Optimising triage of urgent referrals for suspected IBD: results from the Birmingham IBD inception study

Rimmer, Peter; Cheesbrough, Jonathan; Harris, Jane; Love, Melanie; Tull, Samantha; Iqbal, Asif; Regan-Komito, Daniel; Cooney, Rachel; Hazel, Karl; Sharma, Naveen; Dietrich, Thomas; Chapple, Iain; Quraishi, Mohammad Nabil; Iqbal, Tariq H

DOI:
10.1136/flgastro-2023-102523

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Rimmer, P, Cheesbrough, J, Harris, J, Love, M, Tull, S, Iqbal, A, Regan-Komito, D, Cooney, R, Hazel, K, Sharma, N, Dietrich, T, Chapple, I, Quraishi, MN & Iqbal, TH 2024, 'Optimising triage of urgent referrals for suspected IBD: results from the Birmingham IBD inception study', *Frontline Gastroenterology*. https://doi.org/10.1136/flgastro-2023-102523

OPEN ACCESS

Original research

Optimising triage of urgent referrals for suspected IBD: results from the Birmingham IBD inception study

Peter Rimmer ,[1,2] Jonathan Cheesbrough,[1,2] Jane Harris,[2] Melanie Love,[2] Samantha Tull,[3] Asif Iqbal,[1,3] Daniel Regan-Komito,[4] Rachel Cooney ,[1,2] Karl Hazel,[1,2] Naveen Sharma,[2] Thomas Dietrich,[1,5] Iain Chapple,[1,5] Mohammad Nabil Quraishi,[1,2] Tariq H Iqbal[1,2]

[1]Birmingham NIHR Biomedical Research Centre in Inflammation, University of Birmingham, Birmingham, UK
[2]Gastroenterology Department, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
[3]Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
[4]Immunology, Infectious Diseases and Ophthalmology (I2O), Roche Pharma Research & Early Development, F Hoffman La Roche, Basel, Switzerland
[5]The School of Dentistry, University of Birmingham College of Medical and Dental Sciences, Birmingham, UK

Correspondence to
Dr Peter Rimmer, Birmingham NIHR Biomedical Research Centre in Inflammation, University of Birmingham, Birmingham, UK; p.rimmer@bham.ac.uk

Received 7 August 2023
Accepted 6 February 2024

To cite: Rimmer P, Cheesbrough J, Harris J, *et al*. *Frontline Gastroenterology* Epub ahead of print: [please include Day Month Year]. doi:10.1136/flgastro-2023-102523

ABSTRACT

Objective Diagnostic delays in inflammatory bowel disease (IBD) result in adverse outcomes. We report a bespoke diagnostic pathway to assess how best to combine clinical history and faecal calprotectin (FCP) for early diagnosis and efficient resource utilisation.

Methods A rapid-access pathway was implemented for suspected IBD patients referred outside urgent 'two-week wait' criteria. Patients were triaged using symptoms and FCP. A 13-point symptom history was taken prediagnosis and clinical indices, including repeat FCP, collected prospectively.

Results Of 767 patients (January 2021–August 2023), 423 were diagnosed with IBD (208 Crohn's disease (CD), 215 ulcerative colitis (UC)). Most common symptoms in CD were abdominal pain (84%), looser stools (84%) and fatigue (79%) and in UC per-rectal bleeding (94%), urgency (82%) and looser stools (81%). Strongest IBD predictors were blood mixed with stools (CD OR 4.38; 95% CI 2.40–7.98, UC OR 33.68; 15.47–73.33) and weight loss (CD OR 3.39; 2.14–5.38, UC OR 2.33; 1.37–4.00). Repeat FCP testing showed reduction from baseline in non-IBD. Both measurements >100 µg/g (area under the curve (AUC) 0.800) and >200 µg/g (AUC 0.834) collectively predicted IBD. However, a second value ≥220 µg/g considered alone, regardless of the first result, was more accurate (Youden's index 0.735, AUC 0.923). Modelling symptoms with FCP increased AUC to 0.947.

Conclusion Serial FCP measurement prevents unnecessary colonoscopy. Two FCPs >200 µg/g could stream patients direct to colonoscopy, with two >100 µg/g prompting clinic review. A second result ≥220 µg/g was more accurate than dual-result thresholds. Coupling home FCP testing with key symptoms may form the basis of effective self-referral pathways.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Symptom prevalence at inflammatory bowel disease (IBD) onset has been evaluated in a North American Cohort, while symptom prediction has been analysed retrospectively in the UK. No UK study has evaluated symptom prediction prospectively, particularly in a cohort largely preselected using elevated faecal calprotectin (FCP).

WHAT THIS STUDY ADDS

⇒ This study highlights the key symptoms that separate IBD from other, predominantly functional, diagnoses. It also informs our understanding of how best to apply FCP in a health service under increasing strain.

INTRODUCTION

The importance of early diagnosis in inflammatory bowel disease (IBD) is well established. A recent meta-analysis demonstrated that delayed diagnosis is associated with adverse outcomes in both ulcerative colitis (UC) and Crohn's disease (CD).[1] In CD, there is a higher likelihood of stricturing and penetrating disease, translating to an increased risk of need for surgery. In UC, delayed diagnosis increased the risk of colectomy. A collated analysis of all articles in this review found median time to diagnosis was 8 months in CD and 3.7 months in UC.[1] Moreover, these studies took place prior to the advent of the COVID-19 pandemic, where research has demonstrated increases in healthcare avoidance.[2 3]

In IBD diagnostic delays can stem from lateness seeking medical opinion,

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ In its simplest form, this study may help improve referral triage and optimise utilisation of FCP in IBD referral pathways. It has the potential to reduce pressure on endoscopy and outpatient services and ensure the correct patients are prioritised. Our work aligns with 'Getting It Right First Time' but aims to take this one step further. Direct to endoscopy referrals for FCP >200 µg/g on two occasions could reduce strain on outpatient clinics without imposing further delay in diagnosis, while repeat sampling will also prevent unnecessary endoscopy. However, as we plan new care pathways, our data would support a move away from traditional approaches. A second FCP ≥220 µg/g, regardless of the first result, was a more accurate predictor. Coupled with key clinical symptoms, the formulation of reliable self-referral pathways based on patient-initiated home FCP testing could allow even more rapid diagnosis and treatment initiation. This would require further validation with a large prospective cohort, a process that we are currently initiating.

delays after first healthcare interaction (typically primary care) and delays after referral to secondary care. Gastrointestinal symptoms can extend as far back as 5 years prior to diagnosis in IBD patients.[4] In secondary care, COVID-19 has resulted in increased waiting times for diagnostics and treatment.[5] Time from referral to review should still not exceed 4 weeks but the current reality in the UK health service, set against the increasing prevalence of IBD, threatens the consistent achievement of this.[6 7] 'Getting It Right First Time' is a national programme within NHS England, which has set out suggested referral pathways and strategies for IBD diagnosis.[8] National campaigns led by Crohn's and Colitis UK have followed the 2019 IBD UK National report which highlighted delays in diagnosis as one of the major challenges facing patients and clinical services.[9 10] Moreover, time to diagnosis following referral has recently been established as a key performance indicator for IBD.[11]

With burgeoning waiting lists for suspected IBD patients, effective triage of referrals from primary care is vital. A well-established tool to support this is faecal calprotectin (FCP) testing in primary care. Its use has been shown to reduce time to diagnosis and forms a focal point of most contemporary referral pathways.[7 12] Nonetheless, several non-IBD conditions and common medications are associated with elevated FCP. This includes bacterial and viral gastrointestinal infections, diverticular disease, colonic polyps, colorectal cancer, GI bleeding, non-steroidal anti-inflammatory drugs, proton pump inhibitors and obesity.[13] There is no international consensus on specific thresholds for significant elevations in FCP. Implementation studies emphasise that where symptom severity allows, repeating FCP testing more than once would aid diagnostic specificity, particularly after a period of the aforementioned medications.[14 15] Different time intervals have been suggested without clear consensus, ranging from 2 to 8 weeks. Current pressures in primary care mean repeat testing is increasingly rare prior to referral.

However, in this FCP driven environment, the value of the clinical history should not be forgotten. The last large study to prospectively collect presenting symptoms at IBD onset was undertaken by Perler *et al* in North America between 2008 and 2013.[16] However, the discriminant ability of each symptom for IBD was not explored. The most recent UK study

Table 1 Overall cohort demographics

	Crohn's (n=208)		UC (n=215)		Non-IBD (n=339)	
Age, median (IQR), years	34 (22)		37 (22.5)		35 (20)	
BMI, median (IQR), years	25.3 (9.53)		25.5 (6.18)		27.9 (10.4)	
Sex, % male	43		60		40	
Symptom duration, median (IQR), months	10 (19)		4 (8.5)		6 (9)	
Ethnicity, %						
White	67		66		74	
Asian	25		30		17	
Black	8		4		9	
Location (or grouping), n (%)	Ileal (L1)	91 (44)	Proctitis (E1)	69 (32)	Transient infection	54 (16)
	Colonic (L2)	59 (28)	Left sided (E2)	83 (39)	Functional	171 (50)
	Ileocolonic (L3)	58 (28)	Extensive (E3)	63 (29)	Other GI	114 (34)
	Perianal	22 (11)				
Behaviour, n (%)	Non-stricturing, non-penetrating	172 (83)	Mild	87 (41)		
	Stricturing	31 (15)	Moderate	99 (46)		
	Penetrating	5 (2)	Severe	29 (13)		

BMI, body mass index; IBD, inflammatory bowel disease; UC, ulcerative colitis.

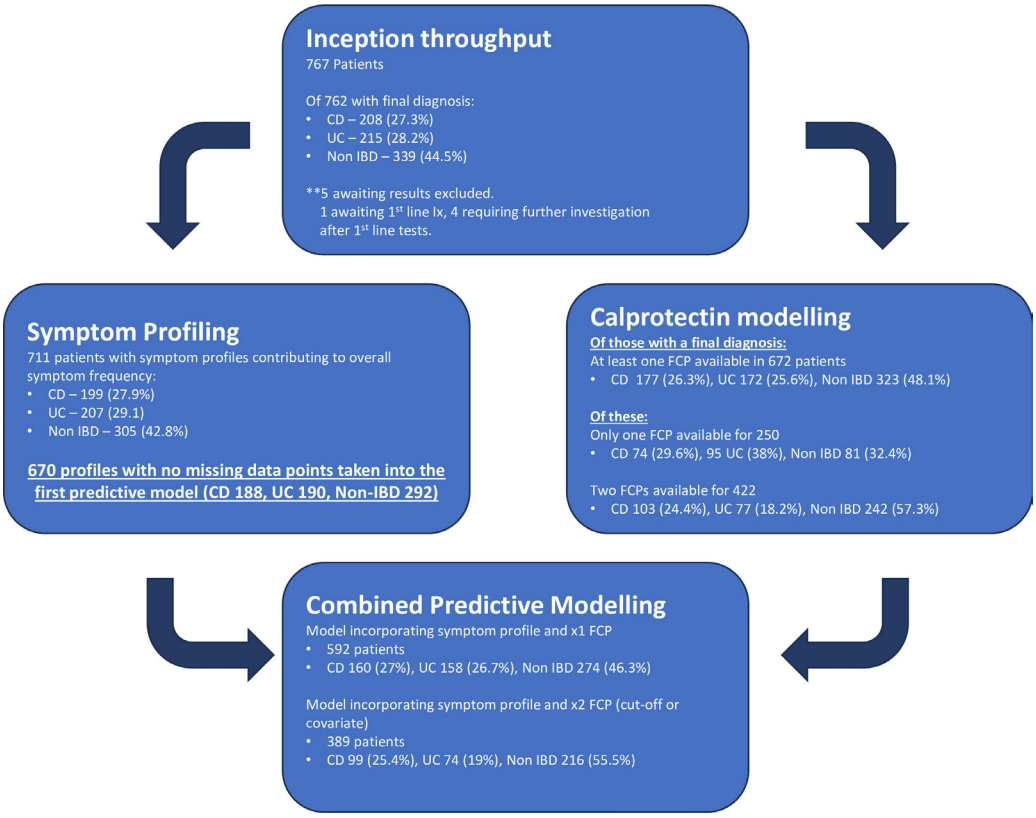

Figure 1 Inception clinic throughput split by diagnosis and overall patient contribution to each component of the analysis and modelling presented in this manuscript. CD, Crohn's disease; FCP, faecal calprotectin; IBD, inflammatory bowel disease; UC, ulcerative colitis.

retrospectively analysed primary care records only.[17] Neither accounted for FCP. To date, an evaluation of the predictive ability of individual symptoms in a cohort selected based on initial FCP result has not been undertaken.

The objectives of this study were therefore to analyse existing tools, with a view to providing clearer guidance to improve referral triage. Specific aims were :

- To identify the preferred FCP thresholds for IBD diagnosis.
- To assess symptom frequency and predictive ability in a cohort preselected by FCP.
- To identify symptom complexes which, when coupled with FCP results, carry the highest pretest probability of IBD.

METHODS

The Birmingham IBD inception pathway is a rapid access clinic for patients over 16 years of age referred with suspected IBD. It was established to address spiralling waiting times for new patient referrals. Consequently, the additional capacity was used pragmatically. Inclusion criteria were not strictly applied and the decision to stream patients to the inception pathway was at the discretion of the triaging clinician. All patients seen were included in the dataset to capture the heterogeneity of real-world practice. In general, patients seen did:

- Not meet criteria for 'straight to test' colorectal cancer referral, whereby existing two-week wait pathways still apply.
- Have symptoms that the triaging clinician felt were compatible with an IBD diagnosis.
- (unless clinical suspicion is very strong) Have supporting evidence in the form of 'elevated' FCP (no set cut-off applied, patients with only one result at referral were still seen).

Patients were sent a stool sample collection kit and asked to bring this to their first appointment to check FCP and minimise delay. FCP was determined using the Buhlmann fCAL turbo test (Buhlmann, Basel-Landschaft). In those with significant symptomatology, failure to return a sample did not prevent progression through the pathway. A standardised 13-point symptom history and duration was obtained from patients by the responsible clinician at the time of the index, pre-diagnosis, appointment. Responses were recorded electronically. Diagnoses were established using history, biochemistry, endoscopy, histological and radiological criteria in line with the European Crohn's and Colitis Organisation guidelines.[18] Clinical and endoscopic severity indices were collected prospectively. Endoscopy was undertaken by IBD physicians on colonoscopy lists ring-fenced for patients on the pathway. Those where IBD was excluded were categorised into broad groups.

Table 2 Differences in FCP at baseline compared between the cohorts providing one and two FCP results, with subsequent comparison of baseline and repeat FCP in those providing two, split by diagnostic subgroup

	Baseline FCP µg/g (one result only)*	Baseline FCP µg/g (two results only)*	Baseline group comparison (within subgroup)	Repeat FCP µg/g (two results only)	Repeated measure comparison (within subgroup)
Crohn's disease					
N	N=74	N=103	Mann-Whitney U p=0.21	N=103	Wilcoxon rank p=0.14
Median (IQR)	664 µg/g (1300)	915 µg/g (1432)		592 µg/g (1347)	
Ulcerative colitis					
N	N=95	N=77	Mann-Whitney U p=0.95	N=77	Wilcoxon rank p=0.09
Median (IQR)	1213 µg/g (1510)	1293 µg/g (1577)		1990 µg/g (1742)	
Non-IBD					
N	N=81	N=242	Mann-Whitney U p=0.002	N=242	Wilcoxon rank p<0.001
Median (IQR)	291 µg/g (444)	391 µg/g (641)		36 µg/g (118)	

*Across all diagnoses, patient age was higher in the cohort providing one FCP (1 FCP median 39 (IQR 25), 2 FCP median 33 (IQR 18), Mann-Whitney U p<0.001), while a higher proportion were male (55.7% vs 38.7%, χ^2 p<0.001).
FCP, faecal calprotectin; IBD, inflammatory bowel disease.

Statistical analyses

All analyses and modelling were undertaken in Jamovi, using R packages.[19–24] JASP was used for additional visualisations.[25] Only those with a final diagnosis were included in the onward analyses. All datasets presented followed a skewed distribution (first FCP Shapiro-Wilk (SW) 0.846 p<0.001, second FCP SW 0.697 p<0.001, difference between FCPs SW 0.926 p<0.002, symptom duration SW 0.589 p<0.001, age SW 0.95 p<0.001, body mass index SW 0.94 p<0.001). As such, non-parametric tests are used throughout. Mann-Whitney U tests were used for two groups, with Kruskal-Wallis for more than two groups (and Dunn test for pairwise comparison, p values use holm correction to account for population wide differences). χ^2 tests (global and pairwise) were used to demonstrate significant differences in the proportion of categorical variables. For repeated measures, a non-parametric Wilcoxon-rank is presented.

The predictive models used a complete-case analysis approach. The sample size was pragmatic and dictated by patient availability. Two models were developed. The first model included only symptom histories. This multinomial logistic regression predicted either a CD, UC or non-IBD diagnosis. ORs predicting both CD and UC over a reference diagnosis of 'non-IBD' were calculated. All parameters presented were modelled as factors with binary 'yes'/'no' responses ('no' given as the reference level for all) bar 'blood type' ('mixed', 'anorectal' or 'none' with 'none' as reference) and 'smoking status' ('current', 'ex', 'non' with 'non' as reference). The second model was a binomial logistic regression predicting 'IBD' versus 'non-IBD'. The same methodology was applied to the symptom profiles. Two approaches were taken to integrating FCP results. In the main model, first and second FCPs were added as covariates. To assess the performance of FCP cut offs alongside specific symptoms, an alternative approach was used. Overall FCP levels were removed as covariates. A binary 'yes'/'no' response to achieving a single FCP cut-off was used instead and then added as a factor. The quality of fit for models is presented using McFadden's PseudoR^2 and the overall model test.

RESULTS

From January 2021 to August 2023, 767 patients were seen. A final diagnosis is currently available for 762 (208 CD, 215 UC, 340 non-IBD). The overall cohort is summarised in table 1. The flow and numbers of patients included in each subsequent analysis is shown in figure 1.

Patients subsequently diagnosed with CD were seen within a median of 14 (IQR 11) days (d) of referral receipt by the inception team, while for UC this was 13d (14.75). However, delays in initial referral processing increased these values to 32.5d (36.3) and 29d (35.5) for CD and UC, respectively from the date the referral was written. At the point of first review, prior to diagnosis, the median duration of symptoms was significantly longer in CD (median 10m (months), UC 4m, non-IBD 5.5m; Kruskal-Wallis (2) 34.54 p<0.001, Dunn pairwise test; UC p_{holm}<0.001, non-IBD p_{holm}<0.001). In CD, longer duration of symptoms significantly increased the likelihood of stricturing or penetrating disease (B1 n=162 median 9m, B2/3 n=35 median 12m, Mann-Whitney U p=0.02).

Predictive value of serial FCP testing

Of 672 patients with at least one FCP result, 86% (581) had a result available at initial referral triage from primary care. Only 38% (258/672) had a paired faecal sample submitted to assess for enteric infection. Of the 422 with two FCP results, only 20% (85/422) had the second sample submitted in primary care pre-referral. Given the second sample was not submitted until review in secondary care, the delay between samples was longer than reported in

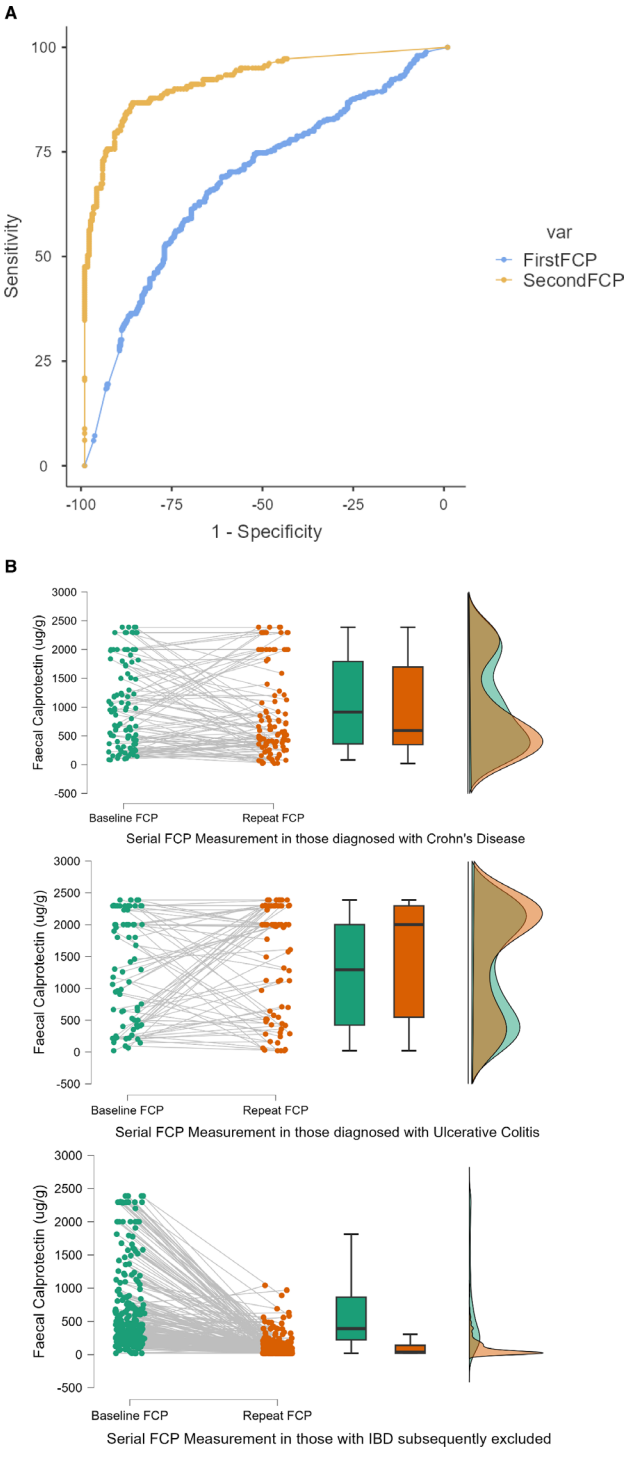

Figure 2 (A) Receiver operating characteristic curves for all first and second faecal calprotectin (FCP) results. For all first FCPs (n=672), the area under the curve (AUC) was 0.683, while for the second FCP (n=422) this was 0.923. The highest Youden's index was at a second FCP cut point of 220 µg/g (Youden's index 0.735). (B) For the cohort with two results (n=422), these plots demonstrate median FCP at referral and on clinic review prediagnosis, split by the condition subsequently diagnosed.

Table 3 Predictive indices of the ability of faecal calprotectin (FCP) to differentiate inflammatory bowel disease (IBD) from non-IBD diagnoses at different cut-offs on the basis on one or two results

FCP cut-offs	Sens (%)	Spec (%)	PPV (%)	NPV (%)	AUC
>50 µg/g	99.4	3.7	44	90	0.516
>100 µg/g	97.2	6.2	44	75	0.517
>150 µg/g	93.9	12.0	44	73	0.530
>200 µg/g	89.5	24.0	47	75	0.567
>250 µg/g	82.3	31.4	47	70	0.569
>50 µg/g×2	94.4	56.7	62	93	0.755
>100 µg/g×2	89.8	70.2	69	90	0.800
>150 µg/g×2	84.8	78.2	74	87	0.815
>200 µg/g×2	79.8	87.0	82	85	0.834
>250 µg/g×2	75.3	89.5	84	83	0.824

This data is presented only for the subgroup who provided two samples (n=422).
AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

previous studies (n=406 where confirmed interval identifiable, median 61d (IQR 63)). Table 2 compares the median baseline FCP between cohorts providing one or two results and displays the median FCP result.

The non-IBD baseline cohort providing a single sample had a significantly lower baseline FCP than those going on to provide a repeat sample, though there was no difference in either the CD or UC cohorts. For those providing a repeat FCP, a significant reduction on re-testing was only seen in the non-IBD cohort. Differences in IBD subtypes were non-significant. Consequently, many patients initially meeting typical thresholds for colonoscopy no longer required investigation following the second test. Receiver operator characteristic (ROC) curves were plotted using all available first and second FCP results. This, alongside plots of baseline and repeat results split by subsequent diagnosis, is displayed in figure 2. All plotted first FCP results carried a grouped area under the curve (AUC) of 0.683, while this increased to 0.923 for the smaller cohort of second FCP results. The optimal FCP threshold was seen at a second FCP level of 220 µg/g (sensitivity 86.7, specificity 86.8, Youden's index 0.735).

Given that most existing referral pathways use binary cut-offs, the predictive ability of several theoretical values was analysed. To ensure validity, these cut-offs were only tested in the 422 patients with two FCP results available. The results are shown in table 3.

The strongest performing cut-off was >200 µg/g×2 (AUC 0.834). However, this was associated with a fall in sensitivity to 79.8%. Applied to our cohort, this threshold missed 36 IBD diagnoses. In 16 of these, significant increases between first/second FCP measurement (median 152.5 µg/g baseline vs 380.5 µg/g repeat)

Table 4 Symptom frequencies at presentation in Crohn's, UC and grouped non-IBD diagnoses (n=711)

	Crohn's (n=199)	UC (n=207)	Non-IBD (n=305)	
	% (n of responses +ve)	% (n of responses +ve)	% (n of responses +ve)	χ^2 p value
Symptom duration >1 m	97% (193/199)	92% (190/207)	89% (279/305)	Global: 0.007
				CD vs UC: 0.005
Gastrointestinal symptoms				
Looser stools	84% (167/199)	81% (169/207)	74% (212/305)	Global: 0.01
				CD vs UC: 0.53
Nocturnal bowel opening	37% (72/194)	46% (91/197)	27% (81/303)	Global: <0.001
				CD vs UC: 0.06
Faecal urgency	63% (126/199)	82% (170/207)	55% (167/305)	Global: <0.001
				CD vs UC: <0.001
Faecal incontinence	10% (20/199)	12% (24/207)	7% (22/305)	Global: 0.22
				CD vs UC: 0.61
Constipation	23% (45/199)	17% (35/207)	34% (104/305)	Global: <0.001
				CD vs UC: 0.16
PR blood (any type)	55% (110/199)	94% (194/207)	49% (149/305)	Global: <0.001
				CD vs UC: <0.001
Type of PR blood				
Anorectal	23% (45/199)	44% (90/207)	38% (117/305)	
Mixed	33% (65/199)	50% (104/207)	11% (32/305)	Global: <0.001
None	45% (89/199)	6% (13/207)	51% (156/305)	CD vs UC: <0.001
PR mucous	16% (32/198)	37% (77/206)	28% (86/304)	Global: <0.001
				CD vs UC: <0.001
Abdominal pain	84% (168/199)	68% (141/206)	87% (266/305)	Global: <0.001
				CD vs UC: <0.001
Extra-intestinal symptoms				
Joint pain	27% (54/197)	18% (36/203)	21% (65/305)	Global: 0.06
				CD vs UC: 0.02
Fatigue	79% (155/196)	63% (126/201)	50% (151/300)	Global: <0.001
				CD vs UC: <0.001
Weight loss	53% (104/198)	36% (75/207)	23% (69/305)	Global: <0.001
				CD vs UC: 0.001
Constitutional symptoms	10% (19/198)	6% (13/206)	9% (27/305)	Global: 0.44
				CD vs UC: 0.23
Mouth ulcers	16% (31/197)	11% (23/204)	10% (31/304)	Global: 0.16
				CD vs UC: 0.20
Risk factors				
Smoking status				
Current	21% (41/199)	7% (14/204)	20% (61/304)	
Ex	18% (36/199)	27% (55/204)	16% (47/304)	Global: <0.001
Never	61% (122/199)	66% (135/204)	64% (196/304)	CD vs UC<0.001
Family history of IBD	26% (52/197)	15% (30/202)	13% (40/301)	Global: <0.001
				CD vs UC 0.005

χ^2 p value presented for global differences and variation between IBD subtypes.
CD, Crohn's disease; IBD, inflammatory bowel disease; PR, per rectal; UC, ulcerative colitis.

highlighted the need for investigation anyhow. This left 20 who could have been missed without scrutiny. Of those, 13 had Crohn's, with limited ileal disease in 9. The remaining 7 had UC, 5 of whom presented only mild proctitis. Overall, proctitis consistently associated with lower FCPs than left sided or extensive disease (first FCP, E1 n=56 median FCP 653.5 µg/g, E2 n=59 1459 µg/g, E3 n=57 1800 µg/g. Dunn pairwise test; E1 vs E2 p_{holm}0.001, E1 vs E3 p_{holm}0.001). In CD, ileal disease behaved similarly (first FCP, L1 n=79 median FCP 400 µg/g, L2 n=47 1403 µg/g, L3 n=51 1052 µg/g. L1 vs L2 $p_{holm}<0.001$, L1 vs L3 $p_{holm}<0.001$).

Table 5 Predictive ability of presenting symptoms for IBD subtypes over non-IBD diagnoses across 670 patients

Multivariate model **Prediction of Crohn's, UC or non-IBD** **Model fit: McFadden's PseudoR^2 0.253, overall model test χ^2 360 df 36 p<0.001**							
Predictor (Y-N)	**(A) Crohn's: non-IBD**			**Predictor (Y-N)**	**(B) UC: non-IBD**		
	OR	**95% CI**	**P value**		**OR**	**95% CI**	**P value**
Model intercept	0.07	0.02–0.24	<0.001	Model intercept	0.07	0.02–0.24	<0.001
Symptom duration >1 month	3.89	1.42–10.62	0.01	Symptom duration >1 month	1.83	0.79–4.27	0.15
Looser stools	1.43	0.77–2.66	0.25	Looser stools	0.38	0.20–0.77	0.01
Constipation	0.78	0.47–1.28	0.32	Constipation	0.65	0.37–1.15	0.14
Nocturnal bowel opening	0.86	0.57–1.60	0.86	Nocturnal bowel opening	1.84	1.05–3.24	0.03
PR bleeding				PR bleeding			
Anorectal—none	0.99	0.60–1.65	0.98	Anorectal—none	11.12	5.41–22.84	<0.001
Mixed—none	4.38	2.40–7.98	<0.001	Mixed—none	33.68	15.47–73.33	<0.001
Mucous	0.39	0.22–0.69	0.001	Mucous	0.83	0.50–1.39	0.49
Faecal urgency	0.84	0.50–1.42	0.51	Faecal urgency	4.37	2.25–8.48	<0.001
Incontinence	1.20	0.55–2.60	0.65	Incontinence	1.27	0.57–2.25	0.55
Weight loss	3.39	2.14–5.38	<0.001	Weight loss	2.33	1.37–4.00	0.002
Abdominal pain	0.51	0.27–0.95	0.03	Abdominal pain	0.22	0.12–0.40	<0.001
Joint pain	0.92	0.56–1.52	0.75	Joint pain	0.72	0.40–1.28	0.26
Mouth ulcers	1.13	0.60–2.12	0.72	Mouth ulcers	0.88	0.42–1.84	0.74
Fatigue	2.94	1.77–4.87	<0.001	Fatigue	1.63	0.96–2.78	0.07
Constitutional symptoms	0.80	0.38–1.71	0.56	Constitutional symptoms	0.57	0.24–1.41	0.23
Family history	2.05	1.21–3.49	<0.001	Family history	0.81	0.43–1.52	0.51
Smoking status				Smoking status			
Curren—non	1.08	0.62–1.86	0.79	Current—non	0.43	0.21–0.90	0.03
Ex—non	1.21	0.69–2.13	0.50	Ex—non	1.50	0.85–2.65	0.16

IBD, inflammatory bowel disease; OR, odds ratio; PR, per rectal; UC, ulcerative colitis.

Revisiting the clinical history: diagnostic discriminators

Symptom profiles were available for 199 CD, 207 UC and 305 non-IBD patients. The number of patients positively reporting symptoms in each criterion is shown in table 4.

The five most common symptoms in CD were abdominal pain, looser stools, fatigue, faecal urgency and per-rectal (PR) bleeding. The most common five symptoms in UC were PR bleeding, faecal urgency, looser stools, abdominal pain and fatigue. As shown in table 4, the biggest differences between CD and UC were the increased prevalence of abdominal pain and fatigue in CD, while rectal bleeding and urgency increased in UC. While common in CD, abdominal pain was also highly prevalent in the non-IBD cohort. The largest differences seen across both IBD subtypes and non-IBD, as shown in table 4, were in the increased frequency of weight loss and the passage of blood per rectum mixed in with stools in IBD. Anorectal bleeding was seen more frequently in people with non-IBD diagnoses than in patients with CD.

The first regression model, based on symptom profiles only, is presented to highlight differences between IBD subtypes and to focus on the symptoms that can separate CD from non-IBD, given the frequent delays to CD diagnosis. The overall model demonstrated adequate fit. Mixed blood in the stool was the most discriminant symptom for both CD (OR 4.38; 95% CI 2.40–7.98) and UC (OR 33.68; 15.47–73.33). Weight loss (OR 3.39; 2.14–5.38), fatigue (OR 2.94; 1.77–4.87) and family history of IBD (OR 2.05; 1.21–3.49) also associated with an increased likelihood of CD over non-IBD diagnoses. The full CD predictive values are shown in table 5A. With regards to UC, weight loss was again a significant predictor (OR 2.33; 1.37–4.00). Faecal urgency (OR 4.37; 2.25–8.48) and nocturnal bowel opening (OR 1.84; 1.05–3.24) were also associated with UC diagnoses. The full UC predictive values are shown in table 5B.

Modelling clinical history with FCP

While table 5 demonstrates the differing presentations of IBD subtypes, any referral pathway needs to be able to separate IBD from non-IBD. To ascertain the overall ability of FCP and symptoms to achieve this, a binomial model was developed using the same symptom indices. This was only applied to the 389 with two FCPs available. By adding FCP, initially modelling two samples >200 µg/g as a factor, the model fit and predictive indices improved significantly. When applied in this way, it was possible to demonstrate a pretest probability of IBD of 92.5% (95% CI 74.6–98.1) in patients with two FCPs >200 µg/g and mixed bleeding. This

Table 6 Ability to predict IBD over non-IBD: clinical history modelled with FCP as a covariate or factor at a set threshold (only for those providing two FCPs and complete symptom history; n=389)

Model A **Set FCP threshold >200 µg/g×2** **McFadden's PseudoR^2 0.511** **Overall model χ^2 311 df 19 p<0.001**				**Model B** **Baseline and repeat FCP as covariates** **McFadden's PseudoR^2 0.580** **Overall model χ^2 305 df 20 p<0.001**			
Predictor (Y-N)	**IBD: non-IBD**			**Predictor (Y-N)**	**IBD: non-IBD**		
	OR	**95% CI**	**P value**		**OR**	**95% CI**	**P value**
Model intercept	0.01	.002–0.09	<0.001	Model intercept	0.01	.001–0.12	<0.001
Symptom duration >1 month	3.43	0.84–14.00	0.09	Symptom duration >1 month	6.58	0.61–71.32	0.12
Looser stools	1.50	0.60–3.74	0.38	Looser stools	1.06	0.41–2.77	0.90
Constipation	0.70	0.33–1.51	0.36	Constipation	0.93	0.42–2.06	0.86
Nocturnal bowel opening	0.84	0.38–1.90	0.68	Nocturnal bowel opening	0.78	0.33–1.85	0.58
PR bleeding				PR bleeding			
Anorectal—none	3.01	1.37–6.61	0.01	Anorectal—none	2.54	1.11–5.80	0.03
Mixed—none	12.23	4.68–31.93	<0.001	Mixed—none	8.09	2.79–23.51	<0.001
Mucous	0.73	0.34–1.60	0.44	Mucous	0.46	0.18–1.14	0.09
Faecal urgency	1.09	0.51–1.42	0.83	Faecal urgency	1.25	0.55–2.84	0.60
Incontinence	1.02	030–3.47	0.97	Incontinence	1.57	0.47–5.22	0.46
Weight loss	3.65	1.74–7.66	<0.001	Weight loss	2.62	1.20–5.73	0.02
Abdominal pain	0.36	0.15–0.92	0.03	Abdominal pain	0.28	0.10–0.74	0.01
Joint pain	0.86	0.39–1.89	0.71	Joint pain	1.13	0.50–2.55	0.76
Mouth ulcers	1.43	0.55–3.72	0.46	Mouth ulcers	1.28	0.48–3.47	0.62
Fatigue	3.11	1.43–6.75	0.004	Fatigue	3.04	1.32–7.00	0.01
Constitutional symptoms	0.33	0.08–1.42	0.13	Constitutional symptoms	0.15	0.02–1.40	0.10
Family history	1.62	0.67–3.96	0.29	Family history	1.39	0.56–3.50	0.48
Smoking status				Smoking status			
Current—non	0.84	0.34–2.08	0.71	Current—non	1.23	0.49–3.06	0.66
Ex—non	1.11	0.47–2.64	0.81	Ex—non	0.98	0.37–2.59	0.97
Both FCP >200 µg/g	27.09	13.73–53.47	<0.001	Baseline and repeat FCPs modelled as covariates			
Overall predictive indices	Optimal cut-off value: 0.40			Overall predictive indices	Optimal cut-off value: 0.35		
	Sensitivity: 0.84				Sensitivity: 0.89		
	Specificity: 0.84				Specificity: 0.88		
	AUC: 0.929				AUC: 0.947		

*These indices outperform the overall outputs generated by the same binomial model based on symptoms alone (McFadden's PseudoR^2 0.278, overall model χ^2 148 df 18 p<0.001. Optimal cut-off value 0.43, sensitivity 0.78, specificity 0.77, AUC 0.837).
AUC, area under the curve; FCP, faecal calprotectin; IBD, inflammatory bowel disease; OR, odds ratio; PR, per rectal.

increased to 95.9% (95% CI 84.0–99.1) if weight loss is also present, or 95.6% (95% CI 83–99) if weight loss is substituted for fatigue. If bleeding is removed but weight loss and fatigue modelled together at the same FCP threshold, the probability remained high at 91.9% (95% CI 73–97.9). However, changing approach and modelling FCP values as covariates, improved model performance and fit further, suggesting this was the optimal approach. The model fit, symptom outputs and predictive values from the integrated FCP models are compared in table 6.

DISCUSSION

Our study is the first to prospectively document presenting IBD symptoms and FCP level in an IBD inception cohort. Interrogating the predictive capacity of individual symptoms in a cohort largely preselected because of FCP is highly relevant given the widespread utilisation of FCP in referral pathways. The prolonged symptom duration at CD onset is again apparent. Longer durations remain associated with higher incidences of stricturing or penetrating disease.

The most common symptoms at presentation in IBD were not always able to discriminate from non-IBD diagnoses. The strongest predictors of IBD were passing blood mixed with stools and weight loss. While abdominal pain remains highly relevant for patients with IBD, it was significantly more prevalent in non-IBD diagnoses (>50% functional).

Obtaining a second FCP result prevented patients undergoing unnecessary colonoscopy. In those being referred outside of two-week wait criteria, and in the absence of fulminant symptoms, a repeat FCP measurement is an accurate predictor of IBD. Of the traditional thresholds currently used in clinical practice, >200 µg/g on two occasions performed best

statistically but missed some mild IBD presentations. A two-sample >100 µg/g threshold to trigger outpatient clinic review with two samples above 200 µg/g to progressing directly to endoscopy would miss fewer IBD cases. However, our data would suggest a substantially higher predictive value is derived from the second FCP result. When plotted as an ROC curve a second FCP of 220 µg/g, regardless of first result, carried a higher sensitivity, specificity and AUC than a two sample >200 µg/g threshold. By combining symptoms and FCP data, this predictive ability can be enhanced. Modelling FCP cut-offs with symptoms improved AUC compared with cut-offs alone and additionally allowed the identification of high-risk symptom complexes. For example, in those with mixed blood per rectum, weight loss and an FCP >200 µg/g×2, the pretest probability of IBD was 95.7%. Clearly, many clinicians would identify the risk here without the need for predictive modelling. However, symptom complexes such as fatigue and weight loss, irrespective of bleeding history, still carried a pretest probability over 90% at this FCP level. When FCP values are modelled as covariates alongside symptom profiles, the model prediction improves further to an AUC of 0.947.

With the growing availability of home FCP testing, symptom complexes and FCP values could feed into algorithms that allow self-referral to secondary care services. The provision of simple symptom questionnaires and home delivered FCP testing could reduce strain on primary care, allow rapid identification and prioritisation of those most likely to have IBD. This aligns with but builds on the recommendations of 'Getting It Right First Time'. Exactly how such a pathway would look at this stage is not defined and would undoubtedly require significant resource, education and validation. At present, our modelling identifies discriminant factors but does not present a readymade pathway to apply at the point of referral. This is something we aim to address with a prospective validation cohort being initiated across sister hospital sites within our trust.

There are several other limitations to the modelling. Symptom histories are currently obtained from one trust and by a small number of clinicians, limiting generalisability. The histories are current obtained via the clinician, and self-reporting of symptoms has not been validated. Furthermore, patient perceptions of this process have not been formally evaluated.

Though throughput has been significant for a single centre, our cohort remains small to base this modelling on.[26] Though this reflects clinical practice, the failure to apply inclusion criteria strictly allows for methodological inconsistency and limits reproducibility. Not mandating an FCP result for entry into the clinical pathway, alongside missing data in 41 patient histories results in patients being removed from the modelling, which includes 670 patients for symptoms alone but drops to 389 patients when two FCPs are required.

When evaluating FCP thresholds, most of our repeat FCP samples were not submitted until the time of initial review in secondary care. The time interval between results is longer than advocated in prior studies and carried a broad spread. With this in mind, it is not possible to robustly comment on the optimal time interval to leave between initial and repeat FCP testing within primary care. Furthermore, the return of stool samples by patients, even when sent in advance, was inconsistent. It is known that patient perspectives around stool samples can be prohibitive.[27] It is therefore an inherent risk, given that two samples were not mandated, that those with more fulminant symptoms are more likely to progress through further investigations despite not returning two FCPs. Of those providing one FCP prior to diagnosis, 67.6% went on to have IBD, while this fell to 42.6% for those providing two FCPs. The biggest drop off (19.8%) was seen in UC. Despite this, there was no significant difference in baseline FCP between the IBD cohorts providing either one or two results. Non-IBD patients providing two FCPs had a higher baseline result than those providing one.

This work represents the first in-depth characterisation of presenting symptoms combined with FCP level in an IBD inception cohort. Through a large validation cohort, we will apply the discriminant signals described here prospectively.

Twitter Peter Rimmer @PeterRimmerIBD, Asif Iqbal @DrAJIqbal, Rachel Cooney @rachelcooney7 and Karl Hazel @karl.h87

Contributors PR collected the majority of the histories, undertook the statistical analysis and wrote the text. JC undertook the second largest portion of the symptom histories and collected associated diagnosis data. JH and ML provided support to the delivery of the clinical pathway. ST, AI and DR-K provided peer review and methodological support. NS and KH undertook patient histories and helped deliver the pathway. RC provided peer review and support with writing the manuscript. IC provided peer review and TD provided support to the statistical methodology. MNQ and THI undertook patient histories, oversaw delivery of the pathway and provided support and feedback to the writing of the manuscript. THI is the overall guarantor of the data presented in this manuscript.

Funding This work has been supported by funding from F. Hoffman La Roche, the Birmingham NIHR Biomedical Research Centre Grant 2023 and a GUTS UK Trainee Research Award (2021 - TRA2021_02).

Competing interests PR, THI, AI and MNQ have received research funding from F. Hoffman La Roche. PR, THI, RC and MNQ have received honoraria from Janssen and Bristol Myers Squibb. THI has received honoraria from Pharmacosmos. RC has received fees from Abbvie, Galapagos and Celltron.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and the data presented here is all standard of care clinical data (clinical history and FCP) collected as part of routine assessment of IBD patients and did not in itself require ethics approval. However, around half of these patients were recruited to active research with additional biological samples obtained. The IRAS for

this was 287279 with approval from London - Bloomsbury Research Ethics Committee (REC reference 21/PR/0515). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Peter Rimmer http://orcid.org/0000-0002-0528-6154
Rachel Cooney http://orcid.org/0000-0003-3710-157X

REFERENCES

1 Jayasooriya N, Baillie S, Blackwell J, *et al*. Systematic review with meta-analysis: time to diagnosis and the impact of delayed diagnosis on clinical outcomes in inflammatory bowel disease. *Aliment Pharmacol Ther* 2023;57:635–52.
2 Splinter MJ, Velek P, Ikram MK, *et al*. Prevalence and determinants of Healthcare avoidance during the COVID-19 pandemic: A population-based cross-sectional study. *PLoS Med* 2021;18:e1003854.
3 Czeisler MÉ, Marynak K, Clarke KEN, *et al*. Delay or avoidance of medical care because of COVID-19–related concerns — United States. *MMWR Morb Mortal Wkly Rep* 2020;69:1250–7.
4 Blackwell J, Saxena S, Jayasooriya N, *et al*. Prevalence and duration of gastrointestinal symptoms before diagnosis of inflammatory bowel disease and predictors of timely specialist review: A population-based study. *J Crohns Colitis* 2021;15:203–11.
5 Pope C. NHS waiting times: a government pledge. *BMJ* 2023;380:71.
6 King D, Reulen RC, Thomas T, *et al*. Changing patterns in the epidemiology and outcomes of inflammatory bowel disease in the United kingdom: 2000-2018. *Aliment Pharmacol Ther* 2020;51:922–34.
7 Lamb CA, Kennedy NA, Raine T, *et al*. British society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* 2019;68(Suppl 3):s1–106.
8 NHS England. Getting it right the first time; IBD pathway Gastroenterology version 1. 2022. Available: https://gettingitrightfirsttime.co.uk/wp-content/uploads/2022/01/Gastroenterology_2022-01-18_Pathway_IBD.pdf [Accessed 30 Jul 2023].
9 IBD UK. Crohn's and colitis care in the UK: the hidden cost and a vision for change. 2021. Available: https://s3.eu-west-2.amazonaws.com/files.ibduk.org/documents/CROJ8096-IBD-National-Report-WEB-210427-2.pdf [Accessed 14 Aug 2023].
10 Crohns and Colitis UK. Cut the crap. Available: https://crohnsandcolitis.org.uk/our-work/campaigns/cut-the-crap [Accessed 14 Aug 2023].
11 Quraishi MN, Dobson E, Ainley R, *et al*. Establishing key performance indicators for inflammatory bowel disease in the UK. *Frontline Gastroenterol* 2023;14:407–14.
12 Hicks A, Hamlin PJ, Selinger CP. The Association of introducing a Faecal Calprotectin testing pathway for suspected inflammatory bowel disease in primary care and time to diagnosis or treatment. *Inflamm Intest Dis* 2020;5:191–9.
13 D'Amico F, Nancey S, Danese S, *et al*. A practical guide for Faecal Calprotectin measurement: myths and realities. *J Crohns Colitis* 2021;15:152–61.
14 Waugh N, Cummins E, Royle P, *et al*. Faecal Calprotectin testing for differentiating amongst inflammatory and non-inflammatory bowel diseases: systematic review and economic evaluation. *Health Technol Assess* 2013;17:xv–xix,
15 Turvill J, Turnock D, Holmes H, *et al*. Evaluation of the clinical and cost-effectiveness of the York Faecal Calprotectin care pathway. *Frontline Gastroenterol* 2018;9:285–94.
16 Perler BK, Ungaro R, Baird G, *et al*. Presenting symptoms in inflammatory bowel disease: descriptive analysis of a community-based inception cohort. *BMC Gastroenterol* 2019;19:47.
17 Stapley SA, Rubin GP, Alsina D, *et al*. Clinical features of bowel disease in patients aged <50 years in primary care: a large case-control study. *Br J Gen Pract* 2017;67:e336–44.
18 Maaser C, Sturm A, Vavricka SR, *et al*. ECCO-ESGAR guideline for diagnostic assessment in IBD part 1: initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis* 2019;13:144–64.
19 The jamovi project. Jamovi (version 2.3). 2022. Available: https://www.jamovi.org [Accessed 4 Jan 2024].
20 R Core Team. R: A Language and environment for statistical computing, . 2021Available: https://cran.r-project.org [Accessed 4 Jan 2024].
21 Lenth R. Emmeans: estimated marginal means, Aka least-squares means. R package; 2020. Available: https://cran.r-project.org/package=emmeans [Accessed 4 Jan 2024].
22 Ripley B, Venables W. R package; nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models, . 2016Available: https://cran.r-project.org/package=nnet [Accessed 4 Jan 2024].
23 Thiele C. Cutpointr: determine and evaluate optimal Cutpoints in binary classification tasks. R package; 2019. Available: https://cran.r-project.org/package=cutpointr [Accessed 4 Jan 2024].
24 Friesen L, Andersen E. Outcomes of collaborative and Interdisciplinary palliative education for health care assistants: A qualitative Metasummary. *J Nurs Manag* 2019;27:461–81.
25 JASP TeamJASP (Version 0.18.1)[Computer software).
26 Riley RD, Ensor J, Snell KIE, *et al*. Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020;368:m441.
27 Lecky DM, Hawking MKD, McNulty CAM, *et al*. Patients' perspectives on providing a stool sample to their GP: a qualitative study. *Br J Gen Pract* 2014;64:e684–93.