Clinical Utility of small bowel ultrasound assessment of Crohn's Disease in adults: A systematic scoping review

Introduction

Due of the vastness of the existing evidence and the main aim of this review being to identify broad generalisable factors related to the application of SBUS into clinical practice, it was decided that a scoping review, rather than a systematic literature review, was more appropriate.[93–95] The objective of this scoping review were to systematically scope the literature on SBUS use in SBCD, identify specific characteristics of clinical utility and generate a comprehensive list of factors relating to the current understanding of the clinical utility of SBUS for patients with SBCD.

For this scoping review the issues identified and extracted were coded and then grouped into themes in relation to the factors of clinical utility presented by Smart[96] (Figure 1).

Clinical Utility

Like any healthcare technology or intervention, diagnostic tests such as SBUS should be thoroughly evaluated before their introduction into routine practice. Increasingly, decision makers and other users of diagnostic tests request more than simple measures of a test's analytical or technical performance and diagnostic accuracy; preferring to see evaluation of the intervention lead to health benefits, and assessment of the extent to which interventions can be applied successfully and cost effectively.[98]

Clinical utility can be described as a multi-dimensional judgement about the usefulness, benefits, and drawbacks of an intervention. A judgement about the clinical utility of a new technology or technique involves asking whether the innovation is appropriate, accessible, practicable, and acceptable for the purposes of the task intended.[96,99,100]

For diagnostic tests such as US, accurately identifying patients with suspected disease is a necessary condition for clinical utility[101], this has previously been addressed by the METRIC study.[75] However, identification of disease does not equate to health benefit; clinical context, severity, clinical consequence and management strategies may vary across cases. The type of evidence of clinical utility must be flexible to accommodate a wide range of questions about the potential benefit, and drawbacks, of interventions.[102]

Concerns about effectiveness and economics are also joined by matters of work practice when considering clinical utility. Work practice includes everyday matters that may affect feasibility of an intervention. HCPs may question how an innovation 'fits' into the existing treatment pathway. In most cases, improvements in health outcomes from diagnostic testing will be generated by the way test results are used to guide downstream management. It is possible that a test has clinical utility without improving the primary health outcomes targeted by subsequent treatment. If the introduction of a test leads to health outcomes comparable to those obtained with current standard of care practice, but these outcomes are achieved in a simpler way, there may be clinical utility.[98]

Clinical utility is frequently a matter of judgement depending on a stakeholders perspective of the supporting evidence.[102] The model of dimensions of clinical utility presented by Smart[96] (Table 2.1) encompasses elements of work practice alongside other factors such as economic considerations, stakeholder acceptability and future planning for interventions and services.

Methods

A preliminary search of MEDLINE, the Cochrane Database of Systematic Reviews and JBI Evidence Synthesis was conducted and no current or underway systematic reviews or scoping reviews on the same topic were identified at that time. Methods for this study were

developed based on Arksey and O'Malley's[94] scoping review methodology, and Levac et al [103] methodological enhancement. Presentation of literature searches, literature selection, PRISMA-SCR[104] diagram and data extraction follows JBI guidance for scoping reviews.[104,105]

Through consultation with the supervisory research team and preliminary searches, the overall main research question developed was defined as: "What evidence is currently available on the clinical utility of US for the diagnosis and management of SBCD?" The question was selected based on the scope of the primary research that will follow on from this review.

Inclusion criteria:

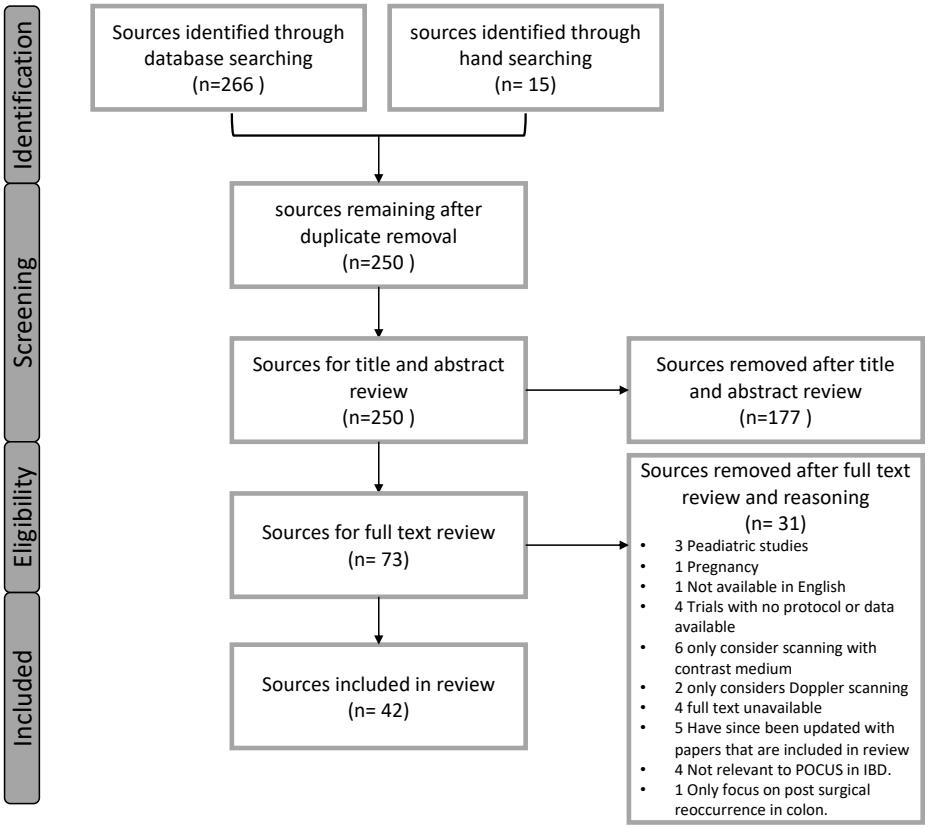
The literature search included any type of study design, including primary research, systematic reviews and meta-analysis. In addition, nonindexed and grey literature were also considered. Searches of electronic databases of the published literature included: MEDLINE, EMBASE, the Cochrane Library, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and PsycINFO. Searches were also conducted of clinicaltrial.gov for current clinical trials, 'TRIP' and Epistemonikos. Reference lists of included studies were hand searched to identify additional sources of relevance.[106,107]

Terms were searched for as both keywords in the title and/or abstract and subject headings (MeSH) as appropriate. Search terms (Table 2.2) were determined through consideration of previously reviewed literature and preliminary searches of Google Scholar. The Boolean operator 'OR' was used within each facet to maximise the searches, with the operator 'AND' used between facets to combine terms, truncation of terms was used to be as inclusive as possible during searches. Searches were performed with the 'suggested search terms' and 'explode' selection, and restricted to 'human', 'adult' and 'English language' publications. No date limits were applied to be as inclusive as possible.

Table 1: Key search terms		
Crohn's Disease (MeSH)	Small Bowel	Ultrasound (MeSH)
Crohn's Disease	ileal	Ultrasound
Crohn's	Ileum	US
CD	Ileitis	Sonography
Crohn*		Echography
Inflammatory bowel disease		Point of care ultrasound
IBD		POCUS
		ultrasonography

The review process consisted of two levels of screening: (1) title and abstract review and (2) full-text review. For the first level of screening, two investigators (Shellie Jean Radford (SJR) and Dr Gordon W. Moran (GWM, Primary Supervisor) independently screened the title and abstract of all retrieved citations against inclusion criteria. No formal quality appraisal process was undertaken; however, quality of sources was considered when reporting findings. Secondly, the two investigators each independently assessed the full-text articles to determine if they met inclusion criteria. There were no disagreements about study eligibility at the full-text review stage that required discussion with a third investigator. Reasons for exclusion of full text sources that were recorded and reported in the PRISMA[104] flow diagram (Figure 2). A Narrative synthesis was conducted in order to explore relationships within and across the included sources.

Figure 2: PRISMA[104] flow diagram



Results

The review included 42 sources (supplementary material) consisting of nine literature reviews, six clinical practice reviews (one including a case report), four clinical practice guidelines, 17 cross-sectional research studies of various designs, one case series, one abstract, two cohort studies, one national survey and one Delphi survey. Sources were from the UK, Europe (Italy, Germany, France, Spain, Hungary, Denmark, Netherlands), Japan, Canada and the USA.

A common statement across 24 sources was that US is non-invasive test that is acceptable and well tolerated by patients, is safe and is inexpensive.[64,74,78,87,108–127]

In central Europe, Canada and some parts of the USA, US is widely used, usually performed by gastroenterologists, GI radiologists or dedicated GI sonographers. US is often performed by gastroenterologists in central Europe and Canada.[127] Outside of the UK, the widespread availability, improved quality and more affordable technology and the increasing expertise of practitioners has boosted the uptake and role of US in assessing patients with CD.[75,122,125,128] The benefits of US being performed by a member of the clinical IBD team are increased capacity for real time interpretation of findings of the clinical questions being posed, expediting decisions concerning disease management and strengthening the rapport between medical teams and the patients in their care.[74,86,121,126,127]

US and MRE both achieve high diagnostic accuracy for the extent and activity of SBCD in newly diagnosed and relapsed CD.[124] A non-inferiority diagnostic study of an adult population of 249 patients with suspected CD, where 120 patients (48%) later had confirmed CD, indicated that US and MRE had comparable diagnostic accuracy in CD detection in terms of sensitivity (94% vs 96%) and specificity (97% vs 94%).[129]

The METRIC study, a prospective UK multicentre trial enrolling 284 CD patients, showed that both US and MRE had a diagnostic accuracy above 90% for detecting SBCD. Sensitivity of US for small bowel disease presence and extent were 92% and 70% respectively.[75] The study found substantial agreement for the presence of SBCD in newly diagnosed patients, and patient with suspected relapse. There was substantial sonographic agreement for the presence of SBCD, both in newly diagnosed and relapsed disease.[76] Agreement for SBCD extent was inferior to that of presence alone; this is in contrast to previous work by Parente et al[130], who reported near perfect agreement for segmental localisation between two experienced sonographers. US has also been shown to have high sensitivity and specificity

in the detection of CD complications such as stenoses, fistulas or abscesses.[64,75,111,113–115,122,127,131]

US does have some limitations when used to assess SBCD. Higher BMI can reduce image quality, although this partly offset by improved probe and software technology. While bowel gas may reduce sensitivity, carefully graded compression and meticulous scanning technique can overcome this.[78]

The most prominent parameter for the detection of inflammation used throughout the reviewed sources was bowel wall thickness (BWT), which correlates well with clinical disease activity markers such as the Harvey Bradshaw Index (HBI) and the CD activity index (CDAI).[64,74,86,87,108,110–115,117–120,123–126,128,132–136] The exact BWT that is considered to be pathological is still a matter of debate. The most common cut off value was BWT exceeding 3mm being considered pathological and a BWT of 2mm or less considered normal for the SB.[122,123,133] Any increase in the cut off value of BWT will increase sensitivity but also subsequently decrease the specificity of detecting SBCD.[126,133] Using a cut off level of 3mm to represent normality, provided a sensitivity and specificity of 88% and 93% respectively, whereas when a cut off level of greater than 4mm was used, sensitivity was 75% and specificity 97%.[125,128,137]

A number of US scores have been developed, but most lack validation, are overly complex, were developed from small sample sizes or are limited to quantification of damage or the risk of surgery.[118,138,139] Novak et al[118] have developed a promising, simple US score for identifying CD activity comparing BWT to endoscopic activity, however this gauges disease activity alone and the results reported have not yet been externally validated.[118]

Multiple authors suggest that US may have a role as a useful examination for predicting disease course monitoring the response to treatment in CD patients.[86,116,120,125,140]

Opportunity of detecting transmural healing has been shown by Paredes et al[141] who used US for assessing changes induced with an anti-tumour necrosing factor (TNF) therapy in patients with CD. The study reported a significant reduction in BWT in patients receiving anti-TNF therapy, however the study reports that 'resolution' of inflammation visible on US was only achieved in 29% of subjects.[125] Results from Ripolles et al [135]showed that US may be able to predict the 1 year response to anti-TNF therapy after 12 weeks of treatment with 85% (22/26) of patients showing a sonographic response at 12 and 52 weeks.

The METRIC[75] study found no major difference between MRE and US on therapeutic decision-making. Both tests agreed with a final therapeutic decision based on all tests in > 75% of cases. Very little further investigation into the impact of the use of US on the clinical decision-making behaviours of clinicians has been undertaken.

Multiple sources refer to US being inexpensive, though there is little empirical evidence within the included sources to support this claim.[75,78,113–116,142] The METRIC[75] study presents data on a cost-utility analysis of MRE vs US indicating a trend towards US over MRE. However, given the small non-significant differences in costs and QALYs between the two options, it was not possible to endorse US or MRE on cost-effectiveness grounds.

Initial results from Grunshaw[78] indicate that almost half of the cases they discussed scans were performed and reported within 7 days, compared to the typical referral-reporting time of MRE of 6-12 weeks. Many centres have standalone IBD US lists.[127] In those centres operating US clinics, appointments typically last around 15-20 minutes, and requires minimal preparation.[78,136] Following the scan the report is immediately available, allowing for immediate decision making in some instances.[78] Those patients requiring consultant review were able to be identified at the point of US examination.[78]

Throughout the included sources the results reported were from US being performed by individuals with extensive experiences of US. [78,108–110,112–114,119–121,134,135,143] For example, Taylor et al [75] reports that the team involved in the METRIC study had an average of 8 years (4-11) experience of interpreting US.

The Canadian association of gastroenterology and British society of gastroenterology, advocates the use of US to ascertain disease activity state, however both note that the US testing is operator depandant.[142,144]

Multiple authors have speculated this is due to lack of training availability and the substantial training and experience requirements of those preforming the test.[125,144] There is a reportedly long learning curve to develop competent US skills in gastroenterology.[78,128,143] However some authors suggest that gastroenterologists are equipped with the knowledge and dexterity to rapidly acquire US competency.[74,127,143,145] Interobserver agreement between sonographers with variable experience in US has been reported in a few preliminary studies showing satisfactory results.[76,108–110,125,127,133,143] Initiatives such as European Crohn's and Colitis organisation (ECCO) imaging workshops and standardised training curricula developed by the international Bowel Ultrasound Group (IBUS) offer opportunities for training.[136] US has been included in the training programme in several medical schools.[146]

Patients experiences and perceptions of test burden and levels of physical and psychological discomfort can impact on compliance, even if the test is diagnostically superior to alternatives.[67] US is very well tolerated and patients prefer to be assessed with US rather than with endoscopic procedures or other more invasive imaging techniques.[64,74,78]

MRE recovery time has been shown to be significantly longer than US.[67] The proportion of participants willing to repeat MRE was 127/147 (91%). This was lower than for US where 133/135 (99%) were happy to repeat the test.[67] Overall 128/145 patients rated MRE as very

or fairly acceptable, while 144/146 (99%) participants rated US as very or fairly acceptable. Issues reported by patients concerning MRE mainly reflected ingesting contrast, repeated breath holds and the after-effects of contrast such as diarrhoea and bloating. Nearly half of the patients (49%) reported US as being fine, with no least acceptable part of the imaging process.[67] One important finding is that patients rated diagnostic accuracy as the most important attribute and more important than the challenges related to investigations.[67,75,76,147]

None of the included sources presented findings related to preferences of clinician or patients as to where and when US should be delivered. Grunshaw[78] described how a direct booking service and dedicated US slots reduced waiting times and improved patients experiences, but there was no comparison between outpatient US clinic appointments and point of care US(POCUS) use. Aside from the findings relating to the differences between US examination being undertaken by gastroenterologists or sonographers, there were no further findings regarding who should undertake the US examinations.

Discussion

The treat-to-target paradigm present in IBD management guidelines is similar in other chronic diseases.[14,148–150] Management strategies in CD reflect a step-up paradigm, where patients clinical symptoms in conjunction with markers of inflammation tend to guide investigation or medical intervention.[121,151] Patient symptoms do not necessarily correspond to inflammatory activity and current guidelines recommend that management should be based on objective evaluations.[52,142,144] Mucosal healing, defined by the absence of ulcerations, is recommended as the therapeutic goal in clinical practice.[34,35,144]

MRE is the current standard for assessing SBCD, however It is expensive, time consuming and a challenging investigation patients.[70,71] In addition, the use of gadolinium as contrast agent is currently being critically discussed as recent data report long term retention of gadolinium in the brain of exposed patients.[68,69]

Meta-analyses suggest that MRE and US have similar accuracy for diagnosing and staging SBCD.[82,83,152–154] US could be a good alternative to more invasive and expensive imaging techniques. Besides being quick, well tolerated, relatively inexpensive and readily available, US is reported and interpreted at the time of scanning and allows for early clinical decision-making in routine IBD care.[74,127]

Multiple sources referred to US as inexpensive.[64,74,78,87,108–127] However none of the included sources presented comprehensive data relating to cost or cost effectiveness of US. Therefore, more data on the cost, cost effectiveness and cost–benefit ratio of US are needed to support the embedding of US in IBD services.[74]

US is often seen as having limited clinical utility due to operator dependence.[127] However, every diagnostic technique, including endoscopy, has a degree of subjectivity and operator dependence and this criticism is perhaps more reflective of a previous lack of identifiable international performance and training standards.[127] The training needs for gastroenterologists are minimal, with a small learning curve, particularly when supported by abdominal radiology specialists and in partnership with radiology departments.[74,78,127] There is no current literature relating to any other IBD healthcare worker undertaking US training.

There are several scoring systems for disease activity assessment using US in CD, however until recently none had been completely validated.[133,155,156] The most widely used scoring system is the Limberg score incorporating BWT and Doppler vascularity.[157,158]

There has not been any further work to investigate patient or HCPs preferences or service delivery. It would seem prudent to investigate broader stakeholder perceptions of the use of US in order to better understand perceived or potential barriers and enablers to US implementation in the world-wide healthcare systems and recognise and manage preferences for future service delivery.

Limitations

Scoping reviews do not formally evaluate the quality of evidence gathering information from a wide range of study designs and methods, providing a descriptive account of available information leading to broad overview of the available literature rather than in-depth analysis. The scope of background information collected, disease activity levels, depth of data relating to types and magnitude of fatigue and its effects appears to vary vastly between sources, making comparison challenging. The outcomes represent an accurate response to the research question. Continuous conversations between authors occurred throughout to ensure a unanimous decision regarding article searches, thus limiting any potential bias.

Conclusions

US has been shown to be as accurate as MRE in detecting presence and extent of SBCD. US is reported as quicker, more acceptable and safer when compared to MRE. US is used widely in central Europe, Canada and some parts of the USA, but has not been embraced in the UK. The resources required in terms of equipment, are readily available in most NHS hospitals, and training needs for gastroenterologists would be minimal, especially if supported by radiology consultants and departments.

Multiple sources reported US as an inexpensive test, however there is scant literature to support this. Further research in this area would better inform decision makers regarding future intervention implementation.

US is reported as having positive influence on clinical IBD practice through expediting clinical decision making, but there is no evidence relating to the impact on the nature of clinical decision making by HCPs. Further research in this area would help us to better understand the impact of US on daily clinical practice, ultimately leading to better understanding of practicable and acceptable aspects of clinical utility.

US was considered highly acceptable by patients when compared with MRE, however further exploration of experiences, perceptions and perceived barriers and enablers to US implementation in the NHS is warranted. This information will help guide researchers to areas of focus to successfully implement US as an NHS service for patients with SBCD.