Diagnostic accuracy of magnetic resonance enterography and small bowel ultrasound for the extent and activity of newly diagnosed and relapsed Crohn's disease (METRIC): a multicentre trial

Stuart A Taylor, Susan Mallett, Gauraang Bhatnagar, Rachel Baldwin-Cleland, Stuart Bloom, Arun Gupta, Peter J Hamlin, Ailsa Antony Higginson, Ilan Jacobs, Sara McCartney, Anne Miles, Charles D Murray, Andrew A Plumb, Richard C Pollok, Shonit Pu Manuel Rodriguez-Justo, Zainib Shabir, Andrew Slater, Damian Tolan, Simon Travis, Alastair Windsor, Peter Wylie, Ian Zealley on behalf of the METRIC study investigators*

Lancet Gastroenterol Hepatol 2018; 3: 548–58

Published Online June 15, 2018

http://dx.doi.org/10.1016/S2468-1253(18)30161-4

See Comment page 521

*Investigators listed in the appendix

Centre for Medicj Imaging, University College LooTon (UCL), LooTon, UK (Prof S A Taylor FRCR, G Bhatnagar FRCR, A A Plumb FRCR, S Punwani FRCR, Prof S Ha 0igan FMedSci); Institute of Applied Health Research, National Institute of

Summary

Background Magnetic resonance enterography (MRE) and ultrasound are used to image Crohn's disease, but their comparative accuracy for assessing disease extent and activity is not known with certainty. Therefore, we did a multicentre trial to address this issue.

Methods We recruited patients from eight UK hospitals. Eligible patients were 16 years or older, with newly diagnosed Crohn's disease or with established disease and suspected relapse. Consecutive patients had MRE and ultrasound in addition to standard investigations. Discrepancy between MRE and ultrasound for the presence of small bowel disease triggered an additional investigation, if not already available. The primary outcome was di erence in per-patient sensitivity for small bowel disease extent (correct identi cation and segmental localisation) against a construct reference standard (panel diagnosis). This trial is registered with the International Standard Randomised Controlled Trial, number ISRCTN03982913, and has been completed.

Findings 284 patients completed the trial (133 in the newly diagnosed group, 151 in the relapse group). Based on the reference standard, 233 (82%) patients had small bowel Crohn's disease. The sensitivity of MRE for small bowel disease extent (80% [95% CI 72–86]) and presence (97% [91–99]) were signi cantly greater than that of ultrasound (70% [62–78] for disease extent, 92% [84–96] for disease presence); a 10% (95% CI 1–18; p=0·027) di erence for extent, and 5% (1–9; p=0·025) di erence for presence. The speci city of MRE for small bowel disease extent (95% [85–98]) was signi cantly greater than that of ultrasound (81% [64–91]), a di erence of 14% (1–27; p=0·039). The speci city for small bowel disease presence was 96% (95% CI 86–99) with MRE and 84% (65–94) with ultrasound (di erence 12% [0–25]; p=0·054). There were no serious adverse events.

Interpretation Both MRE and ultrasound have high sensitivity for detecting small bowel disease presence and both are valid rst-line investigations, and viable alternatives to ileocolonoscopy. However, in a national health service setting, MRE is generally the preferred radiological investigation when available because its sensitivity and speci city exceed ultrasound signi cantly.

Funding National Institute of Health and Research Health Technology Assessment.

Introduction

Small bowel imaging is fundamental for comprehensive phenotyping of Crohn's disease and essential to direct therapeutic strategy[1]. Barium uoroscopy has long been the bedrock of small bowel investigation, providing detailed mucosal assessment[2]. However, in the past 5–10 years enthusiasm has dwindled, and barium uoroscopy is being increasingly replaced by cross-sectional imaging, namely computed tomography enterography (CTE), magnetic resonance enterography (MRE), and ultrasound. Advocates of cross-sectional imaging stress that these techniques assess the bowel wall and beyond, complementing endoscopic visualisation. As barium uoroscopy is ab9oToned, dissemination of the various cross-sectional imaging technologies has been relatively uncontrolled, despite a paucity of supportive data from methodologicj 0y sound prospective multicentre studies. This scarcity of robust evidence is concerning given the pivotal role assumed by small bowel imaging over the lifetime of patients with Crohn's disease.

Of the available modalities, MRE and ultrasound are preferred[3] since they avoid irradiating genera 0y young patients who require repeat imaging[4]. Enteric ultrasound is longer established[5], requires little patient preparation, and the technology is widely available. However,

Department of Psychological Sciences, Birkbeck University of London, London, UK (A Miles PhD); Department of Gastroenterology and Endoscopy, Royal Free London NHS Foundation Trust, London, UK (C D Murray PhD); Department of Gastroenterology, St George's Hospital, London, UK (R C Pollok FRCP); Department of Histopathology, University College Hospital, London, UK (M Rodriguez-Justo FRCPath); Comprehensive Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, Holborn, London, UK (Z Shabir MSc); Department of Radiology, Oxford University Hospitals NHS Trust, Oxford, UK (A Slater FRCR); Department of Radiology, St James's University Hospital, Leeds Teaching Hospitals NHS Trust, Leeds, UK (D Tolan FRCR); Translational Gastroenterology Unit, Oxford University Hospitals, Oxford, UK (Prof S Travis DPhil); Department of Surgery, University College Hospital, London, UK (A Windsor FRCS); Department of Radiology, Royal Free London NHS Foundation Trust, London, UK (P Wylie FRCR); and Department of Radiology, Ninewells Hospital, Dundee, UK (I Zealley FRCR)

Correspondence to:
Prof Stuart A Taylor, Centre for Medical Imaging, UCL, London W1W 7TS, UK
stuart.taylor1@nhs.net

questions remain over accuracy, particularly in the proximal bowel and deep pelvis,[5] and perceived interobserver variability.[7] Conversely, MRE is a newer innovation,[8] requires oral contrast and access to advanced technology imaging platforms, which are comparatively restricted in many health-care settings.

Although meta-analyses[6,9–20] suggest that MRE and ultrasound have similar accuracy for diagnosing and staging Crohn's disease, the primary literature is of questionable quality. Most studies[17,20,21] are small and

representative of institutions likely to implement MRE and ultrasound for patient management (appendix p 1). All sites had an established in ammatory bowel disease service and were already doing MRE and ultrasound as part of usual clinical practice.

Patients were eligible for the newly diagnosed group if they had been diagnosed with Crohn's disease in the 3 months preceding recruitment on the basis of conventional diagnostic criteria, or when Crohn's disease was strongly suspected on the basis of imaging or endoscopic features but pending nal diagnosis. Eligible patients had already had colonoscopy or were awaiting it at recruitment. Patients in whom the nal diagnosis was not Crohn's disease were subsequently excluded.

Patients were eligible for the suspected luminal relapse group if they had established Crohn's disease (>3 months) and there was a strong clinical suspicion of luminal relapse based on either objective markers of in ammatory activity (C-reactive protein [CRP] concentration >8 mg/L or faecal calprotectin concentration >100 µg/g), symptoms suggestive of luminal stenosis (including obstructive symptoms, such as colicky abdominal pain, vomiting), or abnormal endoscopy. Eligible patients for both groups were aged 16 years or older. Patients were ineligible if they were pregnant or if they had contra indications to MRI. Those with psychiatric or other disorders who were unable to give informed consent were also excluded, as were those with evidence of severe or uncontrolled systemicdisease. Patients in the newly diagnosed group were excluded if they had surgical resection before colonoscopy.

Members of the local research team identi ed suitable patients from outpatient clinics, multidisciplinary team meetings, and inpatient wards, and they took informed consent from consecutive, unselected, eligible patients. A screening log detailed all approached patients and reasons for non-participation, if applicable. We collated patient demographics and clinical data (eg, age, sex, Montreal classi cation [relapse group only], disease or symptom duration, medication, and surgical history).

Procedures

Patients had MRE and ultrasound in addition to any other enteric imaging or endoscopic investigations done during their usual clinical care.

MRE was done according to local standard clinical protocols (including the choice of oral contrast agent) on either 1·5 T or 3 T MRI platforms. We acquired a minimum dataset of sequences (appendix p 2). Ultrasound was done by radiologists or sonographers using standard platforms and both curvilinear and high-resolution probes, without oral or intravenous contrast agents (appendix p 3).

index at recruitment and repeated between 10 and 20 weeks later. We asked patients if they found MRE and ultrasound acceptable and which test attribute they considered to be the most important.

We used the construct reference standard model (panel diagnosis), incorporating the concept of clinical test validation—ie, whether test results are meaningful in practice.[23] Specifically, we followed patients' clinical course for 6 months to assess the effect of MRE and ultrasound findings on clinical decision making and patient outcomes. Each recruitment site convened a series of consensus panels consisting of at least one local gastroenterologist and two radiologists (one local and one from another site); a histopathologist was available if required and a member of the trial management group attended to ensure uniformity of process. For each patient, the panel considered the images and results of all small bowel investigations (including MRE and ultrasound) and all additional information accrued over the follow-up period, including endoscopies, surgical findings, histopathology, Harvey Bradshaw index, CRP concentration, calprotectin concentration (and changes thereof), and clinical course. The panel recorded its opinion as to whether small bowel or colonic Crohn's disease was present, and, if so, whether disease was active. All panel decisions were recorded as present or absent, active or inactive, with no option of an indeterminate outcome. Disease could only be categorised as active if at least one objective marker was present (ulceration as seen at endoscopy, measured CRP concentration >8 mg/L, measured calprotectin concentration >250 µg/g or histopathological evidence of acute inflammation based on a biopsy sample or surgery within 2 months of trial imaging).

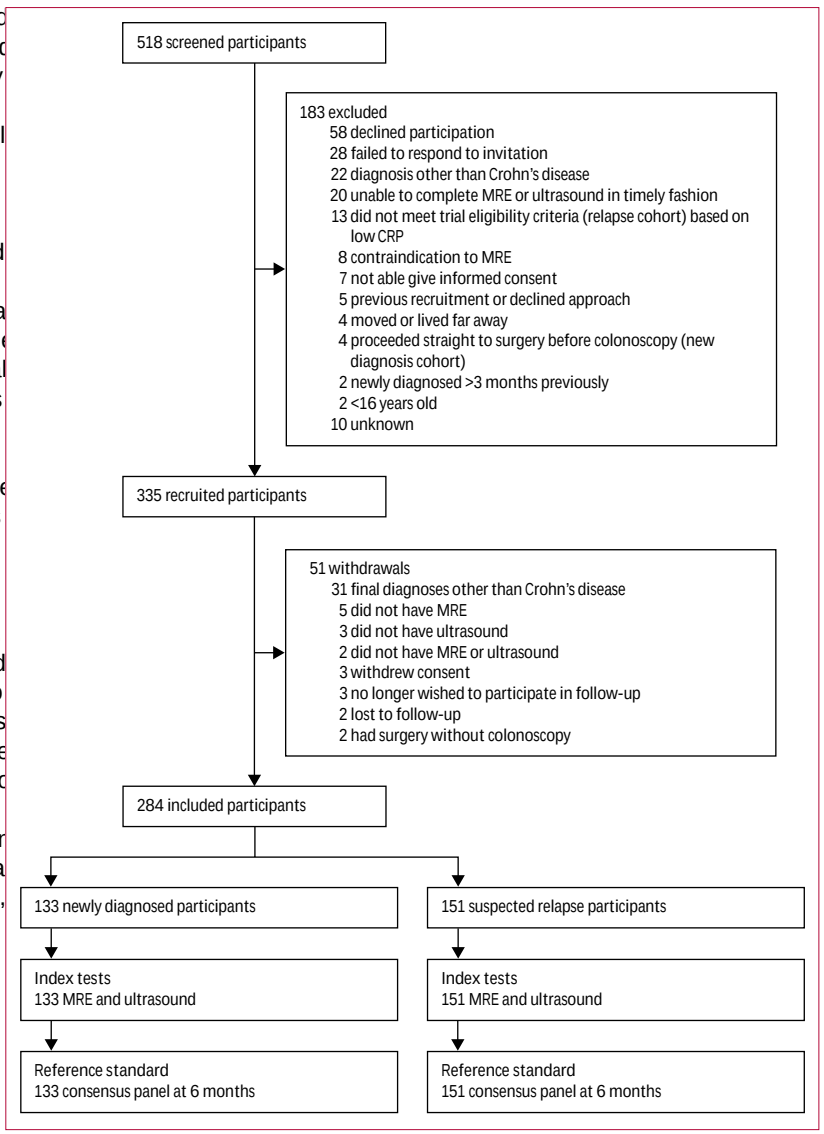

Figure : Trial profile
CRP=C-reactive protein. MRE=magnetic resonance enterography.

Outcomes

The primary outcome was the per-patient difference in sensitivity between MRE and ultrasound for correct identification and localisation of small bowel Crohn's disease, irrespective of activity—ie, the extent of small bowel disease. To be truly positive for disease extent, the index test had to correctly locate the presence and segmental location of disease (terminal ileum, ileum, jejunum, or duodenum). Secondary outcomes reported here were specificity for disease extent, sensitivity and specificity for small bowel disease presence, the difference in per patient sensitivity and specificity for colonic disease presence and extent, and identification of active disease and comparative patient experience. Secondary outcomes also included comparative impact of MRE and ultrasound on clinician diagnostic confidence for presence of Crohn's disease and their influence on management, cost-effectiveness of MRE and ultrasound (compared to each other), diagnostic impact of novel MRE sequences (eg, diffusion-weighted imaging), influence of sequence selection on MRE diagnostic accuracy, diagnostic accuracy of small intestine contrast enhanced ultrasonography (SICUS) compared with standard ultrasound, influence of oral contrast agent and ingested volume on small bowel distension and patient experience during MRE, and interobserver variation, which will be reported elsewhere. We reported most outcomes for the newly diagnosed and suspected luminal relapse groups individually, and for the terminal ileum and colon using colonoscopy as a standalone reference standard (when available) because of its robustness for identifying disease.

We prespecified all outcomes in the protocol[21] except accuracy for individual small bowel segments (duodenum, jejunum, ileum), accuracy for disease presence and extent in the colon, and per-patient disease activity

of 284, 133 were included in the newly diagnosed group and 151 in the relapse group (figure 1; table 1), including 154 (54%) women. Based on the reference standard, 233 (82%) of 284 patients had small bowel Crohn's disease (thereby meeting sample size stipulations), which was active in 209 (90%) patients (table 2). 112 (45%) of 284 patients had colonic disease, which was active in 126 (98%) patients. No data were missing for per-patient diagnosis of disease presence or disease extent, for the reference standard, MRE, or ultrasound.

In 53 patients (24 from the newly diagnosed group and 29 from the relapse group), MRE and ultrasound were discrepant for small bowel disease presence or location, of whom 48 (91%) patients had an additional small bowel imaging test available to the consensus panel. The range of imaging, endoscopic, and biochemical data available to the consensus panels is shown in the appendix (p 5).

The sensitivity of MRE for the extent of small bowel disease (ie, presence and correct segmental location) was 80% (95% CI 72–86) compared with 70% (62–78)

MRE and 92% sensitivity for ultrasound. Barium fluoroscopy has long been advocated as a sensitive test for mucosal disease inaccessible to endoscopy, although its support is limited to a handful of small studies[3] and

the same patients is advocated as the optimal method for diagnostic accuracy studies[22] because differences are attributable directly to the tests and not to differences between participants or study methods. Such head-to-head comparisons are rare in the medical literature.[15] Reference standards might also be applied inconsistently, with endoscopy, surgery, and imaging all variably employed. For example, in a comparative study with ultrasound, Castiglione and colleagues[38] used MRE without any additional reference standard in many recruits, which introduces the potential for incorporation bias.

We used the construct reference standard model (panel diagnosis), which incorporates multiple data sources with clinical outcome.[23] Although such an approach does have limitations, including potential panel bias, it is considered a very robust method for diagnostic accuracy studies in which a single external reference standard is elusive.[23] To reduce incorporation bias, patients without supplementary small bowel imaging had a third small bowel investigation whenever discrepancy between MRE and ultrasound arose. Notably, when our analysis was limited to an ileocolonoscopic reference standard, any differences in accuracy between MRE and ultrasound closely mirrored those found using the consensus panel reference.

We recruited approximately equally from two patient groups: newly diagnosed Crohn's disease and established disease with relapse. Both groups are clinically distinct and important, and might manifest with differing disease phenotypes; prevalence of stricturing and penetrating disease increases with time.[29] Noting that the METRIC trial was not powered to detect differences between these two patient groups, we found that sensitivity for small bowel disease was similar, although specificity tended to be lower in patients in the relapse group. Conversely, sensitivity for colonic disease was higher in the relapse group, but was still poor for colonic disease extent (about 30%).

In newly diagnosed patients, ultrasound achieved significantly greater sensitivity for colonic disease than MRE (67% *vs* 47%). Optimised colonic assessment with MRE requires purgation and fluid distension,[30] which are both omitted from routine MRE protocols; however, ultrasound generally relies on assessing the manually compressed uncleansed colon wall. Accuracy for both techniques in the colon still falls short of colonoscopy, and accuracy with MRE is somewhat lower than

activity of small bowel Crohn's disease in newly diagnosed patients and those who have relapsed, and both tests are valid rstline investigations. In an NHS setting, MRE is generally the preferred radiological investigation when available because its sensitivity and speci city exceed ultrasound signi cantly.

Contributors

All authors made substantial contributions to the conception or design of the work, drafted the work, or revised it critically for important intellectual content, agree to be accountable for all aspects of the work, and gave nal approval of the version to be published. SAT and GB contributed to the literature search, data collection, clinical studies, and patient recruitment. RB-C, AM, and ZS contributed to data collection. SAT, LQ, SMa, and SH contributed to data interpretation. RB-C, SB, SMc, AG, PJH, AH, ALH, CDM, AAP, RCP, SP, MR-J, ZS, AS, DT, ST, AW, PW, and IZ contributed to clinical studies and patient recruitment. LQ and SMa contributed to the statistical analysis. IJ acted as a public and patient representative. SAT, SMa, and SH wrote the initial manuscript draft. SAT is the study guarantor.

Declaration of interests

SAT, SH, ST, SMa, SB, DT, PJH, AG, ALH, AS, SMc, IZ, AH, AW, IJ, MR-J, SP, PW, and CDM report research grant support from the funder (National Institute of Health and Research Health Technology Assessment [NIHR HTA]) to do the study. SAT reports personal fees from Robarts plc. ST reports directorships of charities IBD 2020 (UK 09762150), Crohn's Colitis Cure (ABN 85 154 588 717), Truelove Foundation (UK 11056711); grants and research support from AbbVie, International Organization for the Study of In ammatory Bowel Diseases, Lilly, Union Chimique Belge (UCB), Vifor, and Norman Collisson Foundation; consulting Fees from AbbVie, Centocor, Schering-Plough, Bristol-Myers Squibb, Chemocentryx, Cosmo, Elan Pharma Inc, Genentech, Giuliani SpA, Merck, Takeda, Otsuka Pharmaceuticals, PDL Biopharma, P zer, Shire, Glenmark Pharma, Synthon Pharma, Glenmark Pharma, NPS Pharmaceuticals, Lilly, Warner Chilcott, Proximagen, VHsquared, Topivert, Ferring, Celgene, GlaxoSmithKline, Takeda, Amgen, Biogen, Enterome, Immunocore, Immunometabolism, Bioclinica, Boehringer Ingelheim, Gilead, Grunenthal, Janssen, Novartis, Celgene, Receptos, PharmOlam, Sigmoid Pharma, and Theravance; expert testimony for santarus Inc, Cosmo Technologies, and Tillotts Pharma; grants from Ferring, AbbVie, Schering-Plough, Merck Sharpe and Dhome (MSD), Proctet and Gamble, Warner Chilcott, International Organisation of In ammatory Bowel Disease, Lilly, UCB, Vifor, and Norman Collisson Foundation; lecture fees for AbbVie, Schering Ploufh, Centocor, Merck, Given Imaging, UCB Pharma, Ferring Pharmaceuticals, Tillotts Labs, Shire, Sano Aventis, Vifor, Takeda, Amgen, and Biogen; payment for manuscript preparation for Ferring; royalties from Wiley Blackwell, Elsevier, Oxford University Press; payment for development of educational presentations for Abbott Labs, Procter and Gamble, and Warner Chilcott; travel or accommodation expenses from AbbVie, United European Gastroenterology, and Shire; research nurse salary support from Schering Plough, Procter and Gamble, MSD, and Vifor; and unrestricted education grants from Abbott laboratories, Procter and Gamble, PDL BioPharma, Takeda, and International Consortium for Health Outcomes Measurement. ALH reports personal fees from AbbVie, Atlantic, Bristol-Myers Squibb, Celltrion, Falk, Ferring, Janssen, MSD, Napp Pharmaceuticals, P zer, Pharmacosmos, Shire, and Takeda; and non- nancial support from Genentech. AW reports grants from Allergan; personal fees from Allergan, and Cook; and grants and personal fees from Bard. AAP reports grants from NIHR HTA; grants from NIHR Fellowships programme; and personal fees from Acelity, Actavis, Dr Falk, Janssen-Cilag, and Takeda. IJ reports shared ownership in General Electric and sells GE equipment. CDM reports personal fees from Abbvie, MSD, and Janssen. AH reports personal fees from Toshiba. SH reports non- nancial support from iCAD. SAT and ST are NIHR senior investigators. All other authors declare no competing interests. This project was funded by the NIHR HTA programme (project number 11/23/01) and will be published in full in Health Technology Assessment. The project is supported by researchers at the NIHR University College London Hospitals Biomedical Research Centre, and NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. This report presents independent research commissioned by NIHR. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily re ect those of the NHS, NIHR, NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC), HTA programme, or the Department of Health. The trial steering committee and independent data monitoring committee met at least annually and included Vicky Goh (chair), Andrea Marshall (statistician), Ilan Jacobs (patient representative), James Lindsay (subject expert), Tim Orchard (chair), Doh-Mu UK (subject expert), and Chris Rogers (statistician).

References

1

19