The IBD Control Questionnaire: the development and psychometric validation of a questionnaire for measuring inflammatory bowel disease control from the patient's perspective.

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Medicine by

Clare Rachel Ormerod

December 2018

<u>Statement</u>

I declare this thesis to be comprised of my own work. Patient recruitment and data extraction was assisted by Mrs Daniela Shackcloth, Research Nurse. Statistical support was provided by Dr Keith Bodger.

Acknowledgements

With sincere thanks to my supervisor, Dr Keith Bodger. For his support, guidance and endless patience.

With thanks also to my husband, for his support and motivation, and for distracting our children for long enough for me to write this.

Finally, to my children. For making life somewhat harder, but infinitely more enjoyable.

Abstract

The IBD Control Questionnaire: the development and psychometric validation of a questionnaire for measuring inflammatory bowel disease control from the patient's perspective.

Clare Ormerod

Introduction: The importance of patient reported outcome measures (PROMs) is increasingly recognised. However, their use in inflammatory bowel disease (IBD) care is not widespread. The aim of this research was to develop and validate a PROM for use in day to day clinical practice.

Methods: (1) Questionnaire development: Questionnaire specification was determined by a multidisciplinary steering group. Literature review of existing PROMs was undertaken. Focus groups meetings and one-to-one patient interviews were performed. Participants were asked to discuss the concept of "control" of their IBD. Thematic analysis of field notes and transcribed quotes was performed. A draft questionnaire was designed and a pilot study of 30 patients was undertaken. (2) Prospective validation of the IBD Control Questionnaire: Patients completed the IBD-Control Questionnaire and the following established measures: a quality of life questionnaire (UK-IBD-Q), EuroQol (EQ-5D) and the Hospital Anxiety and Depression Score. Disease activity indices were recorded (Harvey-Bradshaw Index or Simple Clinical Colitis Activity Index). A global physician assessment (blinded to questionnaire score) was also performed. Data were recorded at baseline and subsequent hospital visits. The psychometric properties of the questionnaire were determined as detailed below.

Results: Core domains of "physical", "social", "emotional" and "treatment" were identified on the basis of literature review and patient consultation. The pre-defined questionnaire specification was used to guide item selection to represent these domains. The 'IBD-Control' questionnaire comprises 13 items plus a visual analogue scale (VAS) (0–100). 299 patients returned baseline surveys (Crohn's disease, n=160; ulcerative colitis, n=139) and 138 attended for repeat visits. *Completion time* (mean; SD): 1 min 15 s; 25 s; *Internal consistency*: Cronbach's α for all 13 items (0.85); for subgroup of eight questions ('IBD-Control-8'; 0.86). Strong correlation between IBD-Control-8 and IBD-Control-VAS (r=0.81). *Test-retest reliability* (2 week repeat): intra-class correlation=0.97 for IBD-Control-8 and 0.96 for IBD-Control-VAS. *Construct validity*: Moderate-to-strong correlations between IBD-Control-8 and IBD-Control-VAS versus activity indices, UK-IBD-Q and EQ-5D (utility) with r values 0.52–0.86. *Discriminant validity* (mean instrument scores for remission, mild, moderate or severe): p<0.001 (analysis of variance (ANOVA)). *Sensitivity to change*: Effect sizes: 0.76–1.44. *Sensitivity and specificity to identify quiescent patients:* area under the receiver operating characteristic curve 0.90 IBD Control 8, 0.86 IBD Control VAS. Cut off values for identifying quiescent patients: IBD Control 8 - 13 points or more 90.6% specificity, IBD Control VAS – 85 points or more 90% specificity.

Conclusion: The IBD Control Questionnaire is a valid patient reported measure of disease control. Its brevity and generic content make it suited to routine clinical care.

Contents

Preface

Chapters

List of tables

List of figures

List of abbreviations

ADL..Activities of Daily Living

ANOVA..Analysis of Variance

AVVQ...Aberdeen Varicose Vein Questionnaire

AZA...Azathioprine

CCKNOW...Crohn's and Colitis Knowledge Score

CCPKNOW...Crohn's and Colitis Pregnancy Knowledge Score

CD...Crohn's disease

CDAI...Crohn's Disease Activity Index

CDEIS...Crohn's Disease Endoscopic Index of Severity

CGQL...Cleveland Global Quality of Life Instrument

COPD...Chronic Obstructive Pulmonary Disease

CPWDQ.......................................Crohn's Disease Perceived Work Disability Questionnaire

CRP..C Reactive Protein

CT..Computed Tomography

CWAII...Crohn's Work Activity Impairment Index

EIBDQ...Edinburgh IBD Quality of Life Questionnaire

ER..Endoplasmic reticulum

ES..Effect Size

ESR..Erythrocyte Sedimentation Rate

FACIT-F……………………………Functional Assessment of Chronic Illness Therapy – Fatigue

FACIT-G……………Functional Assessment of Chronic Illness Therapy General Questionnaire

FACT……………………………………………………………………………Functional Assessment of Cancer

FDA……………………………………………………………………………US Food and Drug Administration

GP…………………………………………………………………………………………………General Practice

GPA………………………………………………………………………………Global Physician Assessment

HADS………………………………………………………………………Hospital Anxiety and Depression Scale

HBI……………………………………………………………………………………………Harvey Bradshaw Index

HHMQoL-IBD…………………………………………Household Member Quality of Life Questionnaire

HRQoL………………………………………………………………………………Health Related Quality of Life

IBD……………………………………………………………………………………Inflammatory Bowel Disease

IBD-DS………………………………………………………………Inflammatory Bowel Disease Disability Score

IBDQ…………………………………………………………………Inflammatory Bowel Disease Questionnaire

IBDQIP……………………………………Inflammatory Bowel Disease Quality Improvement Project

IBDQ-9……………………………………………………Inflammatory Bowel Disease Questionnaire 9

IBDQ-36………………………………………Extended Inflammatory Bowel Disease Questionnaire

IBD-SES………………………………………………………………………………………IBD Self-efficacy Scale

IBDU………………………………………………………………Inflammatory Bowel Disease Unclassified

ICC………………………………………………………………………………Intraclass Correlation Coefficient

ICF………………………………International Classification of Functioning Disability and Health

IL-10...Interleukin 10

IPAA...Ileo-anal Pouch Procedure

IRGM..Immunity regulated GTPase M

JAK2...Janus kinase 2

KQ...Knowledge Questionnaire

MCS..Mental Component Summary

MCID...Minimal Clinically Important Difference

MID..Minimal Important Difference

MMAS-8...Morisky Medication Adherence Scale

MRI..Magnetic Resonance Imaging

MSRM...Modified Standardised Response Mean

MST1... Macrophage stimulating 1

NCSI...National Cancer Survivorship Initiative

NICE...National Institute for Health and Care Excellence

NOD2...Nucleotide-binding oligomerisation Domain containing 2

NRS...Numeric Rating Scale

OHS...Oxford Hip Score

OKS.. Oxford Knee Score

OMGE..Organization Mondial de Gastro-enterologie

PCS...Physical Component Summary

PDAI..Perianal Disease Activity Index

PGA..Physicians Global Assessment

PROM..Patient Reported Outcome Measure

PSC..Primary Sclerosing Cholangitis

QALY...Quality Adjusted Life Years

QUOTE-IBD...Quality of Care Through the Patient's Eyes IBD

REC..Research Ethics Committee

RFIPC..Rating Form of IBD Patient Concerns

ROC... Receiver Operator Characteristics

SBCE..Small Bowel Capsule Endoscopy

SCCAI..Simple Clinical Colitis Activity Index

SD..Standard deviation

SES-CD..Simple Endoscopic Score for Crohn's Disease

SF-36...Short Form 36 Health Survey Questionnaire

SHS..Short Health Scale

SIP..Sickness Impact Profile

SIBDQ...Short Inflammatory Bowel Disease Questionnaire

SICC-IBD..Social Impact of Chronic Conditions - IBD

SRM..Standardised Response Mean

STAT3... Signal transducer and activator of transcription 3

TNF...Tumour Necrosis Factor

TTO...Time Trade Off

TSQ-C..Treatment Satisfaction Questionnaire for Crohn's Disease

UC...Ulcerative Colitis

UCEIS...Ulcerative Colitis Endoscopic Index of Severity

UK-IBDQ...............................United Kingdom Inflammatory Bowel Disease Questionnaire

US..Ultrasound

VAS...Visual analogue scale

WPAI..Work Productivity and Activity Impairment

WPAI:CD............................ Work Productivity and Activity Impairment in Crohn's Disease

5-ASA..5-aminosalicylic acid

6MP...6-mercaptopurine

6-TGN..6-thioguanine nucleotides

Chapter 1 Introduction

The need to measure health status from the patient's perspective is becoming increasingly recognised. Patient reported outcome measures are now used to monitor quality of healthcare within the National Health Service and there is ongoing work to develop their use further. Patient reported outcome measures might also play a role in other aspects of healthcare such as informing day-to-day clinical decision-making[1].

Inflammatory bowel disease (IBD) is a chronic disorder of the gastro-intestinal tract. Whilst a number of patient reported outcome measures have been developed for use in IBD, to date none have become established in routine clinical care. This study aims to develop and validate a new patient reported outcome measure for use in the routine clinical care of patients with inflammatory bowel disease.

1.1 Inflammatory bowel disease

1.1.1 Definition

Inflammatory bowel disease is a chronic inflammatory disorder of the gastro-intestinal tract and comprises of two main conditions: ulcerative colitis and Crohn's disease.

Ulcerative colitis is defined as a chronic inflammatory condition resulting in continuous mucosal inflammation of the colon. Inflammation involves the rectum as well as the remaining colon to a variable extent. Inflammation is continuous. There is an absence of granulomata in histological specimens[2].

Crohn's disease is characterised by patchy inflammation with skip lesions. It can affect any part of the gastro-intestinal tract and inflammation is transmural. Crohn's disease is defined by location as well as behaviour (for example, the presence of fistulating disease)[3].

A small proportion of patients with colitis cannot be diagnosed with either ulcerative colitis or Crohn's colitis on the basis of standard investigation modalities. These cases are termed "IBD unclassified " (IBDU)[2].

The Montreal classification of inflammatory bowel disease aims to define IBD on the basis of both disease distribution and behaviour. Ulcerative colitis is classified on the basis of distribution (extent): E1-proctitis (disease limited to the rectum), E2-left sided disease (inflammation distal to the splenic flexure) and E3-extensive (involvement extends proximal to the splenic flexure) (Figure 1). Crohn's disease classification involves age, distribution and disease behaviour (Figure 2)[2]. Classification of IBD is of use to guide clinical management and is also of benefit in clinical trials.

Figure 1 The Montreal classification of ulcerative colitis[2]

	Description	Extent
E1	Proctitis	Inflammation limited to rectum
E2	Left-sided (distal) colitis	Inflammation beyond rectum, but distal to the splenic flexure
E3	Extensive colitis	Inflammation proximal to the splenic flexure

Figure 2 The Montreal classification of Crohn's disease[2]

Age	Location	Behaviour
A1 less than 16	L1 Ileal	B1 Non-stricturing, non-penetrating
A2 17 to 40	L2 Colonic	B2 Penetrating
A3 over 40	L3 Ileo-colonic	B3 Stricturing
	L4* Isolated upper GI disease	p** perianal disease

*L4 is added to L1-L3 if upper GI disease is also present
**p is added to B1-B3 if perianal disease is also present

1.1.2 Epidemiology

Inflammatory bowel disease is a disorder of industrialised countries and is most common in areas such as Northern Europe, the United Kingdom and Northern America [4]. The annual incidence of IBD in Europe per 100 000 population per year, has been estimated at 5.4 for Crohn's disease, 8.2 for ulcerative colitis and 1.7 for IBDU (IBD unclassified)[5].

Within industrialised countries, the incidence of ulcerative colitis was first to rise, followed by Crohn's disease. The incidence of ulcerative colitis has now stabilised, whereas that of Crohn's disease has continued to rise. Of note, the incidence of Crohn's disease presenting in childhood is rising [6].

The epidemiology of IBD is changing globally. Whilst incidence and prevalence have stabilised in developed areas, there has been a rise in cases of IBD in developing countries. There seems to be an increase in incidence of IBD as countries transition from developing to industrialised nations. This phenomenon supports the view that environmental risk factors play a key role in the development of IBD [7].

1.1.3   Risk factors

A number of risk factors for IBD have been identified. Smoking is a risk factor for Crohn's disease and meta-analysis suggests a two fold increase risk in smokers[8]. Moreover, smoking will also increase the risk of an adverse clinical course of the disease. Smokers are more likely to have ileal disease[9], less likely to have an inflammatory phenotype[10] and more likely to have recurrent disease following surgery[11]. Smokers are also more likely to need immunosuppression[12].

In contrast to Crohn's disease, smoking and ulcerative colitis are inversely associated[13]. The mechanism for this has not been identified. Smoking also seems to be protective for conditions associated with ulcerative colitis such as primary sclerosing cholangitis[14] and pouchitis[15].

Previous appendicectomy is also inversely related to ulcerative colitis. A meta-analysis has shown a 69% reduction in risk of development of ulcerative colitis following appendicectomy [16]. Data are limited and at times conflicting. However, most studies have suggested that appendicectomy is protective[17].

Previous appendicectomy has been shown to increase the risk of developing Crohn's disease. A large cohort study has shown an increased risk of Crohn's disease with the exception of children undergoing the procedure under the age of 10 (in whom this was a protective factor against Crohn's disease). It was also noted that patients requiring surgery for a perforated appendix subsequently developed more severe Crohn's[18]. The mechanism by which appendicetomy affects the risk of ulcerative colitis and Crohn's disease is not clear.

There is some evidence that the use of oral contraceptives may increase the risk of developing IBD. A metanalysis has shown an association between oral contraceptive use and ulcerative colitis and Crohn's disease (RR for Crohn's disease 1.46 (95% CI 1.26-1.70), RR for UC 1.28 (95% CI 1.06-1.54))[19]. The mechanism of this association is not known.

There have been many observational, population based and case control studies on possible dietary risk factors for IBD. However, investigation into potential dietary risk factors has been limited by difficulties interpreting results, possible recall bias, as well as results not being reproducible in some studies [7].

Several case controlled studies have reported a link between high sugar intake and IBD[20]. Dietary fat has also been linked to IBD in some epidemiological studies [21, 22]. Overall, whilst the role of dietary antigens in IBD is likely clear dietary triggers have not been identified.

1.1.4   Pathophysiology

The pathophysiology of inflammatory bowel disease is complex and as yet, not fully understood. The consensus view is that IBD develops due to a genetically determined abnormal response to commensal organisms within the gastro-intestinal tract. Furthermore, this abnormal response is affected by a number of environmental triggers[23].

There is epidemiological evidence of the role of genetics in the development of IBD. Around 10% of people with IBD report a positive family history. The concordance between monozygous twins is also moderately high[24]. A number of genetic studies have shown that single genetic mutations may result in IBD. However, not all subjects with these mutations will go on to develop IBD[25, 26]. Hence, it is more likely that IBD consists of several different phenotypes, which are affected by genetic and environmental factors to differing extents[23].

As detailed above, inflammatory bowel disease is likely due to an abnormal immune response to the commensal organisms of the gut. A number of genes that have been implicated in IBD play a role in this response. For example, genes involved in the regulation of innate and adaptive immunity including IL10, STAT3 and JAK2. CCR6 and MST1 are involved in regulating inflammation and genes such as ORMDL3 and IRGM are involved in the regulation of autophagy[23].

It is estimated that more than 400 species of bacteria exist within the human intestine[27]. The microbiota provides a huge antigen load and is responsible for driving

the mucosal inflammatory response. The composition of the microbiome can be affected by genetic factors[23].

Paneth cells exist within the intestinal epithelium. These secrete a number of antimicrobial peptides including α-defensins in response to inflammation[28]. Paneth cell function can, therefore, be affected by environmental factors that result in inflammation and can also affect the composition of the microbiome as a consequence of their function. Genetic risk factors for IBD such as NOD2[29] play a role in Paneth cell functioning[23].

NOD2 is also involved in the regulation of autophagy (the lysosomsal destruction of ingested pathogens)[30]. Mutations in NOD2 are also believed to result in a lack of intestinal mucosal tolerance to bacteria[31]. All these points suggest NOD2 plays a role in the interaction between bacteria and the intestinal mucosa and may be affected by environmental factors[23].

Finally, highly secretory cells such as those of the intestinal epithelium are sensitive to ER (endoplasmic reticulum) stress. This tends to be as a result of the accumulation of unfolded or mis-folded proteins. Genes involved in protecting against ER stress have also been implicated in IBD[32]. Environmental factors may affect the cellular response to ER stress.

Given the close link between the gut microbiome and the intestinal immune system, it is hypothesised that the effect of environmental factors on genetically susceptible individuals is via their effect on the microbiome[23]. Altered composition of the gut

microbiota has been identified in individuals with IBD. A proportion of patients with IBD have a reduced abundance and diversity of Bacteroidetes, with a maintenance or bloom of Proteobacteria [33] [34]. F prausnitzii levels are reduced in IBD and low levels F prausnitzii have been linked to an increased risk of post-operative recurrence in Crohn's disease[35].

In summary, the pathogenesis of IBD is a complex process in which the intestinal microbiota play a role. It is likely that IBD arises in genetically susceptible individuals when certain environmental factors are present.

1.1.5 Clinical features

The clinical features of both ulcerative colitis and Crohn's disease are determined by disease location and in the case of Crohn's disease, disease behaviour.

The presence of blood with or without mucous in the stools is the hallmark feature of ulcerative colitis. Onset of symptoms tends to be insidious and the clinical course of UC is that of periods of relapse and remission. Active disease of the rectum is also associated with symptoms such as urgency and tenesmus. Inflammation proximal to this may also lead to symptoms such as chronic diarrhoea and abdominal pain[36]. The prognosis of ulcerative colitis has been shown to be good over the first 10 years following diagnosis with a low risk of colectomy[37].

Crohn's disease may present in a number of ways, although the most common presentation is that of chronic diarrhoea [38]. Other features include abdominal pain,

weight loss and anaemia. Crohn's disease may also present acutely with severe abdominal pain due to acute terminal ileitis[39].

The transmural nature of the inflammation associated with Crohn's disease means it may be complicated by strictures, abscesses or fistulating disease. The presence of perianal disease, a young age at diagnosis and the need for steroids at presentation have been shown to be associated with a high risk of disabling disease within five years of diagnosis[40].

Inflammatory bowel disease may be complicated by extra-intestinal disorders. These can be classified as either reactive manifestations or co-existing autoimmune disorders. Reactive manifestations are associated with active inflammatory bowel disease and include conditions such as arthropathy, uveitis and the cutaneous conditions erythema nodosum and pyoderma gangrenosum. The increased incidence of co-existing autoimmune disorders such as ankylosing spondylitis, autoimmune thyroid disease and alopecia reflects a common susceptibility to autoimmune disorders. Only ankylosing spondylitis is specific to IBD and activity of co-exiting autoimmune disorders do not tend to be linked to activity of IBD[41].

In addition to the reactive skin manifestations of erythema nodosum and pyoderma gangrenosum, other dermatological conditions associated with IBD include vitiligo and psoriasis[42]. Cutaneous adverse effects of medication may also occur and anti-TNF antibody induced psoriasis is a rare side effect of treatment[43].

A number of hepatobiliary conditions are associated with IBD. Typically, these are not related to IBD activity. Primary sclerosing cholangitis (PSC) is the most common hepatobiliary disorder associated with IBD [44]. Small duct PSC[45] and autoimmune hepatitis/primary sclerosing cholangitis overlap syndrome [46] have also been described in patients with inflammatory bowel disease.

In addition to the hepatobiliary disorders described above, cholelithiasis is more common in patients with Crohn's disease[47] and portal vein thrombosis has been described as a complication of IBD [48]. Abnormal liver function tests may also be related to medication and many medications used to treat IBD are associated with a risk of hepatotoxicity[49].

Severely active colitis may be complicated by haemorrhage, toxic megacolon and perforation. Patients with IBD are also at higher risk of thromboembolic complications, particularly whilst their disease is active and in cases of pancolitis [50].

Patients with ulcerative colitis are at increased risk of colorectal cancer compared to the general population. Reported risk estimates are variable and, on balance, it if felt that whilst ulcerative colitis does confer and increased risk of colonic cancer, the risk is probably lower than originally thought[51].

The risk of colorectal cancer is affected by disease duration with a cumulative risk of 2% at 10 years, 8% at 20 years and 18% at 30 years reported. Risk is also affected by disease extent with extensive colitis associated with the highest risk and distal colitis with an intermediate risk. Proctitis does not increase the risk of colorectal cancer

development[52]. The presence of co-existing primary sclerosing cholangitis (PSC) and a family history of colonic cancer are also associated with an increased risk[53].

1.1.6 Management

The management of inflammatory bowel disease requires coordinated input from a multi-disciplinary team including gastroenterologists and colorectal surgeons as well as dietitians and specialist nurses. Medical therapies remain central to the management of IBD. Whilst long established treatments such as steroids still have a role in treatment, the development of new therapies such as biological agents has greatly improved the management of IBD and its complications.

1.1.6.1 Medical management

Mesalazine (5-aminosalicylic acid, 5-ASA) acts on epithelial cells affecting the release of pro-inflammatory factors such as cytokines as well as inflammatory cells. There are a number of oral forms of mesalazine. pH dependent release and time-controlled release preparations exist. A multimatrix delivery system has also been developed. Drugs such as balsalazide and olsalazine consist of mesalazine bound to a carrier molecule. Mesalazine is then released due to the action of bacterial enzymes within the colon[54]. Mesalazine can also be delivered topically in the form of enemas or suppositories.

Whilst 5-ASA can be used to induce remission in ulcerative colitis, its main role is in the maintenance of remission [3]. Long term use of 5-ASAs has been shown to reduce the

risk of colorectal cancer[55] and should therefore be considered in all patients, particularly with extensive disease.

5-ASAs have not been shown to produce a clinically significant improvement in active Crohn's disease[56], nor have they been shown to be of benefit in maintaining remission[57]. Therefore, the use of 5-ASAs as first line therapy is not recommended in Crohn's disease [3]

Corticosteroids are of use in moderate and severe flare-ups of both Crohn's disease and ulcerative colitis. They are not recommended for maintenance of remission. Steroids may be administered parenterally, orally or topically[3].

Oral prednisolone has been shown to be effective in inducing remission in ulcerative colitis[58]. It is recommended in patients with moderately active UC that has not responded to mesalazine. It is recommended that a gradually reducing course is given over an 8-week period in order to reduce the risk of early relapse[3]. Budesonide, a poorly absorbed corticosteroid with less systemic side effects, has been shown to be as effective as prednisolone in the treatment of mild-moderate ulcerative colitis[59]. Topical steroids can be used in UC, but have been shown to be less effective than topical 5-ASA preparations[60].

Prednisolone has been shown to be effective in inducing remission in Crohn's disease[61]. However, the majority of patients do not remain in sustained remission following their first course of steroids[62]. Although slightly less effective than

prednisolone, budesonide is an alternative corticosteroid in cases of moderately active ileocaecal Crohn's disease in view of the favourable side effect profile[3].

The thiopurines: azathioprine (AZA) and 6-mercaptopurine (6MP) are used in ulcerative colitis and Crohn's disease. Azathioprine is metabolised to 6-mercaptopurine, which is subsequently metabolised to the 6-thioguanine nucleotides (6-TGN). 6-TGN is the active end metabolite of AZA and 6MP and acts by inducing T-cell apoptosis[63].

Azathioprine should be considered in steroid dependent ulcerative colitis or when mesalazine has failed or cannot be tolerated[3]. However, the evidence to support the use of thiopurines in UC is not of high quality and further research is needed[64]. The evidence to support the use of thiopurines in Crohn's disease is stronger and it has been shown to be an effective treatment for both inducing and maintaining remission [65, 66]. Thiopurines should also be considered in steroid dependent or frequently relapsing Crohn's disease.

Methotrexate is a cytotoxic agent with anti-inflammatory properties. It has been shown to be effective in Crohn's disease for both induction and maintenance of remission [67] [68]. It may, therefore, be considered as a second line agent for the treatment of Crohn's disease in patients unresponsive or intolerant of thiopurines. Evidence to support the use of methotrexate in UC is not as strong. However, methotrexate has been shown to be effective in patients unresponsive or intolerant to thiopurines[69]. For this reason methotrexate may also be considered as a second line therapy for maintaining remission in ulcerative colitis[3].

Ciclosporin is a calcinurin inhibitor that may be used as a rescue therapy (to avoid or defer colectomy) in acute severe ulcerative colitis, when standard treatment has failed. Current National Institute for Health and Care Excellence (NICE) guidelines recommend the use of ciclosporin in these cases, with infliximab to be used when ciclosporin is contraindicated[70]. However, whist cyclosporin has been shown to be effective in the short term, there are concerns with regard to toxicity and long-term effectiveness[71]. The Comparison of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis Trial (CONSTRUCT) is a large randomised controlled trial aimed at comparing ciclosporin and infliximab in acute severe UC[72]. Ciclosporin has not been shown to be of use in Crohn's disease[73].

A number of biological agents have been developed for use in inflammatory bowel disease. Biological therapies are antibodies able to block inflammatory pathways. The anti-TNF agents, infliximab and adalimumab were the first biological agents to be established in IBD clinical practice. Infliximab is a chimeric antibody consisting of murine and human components and is administered via an intravenous infusion. Adalimumab is a fully humanised antibody and is administered via subcutaneous injection[3]. More recently, golimumab has been developed. This is also a fully humanised anti-TNF antibody treatment, administered subcutaneously on a four weekly basis[74].

There is good evidence to support the use of infliximab and adalimumab in the treatment of Crohn's disease. Infliximab has been shown to be effective in Crohn's disease unresponsive to standard management[75]. It has subsequently been shown to maintain remission in patients responsive to initial treatment[76]. Adalimumab has been shown to be effective both in patients that have not received previous biological agents[77] and

in those who failed treatment with infliximab[78]. It has also been shown to be effective in maintaining remission[79]. Infliximab and adalimumab have also been shown to improve fistulating Crohn's disease, although improvement in non-perianal fistulating disease was modest [80] [81].

Infliximab significantly reduces the 90-day colectomy rate in acute severe ulcerative colitis[82]. It has also been shown to be effective in moderately active ulcerative colitis that has not responded to standard treatment[83]. Adalimumab has also been shown to be effective in inducing and maintaining remission in ulcerative colitis[84].

Golimumab has also been shown to be effective in inducing remission and as maintenance therapy for ulcerative colitis[85, 86] An open label, phase 4 trial in the UK has also confirmed it is both efficacious and safe for the treatment of moderate to severe UC[87].

Infliximab, adalimumab and golimumab have been approved by NICE for use in moderate to severe ulcerative colitis, following failure of conventional therapy or if such therapies are not tolerated or contraindicated [74]. Infliximab is recommended for patients with steroid refractory acute severe colitis in whom ciclosporin is contraindicated[70]. Infliximab and adalimumab have been approved for use in severely active Crohn's disease and infliximab has been approved for use in fistulating Crohn's disease[88].

Vedolizumab is an α4β7 integrin inhibitor. It blocks α4β7 integrins on gut-specific leukocytes, preventing the infiltration of leukocytes into the gastrointestinal submucosa. It is administered by intravenous infusion. It is effective in inducing and maintaining remission in ulcerative colitis [89] and Crohn's disease [90]. It is also approved by NICE

for use in moderately to severely active ulcerative colitis[91] and in moderately to severely active Crohn's disease if an anti TNF treatment has failed, is not tolerated or is contraindicated[92].

Ustekinumab is a fully humanised monoclonal antibody that blocks interleukin-12 (IL-12) and interleukin-23 (IL-23). This therefore inhibits IL-12 and IL-23 mediated cell signalling, activation and cytokine production [93]. It is administered by intravenous loading dose and then by subcutaneous injection for maintenance. It has been shown to be efficacious in the induction of remission and for maintenance therapy in Crohn's disease[94]It is approved by NICE, for use in moderately to severely active Crohn's disease if there has been loss of response to, intolerance of, or contraindications to conventional or anti-TNF treatment [95].

Tofacitinib is an oral, small molecule Janus kinase (JAK) inhibitor. It inhibits all Janus kinases, but preferentially affects JAK1 and JAK3. Phase 3 trials have confirmed tofacitinib to be effective at inducing and maintaining remission, in patients with ulcerative colitis compared to placebo[96]. It has recently been approved by NICE for use in moderately to severely active ulcerative colitis when conventional treatment or biological treatment cannot be tolerated, or if there inadequate or loss of response to these treatments[97].

There is some evidence to support the use of the antibiotics ciprofloxacin and metronidazole as disease modifying agents in Crohn's disease. These antibiotics have a clear role in the management of perianal sepsis in Crohn's disease. However, they have also been shown to improve or induce remission in a proportion of patients with perianal

fistulae. Although not statistically significant, ciprofloxacin appears to be more effective and better tolerated than metronidazole [98]. Metronidazole or ciprofloxacin are recommended as first line therapy for pouchitis (inflammation of the pouch following an ileo-anal pouch procedure)[3]. There is no evidence to support the use of antibiotics as a disease-modifying agent in ulcerative colitis.

Nutrition is an important aspect of IBD care, not least due to the high rates of malnutrition in patients with Crohn's disease and ulcerative colitis. There is also evidence to suggest that an exclusive liquid polymeric diet reduces the inflammatory response and induces remission in Crohn's disease[99]. Exclusive enteral nutrition is not of benefit in ulcerative colitis. The probiotic VSL3 is effective in maintaining remission in pouchitis[100]. There is also some evidence to support its use in ulcerative colitis[101] although this approach is not widely used.

1.1.6.2   Surgical management

Surgery is an important aspect of IBD care and may be required both as an emergency and electively. Whilst the use of "rescue therapy" in acute severe ulcerative colitis may defer the need for emergency colectomy, many patients will need surgery in the future. Sub-total colectomy and end ileostomy is required in acute severe colitis, whilst an ileo-anal pouch procedure (IPAA) should be considered in elective surgery[3].

The risk of requiring surgery for Crohn's disease is increased over time following diagnosis. Surgery is most likely in the case of ileocaecal disease[102]. A large proportion

of these patients will develop recurrent disease following surgery although will not necessarily be symptomatic[103].

1.1.7   Goals of management

The goal of treating IBD is to induce and maintain remission. However, there is no fully validated or internationally agreed definition of remission. The definition of remission depends on the context in which it is being used. For example, remission in clinical trials is often defined on the basis of disease activity scores, whereas remission at a patient level may be defined as improvement of symptoms and better quality of life[104].

Standard management of IBD centres on control of symptoms with stepwise introduction of increasingly potent medical therapy. However, there is increasing evidence that this approach does not improve long term outcomes[105].

Whilst control of symptoms is of great importance to patients and will affect health related quality of life, it has been proposed that new targets for therapy are required in order to prevent long-term bowel damage and complications. This "treat-to-target" approach is based on the management of other inflammatory diseases such as rheumatoid arthritis in which inflammation in treated aggressively in order to prevent joint destruction. Mucosal healing, as defined by disappearance of ulceration, has been shown to be an appropriate target for treatment in order to avoid long-term complications [106]

1.1.8   Disease activity assessment options

The importance of achieving mucosal healing and deep sustained remission are increasingly recognised.   Measuring disease activity, therefore, requires several investigation modalities as well as more traditional measures such as the disease activity indices. Whilst traditional measures such as disease activity indices play a role, there is increasing use of endoscopy and cross-sectional imaging as well as non-invasive biomarkers such as faecal calprotectin.

1.1.8.1   Assessment of disease extent, activity and complications

Ileo-colonoscopy remains the gold standard test for the diagnosis of inflammatory bowel disease as well as for the assessment of disease activity and extent. Its use is of increasing importance with regard to the need to assess mucosal healing and is often a key investigation when initiating or withdrawing treatments such as biological agents. Whilst ileo-colonoscopy has excellent diagnostic properties it is an invasive investigation with some associated risks. It will also miss non-intestinal pathology [107].

Endoscopic scores have been developed in order to standardise the reporting of endoscopic findings. The Crohn's Disease Endoscopic Index of Severity (CDEIS) is a complicated scoring system which is mainly used in clinical trials[108]. The Simple Endoscopic Score for Crohn's Disease (SES-CD) is a simplified score, suitable for routine clinical practice[109] and correlates well with the CDEIS[110]. The Rutgeerts score can be used to grade post operative recurrence of the distal ileum and has been shown to predict clinical recurrence[103].

A number of endoscopic scores exist for use in UC and a systematic review has recently been undertaken[111]. Both the endoscopic component of the Mayo score[112] and the UC endoscopic index of severity[113] have been shown to be reliable scoring systems.

Small bowel capsule endoscopy (SBCE) may be used for the diagnosis of Crohn's disease. However, whilst diagnostic yield is high, detected changes may not be specific to Crohn's disease and there is no means of obtaining histological specimens [107].

Cross sectional imaging also plays a role in the diagnosis and assessment of inflammatory bowel disease. Ultrasound (US) is quick, cheap and does not involve radiation. It is of use for examining the colon and terminal ileum. Results may be affected by factors such as obesity and are dependent on operator expertise[114].

Computed Tomography (CT), on the other hand, is a well-established and readily available investigation modality. CT has a high sensitivity and specificity for detecting intestinal Crohn's disease [115]. It also allows the detection of extra-intestinal complications such as collections. The significant disadvantage of CT is the associated radiation dose, which is of importance given the young age of some patients as well as the potential need for repeated imaging.

CT is also widely used in the management of ulcerative colitis, again to determine disease extent and detect complications such as perforation. Although some studies have shown correlation of CT features of colitis with endoscopic severity scores [116, 117], endoscopy

remains the first line investigation for assessing disease activity in UC. CT should be utilised in cases such as impassable strictures or when complications are suspected[118].

Magnetic resonance imaging (MRI) of the small bowel is now becoming more available. MRI allows visualization of the GI tract without overlapping bowel loops. It also allows real time functional imaging as well as detection of extra-intestinal complications. More importantly, MRI does not involve the use of radiation[119]. Whilst CT and MRI enterography have comparable sensitivity and specificity for diagnosing Crohn's disease [115], MRI pelvis is the modality of choice for assessing pelvic disease [107]. MR colonography has been reported [120], but is not widely used.

Biomarkers are also of use for measuring disease activity of inflammatory bowel disease and have the benefit of being non-invasive and therefore more acceptable to patients. C-reactive protein (CRP) is an acute phase protein and a routinely available blood test. It can be used as a measure of disease activity. In the case of Crohn's disease, it is particularly useful in patients with high levels of CRP at diagnosis. In these patients it is a sensitive marker of disease activity as well as a predictor of relapse [121].

More recently, interest has turned to the use of faecal biomarkers such as calprotectin and lactoferrin. Both have been shown to be able to detect colonic and ileo-colonic disease, but not Crohn's disease confined to the small bowel[122]. Serial measurement of faecal calprotectin can be used to monitor disease activity. A meta-analysis has shown it to be able to predict disease relapse in patients with quiescent disease[123]. Faecal biomarker levels have also been shown to return to normal in patients with endoscopic response to treatment[124].

In summary, whilst ileo-colonoscopy is the gold standard investigation for the diagnosis of Crohn's disease, other investigation modalities are important particularly for detecting complications such as collections or fistulae. There is increasing evidence to support the use of non-invasive biomarkers. Investigation selection should be based on individual patient need.

1.1.8.2   Disease activity indices

A number of disease activity scores exist for use in IBD. These measure clinical and or endoscopic data. It has been acknowledged that a simple disease activity index for use in routine clinical care is need. However, with the exception of the Truelove and Witts score for ulcerative colitis [125], these activity indices are not used routinely and their main use is in clinical trials [126].

Instruments used in ulcerative colitis:

Disease activity indices used in ulcerative colitis can be further classified as those based on clinical and biochemical parameters, endoscopic parameters, or a combination of both.

The Truelove and Witts Severity Index was first described in 1955. It is composed of 6 variables: number of stools per day, blood in stools, temperature, pulse, haemoglobin and erythrocyte sedimentation rate (ESR). Remission was defined as: 1-2 stools per day, lack of fever and tachycardia, normal (or returning to normal) haemoglobin and ESR and

gaining weight[125]. This disease activity index is not quantitative, ie it does not produce a score. Their definition of remission has not been validated.

The Powell-Tuck Index consists of 10 variables and scores range from 0-20. Remission is defined as a score of 0 and improvement an increase in score by 2 or more above baseline[127].

The Simple Clinical Colitis Activity index incorporates 6 variables: bowel frequency during the day, bowel frequency during the night, urgency of defaecation, blood in the stools, general well-being and extra-colonic manifestations. Scores range from 0-19[128]. A cut off value of less than 2.5 correlates with patient defined remission and a reduction in score by 1.5 points or more correlates with patient defined significant improvement[129].

Endoscopic disease activity indices have been subject to a recent systematic review. Many of these indices have not been validated. However, the endoscopic component of the Mayo score as well as the UC endoscopic index of severity (UCEIS) have been shown to be valid[111].

The endoscopic component of the Mayo score describes the rectal mucosa using a 4-point scale (0-inactive disease, 1-mild disease erythema loss of vascular pattern mildly friable, 2-moderate disease marked erythema absent vascular pattern marked friability, 3-severe disease spontaneous bleeding ulceration [112]. It has been shown to be both reliable and responsive[130]. It has also been shown to correlate with health related quality of life[131].

The UC endoscopic index of severity consists of 3 components: vascular pattern, bleeding and erosion/ulceration. Potential scores range from 0 to 11[113]. The UCEIS has been shown to have favourable psychometric characteristics. Score components have been shown to correlate with a global rating of endoscopic severity. The UCEIS is also to have satisfactory intraobserver and interobserver reliability[132].

A number of disease activity indices used in UC incorporate clinical and endoscopic data. These include the Mayo score and the Sutherland Index.

The Mayo score consists of 4 variables: stool frequency, rectal bleeding, sigmoidoscopy findings and a physician's global assessment (PGA). Scores range from 0 to 12. In addition to these variables a patient functional assessment is determined in order to inform the physician's global assessment. Complete response is defined as: normal stool frequency, no rectal bleeding, patient generally well and normal endoscopic findings. The PGA score must also be 0. A partial response is defined as an improvement in the PGA score in association with an improvement in at least one clinical parameter and an absence of deterioration in any clinical parameter. The Mayo score and its definitions have not been validated[112]. These definitions have been modified in subsequent trials. Clinical remission as defined by a score of 2 or less and no individual sub-score of more than one has been shown to correlate with significant improvement in health related quality of life[133].

Finally, the Disease Activity Index (Sutherland Index) consists of 4 variables: stool frequency, rectal bleeding, mucosal appearance and physician's rating of disease activity. Scores range from 0 to 12 [134]. The Sutherland score has not been validated. However,

a score of less than 2.5 points has been shown to correlate with patient defined remission[129].

Instruments used in Crohn's disease:

The Crohn's disease activity index (CDAI) consists of eight variables: number of liquid stools, extent of abdominal pain, general wellbeing, extra-intestinal symptoms, the need for anti-diarrhoeal drugs, abdominal masses, haematocrit and body weight. The variables are weighted and scores range from 0 to approximately 600. This instrument has been validated as a measure of disease activity in Crohn's disease. A cut off value of 150 has been shown to define remission with 90% of patients with scores below this being rated as "very well" by assessing physicians. A cut off value of 450 or more is consistent with severely active disease[135]. Of note, parts of the CDAI score are based on a symptom diary covering the preceding 7 days and therefore, the CDAI is not amenable to use in day-to-day clinical care.

The Harvey Bradshaw index is a simpler disease measure consisting of 5 variables: general well being, abdominal pain, number of liquid stools daily, the presence of abdominal mass and the presence of complications. A score of less than 5 is consistent with remission, 5 to 7-mild disease, 8 to 16-moderate disease and more than 16-severe disease[136]. The Harvey Bradshaw index is not designed to be completed in full by the patients since the variable "abdominal mass" requires clinical examination.

Other clinical indices include the Organization Mondial de Gastro-enterologie (OMGE) index[137] and the Cape Town index[138]. These indices have been shown to correlate with each other [139].

The CDAI is not an appropriate index for use in patients with active fistulating disease since the presence of fistulae contributes a relatively small number of points to the overall score[140]. The perianal disease activity index (PDAI) has been designed to measure disease activity in patients with perianal Crohn's disease. There are 5 variables: discharge, pain, restriction of sexual activity, type of perianal disease and degree of induration[141]. At present there are no identified cut off values.

1.2 Patient reported outcome measures

A patient reported outcome measure (PROM) may be defined as a series of questions that patients are asked in order to measure their views on their own health. A PROM is designed to measure health or health related quality of life, rather than satisfaction or experience of health care[1]

Traditionally, measurement of quality of healthcare has relied on clinical measures such as complication rates or mortality. However, it has been increasing recognised that since our primary aim is to improve our patients' health, measuring change in health from our patients' perspective is an important outcome measure in clinical trials and an important measure of quality within health systems.

1.2.1 Content of patient reported outcome measures

PROMs measure the patient's subjective experience of health and/or the consequences of illness. The content of PROMs is varied and an individual PROM may measure one or several aspects of health status. Content may be grouped into dimensions. Commonly

included dimensions include physical functioning (eg physical activity, activities of daily living), specific symptoms (eg pain, fatigue), overall assessment of health, psychological wellbeing (eg anxiety, depression), social wellbeing (eg social contact, undertaking leisure activities), cognitive functioning (eg concentration, memory), role activities (employment), personal constructs (eg life satisfaction) and satisfaction with health care (Figure 3).

Figure 3 Dimensions assessed by patient reported outcome measures[142]

Dimension	Example
Physical function	Mobility, physical activity, ADLs
Symptoms	Pain, fatigue, diarrhoea
Global judgements of health	
Psychological well-being	Anxiety, depression, self-esteem
Social well-being	Family relations, leisure activities, social contact
Cognition	Alertness, concentration, memory
Role activities	Employment, household management
Personal constructs	Life satisfaction, stigma,
Satisfaction with care	

ADLs-activities of daily living

1.2.2　　　Types of instrument

There are a number of different types of patient reported outcome measures. Most can generally be described as either "disease specific" or "generic". Disease specific instruments are designed for use in a specific condition. For example, the inflammatory

bowel disease questionnaire (IBDQ) is an IBD specific measure of health related quality of life[143]. The most important advantage of using a disease specific measure is that it is more likely to detect clinical important change over serial measurements[144]. However, disease specific instruments cannot be used to compare health status in respondents with different conditions or with a healthy general population sample[142].

Generic measures are instruments that have been intentionally designed to measure several aspects of health status and can therefore be applied to many different diseases. They may also be used to compare treatment effects between different groups of patients. Generic measures can also be used to measure health status in general population samples and therefore generate "normal values" to which study data can be compared[142]. An example of a generic instrument that has been validated for use in IBD is the SF-36 questionnaire. The SF-36 consists of 8 dimensions covering broad areas such as "physical functioning", "social functioning" and "energy/vitality"[145].

Generic measures are more likely to detect unexpected changes in health status. However, they do lack sensitivity and therefore are less likely to be able to detect clinically important change over time [142].

Another important and distinct type of PROM is the utility measure. Utility measures are used to measure an individual's preference for particular health outcomes[146]. Utility measures may employ techniques such as standard gamble or time trade off in order to obtain information on respondent preferences. Standard gamble technique involves providing the individual with the option of a certain health state or a gamble that the health state may be better or worse. The respondent is asked for the probability of a

better outcome that would make them indifferent between staying in the certain health state or gambling on the alternative state. The time trade off technique asks the individual the amount of time (eg life years) they would give up in order to achieve a better health state[147].

Alternatively, utilities may be derived via an indirect method. An individual's reported health measure is obtained in the usual manner (for example a questionnaire). Scores are then weighted using valuations previously obtained from an appropriate population sample to produce a utility index. An example of this is the EuroQol: EQ-5D questionnaire[148].

Utility measures have a number of advantages. They provide a single value, rather than a number of responses, that reflects the respondent's preferences or values regarding their health. Utility measures can be used to generate quality-adjusted life years (QALYs) and can be used for use in cost-utility analysis.

Utility measures involving methods such as standard gamble and time trade off technique are often time consuming and poorly understood by respondents. Data is therefore often obtained via interview rather than self completed questionnaire. This increases the burden both on participant and researcher[142]. The use of indirect methods to derive utility measures, such as with the EuroQol questionnaire, avoids these problems[148].

Disease–specific and generic PROMs are associated with advantages and disadvantages. It is recommended that both generic and disease-specific instruments are included in trials. This allows the collection of data relevant to the condition under investigation and

detection of clinically important change, whilst ensuring unexpected effects are also identified [149].

1.2.3   Criteria for the assessment of PROMs

When assessing a patient reported outcome measure the following properties should be considered: appropriateness, reliability, validity, responsiveness, precision, interpretability, acceptability and feasibility[142].

Appropriateness refers to whether the instrument has the appropriate content to address the aims of the trial in which it will be used. Both the patient group and the content of potential instruments must be assessed in order to ensure the most appropriate instrument is chosen [149]. A further means of ensuring appropriate outcome measures are used is to include at least one disease specific measure and one generic measure[150]. As detailed above, this strategy is to increase the likelihood of capturing anticipated effects as well as unanticipated effects [142].

Reliability refers to whether an instrument is free from random error. There are two components of reliability: reproducibility and internal consistency.

Reproducibility assesses whether similar results are obtained, on repeated testing, in a stable or unchanged group of participants. It is measured by test-retest reliability - the amount of correlation between initial score and a subsequent assessment. Most investigators will choose a time frame of 2 -14 days between tests. This is trade-off

between the risk of recollection bias and the potential for change in clinical condition.[151]

Although often used, correlation coefficients have the potential to exaggerate reproducibility. Therefore, measurement of intra-class correlation coefficients (ICC) is more appropriate. ICC determines the amount of variability that is due to true differences in individuals vs the amount due to variability in measurement [142].

Internal consistency can also be used to measure the reliability of questionnaires in which a number of items are used to measure the same construct. Internal consistency refers to whether individual items measuring the same construct correlate. This inter-correlation of items within a domain is commonly determined using Cronbach alpha statistic. Whilst a high Cronbach alpha statistic signifies internal consistency, a very high alpha statistic may indicate redundancy of some items[142]. Therefore, it is recommended that Cronbach alpha values should be over 0.7, but no more than 0.9 [152].

Validity is an assessment of whether an instrument measures the construct it was designed to measure. It can be categorised further as: content validity, face validity and construct validity.

Face and content validity are qualitative measures. Face validity is an assessment of whether an instrument appears to be measuring the construct in question. Content validity refers to whether all components of the subject to be tested are adequately covered by the instrument[153].

Construct validity is a quantitative measure of whether an instrument is able to measure what it is designed to measure. It involves correlating instrument scores with scores from other measures that are known to or hypothesised to measure the construct in question[142]. Rather than a single observation, several variables should be used to prove construct validity[154].

Construct validity may also be measured using convergent and discriminant validity [155]. This requires hypothesising that a measure will correlate strongly with some variables and weakly with others. Correlations are expected to be strongest with closely related variables and weakest with those not closely related[142].

An instrument must also possess responsiveness - the ability to detect changes over time. Responsiveness can be assessed in a number of ways. The first is the calculation of effect size. Effect size is calculated by dividing the change in mean score from baseline to follow up, by the standard deviation of the baseline scores. Effect sizes can be used for assessing relative size of change. An effect size of 0.2 is considered small where as an effect size of 0.5 and more than 0.8 are considered medium and large respectively[156].

Alternatively, the standardised response mean can be calculated. This differs from effect size in that the denominator is the standard deviation of the change in scores[157]. Furthermore, for calculation of the modified standardized response mean, the denominator is the standard deviation of the change in scores in stable patients[158].

Receiver-operating characteristics (ROC) can also be used to assess responsiveness. This method requires the existence of a gold standard test for detecting whether a true change

in clinical condition has occurred. Sensitivity (the true positive rate) is plotted against 1-specificity (the false positive rate) to produce a ROC curve. Of note, calculation of sensitivity and specificity of an instrument requires identification of a change in score (between two visits) that is considered a "positive result" i.e. indicates a change in clinical condition.

The area under the ROC curve indicates the instruments responsiveness. If ROC curves are plotted for variously selected cut off scores, the optimal cut off score can be identified[159].

Floor and ceiling effects may affect the responsiveness of a questionnaire. These occur when questionnaire design or wording results in respondents not being able to report improved health state (ceiling effect) of worsened health state (floor effect).

Responsiveness may also be affected by the distribution of baseline scores. If most items within a domain are easily obtainable, then a large change in score could occur despite only a small real improvement[142]

Precision is the ability of a measure to make a number of distinctions. Hence, an instrument that can only distinguish between two health states lacks precision, whereas one which can distinguish between many health states is precise [160].

Interpretability refers to how meaningful the score obtained from an instrument is. Minimal clinically important difference (MCID) may be used to assess interpretability.

A PROM must also possess acceptability, ie is it acceptable to patients? Acceptability is important in order to ensure high completion rates. Acceptability may be proven by measuring completion rates and the time to complete questionnaires [142].

Finally, feasibility must also be considered. This refers to the burden on clinical staff and researchers. For example, directly administering a questionnaire will have a higher burden on staff than a postal questionnaire.[142]

1.3 PROMs and the National Health Service

There has been a significant shift in approach to health policy over the last few years. The government report, "High Quality Care for All", by Lord Darzi was published in 2008. Based on a year long review of the NHS, the report sets out the government's aim to improve quality of care. It is an extensive report covering many aspects of the NHS. However, it is the measurement of healthcare quality that is most relevant to PROMs. By improving the ability to measure quality within the NHS, it is felt that staff will have the means to develop quality improvement strategies.

The report defines quality of care in terms of patient safety, patient experience and effectiveness of care. The report recommends that effectiveness of care should be measured in a number of ways including traditional indicators such as complication rates and survival rates. However, it also states that patient reported outcome measures should also be used in order to ensure quality of healthcare is measured from the patients' perspective [161].

A further white paper, "Equity and excellence: liberating the NHS" was published in 2010. This document set out proposed changes to the NHS including improving patient choice, patient involvement in decision-making and the empowerment of clinicians to make decisions at a local level rather than being managed by government. There was also emphasis on improving healthcare quality. Again, the need to capture patient reported data was recognised as well as the need for improved data reporting and transparency[162].

Following on from this the NHS Outcomes Framework was developed in order to provide a national overview of healthcare performance. It is the primary accountability mechanism between the government and the NHS and one of the strategies to improving healthcare quality.

Indicators fall within five domains of outcomes that the NHS should aim to improve. These domains are: preventing people for dying prematurely, enhancing quality of life in people with long term conditions, helping people to recover from episodes of ill health or following injury, ensuring that people have a positive experience of care, and treating and caring for people in a safe environment and protecting then from avoidable harm. Each domain contains a small number of "overarching indicators" as well as a number of improvement areas. For example, within the NHS Outcomes Framework 2014/15 domain "Enhancing quality of life for people with long term conditions", the overarching indicator is "Health-related quality of life for people with long term conditions". Improvement areas include: "ensuring people feel supported to manage their condition" and "improving functional ability in people with long term conditions"[163]

The "Mandate to the NHS" is published by the government each year. The objectives within the mandate are set around these five domains and therefore, the NHS Outcomes Framework is used to assess whether the objectives have been met [163].

Routine collection of patient reported outcome measures has been mandatory within NHS England since 2009[164]. At present, all hospitals undertaking elective hernia, knee, hip and varicose vein surgery are collecting PROMs data. There is ongoing work in order to extend the PROMs programme to cover more conditions such as mental health, cancer and long term conditions such as asthma[1].

Utility measures, a form of patient reported outcome measures, are used for cost-effectiveness analysis. The National Institute for Health and Care Excellence (NICE) is responsible for the appraisal of healthcare interventions. Recommendations are based on both clinical evidence and cost-effectiveness analysis.

Quality adjusted life years (QALYs) are the preferred outcome measure of cost effectiveness used by NICE for health intervention appraisal. QALYs combine data on length of life and the values placed on various levels of quality of life. The lowest possible value of a QALY is 0 (health rating is as bad as being dead) and the highest is 1 (good health for 1 year). The EQ-5D utility index is the preferred utility index for calculation of QALYs [165]. When comparing a new treatment to an existing one, the incremental cost-effectiveness ratio can be calculated which reflects the extra cost associated with one gained QALY [166].

"Cost per QALY" values are used by NICE to help assess whether an intervention is good value for money and therefore to be recommended. If an intervention is estimated to cost less than £20000 per QALY it is judged to be cost effective. If over £20000, other evidence such as the whether there is a possibility that quality of life data has been mis-captured, should be considered before recommendations are made. If an intervention costs more than £30000 per QALY, there must be increasingly strong evidence to support the use of the treatment in order for it to be recommended[167]. Whilst these thresholds are not backed by evidence, they are a means for NICE to use economic evidence to make decisions with regard to limited resource allocation[1].

1.3.1 The use of PROMs in the management of chronic disease

Chronic diseases are becoming an increasing burden on health services and account for a large proportion of NHS spending. It is estimated that the care of patients with long term conditions accounts for 70% of the total health and social spending in England[168]. Chronic conditions do not commonly lead to death and can be managed on a long-term basis, often requiring ongoing treatment and monitoring. However, long term conditions have been shown to affect quality of life[169]. Therefore, improving or maintaining quality of life is a key goal in this patient group. The development of systems to measure quality of life of patients with chronic diseases is therefore required. As detailed above, the quality of life of patients with long term conditions is a component of the NHS Outcomes Framework[163].

The use of PROMs to measure effective management of chronic diseases is challenging. In contrast to elective surgical care, in which there is a clear intervention at a single point

in time, chronic diseases are complex, may co-exist with other disorders and may involve multiple specialities and multiple interventions over time[170].

The Department of Health commissioned a pilot study to determine the feasibility of the routine collection of patient-reported outcome measures for long-term conditions in primary care. The following 6 chronic conditions were included: asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure and stroke.[170]. The study was conducted by post in general practice and included around 4500 participants. The primary aim of the study was to investigate the feasibility of collecting PROMs in long-term conditions. The change in EQ-5D scores between visits was a secondary endpoint.

The questionnaire pack sent to potential respondents contained a generic PROM and a disease-specific PROM (specific to their conditions). The included instruments were chosen on the basis of extensive literature review of PROMs used in chronic diseases. Overall, there was a 38.4% questionnaire return rate. 71% of those who returned the initial questionnaire agreed to complete a follow up assessment. This response rate is similar to other primary care based surveys[171] and indicates collection of PROMs in this patient group is feasible. Interestingly there was no change in health-related quality of life (as measured using the EQ-5D questionnaire) between baseline and follow-up questionnaire (one year later). This raises questions with regard to optimal follow up interval as well as instrument choice if chronic diseases are incorporated into the current NHS PROMs programme.

1.3.2 PROMs currently used in the National Health Service

A small number of patient reported outcome measures have been incorporated into the NHS Outcomes Framework and it is planned for further PROMs to be used in the future. Within domain 2, "Enhancing quality of life for people with long-term conditions", health related quality of life is measured using the EQ-5D[172]. This is obtained from the GP survey, which includes the EQ-5D questionnaire.

The EQ-5D questionnaire is a widely used generic patient reported outcome measure. It consists of 2 parts: the EQ-5D descriptive system and a visual analogue scale. The descriptive system consists of 5 dimensions: mobility, self-care, usual activities, pain and discomfort and anxiety and depression. Each response is scored between 1 and 3 (1 indicates no problem and 3 indicates severe problems). These numbers are then expressed as a 5-digit code. The visual analogue scale (VAS) is a vertical, thermometer scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). The EQ-5D descriptive system can be converted to a summary utility index using value sets[172].

Improving health related quality of life of people with dementia is an indicator within domain 2 of the NHS outcomes framework. Research is currently underway into the possibility of using a quality of life measure on a routine basis in patients with dementia. If a suitable instrument is identified this may also be incorporated in the Outcomes Framework[173].

"Total health gain as assessed by patients for elective procedures" is an indicator within domain 3 of the NHS Outcomes Framework ("Helping people to recover from episodes of ill health or following injury"). The NHS PROMs programme collects data related to elective hip and knee replacement surgery, elective hernia repair and elective varicose vein surgery. However, as the PROMs programme develops, PROMs related to additional elective procedures may be included.

The instruments currently used in the PROMs programme include the EQ-5D[172] as described above, the Aberdeen varicose vein questionnaire (AVVQ)[174], the Oxford hip score (OHS)[175] and the Oxford knee score (OKS)[176]. All patients complete the EQ-5D before and after surgery as well as the appropriate procedure specific instrument.

The Aberdeen varicose vein questionnaire consists of 13 questions related to subjects such as pain, swelling and interference with daily activities. Overall scores range from 0-100[174]. The Oxford hip and knee scores contain 12 questions related to pain, joint movement and ability to undertake daily activities. Scores range from 0-48[175, 176].

There is ongoing work with regard to the incorporation of patient reported outcome measures following elective psychological therapies, stroke and trauma into the NHS Outcomes Framework[173]

The National Cancer Survivorship Initiative (NCSI) was launched in 2007. Its aim was to improve the health and quality of life of people who had survived cancer[177].

As part of this initiative, the Department of Health undertook a study of the quality of life of cancer survivors [178]. Quality of life was measured using the EQ-5D questionnaire[172], the Social Difficulties Inventory[166] and cancer site-specific questions extracted from the relevant Functional Assessment of Cancer Therapy (FACT) questionnaires (http://www.facit.org/FACITOrg). This study provided important information with regard to the quality of life of people with cancer, which has informed recent recommendations by the NCSI. It also identified a high response to the survey indicating the routine capture of patient reported data using PROMs is both feasible and acceptable in this population.

The National Cancer Survivorship Initiative recommends the routine use of PROMs in order to collect patient reported data in many aspects of cancer care including symptoms, quality of life, patient concerns and unmet needs[177].

1.3.3 Potential uses, benefits and disadvantages of using PROMs in the NHS

The potential use for PROMs within the NHS is wide ranging. Firstly, patient reported data available to patients might allow them to make informed decisions on whether they wish to receive a certain treatment and where they would like it to be done[1]. Patient surveys have confirmed that factors such as "impact on health as a result of treatment" is an important factor when choosing a hospital[179]. It is hoped that the NHS PROMs programme will facilitate patient choice. However, a large survey of NHS patients has revealed that only a small proportion sought performance data when deciding on a hospital for treatment[180].

PROMs data will be of great use in managing clinical quality in hospitals. Data will allow local benchmarking against other hospitals as well as national targets. As detailed above, patient reported data are an important aspect of the NHS Outcomes Framework. PROMs data are likely to be incorporated into trusts' annual quality accounts[1].

PROMs are also likely to be used by health care commissioners. PROMs may be used to monitor the performance of existing care providers or provide incentives via linkage of payment to defined PROMs scores. PROMs, and specifically utility measures may also inform commissioners with regard to obtaining good value for money[1].

Finally, PROMs may be of use at the individual clinician level and may help with clinical decision-making. Serial patient reported data could be recorded in order to monitor response to interventions or treatments[181]. PROMs may be used to screen for certain conditions (for example depression) and a certain score prompt earlier follow up[182]. Completion of PROMs may serve to reassure patients that their views are of importance and may also serve to highlight issues that would otherwise not have been covered during a consultation[183].

PROMs data are subjective and are affected by a patient's view. This is the purpose of PROMs, but it highlights that PROMs should not be used in isolation when making decisions. PROMs are not a replacement for traditional clinical measures, but should be used in conjunction with them.

Conditions such as dementia, stroke, learning difficulties and illiteracy present a problem with regard to capturing patient reported data[1]. However, a number of PROMs have been designed to be completed by caregivers such as parents[184].

1.4 The use of PROMs in inflammatory bowel disease

1.4.1 Current instruments used in IBD

A number of IBD specific patient reported outcome measures have been developed[143, 185, 186], mainly for use in clinical trials. Of these, only the Inflammatory Bowel Disease Questionnaire has undergone extensive psychometric evaluation. It is the most commonly used IBD specific PROM and is a measure of health-related quality of life designed for use in clinical trials. It consists of 32 questions within 4 domains (gastro-intestinal symptoms, systemic symptoms, emotional function and social function). Response options are on a 7-point Likert scale with 1 indicating a severe problem and 7 indicating no problem. Hence a low IBDQ score indicates poor health related quality of life[143]. It has been widely validated for use in many countries including the United Kingdom[187].

The Short Inflammatory Bowel Disease Questionnaire (SIBDQ) was developed in order to provide a measure of quality of life feasible for use in day-to-day clinical practice. It is composed of 10 questions taken from the original IBDQ. Whilst this has been shown to be a valid and reliable instrument[188], it has not been widely established in routine clinical practice.

A small number of generic patient reported outcome measures have been validated for use in inflammatory bowel disease. Of these the EQ-5D is widely used generic measure of health status. The EQ-5D questionnaire has been tested in many diseases including IBD and has been shown to be valid[189, 190]. Again, the use of these generic measures has not crossed over into clinical practice.

1.4.2   The need for IBD specific PROMs for use in clinical practice

The first UK national IBD audit was undertaken in 2006. A total of 212 hospitals submitted data. This audit identified a large variation in the care provided to patients with IBD. For example, only 33% of hospitals had a dedicated gastroenterology (medical or surgical) ward and 44% of sites did not have a IBD specialist nurse[191].

Following on from this, the IBD Standards Group was formed and subsequently published "Quality Care: Service Standards for the healthcare of people who have Inflammatory Bowel Disease (IBD)" in 2009. The aim of this document was to ensure safe, effective and consistently high standards for people with IBD. This comprehensive guideline outlines six standards related to the care of patients with IBD:

High quality clinical care – the provision of high quality care via a multidisciplinary approach.

Local delivery of care – the provision of local secondary care services as well as the development of shared care protocols with general practice.

Maintaining a patient-centred service – care should be patient-centred and responsive to individual needs. Patients should be able to make choices related to their treatment.

Patient education and support – Patients should be well informed with regard to their condition and provided information on the services available to them.

Information technology and audit – A register of all patients with IBD must be kept, a database should be developed and the IBD team should participate in regular audit.

Evidence based practice and research – High level training should be provided to all members if the IBD team [192].

The Inflammatory Bowel Disease Quality Improvement Project (IBDQIP) was established in order to support the implementation of the IBD Standards. It provided a web-based self-assessment tool in order to allow local IBD teams to benchmark their services against national standards. Participation improved overtime and IBDQIP has now merged with the National IBD Audit[193].

A further two rounds of the national IBD audit have been completed and round four is in progress. There has been a significant improvement over a relatively short period of time. There has been improvement in a number of areas such as contact with an IBD nurse during admission and a reduction on mortality associated with ulcerative colitis.

However, there remains room for improvement in some areas. For example, the collection of stool samples was reduced in the most recent audit round[194].

Both the IBD audit programme and the IBD standards have driven up the quality of care provided to patients with IBD. It is hoped that the ongoing implementation of the IBD standards will continue to improve IBD care across the country. Whilst certain IBD standards can be audited using traditional methods, PROMs may have a role to play in the monitoring of these standards, particularly standards related to patient-centred care.

Additionally, PROMs may also play a role in routine clinic care. Patient reported information captured using PROMs might serve to inform physician and patient decisions. They may also have a role in monitoring response to treatment as well as aiding communication between patients and physicians, particularly with regard to identifying unmet needs.

Whilst a small number of IBD specific PROMs exist, none have become established in routine clinical care. The majority of these instruments focus on the measurement of health-related quality of life. The routine use of PROMs in clinical practice is limited by a number of factors. There is generally a lack of knowledge related to PROMs and their potential advantages. Concerns with regard to burden on the patient and clinical team are also barriers to their use.

Therefore, there is a case for developing a PROM that measures an aspect of health status relevant to day to day care of patients with IBD, rather than focusing solely on health-

related quality of life. A PROM for use in clinical care must also be acceptable to patient and physician alike as well.

Therefore, the aim of this study is to develop a patient reported outcome to measure "disease control" and for this to be rapid, user friendly and applicable to routine clinical care.

1.5   Summary

In summary, inflammatory bowel disease is a chronic disorder of the gastro-intestinal tract. It is a disease predominantly of the western world and the incidence of Crohn's disease is continuing to rise. The pathophysiology is complex and most likely due to a genetically determined abnormal immune response to the commensal organisms within the GI tract. Clinical features are varied and determined by disease type, extent and behaviour. Medical management remains a central component to IBD care and the development of new treatments such as biological agents has improved treatment options. Endoscopy remains the gold standard investigation modality, but other investigation modalities as well as disease activity indices play a role.

Patient reported outcome measures allow the measurement of health status from the patient's perspective. PROMs are increasingly recognised as an important outcome measure within clinical trials as well as quality measures for health care systems. PROMs may be generic or disease specific. Utility measures are a distinct type of PROM and are used in cost-utility analysis. It is recommended that the psychometric properties of patient reported outcome measures are determined in order to validate the instrument.

Appropriateness, reliability, validity, responsiveness, precision, interpretability, acceptability and feasibility should be reported.

Changes in health policy have resulted in the need to incorporate the use of PROMs into the National Health Service. Routine collection of patient reported data has been mandatory for NHS trusts undertaking certain elective surgical procedures since 2009. There is ongoing work to determine whether it is feasible to collect PROMs data from other patient groups such as those with chronic diseases. The reporting of PROMs data is a component of the NHS Outcomes Framework and utility measures are used by the National Institute for Health and Clinical Excellence (NICE) when assessing cost effectiveness of healthcare interventions.

The National IBD audit identified a significant variation in the quality of care received by patients with IBD across the UK. Following on from this the IBD Standards were developed in order to standardise and improve care. PROMs could play an important role in the measurement of IBD care from the patient's perspective allowing benchmarking with the IBD standards. Moreover, an IBD PROM may also serve to aid clinical decision-making and improve communication between patients and their healthcare team.

Whilst a small number of IBD specific instruments exist, the majority focus on health-related quality of life. To date, none have become established in routine clinical care. This study aims to develop and validate a novel patient reported outcome measure of IBD "control" for use in the routine care of patients with inflammatory bowel disease. We hypothesised that it should be possible to create a short, simple, pragmatic questionnaire by drawing upon the existing literature for both "generic" and "condition-specific"

PROMs and seeking patient views on content and format. Given the challenges for PROMs in gaining traction in routine settings, our design criteria would place emphasis on generating a simple tool that offers a rapid assessment of overall health status in IBD.

Chapter 2 Literature review

2.1   Introduction

There has been increasing interest in the measurement of patient reported data in IBD over the last 10-20 years. IBD specific, as well as generic patient reported outcome measures have been developed, most commonly to measure health-related quality of life[143, 185, 195]. However, instruments designed to measures other aspects of health status have also been developed[196-198]. Whilst PROMs are most commonly used within clinical research, there is a drive to bring their use into clinical practice [161]. As yet, they have not become established within this role for inflammatory bowel disease.

The primary aim of this chapter was to review current patient reported outcome measures for use in adults with IBD. A secondary aim of the literature review was to determine whether the construct of 'control' of IBD had been used in the development of previous PROMs. Our goal was to use the existing literature to guide the development of a novel instrument.

2.2   Methods

A literature search was undertaken of the PubMed database using the following search terms: ["inflammatory bowel disease" OR "IBD" OR "Crohn's disease" OR "ulcerative colitis"] AND ["patient reported outcome measure" OR "PROM" OR "control" OR "questionnaire"]. Articles published up to and including September 2013 were included. Non-English language articles were excluded. Further PubMed searches of identified

outcome measure names were undertaken. Review article reference lists were also screened for additional instruments.

For each PROM, information was extracted, if possible, on the following: instrument name, details of domains and individual items, item scoring method, and summary score system. Evidence for testing of the psychometric properties in patients with inflammatory bowel disease was extracted and summarised.

2.3   Results – description of patient reported outcome measures

A total of 10492 references were identified. 81 relevant articles were identified of which 9 related to the paediatric population and were therefore excluded. These articles related to 27 IBD specific instruments (the majority of which measure health-related quality of life) and 5 generic instruments validated for use in IBD. No instruments were focused specifically on the construct of disease 'control', confirming the novel nature of our proposed new PROM.

Inflammatory bowel disease questionnaire (IBDQ) and its adaptations

- Inflammatory Bowel Disease Questionnaire (IBDQ)[143]
- Extended Inflammatory Bowel Disease Questionnaire (IBDQ-36)[199]
- UK-IBDQ[195]
- Short Inflammatory Bowel Disease Questionnaire (SIBDQ)[188]
- IBDQ shorted to nine items (IBDQ-9)[200]

Other measures of health-related quality of life

- Rating Form of IBD Patient Concerns (RFIPC)[185]

- Short Health Scale (SHS)[201]

- Cleveland Clinic Questionnaire[202]

- Edinburgh IBD Quality of Life Questionnaire (EIBDQ)[203]

- Household Member Quality of Life Questionnaire (HHMQoL-IBD)[204]

- Social Impact of Chronic Conditions-IBD (SICC-IBD) [205]

- Cleveland Global Quality of Life Instrument (CGQL)[206]

Measures of other aspects of health status

- Health Status Scales[186]

- Burden of Symptoms "feeling thermometer"[207]

- IBD Self-efficacy Scale (IBD-SES)[197]

- Numeric Rating Scale[208]

Measures of patient satisfaction with healthcare

- Treatment Satisfaction Questionnaire for Crohn's Disease (TSQ-C)[198]

- Quality of Care Through the Patient's Eyes IBD (QUOTE-IBD)[209]

Measures of patient knowledge

- Crohn's and Colitis Knowledge Score (CCKNOW)[210]

- Knowledge Questionnaire (KQ)[211]

- Crohn's and Colitis Pregnancy Knowledge Score (CCPKnow)[212]

- Short measure of patient knowledge[213]

Measures of work, productivity and disability

- Work Productivity and Activity Impairment in Crohn's Disease (WPAI:CD)[196]

- Crohn's Disease Perceived Work Disability Questionnaire (CPWDQ)[214]

- Inflammatory Bowel Disease Disability Index[215]

- Inflammatory Bowel Disease Disability Score (IBD-DS)[216]

Generic instruments validated for use in IBD

- EuroQol (EQ-5D)[172]

- Short Form-36 (SF-36)[145]

- 15D Questionnaire[217]

- Morisky Medication Adherence Scale (MMAS-8)[218]

- Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F)[219]

2.3.1 Disease specific measures

2.3.1.1 The Inflammatory Bowel Disease Questionnaire and its adaptations

The most widely used disease-specific instrument is the inflammatory bowel disease questionnaire (IBDQ) [143]. The IBDQ measures the health-related quality of life of IBD patients. Originally developed and validated in Canada, it has also been validated in many other countries and languages[187, 220, 221]. It has also been validated for use in patients following ileal pouch anal anastomosis for ulcerative colitis[222, 223].

The instrument consists of 32 questions within 4 domains (gastro-intestinal symptoms, systemic symptoms, emotional function and social function). Response options are on a 7-point Likert scale with 1 indicating a severe problem and 7 indicating no problem. Hence a low IBDQ score indicates poor health related quality of life[143].

The IBDQ has also been modified further to produce an extended version (The IBD quality of life questionnaire)[199], the short inflammatory bowel disease questionnaire (SIBDQ)[188], the IBDQ-9 [200] and the UK-IBDQ [195].

The extended IBDQ contains 36 questions. This extended instrument was developed to be self-administered by patients with stable IBD being managed within the outpatient setting. The additional questions were derived by splitting some of the original questions into several shorter questions as well as the identification of further important concerns by a group of patients. It must be noted that although this instrument has been used to

compare controls to IBD patients, there are no data with regard to psychometric properties [199].

The short inflammatory bowel disease questionnaire (SIBDQ) was developed for use in routine clinical practice. It consists of 10 questions taken from the original 32-item IBDQ. Items for inclusion were selected using regression analysis techniques[188]. An alternative shortened form, the IBDQ-9, has also been described. Nine items for inclusion were identified using Rasch analysis of the 36 item extended IBDQ (IBDQ-36). Scores for the IBDQ-9 are obtained by adding the 9 responses together and then converting this value to a scale from 0-100[200].

The UK-IBDQ is based on the original IBDQ having been modified in order to change wording, remove questions and simplify response options. The UK IBDQ consists of 30 questions within 5 domains (gastro-intestinal I, gastrointestinal II, systemic symptoms, emotional function and social function). Response options range from 1-4 as opposed to 1-7, with low scores again indicating poor quality of life[195].

2.3.1.2 Other disease specific measures of health-related quality of life

The Rating Form for IBD Patient Concerns (RFIPC) was designed in order to measure the worries and concerns of patients with inflammatory bowel disease. The instrument contains 25 items. Responses are on a visual analogue scale (VAS) ranging from 0-100 with 0 representing "no problem" and 100 representing "a great deal". It was designed to be self-administered[185].

A number of other disease specific instruments have been developed to measure health related quality of life. However, they are not widely used. The Short Health Scale is a 4-item questionnaire originally developed for use in ulcerative colitis. The 4 items were designed to cover 4 dimensions of health status: symptom burden, social function, disease related worry and general well being. Responses are on a 0-100 mm visual analogue scale and the domain scores are presented rather than a total score[201].

The Cleveland Clinic questionnaire, another measure of quality of life, was designed to be used by patients "functioning in society" and therefore emphasis was placed on measuring effects on activities of daily living rather than medical or psychological aspects of health status. The questionnaire consists of 47 questions within 4 domains: functional/economic, social/recreational, affect/life in general and medical/symptoms. 2 questions are open ended and the remaining questions are answered on a 5-point Likert scale. This questionnaire was also used by the authors to develop a quality of life index. The index consists of 18 questions selected from the original questionnaire using univariate analysis of responses[202].

The Edinburgh Inflammatory Bowel Disease (EIBDQ) was specifically designed to be both user friendly and practical. It consists of 15 questions. Response options are either yes/no or a little/moderately/severely[203]

A questionnaire has also been developed to measure the quality of life of household members of people with IBD. It is a 14-item questionnaire. The questions are organised into 2 domains: daily living activities and mental health. Responses are recorded on a 7-point Likert scale[204].

The Social Impact of Chronic Conditions-Inflammatory Bowel Disease (SICC-IBD) questionnaire was designed to measure the effect of IBD on social aspects of quality of life (social morbidity). Originally validated for use in ulcerative colitis, the SICC-IBD consists of 8 questions. Each question consists of 2 parts. The first relates to "disruption of life" and is answered either yes or no. If answered positively, the second part of the question, which relates to "severity of disruption" is completed. This is recorded on a 5-point Likert scale ranging from "extremely" to "not at all". A SICC-IBD score can be derived from the questionnaire. This is a weighted sum of items answered positively. Hence, a low score represents a low level of social morbidity [205].

The Cleveland Global Quality of Life (CGQL) instrument was originally developed and validated to measure health related quality of life in patients following ileo-anal pouch procedures (IPPA)[206]. It has since been validated for use in patients with Crohn's disease[224].

The CGQL instrument is a short questionnaire consisting of 3 items: current quality of life, current quality of health and current energy level. It is designed to be self-completed and respondents are asked to rate each item between 0-10 with 0 indicating worst state and 10 indicating best state. These scores are summed and then divided by 30 to produce a CGQL utility score[206]

2.3.1.3 Measures of other aspects of health status

The Ulcerative Colitis and Crohn's Disease Health Status Scales were developed to measure health status. Health status is described as a concept that incorporates a

patient's perception of illness, functional status, psychological factors and disease activity.

Multiple regression analyses were performed in order to determine the components of the final health status scales. In order to ensure the scales were easy to interpret, variable weights were used to produce potential scores ranging from 0-100. The resultant health status scale for UC consists of 9 components to which a constant is added to produce the final value. The scale for Crohn's disease consists of 10 variables and does not require the addition of a constant value[186].

An instrument consisting of simple visual analogue scales has been developed to measure the burden of Crohn's disease and its treatment, on patients. It consists of 2 "feeling thermometers" ranging from 0 (death) to 100 (perfect health). The first scale relates to burden due to disease. Respondents are instructed to place a "C" on the thermometer to indicate current health and a "P" on the same scale to indicate perceived health in the absence of symptoms related to Crohn's disease. The difference between these values represents symptom burden.

The second scale is related to treatment associated burden. As before, patients are asked to place a "C" on the scale to indicate current health. They are asked to place a "P" to represent their perceived health if all aspects related to treatment could be stopped. The difference between the two values indicates treatment burden. This instrument is designed to be self-completed and can be undertaken rapidly and therefore be used in the clinical setting[207].

A questionnaire to determine self-efficacy in IBD has been developed[197]. Self-efficacy has been shown to affect health outcomes in other chronic diseases [225] and it was therefore hypothesised that it would also have an effect on patients with inflammatory bowel disease.

The Inflammatory Bowel Disease Self-Efficacy Scale (IBD-SES) consists of 29 items within 4 domains (managing stress and emotions, managing medical care, managing symptoms of disease and maintaining remission). Answers are recorded on a 10-point Likert scale with 1 anchored at "not sure at all" and 10 at "totally sure". Hence, a low score indicates low self-efficacy and vice versa [197].

A simple, numeric rating scale has been developed to measure overall health status and general well being in IBD. It has been designed to be self-completed and to be used by patients with both Crohn's disease and ulcerative colitis. It is a horizontal 11-point scale ranging from 0 ("as bad as being dead") to 10 ("perfect health"). It is a rapidly completed and convenient tool, that can give an indication of overall health status from the patient's perspective[208].

2.3.1.4 Measures of patient satisfaction with healthcare

The treatment satisfaction questionnaire for Crohn's disease (TSQ-C)[198] was designed to measure patient satisfaction with medical therapy in people with Crohn's disease. The TSQ-C was developed through adaptation of an existing satisfaction questionnaire used in gastro-oesophageal reflux disease [226].

The questionnaire consists of 32 items within 6 domains: symptoms, satisfaction, expectations, "physician relationships", bother and cost. Response options range from "very strongly agree" to "very strongly disagree" and are recorded on a 6 point Likert scale[198].

The QUOTE (quality of care through the patient's eyes) questionnaires measure quality of healthcare from the patient's perspective and are designed to be self-administered. They have been designed to explore the importance patients convey to particular aspects of care, the performance of the healthcare service used as well as the combined effect of both importance and performance (termed the "quality impact" and calculated using the importance and performance scores)[227].

A series of questionnaires exist for use in a number of specific diseases including inflammatory bowel disease[209]. Each questionnaire consists of 10 core generic questions in addition to disease specific items. The QUOTE-IBD consists of 10 generic questions and 13 IBD specific items[209]

2.3.1.5   Measures of patient knowledge

Providing patient education is an important aspect of chronic disease management and is a recommended standard of IBD care in the UK[192]. It has been shown that patient education positively affects compliance with treatment and long term health outcome[228].

The Crohn's and Colitis Knowledge Score (CCKNOW) is a self-administered questionnaire designed to measure patient knowledge related to IBD and its treatment. It contains 24 multiple choice questions related to 1) general IBD knowledge, 2) medication, 3) diet and 4) complications related to IBD. All but 2 of the questions relate to IBD in general and can therefore be answered by patients with either ulcerative colitis or Crohn's disease.

Each correct answer yields one mark with no negative marking. Therefore, scores range from 0-30 [210]. The CCKNOW has been used to assess patient knowledge in a number of surveys[229, 230].

The Knowledge Questionnaire (KQ) is a similar questionnaire for the assessment of patient knowledge of IBD[211]. A patient knowledge questionnaire related to pregnancy and IBD has also been developed and validated[212]. However, these instruments have not been used as extensively as the CCKNOW questionnaire.

The instruments described above are used to assess patient knowledge over a wide range of areas. More recently, a short outcome measure for determining patient knowledge has been developed. It was designed to cover knowledge of disease risk and severity only. It consists of 10 questions with 8 related to the risk of developing IBD and 2 related to factors that may affect severity of IBD. It has been shown to be a rapid and acceptable measure of patient knowledge[213].

2.3.1.6   Measures of work, productivity and disability

Inflammatory bowel disease results in a significant economic burden and costs are both direct medical costs as well as indirect costs, as a result of loss of productivity[231].

The Work Productivity and Activity Impairment Questionnaire (WPAI) measures time missed from work as well as work and activity impairment due to health problems [232]. The WPAI has been modified for use in a number of chronic diseases [233, 234] including Crohn's disease (WPAI:CD)[196].

The WPAI:CD questionnaire consists of 6 questions relating to work and activity over the preceding 7 days. The questions cover aspects such as employment status, hours missed due to Crohn's disease, hours missed due to other reasons, hours worked, degree to which work productivity was affected (scale 0-100) and degree to which regular activities were affected.   The overall score is expressed as a percentage of impairment/productivity loss. High scores indicate greater impairment [196].

Chronic diseases such as IBD may also result in "work disability": the partial or total inability to perform work activities. The Crohn's Disease Perceived Work Disability Questionnaire (CPWDQ) was developed in order to measure work disability from the patient's perspective.

The CPWDQ consists of 2 domains: clinical determinants of work impairment (containing 11 questions) and social determinants of work impairment (5 questions). Responses are

recorded on a 4-point Likert scale. Potential scores range from 0-64 with higher scores indicating worse perceived disability[214].

It was recognised that, in contrast to other chronic diseases[235], patient reported outcome measures of disability in IBD were lacking[236]. Therefore, a disability index for use in IBD was developed [215]. This index is based on an ICF (International Classification of Functioning, Disability and Health) core set for IBD[237].

The IBD Disability Index is designed to be competed via interview. It consists of 28 questions. The majority of which are answered on a 5-point Likert scale (1=no, 5=extreme). There are also a number of symptoms-based questions such as number of liquid stools and presence of arthritis or arthralgia[215].

A second disability related instrument, the inflammatory bowel disease disability score (IBD-DS) has also been developed. This was also based on the ICF checklists for disability and functioning, as well as literature review and a survey of experts. It consists of 7 domains: demographics, mobility, gastro-intestinal related problems, self-care, major life activities, mental health and interaction with the environment. There are 49 questions; the majority of questions are answered on a 5-point Likert scale and a 0-10 visual analogue scale. Higher scores indicate more significant IBD related disability[216].

2.3.2 Generic instruments that have been validated for use in inflammatory bowel disease

There are a number of generic measures of health-related quality of life that have been validated for use in inflammatory bowel disease. The EQ-5D questionnaire was developed by the EuroQol group - a multidisciplinary, multinational group of experts, in order to measure health related quality of life[172]. It consists of 2 parts: the EQ-5D descriptive system and a visual analogue scale.

The descriptive system consists of 5 dimensions: mobility, self-care, usual activities, pain and discomfort and anxiety and depression. Each response is scored between 1 and 3 (1 indicates no problem and 3 indicates severe problems). These numbers are then expressed as a 5-digit code. The visual analogue scale (VAS) is a vertical, thermometer scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).

The EQ-5D descriptive system can be converted to a summary index using value sets. These values sets, or weightings, have been derived in various countries using the EQ-5D VAS valuation technique or the time trade off (TTO) technique. These are generally derived using a sample of the general population. The EQ-5D is widely used and has been translated into many languages[238].

The Short Form-36 (SF-36) survey is another widely used generic instrument. Originally designed for use in the Medical Outcomes Study, it also measures health related quality of life[145]. It consists of 36 items within 8 domains: physical functioning, role

limitations due to physical health problems, bodily pain, social functioning, general mental health, role limitation due to emotional problems, vitality and general health perceptions.

The scales can be used to produce two summary scores: the physical component summary (PCS) and the mental component summary (MCS)[145]. The SF-36 was originally validated in the United States of America[239] and has since been validated in the UK[240]. The SF-36 was revised in 1996. The main adjustment was to 5-point response options[241].

The 15D questionnaire is a generic instrument for the measurement of health-related quality of life. It consists of 15 dimensions which each contain 1 question. Dimensions include: breathing, mental function, speech, vision, mobility, usual activities, vitality, hearing, eating, elimination, sleeping, distress, discomfort/symptoms, sexual activity and depression[217]. Responses produce a 15D profile that can then be converted to a single index using value sets. It must be noted that the value sets available are based on the Finnish population[242].

The Morisky Medication Adherence Scale (MMAS-8) is a generic tool designed to measure medication adherence. It consists of 8 questions, the majority of which require a yes or no answer. The final question has 5 response options. A total score is produced and on the basis of this, respondents can be classified as low, medium or high adherers (<6 =low, 6-7=medium, 8=high)[218]. It has been validated for use in patients with inflammatory bowel disease[243].

The FACIT-fatigue questionnaire (FACIT-F) is a measure of fatigue[219]. It is a subscale of the Functional Assessment of chronic Illness Therapy general questionnaire (FACIT-G)-a measure of health-related quality of life[244]. Initially developed to measure fatigue associated with anaemia, it has since been validated for use in the general population as well as a number of chronic diseases including IBD[245].

The FACIT-F questionnaire consists of 13 questions that are answered on a 5-point Likert scale. Total scores can range from 0 to 52 and lower scores are associated with higher levels of fatigue[219].

2.4   Results - Psychometric properties of PROMs used in IBD

Criteria for the assessment of patient reported outcome measures have been recommended[142] and are described fully in chapter 1. Furthermore, eight key attributes of outcome measures have been defined. These are: measurement model (measurement characteristics), burden (time to administer), alternative forms, cultural and language adaptations, reliability, validity, responsiveness and interpretability[246]

2.4.1   Disease specific measures of quality of life including the IBDQ and its adaptations

The IBDQ has been shown to have good test-retest reliability with co-efficient of variations ranging from 0.06-0.15 (bowel domain 0.07, systemic domain 0.15, emotional domain 0.11, social function 0.06) reported in its original descriptive study[143]. This has been echoed in subsequent validation studies. High intra-class correlation

coefficients (ICC) across all domains have been reported in patients stable at follow up[187, 247]. A study validating the IBDQ in the UK population has reported similar ICCs (0.73-0.93) for individual domains[187]. Internal consistency has also been shown to be good (Cronbach's alpha 0.72-0.89 for individual domains)[187].

The UK IBDQ, a modified version of the original IBDQ, has also been shown to be reliable (ICC 0.73 to 0.93, Cronbach alpha 0.78 to 0.86)[195]. Similarly the Short Inflammatory Bowel Disease Questionnaire (SIBDQ) remains reliable despite its shortened length: test-retest correlation was good (r=0.65) and internal consistency was high (Cronbach alpha 0.78)[188]. Finally, the testing of the IBDQ-9 has shown good correlation of scores in stable patients at follow up and high ICCs indicating test-retest reliability. However, Cronbach alpha statistic was very high (UC-0.95, CD-0.91) [200] indicating possible redundancy of included items.

Original validation of the rating form for IBD patient concerns (RFIPC) has also proven test-retest reliability. Test-retest reliability was high (test-retest correlation 0.87 for overall sumscore)[185]. Subsequent studies have confirmed this in Crohn's disease (test-retest correlation 0.90, ICC 0.91)[248] and ulcerative colitis (test-retest correlation 0.87)[249]. Cronbach alpha statistics were, again, very high (CD 0.96[248], UC 0.95[249]).

The Short Health Scale has been shown to be reliable in Crohn's disease[250] and ulcerative colitis[201] with good test-retest correlation, and good internal consistency (Cronbach alpha 0.85) [251].

There is some evidence to support the reliability of the remaining quality of life instruments. However, this is less comprehensive than that presented above and, in most cases, relates to the original descriptive study.

IBDQ scores have been shown to correlate with other related measures such as patient global assessment (r=0.36 to 0.52)[143]. Subsequent full validation of the IBDQ in a large clinical trial of patients with Crohn's disease revealed it to correlate well with the Crohn's disease activity index (r=-0.67, p<0.0001)[133] and relevant domain scores of a generic measure of health status (SF-36, r=0.45 to 0.840)[187].

The UK-IBDQ has also been shown to correlate with the SF-36[195] and the SIBDQ with relevant disease activity indices[252]. The IBDQ-9 correlates well with the IBDQ-36 from which it is derived (r=0.91) and also with disease activity indices[200].

The IBDQ has also been shown to be able to discriminate between patient groups (differing disease activity as defined by CDAI scores with significantly different mean scores between groups, p<0.001 ANOVA). Significantly different IBDQ scores have been demonstrated between ulcerative colitis patients in remission and with active disease[253]. Similarly, the UK-IBDQ, SIBDQ and IBDQ-9 have also been shown to possess discriminant validity[188, 195, 200].

Regression analysis has indicated that RFIPC scores are related to disease activity and generic wellbeing instruments[185]. RFIPC scores of patients with Crohn's disease have been shown to correlate with domain scores of the IBDQ, the short health scale (SHS) and generic questionnaires[248]. Correlation of RFIPC scores of patients with ulcerative

colitis with a patient reported visual analogue scale of general wellbeing and the Sickness Impact Profile (SIP, a generic measure of health status) has also been reported [249]. The RFIPC correlates poorly with traditional measures of disease activity, but this is not unexpected as worries and concerns of patients are not necessarily related to current disease activity[248].

The Cleveland Clinic Questionnaire has also been shown to correlate with related components of the SIP (Sickness Impact Profile) questionnaire. It has been shown to be able to distinguish between its 4 test groups (Crohn's disease surgical, Crohn's disease non-surgical, ulcerative colitis surgical, ulcerative colitis non-surgical) [202].

Some of the components of the health status scales correlated strongly with the Crohn's disease activity index (CDAI)[186].

The Short Health Scale item responses correlate with domain scores of other health related quality of life scores (IBDQ, SF-36, RFIPC, psychological general well-being) in patients with Crohn's disease[250] and ulcerative colitis[201]. Scores from the remaining instruments have also been shown to correlate either with disease activity scores or other patient reported measures.

The RFIPC has been show to possess discriminant validity. Mean sum-scores from patients with active disease and disease in remission have been shown to be significantly different[248, 249]. This has also been shown to be the case with the Short Health Scale[250], house hold member measure (HHMQoL-IBD)[204] and the Social Impact of Chronic Conditions (SICC-IBD)[205].

The original validation study of the IBDQ showed that mean IBDQ scores changed in patients with a change in clinical condition. This was statistically significant for bowel, systemic and emotional domains.[143]. The difference in scores of patients with clinical change over the standard deviation of the difference in scores of stable patients was greater than 1 for all domains indicating responsiveness. A subsequent validation study calculated large effect sizes for patients relapsing (UC 1.70, CD 8.04) and patients entering remission (UC -1.88, CD -1.81) again indicating the IBDQ is responsive to clinical change[254].

The responsiveness ratios (mean change in scores of patients experiencing a change divided by standard deviation of the scores of stable patients) of sum-scores and domain scores for bowel function 1, bowel function 2 and social function of the UK-IBDQ were all greater than half a standard deviation in stable patients indicating responsiveness [195]. SIBDQ scores of patients with ulcerative colitis have been shown to change significantly in patients experiencing a change in disease status[252]. The IBDQ-9 is also responsive. Median scores were shown to change in patients experiencing relapse and the effect size was large (UC-2.67, CD -5.29).

Whilst the RFIPC has been shown to be responsive to clinical change in a study of patients with Crohn's disease (responsiveness ratio 0.84)[248], other studies have not been able to prove responsiveness [185, 249].

Significant changes to SHS scores have been shown in patients experiencing change in clinical status (CD: responsiveness ratio 1.06-1.98[250], UC 0.45-1.95[201]. Effect size

was moderate to large for patients with improving symptoms, but only small to moderate in patients with deteriorating symptoms [251].

The Cleveland Global Quality of Life measure (CGQL) has been shown to be responsive in patients following restorative proctocolectomy (significant change in median score in patients experiencing clinical change)[206] and in patients with Crohn's disease (significant change in median scores following surgery)[224].

Minimal clinically important difference (MCID), that is the smallest change in score perceived as beneficial to patients, is a measure of interpretability[255]. A change in overall IBDQ score of between 16 and 30 points has been shown to correspond to relapse (as defined by CDAI score) or change in therapy by the physician[133]. Regression analysis of data obtained from patients with Crohn's disease treated with infliximab has estimated cut off values for the IBDQ. An IBDQ score of 170 or more suggests clinical remission (as defined by a CDAI score of 150 or less). A change in score of 32 or more points suggests response (as defined by a change in CDAI score of -70 to -100)[256].

The MCID of the short IBDQ is 9 points which corresponds to a change of the CDAI of 100 points[188]. There are no reported data on minimal clinical important difference for the remaining disease specific measures of quality of life.

2.4.2 Disease specific measures of other aspects of health status

There are limited data with regard to the reliability of these instruments. The IBD self-efficacy scale (IBD-SES) has been shown to have excellent test-retest correlation of scores

at follow up (r=0.9). However, Cronbach alpha was very high (0.96) indicating possible item redundancy[197]. Cronbach alpha values for the health status scale scores ranged from 0.59-0.84 and hence the internal consistency of the components of the health status scale is acceptable to good. Test-retest reliability has not been reported[186].

Excellent correlation of the 2 "health status" questions within the burden of symptoms "thermometer" measure has been shown and this is an indicator of internal consistency[207]. There are no reported reliability data related to the numeric rating scale[208].

A number of components of the health status scales correlate with CDAI scores[186]. NRS (numeric rating scale) scores have also been shown to correlate with CDAI as well as HBI scores[208]. Burden of symptoms "thermometer" scores also correlates with HBI scores as well as short IBDQ scores[207]. IBD-SES and NRS scores have been shown to correlate IBDQ scores[197, 208]. Hence, there is evidence to support the construct validity of these instruments.

There is little evidence with regard to the responsiveness of the above instruments and none related to interpretability.

2.4.3 Disease specific measures of patient satisfaction with healthcare

The QUOTE questionnaire has been shown to have test-retest reliability with high correlation between scores at baseline and follow up (4 week interval)[209]. Cronbach

alpha statistics for the treatment satisfaction questionnaire (TSQ-C) ranged from 0.63-0.94[198].

TSQ-C scores have been shown to correlate with related subscale scores of the IBDQ[198]. In contrast, the QUOTE questionnaire has not been shown to correlate with health related quality of life[257]. Scores do, however, correlate with related visual analogue scale (VAS) scores[209].

The TSQ-C has also been shown to correlate with the Crohn's Work Activity Impairment Index (CWAII). It can also discriminate between different patient groups (defined by the number of flare ups over the last year and by disease activity) with significantly different scores between these groups[198].

There is no evidence with regard to the responsiveness and interpretability of these instruments.

2.4.4 Disease specific measures of patient knowledge

The CCKNOW, Knowledge Questionnaire (KQ) and CCPKNOW have all been shown to have very high Cronbach alpha statistics (>0.90)[210-212]. The short measure of knowledge has been shown to have good internal consistency (Cronbach alpha 0.73)[213]. There was high correlation of Knowledge Questionnaire scores between baseline and 4-week follow-up indicating test-retest reliability[211]. Test-retest reliability has not been reported for the remaining knowledge questionnaires.

The CCKNOW, CCPKNOW and short measure of knowledge have been shown to be able to discriminate between groups of respondents with different levels of knowledge. Significantly different scores were observed between medical, nursing and administrative staff[210, 212, 213]. CCPKNOW and short measure of knowledge have also been shown to correlate with CCKNOW scores indicating construct validity[212, 213].

Receiver operating characteristics (ROC) curve analysis has been used to identify cut-off scores for the CCPKNOW questionnaire. A score of 14 indicates a very good knowledge of IBD (sensitivity 94%, specificity 95%) and a score of 8 indicates "at least adequate knowledge" (sensitivity 78%, specificity 70%)[212]. Interpretability data have not been reported for the remaining knowledge instruments. Responsiveness has not been reported for any of the knowledge questionnaires.

2.4.5 Disease specific measures of work, productivity and disability

Whilst initial validation studies of the WPAI:CD (Work Productivity and Activity Impairment in Crohn's disease Questionnaire) did not prove it to be reliable, a more recent validation has shown that WPAI:CD scores were similar at 2 to 4 week follow up of stable patients. Intraclass coefficients were higher than 0.7 for all domains of the questionnaire[214].

The CPWDQ (Crohn's Perceived Work Disability Questionnaire) and the IBD Disability index have also been shown to be reliable as evidenced by non-significant changes of scores at follow up (clinically stable patients) and high ICCs (0.89 in both cases) [214,

258]. Analysis of the IBD-DS (IBD Disability score) has shown good repeatability of scores between first and second questionnaires (Bland-Altman analysis)[216]. Cronbach alpha statistics have been reported for the CPWDQ (0.89)[214], IBD Disability Index (0.94) [258] and IBD-DS (0.826 to 0.938)[216].

All instrument scores have been shown to correlate with disease activity indices. WPAI:CD and CPWDQ scores correlate with Harvey Bradshaw Index (HBI) scores [214] [258] [259], Disability Index scores correlate with CDAI scores (Crohn's disease) and partial Mayo scores (ulcerative colitis) [258] and IBD-DS scores also correlate with CDAI scores[216]. The WPAI:CD, CPWDQ and IBD-DS also correlate with measures of health related quality of life[214, 216].

The WPAI:CD is able to discriminate between patients in "best health" and "worst health" (using cut off values of the CDAI, IBDQ and SF-36 to define health status). Significantly different mean scores were observed between these groups[196]. Significantly different IBD Disability Index scores were observed between IBD patients and controls [258].

Reduction of WPAI:CD score has been shown to be larger in patients entering remission compared to those that did not. In addition, effect sizes for patients that failed to enter remission were small, but were moderate to large in patients that did enter remission (0.54-1.02)[196]. This indicates the WPAI:CD is responsive to clinical change. Significant difference in IBD disability index scores of patients experiencing a change in clinical condition has also been reported [258]. Again, this indicates it is responsive. There are no data with regard to responsiveness for either the CPWDQ or the IBD-DS.

Minimally important difference (MID) values for the WPAI:CD have been estimated using data from the PRECiSE 1 trial. Change in WPAI:CD scores of over 7% may be considered to represent substantial change in workplace productivity[260]. Cut-off values for IBD Disability index scores have been derived using ROC curve analysis. A value of more than 2.5 identified controls (sensitivity 94%, specificity 79%) [258].

2.4.6   Generic instruments validated for use in IBD

The EQ-5D has been validated for use in IBD. Test-retest reliability has been shown to be good. In patients reporting no change in condition at follow up, agreement of answers was be good (Kappa statistic 0.6-1.0). The ICC for the EQ-5D VAS score was 0.89[189]. A subsequent larger study has also shown test-retest reliability of the EQ-5D. Again, agreement of responses to the descriptive system was good (Kappa statistic 0.39-1.00). The ICC for the EQ-5D VAS was 0.77 and the EQ-5D index 0.89[190]. Internal consistency has not been reported.

ICCs for the dimension scores of the SF-36 ranged between 0.56 and 0.92. There were similar scores at baseline and follow up in stable patients in most cases. However, differences in scores were only statistically different in 2 dimensions (ulcerative colitis patients: physical functioning and bodily pain). Internal consistency was adequate for all dimensions in both Crohn's disease and ulcerative colitis (Cronbach's alpha 0.72-0.91)[261].

Test-retest reliability has been determined for the FACIT-F questionnaire. The ICC was 0.81 for all IBD patients with stable symptoms at follow up (assessments were completed

within 180 days). Cronbach alpha statistic was very high for ulcerative colitis and Crohn's disease groups[245]. This may indicate item redundancy.

Reliability of the 15D questionnaire and the MMAS-8 in IBD has not been reported.

EQ-5D VAS scores correlate with disease activity indices for Crohn's disease (r=-0.65[189], r=-0.69[190]) and ulcerative colitis (r=-0.71[189], r=-0.67[190]). The EQ-5D index also correlates with disease activity[189, 190]). Discriminative validity has also been shown: response levels of the EQ-5D items were significantly higher in patients with active disease. However, as expected for a generic measure, a ceiling effect was noted for the EQ-5D descriptive system and the EQ-5D index[189].

Correlation with related domains of other measures has been shown. There is good correlation between the EQ-5D VAS and the general health domain of SF-36 (r=0.64), the overall IBDQ score (r=0.73) and most of the SF-36 and IBD sub-scores [189].

The SF-36 questionnaire correlates moderately with disease activity in ulcerative colitis (r=-0.23 to -0.53)[262]. It has also been shown to have good discriminative ability in IBD[261].

The 15D questionnaire correlates strongly with total IBDQ scores (r=0.733, p<0.001) and frequency of IBD symptoms. It is also able to discriminate between active and inactive disease, with mean scores differing significantly[263].

The Morisky Medication Adherence Scale (MMAS-8) has been shown to correlate with other indicators of medication adherence (continuous single medication availability and mean possession interval). Low adherers (as defined by previously set cut-off values) had lower indicators of medication adherence as compared to medium adherers; medium adherers had lower scores than high adherers. There was a significant difference between these 3 groups (ANOVA p<0.001). Harvey Bradshaw Index (HBI) scores were also statistically different between these 3 groups. Lower scores were associated with the high adherer group and higher scores with the low adherer group[243].

FACIT-F scores correlate with disease activity indices (CD: HBI r=-0.49 p<0.001, UC: SCCAI r=-0.59, p<0.001). Scores also correlate with inflammatory markers and haematocrit in ulcerative colitis patients, but not those with Crohn's disease. Significantly different means scores were observed when comparing IBD patients with control respondents from the general population (38.9 vs 43.6, p<0.001)[245].

EQ-5D VAS scores have been shown to change significantly in IBD patients with a change in health status as determined by a transition question[189, 190]. Changes in EQ-5D index scores have also been shown in patients with improved health (if not already in remission). Standardised response means (SRM) were large in all subgroups of patients with the exception of those already in remission reporting an improvement in health status[190].

The responsiveness of SF-36 was adequate for most domains, although Guyatt's responsiveness statistic values were generally low (-0.18 to -0.5 for ulcerative colitis and -0.14- -0.63 for Crohn's disease). Values were lowest for the "role-physical" domain[261].

A trial involving patients with ulcerative colitis identified large effect sizes for the physical component score (PCS, 1.21) and the physical functioning sub-score (1.89)[262]

Significant change in FACIT-F scores has been shown to be associated with change in clinical condition determined by a physician global assessment[245]. Significant change in FACIT-F score has also been observed in patients entering remission following treatment with adalimumab[264].

Responsiveness has not been reported for the 15D questionnaire or the MMAS-8.

Data from two phase III randomised controlled trials of certolizumab[265, 266] have been used to estimate minimal clinically important difference (MCID) values for the EQ-5D VAS and the SF-36 for patients with Crohn's disease. MCID value estimates (using IBDQ for anchor-based estimates) were 4.1 for the mental component summary (MCS) and 3.9 for the physical component summary (PCS). The MCID estimate for the EQ-5D VAS was 9.2 [267].

Estimates of meaningful difference for the components of the EQ5D in IBD have been obtained with a regression model (by regression analysis using the transition question for anchoring). Improvement in the UK EQ-5D index score by 0.08 and the EQ-5D VAS of 11 was associated with patient perception of better health. A fall in UK EQ-5D index of 0.11 and the EQ-5D VAS of 14 was associated with perception of worsening health[190].

Cut-off values have been determined for the MMAS-8 questionnaire in its original validation study[218]. However, this study relates to patients with hypertension and cut-off values have not been determined for patients with IBD.

Interpretability has not been reported for the 15D instrument or the FACIT-F questionnaire.

2.5   Summary

There has been increasing interest in the use of patient reported outcome measures over the last two decades[161, 162]. A number of instruments have been developed for use in IBD[143, 185, 196], but as yet none have become established in routine clinical care.

A literature search of the PubMed database identified 27 IBD specific instruments, the majority of which are designed to measure health related quality of life. A number of instruments exist to measure other constructs such as patient satisfaction with healthcare, patient knowledge and disability. 5 generic outcome measures have been validated for use in inflammatory bowel disease.

Of the disease specific instruments, the IBDQ[143] is the most well established. There has been extensive work related to its validation including its translation into a variety of languages. It has also undergone a number of adaptions including shortening to for the short IBDQ (SIBDQ[188]). It has also been adapted to form an Anglicised version[195]. Other measures of quality of life exist, although they have not been as comprehensively validated as the IBDQ.

A number of instruments have been designed to measure other aspects of health status such as self-efficacy[197] and burden of symptoms and treatment[207]. However, these instruments have not been validated beyond their original descriptive studies. Questionnaires to determine patient satisfaction with healthcare and patient knowledge are described. The QUOTE questionnaire[209] and the CCKNOW[210] questionnaire are the most established of these instruments.

IBD results in significant economic burden[231] and the Work Productivity and Activity Impairment Questionnaire[196] has been designed to measure aspects related to this. A further instrument to measure "work disability" specifically (the partial or total inability to perform work activities) also exists[214]. More recently there has been a move to develop PROMs to assess disability related to IBD. The IBD Disability index[215] and the IBD Disability Score[216] are both based on ICF checklists for disability and functioning.

A number of generic outcome measures exist. Both the EQ-5D and SF-36 are well established and extensively used. They have both been validated for use in IBD[189, 261].

Health-related quality of life is the most common construct measured by IBD related outcome measures. The IBDQ is the most established disease specific instrument. It has been translated and validated in many languages and there is evidence to support its reliability, validity, responsiveness and interpretability. A number of other disease specific measures do exist, but are neither as widely used nor as thoroughly validated.

However, the majority are comprehensive questionnaires that are too burdensome to be used in routine clinical care. They are, therefore, mainly of use in clinical trials.

There are favourable data related to the use of a small number of generic measures in IBD. Whilst some are short and can be completed rapidly, the use of generic measures in isolation is not recommended as they may lack sensitivity to clinically important change[142]. A few shorter IBD specific instruments exist, an example of which is the short IBDQ. However, none have become established in day-to-day clinical care. This led us to conceive the idea that for 'IBD Control' to offer a novel approach and find a place in busy routine settings, it should aim to be a short and generic instrument – taking the simplicity and wide applicability of a general measure but tailored specifically for inflammatory bowel disease.

Given the recognised importance of the use of PROMs in clinical practice, there is a need to develop outcome measures that can be used in routine clinical care. In order to be of use in this role, an instrument must be rapid and easy to complete and must not be overly burdensome to administer. The above review of current PROMs highlights the need to develop such instruments.

Chapter 3 Methods

In the present chapter, I describe the methods adopted for the development, piloting and prospective psychometric testing of the IBD Control questionnaire. The results of the development phase and description of the final instrument are reported in chapter 4, and the data relating to the prospective validation study are presented in chapter 5.

3.1   Development phase

3.1.1   Questionnaire specification

A steering group was convened in order to formulate a specification for the IBD Control questionnaire. Steering group members included 2 Consultant Gastroenterologists, 1 Specialist Registrar in Gastroenterology and 2 IBD Specialist Nurses. With the aim of developing a patient reported measure of control of disease for use in routine clinical care in mind, the group defined a number of desirable and undesirable characteristics. The details of the specification are presented in chapter 4 (see results section 4.1).

3.1.2   Literature review

The design of IBD Control was informed by the literature review of existing patient reported outcome measures (PROMs) used in IBD. The methods for the literature review, and results, were described in chapter 2. My review of the literature confirmed a wealth of existing knowledge derived from IBD patients about the impact of these conditions on health and functional status, and informed the identification of several core domains that

were common to the existing generic and disease-specific instruments (see results section 4.6).

3.1.3 Focus group meetings and one-to-one interviews

3.1.3.1 Research team and reflexivity

Focus groups and interviews were facilitated by me, a Gastroenterology Specialist Registrar undertaking research during an approved break in clinical training for research purposes. My qualifications at the time were Bachelor of Science and MB BS. I had undertaken the required "Good Clinical Practice" training. The meetings were also attended by a member of the IBD specialist nurse team who assisted with collection of field notes. I had not had contact with the participants before the study. An informal introduction was given to the participants at the time of meeting or interview as to my role in the research team and the reason for undertaking the research.

The study was approved by the NHS Research Ethics Committee (REC) and all patients had received a patient information leaflet and provided written informed consent.

3.1.3.2 Qualitative study design

Participants were invited to take part in the qualitative study in one of two ways. The study was publicised at a Patient Education Day that was run at the recruiting hospital. They were asked to provide contact details if they were willing to participate in focus groups or interviews. Secondly, patients were approached face-to-face, during scheduled

clinical visits to the specialist nurse led clinics. Again, they were asked to provide their contact details if they were happy to participate.

Two focus group meetings were undertaken, with 6 patient volunteers in each group. One focus group included patients with ulcerative colitis and the other included patients with Crohn's disease. 13 individual interviews were also performed.

The focus group meetings were held at the recruiting hospital and did not coincide with the patients' routine clinical visits. One-to-one interviews were also performed at the recruiting hospital. These were arranged to occur at the time of clinic attendance and were performed in clinic either before or after their scheduled clinical appointment.

The patients were asked to discuss the concept of "control" of their inflammatory bowel disease and particularly the issues that they felt were associated with good and bad control. They were also shown a number of questionnaire types and asked to comment on these. Other issues discussed included ideal length of the questionnaire and when and how they would wish to complete it.   Field notes were taken during and immediately after the meetings and interviews.

Once a point was reached where new themes were not arising from interviews, and therefore data saturation had been reached, qualitative data collection was brought to a close.

To identify core domains and candidate items for the questionnaire, I undertook thematic analysis of field notes which included the transcription of verbatim patient quotes. This

was conducted deliberately as a top-down or theoretical thematic analysis, driven by my researcher-defined aims. My explicit aims were: (a) To verify that the themes expressed by our patients mapped to the broad domains elicited already from the literature review of PROMs; (b) To confirm that these mapped onto the construct of disease 'control' from the patient perspective; (c) To identify any new themes expressed by patients in relation to the construct of IBD 'control'.

In analysing the data from the interviews and focus groups, I broadly followed the six-phase process outlined by Braun & Clarke[268]. This involved: 1) familiarising myself with the content of all data transcripts to allow reflection on the overall body of data. This included making draft notes and "jotting down" initial impressions. 2) generating initial 'codes' within the field notes. This was informed by my knowledge of the existing literature review. Field notes and verbatim quotes were reviewed manually. Data felt relevant to, or capturing something interesting related to the research question, were coded. 3) searching for themes. Again, this process was informed by the literature review. Patient derived themes were then mapped onto the candidate themes derived from the literature review. 4) Reviewing themes in order to ensure that they made sense in relation to the design specification for the PROM. 5) Defining final themes. This was a final review of the themes and how they related to each other. 6) Write up of final results (see chapter 4 sections 4.3, 4.4 and 4.6).

3.1.4   Questionnaire development

Literature review and thematic analysis of focus group and one to one interview data resulted in the identification of a number of recurring themes. These were used, in

conjunction with the previously defined questionnaire specification, to select items for inclusion (see chapter 4, section 4.6).

The draft questionnaire was completed by 30 patients in order to test acceptability. The patients were also asked to comment on the questionnaire. On the basis of this pilot study, changes to layout and wording were made.

3.2 Prospective validation of the IBD Control Questionnaire

Patients were recruited from a single centre. This was a large teaching hospital serving approximately 330 000 people. Patients were recruited at routine visits to outpatient clinics, drug-monitoring clinics and admissions to the day-case unit. Patients admitted to hospital due to inflammatory bowel disease were also recruited.

Inclusion criteria included adult patients with a confirmed diagnosis of inflammatory bowel disease based on standard clinical, endoscopic, radiological or histological criteria of 6 months or more. Patients were excluded it they did not speak English or if they had cognitive impairment or active psychiatric disease. Informed consent was obtained.

Patients were asked to complete a questionnaire pack (before their consultation in the case of outpatients). This pack included the IBD Control Questionnaire as well as the following established outcome measures:

The UK Inflammatory Bowel Disease Questionnaire (UK IBDQ) is a measure of health-related quality of life for use in patients with IBD. It is an Anglicised, modified version of

the original inflammatory bowel disease questionnaire (IBDQ) and has been validated in the UK population. It consists of 32 questions, each with 4 response options[195]. If more than half of items were missing from one or more subdomains, the questionnaire was disregarded. Otherwise, missing items were substituted with the mean of the remaining items within the subdomain[269].

The EQ-5D (Euro-Qol, EQ-5D-3L) Questionnaire is a generic measure of health status. It consists of 2 parts: the EQ-5D descriptive system and a visual analogue scale. The descriptive system consists of 5 dimensions, which are expressed as a 5-digit code. The visual analogue scale (VAS) is a vertical, thermometer scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state)[172]. The EQ-5D utility index values are calculated using available value sets.

The Hospital Anxiety and Depression Scale (HADS) is a validated questionnaire for the measurement of anxiety and depression. It contains 14 questions within 2 domains (anxiety and depression). Separate scores for each domain are obtained[270].

A clinical assessment of current disease activity was undertaken by a member of the research team using the following disease activity indices:

The Harvey Bradshaw Index (for patients with Crohn's disease) is an established, simple measure of disease activity. It consists of 5 questions related to symptoms, signs and complications. A score of less than 5 indicates remission, whereas a score over 16 indicates severe disease[136].

The Simple Clinical Colitis Index Activity (SCCAI) (for patients with ulcerative colitis and IBD unclassified) is a simple disease activity measure based on patient symptoms[128].

If the patient was due to have a clinical review (eg clinic appointment), the doctor or specialist nurse in clinic was asked to record a "Global Physician Assessment" of the patient's current disease activity (remission, mild, moderate or severe) and whether disease activity had changed (better, the same or worse). Information, with regard to changes to treatment made as a consequence of the clinical encounter, was also recorded.

Hospital records were also used to obtain information with regard to diagnosis, disease duration, disease extent and phenotype. Previous hospital admissions, surgery and co-morbidity were also recorded. Current therapies were also documented.

3.2.1   Assessment of the psychometric properties of the IBD Control Questionnaire

3.2.1.1   Acceptability and feasibility

Ten patients were timed whilst completing the IBD Control Questionnaire in order to obtain an average completion time. All questionnaires completed at baseline visits were assessed for completion, and completion rates for individual questions were calculated.

3.2.1.2   Reliability

Reliability was assessed in 2 groups of patients. In both cases, correlation of serial scores and intraclass correlation coefficients were calculated.

Firstly, 20 patients were sent a further questionnaire pack (containing the IBD Control and UK IBDQ questionnaires) by post, 2 weeks after their first questionnaire completion, in order to determine test-retest reliability. The IBD Control scores of stable patients were compared. Patients with stable disease were defined on the basis of: 1) their response to the transition question within IBD Control (Question 2 "Over the past 2 weeks have your bowel symptoms been getting worse, getting better, or not changed") and 2) no more than a 10-point change in their total UK IBDQ score.

Secondly, scores obtained from stable patients who completed questionnaires on more than one occasion (due to multiple clinic visits) were also compared. In this case, stable patients were identified if there was no change in the Global Physician Assessment.

3.3.3 Internal consistency

Internal consistency was assessed for individual questions and sub-scores using Cronbach alpha[271]. Spearman's correlation coefficient was also used to compare individual question responses and sub-scores to the IBD Control visual analogue scale (VAS) score.

3.3.4 Construct validity

There is no "gold standard" measure of control of IBD. Therefore, construct validity was determined by comparing IBD Control scores with scores of a number of established and validated measures using Spearman's correlation. These established measures included

a disease specific measure of health related quality of life (UK IBDQ[195]), a generic measure of health status (EQ-5D[172]) as well as the Hospital Anxiety and Depression Scale (HADS[270]). Scores were also compared with disease activity indices (Harvey Bradshaw Index[136] and the Simple Clinical Colitis Activity Index[128]).

The ability of the IBD Control Questionnaire to differentiate between clinically different groups was tested (discriminant validity). Patients were categorised on the basis of Physician Global Assessment (remission, mild, moderate, severe). Mean scores were compared using analysis of variance (ANOVA).

Multiple variable linear regressions were performed in order to identify patient factors, other than disease activity, independently associated with IBD Control scores.

3.2.1.5 Responsiveness

Responsiveness was assessed in patients that completed follow up assessments and that had experienced a change in clinical condition. Responsiveness was tested in two groups. In the first, change in clinical condition was defined as a change of total UK IBDQ score of more than 10 points. In the second, the Physician Global Assessment was used to identify change in clinical state from baseline.

Change scores were calculated (i.e. the difference in scores between visit 1 and visit 2). Change scores for the IBD Control were correlated with change scores of the established outcome measures and disease activity indices.

The following responsiveness statistics were also calculated:

Effect size was calculated by dividing the difference between mean scores between assessments by the standard deviation of baseline scores.

The standardised response mean (SRM) was calculated by dividing the difference in mean scores between assessments by the standard deviation of change scores.

The modified standardised response mean (MSRM) was calculated by dividing the difference in mean scores between assessments by the standard deviation of change scores of clinically stable subjects. No change in Physician Global Assessment and a change in UK IBDQ score of no more than 10 points was used to identify stable patients.

3.2.2 Definition of cut-off values of IBD-Control scores for the detection of quiescent IBD.

Patients were defined as having quiescent IBD if all of the following criteria were met:

1. Clinical remission defined by 2 or more of the following disease activity/ patient reported/physician reported measures: simple clinical colitis activity index (SCCAI) less than 4 or Harvey Bradshaw Index (HBI) less than 5, UK-IBDQ score of greater or equal to 90 and a Physician Global Assessment of "remission".

2. The patient was not receiving a course of oral steroids, antibiotics (as a treatment for IBD), therapeutic liquid diet, induction with biological therapies or awaiting surgery for IBD.

3. The patient had not answered "worse" to the transition question within the IBD Control Questionnaire (Question 2 – Over the last two weeks have your bowel symptoms been getting worse, getting better, or not changed?).

4. There had been no escalation in treatment following consultation at the time of questionnaire completion.

Receiver operating characteristic curve analysis was performed to identify whether the IBD Control Questionnaire could reliably predict remission. Given the intended use to identify patients in remission, emphasis was on ensuring high specificity (of 85% or more). Thus, aiming to minimise the risk of identifying a patient with active disease as being in a quiescent state. Within the cut-off values associated with specificities of 85% or more, the value with the highest sensitivity × specificity product was chosen[129].

3.4   Summary

A multi-disciplinary steering group determined the key design specifications for the IBD Control Questionnaire.

Questionnaire development included a literature review of current patient reported outcome measures used in IBD, identifying key domains and candidate items to be

covered by the new PROM. Patient focus group meetings and one-to-one interviews were undertaken to confirm that these domains/items were relevant to the construct of 'disease control from the patient's perspective', and to capture any new domains not covered by existing generic or disease specific instruments. These continued until data saturation was reached.

Patient views from the qualitative work and a pragmatically-defined specification for the new PROM were used to guide the development of the IBD Control Questionnaire. This was piloted on 30 patients and minor changes were made following this.

Prospective validation occurred at a single centre with an established IBD service. Patients completed the questionnaire as well as a number of established measures at the time of a planned clinical encounter (e.g. clinic appointment). Disease activity indices were recorded, as were physician global assessments and changes to treatment. Clinical staff blinded to responses to the IBD Control Questionnaire.

The following psychometric properties were determined using established statistical methods: acceptability and feasibility, reliability, construct validity and responsiveness.

Patients with quiescent disease were identified if they met rigorous criteria on the basis of disease activity indices, patient reported measures and physician reported assessments. This was to establish the potential for the new PROM to reliably identify individuals with optimal disease control. Optimal disease control is not only a key therapeutic goal, but a key performance characteristic for any PROM that might be used in the future to support self-directed care or non-face-to-face disease monitoring via

emerging e-Health solutions such as patient portals or apps. Data from these subjects were used to identify cut-off values for the identification of patients with quiescent IBD. Optimal cut off values were determined using receiver operating characteristic curve analysis.

Chapter 4 Results of Questionnaire Development Phase

4.1   IBD Control Questionnaire specification

Our steering group of Gastroenterologists and Specialist Nurses determined the key design criteria for the IBD Control Questionnaire, based on informal consensus. A number of desirable and non-desirable properties were identified (table 1). It was agreed that in order to be included, items must be 1) able to capture the patient's assessment of their overall disease control, 2) relevant to patients with both ulcerative colitis and Crohn's disease, 3) relatively generic and not focused on specific symptoms, thus ensuring the questionnaire is applicable to all patients with IBD regardless of disease location and behaviour.

The latter criterion represented a new approach to PROM development for IBD, since we made an *a priori* decision to avoid a lengthy, itemised listing of individual gastrointestinal symptoms. We did not wish to re-invent a traditional multiple item health related quality of life (HRQoL tool), such as the existing IBD-Q. We rather aimed to explore a deliberately broad and generic approach to item selection. The PROM was intended to serve as a rapid tool to capture overall control for any patient with IBD, not to capture and quantify a list of discrete symptoms.

We concluded that symptom assessment for patients with Crohn's disease and ulcerative colitis, including those with specific complications or post-operative status, requires an individualised symptom checklist or formal clinical review. By taking this "generic" approach to PROM development for IBD, we would be attempting to produce a novel tool

that combined the broad thematic questions typical of a generic instrument (e.g. EQ5D) but tailored to the specifics of IBD.

Table 1 Questionnaire specification

Desirable	Undesirable
Short completion time	Lengthy questionnaire
Suitable for use in routine clinical practice	Replicates existing IBD questionnaires
Applies to all forms of IBD	Different variations required depending on disease type/location
Measures patient reported "control" of IBD	Measures disease activity alone
Simple scorings system	Complex scoring system
Identifies patients in need of intervention by the IBD team	
Identifies patients with unmet needs or dissatisfaction with healthcare	
Supports self-management	
Supports service improvement and quality assurance	

4.2 Literature review

A literature review of existing patient reported outcome measures was performed and this is reported in chapter 2. A number of disease specific and generic instruments were described in the literature. On review of these measures, a number of common domains were noted within the instruments. Common themes included questions related to physical symptoms such as bowel symptoms, abdominal pain and fatigue, emotional impact and social impact of symptoms.

4.3   Focus group meetings and one-to-one interviews with patients.

Following on from literature review, 2 focus group meetings with patient volunteers were held - one of patients with Crohn's disease and one of patients with ulcerative colitis. 13 one-to-one interviews were also performed. Saturation was reached at this point and no new themes were identified.

Coding of field notes and transcribed quotes identified a number of general themes. Firstly, participants described a number of symptoms that affected their perception of whether their disease was controlled. A number of these were physical symptoms. These symptoms were both disease specific and general. Other emerging themes were of impact of disease on mood and emotions, impact on day to day activities and treatment concerns (table 2).

Bowel symptoms and abdominal pain

Some patients reported very specific symptoms related to their disease such as perianal symptoms and joint symptoms. It was apparent that whilst these symptoms were important to individual patients, they would not be appropriate for inclusion in the IBD Control questionnaire given the aim for it to be used by all patients with IBD.

Frequency of opening bowels, urgency and the presence of nocturnal symptoms were commonly reported as indicators of deteriorating disease control. Patients felt that quantifying the number of times they opened their bowels a day was not helpful. Most patients with colitis reported that going more frequently was associated with a

deterioration in disease control. However, two patients with Crohn's disease reported opening their bowels less frequently indicated to them that their Crohn's disease was becoming more active. This highlighted that if the IBD Control instrument was to be of use to patients with all forms of IBD, items related specifically to bowel frequency should be avoided.

Abdominal pain and discomfort were commonly reported by patients as indicators of deteriorating control. Both patients with Crohn's disease and colitis reported these symptoms.

Fatigue

Fatigue was very commonly reported by patients. Although a number acknowledged that fatigue was an issue when their disease was controlled, it was reported to become much worse during active disease. It was described as a "major symptom" (patient 1, female, Crohn's disease). Another stated that she "cannot do anything at all due to tiredness" (patient 7, female, Crohn's colitis). Another patient commented "I feel drained and unwell all the time and I know my colitis is flaring" (patient 16, female, ulcerative colitis).

Other patients reported "I can't do anything, can't go out, can't do simple tasks" (patient 18, female, Crohn's disease) and "I can fall asleep if I just sit down for a short time" (patient 20, male, ulcerative colitis). More comments on this symptom included "I can't be bothered to get up and have a bath in the morning" (patient 23, male, Crohn's colitis) and "I have to sleep all day in order to get up for work in the evening" (patient 24, female, Crohn's disease).

Difficulty sleeping

A number of patients reported that difficulties sleeping occurred when their disease was not controlled. Sleep was reported to be disturbed due to nocturnal symptoms with one patient reporting that he could only sleep for 4 hours a night.

Psychological and emotional issues

Psychological and emotional issues were reported by many patients as indicators of change in disease control. A patient reported impact on others and a lack of sympathy for other people as indicators that her disease was becoming active. She stated that when she noticed that her symptoms were impacting other people, she realised her IBD was not controlled (patient 1, female, Crohn's disease). Another patient explained that she was more "snappy with the children" during flares (patient 22, female, Crohn's colitis).

Deteriorating stress levels and mood were commonly reported during active episodes. One patient reported being "often close to tears" and felt that emotional symptoms were just as important as bowel symptoms during a flare up (patient 7, female, Crohn's colitis).

Effect on daily activities

This was also a common theme of problems reported by patients. Active symptoms were reported to affect home life, travel and work. Patients reported having to plan carefully before going out. A patient reported avoiding driving during flare ups (patient 22, female, Crohn's disease).

One patient reported having to work from home when her disease was not controlled. She also explained that she had often had to cancel commitments and trips due to her symptoms being active (patient 18, female, Crohn's disease). Another patient was not able to leave the house when symptoms were not controlled (patient 1, female, Crohn's disease). Another reported that "I become a recluse" (patient 16, female, ulcerative colitis).

Impact of treatment on perceived symptom control

Focus group members and interviewees were asked whether treatment concerns affected their perception of whether their IBD was controlled. One focus group felt that concerns about side effects might affect stress levels and therefore overall health. This was confirmed by a patient that reported that she experienced side effects related to her medical treatment and that this affected her general wellbeing (patient 16, female, ulcerative colitis). Another patient reported similar side effects and again reported that this affected her generally, but did not specifically indicate to her that her IBD was not controlled (patient 22, female Crohn's disease). Some patients did not experience side effects and therefore did not have strong concerns about medications.

Table 2 Themes identified following coding of focus group meeting and interview field notes

Domain	Theme	Codes
Physical	Bowel symptoms and abdominal pain	*Frequency of bowel movements* *Opening bowels more frequently* *Opening bowels less frequently* *Urgent need to open bowels* *How often I go to the toilet* *Nocturnal symptoms* *Abdominal pain* *Wind* *Cramps* *Avoiding food to reduce abdominal pain* *Reduced appetite*
	Fatigue	*Fatigue is a major symptom* *Can't do anything due to tiredness* *Feel drained* *Can fall asleep at any time* *I have to sleep all day* *I can't do anything* *I can't do simple tasks* *Lie down and rest as soon as I have done anything*
	Difficulty sleeping	*Symptoms at night* *Disturbed sleep* *Only able to sleep for 4 hours*
Social	Effect on daily activities	*Plan carefully before going out* *Avoiding driving* *Having to work from home* *Cancelling trips and commitments* *I become a recluse* *Not leaving the house* *Have to plan trips based on where toilets are* *Doesn't go on recreational walks due to lack of toilet facilities* *Lost job because of condition*
Emotional	Psychological and mood issues	*Lack of sympathy* *Symptoms impacting on others* *Snappy* *Worsening stress levels* *Low mood* *Often close to tears* *Embarrassment* *Takes over your life* *More emotional* *Cries a lot*
Treatment	Impact of treatment of perceived symptom control	*Concerns about side effects* *Side effects affecting general wellbeing*

4.4   Patient views on questionnaire design and format

Patients were also asked their views on the design and format of the questionnaire. They were shown some existing patient reported outcome measures such as the UKIBDQ[195] as well as examples of visual analogue scales and asked to comment.

Many felt that completing a questionnaire before a clinic appointment would help guide the consultation. Some felt this could be in the waiting room whilst others felt they would prefer to complete it at home. One focus group decided they would be happy to spend around 5 minutes completing a questionnaire. However, another patient felt she would be happy to spend 10-15 minutes doing so. Some patients felt that the questionnaire should be no longer than the equivalent of one side of A4 paper.

With regard to visual analogue scales (VAS), patients felt that horizonal scales were preferable and that a phrase at either end of the scale would be helpful. It was commented that visual analogues such as "smiley faces" were too simplistic. Most patient felt a combination of questions and visual analogue scales would be optimal and some suggested the inclusion of a free text box.

Patients suggested developing a questionnaire that could be completed by email, text message or on a website. One explained that she would not like to return it by post.

4.5   Questionnaire design

Core domains of "physical", "social", "emotional" and "treatment" were identified on the basis of literature review and the above qualitative data (figure 4). The previously defined questionnaire specification was used to guide the choice of broad, non-specific questions to represent these domains. A visual analogue scale (VAS) was also included. This draft questionnaire was piloted on a sample of 30 patients. Following this, changes were made to wording and layout.

Figure 4 Flow chart representing the development of the IBD Control Questionnaire
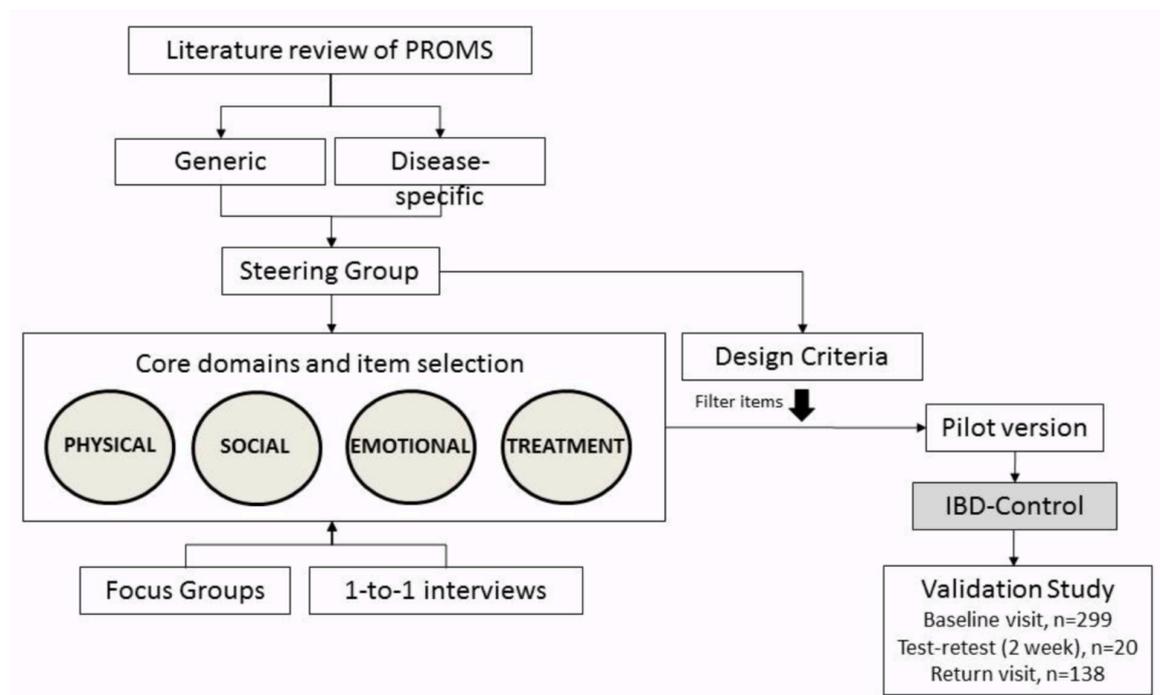
4.5.1 The IBD Control Questionnaire description

The IBD Control Questionnaire (figure 4.2) consists of 5 sections. It contains 13 questions and a visual analogue scale. When completed on paper, it fills one side of A4 paper.

Each question has 3 response options. In all but one question, these options are yes, no and not sure. In the case of question 2 ("Over the past 2 weeks, have your bowel symptoms been getting worse, getting better or not changed") the response options are: better, no change or worse. This was included as a 'transition' question, to determine a patient's sense of change over time.

Question responses are allocated a score of 0 for the least favourable response, 1 for the intermediate response or not sure and 2 for the most favourable response. The scores are simply added to produce a total score. The IBD Control 8 sub-score is calculated by adding the scores of the 8 questions within sections 1 and 3. The methodology for the development of this sub-score is presented in chapter 5.

IBD Control Questionnaire scores range from 0 to 26 and IBD Control 8 scores from 0 to 16. Low scores indicate poor patient reported control of IBD and high scores indicate good control.

The visual analogue scale is a horizontal scale ranging from 0 to 100. 0 is labelled "worst possible control" and 100 is labelled "best possible". Respondents are asked to rate overall control of their IBD symptoms over the last two weeks.

Figure 5 The IBD Control Questionnaire

IBD Control

Inflammatory Bowel Disease Control Questionnaire

---

1 Do you believe that:

	Yes	No	Not sure
a. Your IBD has been well controlled in the past *two weeks* ?	☐	☐	☐
	Yes	No	Not sure
b. Your *current treatment* is useful in controlling your IBD?	☐	☐	☐

(If you are not taking any treatment, please tick this box ☐)

---

2 Over the past 2 weeks, have your bowel symptoms been getting worse, getting better or not changed?

Better	No change	Worse
☐	☐	☐

---

3 In the past 2 weeks, did you:

	Yes	No	Not sure
a. Miss any planned activities because of IBD? (e.g. attending school/college, going to work or a social event)	☐	☐	☐
b. Wake up at night because of symptoms of IBD?	☐	☐	☐
c. Suffer from significant pain or discomfort?	☐	☐	☐
d. Often feel lacking in energy (fatigued) (by 'often' we mean more than half of the the time)	☐	☐	☐
e. Feel anxious or depressed because of your IBD?	☐	☐	☐
f. Think you needed a change to your treatment?	☐	☐	☐

---

4 At your next clinic visit, would you like to discuss:

	Yes	No	Not sure
a. Alternative types of drug for controlling IBD	☐	☐	☐
b. Ways to adjust your own treatment	☐	☐	☐
c. Side effects or difficulties with using your medicines	☐	☐	☐
d. New symptoms that have developed since your last visit	☐	☐	☐

---

5 How would you rate the OVERALL control of your IBD in the past *two weeks*?

Please draw a vertical line (|) on the scale below

Worst possible control	|——————————————————————|	Best possible

---

4.6   Summary

The IBD Control questionnaire specification was defined with the need to develop a measure for use in routine clinical care in mind. Literature review of existing patient reported outcome measures used in IBD was performed and a number of common themes were apparent.

Focus group meetings and one to one interviews were convened and patients described a number of symptoms they related to their IBD being controlled. Thematic analysis of field notes and transcribed quotes identified a number of themes. These included symptoms related to bowel function and abdominal pain. Fatigue and difficulty sleeping were also commonly reported issues. Impact on emotional wellbeing and daily activities were also described. Effect of concerns about treatment on control was also explored and concerns over side effects of medications were often raised.

A draft questionnaire was produced in line with the defined questionnaire specification and included questions related to the key themes described by patients and in the existing literature. It was then a piloted on 40 patients and minor changes to wording and format were made.

The IBD Control questionnaire is a short questionnaire consisting of 13 questions and one visual analogue scale set within 5 sections. Each question has 3 response options. A total IBD Control score is calculated by adding the scores of all question responses. The IBD Control 8 sub-score, the rationale for which is described in chapter 5, is a total score of responses to the 8 questions within section 1 and 3 of the IBD Control Questionnaire.

Chapter 5 Prospective validation of the IBD Control Questionnaire

5.1   Patient demographics

Patients were recruited between March 2011 and June 2012. A total of 299 patients completed a baseline assessment (table 3). 160 (53.5%) patients had a diagnosis of ulcerative colitis and 139 (46.5%) had a diagnosis of Crohn's disease. The mean age of patients completing baseline assessment was 43 (SD 16). 169 (56.5%) of patients were female and 130 (43.5%) were male. The mean duration of disease was 10 years (SD 10) and this was similar between the two diagnosis types (Crohn's disease 10 [10], ulcerative colitis 9 [10]).

83 (27.8%) patients had undergone surgery and 10 (3.3%) had a stoma. 27 (9.0%) patients had perianal disease. As to be expected, more patients with Crohn's disease had undergone surgery, had a stoma or had perianal disease than patients with ulcerative colitis.

26 (8.9%) patients were being treated with oral corticosteroids. 104 (34.8%) patients were on standard immunomodulator drugs (azathioprine, 6-mercaptopurine or methotrexate) and 38 (12.7%) were treated with biological therapies (infliximab and adalimumab). More patients with Crohn's disease were treated with immunomodulators (71 vs 33 p=0.001), biological therapies (30 vs 8, p=0.001) and dietary therapy (8 vs 0, p=0.008) than patients with ulcerative colitis. The frequency of steroid use was similar in both disease type groups (10 vs 16, p=0.107).

The mean EQ-5D utility index and EQ-5D visual analogue scale score[172] (measures of health status) were 0.68 [0.30] and 65 [23]. There were no significant differences in the mean EQ-5D scores of patients with Crohn's disease and ulcerative colitis (EQ-5D utility index 0.65 vs 0.70 p=0.12, EQ-5D VAS 65 in both groups). The mean UK-IBDQ score (an IBD specific measure of health related quality of life) was 86 [20]. Again, there was no significant difference in mean scores between the Crohn's disease and ulcerative colitis groups (85 vs 88 p=0.23).

Clinical disease activity was measured using the Harvey Bradshaw Index (HBI)[136] for Crohn's disease and the Simple Clinical Colitis Activity Index[128] (SCCAI) for ulcerative colitis. The mean HBI was 5[5] and the mean SCCAI was 4 [3].

161 (61.9%) patients were recorded to be in remission using the physician global assessment. 58 (22.3%) had mild disease, 30 (11.5%) had moderate disease and 11 (4.2%) had severe disease at the time of assessment. This distribution of disease activity is in keeping with the fact that most patients were assessed in outpatient or day unit settings.

138 patients undertook a further assessment at scheduled follow up visits (table 4). 82 (59.4%) patients had Crohn's disease and 56 (40.6%) patients had ulcerative colitis. The mean age of follow up patients was 41 [15] and the mean duration of disease was 8 [9]. 45 (32.6%) patients had undergone surgery, 5 (3.6%) patients had a stoma and 20 (14.4%) patients had perianal disease. Again, more patients with Crohn's disease had undergone surgery, had a stoma or had fistulating disease than patients with ulcerative

colitis. 16 (11.6%) patients were on oral corticosteroids, 61 (44.2%) were on immunosuppressants and 29 (21.0%) were on biological therapies.

The mean EQ-5D utility index score was 0.61 [0.34] and mean EQ-5D VAS score was 0.60 [0.35]. The mean UK-IBDQ[195] score was 83 [19]. 49 (44.1%) patients were in remission on the global physician assessment, 38 (34.2%) had mild disease, 18 (16.2%) moderate disease and 6 (5.4%) had severe disease.

Table 3 Characteristics of patients at baseline (n=299)

Characteristic	All patients	Crohn's	UC
Number of cases	299	160	139
Age, yrs	43 [16]	41 [15]	46 [16]
Gender - Male - Female	 130 (43.5%) 169 (56.5%)	 68 (57.5%) 92 (42.5%)	 77 (55.4%) 62 (44.6%)
Disease duration, yrs	10 [10]	10 [10]	9 [10]
Previous surgery - Yes - No	 83 (27.8%) 216 (72.2%)	 76 (47.5%) 84 (52.5%)	 7 (5.0%) 132 (95.0%)
Perianal disease (fistula) - Yes - No	 27 (9.0%) 272 (91.0%)	 27 (16.9%) 133 (83.1%)	 0 (0.0%) 139 (100.0%)
Stoma present - Yes - No	 10 (3.3%) 289 (96.7%)	 8 (5.0%) 152 (95.0%)	 2 (1.4%) 137 (98.6%)
Medication - Topical 5-ASA - Topical steroid - Oral 5-ASA - Oral corticosteroid -Immunosuppressants - Biological agent - Dietary therapy	 36 (12.0%) 9 (3.0%) 177 (59.2%) 26 (8.9%) 104 (34.8%) 38 (12.7%) 8 (2.7%)	 2 (1.3%) 2 (1.3%) 68 (42.5%) 10 (6.3%) 71 (44.4%) 30 (18.8%) 8 (5.0%)	 34 (24.5%) 7 (5.0%) 109 (78.4%) 16 (11.5%) 33 (23.7%) 8 (5.8%) 0 (0.0%)
Disease Activity Indices - Harvey-Bradshaw Index - SCCAI	 n/a n/a	 5 [5] n/a	 n/a 4 [3]
Quality of life questionnaires - EQ-5D Utility Score - EQ-5D Visual Analogue Scale - UK-IBD-Q	 0.68 [0.30] 65 [23] 86 [20]	 0.65 [0.30] 65 [22] 85 [18]	 0.70 [0.29] 65 [24] 88 [21]
Physician Global Assessment - Remission - Mild - Moderate - Severe	 161 (61.9%) 58 (22.3%) 30 (11.5%) 11 (4.2%)	 80 (60.6%) 37 (28.0%) 11 (8.3%) 4 (3.0%)	 81 (63.3%) 21 (16.4%) 19 (14.8%) 7 (5.5%)

Continuous variables expressed as mean [sd] and categorical variables as number (%) where appropriate)

Table 4 Characteristics of patients assessed at follow up visits (n=138)

Characteristic	All patients	Crohn's	UC
Number of cases	138	82	56
Age, yrs	41 [15]	38 [13]	45 [16]
Gender - Male - Female	65 (47.1%) 73 (52.9%)	36 (43.9%) 46 (56.1%)	29 (51.8%) 27 (48.2%)
Disease duration, yrs	8 [9]	9 [8]	8 [9]
Previous surgery -Yes - No	45 (32.6%) 93 (67.4%)	43 (52.4%) 39 (47.6%)	2 (3.6%) 54 (96.4%)
Perianal disease (fistula) - Yes - No	20 (14.5%) 118 (85.5%)	20 (24.4%) 62 (75.6%)	0 (0.0%) 56 (100.0%)
Stoma present - Yes - No	5 (3.6%) 133 (96.4%)	4 (4.9%) 78 (95.1%)	1 (1.8%) 55 (98.2%)
Medication - Topical 5-ASA - Topical steroid - Oral 5-ASA - Oral corticosteroid -Immunosuppressants - Biological agent - Dietary Therapy	11 (8.0%) 5 (3.6%) 82 (59.4%) 16 (11.6%) 61 (44.2%) 29 (21.0%) 5 (3.6%)	0 (0.0%) 0 (0.0%) 34 (41.5%) 6 (7.3%) 44 (53.7%) 24 (29.3%) 5 (6.1%)	11 (19.6%) 5 (8.9%) 48 (85.7%) 10 (17.9%) 17 (30.4%) 5 (8.9%) 0 (0.0%)
Disease Activity Indices - Harvey-Bradshaw Index - SCCAI	n/a n/a	5 [5] n/a	n/a 5 [4]
Quality of life questionnaires - EQ-5D Utility Score - EQ-5D Visual Analogue Scale - UK-IBD-Q	n/a 61 [23] 83 [19]	n/a 63 [22] 84 [18]	n/a 56 [25] 82 [22]
Physician Global Assessment - Remission - Mild - Moderate - Severe	49 (44.1%) 38 (34.2%) 18 (16.2%) 6 (5.4%)	27 (44.3%) 26 (42.6%) 6 (9.8%) 2 (3.3%)	22 (44.0%) 12 (24.0%) 12 (24.0%) 4 (8.0%)

Continuous variables expressed as mean [sd] and categorical variables as number (%) where appropriate)

5.2   Psychometric properties of the IBD Control Questionnaire

5.2.1   Acceptability and feasibility

Acceptability of the IBD Control Questionnaire, or whether or not the questionnaire was acceptable to patients, was determined by measuring administration time and response rates.

10 patients, aged between 25 and 64 years old, were timed completing the IBD Control Questionnaire. The average time to complete the questionnaire was 1 minute and 15 seconds (SD 25 seconds). Completion times ranged from 42 seconds to 2 minutes and 1 second.

259 (86.6%) patients completed all 13 questions within the IBD Control Questionnaire and 272 (91.0%) patients completed the IBD Control visual analogue scale. Completion rates of individual question items was high and ranged from 93.3% (Q2 "Over the past two weeks have your bowel symptoms been getting better, getting worse or not changed?") to 99.0% (Q3a "In the past two weeks did you miss any planned activities because of IBD?" and Q3d "In the past two weeks did you often feel lacking in energy?").

It was planned that items answered with the same response by 80% or more of the baseline sample would be removed from the questionnaire. This was due to the fact that these questions would be unlikely to have the necessary sensitivity to differentiate between different levels of symptom control. Exclusion of these questions would therefore ensure the questionnaire was not over burdensome. However, no

questionnaire item was found to fulfil this pre-specified cut-off. Therefore, all question items remained in the IBD Control Questionnaire.

5.2.2 Reliability

Reliability was determined using two methods. Firstly, test-retest reliability was assessed. 20 patients completed a questionnaire 2 weeks following baseline assessment.

13 of these patients were determined to have stable disease on the basis of: 1) stable UK-IBDQ score (score within 10 points of baseline assessment) and 2) an "unchanged" response to the transition question within the IBD Control Questionnaire (Q2 "Over the past two weeks have your bowel symptoms been getting better, getting worse or not changed"). These patients were therefore used to assess test-retest reliability (table 5A).

There was no statistical difference between IBD Control 8 or IBD Control VAS scores at baseline and 2 weeks later ($p < 0.01$). The mean difference in IBD Control 8 score was -0.45 (SD 1.81) and the mean difference in IBD Control VAS score was +2.25 (SD 9.79). The intraclass correlation co-efficients (ICC) were 0.97 for the IBD Control 8 sub-score (95% CI 0.90 to 0.99) and 0.96 for the IBD Control VAS score (95% CI 0.88 to 0.99).

Secondly, stable patients undertaking an assessment at a planned follow-up visit were identified on the basis of no change in condition reported on their global physician assessment. 32 patients were identified, and the mean interval between follow-up assessments was 131 days (table 5B).

There was no difference in mean IBD Control 8 and IBD Control VAS scores between visits (p<0.01). The mean difference between IBD Control 8 scores was 0.00 (SD 3.72) and the mean difference between IBD Control VAS scores was -0.56 (SD 23.00). The ICCs were 0.87 (95% CI 0.71 to 0.94) and 0.81 (95% CI 0.56 to 0.91) respectively.

Tables 5A and 5B Reliability of IBD Control summary scores

A) Patients returning at 2 weeks (test-retest group)

Instrument score (scale)	Mean difference (Visit 2 – Visit 1)	SD of difference	Intraclass Correlation (95% CI)
IBD-Control-8 sub-score (0-16)	-0.45	1.81	0.97 (0.90-0.99)
IBD-Control-VAS (0-100)	+2.25	9.79	0.96 (0.88-0.99)

Notes: n=13 patients responding 'not changed' to the IBD-Control transition question ('Over the past two weeks have your bowel symptoms been getting worse, getting better or not changed') and with stable UK-IBD-Q total scores (less than 10-point change between two visits). Two-way fixed effects model for consistency, p<0.001.

B) Patients returning for a scheduled clinic visit (routine care group)

Instrument score (scale)	Mean difference (Visit 2 – Visit 1)	SD of difference	Intraclass Correlation (95% CI)
IBD-Control-8 sub-score (0-16)	0.00	3.72	0.87 (0.71-0.94)
IBD-Control-VAS (0-100)	-0.56	23.00	0.81 (0.58-0.91)

Notes: n=32 patients with a stable (unchanged) Physician Global Assessment rating between two visits (mean interval between visits 131 days). Two-way fixed effects model for consistency, p<0.001.

5.2.3   Internal consistency

Internal consistency of the 13 question items within the IBD Control Questionnaire was assessed by Cronbach alpha statistic [271]. Cronbach alpha was 0.85 for the overall total score (13 questions) and 0.86 for the IBD Control 8 sub-score. A Cronbach alpha statistic of between 0.7 and 0.9 is desirable [272], indicating the IBD Control questionnaire is reliable.

Responses to the 13 categorical questions of the IBD Control Questionnaire correlated positively with IBD Control VAS scores. Correlation coefficients ranged from 0.24 to 0.70. Finally, the IBD Control 8 sub-score and IBD Control VAS correlated positively. This was the case for both IBD overall (r=0.78 p<0.001) as well as Crohn's disease (r=0.78, P<0.001) and ulcerative colitis (r=0.83, p<0.001) when considered separately. These correlations suggest that both individual questions as well as summary scores are measuring the same construct (in this case patient reported control of IBD) This therefore, indicates that the questionnaire possesses internal consistency.

5.2.4 Construct validity

There is no gold standard measure of patient reported control of IBD. Therefore, the IBD Control questionnaire was compared to a number of established measures with which it was hypothesised there would be correlations.

5.2.4.1 Validity of individual question items

Individual questions of the IBD Control Questionnaire correlated well with UK-IBDQ[195] scores (a disease specific measure of health related quality of life). Correlation coefficients ranged from 0.24 to 0.70 for the total UK-IBDQ score. Individual item scores also correlated with the IBD Control VAS score. Correlation coefficients ranged from 0.38 to 0.70 (table 6).

Table 6 Correlation between individual question items within the IBD-Control Questionnaire and the IBD-Control-VAS and disease-specific quality of life (UK-IBDQ)

Item	Question	IBD-Control-VAS	UK-IBD-QoL
Q1a	… your IBD has been well controlled in the past 2 weeks *	0.70	0.63
Q1b	… your current treatment is useful in controlling your IBD *	0.39	0.38
Q2	… have your bowel symptoms better/worse/not changed	0.36	0.24
Q3a	… miss any planned activities because of IBD *	0.63	0.67
Q3b	… wake up at night because of symptoms of IBD *	0.61	0.67
Q3c	… suffer from significant pain or discomfort *	0.64	0.66
Q3d	… often feel lacking in energy (fatigued) *	0.51	0.63
Q3e	… feel anxious or depressed because of your IBD *	0.58	0.70
Q3f	… think you need a change to your treatment *	0.60	0.50
Q4a	… would you like to discuss alternative types of drug	0.42	0.26
Q4b	… would you like to discuss ways to adjust your own treatment	0.33	0.31
Q4c	… would you like to discuss side effects	0.24	0.29
Q4d	… would you like to discuss new symptoms that have developed	0.30	0.29

Item responses were scored as 0 (worst option), 1 (intermediate, e.g. 'Not sure') or 2 (best option), so a higher score should indicate better disease control. Correlation coefficients are expressed as Spearman Rho values ($p<0.01$ in all cases).
*Question items included in the IBD-Control-8 sub-score

Furthermore, correlation between individual item scores and generic measures of health status was also observed (table 7). Correlation coefficients with the EQ-5D VAS ranged from 0.18 to 0.54 and with the EQ-5D utility index ranged from 0.16 to 0.62. Correlations of individual question scores with the global physician assessment of disease status (remission, mild, moderate or severe) were also seen (values ranging from -0.23 to -0.48, $p<0.01$), although correlations were weak (table 8).

Table 7 Correlation between individual question items within the IBD-Control questionnaire and EQ-5D VAS and EQ-5D utility index (a generic measure of health status)

| Item	Question	EQ5D VAS	EQ5D Utility index
Q1a	… your IBD has been well controlled in the past 2 weeks *	0.49	0.41
Q1b	… your current treatment is useful in controlling your IBD *	0.29	0.29
Q2	… have your bowel symptoms better/worse/not changed	0.18	0.17
Q3a	… miss any planned activities because of IBD *	0.54	0.52
Q3b	… wake up at night because of symptoms of IBD *	0.47	0.52
Q3c	… suffer from significant pain or discomfort *	0.52	0.51
Q3d	… often feel lacking in energy (fatigued) *	0.48	0.56
Q3e	… feel anxious or depressed because of your IBD *	0.54	0.51
Q3f	… think you need a change to your treatment *	0.42	0.62
Q4a	… would you like to discuss alternative types of drug	0.25	0.36
Q4b	… would you like to discuss ways to adjust your own treatment	0.26	0.16
Q4c	… would you like to discuss side effects	0.24	0.19
Q4d	… would you like to discuss new symptoms that have developed	0.23	0.16

Item responses were scored as 0 (worst option), 1 (intermediate, e.g. 'Not sure') or 2 (best option), so a higher score should indicate better disease control. Correlation coefficients are expressed as Spearman Rho values (p<0.01 in all cases).
*Question items included in the IBD-Control-8 sub-score

Table 8 Correlation between individual question items within the IBD-Control Questionnaire and the Global Physician Assessment (remission, mild, moderate, severe)

Item	Question	Global Physician Assessment
Q1a	… your IBD has been well controlled in the past 2 weeks *	-0.48
Q1b	… your current treatment is useful in controlling your IBD *	-0.34
Q2	… have your bowel symptoms better/worse/not changed	-0.23
Q3a	… miss any planned activities because of IBD *	-0.44
Q3b	… wake up at night because of symptoms of IBD *	-0.36
Q3c	… suffer from significant pain or discomfort *	-0.39
Q3d	… often feel lacking in energy (fatigued) *	-0.34
Q3e	… feel anxious or depressed because of your IBD *	-0.36
Q3f	… think you need a change to your treatment *	-0.42
Q4a	… would you like to discuss alternative types of drug	-0.30
Q4b	… would you like to discuss ways to adjust your own treatment	-0.26
Q4c	… would you like to discuss side effects	-0.21
Q4d	… would you like to discuss new symptoms that have developed	-0.19

Item responses were scored as 0 (worst option), 1 (intermediate, e.g. 'Not sure') or 2 (best option), so a higher score should indicate better disease control. Correlation coefficients are expressed as Spearman Rho values (p<0.01 in all cases).
*Question items included in the IBD-Control-8 sub-score

Question 3e of the IBD Control Questionnaire ("In the past two weeks did you feel anxious or depressed because of your IBD?") correlated with the Hospital Anxiety and Depression Scales (HADS, a patient reported outcome measure of anxiety and depression [270]). The correlation coefficient for the anxiety sub-score was -0.57 (p<0.001) and for the depression subscale was -0.62 (p<0.001).

Individual item scores were also correlated with sub-scores of the UK IBDQ. Question 3e correlated with the "emotional" sub-score of the UK IBDQ (r=-0.68, P<0.001). Question 3a ("in the past two weeks did you miss any planned activity because of IBD") correlated with the "social" sub-score (r=-0.64, p<0.001). Question 3d ("in the past two weeks did you feel lacking in energy") correlated with the "systemic sub-score (r=-0.68, p<0.001) (figure 6).

Figure 6 Validity of Question 3e of the IBD Control Questionnaire ("In the past two weeks have you felt anxious or depressed because of your IBD?") in relation to the Hospital Anxiety and Depression sub-scores.

Error bars: +/− 2 SE

Error bars: +/− 2 SE
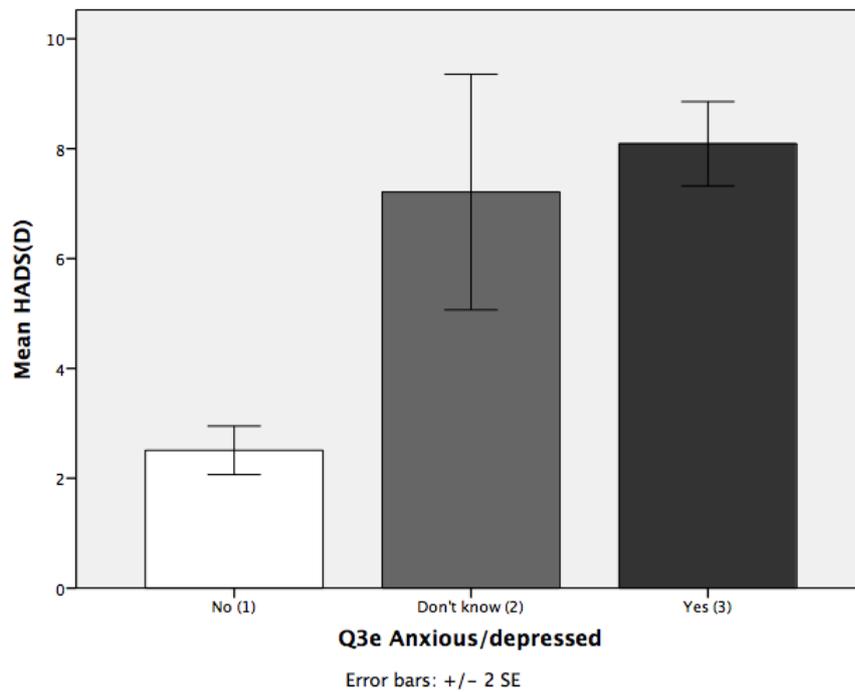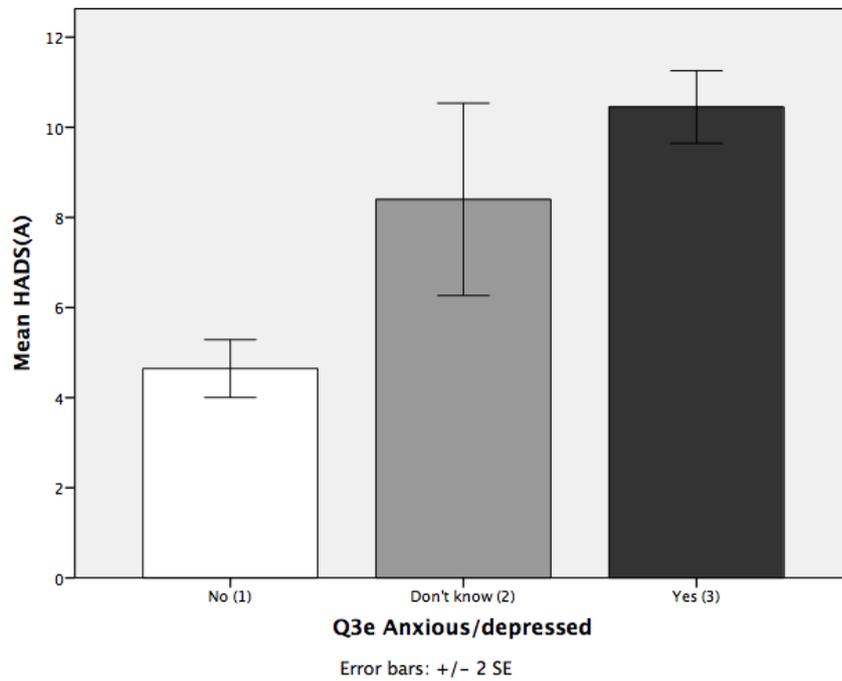
HADS(A)-Hospital anxiety and depression scales anxiety sub-score, HADS(D)-Hospital anxiety and depression scales depression sub-score

5.2.4.2 Validity of IBD Control summary scores

IBD Control 8 summary scores correlated positively with UK-IBDQ scores (figure 7) and EQ-5D scores (the EQ-5D utility index score and EQ-5D VAS score, generic measures of health status). Correlation coefficients for the UK-IBDQ scores were 0.86 (all IBD), 0.81 (Crohn's disease) and 0.90 (ulcerative colitis). Correlation coefficients for the EQ-5D utility score were 0.70 (all IBD), 0.68 (Crohn's disease) and 0.67 (ulcerative colitis) and for the EQ-5D VAS score were 0.68 (all IBD), 0.65 (Crohn's disease) and 0.71(ulcerative colitis) (table 9).

A linear regression model containing the eight items of the IBD Control 8 sub-score accounted for 75% of the variance in UK-IBDQ score (r square 0.747).

IBD Control scores were correlated with disease activity indices (the Harvey Bradshaw Index for Crohn's disease and Simple Clinical Colitis Activity Index for ulcerative colitis). Since a high IBD Control score indicates good disease control, it was predicted that negative correlations with disease activity indices would exist. Strong negative correlations were observed (Crohn's disease: IBD Control 8 -0.68, IBD Control VAS -0.60, UC: IBD Control 8 -0.72, IBD Control VAS -0.75, p<0.01).

IBD Control scores also correlated negatively with "current disease activity" within the global physician assessment (GPA). Moderate negative correlations were observed for all IBD patients (IBD Control 8 -0.58, IBD Control VAS -0.58, p<0.01) and patients with Crohn's disease (IBD Control 8 -0.45, IBD Control VAS -0.47, p<0.01). The IBD Control

summary scores correlated highly with GPA in patients with ulcerative colitis (IBD Control 8 -0.67, IBD Control VAS -0.65, p<0.01) (table 9).

Table 9 Construct validity of IBD Control summary scores: Correlation with quality of life, health status and disease activity at baseline assessment

External measure	IBD-Control-8			IBD-Control-VAS		
	All	CD	UC	All	CD	UC
UK-IBD-QoL score	0.86	0.81	0.90	0.76	0.65	0.85
EQ-5D Utility Score	0.70	0.68	0.67	0.52	0.50	0.58
EQ-5D VAS	0.68	0.65	0.71	0.70	0.69	0.71
Physician Global Assessment	-0.58	-0.45	-0.67	-0.58	-0.47	-0.65
Harvey Bradshaw Index	…	-0.68	…	…	-0.60	…
Simple Clinical Colitis Activity Index	…	…	-0.72	…	…	-0.75

Correlation expressed as Spearman Rho values (p<0.01 in all cases)

Figure 7 Construct validity of IBD Control summary scores: correlation with quality of life (UK-IBDQ)
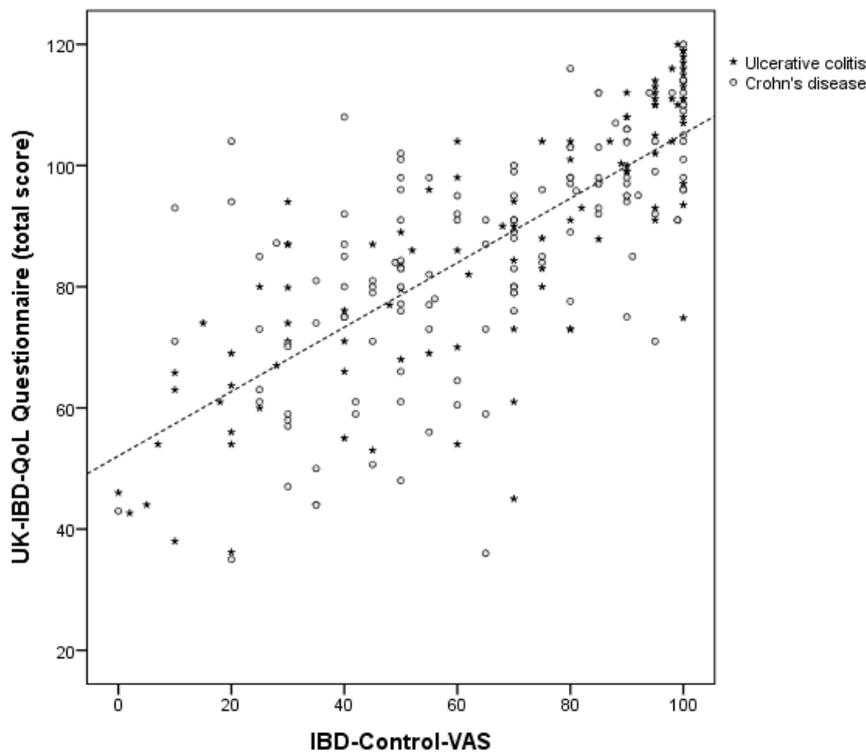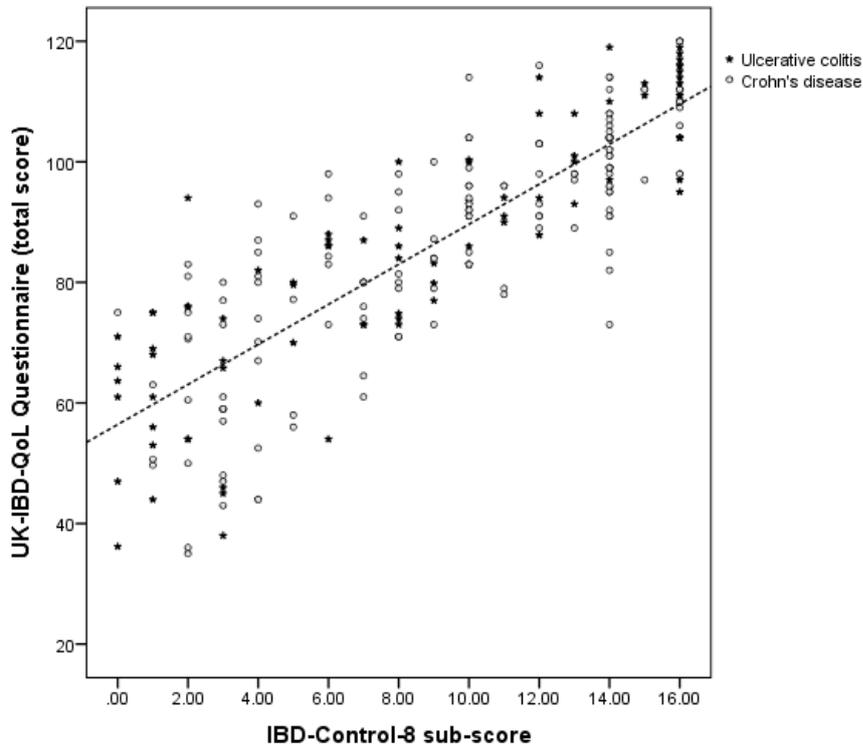
The IBD Control Questionnaire possesses discriminant validity (the ability to differentiate between clinically different patient groups). Patients were categorised on the basis of disease activity score with the global physician assessment (remission, mild, moderate and severe). Means scores were significantly different between these groups for both the IBD Control 8 and IBD Control VAS scores (figure 8).

Figure 8 Discriminant validity of IBD Control summary scores in relation to global physician assessment of disease activity.
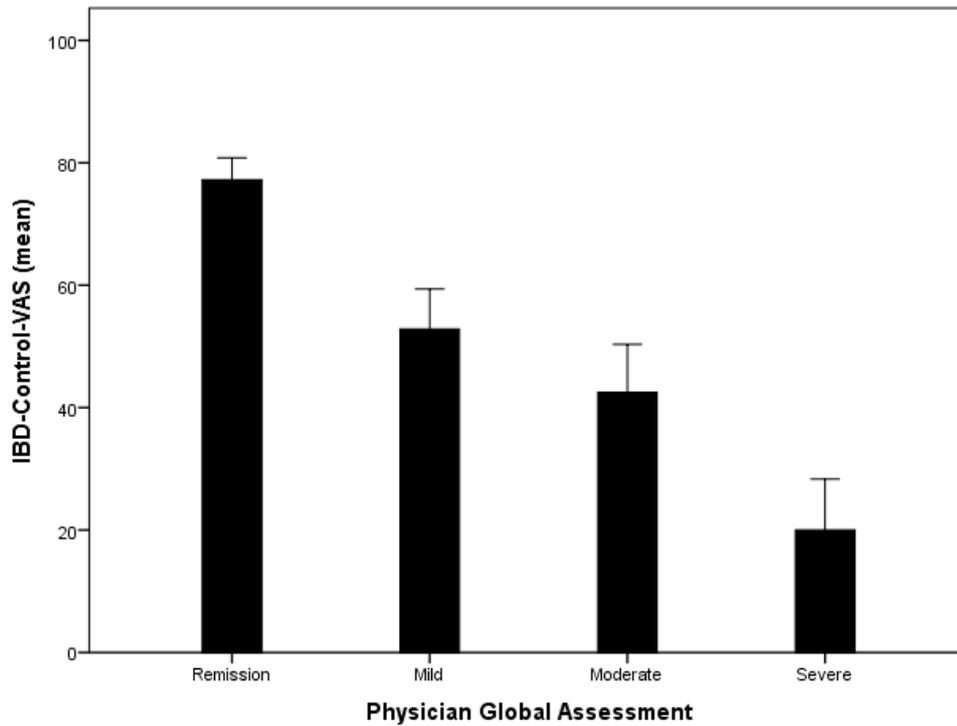
Error bars: +/- 2 SE
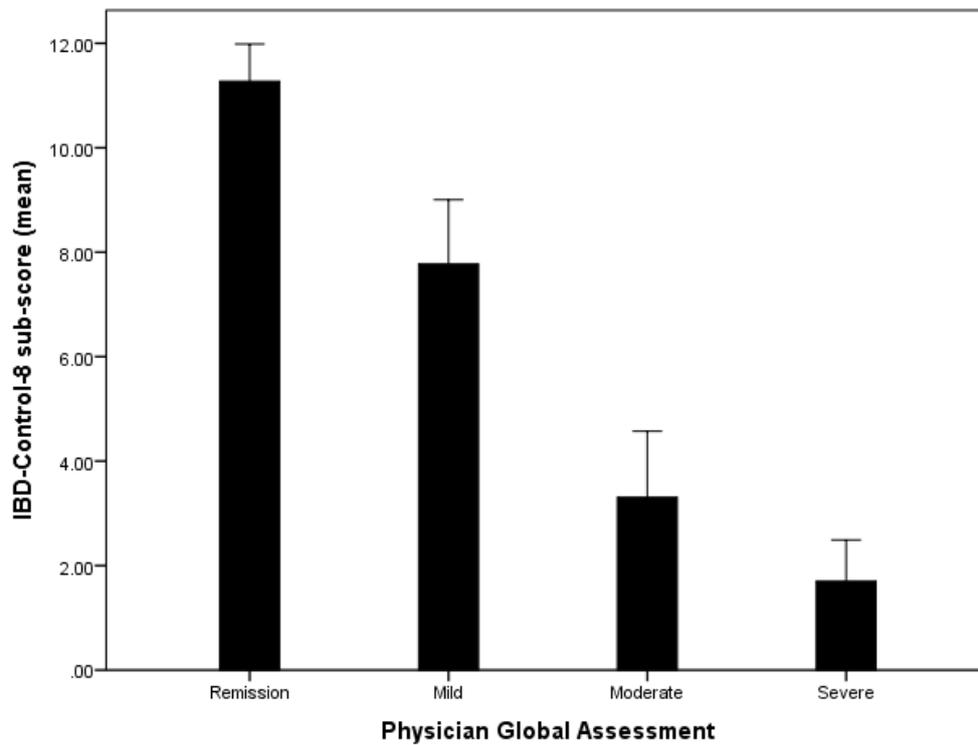
Error bars: +/- 2 SE

Multiple linear regression models examined whether patient factors such as age, gender and disease type were independently associated with IBD Control summary scores after taking disease activity into account. There was a strong association between IBD Control scores and disease activity indices, quality of life and physician global assessment. However, no additional factors were found to be independently associated with IBD Control summary scores.

5.2.4.3   Mapping of the IBD Control 8 sub-score to EQ-5D utility index

EQ-5D health states were converted to EQ-5D single index values. EQ-5D index values are country specific value sets and allow the calculation of quality-adjusted life years (QALYs)[273]. EQ-5D utility index values correlated with IBD-Control 8 sub-scores (figure 9). This indicates that IBD Control scores may be used as an estimate of utility. Therefore, the IBD-Control questionnaire, may be used in cost-utility analysis.

Figure 9 Validity of the IBD Control 8 sub-score in relation to EQ-5D utility index
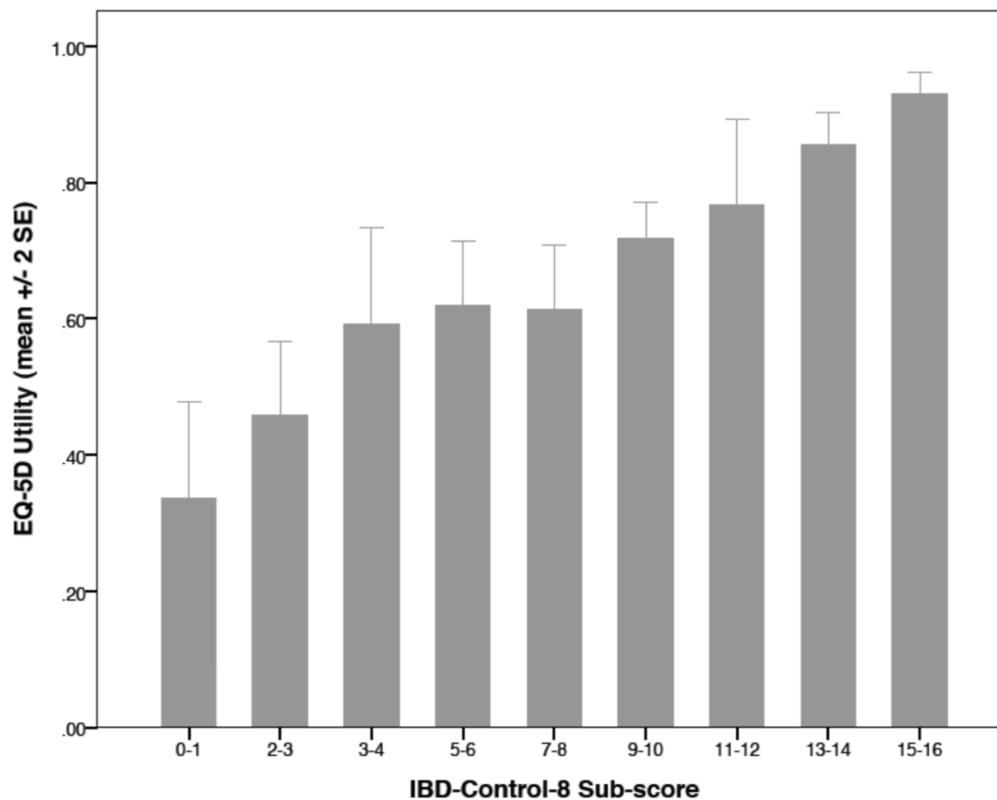
5.2.5 Development of the IBD Control 8 sub-score

5 categorical items were excluded from the IBD Control questionnaire to produce the IBD Control 8 sub-score. Firstly, the questionnaire's transition question (Q2 "Over the past two weeks have your bowel symptoms been getting better, getting worse or not changed") was not included as it measures change in IBD symptom control, rather than current IBD control.

Secondly, questions 4a to 4d had weaker correlations with the IBD Control VAS and UK-IBDQ, although the correlations were significant (table 6). These question items did not contribute significantly to the performance of linear regression models predicting IBD

Control VAS and UK-IBDQ. Whilst not included in the IBD Control 8 sub-score, these questions did provide important clinical information and were therefore retained in the overall questionnaire. Question 4d relates to the development of new symptoms over the preceding 2 weeks. The development of new symptoms is an important change to capture and may be used to prompt further clinical assessment. This question also remained in the questionnaire.

Questions 4a to 4c are related to patient reported concerns regarding treatment. Again, this is important information to guide clinical management and a positive response to these items was predictive of treatment escalation. 67 patients reported no treatment concerns (no positive response to questions 4a to 4c). Of these patients 5 (7.5%) had their treatment escalated. 192 patients recorded either a positive or indeterminate response to one or more of the treatment concern questions. Of these, 59 (30.7%) had their treatment escalated (p<0.001 chi squared test; OR 4.1, 95% CI 1.7 to 9.8).

5.2.6 Responsiveness

Responsiveness was tested in the 138 patients that completed serial questionnaires and clinical assessments. The mean changes in IBD Control summary scores (IBD Control 8 and IBD Control VAS) were correlated with the mean change in scores of quality of life and disease activity measures. Correlations were significant, and ranged from 0.25 to 0.78. The strongest correlations were with changes in UK-IBDQ score.

Responsiveness statistics were also calculated (tables 10A and 10B). The effect size (ES) was calculated by dividing the difference in mean scores (between baseline and follow-up assessment) by the standard deviation of baseline scores. An effect size of 0.2 is considered small where as an effect size of 0.5 and more than 0.8 are considered medium and large respectively[274].

The effect size for IBD Control 8 was 1.44 in patients with an improvement in UK-IBDQ of more than 10 points and 0.99 in patients with a reduction in UK-IBDQ score of more than 10 points. The IBD Control 8 also showed large effect size in patients with a change of more than 10 points on their UK-IBDQ score at follow up visit (>10 point improvement ES 1.09, >10 point deterioration ES 0.90). When change in clinical condition was defined by a change of one or more points on the global physician assessment, the IBD Control 8 sub-score was also found to be responsive (ES range from 0.76 to 1.14).

The standardised response mean (SRM) was also calculated by dividing the difference in mean scores by the standard deviation of change in scores. The SRM for IBD Control 8 was 1.17 in patients with a positive change in UK-IBDQ score of more than 10 points and 1.27 in patients with a negative change in UK-IBDQ score of more than 10 points. Again, SRM values were also large for the IBD Control VAS in patients with a more than 10-point change in UK-IBDQ score (>10 point improvement SRM 0.9, >10 deterioration SRM 0.95). When change in physician global assessment score was used to define change, the SRM values were moderate to large.

Finally, modified standardised response means (MSRM) were calculated by dividing the difference in mean scores by the standard deviation of change in scores, in clinically

stable patients. A clinically stable subgroup of patients with no change in physician global assessment and with a stable UK-IBDQ score (a change of less than 10 points between baseline and follow-up) was used for this calculation. The MSRM values were large for both the IBD Control 8 and IBD Control VAS, whether change in condition was defined by change in UK-IBDQ score or physician global assessment.

Table 10 Responsiveness statistics for the IBD Control summary scores

(A)   IBD-Control-8 sub-score

Criteria for change in state	n	Mean difference (sd)	ES	SRM	MSRM
UK-IBD-QoL					
- Improved by >10 points	19	+5.50 (4.69)	1.44	1.17	2.72
- No change (+/- 10 points)	63	+0.13 (2.02)	0.02	0.06	0.06
- Deteriorated by >10 points	26	-4.92 (3.88)	0.99	1.27	2.44
Physician Global Assessment					
- Improved by ≥1 point	27	+4.52 (5.56)	1.08	0.81	1.22
- No change	32	0.00 (3.72)	0.00	0.00	0.00
- Deteriorated by ≥1 point	27	-3.32 (4.81)	0.76	0.69	0.89

(B)   IBD-Control-VAS

Criteria for change in state	n	Mean difference (sd)	ES	SRM	MSRM
UK-IBD-QoL					
- Improved by >10 points	19	+28.21 (31.33)	1.09	0.90	1.84
- No change (+/- 10 points)	63	+0.84 (15.34)	0.03	0.05	0.05
- Deteriorated by >10 points	26	-26.85 (28.20)	0.90	0.95	1.75
Physician Global Assessment					
- Improved by ≥1 point	27	+29.55 (31.78)	1.14	0.93	1.28
- No change	32	-0.56 (23.00)	0.02	0.02	0.02
- Deteriorated by ≥1 point	27	-19.71 (30.53)	0.76	0.65	0.86

Effect size (ES) = Mean difference divided by the SD of the baseline absolute scores for the group.
Standardised Response Mean (SRM) = Mean difference divided by SD of the difference for the group.
Modified SRM (MSRM) = Mean difference for the group divided by SD of the difference for unchanged patients.

5.2.7 Sensitivity and specificity of the IBD Control Questionnaire to identify quiescent patients

A subgroup of patients with quiescent IBD at baseline assessment were identified on the basis of the following criteria: 1) clinical remission as defined by 2 or more measures (SCCAI <4, HBI<5, UK-IBDQ ≥ 90 points, physician global assessment rating of "remission"), 2) not receiving oral corticosteroids, therapeutic liquid diet, induction with biological therapy or awaiting surgery, 3) "unchanged" or "improved" response to the transition question of the IBD Control questionnaire and 4) no decision to escalate treatment at clinical review. Of these 80 patients, none required subsequent treatment escalation.

Receiver operating characteristic (ROC) analysis was undertaken. The area under the curve was 0.90 for the IBD Control 8 sub-score (p<0.001) and 0.86 for the IBD Control VAS (p<0.001) (figure 10). This indicates both summary scores to be highly accurate at detecting patients with quiescent disease.

A cut-off value for the IBD Control 8 of 13 points or more identified quiescent disease status with 67.5% sensitivity and 90.6% specificity. A cut-off value for the IBD Control VAS of 85 points or more, identified quiescent disease status with 64.3% sensitivity and 90% specificity. Review of the patients falsely categorised to have quiescent disease (the "false positives") revealed that they were all rated as in remission or to have mildly active disease on the global physician assessment. None of the patients required a change to their treatment at the time of clinical assessment.

Figure 10 Receiver operating characteristic (ROC) curves for the IBD Control 8 sub-score and the IBD Control VAS
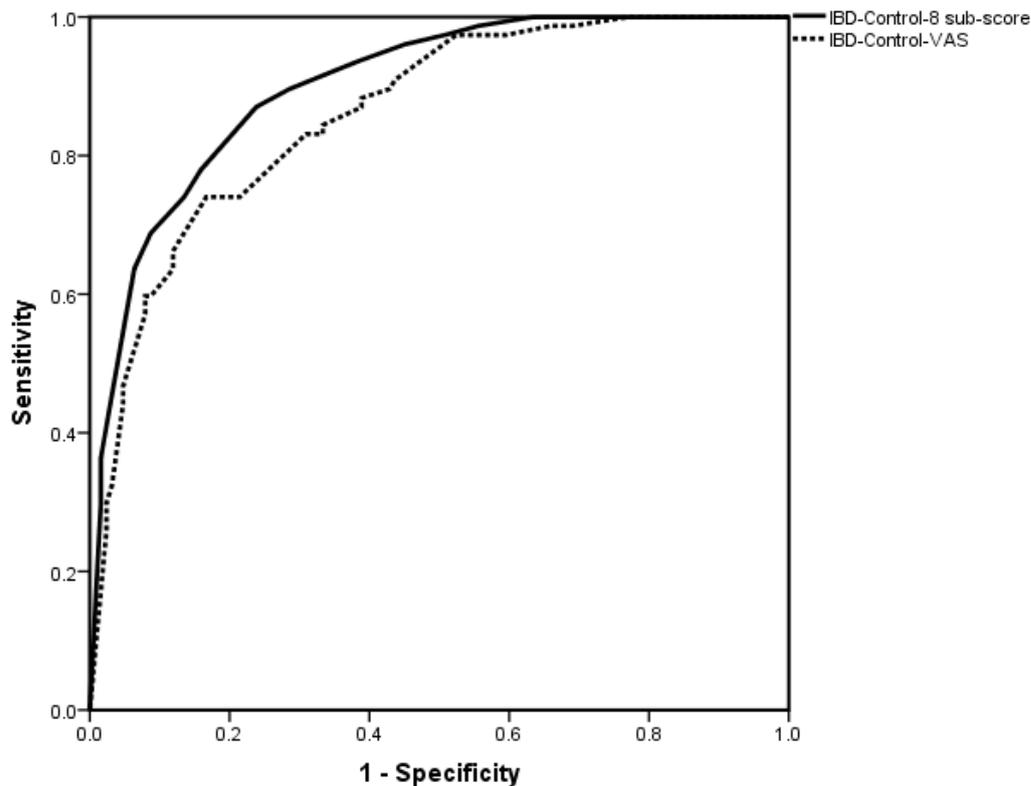
5.3   Summary

The IBD Control questionnaire was prospectively validated in a sample of 299 patients recruited from a single centre. It was shown to have favourable psychometric properties.

The questionnaire was acceptable to patients, as demonstrated by high completion rates (86.6% patients completed all 13 items). It could also be completed rapidly (average completion time 1 minute 15 [25 seconds]) demonstrating that it is feasible to use in routine clinical care.

Test-retest reliability was present. The was no significant difference in mean scores at baseline and 2 week follow up, in patients with stable disease. Intra class correlation co-efficients were high for the IBD Control 8 sub-score (0.97) and the IBD Control VAS (0.96) indicating excellent reliability[275].

Internal consistency, another measure of reliability, was determined by Cronbach alpha statistic[271]. Cronbach alpha statistics were favourable (0.85 for IBD Control 8 and 0.86 for IBD Control VAS). The IBD Control 8 sub-score and IBD VAS scores correlated (r=0.78, all IBD), again indicating internal consistency.

Construct validity was determined by correlating scores with a variety of established patient reported outcome measures and disease activity indices. Individual item scores correlated with UK-IBDQ scores, EQ-5D scores and global physician assessment. Summary scores (IBD Control 8 and IBD Control VAS) also correlated with these established scores as well as disease activity indices.

The IBD Control 8 and IBD Control VAS were also shown to possess discriminant validity, that is an ability to differentiate between different health states. Mean scores of patients categorised by global physician assessment (remission, mild, moderate or severe) were significantly different.

EQ-5D index values (which can be used to determine QALYs) correlated with IBD Control 8 scores. Therefore, the IBD Control 8 may be of use in cost utility studies.

The IBD Control 8 sub-score was developed by removing categorical items related to change in control, rather than control (Q2) and those related to other aspects of IBD care such as treatment (Q4). Although significant, these item scores correlated less strongly with IBD Control VAS and established measure scores. The excluded items were retained in the questionnaire, however, as they captured important patient reported information.

A range of responsiveness statistics were calculated. Clinical change was determined on the basis of a change in UK-IBDQ score and the global physician assessment. Effect size, standardised response mean and modified standardised response mean of the IBD Control 8 and IBD Control VAS were large. This indicated the instrument to be responsive to clinical change.

Finally, the sensitivity and specificity of the IBD Control Questionnaire to detect quiescent disease state were determined. A subgroup of patients with quiescent disease was identified on the basis of multiple criteria and receiver operator characteristic analysis was performed. A cut off value of the IBD Control 8 of 13 identified quiescent patients with 90.6% specificity and a cut of value of the IBD Control VAS of 85 identified quiescent patients with 90% specificity. Therefore, the IBD Control Questionnaire and the use of these cut off values, may be of use in screening for patients with controlled disease.

Chapter 6 Discussion

In this chapter I discuss the results presented in chapters 4 and 5 and how they relate to the aim of this study to develop a rapid and valid patient reported measure of IBD control. I will also describe further validations and uses of the IBD Control since its original development as well as other PROMs that have since been developed for use in IBD. Finally, I will provide my thoughts on the future role of the IBD Control Questionnaire in clinical practice.

6.1 The development and validation of the IBD Control Questionnaire

The IBD Control Questionnaire was designed with an aim to measure disease control from the patient's perspective. Although a number of PROMs existed, none had been incorporated into routine clinical care. Therefore, we aimed to develop and then robustly validate, a concise, rapid tool for use in clinical practice.

Literature review identified a number of disease specific instruments, many of which measured health related quality of life, as well as generic measures of health status. A number of common themes or domains emerged. Patient consultation on the concept of "control", and subsequent thematic analysis of qualitative data found a number of themes that could be mapped to the overarching domains present in the literature review. An additional domain related to IBD 'treatment concerns' emerged. Given the similarity of patient reported themes to those within the literature, a process of item list generation followed by factor analysis and item reduction was not performed. Rather, a pragmatic

approach was adopted, and a small number of deliberately broad questions were chosen to represent the identified themes.

Despite this approach, the data presented in chapter 5 confirm that the IBD Control Questionnaire has favourable psychometric properties. The questionnaire has been shown to be acceptable to patients by rapid average completion time and a high questionnaire completion rate. Acceptability is of a particularly importance, given the aim for it to become established in day to day clinical care.

It is a reliable measure: it possesses internal consistency (Cronbach alpha statistic between 0.7 and 0.9 for both summary scores) and test-retest reliability (intraclass correlation coefficients 0.97 for IBD Control 8 and 0.96 for total IBD Control score). Construct validity was tested by correlating scores with a range of existing measures. Scores correlated with measures of disease activity, patient reported outcome measures and physician assessment. It was also shown to be responsive with large effect sizes, SRM and MSRM values. Cut off values for identifying patients with controlled disease were also derived. Defined cut off values were able to predict patients with quiescent disease with 90% specificity.

This comprehensive psychometric assessment of the IBD Control Questionnaire indicates that it is a valid tool for collecting patient reported data and that it should be easy to use in routine clinical care. The strengths of this study include the participation of patients in the development of the questionnaire, which is recommended by both researchers[142] and the US Food and Drug Administration (FDA)[276]. Its strong psychometric properties and acceptability support its use in routine clinic care. The

definition of cut-off values to detect quiescent disease further adds to the benefit of using the IBD Control Questionnaire in clinical practice. The use of these cut of values may be of particular use in settings such as self-management programmes, non-medical led consultations and virtual consultations.

The comprehensive validation of this PROM not only supports its use in clinical care, but also pragmatic clinical trials. Furthermore, IBD Control 8 scores correlated with EQ-5Q utility index values. Therefore, IBD Control scores could possibly be converted to a utility estimate and may also be used in health economic studies.

Limitations of this study include the self-selection bias associated with recruiting patients to studies involving the completion of questionnaires. However, the study sample size was large and contained a range of patient demographics and clinical features. It was also a single centre study, which may have affected the ability to generalise these results to other subjects. The majority of patients were recruited from outpatient areas and therefore, only around 1/5 of patients had moderate to severe disease. Whilst it is more likely that the IBD Control Questionnaire will be used in outpatient or remote management settings, further validation of subgroups such as patients with more active disease, may be of use. Although we included patients with a history of prior surgery, case numbers were not sufficient to undertake detailed analysis of psychometric properties for selected sub-groups such as stoma patients. Furthermore, the identification of patients with more complex disease, such those with stricturing or penetrating phenotypes, may have allowed subgroup analysis to ensure the validity of the IBD Control Questionnaire in these patient groups. These are areas for future research.

6.2 Further validation of the IBD Control Questionnaire

Since we published the original paper describing the IBD Control Questionnaire, its use in routine clinical care has been reported. For example, it has been used to support virtual IBD clinics conducted via telephone. Patients were invited to complete and return a questionnaire via the post, followed by a telephone follow up. 59% of patients returned the questionnaire indicating that the IBD Control Questionnaire was acceptable to patients and a feasible tool for use in virtual clinical contacts [277]. Electronic capture of the PROM, via an online survey tool, patient portal or 'apps' could further improve the process.

Indeed, the IBD Control 8 sub-score of the IBD Control Questionnaire has been tested for electronic capture in a large web-based survey. The eight items were incorporated into the IBD2020 survey, an online survey of patient satisfaction and experiences of IBD care across the UK and Europe. 818 UK participants completed the survey and of these 85.7% (701) completed all items of the IBD Control 8. This high completion rate suggests the questionnaire is an acceptable tool for patients to complete within an electronic survey – in this instance the IBD-Control-8 was placed towards the end of a lengthy online questionnaire.

Within the IBD2020 survey, the questionnaire was reported to have favourable psychometric properties. Internal consistency was strong (Cronbach alpha 0.82). Scores correlated negatively with patient self-report disease activity in the last year (-0.68) and with number of relapses (-0.61). Mean scores were significantly higher in patients

reporting to be in remission than those reporting active disease. Finally, the IBD Control 8 score was confirmed to be of use in screening for patients with quiescent disease. A score of more than 13 points detected patients in self-reported remission with 88.2% specificity and patients with self-reported remission or minimally active disease with 95% specificity[278]. These data confirm the IBD Control 8 sub-score to be an acceptable and valid tool for use by patients electronically.

Other validations, including for use in a virtual biologics clinic, have been reported in abstract form[279, 280].

6.3   Current use of the IBD Control Questionnaire

The IBD Registry (ibdregistry.org.uk) is a UK wide repository of anonymised IBD patient data, for prospective audit and research purposes. It facilitates local data collection and can provide reports on performance compared with aggregated data from all contributing centres. It aims to drive improvement in clinical are, inform commissioning and service development, improve understanding of long-term outcomes and to support IBD research. The IBD Control Questionnaire is included in the IBD Registry dataset and, therefore, is being used to record patient reported outcomes at a number of centres in the UK. Although this currently requires manual entry of questionnaire responses by clinicians into the relevant registry data fields, the IBD Registry is seeking to roll-out patient-facing electronic tools to capture the PROM in 2019.

The IBD Control Questionnaire has been integrated into the Salford group's electronic patient portal, "My IBD Portal". This portal is a web-based system designed with the aim

to facilitate IBD patient engagement with the management of their condition. It provides a range of facilities including access to information such as blood results, a means to contact the IBD team and an ability to record patient reported information [281]. Over 700 patients are now using the My IBD Portal. It has been shown to have resulted in high satisfaction with self-management, improved perceived support and reduction in outpatient clinic attendances. The portal has also been adapted for integration with the IBD Registry [282].

"TrueColours ulcerative colitis" (TCUC), developed by researchers in Oxford, is another web-based programme and is designed for use by patients with ulcerative colitis. This too, includes a number of questionnaires including the IBD Control 8. A pilot study to determine feasibility and usability has been performed. 95% of patients adhered to completing the IBD Control 8 questionnaire on a fortnightly basis as directed[283]. This study illustrated that patients could sustain electronic collection of PROMs at regular (fortnightly) intervals over time, although we would advocate less frequent administration of the questionnaire in routine practice to avoid survey fatigue.

The need for value driven healthcare is increasingly recognised. Measuring outcomes that matter to patients is important when determining value[284]. A standard set of patient-centred outcome measures has been developed through the collaboration of the Oxford Academic Research Network with the International Consortium of Health Outcomes Measurement (ICHOM).

It was developed with the aim to allow standardised capture of patient reported data in routine clinical practice, thus facilitating value driven care. The set of outcome measures

includes a number of measures set out within domains that encompass all aspects of IBD. The IBD Control 8 items have been recommended for inclusion in the IBD Standard Set, providing a tool to capture data in the domain related to "symptoms, function and quality of life"[285]. A number of translations of the tool are in progress for non-English speaking settings.

IBD Control 8 has been used to capture patient reported data in a number of published abstracts[286-288]

6.4   Other IBD related PROMs designed for use in routine clinical care

During the period of my research, there was increasing interest in the collection of patient reported outcome data in order to facilitate routine clinical care. For example, the Crohn's and Ulcerative Colitis Questionnaire (CUCQ) is a new measure of quality of life and was designed to be suitable for use by outpatients with stable or mildly active disease as well as patients with severe disease[289].

The CUCQ-32 is a 32-item questionnaire and was developed along conventional lines as a disease-specific quality of life instrument. An overall score is obtained by adding item scores together, with potential scores ranging from 0-272. A high score is associated with poor quality of life. Comprehensive psychometric evaluation was performed. The PROM was shown to be reliable: Internal consistency was present (Cronbach alpha 0.88) and test-retest reliability was also proven. Scores correlated well with EQ-5D and SF-12 scores. Responsiveness was also very good (responsiveness ratio 0.85, SRM 0.99).

Stepwise regression analysis was performed in order to identify items for inclusion in a short version of the questionnaire (as a short version might be appropriate for routine settings). 8 items were identified that explained 95% of variance and these were used to form the CUCQ-8 questionnaire. This shortened version retained favourable psychometric properties[289].

The CUCQ has also been validated separately in patients with severe IBD. The 32-item questionnaire again had good psychometric properties in this context. Stepwise regression was performed, and in this patient group 16 items accounted for >95% of variance[290].

A simple rapid measure of patient reported symptom control, the IBD-10, has been reported [291]. This did not undergo formal developmental processes, or qualitative work with patients. The PROM was administered by asking patients to verbally report how their IBD was, on a scale of 0 to 10 (with 10 being the best).

Scores were shown to correlate strongly with disease activity indices (Harvey Bradshaw Index for Crohn's disease and SCCAI for ulcerative colitis). IBD-10 scores were also statistically different between remission and active disease groups. Receiver operator characteristics analysis identified IBD-10 cut off values for the detection of remission. A score of 7 or more was reported to predict remission with 90% sensitivity and 75% specificity.

IBD-10 scores correlated with CRP in patients with ulcerative colitis, but not Crohn's disease. Patients requiring treatment escalation had significantly lower IBD-10 score

than those that did not. A cut off value of <7 was reported to predict treatment escalation with 81% sensitivity and 70% specificity[291].

6.5 The role of IBD specific PROMs in clinical trials

There is increased recognition that patient reported outcome measures should be used to measure end points in clinical trials. The Food and Drug Administration (FDA) has published guidance on the development and use of PROMs in clinical trials to support labelling claims[276]. In contrast to traditional disease activity indices, it was recommended that patient reported outcome measures used in clinical trials were validated. The FDA recommended comprehensive validation and at the time of the guidance no IBD specific PROM met these stringent requirements[292]. Although there are a number of IBD specific outcome measures described in the literature, a recent systematic review has identified that very few have been adequately validated[293].

The development and validation of instruments to standards required by the FDA was predicted to be a lengthy and complex process. It was therefore proposed that interim PROMs be designed in order to allow clinical trials and therefore medicines development to continue. It was also proposed that these interim PROMs be used to measure symptoms, rather than outcomes, and to be based on patient reported components of existing disease activity indices[292].

PRO-2 and PRO-3 measures have been derived from patient reported components of the CDAI disease activity index. Optimal cut off values of individual diary card items to detect remission (CDAI score <150) were obtained and this generated a 2 item (PRO-2) and 3

item (PRO-3) measure. These measures were subsequently tested. They correlated moderately with CDAI scores and were also shown to be responsive [294].

Patient reported components of the Mayo Clinic Score (MCS) have also been used to generate interim patient reported measures for use in ulcerative colitis. Again, receiver operator characteristics were used to determine cut off values for the identification of remission (as defined by a Mayo endoscopic score of 0 or less than/equal to 1). Patient reported items alone or in combination (PRO-2) were assessed. A PRO containing both patient reported items (rectal bleeding and stool frequency) performed better than either item in isolation. The area under the curve (AUC) was 0.80 for remission defined by a Mayo endoscopy score of 0 and 0.90 for remission defined by a score of less than/equal to 1. Single patient reported items, items in combination (PRO -2) and PROs combined with endoscopic score were able to differentiate a treatment effect in most scenarios. However, it was noted that using PRO scores without endoscopy scores resulted in higher placebo rates of remission[295].

The Crohn's disease patient reported outcomes signs and symptoms (CD-PRO/SS) diary was developed in line with FDA recommendations. This comprehensive instrument consists of 6 modules: bowel signs and symptoms, abdominal symptoms, systemic symptoms, coping mechanisms, daily life impact and emotional impact. The first 2 modules, bowel signs and symptoms and abdominal symptoms form the CD-PRO/SS measure. It has been shown to have adequate internal consistency (Cronbach alpha statistic 0.74 for bowel signs and symptoms and 0.67 for abdominal symptoms). Test retest reliability was also evident with intraclass correlation coefficients of more than 0.8 for both modules[296].

The Ulcerative Colitis Patient reported outcome (UC-PRO) instrument has also been designed in line with the FDA guidelines. Again, this consists of 6 modules. Modules 1 (bowel signs and symptoms) and 2 (abdominal symptoms) form the UC-PRO/SS instrument. The remaining modules are systemic symptoms, coping strategies, daily life impact and emotional impact. The instrument was designed for use by patients with moderate to severe ulcerative colitis in the outpatient setting.

Internal consistency of the UC-PRO/SS was adequate. Cronbach alpha statistic was 0.8 for bowel signs and symptoms and 0.66 for abdominal signs. Test re-test reliability was evident with high intraclass correlation coefficients (ICC) for bowel signs and symptoms and abdominal signs (0.81, 0.71). UC-PRO/SS scores correlated with a number of established measures such as the IBDQ, partial Mayo score and a global physician score. Scores were also able to differential between groups based on partial Mayo score, global patient assessment and global physician assessment[297].

Both the CD-PRO and UC-PRO are valid patient reported outcome measures with similar measurement properties. They have been designed in line with FDA recommendations.

6.6 Future areas for development

Given the aim of this research was to develop at PROM for use in routine care, future plans relate to the establishment of its role day to day clinical practice. With increasing burden on outpatient clinics as well as the changing needs of IBD patients, novel means to interact with our patients are required.

The IBD Control Questionnaire could be used to record patient reported data for virtual reviews such as annual review of biological treatments. The ability of the IBD Control Questionnaire to predict patients in remission with high specificity, makes it an ideal measure for use in self-management programmes for patients with stable disease; particularly given recent advances in the provision of web-based tools such as My IBD Portal.

Practical applications of the instrument to support care decision-making should be tested, ideally in a future randomised-controlled trial. However, the adoption of PROMs into healthcare is progressing rapidly and often without formal experimental testing of the benefits and harms of using individual instruments for defined applications. At this point, we do not advocate that the IBD Control should be used as a 'device' to influence self-care decisions by patients. Rather, its adoption may allow services to gather information that compliments their existing assessments.

The proposed application is for IBD Control to serve as a 'screening' tool – with particular utility in identifying patients who are satisfied that their condition is controlled. It is, therefore, ideally placed for the management of patients in remission, possibly within a self-management programme. Patients with sub-optimal scores, or 'flags' that suggest concerns, would require further assessment. In the emerging Digital and e-Health landscape it is envisaged that the IBD-Control could serve as an initial screening tool, with supplemental questions, additional PROM completion, or a formal clinical review triggered for patients indicating sub-optimal control. Hence, a patient with good control (e.g. IBD-Control-8 score of 15) would not be required to complete additional items,

whereas a patient indicating sub-optimal control (e.g. IBD-Contro-8 score of 10) might be prompted to answer an individualised set of questions relevant to their disease. At a simple level, the supplementary questions could be akin to the interim 'PRO-2' approach – with a set of specific symptom questions for Crohn's disease, ulcerative colitis or tailored to the individual patient. These are all areas for future research and development. The developers of the TrueColoursUC app have piloted this approach, using IBD-Control-8 as the overall measure of health status, combined with patient-facing question items for a set of specific symptoms for ulcerative colitis.

It is acknowledged that IBD Control scores may be affected by the co-existence of symptoms related to irritable bowel syndrome (IBS). The proportion of IBD patients in biochemical remission with IBS symptoms has been reported to be between 19.8-27.7%[298]. However, any new symptoms, regardless of cause, should prompt further review in order to exclude active inflammation. If entering patients into a self-management programme involving the use of the IBD Control Questionnaire, there should be clear criteria for inclusion. It may be the case that patients with significant IBS symptoms are not be suitable for remote follow up in view of the unpredictability of their symptoms. Alternatively, the cut-off value for triggering addition measures, clinical review etc could be adjusted for patients with IBS. Again, this is an area for further investigation.

There is also a subset of patients with IBD who normalise their symptoms and "put up with" quite significant issues such as high stool frequency. Again, this may affect IBD Control scores. However, as mentioned above, if a patient is adequately screened before being considered for remote management, the risk of this, too, should be minimised.

The primary goal of this research was to produce a PROM that was tailored to the specific needs of busy, routine IBD services – a simple, generic tool that would be suited to all patients in the service, regardless of disease type, severity or a particular set of dominant symptoms. The questionnaire is meant to rapidly establish a global sense of disease control, and to highlight specific 'flags' for attention by healthcare staff. By explicitly excluding a 'check list' of individual gastrointestinal symptoms, our generically-focused development process avoided the replication of a traditional HRQoL instrument. We did not seek to re-invent the IBD-Q and were aware that a number of other groups were producing updated versions of traditional instruments for IBD.

The endorsement of IBD Control by ICHOM, which included international experts and strong patient representation, suggests this generic simplicity has a broad appeal. It will be interesting in future to determine how the growing range of PROMs can be used to support patient care. Future strategies might deploy PROMs individually to support care, use specific combinations of different instruments according to patient characteristics, or use electronic tools to capture a tailored sequence of PROMs according to a patient's current status (e.g. IBD-Control +/- individualized symptom survey tool +/- clinical assessment).

6.7   Summary

The IBD Control Questionnaire has been shown to be a valid patient reported outcome measure of disease control from the patient perspective. Its brevity and simple generic content make it suited to use in routine care. Independent groups have shown it can be

deployed into routine settings, online electronic surveys, patient-facing apps and patient portals.

Its inclusion in the data submission framework for the UK IBD Registry has allowed feasibility to be established for capturing the relevant items as part of hospital uploads to NHS-Digital's audit platform. This sets the scene for future electronic capture. International endorsement by ICHOM, and ongoing translations into other languages, suggest that this research has produced an instrument of genuine value to the international IBD community and filled a gap in existing PROMs.

References

1. Devlin, N. and J. Appleby. *Getting the most out of PROMs*. 2010.
2. Silverberg, M.S., et al., *Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology.* Canadian Journal of Gastroenterology and Hepatology, 2005. **19**(Suppl A): p. 5A-36A.
3. Mowat, C., et al., *Guidelines for the management of inflammatory bowel disease in adults.* Gut, 2011. **60**(5): p. 571-607.
4. Loftus, E.V. and W.J. Sandborn, *Epidemiology of inflammatory bowel disease.* Gastroenterology Clinics, 2002. **31**(1): p. 1-20.
5. Burisch, J., et al., *East–West gradient in the incidence of inflammatory bowel disease in Europe: the ECCO-EpiCom inception cohort.* Gut, 2014. **63**(4): p. 588-597.
6. Martín-de-Carpi, J., et al., *Increasing Incidence of Pediatric Inflammatory Bowel Disease in Spain (1996–2009) The SPIRIT Registry.* Inflammatory bowel diseases, 2012. **19**(1): p. 73-80.
7. Ng, S.C., et al., *Geographical variability and environmental risk factors in inflammatory bowel disease.* Gut, 2013. **62**(4): p. 630-649.
8. Calkins, B.M., *A meta-analysis of the role of smoking in inflammatory bowel disease.* Digestive diseases and sciences, 1989. **34**(12): p. 1841-1854.
9. Lindberg, E., G. Järnerot, and B. Huitfeldt, *Smoking in Crohn's disease: effect on localisation and clinical course.* Gut, 1992. **33**(6): p. 779-782.
10. Picco, M.F. and T.M. Bayless, *Tobacco consumption and disease duration are associated with fistulizing and stricturing behaviors in the first 8 years of Crohn's disease.* The American journal of gastroenterology, 2003. **98**(2): p. 363-368.
11. Sutherland, L.R., et al., *Effect of cigarette smoking on recurrence of Crohn's disease.* Gastroenterology, 1990. **98**(5 Pt 1): p. 1123-8.
12. Cosnes, J., et al., *Effects of current and former cigarette smoking on the clinical course of Crohn's disease.* Alimentary pharmacology & therapeutics, 1999. **13**(11): p. 1403-1411.
13. Harries, A.D., A. Baird, and J. Rhodes, *Non-smoking: a feature of ulcerative colitis.* British medical journal (Clinical research ed.), 1982. **284**(6317): p. 706.
14. Loftus, E.V., Jr., et al., *Primary sclerosing cholangitis is associated with nonsmoking: a case-control study.* Gastroenterology, 1996. **110**(5): p. 1496-502.
15. Merrett, M.N., et al., *Smoking may prevent pouchitis in patients with restorative proctocolectomy for ulcerative colitis.* Gut, 1996. **38**(3): p. 362-4.
16. Koutroubakis, I.E., I.G. Vlachonikolis, and E.A. Kouroumalis, *Role of appendicitis and appendectomy in the pathogenesis of ulcerative colitis: a critical review.* Inflammatory bowel diseases, 2002. **8**(4): p. 277-286.
17. Gardenbroek, T.J., et al., *The effect of appendectomy on the course of ulcerative colitis: a systematic review.* Colorectal Disease, 2012. **14**(5): p. 545-553.
18. Andersson, R.E., et al., *Appendectomy is followed by increased risk of Crohn's disease.* Gastroenterology, 2003. **124**(1): p. 40-46.
19. Cornish, J.A., et al., *The risk of oral contraceptives in the etiology of inflammatory bowel disease: a meta-analysis.* The American journal of gastroenterology, 2008. **103**(9): p. 2394.

20.   Riordan, A.M., C.H.S. Ruxton, and J.O. Hunter, *A review of associations between Crohn's disease and consumption of sugars.* European journal of clinical nutrition, 1998. **52**(4): p. 229.

21.   Geerling, B.J., et al., *Diet as a risk factor for the development of ulcerative colitis.* The American journal of gastroenterology, 2000. **95**(4): p. 1008-1013.

22.   Shoda, R., et al., *Epidemiologic analysis of Crohn disease in Japan: increased dietary intake of n-6 polyunsaturated fatty acids and animal protein relates to the increased incidence of Crohn disease in Japan.* The American journal of clinical nutrition, 1996. **63**(5): p. 741-745.

23.   Kaser, A., S. Zeissig, and R.S. Blumberg, *Genes and environment: how will our concepts on the pathophysiology of IBD develop in the future?* Digestive diseases, 2010. **28**(3): p. 395-405.

24.   Halme, L., et al., *Family and twin studies in inflammatory bowel disease.* World journal of gastroenterology: WJG, 2006. **12**(23): p. 3668.

25.   Maillard, M.H., et al., *The Wiskott-Aldrich syndrome protein is required for the function of CD4+ CD25+ Foxp3+ regulatory T cells.* Journal of Experimental Medicine, 2007. **204**(2): p. 381-391.

26.   Bennett, C.L., et al., *The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3.* Nature genetics, 2001. **27**(1): p. 20.

27.   Garrett, W.S., J.I. Gordon, and L.H. Glimcher, *Homeostasis and inflammation in the intestine.* Cell, 2010. **140**(6): p. 859-870.

28.   Wehkamp, J., K. Fellermann, and E.F. Stange, *Human defensins in Crohn's disease*, in *Mechanisms of Epithelial Defense.* 2005, Karger Publishers. p. 42-54.

29.   Fritz, T., et al., *Crohn's disease: NOD2, autophagy and ER stress converge.* Gut, 2011. **60**(11): p. 1580-1588.

30.   Cooney, R., et al., *NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation.* Nature medicine, 2010. **16**(1): p. 90.

31.   Watanabe, T., et al., *Nucleotide binding oligomerization domain 2 deficiency leads to dysregulated TLR2 signaling and induction of antigen-specific colitis.* Immunity, 2006. **25**(3): p. 473-485.

32.   Kaser, A., et al., *XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease.* Cell, 2008. **134**(5): p. 743-756.

33.   Frank, D.N., et al., *Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases.* Proceedings of the National Academy of Sciences, 2007. **104**(34): p. 13780-13785.

34.   Sartor, R.B., *Microbial influences in inflammatory bowel diseases.* Gastroenterology, 2008. **134**(2): p. 577-594.

35.   Sokol, H., et al., *Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients.* Proceedings of the National Academy of Sciences, 2008. **105**(43): p. 16731-16736.

36.   Danese, S. and C. Fiocchi, *Ulcerative Colitis.* New England Journal of Medicine, 2011. **365**(18): p. 1713-1725.

37.   Solberg, I.C., et al., *Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN Study).* Scandinavian journal of gastroenterology, 2009. **44**(4): p. 431-440.

38.   Sands, B.E., *From symptom to diagnosis: clinical distinctions among various forms of intestinal inflammation.* Gastroenterology, 2004. **126**(6): p. 1518-1532.

39. Van Assche, G., et al., *The second European evidence-based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis.* Journal of Crohn's and Colitis, 2010. **4**(1): p. 7-27.

40. Beaugerie, L., et al., *Predictors of Crohn's disease.* Gastroenterology, 2006. **130**(3): p. 650-656.

41. Danese, S., et al., *Extraintestinal manifestations in inflammatory bowel disease.* World Journal of Gastroenterology, 2005. **11**(46): p. 7227.

42. Levine, J.S. and R. Burakoff, *Extraintestinal manifestations of inflammatory bowel disease.* Gastroenterology & hepatology, 2011. **7**(4): p. 235-241.

43. Fiorino, G., et al., *Review article: anti TNF-α induced psoriasis in patients with inflammatory bowel disease.* Alimentary Pharmacology & Therapeutics, 2009. **29**(9): p. 921-927.

44. Rojas-Feria, M., et al., *Hepatobiliary manifestations in inflammatory bowel disease: the gut, the drugs and the liver.* World journal of gastroenterology, 2013. **19**(42): p. 7327-7340.

45. Angulo, P., Y. Maor-Kendler, and K.D. Lindor, *Small-duct primary sclerosing cholangitis: a long-term follow-up study.* Hepatology, 2002. **35**(6): p. 1494-1500.

46. Woodward, J. and J. Neuberger, *Autoimmune overlap syndromes.* Hepatology, 2001. **33**(4): p. 994-1002.

47. Parente, F., et al., *Incidence and risk factors for gallstones in patients with inflammatory bowel disease: a large case-control study.* Hepatology, 2007. **45**(5): p. 1267-1274.

48. Talbot, R.W., et al. *Vascular complications of inflammatory bowel disease.* Elsevier.

49. Gisbert, J.P., et al., *Liver injury in inflammatory bowel disease: Long-term follow-up study of 786 patients.* Inflammatory bowel diseases, 2007. **13**(9): p. 1106-1114.

50. Danese, S., et al., *Inflammation and coagulation in inflammatory bowel disease: the clot thickens.* The American journal of gastroenterology, 2007. **102**(1): p. 174.

51. Van Assche, G., et al., *Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 3: special situations.* Journal of Crohn's and Colitis, 2013. **7**(1): p. 1-33.

52. Eaden, J.A., K.R. Abrams, and J.F. Mayberry, *The risk of colorectal cancer in ulcerative colitis: a meta-analysis.* Gut, 2001. **48**(4): p. 526-535.

53. Bergeron, V., et al., *Risk factors for neoplasia in inflammatory bowel disease patients with pancolitis.* The American journal of gastroenterology, 2010. **105**(11): p. 2405.

54. Sandborn, W.J. and S.B. Hanauer, *The pharmacokinetic profiles of oral mesalazine formulations and mesalazine pro-drugs used in the management of ulcerative colitis.* Alimentary pharmacology & therapeutics, 2003. **17**(1): p. 29-42.

55. Eaden, J., et al., *Colorectal cancer prevention in ulcerative colitis: a case-control study.* Alimentary pharmacology & therapeutics, 2000. **14**(2): p. 145-153.

56. Hanauer, S.B. and U. Strömberg, *Oral Pentasa in the treatment of active Crohn's disease: a meta-analysis of double-blind, placebo-controlled trials.* Clinical Gastroenterology and Hepatology, 2004. **2**(5): p. 379-388.

57. Akobeng, A.K. and E. Gardener, *Oral 5-aminosalicylic acid for maintenance of medically-induced remission in Crohn's disease.* Cochrane database of systematic reviews, 2005(1).

58. Baron, J.H., et al., *Out-patient treatment of ulcerative colitis.* British medical journal, 1962. **2**(5302): p. 441.

59. Lofberg, R., et al., *Oral budesonide versus prednisolone in patients with active extensive and left-sided ulcerative colitis.* Gastroenterology, 1996. **110**(6): p. 1713-1718.

60. Lee, F.I., et al., *A randomised trial comparing mesalazine and prednisolone foam enemas in patients with acute distal ulcerative colitis.* Gut, 1996. **38**(2): p. 229-233.

61. Malchow, H., et al., *European Cooperative Crohn's Disease Study (ECCDS): results of drug treatment.* Gastroenterology, 1984. **86**(2): p. 249-266.

62. Munkholm, P., et al., *Frequency of glucocorticoid resistance and dependency in Crohn's disease.* Gut, 1994. **35**(3): p. 360-362.

63. Tiede, I., et al., *CD28-dependent Rac1 activation is the molecular target of azathioprine in primary human CD4+ T lymphocytes.* The Journal of clinical investigation, 2003. **111**(8): p. 1133-1145.

64. Timmer, A., et al., *Azathioprine and 6-mercaptopurine for maintenance of remission in ulcerative colitis.* Cochrane database of systematic reviews, 2012(9).

65. Prefontaine, E., J.K. MacDonald, and L.R. Sutherland, *Azathioprine or 6-mercaptopurine for induction of remission in Crohn's disease.* Cochrane database of systematic reviews, 2010(6).

66. Prefontaine, E., et al., *Azathioprine or 6-mercaptopurine for maintenance of remission in Crohn's disease.* Cochrane Database of Systematic Reviews, 2009(1).

67. Alfadhli, A.A.F., J.W.D. McDonald, and B.G. Feagan, *Methotrexate for induction of remission in refractory Crohn's disease.* Cochrane Database of Systematic Reviews, 2004(4).

68. Patel, V., et al., *Methotrexate for maintenance of remission in Crohn's disease.* Cochrane Database of Systematic Reviews, 2009(4).

69. Fraser, A.G., et al., *The efficacy of methotrexate for maintaining remission in inflammatory bowel disease.* Alimentary pharmacology & therapeutics, 2002. **16**(4): p. 693-697.

70. *Infliximab for acute exacerbations of ulcerative colitis (TA163).* 2008, National Institute for Health and Care Excellence.

71. Haslam, N., S.D. Hearing, and C.S. Probert, *Audit of cyclosporin use in inflammatory bowel disease: limited benefits, numerous side-effects.* European journal of gastroenterology & hepatology, 2000. **12**(6): p. 657-660.

72. Seagrove, A.C., et al., *Randomised controlled trial. Comparison Of iNfliximab and ciclosporin in STeroid Resistant Ulcerative Colitis: Trial design and protocol (CONSTRUCT).* BMJ open, 2014. **4**(4): p. e005091.

73. McDonald, J.W.D., et al., *Cyclosporine for induction of remission in Crohn's disease.* Cochrane Database of Systematic Reviews, 2005(2).

74. *Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (TA329).* 2015, National Institute for Health and Care Excellence.

75. Targan, S.R., et al., *A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor α for Crohn's disease.* New England Journal of Medicine, 1997. **337**(15): p. 1029-1036.

76. Hanauer, S.B., et al., *Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial.* The Lancet, 2002. **359**(9317): p. 1541-1549.

77. Hanauer, S.B., et al., *Human anti–tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I Trial.* Gastroenterology, 2006. **130**(2): p. 323-333.

78. Sandborn, W.J., et al., *Adalimumab induction therapy for Crohn disease previously treated with infliximab: a randomized trial.* Annals of internal medicine, 2007. **146**(12): p. 829-838.

79. Sandborn, W.J., et al., *Adalimumab for maintenance treatment of Crohn's disease: results of the CLASSIC II trial.* Gut, 2007.

80. Sands, B.E., et al., *Infliximab maintenance therapy for fistulizing Crohn's disease.* New England Journal of Medicine, 2004. **350**(9): p. 876-885.

81. Colombel, J.F., et al., *Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial.* Gastroenterology, 2007. **132**(1): p. 52-65.

82. Järnerot, G., et al., *Infliximab as rescue therapy in severe to moderately severe ulcerative colitis: a randomized, placebo-controlled study.* Gastroenterology, 2005. **128**(7): p. 1805-1811.

83. Rutgeerts, P., et al., *Infliximab for induction and maintenance therapy for ulcerative colitis.* New England Journal of Medicine, 2005. **353**(23): p. 2462-2476.

84. Reinisch, W., et al., *Adalimumab for induction of clinical remission in moderately to severely active ulcerative colitis: results of a randomised controlled trial.* Gut, 2011. **60**(6): p. 780.

85. Sandborn, W.J., et al., *Subcutaneous Golimumab Induces Clinical Response and Remission in Patients With Moderate-to-Severe Ulcerative Colitis.* Gastroenterology, 2014. **146**(1): p. 85-95.

86. Sandborn, W.J., et al., *Subcutaneous Golimumab Maintains Clinical Response in Patients With Moderate-to-Severe Ulcerative Colitis.* Gastroenterology, 2014. **146**(1): p. 96-109.e1.

87. Probert, C.S.J., et al., *Golimumab induction and maintenance for moderate to severe ulcerative colitis: results from GO-COLITIS (Golimumab: a Phase 4, UK, open label, single arm study on its utilization and impact in ulcerative Colitis).* BMJ Open Gastroenterology, 2018. **5**(1): p. e000212.

88. *Infliximab and adalimumab for the treatment of Crohn's disease (TA187).* 2010, National Institute for Health and Care Excellence.

89. Feagan, B.G., et al., *Vedolizumab as Induction and Maintenance Therapy for Ulcerative Colitis.* New England Journal of Medicine, 2013. **369**(8): p. 699-710.

90. Sandborn, W.J., et al., *Vedolizumab as Induction and Maintenance Therapy for Crohn's Disease.* New England Journal of Medicine, 2013. **369**(8): p. 711-721.

91. *Vedolizumab for treating moderately to severely active ulcerative colitis (TA342).* 2015, National Institute for Health and Care Excellence.

92. *Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy (TA352).* 2015, National Institute for Health and Care Excellence.

93. Deepak, P. and E.V. Loftus, Jr., *Ustekinumab in treatment of Crohn's disease: design, development, and potential place in therapy.* Drug design, development and therapy, 2016. **10**: p. 3685-3698.

94. Feagan, B.G., et al., *Ustekinumab as Induction and Maintenance Therapy for Crohn's Disease.* New England Journal of Medicine, 2016. **375**(20): p. 1946-1960.

95. *Ustekinumab for moderately to severely active Crohn's disease after previous treatment (TA456).* 2017, National Institute for Health and Care Excellence.

96. Sandborn, W.J., et al., *Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis.* New England Journal of Medicine, 2017. **376**(18): p. 1723-1736.

97. *Tofacitinib for moderately to severely active ulcerative colitis (TA547).* 2018, National Institute for Health and Care Excellence.

98. Thia, K.T., et al., *Ciprofloxacin or metronidazole for the treatment of perianal fistulas in patients with Crohn's disease: a randomized, double-blind, placebo-controlled pilot study.* Inflammatory bowel diseases, 2008. **15**(1): p. 17-24.

99. Fell, J.M., et al., *Mucosal healing and a fall in mucosal pro-inflammatory cytokine mRNA induced by a specific oral polymeric diet in paediatric Crohn's disease.* Alimentary pharmacology & therapeutics, 2000. **14**(3): p. 281-289.

100. Mimura, T., et al., *Once daily high dose probiotic therapy (VSL#3) for maintaining remission in recurrent or refractory pouchitis.* Gut, 2004. **53**(1): p. 108-14.

101. Bibiloni, R., et al., *VSL# 3 probiotic-mixture induces remission in patients with active ulcerative colitis.* The American journal of gastroenterology, 2005. **100**(7): p. 1539.

102. Bernell, O., A. Lapidus, and G. Hellers, *Risk factors for surgery and recurrence in 907 patients with primary ileocaecal Crohn's disease.* British journal of surgery, 2000. **87**(12): p. 1697-1701.

103. Rutgeerts, P., et al., *Predictability of the postoperative course of Crohn's disease.* Gastroenterology, 1990. **99**(4): p. 956-963.

104. Travis, S.P.L., et al., *defining remission in ulcerative colitis.* Alimentary pharmacology & therapeutics, 2011. **34**(2): p. 113-124.

105. Thia, K.T., et al., *Risk factors associated with progression to intestinal complications of Crohn's disease in a population-based cohort.* Gastroenterology, 2010. **139**(4): p. 1147-1155.

106. Bouguen, G., et al., *Treat to target: a proposed new paradigm for the management of Crohn's disease.* Clinical Gastroenterology and Hepatology, 2015. **13**(6): p. 1042-1050.

107. D'incà, R. and R. Caccaro, *Measuring disease activity in Crohn's disease: what is currently available to the clinician.* Clinical and experimental gastroenterology, 2014. **7**: p. 151.

108. Mary, J.-Y. and R. Modigliani, *Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID).* Gut, 1989. **30**(7): p. 983-989.

109. Daperno, M., et al., *Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD.* Gastrointestinal endoscopy, 2004. **60**(4): p. 505-512.

110. Sipponen, T., et al., *Endoscopic evaluation of Crohn's disease activity: comparison of the CDEIS and the SES-CD.* Inflammatory bowel diseases, 2010. **16**(12): p. 2131-2136.

111. Samaan, M.A., et al., *A systematic review of the measurement of endoscopic healing in ulcerative colitis clinical trials: recommendations and implications for future research.* Inflammatory bowel diseases, 2014. **20**(8): p. 1465-1471.

112. Schroeder, K.W., W.J. Tremaine, and D.M. Ilstrup, *Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis.* New England Journal of Medicine, 1987. **317**(26): p. 1625-1629.

113. Travis, S.P.L., et al., *Developing an instrument to assess the endoscopic severity of ulcerative colitis: the Ulcerative Colitis Endoscopic Index of Severity (UCEIS).* Gut, 2012. **61**(4): p. 535-542.

114. Panes, J., et al., *Imaging techniques for assessment of inflammatory bowel disease: joint ECCO and ESGAR evidence-based consensus guidelines.* Journal of Crohn's and Colitis, 2013. **7**(7): p. 556-585.

115. Fiorino, G., et al., *Prospective comparison of computed tomography enterography and magnetic resonance enterography for assessment of disease activity and complications in ileocolonic Crohn's disease.* Inflammatory bowel diseases, 2010. **17**(5): p. 1073-1080.

116. Patel, B., et al., *MDCT assessment of ulcerative colitis: radiologic analysis with clinical, endoscopic, and pathologic correlation.* Abdominal imaging, 2012. **37**(1): p. 61-69.

117.   Andersen, K., et al., *Multi-detector CT-colonography in inflammatory bowel disease: prospective analysis of CT-findings to high-resolution video colonoscopy.* European journal of radiology, 2006. **58**(1): p. 140-146.

118.   Stange, E.F., et al., *European evidence-based consensus on the diagnosis and management of ulcerative colitis: definitions and diagnosis.* Journal of Crohn's and Colitis, 2008. **2**(1): p. 1-23.

119.   Masselli, G. and G. Gualdi, *CT and MR enterography in evaluating small bowel diseases: when to use which modality?* Abdominal imaging, 2013. **38**(2): p. 249-259.

120.   Rimola, J., et al., *Role of 3.0-T MR colonography in the evaluation of inflammatory bowel disease.* Radiographics, 2009. **29**(3): p. 701-719.

121.   Sandor Kiss, L., et al., *High-sensitivity C-reactive protein for identification of disease phenotype, active disease, and clinical relapses in Crohn's disease: a marker for patient classification?* Inflammatory bowel diseases, 2011. **18**(9): p. 1647-1654.

122.   Sipponen, T., et al., *Correlation of faecal calprotectin and lactoferrin with an endoscopic score for Crohn's disease and histological findings.* Alimentary pharmacology & therapeutics, 2008. **28**(10): p. 1221-1229.

123.   Mao, R., et al., *Fecal calprotectin in predicting relapse of inflammatory bowel diseases: a meta-analysis of prospective studies.* Inflammatory bowel diseases, 2012. **18**(10): p. 1894-1899.

124.   Sipponen, T., et al., *Faecal calprotectin and lactoferrin are reliable surrogate markers of endoscopic response during Crohn's disease treatment.* Scandinavian journal of gastroenterology, 2010. **45**(3): p. 325-331.

125.   Truelove, S.C. and L.J. Witts, *Cortisone in ulcerative colitis.* British medical journal, 1955. **2**(4947): p. 1041.

126.   Dignass, A., et al., *Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis.* Journal of Crohn's and Colitis, 2012. **6**(10): p. 965-990.

127.   Powell-Tuck, J., R.L. Bown, and J.E. Lennard-Jones, *A comparison of oral prednisolone given as single or multiple daily doses for active proctocolitis.* Scandinavian journal of gastroenterology, 1978. **13**(7): p. 833-837.

128.   Walmsley, R.S., et al., *A simple clinical colitis activity index.* Gut, 1998. **43**(1): p. 29-32.

129.   Higgins, P.D., et al., *Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis.* Gut, 2005. **54**(6): p. 782-8.

130.   Levesque, B., et al., *Tu1119 Responsiveness of Central Endoscopic Assessment of Disease Activity Using the Modified Mayo Clinic Score in Ulcerative Colitis.* Gastroenterology, 2013. **144**(5): p. S-767.

131.   Feagan, B.G., et al., *The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients.* Am J Gastroenterol, 2007. **102**(4): p. 794-802.

132.   Travis, S.P.L., et al., *Reliability and initial validation of the ulcerative colitis endoscopic index of severity.* Gastroenterology, 2013. **145**(5): p. 987-995.

133.   Irvine, E.J., et al., *Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group.* Gastroenterology, 1994. **106**(2): p. 287-96.

134.   Sutherland, L.R., et al., *5-Aminosalicylic acid enema in the treatment of distal ulcerative colitis, proctosigmoiditis, and proctitis.* Gastroenterology, 1987. **92**(6): p. 1894-1898.

135.   Best, W.R., et al., *Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study.* Gastroenterology, 1976. **70**(3): p. 439-44.

136.   Harvey, R.F. and J.M. Bradshaw, *A simple index of Crohn's-disease activity.* Lancet, 1980. **1**(8167): p. 514.

137.   Myren, J., et al., *The OMGE Multinational Inflammatory Bowel Disease Survey 1976-1982. A further report on 2,657 cases.* Scandinavian journal of gastroenterology. Supplement, 1984. **95**: p. 1-27.

138.   Wright Jp Fau - Marks, I.N., A. Marks In Fau - Parfitt, and A. Parfitt, *A simple clinical index of Crohn's disease activity--the Cape Town index.* (0256-9574 (Print)).

139.   De Dombal, F.T. and A. Softley, *IOIBD report no 1: Observer variation in calculating indices of severity and activity in Crohn's disease. International Organisation for the Study of Inflammatory Bowel Disease.* Gut, 1987. **28**(4): p. 474-481.

140.   Sandborn, W.J., et al., *A review of activity indices and efficacy endpoints for clinical trials of medical therapy in adults with Crohn's disease.* Gastroenterology, 2002. **122**(2): p. 512-530.

141.   Irvine, E.J., *Usual therapy improves perianal Crohn's disease as measured by a new disease activity index. McMaster IBD Study Group.* Journal of clinical gastroenterology, 1995. **20**(1): p. 27-32.

142.   Fitzpatrick, R., et al., *Evaluating patient-based outcome measures for use in clinical trials.* Health Technol Assess, 1998. **2**(14): p. i-iv, 1-74.

143.   Guyatt, G., et al., *A new measure of health status for clinical trials in inflammatory bowel disease.* Gastroenterology, 1989. **96**(3): p. 804-10.

144.   Patrick, D.L. and R.A. Deyo, *Generic and disease-specific measures in assessing health status and quality of life.* Medical care, 1989: p. S217-S232.

145.   Ware, J.E., Jr. and C.D. Sherbourne, *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.* Med Care, 1992. **30**(6): p. 473-83.

146.   Bakker, C., *Health related utility measurement: an introduction.* The Journal of rheumatology, 1995. **22**(6): p. 1197-1199.

147.   Read, J.L., et al., *Preferences for health outcomes: comparison of assessment methods.* Medical Decision Making, 1984. **4**(3): p. 315-329.

148.   Brooks, R., *EuroQol: the current state of play.* Health Policy, 1996. **37**(1): p. 53-72.

149.   Guyatt, G., D. Feeny, and D. Patrick, *Issues in quality-of-life measurement in clinical trials.* Controlled Clinical Trials, 1991. **12**(4, Supplement): p. S81-S90.

150.   Bombardier, C., et al., *Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery.* Medical care, 1995: p. AS131-AS144.

151.   Marx, R.G., et al., *A comparison of two time intervals for test-retest reliability of health status instruments.* Journal of clinical epidemiology, 2003. **56**(8): p. 730-735.

152.   Streiner, D.L., G.R. Norman, and J. Cairney, *Health measurement scales: a practical guide to their development and use.* 2015: Oxford University Press, USA.

153.   Guyatt, G.H., D.H. Feeny, and D.L. Patrick, *MEasuring health-related quality of life.* Annals of Internal Medicine, 1993. **118**(8): p. 622-629.

154.   Bergner, M. and M.L. Rothman, *Health status measures: an overview and guide for selection.* Annual review of public health, 1987. **8**(1): p. 191-210.

155.   Campbell, D.T. and D.W. Fiske, *Convergent and discriminant validation by the multitrait-multimethod matrix.* Psychological bulletin, 1959. **56**(2): p. 81.

156.   Kazis, L.E., J.J. Anderson, and R.F. Meenan, *Effect sizes for interpreting changes in health status.* Medical care, 1989: p. S178-S189.

157.   Liang, M.H., A.H. Fossel, and M.G. Larson, *Comparisons of five health status instruments for orthopedic evaluation.* Medical care, 1990: p. 632-642.

158.	Guyatt, G., S. Walter, and G. Norman, *Measuring change over time: assessing the usefulness of evaluative instruments.* Journal of chronic diseases, 1987. **40**(2): p. 171-178.

159.	Deyo, R.A. and R.M. Centor, *Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance.* Journal of chronic diseases, 1986. **39**(11): p. 897-906.

160.	Kessler, R.C. and D.K. Mroczek, *Measuring the effects of medical interventions.* Medical Care, 1995: p. AS109-AS119.

161.	*High Quality Care for All NHS next stage review final report*, D.o. Health, Editor. 2008.

162.	*Equity and excellence: liberating the NHS*, D.o. Health, Editor. 2010.

163.	*The NHS Outcomes Framework 2014/15*, D.o. Health, Editor. 2013.

164.	*Guidance on the routine collection of Patient Reported Outcome Measures (PROMs)*, D.o. Health, Editor. 2008.

165.	Richardson, G. and A. Manca, *Calculation of quality adjusted life years in the published literature: a review of methodology and transparency.* Health Economics, 2004. **13**(12): p. 1203-1210.

166.	Wright, E.P., et al., *Development and evaluation of an instrument to assess social difficulties in routine oncology practice.* Quality of Life Research, 2005. **14**(2): p. 373-386.

167.	*Guide to the Methods of Technology Appraisal.* 2013, National Institute for Health and Care Excellence.

168.	*Improving health and well-being of people with long term conditions*, D.o. Health, Editor. 2010.

169.	Heyworth, I.T.M., et al., *How do common chronic conditions affect health-related quality of life?* Br J Gen Pract, 2009. **59**(568): p. e353-e358.

170.	Peters, M., et al., *The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey.* BMJ open, 2014. **4**(2): p. e003968.

171.	Roland, M., et al., *Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England.* Bmj, 2009. **339**: p. b3851.

172.	*EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group.* Health Policy, 1990. **16**(3): p. 199-208.

173.	*The NHS Outcomes Framework 2014/15 Technical Appendix*, D.o. Health, Editor. 2014/2015.

174.	Garratt, A.M., et al., *Towards measurement of outcome for patients with varicose veins.* BMJ Quality & Safety, 1993. **2**(1): p. 5-10.

175.	Dawson, J., et al., *Questionnaire on the perceptions of patients about total hip replacement.* The Journal of bone and joint surgery. British volume, 1996. **78**(2): p. 185-190.

176.	Dawson, J., et al., *Questionnaire on the perceptions of patients about total knee replacement.* The Journal of bone and joint surgery. British volume, 1998. **80**(1): p. 63-69.

177.	*Living with and beyond cancer: taking actions to improve outcome (and update to the 2010 The National Cancer Survivor Inititiative Vision)*, D.o. Health, Editor. 2013.

178.	*Quality of life of cancer survivors in England: Report on a pilot survey using Patient Reported Outcome Measures (PROMS)*, D.o. Health, Editor. 2012.

179.	Burge, P., et al., *Understanding patients' choices at the point of referral.* 2006.

180.	Robertson, R. and A. Dixon, *Choice at the point of referral.* Report, The King's Fund, London, 2009.

181. Fung, C.H. and R.D. Hays, *Prospects and challenges in using patient-reported outcomes in clinical practice.* Quality of Life Research, 2008. **17**(10): p. 1297-1302.

182. Snyder, C.F. and N.K. Aaronson, *Use of patient-reported outcomes in clinical practice.* The Lancet, 2009. **374**(9687): p. 369-370.

183. Feldman-Stewart, D. and M.D. Brundage, *A conceptual framework for patient–provider communication: a tool in the PRO research tool box.* Quality of Life Research, 2009. **18**(1): p. 109.

184. Varni, J.W., C.A. Limbers, and T.M. Burwinkle, *Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL™ 4.0 Generic Core Scales.* Health and quality of life outcomes, 2007. **5**(1): p. 2.

185. Drossman, D.A., et al., *The rating form of IBD patient concerns: a new measure of health status.* Psychosom Med, 1991. **53**(6): p. 701-12.

186. Drossman, D.A., et al., *Ulcerative colitis and Crohn's disease health status scales for research and clinical practice.* J Clin Gastroenterol, 1992. **15**(2): p. 104-12.

187. Han, S.W., et al., *The inflammatory bowel disease questionnaire: a valid and reliable measure in ulcerative colitis patients in the North East of England.* Scand J Gastroenterol, 1998. **33**(9): p. 961-6.

188. Irvine, E.J., Q. Zhou, and A.K. Thompson, *The Short Inflammatory Bowel Disease Questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators. Canadian Crohn's Relapse Prevention Trial.* Am J Gastroenterol, 1996. **91**(8): p. 1571-8.

189. Konig, H.H., et al., *Validation of the EuroQol questionnaire in patients with inflammatory bowel disease.* Eur J Gastroenterol Hepatol, 2002. **14**(11): p. 1205-15.

190. Stark, R.G., et al., *Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany.* Inflamm Bowel Dis, 2010. **16**(1): p. 42-51.

191. *UK IBD Audit 2006: National results of the organisation and process of IBD care in the IK*. 2007.

192. *Quality Care: service standards for the healthcare of people who have inflammatory bowel disease (IBD)*. 2009.

193. Shaw, I. and E. Fernandez, *PTU-079 The Inflammatory Bowel Disease Quality Improvement Project – Progress Two Years on.* Gut, 2013. **62**(Suppl 1): p. A78.

194. *Report of the results for the national clinical audit of adult inflammatory bowel disease inpatient care in the UK. Round 3*. 2012.

195. Cheung, W.Y., et al., *The UK IBDQ-a British version of the inflammatory bowel disease questionnaire. development and validation.* J Clin Epidemiol, 2000. **53**(3): p. 297-306.

196. Reilly, M.C., et al., *Validity, reliability, and responsiveness of the work productivity and activity impairment questionnaire in Crohn's disease.* Clin Ther, 2008. **30**(2): p. 393-404.

197. Keefer, L., J.L. Kiebles, and T.H. Taft, *The role of self-efficacy in inflammatory bowel disease management: Preliminary validation of a disease-specific measure.* Inflamm Bowel Dis, 2010.

198. Coyne, K., et al., *Validation of the treatment satisfaction questionnaire for Crohn's disease (TSQ-C).* Dig Dis Sci, 2005. **50**(2): p. 252-8.

199. Love, J.R., E.J. Irvine, and R.N. Fedorak, *Quality of life in inflammatory bowel disease.* J Clin Gastroenterol, 1992. **14**(1): p. 15-9.

200. Alcala, M.J., et al., *Shortened questionnaire on quality of life for inflammatory bowel disease.* Inflamm Bowel Dis, 2004. **10**(4): p. 383-91.

201. Hjortswang, H., et al., *The Short Health Scale: a valid measure of subjective health in ulcerative colitis.* Scand J Gastroenterol, 2006. **41**(10): p. 1196-203.

202. Farmer, R.G., K.A. Easley, and J.M. Farmer, *Quality of life assessment by patients with inflammatory bowel disease.* Cleve Clin J Med, 1992. **59**(1): p. 35-42.

203. Smith, G.D., R. Watson, and K.R. Palmer, *Inflammatory bowel disease: developing a short disease specific scale to measure health related quality of life.* Int J Nurs Stud, 2002. **39**(6): p. 583-90.

204. Vergara, M., et al., *Assessing the quality of life of household members of patients with inflammatory bowel disease: development and validation of a specific questionnaire.* Am J Gastroenterol, 2002. **97**(6): p. 1429-37.

205. Smith, J.J., et al., *Development of a social morbidity score in patients with chronic ulcerative colitis as a potential guide to treatment.* (1463-1318 (Electronic)).

206. Fazio, V.W., et al., *Long-term functional outcome and quality of life after stapled restorative proctocolectomy.* Ann Surg, 1999. **230**(4): p. 575-84; discussion 584-6.

207. Wilcox, A.R., et al., *A new tool to measure the burden of Crohn's disease and its treatment: do patient and physician perceptions match?* Inflamm Bowel Dis, 2010. **16**(4): p. 645-50.

208. Surti, B., et al., *Assessing health status in inflammatory bowel disease using a novel single-item numeric rating scale.* Digestive diseases and sciences, 2013. **58**(5): p. 1313-1321.

209. van der Eijk, I., et al., *Quality of health care in inflammatory bowel disease: development of a reliable questionnaire (QUOTE-IBD) and first results.* Am J Gastroenterol, 2001. **96**(12): p. 3329-36.

210. Eaden, J.A., K. Abrams, and J.F. Mayberry, *The Crohn's and Colitis Knowledge Score: a test for measuring patient knowledge in inflammatory bowel disease.* Am J Gastroenterol, 1999. **94**(12): p. 3560-6.

211. Jones, S.C., et al., *A patient knowledge questionnaire in inflammatory bowel disease.* J Clin Gastroenterol, 1993. **17**(1): p. 21-4.

212. Selinger, C.P., et al., *Patients' knowledge of pregnancy-related issues in inflammatory bowel disease and validation of a novel assessment tool ('CCPKnow').* Alimentary pharmacology & therapeutics, 2012. **36**(1): p. 57-63.

213. Keegan, D., et al., *Development, validation and clinical assessment of a short questionnaire to assess disease-related knowledge in inflammatory bowel disease patients.* Scandinavian journal of gastroenterology, 2013. **48**(2): p. 183-188.

214. Vergara, M., et al., *Development and validation of the Crohn's disease perceived work disability questionnaire.* Inflamm Bowel Dis, 2011. **17**(11): p. 2350-7.

215. Peyrin-Biroulet, L., et al., *Development of the first disability index for inflammatory bowel disease based on the international classification of functioning, disability and health.* Gut, 2012. **61**(2): p. 241-247.

216. Allen, P.B., et al., *Development and validation of a patient-reported disability measurement tool for patients with inflammatory bowel disease.* Alimentary pharmacology & therapeutics, 2013. **37**(4): p. 438-444.

217. Sintonen, H. and M. Pekurinen, *A generic 15 dimentional measure of health-related quality of life (15D).* J Soc Med, 1989. **26**: p. 85-96.

218. Morisky, D.E., et al., *Predictive validity of a medication adherence measure in an outpatient setting.* The Journal of Clinical Hypertension, 2008. **10**(5): p. 348-354.

219. Yellen, S.B., et al., *Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system.* Journal of pain and symptom management, 1997. **13**(2): p. 63-74.

220. Ciccocioppo, R., et al., *Validation of the Italian translation of the Inflammatory Bowel Disease Questionnaire.* Dig Liver Dis, 2011. **43**(7): p. 535-41.

221. de Boer, A.G., et al., *Inflammatory Bowel Disease Questionnaire: cross-cultural adaptation and further validation.* Eur J Gastroenterol Hepatol, 1995. **7**(11): p. 1043-50.

222. Hauser, W., et al., *Validation of the inflammatory bowel disease questionnaire IBDQ-D, German version, for patients with ileal pouch anal anastomosis for ulcerative colitis.* Z Gastroenterol, 2004. **42**(2): p. 131-9.

223. Watanabe, K., et al., *Assessment of the Japanese Inflammatory Bowel Disease Questionnaire in patients after ileal pouch anal anastomosis for ulcerative colitis.* J Gastroenterol, 2006. **41**(7): p. 662-7.

224. Kiran, R.P., et al., *Prospective assessment of Cleveland Global Quality of Life (CGQL) as a novel marker of quality of life and disease activity in Crohn's disease.* Am J Gastroenterol, 2003. **98**(8): p. 1783-9.

225. Schiaffino, K.M., T.A. Revenson, and A. Gibofsky, *Assessing the impact of self-efficacy beliefs on adaptation to rheumatoid arthritis.* Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, 1991. **4**(4): p. 150-157.

226. Coyne, K.S., et al., *Development and validation of a disease-specific treatment satisfaction questionnaire for gastro-oesophageal reflux disease.* Alimentary pharmacology & therapeutics, 2003. **18**(9): p. 907-915.

227. Van Campen, C., et al., *Assessing patients' priorities and perceptions of the quality of health care: the development of the QUOTE-Rheumatic-Patients instrument.* British journal of rheumatology, 1998. **37**(4): p. 362-368.

228. Mazzuca, S.A., *Does patient education in chronic disease have therapeutic value?* Journal of chronic diseases, 1982. **35**(7): p. 521-529.

229. Butcher, R.O., et al., *Patient knowledge in inflammatory bowel disease: CCKNOW, how much do they know?* Inflammatory bowel diseases, 2011. **17**(10): p. E131-E132.

230. Rezailashkajani, M., et al., *Knowledge of disease and health information needs of the patients with inflammatory bowel disease in a developing country.* Int J Colorectal Dis, 2006. **21**(5): p. 433-40.

231. Luces, C. and K. Bodger, *Economic burden of inflammatory bowel disease: a UK perspective.* Expert review of pharmacoeconomics & outcomes research, 2006. **6**(4): p. 471-482.

232. Reilly, M.C., A.S. Zbrozek, and E.M. Dukes, *The validity and reproducibility of a work productivity and activity impairment instrument.* Pharmacoeconomics, 1993. **4**(5): p. 353-65.

233. Reilly, M.C., et al., *The validity and accuracy of the Work Productivity and Activity Impairment questionnaire–irritable bowel syndrome version (WPAI: IBS).* Alimentary pharmacology & therapeutics, 2004. **20**(4): p. 459-467.

234. Wahlqvist, P., et al., *Validity of a work productivity and activity impairment questionnaire for patients with symptoms of gastro-esophageal reflux disease (WPAI-GERD)—results from a cross-sectional study.* Value in Health, 2002. **5**(2): p. 106-113.

235. Ebers, G.C., et al., *Disability as an outcome in MS clinical trials.* Neurology, 2008. **71**(9): p. 624-631.

236. Peyrin-Biroulet, L., *What is the patient's perspective: How important are patient-reported outcomes, quality of life and disability?* Dig Dis, 2010. **28**(3): p. 463-71.

237. Peyrin-Biroulet, L., et al., *Disability in inflammatory bowel diseases: developing ICF Core Sets for patients with inflammatory bowel diseases based on the International Classification of Functioning, Disability, and Health.* Inflamm Bowel Dis, 2010. **16**(1): p. 15-22.

238.   Rabin, R., M. Oemar, and M. Oppe. *EQ-5D-3L User Guide. Basic information on how to use the EQ-5D-3L instrument*. 2011; 4:[Available from: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-3L.pdf.

239.   McHorney, C.A., et al., *The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts. Results from the Medical Outcomes Study.* Med Care, 1992. **30**(5 Suppl): p. MS253-65.

240.   Garratt, A.M., et al., *The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS?* BMJ, 1993. **306**(6890): p. 1440-4.

241.   Ware, J.E., M. Kosinski, and J.E. Dewey, *How to Score Version 2 of the SF-36® Health Survey (Standard & Acute Forms)* 2001, QualityMetric, Inc., : Lincoln, RI.

242.   Sintonen, H., *The 15D instrument of health-related quality of life: properties and applications.* Ann Med, 2001. **33**(5): p. 328-36.

243.   Trindade, A.J., et al., *Are your patients taking their medicine? Validation of a new adherence scale in patients with inflammatory bowel disease and comparison with physician perception of adherence.* Inflamm Bowel Dis, 2011. **17**(2): p. 599-604.

244.   Cella, D.F., et al., *The Functional Assessment of Cancer Therapy scale: development and validation of the general measure.* J Clin Oncol, 1993. **11**(3): p. 570-579.

245.   Tinsley, A., et al., *Validation of the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) in patients with inflammatory bowel disease.* Alimentary pharmacology & therapeutics, 2011. **34**(11-12): p. 1328-1336.

246.   Lohr, K.N., *Assessing health status and quality-of-life instruments: attributes and review criteria.* Quality of Life Research, 2002. **11**(3): p. 193-205.

247.   Russel, M.G., et al., *Validation of the Dutch translation of the Inflammatory Bowel Disease Questionnaire (IBDQ): a health-related quality of life questionnaire in inflammatory bowel disease.* Digestion, 1997. **58**(3): p. 282-8.

248.   Stjernman, H., et al., *Worries and concerns in a large unselected cohort of patients with Crohn's disease.* Scand J Gastroenterol, 2010. **45**(6): p. 696-706.

249.   Hjortswang, H., et al., *Evaluation of the RFIPC, a disease-specific health-related quality of life questionnaire, in Swedish patients with ulcerative colitis.* Scand J Gastroenterol, 1997. **32**(12): p. 1235-40.

250.   Stjernman, H., et al., *Short health scale: a valid, reliable, and responsive instrument for subjective health assessment in Crohn's disease.* Inflamm Bowel Dis, 2008. **14**(1): p. 47-52.

251.   Jelsness-Jørgensen, L.-P., T. Bernklev, and B. Moum, *Quality of life in patients with inflammatory bowel disease: translation, validity, reliability and sensitivity to change of the Norwegian version of the short health scale (SHS).* Quality of Life Research, 2012. **21**(9): p. 1671-1676.

252.   Jowett, S.L., et al., *The short inflammatory bowel disease questionnaire is reliable and responsive to clinically important change in ulcerative colitis.* Am J Gastroenterol, 2001. **96**(10): p. 2921-8.

253.   Hjortswang, H., et al., *Validation of the inflammatory bowel disease questionnaire in Swedish patients with ulcerative colitis.* Scand J Gastroenterol, 2001. **36**(1): p. 77-85.

254.   Lopez-Vivancos, J., et al., *Validation of the spanish version of the inflammatory bowel disease questionnaire on ulcerative colitis and Crohn's disease.* Digestion, 1999. **60**(3): p. 274-80.

255.   Juniper, E.F., et al., *Determining a minimal important change in a disease-specific quality of life questionnaire.* Journal of Clinical Epidemiology, 1994. **47**(1): p. 81-87.

256. Hlavaty, T., et al., *Evaluation of short-term responsiveness and cutoff values of inflammatory bowel disease questionnaire in Crohn's disease.* Inflamm Bowel Dis, 2006. **12**(3): p. 199-204.

257. Pallis, A.G., I.G. Vlachonikolis, and I.A. Mouzas, *Validation of a reliable instrument (QUOTE-IBD) for assessing the quality of health care in Greek patients with inflammatory bowel disease.* Digestion, 2003. **68**(2-3): p. 153-60.

258. Leong, R., et al., *P122 Validation of the International Classification of Functioning, Disability and Health score: a measure of disability in inflammatory bowel diseases.* Journal of Crohn's and Colitis, 2013. **7**: p. S57.

259. Vergara, M., et al., *A new validation of the Spanish Work Productivity and Activity Impairment Questionnaire—Crohn's disease version.* Value in Health, 2011. **14**(6): p. 859-861.

260. Reilly MC, B.M., Brabant Y, et al, *Minimally important difference for WPAI:CD scores: Defining relevant impact on work productivity in active Crohn's disease (Abst 962). ,* in *Presented at the American College of Gastroenterology Annual Scientific Meeting.* 2007: Philadelphia.

261. Bernklev, T., et al., *Health-related quality of life in patients with inflammatory bowel disease measured with the short form-36: psychometric assessments and a comparison with general population norms.* Inflamm Bowel Dis, 2005. **11**(10): p. 909-18.

262. McColl, E., et al., *A comparison of the discriminatory power of the Inflammatory Bowel Disease Questionnaire and the SF-36 in people with ulcerative colitis.* Qual Life Res, 2004. **13**(4): p. 805-11.

263. Haapamaki, J., et al., *Health-related quality of life in inflammatory bowel disease measured with the generic 15D instrument.* Qual Life Res, 2010. **19**(6): p. 919-28.

264. Loftus, E.V., et al., *Effects of adalimumab maintenance therapy on health-related quality of life of patients with Crohn's disease: patient-reported outcomes of the CHARM trial.* Am J Gastroenterol, 2008. **103**(12): p. 3132-41.

265. Sandborn, W.J., et al., *Certolizumab pegol for the treatment of Crohn's disease.* N Engl J Med, 2007. **357**(3): p. 228-38.

266. Schreiber, S., et al., *Maintenance therapy with certolizumab pegol for Crohn's disease.* New England Journal of Medicine, 2007. **357**(3): p. 239-250.

267. Coteur, G., et al., *Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease.* Aliment Pharmacol Ther, 2009. **29**(9): p. 1032-41.

268. Braun, V. and V. Clarke, *Using thematic analysis in psychology.* Qualitative research in psychology, 2006. **3**(2): p. 77-101.

269. Rubin, G.P., et al., *Quality of life in patients with established inflammatory bowel disease: a UK general practice survey.* Aliment Pharmacol Ther, 2004. **19**(5): p. 529-35.

270. Zigmond, A.S. and R.P. Snaith, *The hospital anxiety and depression scale.* Acta Psychiatr Scand, 1983. **67**(6): p. 361-70.

271. L, C., *Coefficient alpha and the internal structure of tests. .* Psychometrika, 1951. **16**: p. 297-334.

272. *Health measurement scales: a practical guide to their development and use (5th edition).* (1753-6405 (Electronic)).

273. Dolan, P., *Modeling valuations for EuroQol health states.* (0025-7079 (Print)).

274. Kazis, L.E., R.F. Anderson Jj Fau - Meenan, and R.F. Meenan, *Effect sizes for interpreting changes in health status.* (0025-7079 (Print)).

275. Koo, T.K. and M.Y. Li, *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research.* Journal of chiropractic medicine, 2016. **15**(2): p. 155-163.

276. Food and D.A.J.F. Regist, *Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims.* 2009. **74**(235): p. 65132-65133.

277. Gledhill T, B.E., Collins B, Subramanian S, Bassi A, Bodger K, *PWE-114 The Ibd-control Questionnaire: Multi-centre Validation Plus Evaluation In Routine Care* Gut, 2014. **63**: p. A174-A175.

278. Bodger, K., et al., *PWE-054 Independent validation of the ibd-control questionnaire: results from a large-scale electronic patient experience survey (ibd2020).* 2015, BMJ Publishing Group.

279. *P370. IBD-Control Questionnaire: validation and evaluation in clinical practice.* Journal of Crohn's and Colitis, 2015. **9**(suppl_1): p. S264-S264.

280. Keogh, A. and M. Burke, *N809 IBD control questionnaire: validation and evaluation as part of a virtual biologic clinic in Galway University Hospital, Ireland.* Journal of Crohn's and Colitis, 2017. **11**(suppl_1): p. S495-S496.

281. Calvert CR, L.S., Stansfield C, McLaughlin J, Robinson A, *OC-082 The Development of a Web-Based System to Enhance Care in Inflammatory Bowel Disease* Gut, 2013. **62**: p. A35-A36.

282. Mclaughlin J, S.S., Lal S, Bodger K, Robinson JA, *OTU-031 My–IBD portal: progress and impact update* Gut, 2018. **67**: p. A223.

283. Walsh, A., et al., *P360 TrueColours ulcerative colitis (TCUC): Will patients with UC complete digital questionnaires in real-time?* Journal of Crohn's and Colitis, 2018. **12**(supplement_1): p. S286-S286.

284. Porter, M.E., S. Larsson, and T.H.J.N.E.J.o.M. Lee, *Standardizing patient outcomes measurement.* 2016. **374**(6): p. 504-506.

285. Kim, A.H., et al., *Developing a standard set of patient-centred outcomes for inflammatory bowel disease—an international, cross-disciplinary consensus.* 2017. **12**(4): p. 408-418.

286. Lim, T.Y., et al., *PTU-074 Vedolizumab in Primary and Autoimmune Sclerosing Cholangitis Associated Inflammatory Bowel Disease Pre and Post Liver Transplantation: A Case Series.* Gut, 2016. **65**(Suppl 1): p. A89.

287. Parr, H., et al., *PTH-054 Home-Testing of Faecal Calprotectin Using the IBDoc™ System: A Comparative Pilot Study.* Gut, 2016. **65**(Suppl 1): p. A245.

288. Rahmany, S., et al., *PTU-073 In Patients with IBD Switching from Originator Infliximab (Remicade) to Biosimilar Infliximab (CT-P13) Is Safe and Effective.* Gut, 2016. **65**(Suppl 1): p. A89.

289. Alrubaiy, L., et al., *Development of a Short Questionnaire to Assess the Quality of Life in Crohn's Disease and Ulcerative Colitis.* Journal of Crohn's and Colitis, 2015. **9**(1): p. 66-76.

290. Hutchings, H.A., et al., *Validation of the Crohn's and Ulcerative Colitis questionnaire in patients with acute severe ulcerative colitis.* United European Gastroenterology Journal, 2016. **5**(4): p. 571-578.

291. Subramanian, S., et al., *Validation of a Simple 0 to 10 Numerical Score (IBD-10) of Patient-reported Inflammatory Bowel Disease Activity for Routine Clinical Use.* Inflammatory Bowel Diseases, 2016. **22**(8): p. 1902-1907.

292. Bojic, D., K. Bodger, and S. Travis, *Patient Reported Outcome Measures (PROMs) in Inflammatory Bowel Disease: New Data.* Journal of Crohn's and Colitis, 2017. **11**(suppl_2): p. S576-S585.

293.   Alrubaiy, L., et al., *Systematic Review of Health-related Quality of Life Measures for Inflammatory Bowel Disease.* Journal of Crohn's and Colitis, 2015. **9**(3): p. 284-292.

294.   Khanna, R., et al., *A retrospective analysis: the development of patient reported outcome measures for the assessment of Crohn's disease activity.* Alimentary pharmacology & therapeutics, 2015. **41**(1): p. 77-86.

295.   Jairath, V., et al., *Development of interim patient-reported outcome measures for the assessment of ulcerative colitis disease activity in clinical trials.* 2015. **42**(10): p. 1200-1210.

296.   Higgins, P.D.R., et al., *Development and validation of the Crohn's disease patient-reported outcomes signs and symptoms (CD-PRO/SS) diary.* Journal of Patient-Reported Outcomes, 2018. **2**(1): p. 24.

297.   Higgins, P.D.R., et al., *Development and validation of the Ulcerative Colitis patient-reported outcomes signs and symptoms (UC-pro/SS) diary.* Journal of patient-reported outcomes, 2018. **2**(1): p. 26-26.

298.   Gracie, D.J., et al., *Negative Effects on Psychological Health and Quality of Life of Genuine Irritable Bowel Syndrome–type Symptoms in Patients With Inflammatory Bowel Disease.* Clinical Gastroenterology and Hepatology, 2017. **15**(3): p. 376-384.e5.

Appendix 1 The IBD Control Questionnaire scoring system

Items and scoring for the IBD-Control-8 summary score

IBD Control

Inflammatory Bowel Disease Control Questionnaire

> The IBD-Control-8 score is based on sum of responses to eight items (1a, 1b, 3a to 3f). Each item is allocated a score of 0, 1 or 2. The score ranges from 0 (worst control) to 16 (best control).

1 Do you believe that:

	Yes	No	Not sure
a. Your IBD has been well controlled in the past *two weeks*	☐ 2	☐ 0	☐ 1
b. Your *current treatment* is useful in controlling your IBD?	☐ 2	☐ 0	☐ 1

(If you are not taking any treatment, please tick this box ☐) 1

2 Over the past 2 weeks, have your bowel symptoms been getting worse, getting better or not changed?

Better	No change	Worse
☐	☐	☐

This is a stand alone 'transition' question to monitor overall change in status

3 In the past 2 weeks, did you:

	Yes	No	Not sure
a. Miss any planned activities because of IBD? (e.g. attending school/college, going to work or a social event)	☐ 0	☐ 2	☐ 1
b. Wake up at night because of symptoms of	☐ 0	☐ 2	☐ 1
c. Suffer from significant pain or discomf	☐ 0	☐ 2	☐ 1
d. Often feel lacking in energy (fatigued) (by 'often' we mean more than half of the the time)	☐ 0	☐ 2	☐ 1
e. Feel anxious or depressed because of your IBD	☐ 0	☐ 2	☐ 1
f. Think you needed a change to your treatment?	☐ 0	☐ 2	☐ 1

4 At your next clinic visit, would you like to discuss:

	Yes	No	Not sure
a. Alternative types of drug for controlling IBD	☐	☐	☐
b. Ways to adjust your own treatment	☐	☐	☐
c. Side effects or difficulties with using your medicines	☐	☐	☐
d. New symptoms that have developed since your last visit	☐	☐	☐

These questions are additional flags to highlight medication-related concerns or the development of new symptoms

5 How would you rate the OVERALL control of your IBD in the past *two weeks*?

Please draw a vertical line (|) on the scale below

Worst possible control		Best possible

The IBD-Control-VAS is a stand alone summary score

Appendix 2 Publication arising from this work

OPEN ACCESS

ORIGINAL ARTICLE

Development and validation of a rapid, generic measure of disease control from the patient's perspective: the IBD-Control questionnaire

Keith Bodger,[1,2] Clare Ormerod,[1,2] Daniela Shackcloth,[2] Melanie Harrison,[2] on behalf of the IBD Control Collaborative

Open Access
Scan to access more
free content

CrossMark

ABSTRACT

Introduction The use of patient reported outcome measures to support routine inflammatory bowel disease (IBD) care is not widespread and suggests that existing questionnaires lack relevance to day-to-day decisions or are too cumbersome to administer. We developed a simple, generic tool for capturing disease control from the patient's perspective to address these barriers.
Methods Development based on literature review, patient focus groups/interviews and a steering group, defining a limited set of generic questions. The 'IBD-Control' questionnaire comprises 13 items plus a visual analogue scale (VAS) (0–100). Prospective validation involved baseline completion of IBD-Control, quality of life (QoL) questionnaire (UK-IBD-Q), EuroQol (EQ-5D), Hospital Anxiety and Depression Score; and clinician assessment (blinded to questionnaire; recording Harvey-Bradshaw Index or Simple Clinical Colitis Activity Index; Global Clinician Rating; treatment outcome).
Results 299 patients returned baseline surveys (Crohn's disease, n=160; ulcerative colitis, n=139) and 138 attended for repeat visits. *Completion time* (mean; SD): 1 min 15 s; 25 s; *Internal consistency*: Cronbach's α for all 13 items (0.85); for subgroup of eight questions ('IBD-Control-8'; 0.86). Strong correlation between IBD-Control-8 and IBD-Control-VAS (r=0.81). *Test-retest reliability* (2 week repeat): intra-class correlation=0.97 for IBD-Control-8 and 0.96 for IBD-Control-VAS. *Construct validity*: Moderate-to-strong correlations between IBD-Control-8 and IBD-Control-VAS versus activity indices, UK-IBD-Q and EQ-5D (utility) with r values 0.52–0.86. *Discriminant validity* (mean instrument scores for remission, mild, moderate or severe): p<0.001 (analysis of variance (ANOVA)). *Sensitivity to change*: Effect sizes: 0.76–1.44.
Conclusions The IBD-Control is a rapid, reliable, valid and sensitive instrument for measuring overall disease control from the patient's perspective. Unlike existing patient reported outcome measures, its simplicity, ease-of-use and generic applicability make it a candidate for supporting routine care.

INTRODUCTION

The goal of therapy for inflammatory bowel diseases (IBD) is to achieve and maintain disease control and thereby optimise quality of life (QoL). Hence, assessment of disease control is a core component of care and underpins management decisions. Surprisingly,

Significance of this study

What is already known about this subject?
► The routine use of patient-reported outcome measures (PROMs) in clinical practice is increasingly advocated as a means of supporting patient-centred care, informing decisions and driving service quality.
► A range of PROMs has been developed to measure quality of life in inflammatory bowel disease but none has found widespread practical use in routine settings.
► Barriers to adoption of PROMs include patient and administrator burden and lack of tangible relevance to day-to-day decision-making or quality assurance of services.

What are the new findings?
► The IBD-Control questionnaire is the first PROM to capture disease control from the patient's perspective using a simple set of generic items applicable to all patients with IBD.
► Summary scores from the instrument show strong validity versus more complex quality of life questionnaires, generic utility measures, disease activity scores and global physician assessment.
► As a patient-completed screening tool, the IBD-Control provides a rapid and reliable means of identifying those in a quiescent state and the individual items and summary scores have potential to support efficient models of disease follow-up.

How might it impact on clinical practice in the foreseeable future?
► In routine practice, IBD-Control has potential to provide patients and healthcare teams with a practical means of capturing simple but reliable outcome data in a standardised and easily interpretable form.
► The strong measurement properties suggest potential for use in clinical studies, particularly in long-term or large-scale trials, surveys or registries where its minimal user and administrator burden will favour serial capture of PROM data in routine settings.

185

no survey instrument has been developed for routine clinical practice with the specific aim of capturing disease control from the patient's perspective.

The routine use of patient reported outcome measures (PROMs) in healthcare is gaining increasing political and professional support as a means of informing day-to-day decisions and driving service quality.[1][2] PROMs are standardised, validated questionnaires intended for completion by patients in order to measure their perceptions of their own functional status and well-being. PROMs have begun to find a role in national audits and registers and there is rapidly growing interest in their potential to inform individual care.[1]

No single PROM instrument has gained widespread popularity in the UK or internationally for routine use in IBD. When formal measurement of health status is undertaken (eg, for clinical trials) the traditional emphasis has been on clinician-reported indices assigning scores for selected symptoms, vital signs and laboratory or endoscopic parameters.[3–7] However, it has long been recognised that clinician-reported indices are not entirely objective[8] and fail to capture the impact of disease from the patient's perspective.

The last decade has seen a wealth of qualitative research employing focus groups, interviews, expert panels and factor analysis to identify key issues for patients with IBD.[9–17] This has resulted in a number of multidomain PROMs being developed as measures of disease-specific QoL, typically combining symptom questions with more generic questions that relate to overall well-being, energy/vitality, bodily pain and impacts of disease on physical, social and emotional function.[9–17] Only a limited number of these instruments has found use in clinical trials[10][16] but none has established a significant place in routine practice. This may reflect, in part, the patient and administrator burden inherent in using lengthy multidomain questionnaires. Shortened forms may have more potential for routine clinical applications[18][19] but are not used widely.

If PROMs are to find a place in routine care, instruments need to be acceptable to patients and healthcare teams, demonstrating added value to normal practice and with simple and clinically-relevant interpretation. This means combining a high level of user-friendliness (ie, short and rapidly completed tools with low burden and easy interpretation) with strong and explicitly stated measurement properties to support decisions at the level of the individual patient. The objective of this study was to develop and validate a novel questionnaire intended to rapidly capture disease control from the patient's perspective ('IBD-Control'). Our design criteria mandated that the tool should be short, simple and generic in content in order to maximise its potential for routine use across the full spectrum of patients with IBD. As there is no gold standard measure of disease control in IBD, we used a wide array of external measures to validate the questionnaire, including disease activity indices, generic and disease-specific QoL questionnaires and pragmatic clinical end points (global physician assessment and treatment outcome).

METHODS
Development phase
Design criteria: Based on the intended practical clinical application for the IBD-Control questionnaire, we formulated a number of guiding principles for its development. These design criteria were used as a filter to select items generated from literature review and the qualitative patient study. Essential attributes for the choice of individual question items were specified as follows: (A) Include items that capture a patient's *global* self-assessment of overall disease control; (B) Items must reflect *generic* areas of concern, function or impact that are important to patients with both main forms of IBD; (C) Items should *not* focus on individual gastrointestinal symptoms since their relevance and relative importance will vary depending on disease type and extent.

Criteria for the overall design of the instrument were: (A) Minimal user burden (ie, a rapid completion time with a limited total number of question items and simple response categories); (B) The tool should generate overall summary scores with clearly defined interpretation at the individual patient level; and (C) The development and validation process should provide evidence for its acceptability, reliability, validity and responsiveness based on modern psychometric standards.[20–22] The project was guided by a steering group of medical and nursing professionals with expertise in the management of IBD.

Literature review of existing PROMS used in IBD: A detailed review of the English-language literature was undertaken using the PubMed electronic database supplemented with manual searching of bibliographies and conference proceedings within relevant journals. We used combinations of the search terms 'inflammatory bowel disease', 'ulcerative colitis', 'Crohn's disease', 'questionnaire', 'validation', 'quality of life', 'patient-reported outcome' and 'control'. We excluded clinician-reported measures (ie, disease activity indices) and those developed specifically for paediatric patients. Self-administered PROMS reported for adult patients with patients with IBD where categorised as either generic (instruments developed originally for general use across different diseases) or disease-specific (tools developed specifically for patients with IBD). From the literature review we generated a list of core domains (figure 1) to serve as a framework for analysing outputs of the qualitative work with patients.

Qualitative study: We undertook two focus groups (six patients), one for each condition (ulcerative colitis (UC) or Crohn's disease (CD)) plus 13 one-to-one interviews with informed consent. A semistructured approach was taken to encourage patients to talk about the general concept of 'control' of IBD and the things they associated with 'good control' or 'poor control'. Verbatim contributions were recorded and summarised as field notes by the lead interviewer (CO). These items or themes were grouped and mapped onto a framework based on the review of existing IBD questionnaires. At the end of each session, the patients were shown a selection of questionnaires and asked to contribute ideas or preferences about question format, layout, response categories, item number and completion time. 'Saturation' was reached when patient interviews ceased to add new items or themes to the existing framework.

Selection of items: First, we selected a set of generic questions to represent each of the core domains (figure 1), identifying where possible individual question items that have broad, general coverage rather than specific wording to conform to our design criteria. For example, the item selected for social impact would ask about missing 'planned activities' (rather than a more specific question about missing 'work', 'study' or 'a social event'). To qualify for potential inclusion in the IBD-Control, each item needed to be identified as relevant to the construct of disease control in the patient contributions and validated in piloting.

Piloting and modification: To test for acceptability, lack of ambiguity and content validity we pretested the draft instrument in a pilot group of 30 patients who were asked to complete the IBD-Control and provide verbal or written feedback or annotations regarding content and format. This resulted in changes to item wording or layout.

Prospective validation phase (psychometric testing)
The prospective validation study recruited patients at the investigating centre, a university hospital serving a population of

Figure 1 Flow chart summarising development and validation of IBD-Control. IBD, inflammatory bowel disease.
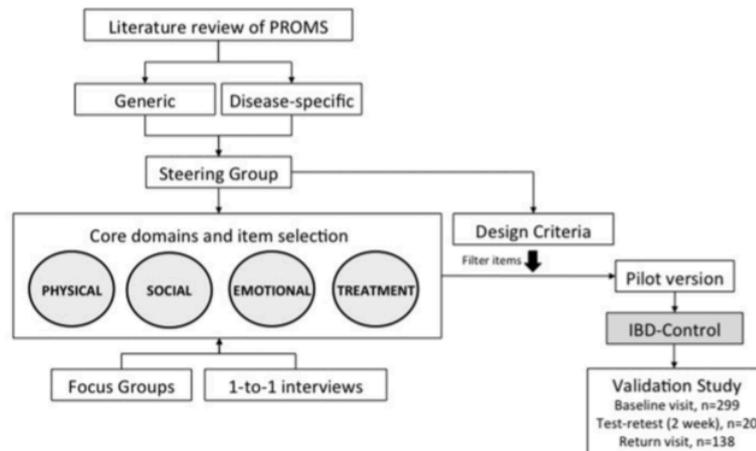
approximately 330 000 people with an established secondary care IBD service providing outpatient and inpatient medical and surgical care.[23] Patients were recruited during routine visits to the outpatient clinics prior to consultation with the doctor or specialist nurse, during other treatment-related visits (eg, azathioprine monitoring clinic or infusion visit for biologics) or at the time of admission for inpatient care. Inclusion criteria specified a confirmed diagnosis of IBD on the basis of clinical, endoscopic, radiological and/or histological criteria with disease duration of at least 6 months. Exclusion criteria specified non-English speaking subjects, cognitive impairment or serious active psychiatric disease.

After informed consent, the patients completed the IBD-Control questionnaire and were then asked to complete a questionnaire pack comprising a disease-specific QoL questionnaire (the UK-IBD-QoL),[16 17] a generic health status instrument (EuroQol, EQ-5D-3L),[24] and the Hospital Anxiety and Depression Scale.[25] The research team undertook a simultaneous assessment of current disease activity, using the Harvey-Bradshaw Index (HBI) for CD [5] and the Simple Clinical Colitis Activity Index (SCCAI) for UC.[6] The research team and clinicians were blinded to the results of patient-completed questionnaires.

Where the study visit was taking place at a scheduled clinical review (eg, outpatient attendance), the treating clinician or specialist nurse was asked to complete a questionnaire at the end of the consultation to indicate the current state of IBD (Global Physician Assessment) using a categorical scale (remission, mild, moderate or severe), blinded to patient surveys. All treatment decisions were recorded, capturing whether new therapies were started, existing drug doses changed, therapies discontinued or surgery recommended. The research team reviewed the hospital case records and clinical information systems to extract background clinical information regarding diagnosis, duration of disease, previous hospitalisation and surgery, disease extent, presence of stoma or perianal fistulae, major comorbid illness and current therapy for their IBD.

Assessment of psychometric (measurement) properties of IBD-Control

Internal consistency was assessed for categorical questions and any subscales derived from them using Cronbach's α.[26] In addition, we used Spearman's correlation to study the association between individual categorical questions and the rating from the instrument's visual analogue scale (IBD-Control-VAS).

Reproducibility was assessed under two sets of conditions using serial observations and the Intraclass Correlation Coefficient. First, a group of 20 patients underwent a repeat assessment at 2 weeks from baseline, completing the IBD-Control and UK-IBD-QoL on both occasions (test-retest group). In this group, we defined a stable health state based on the patient's response to the IBD-Control transition question item (Q2, response='No change') plus ≤10-point change on the UK-IBD-Q total score. Second, we compared IBD-Control scores obtained from a group of patients who completed assessments at baseline and then at a subsequent routinely scheduled visit (routine care group). In this group, we defined a stable state on the basis of there being no change in the physician global assessment between visits.

Construct validity was assessed by comparing the scores derived from the questionnaire with several independent external measures of patient health state. These measures included generic QoL (EQ-5D utility index and EQ-5D visual analogue score), disease specific QoL (UK-IBD-QoL), physician global assessment and clinical disease activity indices. Bivariate correlations were expressed as Spearman's correlation coefficients and between-group mean scores were compared by t test or analysis of variance (ANOVA). Multiple variable linear regression was used to explore the relative strength and independence of individual question items as predictors of the various external measures of health state and to select items to include in subscores.

Responsiveness was assessed in those patients who completed baseline and follow-up assessments and experienced a change in their health state. There is no true external gold standard to define a change in disease control and so we performed responsiveness analyses using two alternative external criteria to define a change in health state between visits. First, we defined improvement or deterioration on the basis of self-reported UK-IBD-QoL total score, with a change of >10 points taken as a cut-off. As an alternative, we used the physician global rating to identify a change in clinical state (any change from baseline category, eg, mild to moderate). A range of responsiveness statistics were calculated including correlations in change scores,

effect size, standardised response mean (SRM) and modified SRM (MSRM).

Definition and evaluation of cut-off values for detecting the 'quiescent' state: A key potential application for IBD-Control was to serve as a screening test to rapidly and reliably identify patients with optimal self-reported disease control. Such patients are suitable for alternative forms of chronic disease follow-up (eg, self-care, telephone or virtual clinics) rather than traditional, costly and inconvenient office or clinic attendance. We defined a state of 'quiescent' disease using a combination of parameters, as follows:

1. Disease-specific measures consistent with clinical remission as defined by disease activity indices within reported remission ranges (SCCAI<4; HBI<5), UK-IBD-QoL total score ≥90 points and Physician Global Assessment rated as 'remission'. To allow for missing or incomplete data, patients were required to have sufficient data to satisfy any two of the criteria.
2. Not currently receiving oral corticosteroid therapy, a course of antibiotics, therapeutic (polymeric) diet, induction with biologics or awaiting surgical treatment.
3. IBD-Control transition question (Q2) indicates condition was unchanged or improved in the last 2 weeks (ie, excluding patients who responded that bowel symptoms were 'worse').
4. No escalation of therapy occurred as a consequence of the clinical consultation.

Individuals satisfying all four of the above criteria would be regarded as well-controlled. Patients who did not meet these criteria where designated as 'not quiescent'. We used receiver-operating characteristics analysis to assess the performance of the IBD-Control as a tool for identifying this subgroup of quiescent patients. Given the intended clinical application, we defined optimum cut-off scores with an emphasis on achieving high specificity (≥85%) to minimise the risk of false positives (identifying a non-quiescent patient as quiescent). The cut-off value with the highest sensitivity×specificity product was selected, as described.[27]

RESULTS
Development phase
Literature review: A PubMed search identified 7244 references with a combination of search terms: ['Patient reported outcome measure' OR 'PROM' OR 'quality of life' OR 'questionnaire' OR 'outcome measure'] AND ['Inflammatory bowel disease' OR 'ulcerative colitis' OR 'Crohn's disease' OR 'colitis' OR 'IBD']. Of these, 682 articles included the additional terms: ['development' or 'validation' OR 'psychometric']. We identified 71 relevant articles relating either to IBD-specific questionnaires (23 instruments for adults including variants or short forms; 4 for paediatrics were excluded) or to the validation of existing generic PROMS in patients with IBD (3 instruments). Of the disease-specific PROMS for adults with IBD, only the original McMaster IBD-Q (32 items) [10] and a similar instrument developed in the UK [17] were found to have undergone detailed psychometric evaluation in accordance with modern guidelines [20] [21] and subsequent use as outcome measures in clinical trials. Hence, we chose the UK IBD-Q as our measure of QoL in the validation study. Of the generic instruments, the EuroQoL EQ-5D [24] had the most publications and so was our choice of utility measure. No instrument was identified that measured IBD disease control from the patient perspective but we reviewed questionnaires for other chronic diseases.

Focus groups and one-to-one interviews: Saturation was reached after completing two focus groups and 13 one-to-one interviews. Consistently, the idea of disease control was expressed by contributors as their overall sense that disease-related symptoms were minimised or absent (ie, a global sense of control). As expected, the control of a wide variety of specific gastrointestinal symptoms was identified as important by individual patients but these did not meet our criteria for generic item selection. However, a common theme across all participants was the idea that good control implied that their *particular* array of bowel symptoms was minimised and not impacting on day-to-day physical functioning (overall energy and vitality; and sleep), social functioning (daily activities, work and recreational) or emotional functioning (feelings of anxiety or depression). In addition, the control of symptoms of discomfort or pain was a distinct factor. Loss of disease control was expressed as either a worsening of overall bowel symptoms or the development of an unfamiliar new symptom.

Control of IBD was linked also with ideas about the perceived effectiveness or acceptability of current treatments. This is not a core theme that appears in the established model for QoL in IBD.[9–19] Good control would imply feelings that treatment was 'right' or 'was working' whereas poor control would invoke the sense that treatment was ineffective. For some patients, the ability to control the condition themselves (ie, by making adjustments to medication according to symptoms) was seen as part of the concept of disease control. Additional treatment-related themes included perceived difficulties with using certain drugs or side effect. Good control would imply a lack of such treatment concerns.

Description and specification of the IBD-Control instrument
The IBD-Control instrument is shown in online supplementary appendix 1. The paper version comprises a single-sided A4 document that includes five sections numbered 1 through 5. The first four sections contain a series of 13 categorical questions, each of which has three response options. Of the categorical questions, 12 items have response options of 'Yes', 'No' or 'Not sure' and the remaining question is a transition question for overall bowel symptoms with options of 'Better', 'No change' or 'Worse'. The fifth section contains the horizontal VAS, anchored between zero (worst possible control) and 100 (best possible control). The time horizon for disease control assessment is the 'past 2 weeks'.

Scoring: Each response to the 13 individual items is scored as follows: zero points for least favourable reply; one point for intermediate or indeterminate reply; two points for most favourable reply. The IBD-Control-8 subscore is calculated by summing scores for Q1a, Q1b, Q3a, Q3b, Q3c, Q3d, Q3e and Q3f resulting in a range of 0–16 (0=worst control). The rationale for selecting this subgroup of eight questions is explained later in the results section. The IBD-Control-VAS scores are in the range 0–100 (0=worst control).

Validation phase
Patient sample
The characteristics of the patient sample at baseline are summarised in table 1. There were 299 patients who completed the baseline assessment, of whom 160 (53.5%) had a diagnosis of CD and 139 (46.5%) were suffering from UC. The spectrum of disease severity was very similar between the two forms of IBD (CD vs UC), with no significant difference in mean scores for generic QoL (EQ-5D utility or VAS), disease-specific QoL (UK-IBD-QoL total score) or proportion of patients in remission

Table 1 Characteristics of 299 patients with inflammatory bowel disease (IBD) at baseline assessment

Characteristic	All patients	Crohn's disease	Ulcerative colitis
Number of cases	299	160	139
Age, years	43 [16]	41 [15]	46 [16]
Gender			
Male	130 (43.5%)	68 (57.5%)	77 (55.4%)
Female	169 (56.5%)	92 (42.5%)	62 (44.6%)
Disease duration, years	10 [10]	10 [10]	9 [10]
Previous surgery			
Yes	83 (27.8%)	76 (47.5%)	7 (5.0%)
No	216 (72.2%)	84 (52.5%)	132 (95.0%)
Perianal disease (fistula)			
Yes	27 (9.0%)	27 (16.9%)	0 (0.0%)
No	272 (91.0%)	133 (83.1%)	139 (100.0%)
Stoma present			
Yes	10 (3.3%)	8 (5.0%)	2 (1.4%)
No	289 (96.7%)	152 (95.0%)	137 (98.6%)
Medication			
Topical 5-ASA	36 (12.0%)	2 (1.3%)	34 (24.5%)
Topical steroid	9 (3.0%)	2 (1.3%)	7 (5.0%)
Oral 5-ASA	177 (59.2%)	68 (42.5%)	109 (78.4%)
Oral corticosteroid	26 (8.9%)	10 (6.3%)	16 (11.5%)
Standard immunosuppressants	104 (34.8%)	71 (44.4%)	33 (23.7%)
Biological agent	38 (12.7%)	30 (18.8%)	8 (5.8%)
Dietary therapy (polymeric diet)	8 (2.7%)	8 (5.0%)	0 (0.0%)
Disease activity indices			
Harvey-Bradshaw Index	n/a	5 [5]	n/a
Simple Clinical Colitis Activity index	n/a	n/a	4 [3]
Quality of life questionnaires			
EQ-5D utility score	0.68 [0.30]	0.65 [0.30]	0.70 [0.29]
EQ-5D visual analogue scale	65 [23]	65 [22]	65 [24]
UK-IBD-Q	86 [20]	85 [18]	88 [21]
Physician global assessment			
Remission	161 (61.9%)	80 (60.6%)	81 (63.3%)
Mild	58 (22.3%)	37 (28.0%)	21 (16.4%)
Moderate	30 (11.5%)	11 (8.3%)	19 (14.8%)
Severe	11 (4.2%)	4 (3.0%)	7 (5.5%)

Continuous variables expressed as mean [SD] and categorical variables as number (%) where appropriate.

Table 2 Characteristics of 138 patients with inflammatory bowel disease (IBD) at second visit (returning patients)

Characteristic	All patients	Crohn's disease	Ulcerative colitis
Number of cases	138	82	56
Age, years	41 [15]	38 [13]	45 [16]
Gender			
Male	65 (47.1%)	36 (43.9%)	29 (51.8%)
Female	73 (52.9%)	46 (56.1%)	27 (48.2%)
Disease duration, years	8 [9]	9 [8]	8 [9]
Previous surgery			
Yes	45 (32.6%)	43 (52.4%)	2 (3.6%)
No	93 (67.4%)	39 (47.6%)	54 (96.4%)
Perianal disease (fistula)			
Yes	20 (14.5%)	20 (24.4%)	0 (0.0%)
No	118 (85.5%)	62 (75.6%)	56 (100.0%)
Stoma present			
Yes	5 (3.6%)	4 (4.9%)	1 (1.8%)
No	133 (96.4%)	78 (95.1%)	55 (98.2%)
Medication			
Topical 5-ASA	11 (8.0%)	0 (0.0%)	11 (19.6%)
Topical steroid	5 (3.6%)	0 (0.0%)	5 (8.9%)
Oral 5-ASA	82 (59.4%)	34 (41.5%)	48 (85.7%)
Oral corticosteroid	16 (11.6%)	6 (7.3%)	10 (17.9%)
Standard immunosuppressants	61 (44.2%)	44 (53.7%)	17 (30.4%)
Biological agent	29 (21.0%)	24 (29.3%)	5 (8.9%)
Dietary therapy (polymeric diet)	5 (3.6%)	5 (6.1%)	0 (0.0%)
Disease activity indices			
Harvey-Bradshaw index	n/a	5 [5]	n/a
Simple Clinical Colitis Activity index	n/a	n/a	5 [4]
Quality of life questionnaires			
EQ-5D utility score	0.61 [0.34]	0.60 [0.35]	0.62 [0.33]
EQ-5D visual analogue scale	61 [23]	63 [22]	56 [25]
UK-IBD-Q	83 [19]	84 [18]	82 [22]
Physician global assessment			
Remission	49 (44.1%)	27 (44.3%)	22 (44.0%)
Mild	38 (34.2%)	26 (42.6%)	12 (24.0%)
Moderate	18 (16.2%)	6 (9.8%)	12 (24.0%)
Severe	6 (5.4%)	2 (3.3%)	4 (8.0%)

Continuous variables expressed as mean [SD] and categorical variables as number (%) where appropriate.

(physician global assessment). Of the baseline cohort, 138 patients were recruited to undertake a repeat assessment at a scheduled follow-up visit (returning patients), summarised in table 2. Patients were recruited between March 2011 and June 2012.

Acceptability and feasibility

Completion time was measured in a sample of 10 patients (aged between 25 years and 64 years), with a mean completion time for the IBD-Control of 1 min and 15 s (SD: 25 s). The range was between 42 s and 2 min 1 s. Completion rates for the 13 individual questions ranged between 93.3% (Q2) and 99% (Q3a and Q3d). Of the 299 patients who completed the IBD-Control at baseline visit, 259 (86.6%) provided responses to all 13 question items and 272 (91%) completed the VAS

question. We prespecified that any question returning the same response from 80% or more of the patients would be considered for exclusion from the questionnaire. Such questions are unlikely to be sensitive to different levels of severity. However, none of the question items yielded the same response from 80% or more of the patients.

Internal consistency

Scored responses to each of the 13 individual questions showed significant positive correlations with the IBD-Control-VAS score, with Spearman's ρ values ranging from 0.24 to 0.70 (table 3). These findings suggest that the individual items within the questionnaire are measuring aspects of the same construct ('disease control') and that there is internal consistency within the questionnaire. A linear regression model retaining eight of the

Table 3 Correlation between individual question items within the IBD-Control instrument and visual analogue scale for disease control (IBD-Control-VAS) and disease-specific quality of life total score measured using the UK-IBD-QoL questionnaire

Item	Question	IBD-Control-VAS	UK-IBD-QoL
Q1a	… your IBD has been well controlled in the past 2 weeks*	0.70	0.63
Q1b	… your current treatment is useful in controlling your IBD*	0.39	0.38
Q2	… have your bowel symptoms been getting better, getting worse or not changed	0.36	0.24
Q3a	… miss any planned activities because of IBD*	0.63	0.67
Q3b	… wake up at night because of symptoms of IBD*	0.61	0.67
Q3c	… suffer from significant pain or discomfort*	0.64	0.66
Q3d	… often feel lacking in energy (fatigued)*	0.51	0.63
Q3e	… feel anxious or depressed because of your IBD*	0.58	0.70
Q3f	… think you need a change to your treatment*	0.60	0.50
Q4a	… would you like to discuss alternative types of drug	0.42	0.26
Q4b	… would you like to discuss ways to adjust your own treatment	0.33	0.31
Q4c	… would you like to discuss side effects or difficulties with your medicines	0.24	0.29
Q4d	… would you like to discuss new symptoms that have developed	0.30	0.29

Item responses were scored as 0 (worst option), 1 (intermediate, eg, 'Not sure') or 2 (best option), so a higher score should indicate better disease control. Correlation coefficients are expressed as Spearman's ρ values (p<0.01 in all cases).
*Question items included in the IBD-Control-8 subscore.
IBD, inflammatory bowel disease; QoL, quality of life.

question items (IBD-Control-8) accounted for 70% of the variance in the IBD-Control-VAS score (R-Square 0.704). Bivariate correlation coefficients for each of the eight items within the IBD-Control-8 subscore ranged from 0.39 to 0.70.

We assessed degree of agreement between the 13 categorical question items by Cronbach's α statistic.[26] The higher the α value, the higher the internal consistency between the individual questions. A score above 0.70 is regarded as desirable but values in excess of 0.90 are suggestive of excessive overlap or redundancy between individual items. In the present study, Cronbach's α was 0.85 overall for all 13 question items and 0.86 for the subscore of eight questions (IBD-Control-8). This provides further evidence for internal consistency for the 'disease control' construct.

The IBD-Control-8 subscore and the IBD-Control-VAS score are intended to represent a summary measure of the same construct (ie, an overall rating of disease control from the patient's perspective). Hence, the two summary scores should show a strong positive correlation with each other. This was confirmed for patients with IBD overall (r=0.811; p<0.001), and for CD (r=0.78; p<0.001) and UC (r=0.83; p<0.001) separately.

Reliability

This was tested under two sets of conditions. First, we evaluated reliability in a test-retest group of patients who completed evaluations 2 weeks apart. Of the 20 subjects, 13 reported no change in their disease control over the past 2 weeks (IBD-Control transition question, Q2) and had stable UK-IBD-QoL scores (±10-points from baseline). There was no significant difference in mean IBD-Control-8 or IBD-Control-VAS scores between visits and the intraclass correlation coefficient was 0.97 for the IBD-Control-8 subscore and 0.96 for IBD-Control-VAS (table 4A; p<0.01).

Second, reliability was assessed in 32 patients who completed questionnaires at consecutive clinical visits and were categorised as having no change in their physician global assessment between the two visits. The mean interval between the two visits was 131 days. Again, there was no significant difference in mean summary scores on the IBD-Control questionnaire between visits. For these patients, intraclass correlation

coefficient was 0.87 for the IBD-Control-8 subscore and 0.81 for IBD-Control-VAS (table 4A; p<0.01).

Construct validity

There is no true gold standard for measuring disease control from the patient's perspective. However, it is possible to test a number of hypotheses concerning the logical relationships that should exist between the item scores or summary scores obtained with the IBD-Control instrument and a range of other external criteria, measures or patient characteristics.

Validity of individual question items: Good disease control would be expected to be associated with better QoL. Consistent with this, each individual question item within the IBD-Control demonstrated a significant positive correlation with the total score for the UK-IBD-QoL (table 3). Hence, 'positive' responses to questions within the IBD-Control are correlated with better overall

Table 4 Reproducibility of IBD-Control summary scores for stable patients

Instrument score (scale)	Mean difference (Visit 2—Visit 1)	SD of difference	Intraclass correlation (95% CI)
(A) Patients returning at 2 weeks (test-retest group)*			
IBD-Control-8 subscore (0–16)	−0.45	1.81	0.97 (0.90 to 0.99)
IBD-Control-VAS (0–100)	+2.25	9.79	0.96 (0.88 to 0.99)
(B) Patients returning for a scheduled clinic visit (routine care group)†			
IBD-Control-8 subscore (0–16)	0.00	3.72	0.87 (0.71 to 0.94)
IBD-Control-VAS (0–100)	−0.56	23.00	0.81 (0.58 to 0.91)

*Data for 13 patients responding 'not changed' to the IBD-Control transition question ('Over the past 2 weeks have your bowel symptoms been getting worse, getting better or not changed') and with stable UK-IBD-Qol total scores (<10-point change between two visits).
†Data for 32 patients with a stable (unchanged) physician global assessment rating between two visits (mean interval between visits 131 days).
Two-way fixed effects model for consistency, p<0.001.
IBD, inflammatory bowel disease; QoL, quality of life; VAS, visual analogue scale.

disease-specific QoL. The correlation coefficients for each of the IBD-Control-8 subscore items and total UK-IBD-QoL score ranged from 0.38 to 0.70. These findings suggest the individual items within the instrument are valid.

We used a single question item to capture feelings of anxiety or depression ('… felt anxious or depressed because of your IBD'). Responses to this question (No, Not Sure, Yes) showed moderate-to-strong correlation with the summary scores for the multi-item Hospital Anxiety and Depression Scale questionnaire [25] for anxiety and depression subscores (r=0.57 and r=0.62, respectively, p<0.001).

Validity of IBD-Control summary scores: The summary scores generated by the disease control measure are predicted to correlate positively with disease-specific (UK-IBD-QoL) and generic (EQ-5D utility and VAS) QoL scores. Consistent with this, we found moderate-to-strong positive correlations between the IBD-Control summary scores (IBD-Control-8 and IBD-Control-VAS) and each of the QoL scales (table 5 and figure 2). A linear regression model containing the eight items from the IBD-Control-8 subscore accounted for 75% of the variance in UK-IBD-QoL total score (R Square 0.747).

Conversely, negative correlations would be expected to exist between IBD-Control summary scores and traditional clinical measures of disease activity or severity. In the present study, we obtained clinician-reported outcomes in the form of traditional symptom-based activity indices (HBI or SCCAI) and also as a global physician rating of current disease activity. We found highly significant moderate-to-strong correlations between the IBD-Control scores and the disease severity measures in the expected direction (ie, negative coefficients) (table 5).

These associations provide strong evidence to support the construct validity of the IBD-Control instrument. In addition, a valid measure of disease control would be expected to differentiate between mutually exclusive groups of patients categorised using an accepted external clinical criterion. Hence, mean scores for disease control would be expected to be significantly different between patients 'in remission' and those with 'active disease'. Comparison of mean IBD-Control-8 and IBD-Control-VAS scores across the four physician global assessment categories (remission, mild, moderate or severe) confirmed good discriminant validity across the spectrum of disease (figure 3).

Multiple linear regression models examined whether additional patient factors (age, gender, disease type, previous surgery) were independently associated with IBD-Control scores

after taking account of current disease severity. All models confirmed the strong association between IBD-Control summary scores and disease activity indices, generic and disease-specific QoL and global assessment. However, no consistent independent associations were observed for the other predictor variables.

Mapping of IBD-Control-8 to EQ-5D utilities: Ease of use makes the IBD-Control a candidate outcome measure for large scale observational clinical studies, trials, surveys or registries. Hence, the ability to convert IBD-Control summary scores to an estimate of utility could support health economic evaluations. Figure 4 illustrates the association between IBD-Control-8 subscore and the EQ-5D utility index.

Table 5 Construct validity: Correlations between the IBD-Control summary scores and external measures of health status in patients with inflammatory bowel disease (IBD) at baseline assessment

External measure	IBD-Control-8			IBD-Control-VAS		
	All	CD	UC	All	CD	UC
UK-IBD-QoL score	0.86	0.81	0.90	0.76	0.65	0.85
EQ-5D utility score	0.70	0.68	0.67	0.52	0.50	0.58
EQ-5D VAS	0.68	0.65	0.71	0.70	0.69	0.71
Physician global assessment	−0.58	−0.45	−0.67	−0.58	−0.47	−0.65
Harvey-Bradshaw Index	n/a	−0.68	n/a	n/a	−0.60	n/a
Simple Clinical Colitis Activity Index	n/a	n/a	−0.72	n/a	n/a	−0.75

Correlation expressed as Spearman's ρ values (p<0.01 in all cases).
CD, Crohn's disease; QoL, quality of life; UC, ulcerative colitis; VAS, visual analogue scale.

Figure 2 Validity of IBD-Control summary scores in relation to UK-IBD-QoL Questionnaire. IBD, inflammatory bowel disease; QoL, quality of life.

Figure 3 Validity of IBD-Control summary scores in relation to physician global assessment.

Responsiveness

A vital measurement property for any PROM is that the instrument is sensitive to gains or losses in the measurement concept of interest. We evaluated the responsiveness of the instrument by comparing serial data collected in the patients who returned for a second visit.

Change scores: A crude method for assessing responsiveness is to calculate 'change scores' for the instrument (ie, subtract score at visit 1 from the score at visit 2) and to examine the correlation between this value and a change score derived from another relevant clinical measure. For change scores in IBD-Control-8 subscore and IBD-Control-VAS, there were significant correlations with change scores for each of the key external outcome measures (r values ranging from 0.25 to 0.78). The strongest correlation for change in IBD-Control scores were versus the change in UK-IBD-QoL total score (IBD-Control-8, r=0.72; IBD-Control-VAS, r=0.55; p<0.001) and the weakest for HBI (IBD-Control-8, r=−0.25; IBD-Control-VAS, r=−0.32; p<0.001).

Responsiveness statistics: Several methods have been proposed to provide more robust quantitative expressions of the magnitude and meaning of changes in instrument scores. *Effect size* (ES) is calculated as the size of the change in scores (ie, the difference between mean scores at the baseline and repeat assessments) divided by the SD of baseline scores. Benchmarks for assessing the relative size of change have been proposed whereby an ES of 0.2 is considered small, 0.5 as medium and ≥0.8 as large. The IBD-Control-8 and IBD-Control-VAS showed moderate-to-strong ES with values ≥0.76 (table 6) indicating good responsiveness.

The *SRM* differs from ES in that the denominator is the SD of change scores (rather than of baseline scores). This takes account of variability in change per se rather than variability in absolute baseline scores. Responsiveness of the IBD-Control summary scores measured by SRM was strong (table 6).

Finally, in the *modified SRM* (MSRM) the numerator remains the mean change in absolute scores for the group but for this statistic the denominator is the SD of change scores taken specifically from individuals who are identified by other means as clinically stable. In the present study, this was taken from a subgroup of patients who showed no change in physician global rating between assessments and had a stable disease-specific

Items not included in the IBD-Control-8 subscore: Five questions were excluded from the IBD-Control-8 summary subscore. The transition question (Q2) was excluded since this asks about *change* in status over 2 weeks rather than measuring current health status. The remaining questions (grouped together as items Q4a to 4d in the final questionnaire) had significant but weaker correlations with IBD-Control-VAS and/or UK-IBD-QoL and did not add significantly to performance of linear regression models predicting these measures. However, the development of new symptoms (Q4d) was regarded as an important 'stand-alone' item to retain in the IBD-Control instrument since a new symptom should serve as an automatic trigger for traditional clinical evaluation.

The three remaining questions (Q4a to 4c) are related to treatment concerns. These items were found to be predictive of the clinical outcome of the consultation in terms of likelihood of treatment escalation (note that clinicians were blinded to questionnaire responses). Of 67 patients reporting no treatment concerns on the three items, only 5 (7.5%) had their treatment escalated, whereas 59 of 192 (30.7%) patients with indeterminate or positive responses to at least one of the three questions had treatment escalated (p<0.001, χ^2 test; OR: 4.1 (95% CI 1.7 to 9.8)).

Figure 4 Validity of IBD-Control-8 subscore in relation to EuroQol (EQ-5D) utility index. IBD, inflammatory bowel disease; QoL, quality of life.

Table 6 Responsiveness statistics for the IBD-Control summary scores in subjects who improved or deteriorated. Data for stable patients (no change) are included for reference

Criteria for change in state	n	Mean difference (SD)	Effect size*	Standardised response mean†	Modified standardised response mean‡
(A) IBD-Control-8 sub-score					
UK-IBD-QoL					
Improved by >10 points	19	+5.50 (4.69)	1.44	1.17	2.72
No change (±10 points)	63	+0.13 (2.02)	0.02	0.06	0.06
Deteriorated by >10 points	26	−4.92 (3.88)	0.99	1.27	2.44
Physician global assessment					
Improved by ≥1 point	27	+4.52 (5.56)	1.08	0.81	1.22
No change	32	0.00 (3.72)	0.00	0.00	0.00
Deteriorated by ≥1 point	27	−3.32 (4.81)	0.76	0.69	0.89
(B) IBD-Control-VAS					
UK-IBD-QoL					
Improved by >10 points	19	+28.21 (31.33)	1.09	0.90	1.84
No change (±10 points)	63	+0.84 (15.34)	0.03	0.05	0.05
Deteriorated by >10 points	26	−26.85 (28.20)	0.90	0.95	1.75
Physician global assessment					
Improved by ≥1 point	27	+29.55 (31.78)	1.14	0.93	1.28
No change	32	−0.56 (23.00)	0.02	0.02	0.02
Deteriorated by ≥1 point	27	−19.71 (30.53)	0.76	0.65	0.86

*Effect size (ES)=Mean difference divided by the SD of the baseline absolute scores for the group.
†Standardised response mean (SRM)=Mean difference divided by SD of the difference for the group.
‡Modified SRM (MSRM)=Mean difference for the group divided by SD of the difference for *unchanged* patients.
IBD, inflammatory bowel disease; QoL, quality of life.

QoL score (UK-IBD-Q total score changed by no more than 10 points between visits). The IBD-Control-8 and the IBD-Control-VAS scores have MSRM values ≥0.86 (table 6).

Sensitivity and specificity of IBD-control in identifying 'quiescent' patients

Of 217 patients at baseline who had formal clinical assessment and all relevant data, 80 individuals could be categorised as quiescent at the time of their routinely scheduled clinic visit. None required treatment escalation (including induction with biologics or referral for surgery). Receiver-operating characteristics analysis for IBD-Control-8 subscore and IBD-Control-VAS indicated strong performance could be achieved with either measure as a test to identify quiescent patients from the remaining cases (figure 5). For IBD-Control-8, a cut-off of ≥13 points identified patients with quiescent IBD with 67.5% sensitivity and 90.6% specificity, whereas for IBD-Control-VAS a cut-off of ≥85 points achieved 64.3% sensitivity and 90% specificity. These cut-off scores would mean that only 1 in 10 patients failing to meet our stringent clinical criteria for quiescent IBD would be falsely categorised as quiescent but none of these 'false positives' required any treatment change at the time of consultation and all were considered to be either in remission or only mildly active by the attending clinician. These data confirm the potential clinical application of IBD-Control as a rapid triage tool to categorise low-risk (quiescent) patients suitable for alternative forms of follow-up.

DISCUSSION

The IBD-Control questionnaire is a novel PROM designed to provide a rapid, user-friendly assessment of disease control from a patient's perspective. It was developed with patient involvement in defining the measurement construct (ie, what is 'disease control'?) and in generating patient-centred items. Key requirements were that IBD-Control should be applicable to routine clinical care and that it would contain items relevant to both

main forms of IBD. This resulted in the generation of a short and deliberately generic set of questions.

As expected, there was considerable consensus and overlap between items or domains identified by patient discussions of
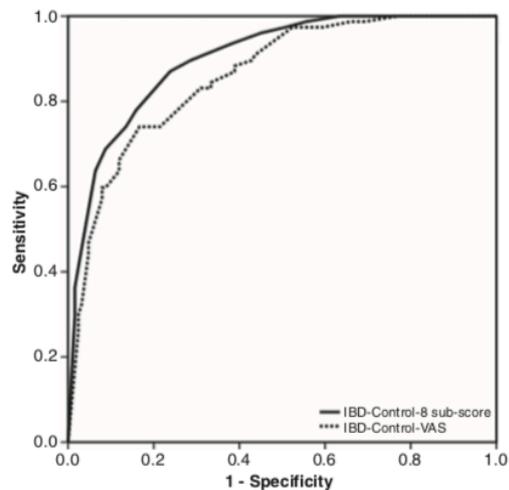

Figure 5 Performance of IBD-Control as a screening test to identify 'quiescent' patients: Receiver operating characteristic (ROC) curves for IBD-Control-8 subscore (solid line; area under curve: 0.90; p<0.001) and IBD-Control-VAS (broken line; area under curve: 0.86; p<0.001). For IBD-Control-8, a cut-off of ≥13 points identified patients with quiescent IBD with 67.5% sensitivity and 90.6% specificity, whereas for IBD-Control-VAS a cut-off of ≥85 points achieved 64.3% sensitivity and 90% specificity. The quiescent state was defined using strict composite criteria (see text). IBD, inflammatory bowel disease.

disease control and the themes reported in the extensive literature on QoL in IBD.[9–19] [28–30] Hence, good control implies the absence of disease impacts on energy/vitality, ability to perform daily activities, sleep, mood and freedom from pain or discomfort. Question items on these topics feature prominently in short forms of existing disease-specific QoL questionnaires and/or in the generic instruments that have been applied to IBD. In addition, our qualitative study identified a discrete domain of 'treatment concerns', encompassing perceptions of effectiveness of treatment, satisfaction with choice of therapy, ability to adjust own treatment and concern about side effects and we incorporated these items into the IBD-Control.

This study shows that IBD-Control is acceptable to patients based on its rapid completion time and high completion rates. Acceptability is a vital characteristic for any PROM if the intended application is to support routine care—a property that may be lacking with the current multi-item, multidomain QoL measures. Despite its simplicity, the IBD-Control exhibits strong measurement properties.

Strengths of the present study include our emphasis on patient involvement in item generation and questionnaire design to complement literature review and clinical opinion. Furthermore, we adopted a comprehensive approach to assessing the full range of psychometric (measurement properties) of the questionnaire in order to conform to current international standards.[20–22] This required the simultaneous collection of a wide array of patient-reported and clinician-reported outcome measures in a large cohort of patients. A key strength of the IBD-Control tool is that it has been developed primarily for clinical application and to support patient-centred care. Hence, our validation has not rested simply on showing correlations with traditional disease measures across the patient population but has considered the performance of summary scores at individual level and for a specific clinical purpose.

We were particularly interested in generating a tool that would reliably identify a patient in a 'quiescent' state, defined stringently using a composite of criteria including physician rating and treatment outcome. This has application to models of care that seek to reduce the need for traditional face-to-face consultation (eg, 'virtual' clinics). The instrument produces scores relevant to the planning of day-to-day care, serving as a potential trigger or triage tool to guide the necessity, timing and nature of more formal clinical assessment. The inclusion of items relating to concern about treatment or new symptoms further distinguishes the IBD-Control from traditional QoL tools. Hence, it is intended to provide information to support patient-focused consultations and summary scores to support traditional clinical assessment and decision-making.

Potential limitations of our approach to PROM development was the adoption of rigid, prespecified design criteria that mandated the IBD-Control should contain only a small number of 'general' items with simple response categories to optimise acceptability. This meant a pragmatic approach was taken to item selection to favour broad, generically worded questions covering core domains with active exclusion of items about individual gastrointestinal symptoms. We did not undertake a process of generating a very large number of items followed by factor analysis and item reduction. We believe our approach was justified by the wealth of existing published information identifying key areas of concern for patients with IBD. We used qualitative data from patients to confirm that each item was relevant to the concept of disease control. Nevertheless, there was a risk that producing an instrument to this strict specification could have resulted in suboptimal measurement properties. However,

the findings of our prospective evaluation suggest quite the opposite—the IBD-Control questionnaire has very strong properties when compared with a wide array of external measures.

There is inevitable selection bias inherent in recruiting patients willing to complete multiple questionnaires for a validation study. However, our patient sample size was large and contained individuals with a broad spectrum of demographic and clinical attributes that are similar to our local denominator IBD population.[23] Although the proportion of cases with moderate-to-severe disease was only 15–20%, the linear trends and consistent confidence limits for IBD-Control scores observed across the disease spectrum suggests sufficient cases in the study. By recruiting patients at the time of 'real-life' clinical visits and asking for IBD-Control completion before other assessments we sought to reproduce the intended application for the tool. Clearly, further validation in other settings and in specific patient subgroups is desirable.

Current guidelines for IBD do not endorse the use of any specific PROM to guide the organisation or monitoring of care nor to support patient-level decision-making. Hence, currently available PROMs serve largely as research tools. This contrasts with some other chronic diseases where the use of specific patient-centred outcome measures has been incorporated into guidelines and care pathways.[1] Recently, quality improvement programmes for IBD in a range of healthcare settings have highlighted the potential for routine capture of PROMs to help drive better patient-centred services.[31] [32] We suggest that the simplicity of the IBD-Control in combination with its strong measurement properties and relevance to both forms of IBD make it ideally suited to support such initiatives, in addition to having a potential role as an outcome measure in pragmatic clinical trials. As an addition to routine clinical care, the instrument offers the realistic possibility of capturing serial PROM data with minimal user burden and its content is eminently suited to adaptation to electronic capture via web-based systems, portals or mobile devices.

194

REFERENCES

1. Dawson J, Doll H, Fitzpatrick R, et al. The routine use of patient reported outcome measures in healthcare settings. *Br Med J* 2010;340:c186.
2. Department of Health. The NHS Outcomes Framework 2012/13. London, 2011.
3. Truelove SC, Witts LJ. Cortisone in ulcerative colitis; final report on a therapeutic trial. *Br Med J* 1955;2:1041–8.
4. Best WR, Becktel JM, Singleton JW, et al. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology* 1976;70:439–44.
5. Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. *Lancet* 1980;1:514.
6. Walmsley RS, Ayres RC, Pounder RE, et al. A simple clinical colitis activity index. *Gut* 1998;43:29–32.
7. Cooney RM, Warren BF, Altman DG, et al. Outcome measurement in clinical trials for Ulcerative Colitis: towards standardisation. *Trials* 2007;8:17.
8. de Dombal FT, Softley A. IOIBD report no 1: Observer variation in calculating indices of severity and activity in Crohn's disease. International Organisation for the Study of Inflammatory Bowel Disease. *Gut* 1987;28:474–81.
9. Irvine EJ. Development and subsequent refinement of the inflammatory bowel disease questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 1999;28:S23–27.
10. 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.
11. Drossman DA, Li Z, Leserman J, et al. Ulcerative colitis and Crohn's disease health status scales for research and clinical practice. *J Clin Gastroenterol* 1992;15:104–12.
12. Farmer RG, Easley KA, Farmer JM. Quality of life assessment by patients with inflammatory bowel disease. *Cleve Clin J Med* 1992;59:35–42.
13. Irvine EJ, Feagan B, Rochon J, et al. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group. *Gastroenterology* 1994;106:287–96.
14. Hjortswang H, Strom M, Almeida RT, et al. Evaluation of the RFIPC, a disease-specific health-related quality of life questionnaire, in Swedish patients with ulcerative colitis. *Scand J Gastroenterol* 1997;32:1235–40.
15. Achleitner U, Coenen M, Colombel JF, et al. Identification of areas of functioning and disability addressed in inflammatory bowel disease-specific patient reported outcome measures. *J Crohn's Colitis* 2012;6:507–17.
16. Williams JG, Cheung WY, Russell IT, et al. Open access follow up for inflammatory bowel disease: pragmatic randomised trial and cost effectiveness study. *Br Med J* 2000;320:544–8.
17. Cheung WY, Garratt AM, Russell IT, et al. The UK IBDQ—A British version of the inflammatory bowel disease questionnaire: development and validation. *J Clin Epidemiol* 2000;53:297–306.
18. Irvine EJ, Zhou Q, Thompson AK. The Short Inflammatory Bowel Disease Questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators. Canadian Crohn's Relapse Prevention Trial. *Am J Gastroenterol* 1996;91:1571–8.
19. Casellas F, Alcala MJ, Prieto L, et al. Assessment of the influence of disease activity on the quality of life of patients with inflammatory bowel disease using a short questionnaire. *Am J Gastroenterol* 2004;99:457–61.
20. Fitzpatrick RD, Davy C, Buxton MJ, et al. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:3–47.
21. United States Department for Health. FDA Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;4:79.
22. Frost MH, Reeve BB, Liepa AM, et al. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health* 2007;10(Suppl 2):S94–S105.
23. Bassi A, Dodd S, Williamson P, et al. Cost of illness of inflammatory bowel disease in the UK: a single centre retrospective study. *Gut* 2004;53:1471–8.
24. The EuroQol Group. A new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
25. Snaith RP, Zigmond AS. The hospital anxiety and depression scale. *Br Med J* 1986;292:344.
26. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
27. Higgins PD, Schwartz M, Mapili J, et al. Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis. *Gut* 2005;54:782–8.
28. Irvine EJ. Quality of life issues in patients with inflammatory bowel disease. *Am J Gastroenterol* 1997;92:18S–24S.
29. Cohen RD. The quality of life in patients with Crohn's disease. *Aliment Pharmacol Ther* 2002;16:1603–9.
30. Drossman DA. Measuring quality of life in inflammatory bowel disease. *PharmacoEconomics* 1994;6:578–80.
31. Royal College of Physicians. The United Kingdom Inflammatory Bowel Disease Audit. http://www.rcplondon.ac.uk/resources/inflammatory-bowel-disease-audit (accessed 1 Jul 2013).
32. Siegel CA, Allen JI, Melmed GY. Translating improved quality of care into an improved quality of life for patients with inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2013;11:908–12.

195