Published in final edited form as:

Clin Gastroenterol Hepatol. 2009 February ; 7(2): 168–174.e1. doi:10.1016/j.cgh.2008.08.029.

CONTROVERSIES IN ULCERATIVE COLITIS: A SURVEY COMPARING DECISION-MAKING OF EXPERTS *VERSUS* COMMUNITY GASTROENTEROLOGISTS

Brennan M.R. Spiegel, MD, MSHS^{1,2,3,4,8}, Wayne Ho, MD¹, Eric Esrailian, MD, MPH^{2,8},
Stephan Targan, MD⁵, Peter Higgins, MD, PhD, MSc⁶, Corey A. Siegel, MD⁷, Marla
Dubinsky, MD⁵, and Gil Melmed, MD⁵

¹Veteran's Administration Greater Los Angeles Healthcare System

²David Geffen School of Medicine at UCLA

³UCLA School of Public Health

⁴CURE Digestive Diseases Research Center

⁵Cedars-Sinai Medical Center

⁶University of Michigan School of Medicine

⁷Dartmouth Hitchcock Medical Center

⁸UCLA/VA Center for Outcomes Research and Education (CORE)

Abstract

Background—Despite the development of consensus guidelines in ulcerative colitis (UC), there remain several areas of uncertainty in the everyday management of this incompletely understood disease. We performed a national vignette survey to measure variations in decision-making in areas of controversy.

Methods—We constructed a survey with 3 vignettes to measure decision-making in 4 areas of controversy in UC: (1) dysplasia management, (2) 5-ASA dosing, (3) diagnostic testing for underlying Crohn's Disease, and (4) treatment of steroid-refractory inpatient UC. We compared responses between a group of community gastroenterologists and UC experts.

Results—We received 192 responses (36% response). Compared to community gastroenterologists, UC experts were more likely to endorse colectomy for both unifocal and multifocal low grade dysplasia, use narrow band imaging and chromoendoscopy for surveillance colonoscopy, use high-dose 5-ASA for inducing remission, use long-term 5-ASA for cancer chemoprevention, order CT enterography to evaluate for Crohn's disease, and to have a lower threshold to call for surgery consultation in steroid-refractory UC. There was little agreement

© 2008 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

Corresponding Author Information: Brennan M.R. Spiegel, MD, MSHS, 11301 Wilshire Blvd, Bldg 115, Room 215, Los Angeles, California 90073, Office Phone: 310-268-3256, Office Fax: 310-268-4510, bspiegel@mednet.ucla.edu.

Disclaimer: The opinions and assertions contained herein are the sole views of the authors and are not to be construed as official or as reflecting the views of the Department of Veteran Affairs.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

regarding the optimal frequency of surveillance colonoscopy, even among experts. Most respondents favored using infliximab over cyclosporin in steroid-refractory UC.

Conclusions—Community gastroenterologists and UC experts vary dramatically in their approach to many areas of uncertainty in UC. The only area of consensus between groups is the use of infliximab over CSA in steroid-refractory UC – itself a controversial decision. These data suggest that current practice patterns are highly disparate, and focus attention on specific areas of disconnect that should be further investigated.

Keywords

Ulcerative Colitis; Provider Survey

INTRODUCTION

Ulcerative colitis (UC) is a prevalent and expensive condition that impairs quality of life.^{1,2,3} The large burden of illness of UC has spurred interest in developing improved diagnostic and therapeutic approaches – a process that has yielded novel diagnostic testing strategies, improved dysplasia surveillance practices, development of new therapies, and enhancements of older treatments. There is now sufficient evidence to support a set of best practices for the management of UC.^{4,5,6,7} Yet despite the development of consensus guidelines for UC,^{3–6} there remain several areas of uncertainty in the everyday management of this incompletely understood disease.

Four areas, in particular, are marked by clinical uncertainty: (1) screening and management of dysplasia, including the appropriateness of colectomy for flat low-grade dysplasia, use of chromoendoscopy or narrow band imaging (NBI) in surveillance, and the optimal frequency of surveillance colonoscopies; (2) use of mesalamine (5-ASA) products, including optimal dosing and duration, optimal formulation, and use of 5-ASA for cancer chemoprevention; (3) appropriateness of searching for underlying Crohn's disease in UC; and (4) treatment of steroid-refractory inpatient UC, including selection between cyclosporine (CSA) vs. infliximab, and the optimal timing of surgery.

Each area is marked by uncertainty about what constitutes best practice, yet gastroenterologists face these issues daily and must make decisions in the absence of evidence-based guidelines. Without explicit guidance, different physicians may employ different diagnostic tests, prescribe different therapies, and follow disparate management algorithms when faced with the same set of facts – a phenomenon that occurs in all areas of medicine, even when practice guidelines *are* available.^{8,9,10} In the field of gastroenterology, for example, we have demonstrated wide practice variations in Crohn's disease,⁸ irritable bowel syndrome,¹¹ and non-variceal hemorrhage.¹² Although minor variations in the process of care are an expected byproduct of modern healthcare, extreme variations denote poor consensus and poor quality of care – a fundamental principle of quality improvement initiatives.^{13,14,15} Wide practice variations may have several explanations, including the need for more evidence to determine the best course of action, the possibility that multiple approaches may be equally effective for a clinical scenario, or the need for existing evidence to be more effectively consolidated into guidelines and disseminated into practice. Recent efforts to improve the healthcare quality have focused on reducing extreme variations in diagnostic and therapeutic decision-making.^{12–14} As an initial effort to quantify the degree of variation in UC decision-making, we performed a national survey to measure beliefs of UC experts and community gastroenterologists regarding areas of uncertainty.

METHODS

Overview of Clinical Vignette Survey Methodology

Vignette Survey Design—We developed an online survey with 3 vignettes to evaluate scenarios in the diagnosis and management of UC. We developed the vignettes in concert with experts in UC and survey design specialists to ensure face validity, comprehensibility, and comprehensiveness. Each vignette began with a standardized patient presentation, and was followed by a management questions regarding diagnostic testing, treatment, and follow-up. The questions included vertical single best answers, horizontal matrix items, and open-ended items. The first vignette described an outpatient with a history of limited proctitis, presenting with a flare that partially responded to 2.4g of oral mesalamine. The second vignette depicted a patient with left-sided colitis, low-grade fever, and a symptom flare. The third vignette described an inpatient with UC and a flare persisting despite 6 days of intravenous steroids. Refer to the Appendix for the full vignettes.

Sampling Frame—We surveyed 2 groups of physicians (Table 1). The first group was a sample of 40 experts in UC. We identified these experts in a qualitative manner on the basis of their publication records in the last five years, leadership roles on advisory councils of the Crohn's and Colitis Foundations of America, and memberships on guideline committees. We use the term “expert” in reference to this group. The second group was a random sample of 500 gastroenterologists listed in the American Medical Association Masterfile. We use the term “non-expert” in reference to this largely community-based group.

Sample Size Considerations—Assuming 15 subjects for each of 10 potential independent predictors in multivariable regression analysis (see Analytic Statistics, below), we required a minimum of 150 to complete the survey in order to avoid overmatching of the regression models. Assuming a 30% response rate, we required at least 500 physicians to survey in the sampling frame.

Survey Distribution and Follow-up Procedures—Respondents received the survey electronically using an online questionnaire platform (Survey Monkey software). Physicians received e-mails with cover letters and a link to online survey. Non-responders after 2 weeks received a follow-up e-mail. One week after the second e-mail correspondence, a paper version of the questionnaire was mailed to non-responders. Using baseline data from the AMA Masterfile, we compared responders to non-responders for age, gender, region, practice setting, and years in practice.

Analyses

We measured the level of agreement between expert and community provider groups within 4 areas of controversy:

1) Controversies in Surveillance and Management of Dysplasia

Appropriateness of Colectomy: Although there are areas of consensus regarding the indication for colectomy in dysplasia,⁵ there are scenarios where the appropriateness of colectomy remains uncertain. We included 4 “mini” vignettes of young (35 years old) UC patients with varying dysplasia scenarios that are not explicitly addressed in published guidelines, including 2 cases with flat low-grade dysplasia, 3 cases with a solitary 4 mm adenoma, and 1 case of diffuse pseudopolypoidosis (Table 2). Respondents were informed that the patients had quiescent UC, no primary sclerosing cholangitis, and no family history of colon cancer. In each scenario, respondents rated the appropriateness of colectomy using a standard 9-point RAND Appropriateness Scale (RAS), with the following interpretation:

scores 1–3 = “generally inappropriate;” scores 4–6 = “unsure;” scores 7–9 = “generally appropriate.”¹⁶ We compared mean RAS scores between experts and non-experts using a 2-sided t-test and employing a p-value <0.05 as significant. In order to quantify the level of agreement, we also calculated the RAND/UCLA Disagreement Index (DI) for each factor.¹⁶ A DI>1.0 indicates “extreme variation” in beliefs. DI values below 1.0, including negative values, reflect increasing levels of agreement.²⁶ The DI is calculated using a standard published equation.^{8,16} In addition, we defined dichotomous “endorsement” of colectomy for each scenario as an RAS score of 7–9, as has been performed in previous studies,⁸ and compared endorsement between groups with a chi-squared test. We then performed multivariate logistic regression analysis to determine if any provider or practice-type characteristics were associated with endorsement of colectomy.

Frequency of Surveillance Colonoscopies: We measured beliefs about competing dysplasia surveillance guidelines. We asked the following question: “*Once you begin surveillance for dysplasia in left-sided colitis, how frequently would you do it?*” There were 6 choices (Table 3), including 2 variants of the American Gastroenterological Association (AGA) guidelines (annual or bi-annual surveillance indefinitely),¹⁷ the British Society of Gastroenterology guidelines (surveillance every 3 years, then every 2 years after 20 years of disease, then annually after 30 years of disease),¹⁸ surveillance every 3 years indefinitely, and surveillance in 1 year, if negative repeat every 3 years, then annually after 20 years. Respondents could also specify alternative surveillance practices. We compared endorsement of each guideline between groups with a chi-squared test, and employed a p-value of <0.05 as evidence for significance.

Use of Chromoendoscopy and NBI: We posed the following question about dysplasia surveillance: “*When performing surveillance colonoscopy in a UC patient with quiescent pancolitis, do you routinely use targeted biopsy techniques such as chromoendoscopy or narrow-band imaging?*” We compared binary responses between groups using a chi-squared test, and performed multivariable logistic regression to identify predictors of chromoendoscopy or NBI use in UC.

2) Controversies Regarding Use of 5-ASA in UC

Dose and Delivery of 5-ASA for Induction: There is controversy regarding the appropriateness of using low-dose (*i.e.* 2.4 g mesalamine equivalent) *vs.* high-dose (*i.e.* 4.8 g) 5-ASA for induction of remission in mild-to-moderate UC.^{19,20} Moreover, there are several 5-ASA delivery systems available. Respondents endorsing the use of 5-ASA in vignette #2 were asked to select among the following: balsalazide 6.75g, mesalamine 2.4g, mesalamine 4.8g, micronized mesalamine 4g, delayed-release (*i.e.* multi-matrix system [MMX]) mesalamine 2.4g, delayed-release mesalamine 4.8 g, and sulfasalazine titrated up to 4g daily as tolerated. We compared endorsement of each agent between the expert and non-expert groups, and compared endorsement of high *vs.* low dose mesalamine between groups using a chi-squared test. We subsequently performed multivariable logistic regression to identify predictors of high *vs.* low-dose mesalamine for induction.

Use of Indefinite 5-ASA After Induction: Once remission is induced with 5-ASA therapies, it remains unclear whether indefinite therapy is warranted to maintain remission, especially among those already on maintenance immunomodulators.^{19,20} Similarly, it is uncertain whether long-term 5-ASA has important benefits for cancer chemoprevention.^{19,21,22} We asked the following question: “*In a patient with UC in remission on 6MP/AZA, do you continue 5-ASA products?*” Respondents were instructed to select all that apply among the following choices: (a) “No, I taper them off to reduce pill burden;” (b) “No, I taper them off because it is not incrementally effective;” (c) “Yes, for

purposes of treatment of disease;" and (d) "Yes, for purposes of cancer chemoprevention." We compared responses between groups using chi-squared tests.

3) Appropriateness of Testing for Underlying Crohn's Disease in UC—Because Crohn's Disease involving the colon may be phenotypically similar to UC, there is often uncertainty about when and how to search for possible underlying Crohn's disease in patients carrying a diagnosis of "UC." We therefore compared diagnostic decision-making between and within groups regarding testing for underlying Crohn's disease. We explicitly designed vignette #2 to represent a "standardized" patient with presumptive left-sided UC, and minimized signs or other clinical clues for underlying Crohn's disease. Following a set of management questions, we posed the following: "*As part of the subsequent diagnostic evaluation, what, if any, additional diagnostic tests would you order?*" Respondents viewed a set of diagnostic tests that could be used to identify possible underlying Crohn's disease (see Table 4). Each test was rated using the RAS. We calculated DI statistics for each group of ratings, and compared endorsement between groups using a chi-squared test.

4) Management of Steroid-Refractory Inpatient UC—Patients with severe, steroid-refractory UC pose a clinical challenge, and the optimal management remains uncertain.²³ We asked respondents to select from a list of treatment options (see Table 5) for vignette #3 (severe steroid-resistant UC). We compared endorsement of each option between groups using a chi-squared test.

RESULTS

Sample Characteristics

Table 1 displays the characteristics of the survey respondents. One hundred and ninety-two respondents returned their surveys, including 25 of 40 UC experts (63% response) and 150 of 500 non-experts (30% response). The expert group was younger, more likely to practice in a University-based setting, more likely to be a member of the CCFA, and more likely to be engaged in conducting clinical research. Although the expert group had a smaller proportion of time dedicated to clinical care, they evaluated significantly more UC patients per month *vs.* the non-expert group. There were no significant differences between responders and non-responders for age (49.7 ± 9.5 *vs.* 50.4 ± 8.2 ; $p = \text{NS}$), male gender (91.4% *vs.* 89.2%; $p = \text{NS}$), years in practice (25.5 ± 12.4 *vs.* 26.2 ± 10.8 ; $p = \text{NS}$), practice setting, or region of practice (e.g. West Region=32.7% *vs.* 30.1%; South=26.7% *vs.* 28.2%; Northeast=19% *vs.* 17.6%, etc; $p = \text{NS}$ for all comparisons).

Controversies in Surveillance and Management of Dysplasia

Appropriateness of Colectomy—Table 2 lists the results of the expert and non-expert ratings regarding the appropriateness of colectomy in each of the 6 "mini" clinical vignettes. In the unifocal low grade dysplasia scenario, 75% of experts believed colectomy is "generally appropriate" (*i.e.* RAS score of 7–9) *vs.* 47% of non-experts ($p = 0.01$). Although experts were consistent in their shared belief about colectomy for unifocal low grade dysplasia ($\text{DI} = -3.8$), community GIs exhibited "extreme variation" in their belief ($\text{DI} = 1.9$). In the multifocal low grade dysplasia scenario, 100% of experts endorsed colectomy *vs.* 77% of non-experts ($p = 0.008$). Both experts and non-experts were internally consistent in their beliefs about multifocal LGD (expert $\text{DI} = -4.0$; non-expert $\text{DI} = -2.7$). In logistic regression adjusting for provider demographics, practice setting, expert status, society memberships, and clinical UC load, experts remained over 14 times more likely to endorse colectomy for unifocal flat LGD than non-experts ($\text{OR} = 14.7$; 95% $\text{CI} = 2.6\text{--}81.6$). The differences in beliefs regarding colectomy were less pronounced in the remaining scenarios.

Frequency of Surveillance Colonoscopy—Table 3 provides the data regarding preferred surveillance strategy between experts and non-experts. Within groups, non-experts were more likely to prefer the AGA guidelines (annual or bi-annual surveillance indefinitely) ($p<0.001$). This was repeated in the expert group, although the preference was not statistically significant ($p=0.08$). The least preferred strategy in both groups was surveillance every 3 years indefinitely.

Use of NBI and Chromoendoscopy—The expert group was more likely to endorse the use of NBI or chromoendoscopy as part of surveillance colonoscopy compared to the non-expert group (32% vs. 11%; $p=0.01$). This endorsement remained significant after adjusting for key covariates in regression analysis.

Controversies Regarding Use of 5-ASA in UC

Dose and Delivery of 5-ASA for Induction—Ninety-four percent of responders endorsed using a 5-ASA product vignette #1. Of these, 54% endorsed mesalamine, 17% delayed-release (MMX) mesalamine, 11% azulfidine, 9% balsalazide, and 3% micronized mesalamine. Fifty-six percent endorsed high-dose 5-ASA (4.8g mesalamine equivalent) to induce remission. There was a trend towards more experts vs. non-experts endorsing high-dose 5-ASA (71% vs. 53%; $p=0.1$).

Use of Indefinite 5-ASA After Induction—Twenty-one percent of responders endorsed indefinite 5-ASA following induction of remission with 5-ASA. However, if the patient subsequently required 6MP to induce remission, then 66% endorsed long-term 5-ASA. Of those endorsing indefinite 5-ASA, 33% believed 5-ASA is indicated for cancer chemoprevention. In logistic regression adjusting for provider characteristics, experts remained over 5 times more likely to endorse indefinite 5-ASA in remission ($OR=5.3$; 95% $CI=1.5-18.5$), and over 4 times more likely to use 5-ASA for chemoprevention (4.5; 1.3–15.5) vs. non-experts.

Appropriateness of Testing for Underlying Crohn's Disease in UC

Table 4 displays the results regarding diagnostic testing for Crohn's disease in UC. Sixty-five percent of respondents endorsed at least one additional test to evaluate for underlying Crohn's Disease in vignette #2, including either SFBT, EGD, WCE, MR enterography, CT enterography, or serologies. There was no difference between expert (75%) and non-expert (63%) groups for ordering at least one test ($p=0.24$). However, experts were more likely to endorse MR or CT enterography vs. non-experts ($p<0.001$ for both). In contrast, experts were significantly less likely to endorse IBD serologies (8.3% vs. 24%; $p=0.0001$). The DI results revealed that both experts and non-experts were internally consistent in their beliefs regarding all tests except use of MRI enterography and CT enterography. For these tests, the expert group revealed extreme variation in their beliefs (DI for MI=1.3; DI for CT=1.8).

Management of Steroid-Refractory Inpatient UC

Table 5 provides the data regarding management of steroid-refractory inpatient UC. Across groups, respondents favored the use of infliximab over CSA (infliximab=62%; CSA=26%; $p<0.0001$). This belief did not vary significantly between expert and non-expert groups. Experts were more likely to consult surgery vs. non-experts (50% vs. 17%; $p=0.0003$). Both groups were consistent that increasing the steroid dose further was not warranted, and that other maneuvers were required instead.

DISCUSSION

There is uncertainty about what constitutes best practice in UC. In order to proceed with additional guideline development, it is important to first understand current practice patterns, measure process of care, and identify areas of consensus and variation in UC. If current practice is consistent among providers, then it would indicate that future efforts to develop guidelines and quality indicators will be straightforward and uncontroversial. However, if current practice varies widely among providers, then it would suggest that more evidence is needed to determine the best course of action in areas of uncertainty. This variation might also reflect a belief that there are multiple effective approaches for a clinical situation, or that existing evidence needs to be more effectively disseminated to practitioners. Moreover, if providers cannot agree on how best to manage UC, then establishing additional best practices for UC will be an exercise fraught with pitfalls, as extreme variation is an indicator of inadequate evidence, poor consensus, and even poor healthcare quality.^{12–14} As a preliminary step, we sought to measure the current process of care and to identify the areas of consensus and variation in UC.

We found important differences regarding the use of colectomy in young patients with LGD. Whereas three-quarters of UC experts recommend colectomy in unifocal LGD, less than half of community gastroenterologists share this view. However, we also found that community gastroenterologists have highly polarized beliefs about this clinical scenario. For multifocal LGD, experts in this survey unanimously endorse colectomy, whereas only 77% of non-experts share this recommendation. This disconnect suggests that efforts should be made to develop and disseminate evidence-based guidelines regarding management of flat LGD among patients with colitis – a goal with immediacy given the high-stakes and severity of developing potentially avoidable colorectal cancer.

We found a wide variety of beliefs about the optimal guideline for dysplasia surveillance. Although both expert and non-expert groups tended to favor either annual or bi-annual surveillance indefinitely, nearly half of each group endorsed one of several alternative published guidelines. Moreover, 1 in 5 respondents supported one of several unpublished surveillance strategies, indicating that many providers “march to the beat of their own drum” with surveillance practices. This wide variation in practices may indicate that further research is necessary to better establish the evidence for competing approaches, and that it is premature to endorse any one guideline as representative of best practice. Moreover, it suggests that the wide variety of competing guidelines, in and of itself, is confusing to providers and a source of practice variation.

Experts in this survey were more likely to endorse the use of chromoendoscopy or NBI as part of dysplasia surveillance in UC. However, even among experts, the penetration of these techniques remains low. Although the disparity between experts and community providers may simply reflect experts’ proximity to the technology, the disconnect also indicates that existing and emerging evidence supporting these techniques should be disseminated in literature summaries, educational seminars, and demonstrations to practicing community gastroenterologists.

We identified important differences between expert and non-expert groups regarding the use of 5-ASA compounds in UC. In particular, compared to non-experts, UC experts are more likely to endorse indefinite 5-ASA after induction of remission, to believe 5-ASA has cancer chemopreventive properties, and to use high-dose 5-ASA to induce remission in moderate UC. These variations indicate that further development and dissemination of evidence-based 5-ASA guidelines is warranted to assist decision-making in these controversial areas. In addition, future research should measure the resource utilization implications and cost-

effectiveness of long-term 5-ASA use in UC – both in terms of treatment effectiveness and cancer chemoprevention.

Despite purposefully creating a UC vignette without obvious clues for underlying Crohn's disease, most respondents ordered one or more additional diagnostic tests to evaluate for Crohn's disease. It has been established that some patients with an initial diagnosis of UC undergo a subsequent change of diagnosis to CD. While clinical predictors of a change in diagnosis have been identified (*e.g.* nonbloody diarrhea, weight loss at diagnosis),²⁴ none of these were apparent in the vignette. With the exception of serologic markers, all the diagnostic tests for Crohn's disease involved assessment of the small intestine, suggesting that small bowel evaluation may be an important component of the evaluation of moderate UC unresponsive to conventional therapy. In addition, we found that experts were more likely to endorse MR or CT enterography than non-experts – an indication that they are either more familiar with these relatively new technologies, that these modalities are more readily available at the experts' institutions, or that non-experts are less eager to embrace these tests which yield higher quality images at the expense of radiation exposure (CT) or higher cost (MR) than conventional methods. Of note, even the expert group was internally conflicted regarding the appropriateness of MR and CT enterography, suggesting that experience with these technologies varies widely among key opinion leaders. This suggests that use of MR and CT enterography will remain inconsistent until, at the very least, experts themselves forge consensus about the appropriate use of these technologies.

There is debate about whether to use CSA or infliximab in patients with steroid-refractory inpatient UC.²³ Randomized controlled trials support both therapies,^{25,26,27,28} although these studies vary in terms of sample size, previous steroid exposure, and patient illness severity. There are no guidelines to support one therapy over the other. We found that most respondents endorsed the use of infliximab over CSA – a finding that was consistent across groups. In contrast to CSA, there are limited long-term outcomes data on the use of infliximab in the setting of the hospitalized UC patient unresponsive to intravenous steroids, yet infliximab was the preferred treatment among both experts and non-experts. This suggests that factors other than published evidence may guide decision-making in this tenuous sub-population, perhaps influenced by the availability or burden of daily CSA monitoring, influence of marketing campaigns for infliximab, or increased experience and physician comfort with infliximab. More long-term data for infliximab use in this setting is needed before an evidentiary basis can justify what appears to be the favored current practice.

A limitation of this study is that survey responses may not be reflective of actual decision-making in clinical practice. Directly observing patient-provider interactions is the gold standard for assessing process of care. However, this approach is also limited because of the Hawthorne effect in which providers artificially alter their practice when they are knowingly observed. This undermines the efforts to capture the true process of care. Standardized patients²⁹ and medical record data abstraction³⁰ are alternatives. Notably, survey-based clinical vignettes are validated as an accurate surrogate for both chart abstraction and standardized patients.³¹ An additional limitation is that our vignettes do not represent all possible scenarios in UC. Other investigators may have developed different vignettes with different details. However, we followed several steps to ensure adequate content validity of the vignettes, including consultation with key opinion leaders in UC, review by experts in survey design, and pilot testing for comprehensibility. Moreover, regardless of the precise content of the vignettes, all providers were faced with the same clinical data, yet came to different conclusions on many occasions. Thirdly, although we operationalized “extreme variation” using a validated and quantitative method,¹⁶ there is no gold standard for this construct. Moreover, variation itself is a natural consequence of medicine being as much art

as science, and thus some basal level of variation is to be expected. However, our analyses reveal that the process of care in UC may vary widely between providers – more than would be expected by chance alone. Evidence has shown that extreme variation in care is an indicator of inadequate evidence, poor consensus, and even poor healthcare quality.¹³ Although it may be disquieting to reveal these wide variations in UC, we believe it is important to highlight these variations because of the potential implications for current and future quality of care. Quantifying the current process of care allows both expert and community physicians to recognize the extreme variations in our approaches and to consider next steps to streamline care – always while recognizing that no two physicians will approach the same patient in the exact same way, nor should they.

Our survey is further limited by the relatively low response rate (30%). There may be systematic differences between responders and non-responders. Non-responders typically cite a wide range of reasons for their failure to participate, including not receiving the survey, believing the survey did not pertain to their practice, lacking sufficient time, or simply choosing not to participate. We cannot know the specific reasons for non-response to this survey. It is possible that non-responders were too busy, did not receive the email or paper surveys, or had other systematic differences – i.e. do not evaluate enough IBD patients to feel comfortable taking the survey, do not participate in academic studies, are not clinically active, etc. Using data from the AMA Masterfile, we were able to confirm that non-responders were similar to responders in terms of age, gender, years in practice, practice settings, and geographic locations. This suggests that respondents were not systematically different from non-respondents for key measured characteristics. Nonetheless, there may be unmeasured factors, including those described above, that remain systematically different between groups. As with any provider survey, a higher response rate could have potentially altered the results, although it is difficult to accurately predict in what direction, if any, the results would be different.

In conclusion, we found wide variations in practice patterns in UC. In particular, compared to community gastroenterologist, UC experts are more likely to endorse colectomy for LGD, use NBI and chromoendoscopy for surveillance colonoscopy, use high-dose 5-ASA for inducing remission, use long-term 5-ASA for cancer chemoprevention, order CT and MR enterography to evaluate for underlying Crohn's disease, and to have a lower threshold to consult surgery in steroid-refractory UC. These variations indicate that best practices in these areas remain uncertain and should be subjected to further research and guideline development.

Acknowledgments

Grant support: Dr. Spiegel is supported by a Veteran's Affairs Health Services Research and Development (HSR&D) Career Development Transition Award (RCD 03-179-2), and the CURE Digestive Disease Research Center (NIH 2P30 DK 041301-17). Dr. Corey Siegel is supported by a Crohn's and Colitis Foundation Career Development Award and NIH 1K23DK078678-01A1. This study was supported by an investigator-initiated research grant by Procter & Gamble.

References

1. Pallis AG, Mouzas IA. Instruments for quality of life assessment in patients with inflammatory bowel disease. *Dig Liver Dis.* 2000; 32:682–8. [PubMed: 11142577]
2. Caellas F, Lopez-Vivancos J, Vergara M, Malagelada J. Impact of inflammatory bowel disease on health-related quality of life. *Dig Dis.* 1999; 17:208–18. [PubMed: 10754360]
3. Irvine EJ. Quality of life of patients with ulcerative colitis: Past, present, and future. *Inflamm Bowel Dis.* 2007 Epub ahead of print.

4. Kornbluth A, Sachar D. Ulcerative colitis practice guidelines in adults (update): American College of Gastroenterology, Practice Patterns Committee. *Am J Gastroenterol*. 2004; XX:1371. [PubMed: 15233681]
5. Society for Surgery of the Alimentary Tract. SSAT patient care guidelines: management of ulcerative colitis. *J Gastrointest Surg*. 2007; 11:1203–6. [PubMed: 18062071]
6. Itzkowitz SH, Present DH. Crohn's and Colitis Foundation of America Colon Cancer in IBD Study Group. Consensus conference: Colorectal cancer screening and surveillance in inflammatory bowel disease. *Inflamm Bowel Dis*. 2005; 11:314–21. [PubMed: 15735438]
7. Carter MJ, Lobo AJ, Travis SP. IBD Section, British Society of Gastroenterology. Guidelines for the management of inflammatory bowel disease in adults. *Gut*. 2004; 53 (Suppl 5):V1–16. [PubMed: 15306569]
8. Esrailian E, Spiegel BMR, Targownik L, Dulai GS, Gralnek IM. Differences in the management of Crohn's disease among experts and community providers, based on a national survey of sample case vignettes. *Aliment Pharmacol Ther*. 2007; 26:1005–18. [PubMed: 17877507]
9. Chassin MR, Kosecoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services: a study of three procedures. *JAMA*. 1987; 258:2533–37. [PubMed: 3312655]
10. Tobler HG, Sethi GK, Grover FL, Shroyer AL, Moritz TE, Henderson WG, Hammermeister KE. Variations in Processes and Structures of Cardiac Surgery Practice. *Med Care*. 1995; 33:OS43–OS58. [PubMed: 7475411]
11. Spiegel BM, Farid M, Esrailian E, Talley J, Chang L. Appropriateness of diagnostic testing in IBS: provider beliefs do not match published guidelines. *Gastroenterology*. 2006; 130:S1134.
12. Spiegel BM, Gralnek IM, Laine L, Jensen D, Dulai GS, Eisen G, Esrailian E. Adherence to “best practices” in nonvariceal upper GI hemorrhage (NVUGIH): a national vignette survey of experts versus community providers. *DDW*. 2007
13. Brook RH, Lohr KN. Efficacy, effectiveness, variations, and quality boundary-crossing research. *Med Care*. 1985; 23:710–22. [PubMed: 3892183]
14. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. *Int J Qual Health Care*. 2000; 12:281–95. [PubMed: 10985266]
15. Wennberg JE. Practice variations and health care reform: connecting the dots. *Health Aff (Millwood)*. 2004; VAR140–4. Suppl WebExclusive. [PubMed: 15471778]
16. Fitch, K.; Bernstein, SJ.; Aguilar, M., et al., editors. *RAND/UCLA Appropriateness Method User's Manual*. Santa Monica: RAND; 2001.
17. Winawer S, Fletcher R, Rex D, et al. AGA guidelines: colorectal cancer screening and surveillance. *Gastroenterol*. 2003; 124:544.
18. Eaden JA, Mayberry JF. Guidelines for screening and surveillance of asymptomatic colorectal cancer in patients with inflammatory bowel disease. *Gut*. 2002; 51 (Suppl 5):V10. [PubMed: 12221032]
19. Bergman R, Parkes M. Systematic review: the use of mesalazine in inflammatory bowel disease. *Aliment Pharmacol Ther*. 2006; 23:841–55. [PubMed: 16573787]
20. Sutherland L, MacDonald JK. Oral 5-aminosalicylic acid for induction of remission in ulcerative colitis. *Cochrane Database Syst Rev*. 2003; 3:CD000543. [PubMed: 12917894]
21. Munkolm P, Loftus EV Jr, Reinacher-Schick A, Kornbluth A, Mittman U, Esendal B. Prevention of colorectal cancer in inflammatory bowel disease: value of screening and 5-aminosalicylates. *Digestion*. 2006; 73:11–9. [PubMed: 16410688]
22. Giannini EG, Kane SV, Testa R, Savarino V. 5-ASA and colorectal cancer chemoprevention in inflammatory bowel disease: can we afford to wait for ‘best evidence’? *Dig Liver Dis*. 2005; 37:723–31. [PubMed: 16023905]
23. Velayos F, Mahadevan U. Management of steroid-dependent ulcerative colitis: immunomodulatory agents, biologics, or surgery? *Clin Gastroenterol Hepatol*. 2007; 5:668–71. [PubMed: 17544992]
24. Melmed GY, Elshoff R, Chen GC, et al. Predicting a change in diagnosis from ulcerative colitis to Crohn's disease: a nested, case-control study. *Clin Gastroenterol Hepatol*. 2007; 5:602–8. [PubMed: 17478347]

25. Lichtiger S, Present DH, Kornbluth A, et al. Cyclosporine in severe ulcerative colitis refractory to steroid therapy. *N Engl J Med.* 1994; 330:1841–5. [PubMed: 8196726]
26. Jarnerot G, Hertevig E, Friis-Liby I, et al. Infliximab as rescue therapy in severe to moderately severe ulcerative colitis: a randomized, placebo-controlled study. *Gastroenterology.* 2005; 128:1805–11. [PubMed: 15940615]
27. Van Assche G, D'Haens G, Noman M, et al. Randomized, double-blind comparison of 4 mg/kg versus 2 mg/kg intravenous cyclosporine in severe ulcerative colitis. *Gastroenterology.* 2003; 125:1025–31. [PubMed: 14517785]
28. Kohn A, Daperno M, Armuzzi A, et al. Infliximab in severe ulcerative colitis: short-term results of different infusion regimens and long-term follow-up. *Aliment Pharmacol Ther.* 2007; 26:747–56. [PubMed: 17697208]
29. Colliver JA, Swartz MH. Assessing clinical performance with standardized patients. *JAMA.* 1997; 278:790–1. [PubMed: 9286845]
30. Luck J, Peabody JW, Dresselhaus TR, et al. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *Am J Med.* 2000; 108:642–9. [PubMed: 10856412]
31. Peabody JW, Luck J, Glassman P, Hansen J, Spell M, Lee M. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med.* 2004; 141:771–80. [PubMed: 15545677]

APPENDIX

CLINICAL VIGNETTE # 1

PATIENT HISTORY

A 44 year-old woman with a history of ulcerative colitis, first diagnosed with limited proctitis 5 years ago, now presents to your office with a 2-month history of intermittent, crampy abdominal pain and loose, bloody stools up to 3 times daily. She has experienced infrequent flares that were successfully treated with Canasa™ suppositories roughly once every other year; these were discontinued shortly after symptom resolution. Over the past year she has flared twice. One month ago she was placed on mesalamine (Asacol®) 2.4 g PO daily in addition to nightly Canasa™. Her bowel movement frequency subsequently decreased from 4 loose bowel movements daily to 2 BMs per day. However, she continues to have blood in her stool and lower abdominal cramping that generally improves with BMs.

She does not have weight loss, fevers, or chills. A total colonoscopy was performed 1 year ago while the patient was asymptomatic and showed mild chronic inflammation confined to 20 cm from anal verge, but no acute inflammation.

She has no allergies and takes no other medications. She does not smoke. She confirms that she does take her medication as prescribed.

PHYSICAL EXAMINATION

Examination is only remarkable for mild LLQ tenderness to palpation.

LABORATORY TESTS

Complete blood count, chemistries, and stool studies for infectious etiologies, including *C. diff*, are negative; ESR is 18 (normal=0–15)

CLINICAL VIGNETTE # 2

PATIENT HISTORY

A 32 year-old man with a history of UC, first diagnosed 4 years ago, is now referred to you by his primary doctor for complaints of recurrent, bloody diarrhea for the past 3 months. The patient reports having 3–5, loose, bloody bowel movements daily with accompanying moderate lower abdominal pain. He reports intermittent subjective fevers over the past 1–2 weeks. His primary care physician obtained stool studies (including stool leukocytes, ova & parasites, bacterial cultures, *C. diff* toxin x 3), which thus far have all been unrevealing. The patient was started on 5-ASA enemas and has been compliant with them.

PHYSICAL EXAMINATION

On physical examination, he has a low-grade fever of 100.3 F, has a heart rate of 86, and a blood pressure of 136/84. He appears euvolemic. Abdominal exam reveals mild tenderness in the left lower quadrant with palpation, without rebound or guarding. Rectal exam reveals brown stool mixed with red blood.

ENDOSCOPIC FINDINGS

Your endoscopic evaluation reveals diffuse moderate inflammation to the mid-transverse colon (see image below and accompanying **color plate** image on the **back cover** of this folder), with normal-appearing proximal colon; histopathology confirms mild to moderate acute and chronic inflammatory changes without granulomas, dysplasia, or infection, “consistent with inflammatory bowel disease” in the endoscopically affected segments, with normal histology proximally.

CLINICAL VIGNETTE # 3

PATIENT HISTORY

A 38 year-old woman with pANCA-positive ulcerative colitis is admitted for a flare with 8–10 bloody bowel movements daily despite longstanding 6MP (with therapeutic 6TG levels) and 1 week of outpatient prednisone 40mg daily. After 6 days of inpatient IV steroid therapy (methylprednisolone 20 mg TID) she still has 6–8 bloody bowel movements daily. Her blood pressure remains stable and her mental status remains fully intact. Her examination is notable for diffuse tenderness to palpation but no peritoneal signs, rebound, or guarding. CMV and other enteric infections are not evident with stool studies and colonic biopsies. ESR is 46, WBC count is 11,000 with 10% band forms, and electrolytes are within normal

limits. CT scan reveals pan-colonic thickening, but no bowel dilation. She says she does “not want to lose the colon if possible.”

Table 1

Demographic and Practice-Pattern Information of Respondents

Variable	Mean $\pm$ SD Experts N=25	Mean $\pm$ SD Community Providers N=150	p-value
Age (mean years)	45.8 $\pm$ 1.3	50.4 $\pm$ 9.3	0.03
Male Gender	95.8%	90.5%	0.4
Years in Practice	25.1 $\pm$ 4.2	25.5 $\pm$ 1	0.88
Clinical GI Practice	57.6%	80.1%	0.002
Clinical Research Practice	26.7%	7.5%	<0.0001
Member of Crohn's and Colitis Foundation of America (CCFA) (%)	75%	32%	0.0001
Number of UC patients seen per month	128 $\pm$ 17	25 $\pm$ 4	<0.0001
Practice Setting (%)			
Private Practice	8%	53%	<0.001
Health Maintenance Organization	0%	3%	
University-Based Practice	88%	16%	
Veteran's Affairs Hospital	0%	12%	
Other	4%	16%	
Geographic Location (%)			
West	13	36	<0.001
Midwest	26	19	
South	19	28	
Northeast	31	17	
International	5	0	

Expert vs. Non-Expert Beliefs about Appropriateness of Colectomy in 6 Different Clinical Scenarios. Appropriateness is measured on a 9-point RAND Appropriateness Scale (RAS), where 1–3=“generally inappropriate,” 4–6=“unsure,” and 7–9=“generally appropriate.” “Endorsement” is defined as scoring an RAS of 7–9. For example, 75% of experts endorsed colectomy in scenario 2 – a 35 year old with unifocal flat low grade dysplasia (LGD). Variation within group is measured using the RAND Disagreement Index (DI), where a DI 1.0 indicates “extreme variation” in ratings.

Table 2

Patient Scenarios	Experts			Non-Experts		
	Mean RAS	% Endorse	DI	Mean RAS	% Endorse	DI
35 year old with single, 4mm adenoma in involved area, removed with snare polypectomy. Adjacent biopsies normal.	2.68	12.5%	0.49	2.24	14.2%	0.29
35 year old with unifocal flat LGD	6.80	75.0%	–3.80	5.60	47.5%	1.90
35 year old with 50–60 pseudopolyps, with sampling of 5 all showing inflammation but no dysplasia	3.10	12.5%	0.68	2.60	14.1%	0.29
35 year old with multifocal flat LGD	8.70	100.0%	0.00	7.40	77.0%	–2.10
35 year old with single, 4mm sessile adenoma in involved area, removed with polypectomy. Adjacent biopsies are normal.	1.60	12.5%	0.03	2.40	12.5%	0.29
35 year old with single, 4mm adenoma in uninvolved area, removed with polypectomy. Adjacent biopsies are normal.	3.70	15.0%	0.76	2.30	21.0%	0.28
						0.001

Table 3

Expert vs. Non-Expert Endorsement of Competing Dysplasia Surveillance Guidelines. There were no significant differences across groups. Within groups, there was a higher preference for performing annual or every-other year surveillance (expert ANOVA p-value=0.08; non-expert ANOVA p-value<0.001).

Surveillance Guideline	Experts (% Endorsing)	Non-Experts (% Endorsing)	p-Value
Annual surveillance indefinitely	17%	23%	0.53
Surveillance every 2 years indefinitely	21%	28%	0.45
Surveillance every 3 years indefinitely	4%	3%	0.84
Surveillance in 1 year, if negative repeat every 3 years, then annually after 20 years	21%	12%	0.28
Surveillance every 3 years, then every 2 years after 20 years of disease, then annually after 30 years of disease	16%	14%	0.75
“Other” guidelines (open ended responses – variable)	21%	20%	0.84

Table 4
Expert vs. Non-Expert Beliefs about Appropriateness of Various Diagnostic Tests for Underlying Crohn's Disease in UC.

Diagnostic Tests	Experts			Non-Experts			RAS P- Value
	Mean RAS	% Endorse	DI	Mean RAS	% Endorse	DI	
Small Bowel Follow-Through	3.85	25.0%	0.74	3.12	18.3%	0.74	0.18
Upper Endoscopy	2.1	8.3%	0.29	1.95	11.6%	0.29	0.64
Wireless Capsule Endoscopy	1.7	8.3%	0.17	2.3	14.1%	0.29	0.19
MRI Enterography	3.6	33.3%	1.3	2	13.3%	0.29	0.0003
CT Enterography	4.6	33.0%	1.8	2.3	15.0%	0.29	<0.0001
IBD Serologies (e.g. ANCA, ASCA, Omp-C)	1.8	8.3%	0.13	3.8	24.1%	0.65	0.0001

Table 5

Expert vs. Non-Expert Endorsement of Competing Management Strategies in Steroid-Refractory Inpatient UC. Respondents could choose multiple simultaneous treatments (thus totals exceed 100%).

Management Strategy	Experts (% Endorsing)	Non-Experts (% Endorsing)	p-Value
Increase steroid dose	0.0%	3.0%	0.36
Keep Steroid, Add Infliximab	54.0%	53.0%	0.91
Stop Steroid, Add Infliximab	8.3%	9.2%	0.89
Call Surgical Evaluation	50.0%	16.8%	0.0003
Keep Steroid, Add Cyclosporine	33.3%	22.6%	0.27
Stop Steroid, Add Cyclosporine	0.0%	2.5%	0.43