
Article:
Conroy, S., Hale, M.F., Cross, S.S. orcid.org/0000-0003-2044-1754 et al. (4 more authors) (2018) Unrestricted faecal calprotectin testing performs poorly in the diagnosis of inflammatory bowel disease in patients in primary care. Journal of Clinical Pathology, 71 (4). ISSN 0021-9746

https://doi.org/10.1136/jclinpath-2017-204506

Unrestricted Faecal Calprotectin Testing Performs Poorly in the Diagnosis of Inflammatory Bowel Disease in Patients in Primary Care

Samantha Conroy[1], **Melissa F Hale**[1], **Simon S. Cross**[3], **Kirsty Swallow**[2], **Reena Sidhu**[1], **Ravi Sargur**[2], **Alan J. Lobo**[1]

Academic Department of Gastroenterology, Royal Hallamshire Hospital, Sheffield [1], Department of Immunology[2], Sheffield Teaching Hospitals NHS Foundation Trust and Academic Unit of Pathology[3], University of Sheffield.

Author for Correspondence:
Professor Alan J. Lobo, Gastroenterology Unit, P Floor Royal Hallamshire Hospital, Sheffield S10 2JF.
email: alan.lobo@sth.nhs.uk

Competing interests

AJL has received fees for lectures or as a member of advisory boards for Vifor Pharma, Takeda UK, Abbvie, Shield Therapeutics, Janssen and Dr Falk.
Other authors: no competing interests declared.

Abstract

Background: Faecal calprotectin (FC) measurement distinguishes patients with inflammatory bowel disease (IBD) from those with irritable bowel syndrome but evidence of its performance in primary care is limited.

Objective: To assess the yield of IBD from FC testing in primary care.

Methods: Retrospective review of hospital records to assess the outcome following FC testing in primary care. Investigations for all patients undergoing FC testing in a single laboratory for six months from 1.10.13 to 28.2.14 were reviewed.

Results: 410 patients (162 male; median age 42; range 16-91) were included. FC >50µg/g was considered positive (FC+). 148/410 (36.1%; median age 44 (17-91)) were FC+ (median FC 116.5µg/g (51-1770)). 122/148 FC+ patients (82.4%) underwent further investigation. 97 (65.5%) underwent lower GI endoscopy (LGIE), of which 7 (7.2%) had IBD. 49/262 (18.7%) FC negative (FC-) patients (FC ≤50µg/g) (median age 47 (19-76)) also underwent LGIE, of whom 3 (6.1%) had IBD.

IBD was diagnosed in 11/410 (2.7%; 4 ulcerative colitis, 3 Crohn's disease, 4 microscopic colitis). 8/11 were FC+ (range 67–1170) and 3 FC-. At a 50µg/g threshold, sensitivity for detecting IBD was 72.7%, specificity 64.9%, positive predictive value (PPV) 5.41%, negative predictive value 98.9%. Increasing the threshold to 100µg/g reduced the sensitivity of the test for detecting IBD to 54.6%.

Conclusion: FC testing in primary care has low sensitivity and specificity with poor PPV for diagnosing IBD. Its use needs to be directed to those with a higher pre-test probability of disease. Local services and laboratories should advise general practitioners accordingly.

Keywords: Faecal Calprotectin, inflammatory bowel disease, colonoscopy, primary health care

Introduction

Calprotectin is a neutrophil protein, the measurement of which, in faeces, detects intestinal inflammation[1]. Measurement of faecal calprotectin (FC), in people with gastrointestinal symptoms is thought to be useful in identifying those at a higher risk of harbouring inflammatory bowel disease (IBD), including ulcerative colitis and Crohn's disease[2].

FC testing has been recommended for those with new lower gastrointestinal symptoms, to help differentiate IBD and irritable bowel syndrome[2] and when symptoms suspicious for cancer[3] were absent. These guidelines emphasised the importance of appropriate local pathways of care and quality assurance and raised a series of hitherto unanswered questions relating to referral and investigation following FC testing in primary care. These questions included the proportion of patients with a raised FC who undergo further gastrointestinal investigation including which investigations are undertaken, how many of those investigations subsequently lead to a diagnosis of IBD when FC testing is used in primary care, the sensitivity and specificity of FC testing, the numbers of tested patients undergoing colonoscopy and the diagnostic yield of these colonoscopies[2]. Such information is of central importance in understanding the position of FC testing in a diagnostic pathway, utilising laboratory services, and which starts with patients in primary care.

This study was therefore undertaken to assess the diagnostic accuracy of FC in the detection of IBD in a primary care setting following the introduction of the test to the Sheffield Teaching Hospitals Trust in October 2013.

Methods

Patients

Consecutive patients who had a FC measurement undertaken in the Immunology Laboratories of the Sheffield Teaching Hospitals NHS Foundation Trust between 01/10/13 and 28/02/14 were identified. Subsequent inclusion in this evaluation is outlined in Figure 1. Patients aged under 16 years (n=4), those with known IBD (n=23) (as identified from clinic letters and investigations) and where FC testing was initiated in secondary care (n=57), were excluded from the study. Further investigations had therefore not been undertaken at the time of FC testing.

Where patients had raised FC levels but no identifiable further assessment, their general practitioners (GPs) were approached by letter for information about further investigations (n=20). Patients for whom no reply was received were also excluded from the study (n=6).

For those who had repeated FC measurements over the inclusion period, the first FC measurement initiated by the GP was used for analysis.

Measurement of FC

FC was measured by IDK® Calprotectin ELISA (Immunodiagnostik, Germany). The normal range (FC negative) in adults was taken as ≤ 50 µg/g faeces[2]. Results of testing were available to those accepting referrals and at the point of subsequent diagnoses.

Identification of further investigations and hospital assessments following FC testing

Using patient administration systems and laboratory databases, we retrospectively identified investigations undertaken. These included blood tests for erythrocyte sedimentation rate (ESR) and c-reactive protein (CRP) within 2 months before or after FC testing and follow-up imaging or endoscopic evaluation within 18 months. Hospital assessment at our centre within 12 months following the FC test by a gastroenterologist, colorectal or upper gastrointestinal surgeon was also identified from hospital records.

Lower GI Endoscopy and Imaging

Positive findings on lower GI endoscopy were defined as identification of IBD, colorectal cancer or adenomas. Positive findings on further imaging included evidence of diagnoses such as diverticulitis and appendicitis. All other findings were defined as non-significant. A diagnosis of IBD was made on the basis of typical histological features in biopsies taken at lower GI endoscopy, typical radiological features[4,5] or accepted diagnostic appearances at small bowel video capsule endoscopy[6,7]

Statistics

Descriptive statistics are presented as the mean or median (as appropriate). Categorical data was analysed using a chi squared test, continuous data with either t-test (normally-distributed data) or Mann Whitney U test (non-parametric data), where statistical significance was considered as $p<0.05$. Analyses were performed using Statistical Package for the Social Sciences ((SPSS), IBM, Somers, USA. Version 22.0). Bonferroni's correction was used for multiple univariate tests, with $p<0.007$ being regarded as significant for comparison of rates for different investigations between FC positive and FC negative patients (Figure 3). Exploratory assessment of sensitivity, specificity, positive and negative predictive value at differing FC thresholds was also undertaken. Receiver operating characteristic (ROC) curves were also derived for the diagnosis of IBD or organic disease, using the

ROCR package in the R environment for statistical computing and the area under the curve (AUROCC) calculated[8].

The study was undertaken as Sheffield Teaching Hospitals NHS Foundation Trust, Clinical Effectiveness Unit Service Evaluation, Reference 6509, December 2014.

Results

500 patients were identified of whom 410 patients (162 male) were included in the analysis (Figure 1). The median age was 42 (range 16-91). 148 patients had positive FC levels (median FC 116.5 µg/g faeces (range 51-1770)), with a median age of 44 years (range 17-91). 262 were FC negative (median FC 19 µg/g faeces (range 0-50)), with a median age of 40 years (range 16-88).

Faecal Calprotectin Levels

The overall distribution of FC values in the study cohort is shown in Figure 2.

Referral to secondary care

196/410 patients (47.8%) underwent referral to a gastrointestinal specialist. Of these, 126 (64.3%) had a raised FC. FC positive patients were significantly more likely to be referred to secondary care (126/148 (85.1%)) than those with a FC negative result (70/262 (26.7%)) ($p<0.001$).

Investigations

Further investigation in the FC positive and FC negative groups is shown in Figure 3. 97/148 (65.5%) of FC positive patients underwent lower GI endoscopy (90 colonoscopies, 7 flexible sigmoidoscopies) compared to 49/262 (18.7%) FC negative patients (43 colonoscopies, 6 flexible sigmoidoscopies).

49 patients FC negative patients still underwent lower GI endoscopy and were significantly older (median age 47 (range 19-76)) than FC negative patients who did not undergo colonoscopy (median age 39 (range 16-88) ($p=0.002$)).

80/97 (82.5%) of FC positive patients were found to have no significant abnormality at lower GI endoscopy compared to 44/49 (89.8%) FC negative patients who had lower GI endoscopy (p=0.33).

94 (63.5%) FC positive and 142 (54.2%) FC negative patients had CRP testing – with higher values in the FC positive patients (median 3.1 (range 0-87)) than in the FC negative patients (1.3 mg/L (range 0-105)) respectively, (p<0.001). 94 (63.5%) FC positive patients and 151 (57.6%) FC negative patients had ESR measured with respective median values of 7 mm/hr (range 2-41) and 5mm/hr (2-95) (p=0.009).

Diagnosis of IBD

11/410 (2.6%) patients who underwent FC testing were diagnosed with new IBD (Table 1). For newly diagnosed IBD, the sensitivity and specificity for the pre-specified diagnostic threshold are shown in Table 2, alongside values from existing experience in this field, including secondary care studies used to develop NICE guidance.

The performance of the test using different thresholds of abnormality is described in Table 3. Altering the threshold at which FC is regarded as 'positive' did not affect the performance at the thresholds considered.

The ROC curve for the performance of FC in diagnosing IBD is given in Figure 4a. The AUROCC was 0.69.

The accuracy (number of true positive tests + number of true negative tests as a proportion of all tests) in diagnosing IBD is shown in Figure 5, with poor accuracy below a threshold of 250 µg/g faeces. Only 30 of the cohort had FC levels > 250µg/g and only 3 of these had a diagnosis of IBD.

Other diagnoses

In addition to those diagnosed with IBD, a number of positive diagnoses of organic disease were made following endoscopic and radiological investigation. Of those who were FC+, 2 patients had colorectal cancer, 8 colorectal adenomatous polyps, 1 diverticulitis, 1 appendicitis. One FC+ patient had a final diagnosis of diversion proctitis, but it was felt that the FC result obtained from the proximal stoma could not reflect that diagnosis, and so was not included as having detected IBD. Of those who were FC-, 2 had colorectal adenomatous polyps and 1 appendicitis. The ROC curve for the performance of FC in diagnosing organic disease is given in Figure 4b. The AUROCC was 0.76.

Discussion

We have described the experience of primary care FC testing in a single, predominantly urban area. This has highlighted low sensitivity and specificity for the detection of IBD and a high rate of negative endoscopic investigations.

The introduction of FC testing has been considered an opportunity to reduce unnecessary gastrointestinal investigation in people suspected of having IBD[2]. The current study was undertaken to examine the outcome of its use in a routine primary care setting, where there is limited information[7] with only two studies assessing its sensitivity and specificity[9,10]. Pavlidis et al showed high sensitivity and specificity of FC testing in primary care (82% and 77%, respectively)[9]. However, the performance of the test was assessed in differentiating organic from non-organic disease, rather than IBD from non-IBD. Patients were younger, with a median age of 33 compared to 42 in the current study. This may have contributed to the lower sensitivity and specificity of 72.7% and 64.9% respectively in our study, where differentiation of IBD from non-IBD was assessed. If other organic diagnoses are included, the sensitivity increases to 78% and the specificity to 69%. The age range of patients in the current study also included a number of elderly patients where the yield of IBD may be lower and investigation for other pathology, especially cancer, may have been required irrespective of FC result.

We have shown lower sensitivity and specificity for FC testing in primary care than in secondary care, where a meta-analysis of five studies used by NICE to create DG11[2] gave sensitivity and specificity values of 93% and 94%, respectively, when distinguishing IBS and IBD. In addition, the sensitivity values increased to 100% when comparing IBD versus non-IBD. However, eight of eleven studies included in the meta-analysis were paediatric studies. The PPV of 5.4% seen in our study is also lower than other studies. This may be due to the lower prevalence of IBD in our primary care population at 2.6%, than in a secondary care cohort. Importantly however, the NPV was high at 98.9%, suggesting a role for FC testing as a 'rule-out' test for IBD. 64% of FC results were negative

and a further 16% fell between 51 and 100 µg/g. This distribution may have skewed the NPV. The skewed distribution might reflect that GPs are using the test to aid diagnosis, and that those with obvious symptoms of severe IBD may have been referred – appropriately - without testing. 3/11 patients diagnosed with IBD had a negative FC, reinforcing the importance of other clinical and biochemical findings. Further, only 1 of these 3 patients had small bowel inflammatory disease to potentially explain the false negative result.

Over 80% of all people who were investigated by lower GI endoscopy did not have significant pathological findings – in either the FC positive and negative groups. It is possible that FC testing may therefore be increasing the number of unnecessary colonoscopies. There are a number of factors influencing changes in the numbers of patients referred for colonoscopy and it is not possible to determine, from this study, how much the introduction of FC testing contributes to this. A significant number of patients with negative FC values still underwent lower GI endoscopy. This group were significantly older than those with a negative FC who did not undergo lower GI endoscopy, suggesting that the clinical concern was of cancer. Recent UK guidance recommends that a threshold for FC above which colonoscopy should be undertaken is set based on local audit and that measurement should not be undertaken in older patients with altered bowel habit[11]. Increasing demand for unnecessary endoscopy has been identified in patients in a secondary care setting though the same study also confirmed a high NPV for detecting organic disease for an FC between 100 and 200µg/g[11]. It has also been suggested that incorporating FC measurement into a defined pathway of care – including using a threshold of 250µg/g for fast track investigation and repeat testing for those with intermediate results - may improve the PPV for organic disease[10]. However, we have demonstrated a poor accuracy below a FC threshold of 250 µg/g and raising the threshold to 250µg/g reduced sensitivity to 27%, which would have left over half of the IBD patients undetected – and repeat testing has not been supported in national guidance[11]. This poor performance was also demonstrated by the low AUROC. The higher PPV may represent a different tested population – with a lower median age and a pathway specifying that cancer was not suspected. In our series, a change in threshold did not improve the performance of the test.

This was a retrospective assessment. Where investigation results were not immediately identifiable, a good response was obtained with follow-up letters to general practitioners. There may therefore be a

verification bias in this study, as an assumption has been made that those with a negative FC value did not develop IBD, if they underwent no further documented assessment. It is possible that patients underwent investigation at another centre, or privately, but we feel that the risk of this is small given local referral practice.

Details of associated symptoms might have enhanced this study, but would have needed to be collected prospectively. It seems likely that, in practice, more focused use of FC testing in particular symptom groups – ie a cohort with higher pre-test probability of IBD - would help improve the performance of FC testing. Information on the use of proton pump inhibitors or non-steroidal anti-inflammatory drug use, which may elevate FC[13, 14] was also not available, and it is possible that this might have accounted for the normal investigations in those with raised FC. Assessment of FC in combination with other markers of inflammation, such as ESR or CRP may be helpful in distinguishing those with IBD from other diagnoses. Interestingly, although both ESR and CRP were higher in the FC positive patients, the median levels were still close to normal and where IBD was diagnosed, only 2/7 were associated with a CRP greater than 5mg/L.

Unrestricted FC testing in primary care in this study had lower sensitivity and specificity than in secondary care and a poor PPV for diagnosing IBD. Over 80% of all people who were investigated by lower GI endoscopy did not have significant pathology and a substantial number of FC negative patients still underwent lower GI endoscopy. Studies are therefore required to identify and define the population in primary care with a higher pre-test probability of IBD, where FC testing may be more discriminatory. In the interim, laboratory services should agree with clinicians local guidelines to optimise use of the test.

Key Messages

- Unrestricted testing of faecal calprotectin in primary care has a low sensitivity and specificity and a poor positive predictive value for diagnosing inflammatory bowel disease (IBD).
- As a result, over 80% of those with positive faecal calprotectin results had no significant abnormality at colonoscopy, which was not significantly different from those who were negative.
- A raised threshold for colonoscopy to 250µg/g reduces the sensitivity for diagnosing IBD.
- A different strategy for testing, focusing on those with higher pre-test probability of disease, should be developed.

References

[1] Røseth AG, Schmidt PN, Fagerhol MK. Correlation between Faecal Excretion of Indium-111-Labelled Granulocytes and Calprotectin, a Granulocyte Marker Protein, in Patients with Inflammatory Bowel Disease. *Scand J Gastroenterol*. 1999; 34:50-4

[2] NICE DG11 (2013). Faecal calprotectin diagnostic tests for inflammatory diseases of the bowel. London: NICE. ISBN: 978-1-4731-0326-9. https://www.nice.org.uk/guidance/dg11 (accessed 12.11.16)

[3] NICE NG12 (2015). Suspected cancer: recognition and referral. London: NICE. ISBN: 978-1-4731-1164-6. https://www.nice.org.uk/guidance/ng12 (accessed 12.11.16)

[4] Feakins RM. Inflammatory bowel disease biopsies: updated British Society of Gastroenterology reporting guidelines. *J Clin Pathol*. 2013; 66(12):1005-26.

[5] Mowat C, Cole A, Windsor A et al. Guidelines for the management of inflammatory bowel disease in adults. *Gut*. 2011; 60(5): 571-607

[6] Bourreille A, Ignjatovic A, Aabakken L, *et al*. Role of small-bowel endoscopy in the management of patients with inflammatory bowel disease: an international OMED-ECCO consensus. *Endoscopy* 2009;41:618-37.

[7] Mow WS, Lo SK, Targan SR et al. Initial experience with wireless capsule enteroscopy in the diagnosis and management of inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2004; 2(1): 31-40.

[8] Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005; 21 (20): 3940-3941.

[9] Pavlidis P, Chedgy F, Tibble J. Diagnostic accuracy and clinical application of faecal calprotectin in adult patients presenting with gastrointestinal symptoms in primary care. *Scand J Gastroenterol*. 2013; 48:1048-54.

[10] Turvill J, O'Connell S, Brooks A et al. Evaluation of a faecal calprotectin care pathway for use in primary care. *Prim Health Care Res Dev*. 2016; 1-9.

[11] Brookes M, Gaya D, Hawthorne B on behalf of the IBD section committee. BSG Guidance Document on use of Faecal Calprotectin. http://www.bsg.org.uk/clinical-guidance/ibd/bsg-guidance-on-the-use-of-faecal-calprotectin-testing-in-ibd.html (accessed 12.11.16)

[12] Seenan JP, Thomson F, Rankin K, Smith K, Gaya DR. Research: Are we exposing patients with a mildly elevated faecal calprotectin to unnecessary investigations? Frontline Gastroenterol 2015;6:3 156-160

[13] Tibble J, Sigthorsson G, Foster R et al. High prevalence of NSAID enteropathy as shown by a simple faecal test. *Gut*. 1999; 45:362-6.

[14] Poullis A, Foster R, Mendall M. Proton pump inhibitors are associated with elevation of faecal calprotectin and may affect specificity. *Eur J Gastroenterol Hepatol*. 2003; 15:573-4

[15] Silverberg MS, Satsangi J, Ahmad T, 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. Can J Gastroenterol. 2005; 19:5–36.

Patient	Age (years)	UC/CD	Diagnosed by	Disease extent (Montreal Classification[15])	FC (μg/g faeces)	CRP (mg/L)	ESR (mm/hour)
1	29	UC	Colonoscopy	E2	36	3.2	-
2	19	UC	Flexible Sigmoidoscopy	E2	0	-	-
3	26	UC	Colonoscopy	E1	90	0.9	2
4	22	UC	Colonoscopy	E3	1770	-	26
5	67	CD	VCE	L1 B2	0	8.5	13
6	26	CD	VCE	L3 B1	883	2.4	2
7	27	CD	VCE	L4, B1	67	10.6	5
8	37	Collagenous colitis	Colonoscopy	-	157	0.9	2
9	67	Collagenous colitis	Colonoscopy	-	130	8	26
10	44	Lymphocytic colitis	Colonoscopy	-	480	7.6	2
11	32	Lymphocytic colitis	Colonoscopy	-	239	16.7	-
Median	29				130	7.6	3.5
Range	19-67				0-1770	0.9-16.7	2-26

Table 1. Clinical and laboratory details of patients diagnosed with inflammatory bowel disease following FC testing in primary care. (FC, faecal calprotectin; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; CD, Crohn's disease; UC, ulcerative colitis; VCE, small bowel video capsule endoscopy)

			Other Studies		
	Current Study		Pavlidisa 2013[6]	Turvill 2016[8]	NICE DG11[3]: Secondary Care Meta-Analysis
Disease differentiation	For IBD vs no IBD	Organic v non-organic	Organic v non-organic		
Primary or secondary care setting	**Primary**	**Primary**	**Primary**	**Primary**	**Secondary**
Threshold for positive faecal calprotectin (µg/g faeces)	**>50**	**>50**	**>50**	**>250**	**>50**
Sensitivity	**72.7%**	**77.8%**	**82%**	**89%**	**93%**
Specificity	**64.9%**	**66.8%**	**77%**	**-**	**94%**
PPV	**5.41%**	**14.2%**	**28%**	**40%**	**5.9%**
NPV	**98.9%**	**97.7%**	**98%**	**97%**	**Not stated**

Table 2. Summary of sensitivity and specificity of faecal calprotectin testing in current study and other studies in primary and secondary care. (IBD = inflammatory bowel disease; PPV positive predictive value; NPV negative predictive value)

Faecal calprotectin threshold for abnormal test	Sensitivity	Specificity	PPV	NPV
> 50μg/g	72.7%	64.9%	5.41%	98.8%
> 100μg/g	54.5%	80.5%	7.14%	98.5%
> 125μg/g	54.5%	83.5%	8.33%	98.5%
> 150μg/g	45.5%	86.7%	8.62%	98.3%
> 250μg/g	27.3%	93.5%	10.3%	97.9%

Table 3. Performance of faecal calprotectin and testing using sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) at different thresholds of faecal calprotectin for detection of inflammatory bowel disease in patients tested in primary care.

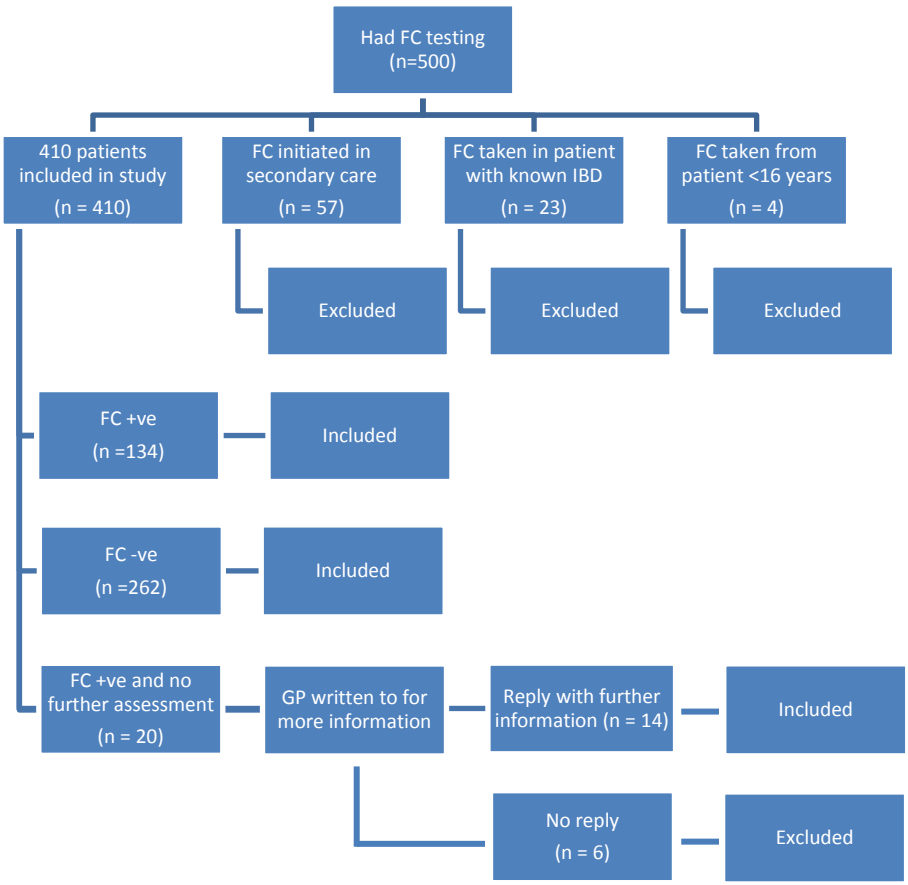

Figure 1. Flow diagram for patient inclusion in study. (FC faecal calprotectin; IBD inflammatory bowel disease; GP general practitioner)

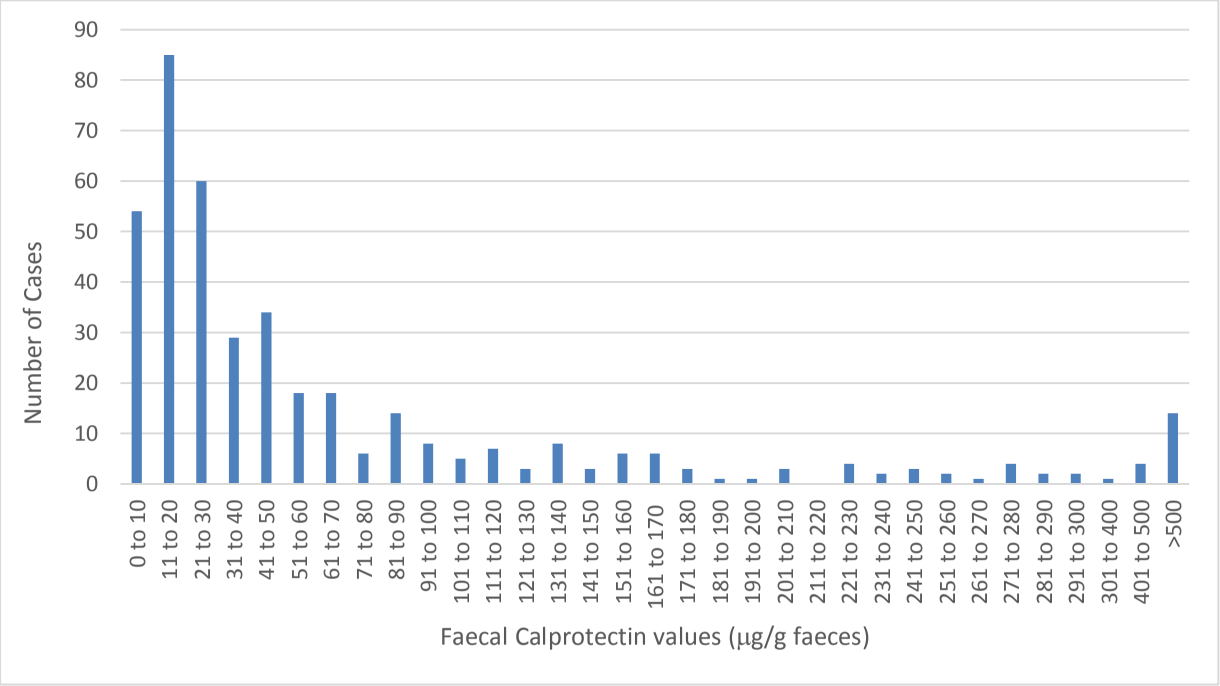

Figure 2. Overall distribution of faecal calprotectin levels

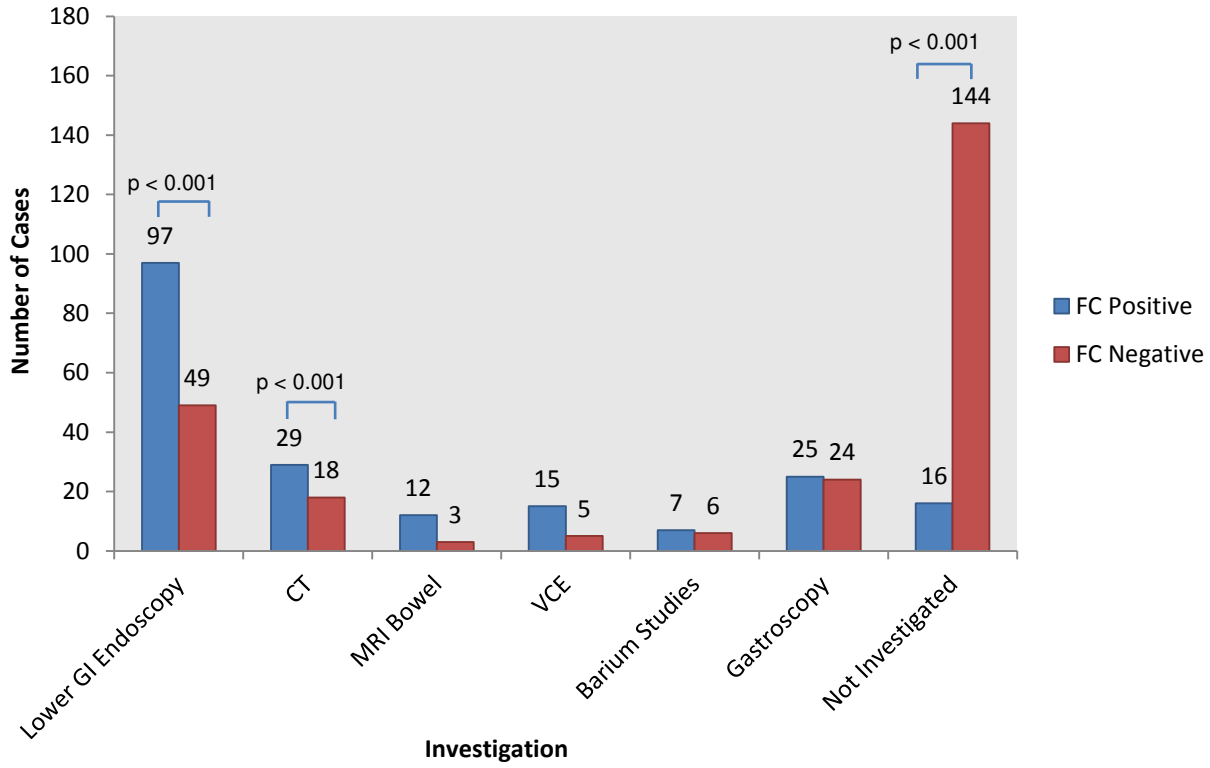

Figure 3. Further investigations undertaken in faecal calprotectin positive and negative patients. Significant differences shown for numbers of patients undergoing lower gastrointestinal (GI) endoscopy, computerized tomography (CT) scan and those not investigated (p <0.007 regarded as significant to account for multiple testing). (MRI magnetic resonance imaging; VCE video capsule endoscopy; FC Faecal calprotectin)

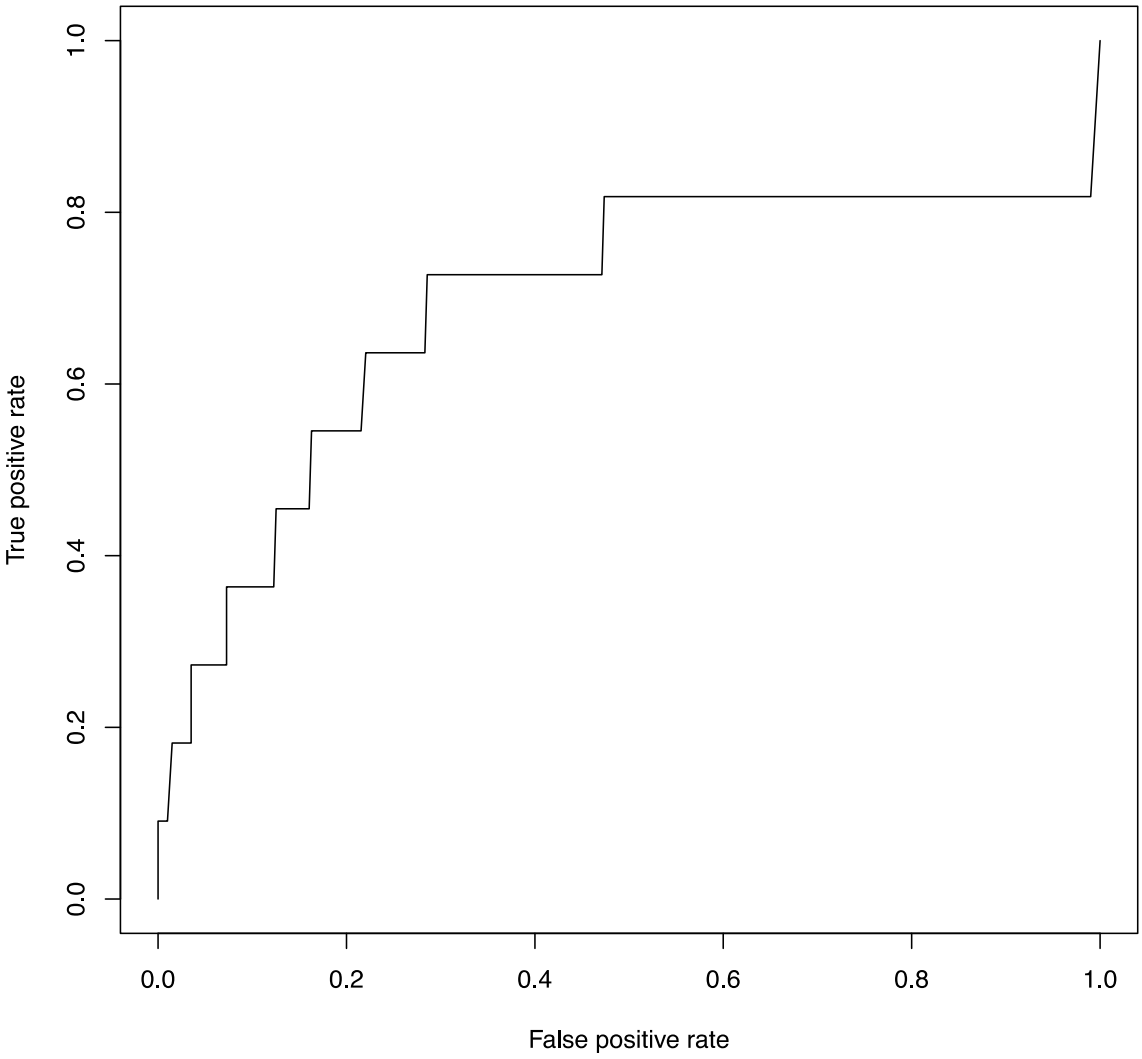

Figure 4a. Receiver operating characteristic (ROC) curve for the diagnosis of inflammatory bowel disease using faecal calprotectin testing. Area under the ROC curve = 0.69.

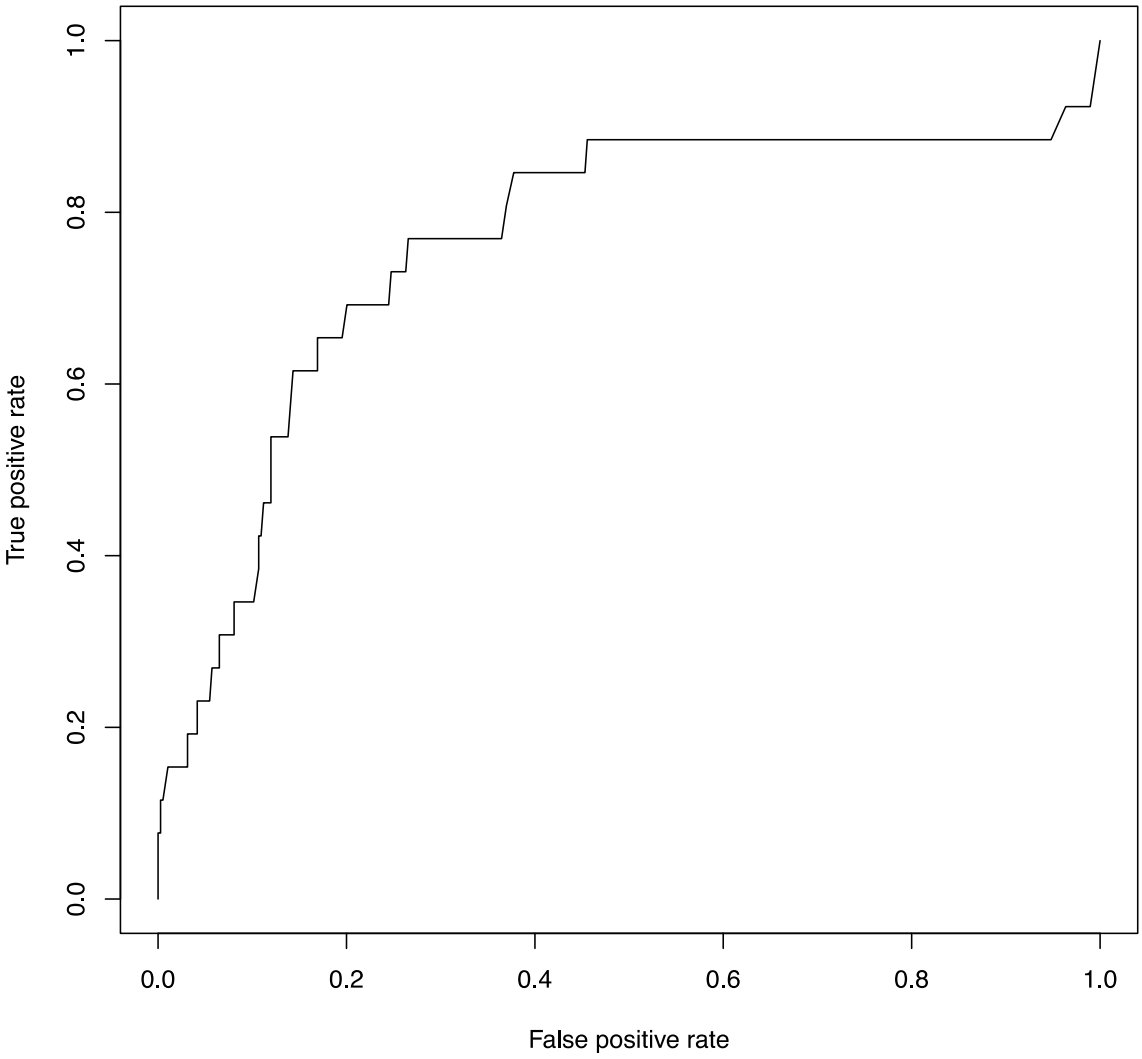

Figure 4b. Receiver operating characteristic (ROC) curve for patients tested by measurement of faecal calprotectin and with a final diagnosis of organic bowel diseases. Area under the ROC curve = 0.76.

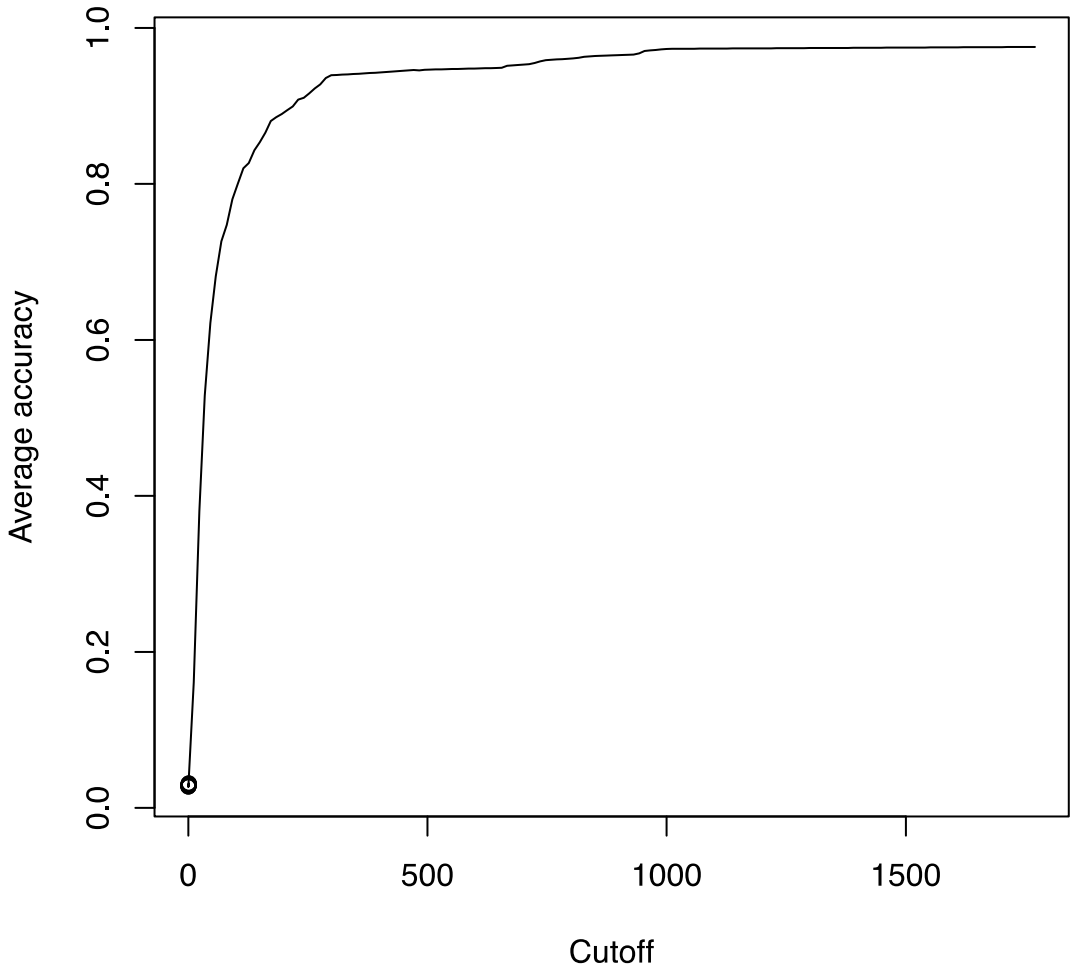

Figure 5. Average accuracy curve for diagnosis of inflammatory bowel disease at differing thresholds of faecal calprotectin ('Cutoff') in µg/g faeces.