

Medical Treatment of Inflammatory Bowel Disease. Studies of medical treatments of inflammatory bowel diseases that use HRQOL outcome measures have

been far less numerous than surgical studies but have used relatively high standards of study design and HRQOL instruments. Irvine *et al.*²¹ used the IBDQ in a prospective, double-blind, placebo-controlled trial of cyclosporine in Crohn's disease patients. The study further validated the usefulness of this instrument in the CD population. The IBDQ discriminated between patients who required surgery and those who did not. Furthermore IBDQ results discriminated the group going to surgery, even when disease activity scores could not.

Greenberg and colleagues⁶⁸ used the IBDQ in a multicenter randomized, prospective, placebo-controlled study of budesonide (Entocort CIR[®], Astra Draco, Lund, Sweden) for active Crohn's disease. Patients were randomly assigned to one of three dosage schedules. Similar improvements in clinical activity were found with 9 mg daily and 15 mg daily (each better than 3 mg daily and much better than placebo). IBDQ scores show greater improvement at 9 mg daily than 15 mg daily. In this study measurement of HRQOL helps to determine optimum dosage.

Robinson *et al.*⁶⁹ used a multidomain visual analog scale HRQOL instrument modified from Somerville *et al.*⁷⁰ with general and disease-specific domains as an outcome measure in a prospective, placebo-controlled trial of mesalamine (Pentasa[®], Marion Merrell Dow, Kansas City, KS) capsules in UC. "The convergent validity, reliability and responsiveness of this instrument were documented," but unfortunately no further information is available in this publication. The significant and dose-dependent improvements in several objective measurements of bowel function and all seven domains of HRQOL provide strong evidence for the usefulness of this treatment.

Somerville *et al.*⁷⁰ used a visual analog scale measure of domains of general HRQOL to compare hydrocortisone acetate foam to prednisolone 21-phosphate enemas in 46 patients with relapses of UC. Their findings emphasize the use of HRQOL measures. Both treatments resulted in equivalent symptomatic improvement and similar improvement in sigmoidoscopic abnormalities, but the hydrocortisone foam was significantly superior to the prednisolone enemas on several domains of HRQOL (sexual function, occupational activity, work/social activity, and routine outdoor activity). Unfortunately validation of their instrument is not included in the study.

Nyman and colleagues⁷¹ prospectively examined patients treated with azathioprine or 6-mercaptopurine for immunosuppression for five years. The study

is uncontrolled, although outcome data are compared with surgical studies carried on at the same center. The quality-of-life measure is the I-III scale criticized above. There would appear to be no problem with the responsiveness of the instrument in this case because the shift from baseline (I = 14/II = 7/III = 21) to follow-up (32/6/3) is very significant. The potential bias relating to timing (assuming all patients were quite ill at the start of treatment, with some quiescent by chance at the time of follow-up) can only be addressed with a control group.

One study is unique in the population it studies. Gavalier *et al.*⁷² examined 23 patients with sclerosing cholangitis in association with ulcerative colitis after liver transplantation and immunosuppression with cyclosporine and prednisone. They used a published, self-designed questionnaire that included only 2 of 29 questions relating to restriction of function and activity because of symptoms. Illness in general was found to be improved postoperatively. Eleven of 14 patients with preoperative activity restriction showed improvement. Any impairment of function postoperatively in the six patients with quiescent disease preoperatively is not reported.

Medical studies demonstrate the value of measuring HRQOL in clinical trials. Somerville *et al.*⁷⁰ found a very significant difference between two methods of conservative treatment of IBD that would not have been apparent using clinical improvement as an endpoint. The Robinson and colleagues⁶⁹ study demonstrates convincingly the usefulness of mesalamine. The multicenter Canadian Crohn's Relapse Prevention Study reported by Irvine and colleagues²¹ is exemplary in its careful use of HRQOL measures in a prospective, controlled study.

Home Parenteral Nutrition. The effects of home parenteral nutrition (HPN) on HRQOL were examined in four studies. Each was an uncontrolled, retrospective study, and in three the HRQOL measures were unsophisticated. The exception is the quality-of-life-adjusted survival analysis by Detsky and colleagues,²⁵ which used the TTOT, DQO, and an index score of the global desirability of the states "before HPN" and "on HPN." Large differences were found between the before HPN and on HPN conditions, and no significant differences were found between measurement techniques. HPN was compared with TPN indirectly through projections extrapolated from measured data. The authors acknowledge the biases of the retrospective technique and the hazards of projected data but, through the use of conservative as-

sumptions, provide tentative evidence for the HRQOL advantages of HPN.

Mughal and Irving⁷³ reviewed 200 patients registered as receiving HPN in the United Kingdom and Ireland. Ninety of these patients had Crohn's disease, the others had non-IBD indications. Quality of life was "crudely assessed" with a four-grade classification based on ability to work and look after home and family. Of the Crohn's patients 52 were in Grade I (best HRQOL), 18 in Grade II, 10 in Grade III, and 1 in Grade IV. This was a better result than patients with malignancy, scleroderma, mesenteric vascular disease, or radiation enteritis (no statistical analysis).

Galandiuk *et al.*⁷⁴ used a three-domain test of HRQOL without validation to retrospectively compare preillness, HPN, and post-HPN HRQOL in 41 patients with Crohn's disease. Quality of life, social activity, and psychologic well-being (all undefined) were graded 1 (no change from preillness), 2 (50 percent reduction from preillness), or 3 (severe disability). They reported on a summed composite score. Information was gathered by telephone interview or chart review in deceased patients. The composite score, but none of the individual domain scores, was significantly worse in the pre-HPN group than the HPN and post-HPN group.

A third study⁷⁵ recently assessed HPN in 18 patients with chronic intestinal failure and 7 with CD. QOL was rated in four grades based on Mughal and Irving's⁷³ criteria. Results during HPN were I = 8, II = 3, III = 4, IV = 3. All of the Crohn's patients were in Grades I or II. Two Crohn's patients stopped HPN because of clinical improvement.

Home parenteral nutrition is an intervention whose justification depends on a reasonable assumption that it promotes both increased quality of life and health cost savings. Detsky and colleague's²⁵ analysis provides support for these assumptions. Other studies⁷³⁻⁷⁵ reviewed do not refute the benefits of HPN but are too methodologically weak to provide strong support and use HRQOL instruments, which do not allow analysis of the particular benefits of HPN.

Psychosocial Interventions. Psychosocial interventions have been recognized as an important component of the treatment of the whole person who suffers from IBD.⁷⁶ Psychotherapy has been advocated since the era of psychoanalytic hypotheses of psychosomatic causation, but studies in the modern era of psychotherapy outcome research have had only very modest success in improving disease activity.⁷⁷ To

date there are no studies of psychosocial interventions that have included HRQOL as an outcome measure.

SUMMARY

Objective quality-of-life measurement is now an established part of the IBD literature. There has been a significant move toward multidomain evaluative HRQOL instruments in clinical trials for the past five to ten years. There remains a lack of consistency in the particular instruments chosen with most trials using self-designed, unvalidated questionnaires. Resulting contradictions and controversies in the literature should be addressed through more uniform methods of testing HRQOL and more rigorously designed clinical trials. Although there is no "gold standard" of HRQOL, validated and responsive evaluative instruments (particularly the RFIPC and the IBDQ) are now available and should be used in future studies. Better instruments may yet be designed, but at this point the onus is on investigators developing them to demonstrate their comparability or superiority to current standards.

REFERENCES

1. Drummond MF. Resource allocation decisions in health care: a role for quality of life assessments? *J Chron Dis* 1987;40:605-16.
2. Stewart AL, Greenfield S, Hays RD, *et al.* Functional status and well-being of patients with chronic conditions: results from the medical outcomes study. *JAMA* 1989;262:907-13.
3. Guyatt GH, Bombardier C, Tugwell PX. Measuring disease-specific quality of life in clinical trials. *Can Med Assoc J* 1986;134:889-95.
4. Gazzard BG. The quality of life in Crohn's disease. *Gut* 1987;28:378-81.
5. Künsebeck HW, Körber J, Freyberger H. Quality of life in patients with inflammatory bowel disease. *Psychother Psychosom* 1990;54:110-6.
6. Garrett JW, Drossman DA. Health status in inflammatory bowel disease: biological and behavioral considerations. *Gastroenterology* 1990;99:90-6.
7. O'Young J, McPeck B. Quality of life variables in surgical trials. *J Chron Dis* 1987;40:513-22.
8. Drossman D. Psychosocial considerations in gastroenterology. In: Sleisenger MH, Fordtran JS, eds. *Gastrointestinal disease: pathophysiology, diagnosis and management*. 5th ed. Philadelphia: WB Saunders, 1993.
9. Spilker B. *Quality of life assessments in clinical trials*. New York: Raven Press, 1990.
10. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622-9.
11. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status measure. *Med Care* 1981;19:787-805.
12. Ware JE, Brook RH, Davies-Avery A. *Conceptualization and measurement of health for adults in the health insurance study: model of health and methodology*. Vol. 1. Santa Monica, CA: RAND Corporation, 1980.
13. Spitzer WO, Dobson AJ, Hall J, *et al.* Measuring quality of life of cancer patients: a concise QL-Index for use by physicians. *J Chron Dis* 1981;34:585-97.
14. Mitchell A, Guyatt G, Singer J, *et al.* Quality of life in patients with inflammatory bowel disease. *J Clin Gastroenterol* 1988;10:306-10.
15. Farmer RG, Easley KA, Farmer JM. Quality of life assessment by patients with inflammatory bowel disease. *Cleve Clin J Med* 1992;59:35-42.
16. Guyatt G, Mitchell A, Irvine EJ, *et al.* A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology* 1989;96:804-10.
17. Love JR, Irvine EJ, Fedorak RN. Quality of life in inflammatory bowel disease. *J Clin Gastroenterol* 1992;14:15-9.
18. Drossman DA, Lesserman J, Mitchell M, Li Z, Zagami EA, Patrick DL. Health status and health care use in persons with inflammatory bowel disease: a national sample. *Dig Dis Sci* 1991;36:1746-55.
19. Ware JE. Standards for validating health measures: definition and content. *J Chron Dis* 1987;40:473-80.
20. Spilker B. *Guide to clinical trials*. New York: Raven Press, 1991.
21. Irvine EJ, Feagan B, Rochon J, *et al.* for the Canadian Crohn's Relapse Prevention Trial. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. *Gastroenterology* 1994;106:287-96.
22. Drossman DA, Lesserman J, Li Z, Mitchell CM, Zagami EA, Patrick DL. The rating form of IBD patient concerns: a new measure of health status. *Psychosom Med* 1991;53:701-12.
23. Drossman DA, Patrick DL, Mitchell CM, Zagami EA, Appelbaum MI. Health-related quality of life in inflammatory bowel disease: functional status and patient worries and concerns. *Dig Dis Sci* 1989;34:1379-86.
24. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Services Research* 1972;Summer:118-33.
25. Detsky AS, McLaughlin JR, Abrams HB, *et al.* Quality of life of patients on long-term total parenteral nutrition at home. *J Gen Intern Med* 1986;1:26-33.
26. McLeod RS, Churchill DN, Lock AM, Vanderburgh S, Cohen Z. Quality of life of patients with ulcerative colitis preoperatively and postoperatively. *Gastroenterology* 1991;101:1307-13.

27. Köhler LW, Pemberton JH, Zinsmeister AR, Kelly KA. Quality of life after proctocolectomy: a comparison of Brooke ileostomy, Kock pouch and ileal pouch-anal anastomosis. *Gastroenterology* 1991;101:679–84.
28. Sagar PM, Lewis W, Holdsworth PJ, Johnston D, Mitchell C, MacFie J. Quality of life after restorative proctocolectomy with a pelvic ileal reservoir compares favorably with that of patients with medically treated colitis. *Dis Colon Rectum* 1993;36:584–92.
29. Fujita S, Kusonoki M, Shoji Y, Utsunomiya J. Quality of life after total proctocolectomy and ileal J-pouch-anal anastomosis. *Dis Colon Rectum* 1992;35:1030–9.
30. Skarsgard ED, Atkinson KG, Bell GA, Pezim ME, Seal AM, Sharp FR. Function and quality of life results after ileal pouch surgery for chronic ulcerative colitis and familial polyposis. *Am J Surg* 1989;157:467–71.
31. Awad RW, el-Gohary TM, Skilton JS, Elder JB. Life quality and psychological morbidity with an ileostomy. *Br J Surg* 1993;80:252–3.
32. Keighley MR, Kmiot W. Surgical options in ulcerative colitis: the role of ileo-anal anastomosis. *Aust N Z J Surg* 1990;60:835–48.
33. Keighley MR, Yoshioka K, Kmiot W. Prospective randomized trial to compare the stapled double lumen pouch and the sutured quadruple pouch for restorative proctocolectomy. *Br J Surg* 1988;75:1008–11.
34. Nasmyth DG, Williams NS, Johnston D. Comparison of the function of triplicated and duplicated reservoirs after mucosal proctectomy and ileo-anal anastomosis for ulcerative colitis and adenomatous polyposis. *Br J Surg* 1986;73:361–6.
35. Öresland T, Fasth S, Norgren S, Hultén L. The clinical and functional outcome after restorative proctocolectomy: a prospective study in 100 patients. *Int J Colorectal Dis* 1989;4:50–6.
36. Pemberton JH, Phillips SF, Ready RR, Zinsmeister AR, Beahrs OH. Quality of life after Brooke ileostomy and ileal pouch-anal anastomosis: comparison of performance status. *Ann Surg* 1989;209:620–8.
37. Köhler LW, Pemberton JH, Hodge DO, Zinsmeister AR, Kelly KA. Long-term functional results and quality of life after ileal pouch-anal anastomosis and cholecystectomy. *World J Surg* 1992;16:1126–32.
38. Tjandra JJ, Fazio VW, Church JM, Oakley JR, Milsom JW, Lavery IC. Similar functional results after restorative proctocolectomy in patients with familial adenomatous polyposis and mucosal ulcerative colitis. *Am J Surg* 1993;165:322–5.
39. Kennedy HJ, Lee EC, Claridge G, Truelove SC. The health of subjects living with a permanent ileostomy. *Q J Med* 1982;51:341–57.
40. Oakley JR, Jagelman DG, Fazio VW, *et al*. Complications and quality of life after ileorectal anastomosis for ulcerative colitis. *Am J Surg* 1985;149:23–30.
41. Perrault J, Telander RL, Zinsmeister AR, Kaufman B. The endorectal pull-through procedure in children and young adults: a follow-up study. *J Pediatr Gastroenterol Nutr* 1988;7:89–94.
42. Anselme PF. Quality of life after restorative proctocolectomy. *Aust N Z J Surg* 1990;60:683–8.
43. Nicholls RJ. Restorative proctocolectomy with ileal reservoir. *Ann Gastroenterol Hepatol (Paris)* 1986;22:193–8.
44. Wexner SD, Jensen L, Rothenberger DA, Wong DA, Goldberg SM. Long-term functional analysis of the ileo-anal reservoir. *Dis Colon Rectum* 1989;32:275–81.
45. Barnett WO. Current experiences with the continent intestinal reservoir. *Surg Gynecol Obstet* 1989;168:1–5.
46. Sasaki I, Funyama Y, Naito H, Toda M, Sato T. Long-term follow-up of surgical treatment for ulcerative colitis—with special reference to recurrence and the quality of life. *Jpn J Surg* 1986;16:22–8.
47. Glymour C. Freud, Kepler and the clinical evidence. In: Wollheim R, Hopkins J, eds. *Philosophical Essays on Freud*. Cambridge: Cambridge University Press, 1982.
48. Drossman DA, Mitchell CM, Appelbaum MI, Patrick DL, Zagami EA. Do ostomates do better? A study of symptoms and health-related quality of life. *Gastroenterology* 1989;96:A130.
49. Gerber A, Apt MK, Craig PH. The improved quality of life with the Kock continent ileostomy. *J Clin Gastroenterol* 1984;6:513–7.
50. Sjö Dahl R, Lemon E, Nystrom PO, Olaison G. Complications, surgical revision and quality of life with conventional and continent ileostomy. *Acta Chir Scand* 1990;156:403–7.
51. McLeod RS, Lavery IC, Leatherman JR, *et al*. Factors affecting quality of life with a conventional ileostomy. *World J Surg* 1986;10:474–80.
52. Martinsson ES, Josefsson M, Ek A. Working capacity and quality of life after undergoing an ileostomy. *J Adv Nurs* 1991;16:1035–41.
53. Küchenhoff J, Wirsching M, Drüner HU, Herrman G, Köhler C. Coping with a stoma—a comparative study of patients with rectal carcinoma or inflammatory bowel disease. *Psychother Psychosom* 1981;36:98–104.
54. Kennedy HJ. Quality of life in patients with an ileostomy. *International Disability Studies* 1988;10:175–6.
55. Bechi P, Tonelli L. Results in the radical surgical treatment of Crohn's disease. *Int Surg* 1982;67:325–8.
56. Lindhagen T, Eklund G, Leandroer L, Hildell J, Lindström C, Wenckert A. Crohn's disease in a defined population: course and results of surgical treatment. Part II. Large bowel disease. *Acta Chir Scand* 1983;149:415–21.
57. Lindhagen T, Eklund G, Leandroer L, Hildell J, Lindström C, Wenckert A. Crohn's disease in a defined population: course and results of surgical treatment. Part I. Small bowel disease. *Acta Chir Scand* 1983;149:407–13.

58. Speranza V, Simi M, Leardi S, Prantera C. Indications, strategy and results of surgical management in 141 cases of Crohn's disease. *Ital J Surg Sci* 1983;13:5-12.
59. Karesen R, Serch-Hanssen A, Thoresen BO, Hertzberg J. Crohn's disease: long-term results of surgical treatment. *Scand J Gastroenterol* 1981;16:57-64.
60. Meyers S. Assessing quality of life. *Mt Sinai J Med* 1983;50:190-2.
61. Cooper JC, Jones D, Williams NS. Outcome of colectomy and ileorectal anastomosis in Crohn's disease. *Ann R Coll Surg Engl* 1986;68:279-92.
62. Halme LE. Results of surgical treatment of patients with Crohn's disease. *Ann Chir Gynaecol* 1992;81:277-83.
63. Scott NA, Hughes LE. Timing of ileocolonic resection for symptomatic Crohn's disease—the patient's view. *Gut* 1994;35:656-7.
64. Farmer RG, Michener WM. Long-term follow-up of patients with Crohn's disease. *Dig Dis Sci* 1979;24:752-7.
65. Sales DJ, Kirsner JB. The prognosis of inflammatory bowel disease. *Arch Intern Med* 1983;143:294-9.
66. Sørensen VZ, Olsen BG, Binder V. Life prospects and quality of life in patients with Crohn's disease. *Gut* 1987;28:382-5.
67. Reed JF, Faust LA, Rosen L. Assessment of disease activity in a registry of Crohn's disease patients in eastern Pennsylvania. *Z Gastroenterol* 1991;29:378-82.
68. Greenberg GR, Feagan BG, Martin F, *et al.* and the Canadian Inflammatory Bowel Disease Study Group. Oral budesonide for active Crohn's disease. *N Engl J Med* 1994;331:836-41.
69. Robinson M, Hanauer S, Hoop R, Zbrozek A, Wilkinson C. Mesalamine capsules enhance the quality of life for patients with ulcerative colitis. *Aliment Pharmacol Ther* 1994;8:27-34.
70. Somerville KW, Langman MJ, Kane SP, MacGilchrist AJ, Watkinson G. Effect of treatment on symptoms and quality of life in patients with ulcerative colitis: comparative trial of hydrocortisone acetate foam and prednisolone 21-phosphate enemas. *BMJ* 1985;291:866.
71. Nyman M, Hansson I, Eriksson S. Long-term immunosuppressive treatment in Crohn's disease. *Scand J Gastroenterol* 1985;20:1197-203.
72. Gavalier JS, Delemos B, Belle SH, *et al.* Ulcerative colitis disease activity as subjectively assessed by patient-completed questionnaires following orthoptic liver transplantation for sclerosing cholangitis. *Dig Dis Sci* 1991;36:321-8.
73. Mughal M, Irving M. Home parenteral nutrition in the United Kingdom and Ireland. *Lancet* 1986;2:383-7.
74. Galandiuk S, O'Neill M, McDonald P, Fazio VW, Steiger EA. A century of home parenteral nutrition for Crohn's disease. *Am J Surg* 1990;159:540-5.
75. Pironi L, Miglioli M, Ruggeri E, *et al.* Home parenteral nutrition for the management of chronic intestinal failure: a 34 patient-year experience. *Ital J Gastroenterol* 1993;25:411-8.
76. Karush A, Daniels GE, Flood C, O'Connor JF. *Psychotherapy in chronic ulcerative colitis*. Philadelphia: WB Saunders, 1977.
77. Schwarz SP, Blanchard EB. Inflammatory bowel disease: a review of the psychological assessment and treatment literature. *Ann Behav Med* 1990;12:95-105.