



This is a repository copy of *Establishment of a validated central reading system for ileocolonoscopy in an academic setting*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/184026/>

Version: Accepted Version

Article:

Raine, T., Pavey, H., Qian, W. et al. (9 more authors) (2022) Establishment of a validated central reading system for ileocolonoscopy in an academic setting. *Gut*, 71. pp. 661-664. ISSN 0017-5749

<https://doi.org/10.1136/gutjnl-2021-325575>

© 2022 The Author(s) (or their employer(s)). This is an author-produced version of a paper subsequently published in *Gut*. Reuse of this manuscript version (excluding any databases, tables, diagrams, photographs and other images or illustrative material included where a another copyright owner is identified) is permitted strictly pursuant to the terms of the Creative Commons Attribution-Non Commercial 4.0 International Licence (<http://creativecommons.org/licenses/by-nc/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Establishment of a validated central reading system for ileocolonoscopy in an academic setting

Tim Raine¹, Holly Pavey², Wendi Qian², Gordon W Moran^{3,4}, Sree Subramanian¹, Lizzie Swaby⁵, Simon Travis⁶, Shahida Din⁷, Peter M Irving^{8,9}, James O Lindsay¹⁰, Miles Parkes^{1,11}, Nicholas A Kennedy¹²

1. Department of Gastroenterology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK tim.raine@addenbrookes.nhs.uk
2. Cambridge Clinical Trials Unit, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
3. NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals and the University of Nottingham.
4. Translational Medical Sciences, School of Medicine, Faculty of Medicine and Health Sciences, University of Nottingham. Nottingham. NG7 2UH.
5. Clinical Trials Research Unit, School of Health and Related Research, University of Sheffield, Sheffield, UK
6. Translational Gastroenterology Unit, NIHR Oxford Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
7. Edinburgh IBD Unit, Western General Hospital, NHS Lothian, Edinburgh, Scotland, UK
8. IBD Unit, Guy's and St Thomas' Hospitals NHS Trust, London, UK
9. School of Immunology & Microbial Sciences, King's College London, London, UK
10. Centre for Immunobiology, Blizard Institute, Barts and the London School of Medicine, Queen Mary University of London
11. Department of Medicine, University of Cambridge, Cambridge, UK

12. Department of Gastroenterology, Royal Devon and Exeter NHS Foundation Trust,
Exeter, UK; Exeter Inflammatory Bowel Disease and Pharmacogenetics Research
Group University of Exeter, Exeter, UK.

Background

Therapeutic interventions in inflammatory bowel disease (IBD) should alleviate patient symptoms and modify long term disease outcomes. This second target is hard to demonstrate within the context of typical observational or interventional studies due to problems of sample size and duration of follow up. Instead, there has been considerable focus on the assessment of mucosal inflammation using validated scoring systems in both Crohn's disease (CD)[1,2] and ulcerative colitis (UC).[3,4] Achievement of endoscopic healing correlates with avoidance of long-term disease complications in both CD and UC,[5,6] and is an established therapeutic target in addition to the achievement of clinical remission.[7]

For these reasons, data on endoscopic outcomes are now required by regulatory agencies for any registration trial.[8] In addition, endoscopic outcome data may be used in earlier phase trials to guide decisions about further asset development, adding biological context to clinical data.[9] Interventions which lead to improvements in symptom scores may not demonstrate improvements in endoscopic outcomes and this additional information may inform or limit placement in treatment paradigms.[10,11] Furthermore, the use of endoscopic activity as an inclusion criterion for clinical trials can reduce the high placebo response rates previously reported in trials reliant purely on clinical inclusion criteria, which may mask efficacy signals.[12]

Key barriers to use of endoscopy in clinical trials include patient acceptability.[13] Additionally, there is a need for blinded central reading of all trial endoscopies to ensure standardisation and avoid inadvertent introduction of investigator bias: local endoscopy reads are typically less reliable than blinded central reading by a recognised expert

reader.[14] Reliance on local readers may even lead to critical errors in trial interpretation.[15] Furthermore, use of central readers alone does not completely mitigate against errors of endoscopy reading that may lead to false conclusions about novel interventions,[16] underscoring the importance of ensuring validity and quality in any reader pool.

Currently, access to a validated pool of expert readers may be beyond the budget of typical investigator-initiated studies or early phase industry studies, which may instead rely on local reads or unvalidated central readers.[10,17] Previous reports have demonstrated the validity and reliability of small expert reader pools.[18] Nevertheless, the methodology and feasibility of undertaking the process of establishing a validated pool of central readers has not previously been reported. We developed a protocol and platform to support the validation of a pool of central readers, using the simple endoscopic score for CD (SES-CD) as an example.[2] This is a relatively straightforward and reliable score that shows good inter- and intra-observer variability after training.[18] We report here the methodology and the results of this exercise as well as a process for maintaining and assuring the quality of the pool of readers for future studies.

Methods and results

Platform

We developed a platform to enable uploading, processing and reading of endoscopic videos using free-to-use software. The core web-based system was implemented using C# and ASP.NET core (.NET 5.0; Microsoft, Redmond, USA), with data stored in MySQL 8 (Oracle, Redwood Shores, USA). Videos were captured locally in various formats using a range of video

recording hardware compatible with local endoscopy systems, then uploaded and transcoded to segmented MP4 files using FFmpeg (FFmpeg developers) and Bento4 (Axiomatic Systems LLC, San Francisco, USA). Minimum video resolution was 720x480 pixels. Video playback was provided for central readers using Video.js (Brightcove Inc., Boston, USA).

Initial establishment of a reader pool

A library of anonymised ileocolonoscopy videos from patients with CD was created on the platform. Fully anonymised videos were used with appropriate informed consent obtained at the time of index examination using internal institutional approvals. Videos were acquired by endoscopists under routine clinical conditions; endoscopists were made aware that the videos were being recorded for scoring purposes and reminded to ensure adequate mucosal visualisation, but not provided with any further instructions or specific training. Endoscopists were requested not to upload videos with bowel preparation judged to be inadequate for mucosal assessment. The platform provides the option for readers to indicate that bowel preparation was inadequate for scoring purposes, but this option was not used by any of the readers in the present exercise.

Video scoring and analysis proceeded according to a pre-specified analysis plan. Seven prospective scorers were identified from a group of experienced clinical trialists, all with prior experience of use of the SES-CD as local readers. Scorers were invited to attend an initial videoconference where the SES-CD was reviewed and the use of the platform demonstrated. In particular, areas of common difficulty were discussed and a series of “Top Tips” agreed and provided as an aid for reference during subsequent scoring (see Supplementary material 1). Although one reader acted as chair and facilitated the discussion (JOL), no one individual

acted as a ‘trainer’ but rather the group discussed use of the score, as well as reviewing a series of endoscopic videos selected to illustrate a range of commonly encountered challenges.

The scorers were then asked individually to score sets of videos taken in random batches of 10 from the library (without replacement), using a standardised template to report the SES-CD score and sub-components (Supplementary material 1). Inter-rater reliability scores were determined by a blinded independent assessor using intraclass correlation coefficients (ICC) and 95% confidence interval (CI), with a two-way mixed effects model (treating readers as fixed and videos as random variables). We wanted the eventual pool of readers to allow a high degree of confidence in the absolute score assigned by any single reader chosen at random from within this fixed pool, hence used an ICC calculation based on ‘single reader, absolute agreement’.[19,20] The final formula used for the ICC calculation was:

$$ICC = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$$

Where MS_R denotes mean square for rows; MS_E denotes mean square for error; MS_C = mean square for columns; n = number of videos; k = number of readers. Analysis was performed using *R* version 4.0.0[21] with the ICC calculated using the *irr* package. As a general rule, ICC values of >0.6 can be regarded as good and >0.75 as excellent.[22]

Video scoring and analysis plan

According to the prespecified plan, should the initial estimate of the ICC for the first set of 10 videos be <0.7, then these scores would be removed from further rounds of ICC calculation

and the first 10 videos regarded as a training exercise. In our study, the ICC after the first ten videos was 0.77 (95% CI 0.56 – 0.92) so these scores were retained in subsequent rounds of analysis.

In addition, the ICC was recalculated dropping each scorer in turn to assess any scorer whose inclusion adversely affected the performance of the group – we did not identify any such outliers. ICC were calculated for each subscore of the SES-CD, with the lowest ICC identified for the stenosis subscore of the SES-CD. These findings were all fed back to the scorers at a further video-conference, at which points of discrepancy were discussed and videos of segments showing high inter-scorer variability reviewed.

Next, further batches of 10 randomly selected videos were scored, with the ICC calculations run after each set. We pre-specified that the scoring exercise would be terminated when any of the following criteria were met:

- ICC >0.7 with 95% CI for ICC >0.6 (using scores for all videos scored to date either with or without the initial 10 video scores according to rules above)
- exhaustion of available videos in library for scoring;
- agreement amongst scorers that exercise should be terminated.

Table 1 shows the estimated number of videos required to be scored, in order to produce a lower limit of the 95% confidence interval of 0.6, at different expected ICC levels and with different numbers of readers in the pool. In our initial sample size calculation, we anticipated that 45 videos would be sufficient, based on an expected ICC of 0.7 and assuming that the ICC of the first batch of 10 videos would be ≥ 0.7 .

In fact, after a total of 20 videos had been scored, the ICC was 0.82 (95% CI 0.71-0.91) which meant that we terminated the exercise without needing further scoring rounds. No single reader significantly adversely affected the performance of the pool (ICC estimates remained in range 0.80-0.83 when each reader was removed). As determined after completion of the scoring exercise, the median SES-CD score was 7 (inter-quartile range 4-11). Table 2 shows further characteristics of the videos used. Notably, despite the modest number of videos ultimately read, a wide range of disease severities was represented. The maximum SES-CD score encountered was 22, whilst 1 video had an SES-CD of 0. One video had an impassable stricture, whilst 10 videos included patients with prior resectional surgery. In 3 patients, the extent of endoscopic examination was less than complete (with a combined total of 5 segments of unexplored bowel).

Ongoing quality assurance

The performance of readers may change over time. To assess this, we invited readers to re-score a selection of videos after a period of at least 6 months since the original reading and without any further training. Videos were selected and allocated to readers at random, with subsequent adjustment to ensure that each video was re-scored by at least one reader. Readers were asked to re-score three videos, based upon a prior power calculation assuming 80% power to detect an intra-rater ICC score of 0.8 (Supplementary material 2). Using a two-way, mixed effects, single rater, absolute agreement approach, we determined the intra-rater ICC as 0.937 (95% CI 0.877, 0.968).

To monitor and assure quality, our platform automatically assigns a minimum of 20% of all videos scored on the platform are automatically assigned to a pair of randomly selected

readers, both blinded to this. After each reader has completed at least 20 such paired readings the ICC is recalculated using a 2-way random-effects model applied to all paired readings available since the last performance check. In the event that the lower limit of the 95% CI for the ICC decreases to ≤ 0.6 , the ICC is then recalculated with the sequential exclusion of each reader in the pool to establish whether the exclusion of any one individual lead to an improvement in the lower limit of the 95%CI to >0.6 . If an underperforming reader is identified as contributing to the change in reliability, this reader would be temporarily removed from the pool and offered the chance to retrain, with parallel reading continuing prospectively for a further 20 videos – at which point ICC would be recalculated using these scores and the same performance criteria applied.

New readers

For new readers wishing to join the pool, it is pre-specified that the reader would be offered training, then invited to read the original 10 videos in the training exercise. Their scores for these videos would then be used to calculate an ICC using the new reader's scores in the place of each reader in turn from the original group of 7 readers. The process specifies that the ICC should not decrease with more than 3 of the combinations applied, meaning that each new reader would need to show consistency in line with the median of the original 7 readers. To test this process, we invited an 8th reader to score the original 10 videos and assessed their read quality in this manner and found that when the new reader was substituted for 3 of the original 7, the ICC decreased (lowest value was 0.73), whilst in the other 4 substitutions, the ICC increased (highest value was 0.79). The new reader was therefore admitted to the pool.

Discussion

Central reading of endoscopy videos has played an ever-increasing role in clinical trial conduct, but to date there have been no reports on the protocol for, or establishment of, a validated pool of central readers suitable for use in the context of investigator-initiated studies. Khanna et al. reported the detailed analysis of expert reader performance using a range of CD endoscopy scores and showed good performance of the SES-CD when using 4 expert central readers.[18] The overall intra-reader ICC reported for SES-CD scores in this study was 0.83 (0.75-0.88), against which our reported ICC of 0.82 (95% CI 0.71-0.91) compares favourably, although it must be noted that Khanna et al. used a slightly different approach to ICC derivation due to the different methodology in their study. Whilst a high ICC reflects strong agreement between readers, a note of caution common to all such studies is the absence of a defined 'gold standard' against which to measure overall performance of the reader pool and exclude the possibility of a systematic error common to all readers.

Previous studies have examined how non-expert readers can be trained, leading to improvements in inter-observer agreement (though with wider confidence intervals around the ICC than those we report here).[23] However, until the present study the method and means to establish a central read pool and the feasibility of doing so, using readers without the same degree of prior experience in central reading, had not been reported. In our study, after just 20 videos and 2 teleconferences, our 7 readers were able to demonstrate a high degree of consistency and an 8th reader could be admitted to the pool after similar training. Indeed, the termination of the scoring exercise after just 20 videos likely reflects the importance of adequate training prior to the exercise for a group of readers with prior experience of using the score, and the prior opportunity for discussion and consensus around areas of common difficulty.

We calculated an SES-CD metric that reflects the confidence that can be placed in any single read made by a reader randomly selected from the pool. We also demonstrate how the pool can be added to and monitored for quality assurance over time. There are arguments in favour of using more than a single reader – either using statistical methods to combine reader scores and detect outliers or via methods of arbitration.[24,25] Whilst this may lead to further increases in confidence in scores, and is supported within our platform for future users who may require this level of confidence, this is not an absolute requirement of our system. Importantly, a minimum percentage of reads must be double-read to support ongoing quality assurance of the reader pool – a process that is implemented in our platform in a blinded manner.

Whilst we have developed and validated our system to work with SES-CD scores, the methodological principles we set out here might be equally applicable to other scoring systems where the outcome can be handled as a continuous variable. These include not only endoscopic scores in UC, but also other systems where human raters are asked to assess a visual image against a validated scoring system, such as occurs, for example, in the scoring of MRI images. Nevertheless, other scoring systems might differ in complexity, require more extensive training and necessitate much larger image sets and more iterations of the scoring and training process during any validation exercise.

This exercise shows that it is feasible to establish a validated pool of central readers to support clinical trials using individuals with previous experience of use of the scoring system. This pool of readers and the supporting platform is available to support investigator-initiated clinical

trials, on a not-for-profit basis. Our methodology is reproducible by other investigators seeking to establish their own validated reader pools.

Patient and public involvement

No patient or public involvement to report.

Conflict of interest statement:

The authors declare no relevant conflicts of interest.

Funding:

No funding received. SD acknowledges the support of NHS Research Scotland through NHS Lothian. JOL acknowledges the support of the NIHR EME programme.

References

- 1 Mary JY, Modigliani R. Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). *Gut* Published Online First: 1989. doi:10.1136/gut.30.7.983
- 2 Daperno M, D'Haens G, Van Assche G, *et al.* Development and validation of a new, simplified endoscopic activity score for Crohn's disease: The SES-CD. *Gastrointest Endosc* Published Online First: 2004. doi:10.1016/S0016-5107(04)01878-4
- 3 Schroeder KW, Tremaine WJ, Ilstrup DM. Coated Oral 5-Aminosalicylic Acid Therapy for Mildly to Moderately Active Ulcerative Colitis. *N Engl J Med* Published Online First: 1987. doi:10.1056/nejm198712243172603
- 4 Travis SPL, Schnell D, Krzeski P, *et al.* Reliability and initial validation of the ulcerative colitis endoscopic index of severity. *Gastroenterology* Published Online First: 2013.

- doi:10.1053/j.gastro.2013.07.024
- 5 Frøslie KF, Jahnsen J, Moum BA, *et al.* Mucosal Healing in Inflammatory Bowel Disease: Results From a Norwegian Population-Based Cohort. *Gastroenterology* Published Online First: 2007. doi:10.1053/j.gastro.2007.05.051
 - 6 Shah SC, Colombel JF, Sands BE, *et al.* Systematic review with meta-analysis: Mucosal healing is associated with improved long-term outcomes in Crohn's disease. *Aliment. Pharmacol. Ther.* 2016. doi:10.1111/apt.13475
 - 7 Turner D, Ricciuto A, Lewis A, *et al.* STRIDE-II: An Update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. *Gastroenterology* Published Online First: 2021. doi:10.1053/j.gastro.2020.12.031
 - 8 Reinisch W, Gottlieb K, Colombel JF, *et al.* Comparison of the EMA and FDA Guidelines on Ulcerative Colitis Drug Development. *Clin. Gastroenterol. Hepatol.* 2019. doi:10.1016/j.cgh.2018.10.032
 - 9 Feagan BG, Sands BE, Rossiter G, *et al.* Effects of Mongersen (GED-0301) on Endoscopic and Clinical Outcomes in Patients With Active Crohn's Disease. *Gastroenterology* Published Online First: 2018. doi:10.1053/j.gastro.2017.08.035
 - 10 Naftali T, Schleider LBL, Benjaminov FS, *et al.* Cannabis is associated with clinical but not endoscopic remission in ulcerative colitis: A randomized controlled trial. *PLoS One* Published Online First: 2021. doi:10.1371/journal.pone.0246871
 - 11 Graham DY, Naser SA, Offman E, *et al.* 643 RHB-104, a Fixed-Dose, Oral Antibiotic Combination Against Mycobacterium Avium Paratuberculosis (MAP) Infection, Is Effective in Moderately to Severely Active Crohn's Disease. *Am J Gastroenterol*

- Published Online First: 2019. doi:10.14309/01.ajg.0000592108.53051.68
- 12 Reinisch W, De Villiers W, Bene L, *et al.* Fontolizumab in moderate to severe Crohn's disease: A phase 2, randomized, double-blind, placebo-controlled, multiple-dose study. *Inflamm Bowel Dis* Published Online First: 2010. doi:10.1002/ibd.21038
 - 13 Harris MS, Wichary J, Zadnik M, *et al.* Competition for Clinical Trials in Inflammatory Bowel Diseases. *Gastroenterology*. 2019. doi:10.1053/j.gastro.2019.08.020
 - 14 Hébuterne X, Lémann M, Bouhnik Y, *et al.* Endoscopic improvement of mucosal lesions in patients with moderate to severe ileocolonic Crohn's disease following treatment with certolizumab pegol. *Gut* Published Online First: 2013. doi:10.1136/gutjnl-2012-302262
 - 15 Feagan BG, Sandborn WJ, D'Haens G, *et al.* The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* Published Online First: 2013. doi:10.1053/j.gastro.2013.03.025
 - 16 Sandborn WJ, Bressler B, Lee S, *et al.* PTG-100, an oral gut-restricted peptide α 4 β 7 antagonist, induces clinical and histologic remission in patients with moderate to severely active ulcerative colitis. *United Eur Gastroenterol J* 2018.
 - 17 Hawkey CJ, Allez M, Clark MM, *et al.* Autologous hematopoietic stem cell transplantation for refractory Crohn disease a randomized clinical trial. *JAMA - J Am Med Assoc* Published Online First: 2015. doi:10.1001/jama.2015.16700
 - 18 Khanna R, Zou G, D'Haens G, *et al.* Reliability among central readers in the evaluation of endoscopic findings from patients with Crohn's disease. *Gut* Published Online First: 2016. doi:10.1136/gutjnl-2014-308973
 - 19 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability.1. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*

- 1979, 86:420–8. *Psychol Bull* 1979.
- 20 Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* Published Online First: 2016. doi:10.1016/j.jcm.2016.02.012
- 21 R Core Team (2020). R: A language and environment for statistical computing. R A Lang. Environ. Stat. Comput. R Found. Stat. Comput. Vienna, Austria. 2020.
- 22 Cicchetti D V. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychol Assess* Published Online First: 1994. doi:10.1037/1040-3590.6.4.284
- 23 Daperno M, Comberlato M, Bossa F, *et al.* Training Programs on Endoscopic Scoring Systems for Inflammatory Bowel Disease Lead to a Significant Increase in Interobserver Agreement Among Community Gastroenterologists. *J Crohns Colitis* Published Online First: 2017. doi:10.1093/ecco-jcc/jjw181
- 24 Panés J, Feagan BG, Hussain F, *et al.* Central endoscopy reading in inflammatory bowel diseases. *J Crohn's Colitis* Published Online First: 2016. doi:10.1093/ecco-jcc/jjv171
- 25 Gottlieb K, Daperno M, Usiskin K, *et al.* Endoscopy and central reading in inflammatory bowel disease clinical trials: Achievements, challenges and future developments. *Gut*. 2021. doi:10.1136/gutjnl-2020-320690

Tables

Table 1

Expected ICC	Readers (k)	N (LCI 0.5)	N (LCI 0.55)	N (LCI 0.6)	N (LCI 0.65)	N (LCI 0.7)
0.70	4	17	29	56	222	-
	5	15	26	50	200	-
	6	14	24	47	187	-
	7	13	23	45	178	-
	8	13	22	43	172	-
	9	12	22	42	168	-
	10	12	21	41	164	-
0.75	4	9	14	23	43	170
	5	8	13	21	39	154
	6	8	12	20	37	145
	7	7	11	19	35	139
	8	7	11	18	34	134
	9	7	10	18	33	131
	10	7	10	17	32	129
0.80	4	6	8	11	17	40
	5	5	7	10	15	36
	6	5	6	9	14	34
	7	5	6	9	14	33
	8	4	6	8	13	32
	9	4	6	8	13	31
	10	4	5	8	13	31
0.85	4	4	5	6	9	11
	5	-	4	6	8	10
	6	-	4	5	7	10
	7	-	-	5	7	9
	8	-	-	5	7	9
	9	-	-	4	6	9
	10	-	-	4	6	8

Number of videos (N) required for k readers scoring for expected ICC's in the range 0.70 to 0.85. Sample size based on a two-way mixed effects model. LCI: Lower limit of 2-way 95% Confidence interval - i.e. value below which the true ICC of the sample of readers is <2.5% likely to fall.

Table 2

Characteristics of the 20 videos forming the validation set as determined by median scores from the 7 original readers.

	Median (range)
Affected surface area	2.4 (0-9.6)
Ulcerated surface area	1.4 (0-6.3)
Ulcer size	3.1 (0-7)
Strictures	0.1 (0-2.1)
Total	7.1 (0-22)

Licence

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in Gut and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.