Volume assessment magnetic resonance imaging technique for monitoring perianal Crohn’s fistulas

Phillip F. C. Lung, Kapil Sahnan, David Burling, James Burn, Phillip Tozer, Nuha Yassin, Samuel O. Adegbola, Rachel Baldwin-Cleland, Janindra Warusavitarne, Arun Gupta, Omar Faiz, Robin K. S. Phillips and Ailsa L. Hart

Abstract

Background: Perianal Crohn’s fistula and their response to anti-tumour necrosis factor (TNF) therapies are best assessed with magnetic resonance imaging (MRI), but radiologist reporting is subjective and variable. This study investigates whether segmentation software could provide precise and reproducible objective measurements of fistula volume.

Methods: Retrospective analysis of patients with perianal Crohn’s fistula at our institution between 2007 and 2013. Pre- and post-biologic MRI scans were used with varying time intervals. A total of two radiologists recorded fistula volumes, mean signal intensity and time taken to measure fistula volumes using validated Open Source segmentation software. A total of three radiologists assessed fistula response to treatment (improved, worse or unchanged) by comparing MRI scans.

Results: A total of 18 cases were reviewed for this pilot study. Inter-observer variability was very good for volume and mean signal intensity; intra-class correlation (ICC) 0.95 [95% confidence interval (CI) 0.91–0.98] and 0.95 [95% CI 0.90–0.97] respectively. Intra-observer variability was very good for volume and mean signal intensity; ICC 0.99 [95% CI 0.97–0.99] and 0.98 [95% CI 0.95–0.99] respectively. Average time taken to measure fistula volume was 202 s and 250 s for readers 1 and 2. Agreement between three specialist radiologists was good [kappa 0.69 [95% CI 0.49–0.90]] for the subjective assessment of fistula response. Significant association was found between objective percentage volume change and subjective consensus agreement of response (p = 0.001). Median volume change for improved, stable or worsening fistula response was -67% [interquartile range (IQR): -78, -47], 0% (IQR: -16, +17), and +487% [IQR: +217, +559] respectively.

Conclusion: Quantification of fistula volumes and signal intensities is feasible and reliable, providing an objective measure of perianal Crohn’s fistula and response to treatment.

Keywords: fistula: Crohn’s disease, MRI

Introduction

One third of all Crohn’s disease (CD) patients develop perianal fistula,1 representing a distinct and aggressive phenotype,2 causing pain and discharge that results in reduced quality of life.3 The disease course is often severe and disabling, involving multiple medical and surgical interventions.1,4,5 Established treatment principles involve a multidisciplinary approach,6 aggressively managing proctitis, judicious use of surgery to drain sepsis, and optimising medical Crohn’s treatment through a combination of antibiotics, immunosuppressants and anti-tumour necrosis factor (TNF) therapies.

Response to treatment can be assessed in a variety of ways. The most common scoring tools and techniques used include: fistula drainage assessment...
(FDA; described below); perianal disease activity index (PDAI); Crohn’s disease activity index (CDAI); inflammatory bowel disease questionnaire (IBDQ) and magnetic resonance imaging (MRI). The current standard of clinical response to treatment is the FDA as used in the ACCENT II trial, which defines remission as ‘complete cessation of drainage from all fistula openings present at baseline despite gentle finger pressure, on at least two occasions over at least a 4 week period, and improvement as a 50% reduction in the number of external openings present or draining. CD is a chronic relapsing and remitting disease as are its fistulising complications, so clinical remission cannot be determined at individual time points, as the very nature of discharge is sporadic. Furthermore, closure of an external opening does not necessarily equate to deep healing of the tracts, or indeed improved quality of life.

Pelvic MRI has been shown to be the most accurate method for classification of the primary tract and secondary extensions; the recognition and treatment of the latter being essential to ensure satisfactory healing. MRI scans inform clinicians about fistula complexity, for example, the number of tracts, presence of undrained sepsis and the patients’ response to treatment interventions. Importantly, MRI can evaluate deep tissue healing, which cannot be clinically assessed.

Assessment of radiological improvement is subjective and wide variation exists between radiologists. Previous attempts to standardize radiological response, such as the Van Assche system, have been criticised due to insensitivity to change and represent descriptive scores rather than monitoring tools. Currently, there is no universally accepted or reliable method of monitoring long-term radiological response to treatment. However, it is our view that clinical monitoring alone is inadequate. Proper evaluation of interventions for perianal Crohn’s fistula requires accurate monitoring of response, as does the development of tools that will enable prediction of successful treatment. Future research and development therefore needs to focus on accurate forms of imaging to provide standardized, responsive and accurate tools to assess radiological response to treatment.

The imaging module of the recently published fistulising CD core outcome set comprises two outcomes, ‘Fistula response on imaging’ and ‘an activity-based MRI score responsive to change’. At present, perianal fistula complexity on MRI is judged based on fistula configuration, involvement of adjacent pelvic spaces, presence of undrained sepsis and volume. Whilst the first three components are based on anatomical assessment, there is currently no method to calculate fistula volume and it is not performed in routine practice. We hypothesize that volume change is an important measure of fistula response that may be incorporated as part of a core measurement set.

This pilot study aims to confirm the feasibility of objectively calculating perianal fistula volume using segmentation software through evaluation of repeatability, reproducibility and comparison with the current standard of subjective assessment of MRI fistula response.

Methods

Ethics

All patients were treated on the basis of clinical need according to standard clinical care at our institution, on an open-label basis and according to licensed or published regimens. After consultation with the local research and development (R&D) department (London North West Healthcare NHS Trust), it was advised that approval by a National Health Service (NHS) research ethics committee or the R&D department was deemed unnecessary owing to the non-patient identifiable retrospective nature of this feasibility study. The study received departmental approval by the research lead of St Mark’s Hospital, London, United Kingdom. The study was registered as a service evaluation and all patients were treated only after full and informed clinical consent as part of routine care, and therefore research consent was not required.

Patients and setting

We performed a retrospective analysis of patients with perianal Crohn’s fistula selected using a random number generator from a cohort treated at St. Mark’s Hospital, UK between 2007 and 2013. Adult patients were selected if they met the following inclusion criteria: perianal Crohn’s fistula; current anti-TNFα treatment (infliximab or adalimumab); and had two MRI studies >1 year apart. MRI scans were performed twice a year following biological therapy and then annually as
part of the St Mark’s fistula protocol. A cohort of
18 patients were selected using a random number
generator by one of the authors (NY) for this pilot
study as a previous paper was able to assess inter-
observer variability on assessment of fistula scor-
ing in a group of this size. Patient demographic
data were retrieved.

**MRI technique**
Standard axial T2-weighted Spectral Attenuated
Inversion Recovery (SPAIR) MRI sequences
were acquired (TR 8000; TE96; thickness 4 mm;
gap 0.2 mm; FOV 240 mm; averages 4, flip angle
150, bandwidth 130 Hz/PX). Digital imaging and
communications in medicine images of the MRI
sequences were retrieved from the picture archiv-
ing and communication system and anonymized.

**MRI evaluation**
Perianal fistulas were defined as tubular struc-
tures that abnormally connected the enteric
lumen to the skin and the borders defined as the
transition between T2 high signal and low signal
fibrosis/fat on fat suppressed sequences. Each fis-
tula tract at baseline and follow up was assessed
and characteristics documented by reader, PL.
Fistula characteristics included: fistula type based
on the Park’s classification; simple or branched
fistula configuration; supralevator or ischioanal
fossa involvement; presence of horseshoe config-
uration; presence of interstitial high signal sur-
rounding the fistula tract; and presence of proctitis.

A total of two specialist gastrointestinal radiolo-
gists (PL and JB with 7 and 3 years’ experience)
measured fistula volumes, signal intensity and
time taken to measure fistula volumes on baseline
and follow-up MRI scans using ITK-SNAP (ver-
sion 2.2.0), a validated Open Source segmenta-
tion software18 (Figure 1). All contents within the
borders of the fistula were considered part of the
fistula volume. Reader PL repeated the readings
for MRI studies after 2 months.

To minimize bias, three experienced, specialist
gastrointestinal radiologists (PL, AG, DB with 7,
12 and 15 years’ experience) also assessed fistula
response to treatment (improved, worse or
unchanged) by comparing baseline and follow-up
MRI scans for each patient.

**Statistics**
Inter and intra-observer correlation for measure-
ment of fistula volumes were assessed using intra-
class correlation (ICC). Subsequent analysis of
fistula volume was performed using data from
reader PL.

Agreement between specialist gastrointestinal
radiologists for fistula response to treatment was
assessed using the kappa statistic. Consensus
agreement was taken if two or more radiologists
agreed. Differences in percentage volume change
for each category of radiologist assessed fistula
response to treatment were assessed using the
Kruskal–Wallis test, for each reader and also for
the consensus.

The Mann–Whitney test was used to compare
percentage volume change with fistula character-
istics with only two categories, whilst the Kruskal–
Wallis test was used where there were more than
two categories.

![Figure 1. Example of manual segmentation of a perianal fistula. (a) An axial fat saturated T2 weighted image showing a complex high signal fistula [arrows]. (b) The same image with the fistula having been manually segmented (now highlighted in red).](image)
Results
A total of 18 patients (12 men, 6 women) with Crohn’s perianal fistula were recruited with median age of 29 years (range 19–52 years). At baseline, they had a total of 23 fistula tracts; 13 with single tracts only and 5 with two fistula tracts. 15 of 23 (65%) tracts were transphincteric, 5 (22%) were intersphincteric and 3 (13%) were extrasphincteric. A total of seven patients had proctitis. Fistula characteristics during the course of this study are detailed in Table 1.

Average time taken to calculate volume measurements was 3 min 22 sec and 4 min 10 sec for readers 1 and 2 respectively with moderate agreement (ICC 0.73) in time taken between readers. Inter-observer variability between readers was very good for volume and mean signal intensity with ICC of 0.95 [95% confidence interval (CI) 0.91–0.98] and 0.95 (95% CI 0.90–0.97) respectively. Intra-observer variability was also very good for volume and mean signal intensity with ICC of 0.99 (95% CI 0.97–0.99) and 0.98 (95% CI 0.95–0.99) respectively.

There was good agreement for assessment of objective fistula response between all three radiologists (kappa 0.69; 95% CI: 0.49–0.90), with all (100%) cases agreed by at least two radiologists and most (13/18; 72%) agreed by all three radiologists.

Manual segmentation measuring percentage change in fistula volume was significantly associated with

Table 1. Patient fistula characteristics and their association with percentage volume change.

| Variable                                | N  | Median % volume change (IQR) | p-value |
|-----------------------------------------|----|-----------------------------|---------|
| Fistula tract type                      |    |                             |         |
| Intersphincteric                        | 3  | 514 [-78, 559]              | 0.63    |
| Transphincteric                         | 12 | ~5 [-61, 301]               |         |
| Extrasphincteric                        | 3  | ~55 [-67, 217]              |         |
| Configuration                           |    |                             |         |
| Solitary                                | 4  | 337 [54, 523]               | 0.34    |
| Branched                                | 14 | ~30 [-67, 217]              |         |
| Suprarelevator/ischioanal involvement   |    |                             |         |
| No                                      | 3  | 515 [-78, 559]              | 0.37    |
| Yes                                     | 15 | ~17 [-67, 217]              |         |
| Horseshoe sepsis                        |    |                             |         |
| No                                      | 9  | 186 [-47, 487]              | 0.20    |
| Yes                                     | 9  | ~42 [-74, 17]               |         |
| Peri-fistula interstitial high signal   |    |                             |         |
| No                                      | 6  | ~47 [-74, 416]              | 0.88    |
| Yes                                     | 12 | 6 [-17, 217]                |         |
| Proctitis                               |    |                             |         |
| No                                      | 13 | 6 [-55, 416]                | 0.59    |
| Yes                                     | 5  | ~42 [-82, 217]              |         |

IQR, interquartile range.
subjective assessment of radiological response to treatment by all three radiologists ($p = 0.001$). The calculated median volume change was $-67\%$ [interquartile range (IQR): $-78$, $-47$], $0\%$ (IQR: $-16$, $+17$), and $+487\%$ (IQR: $+217$, $+559$) for fistulas assessed by radiologists as improved, stable or worsening respectively.

Percentage volume change measured by manual segmentation was not associated with any of the fistula characteristics assessed by radiology (Table 1).

The observed difference in median percentage change in fistula volume between infliximab and adalimumab, $-45\%$ (IQR: $-70\%$, $+216\%$) and $+201\%$ (IQR: $+6\%$, $+559\%$) did not reach significance ($p = 0.16$).

No significant association was found between the categorical (Table 2) and continuous (Table 3) variables relating to participants’ CD and percentage volume change.

### Discussion

MRI is the gold standard test for assessment of perianal Crohn’s fistula and subjective assessment of change in fistula volume is frequently reported to help determine response to treatment, particularly for Gastroenterologists assessing the impact of drug treatment.$^{8,12}$ However, there has been no objective method for quantification of fistula volume and this has therefore not been included in MRI scoring systems to date.$^{8,12,17,19}$ Studies of surgical treatment have also struggled to identify a robust, objective measure of radiological improvement.$^{20}$ The present study has shown that volume of perianal fistula in patients with CD can be reliably measured by MRI with very good inter- and intra-observer agreement. Objective measurement of volume is facilitated by the sharp boundary between high intensity fistula and the surrounding low intensity fat or scar tissue (fibrosis), which can be perceived easily by the human observer.

Manual measurement of fistula volume took a relatively short time (up to 4 min 10 sec) with the difference between readers likely related to experience. The authors acknowledge that any significant delay in reporting times limits generalizability and warrants scrutiny, and therefore determination of the additional clinical benefit of precise volume measurement over subjective assessment will be required. That said, objective fistula volume measurement can be applied to very

### Table 2. Categorical patient and disease factors, and their association with fistula volume change.

| Variable                  | Category                                      | N   | Mean fistula % volume percentage change (SD) | $p$-value |
|---------------------------|-----------------------------------------------|-----|---------------------------------------------|-----------|
| Gender                    |                                               |     |                                             |           |
|                           | Male                                          | 12  | $-5$ [−66, 337]                             | 0.71      |
|                           | Female                                        | 6   | 85 [−67, 416]                               |           |
| Site affected at presentation |                                       |     |                                             |           |
|                           | Perianal disease only                         | 7   | −7 [−67, 217]                               | 0.62      |
|                           | Perianal + luminal disease                   | 11  | 6 [−74, 487]                                |           |
| Biologic therapy          |                                               |     |                                             |           |
|                           | Infliximab                                    | 12  | −45 [−70, 216]                             | 0.16      |
|                           | Adalimumab                                    | 6   | 201 [6, 559]                               |           |
| Thiopurines               |                                               |     |                                             |           |
|                           | No                                            | 5   | −47 [−78, 6]                               | 0.30      |
|                           | Yes                                           | 13  | 17 [−55, 416]                              |           |
| SD, standard deviation.   |                                               |     |                                             |           |
complex disease that involve multiple tracts and collections, which are usually more difficult to assess subjectively. Furthermore, it will be possible to automate measurement of fistula volume in the future, negating any impact on reporting times and enabling less experienced observers to assess fistula volume more accurately.

We found good inter-observer agreement for the subjective assessment of fistula response on MRI indicating that it represents a good standard against which to assess volume as objective measures. We also demonstrated minimal inter and intra-observer variability in volume measurement indicating this represents a consistent and repeatable measure. We investigated the association between change in fistula volume and subjective assessment of fistula response by experienced gastrointestinal (GI) radiologists at St. Mark’s Hospital, UK. We found strong agreement between this objective measure and the subjective assessment (Figure 2).

Having determined that volume demonstrates good inter- and intra-observer variability, is quick to perform and correlates well with the current standard of subjective assessment of improvement when performed by experienced GI radiologists, we plan to investigate the relationship between change in fistula volume and subjective assessment of fistula response by experienced gastrointestinal (GI) radiologists at St. Mark’s Hospital, UK. We found strong agreement between this objective measure and the subjective assessment (Figure 2).

Despite good agreement overall, disagreement between readers for individual patient assessments are noteworthy. One fistula with an objective 17% increase in volume was reported as ‘no change’ by two readers and ‘worse’ by one reader. When reviewed retrospectively, all readers agreed there was a mixed response where different component of the fistula increased and decreased in volume over time. Similarly, another fistula had nearly healed bar a subcutaneous fluid focus, which had increased in size leading to an overall 6% increase in fistula volume. These cases reinforce the need to assess additional fistula characteristics, beyond volume alone, when judging response to treatment. Notably, our study also found that these additional characteristics (Table 1) are independent of percentage volume change (implying that scoring systems based on description of these characteristics may remain insensitive to change, and that specific fistula configurations do not appear to obviate the utility of this technique, although larger studies are required to confirm these).

We believe that accurate assessment of reduction in inflammation and volume is key to the clinical utility of this technique. One hypothesis is that the rate and/or extent of volume reduction in a given period following treatment (at 3 months from induction, for example), will predict the rate and completeness of fistula closure, the duration of remission or the risk of fistula recurrence. This would be beneficial in stratifying patients undergoing expensive treatment with significant adverse effects.

Although we did not find significant correlation between fistula volume and clinical findings, we
suspect this is related to the small sample size and heterogeneous cohort. This pilot work was designed to develop the volume measurement technique and was not powered to detect such a correlation with clinical findings.

The present study has several limitations, including a small sample size and its retrospective design. However, it confirms the feasibility of our technique to measure perianal fistula volumes. Additionally, the standard with which we compared our objective measurement is subjective and complete agreement between all three assessors was not observed. Nevertheless, agreement between them was strong and as experienced, specialist GI radiologists their subjective assessment represents the best currently available standard with which to compare volume assessments.

Conclusion
Quantification of fistula volumes is feasible and reliable, with the potential to provide objective measurement of perianal Crohn’s fistula and response to treatment. This is the first step towards developing an objective and responsive MRI based volume score.

Acknowledgements
All authors contributed to this manuscript. Phillip F. C. Lung and Kapil Sahnan are joint first author. Previous presentation are as follows:

1. Sahnan K, Lung PFC, Adegbola SO, Burling D, Burn J, Tozer PJ, Gupta A, Faiz OD, Phillips RKS, Hart AL. An objective measure of response to treatment for patients with Crohn’s perianal fistulas on anti-TNF treatment. European Crohn’s & Colitis Organisation. Poster Presentation, 15-18th February 2017.

2. Sahnan K, Adegbola SO, Tozer PJ, Faiz OD, Phillips RKS, Hart AL. Novel use of MRI volume measurement & 3D modelling in perianal Crohn’s fistulae. ‘Global Health Innovator Competition’, Institute for Global Health Innovation, Oral Presentation, 7th March 2016.

3. Sahnan K, Phillip RKS. 3D reconstruction of perianal fistula and the use of MRI volumes. Imaging, sensing and digital in GI medicine, Enteric Hackday 2016. Invited podium presentation. Course Convenors – Prof Knowles, Prof Williams, 14th October 2016.

Funding
Kapil Sahnan is supported by a Royal College of Surgeons of England Research Scholarship. The other authors have no conflict of interests or financial ties to disclose.

Conflict of interest statement
The authors declare that there is no conflict of interest.

References
1. Ardizzzone S and Porro GB. Perianal Crohn’s disease: overview. Dig Liver Dis 2007; 39: 957–958.

2. Tang LY, Rawsthorne P and Bernstein CN. Are perineal and luminal fistulas associated in Crohn’s disease? A population-based study. Clin Gastroenterol Hepatol 2006; 4: 1130–1134.
3. Afşarlar CE, Karaman A, Tanr G, et al. Perianal abscess and fistula-in-ano in children: clinical characteristic, management and outcome. Pediatr Surg Int 2011; 27: 1063–1068.

4. Beaugerie L, Seksik P, Nion-Larmurier I, et al. Predictors of Crohn’s disease. Gastroenterology 2006; 130: 650–656.

5. Hughes LE. Surgical pathology and management of anorectal Crohn’s disease. J R Soc Med 1978; 71: 644–651.

6. Sahnan K, Adegbola SO, Tozer PJ, et al. Perianal abscess. BMJ 2017; 356: j475.

7. Irvine EJ. Usual therapy improves perianal Crohn’s disease as measured by a new disease activity index. McMaster IBD Study Group. J Clin Gastroenterol 1995; 20: 27–32.

8. Ng SC, Plamondon S, Gupta A, et al. Prospective assessment of the effect on quality of life of anti-tumour necrosis factor therapy for perineal Crohn’s fistulas. Aliment Pharmacol Ther 2009; 30: 757–766.

9. Colombel J-F, Sandborn WJ, Rutgeerts P, et al. Adalimumab for maintenance of clinical response and remission in patients with Crohn’s disease: the CHARM trial. Gastroenterology 2007; 132: 52–65.

10. Sands BE, Blank MA, Diamond RH, et al. Maintenance infliximab does not result in increased abscess development in fistulizing Crohn’s disease: results from the ACCENT II study. Aliment Pharmacol Ther 2006; 23: 1127–1136.

11. Ingle SB and Loftus EV. The natural history of perianal Crohn’s disease. Dig Liver Dis 2007; 39: 963–969.

12. Tozer P, Ng SC, Siddiqui MR, et al. Long-term MRI-guided combined anti-TNF-α and thiopurine therapy for Crohn’s perianal fistulas. Inflamm Bowel Dis 2012; 18: 1825–1834.

13. Sahni VA, Ahmad R and Burling D. Which method is best for imaging of perianal fistula? Abdom Imaging 2008; 33: 26–30.

14. Limura E. Modern management of anal fistula. World J Gastroenterol 2015; 21: 12.

15. Sahnan K, Tozer PJ, Adegbola SO, et al.; ENiGMA Collaborators. Developing a core outcome set for fistulising perianal Crohn’s disease. Gut 2018.

16. Halligan S and Stoker J. Imaging of fistula in ano. Radiology 2006; 239: 18–33.

17. Van Assche G, Vanbekevoort D, Bielen D, et al. Magnetic resonance imaging of the effects of infliximab on perianal fistulizing Crohn’s disease. Am J Gastroenterol 2003; 98: 332–339.

18. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006; 31: 1116–1128.

19. Shenoy-Bhangle A, Nimkin K, Goldner D, et al. MRI predictors of treatment response for perianal fistulizing Crohn disease in children and young adults. Pediatr Radiol 2014; 44: 23–29.

20. Panes J, Garcia-Olmo D, Van Assche G, et al. Expanded allogeneic adipose-derived mesenchymal stem cells (Cx601) for complex perianal fistulas in Crohn’s disease: a phase 3 randomised, double-blind controlled trial. Lancet 2016; 388: 1281–1290.