Virchows Archiv (2019) 475:573–578
https://doi.org/10.1007/s00428-019-02629-2

ORIGINAL ARTICLE

Effects of subspecialty signout and group consensus on the diagnosis of microscopic colitis

Meenal Sharma[1] · Christa L. Whitney-Miller[1] · Michael G. Drage[1] · Aaron R. Huber[1] · Raul S. Gonzalez[1,2]

Received: 4 April 2019 / Revised: 10 July 2019 / Accepted: 17 July 2019 / Published online: 29 July 2019

Abstract

Microscopic colitis (MC) includes lymphocytic colitis (LC) and collagenous colitis (CC). Microscopic changes are required to establish these diagnoses. While criteria exist, interobserver variability has been reported previously. This has not been evaluated in the context of subspecialty signout (SSSO) or a consensus conference. We identified 133 colon biopsies diagnosed as LC, CC, MC, or normal but with mild changes insufficient for MC. All predated the introduction of SSSO at our institution. They were independently reviewed by three gastrointestinal (GI) pathologists. Cases lacking independent consensus were reviewed by the same pathologists in consensus conference to establish a final diagnosis. Individual diagnoses were compared with the consensus diagnoses, and consensus diagnoses were compared with original diagnoses made by GI and non-GI pathologists. Consensus diagnoses were normal ($n = 34$), LC ($n = 57$), and CC ($n = 42$). "Normal" was the diagnosis most commonly agreed upon independently (27/34 cases, $P = 0.0073$ versus LC, $P = 0.0172$ versus CC). The reviewing pathologists independently agreed with 80%, 80%, and 94% of consensus diagnoses ($\kappa = 0.70$, 0.69, and 0.91). The group consensus agreed with the diagnoses in 49 of 58 (84%) cases originally signed out by non-GI pathologists ($\kappa = 0.77$) and in 44 of 57 (77%) cases originally signed out by GI pathologists ($\kappa = 0.63$). Good interobserver agreement exists for MC, though whether GI subspecialty training improves agreement remains unclear. Group consensus may aid in diagnosis of difficult/borderline MC cases.

Keywords Microscopic colitis · Lymphocytic colitis · Collagenous colitis · Interobserver variability · Subspecialty signout

Microscopic colitis (MC) is a common cause of chronic watery diarrhea that does not cause macroscopic changes in the colon but can be diagnosed based on histopathologic findings [1–4]. The incidence has been steadily rising and has been reported as up to 19 per 100,000 persons per year [5]. The increasing incidence may be due to awareness of the disease category and an increase in biopsy volume.

The two diseases established under the banner of MC are lymphocytic colitis (LC) and collagenous colitis (CC). These entities demonstrate some clinical and histopathologic overlap but can usually be distinguished from one another [6, 7]. LC can be diagnosed when a colon biopsy shows an expanded lamina propria accompanied by more than 20 intraepithelial lymphocytes (IELs) per 100 epithelial cells [8]. (Normal colonic mucosa should have fewer than 5 IELs per 100 epithelial cells.) CC can be diagnosed in the presence of lamina propria expansion and a subepithelial collagen band greater than 10 μm in thickness. (Normal colonic mucosa should have a band 5–7 μm thick.) The thickened collagen band may have entrapped capillaries, red blood cells, and inflammatory cells. The surface epithelium may focally strip off and may show an increase in IELs, though the increase is usually lower than that seen in LC [8].

While the above criteria are generally accepted, including by the European Crohn's and Colitis Organisation and the European Society of Pathology [9], other standards have been proposed. For example, some authors utilize a cutoff of 15 IELs per 100 epithelial cells for a diagnosis of LC, rather than 20 [10]. Additionally, cases with a minor degree of the above changes that do not fulfill criteria for MC are sometimes labeled as "incomplete MC"; these patients are sometimes clinically indistinguishable from patients with MC, suggesting that diagnostic criteria may need to be adjusted [11]. Perhaps

Dr. Gonzalez is currently affiliated with Beth Israel Deaconess Medical Center.

✉ Raul S. Gonzalez
rgonzal5@bidmc.harvard.edu

[1] Department of Pathology and Laboratory Medicine, University of Rochester Medical Center, Rochester, NY, USA

[2] Department of Pathology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, USA

due to these issues, there is some degree of interobserver and intraobserver variability in the diagnosis of both LC and CC, suggesting subjectivity [12, 13]. This has been previously explored by comparing the performance of individual pathologists, but it has not been evaluated in the context of subspecialty signout (SSSO) or a group consensus model. We therefore undertook this study to determine whether either of these factors affects pathologist variability in the diagnosis of MC, including agreement by pathologists with subspecialty gastrointestinal (GI) training with diagnoses originally rendered by GI and non-GI pathologists.

Materials and methods

Case information

With Institutional Research Board approval, we searched our archives for cases of colon biopsies that were either diagnosed as LC, CC, or MC, or that demonstrated mild changes insufficient for one of those three diagnoses. In order to create a study set of manageable size, two time periods were arbitrarily chosen and searched: cases from 2007 and cases from 2012–2013. We identified 133 specimens that met the above criteria and that had available H&E slides. For each case, the original diagnosis was recorded, and it was also noted whether the signout pathologist had subspecialty training in GI pathology. All cases predated SSSO at our institution, meaning some had been signed out by pathologists with GI fellowship training and some had not. Clinical history was available for 91 patients, among whom 84 (92%) had both a history of diarrhea and essentially normal colonic mucosa by colonoscopy (excluding diverticulosis and polyps). These 84 cases were originally diagnosed as follows: 17 normal, 28 LC, 31 CC, 8 MC; the other 7 were diagnosed as having MC (3), LC (2), CC (1), and normal mucosa (1). Among the 84 patients with diarrhea and normal colonoscopy, 40 had symptomatic improvement after treatment for MC, 20 had symptomatic resolution without treatment, 1 had persistent symptoms despite treatment, and the management-related outcome for the remaining 23 was unclear based on available subsequent medical records.

The original H&E slides from all 133 cases were independently reviewed by three pathologists with GI subspecialty training, who rendered a diagnosis of LC, CC, or normal. The three pathologists were not involved in identifying or compiling the cases, they were not given any clinical information about the cases, they did not know the originally rendered diagnoses, and they did not meet beforehand to discuss diagnostic criteria for LC or CC. They had been in practice for 10, 3, and 1 year(s) after completing their subspecialty fellowship, respectively; the three completed GI fellowships at three different institutions. After the review was completed, cases given the same diagnosis by all three pathologists were set aside. The remaining cases (which lacked independent consensus) were then reviewed simultaneously at a multi-headed microscope by the same three pathologists, who discussed each case and established a final consensus diagnosis. The "final" diagnoses for each case (whether achieved by independent agreement or by in-person consensus) were recorded. Ancillary stains were not used for any case, even if they were ordered for signout purposes originally; as most cases were originally diagnosed using only H&E slides, such additional stains were excluded if available in order to grant uniformity to the study material, though the existence of additional stains was noted.

Statistical analysis

Fleiss's kappa (unweighted) was used to compare the consensus diagnoses to each GI pathologist's individual interpretation. It was also used to compare agreement between original and consensus diagnoses for cases originally signed out by GI pathologists, and the same for cases originally signed out by non-GI pathologists. Fisher's exact test was used to compare the percentages of independent consensus among the three diagnostic groups, to compare the cases originally diagnosed by GI pathologists and those originally diagnosed by non-GI pathologists in order to determine whether they were similar in composition, and to compare the rate of usage of the less-specific diagnosis of MC between the 2007 cases and the 2012–2013 cases. All statistics were calculated using GraphPad Software online (https://graphpad.com/quickcalcs, GraphPad Software, San Diego, CA, USA, last accessed 06/24/2019). Statistical significance was set at $P < 0.05$.

Results

Diagnostic results are summarized in Table 1. The final consensus diagnoses for the 133 cases were normal ($n = 34$), LC ($n = 57$; Fig. 1a, b), and CC ($n = 42$; Fig. 1c, d). Of these, "normal" was the diagnosis most commonly agreed upon by the three reviewing pathologists prior to consensus (27 of 34 cases, 79%; $P = 0.0073$ versus LC and $P = 0.0172$ versus CC, Fisher's exact test on 2×2 contingency tables). LC was agreed upon prior to consensus in 28 of 57 cases (49%), and CC in 22 of 42 cases (52%; $P = 0.840$ for LC versus CC). There was a single case where the three pathologists independently rendered three different diagnoses (Fig. 1e, h).

The three reviewing GI pathologists with 10, 3, and 1 year(s) of post-fellowship experience independently agreed with 80%, 94%, and 80% of the final consensus diagnoses ($\kappa = 0.70$, 0.91, and 0.69, respectively, with 95% confidence intervals (CIs) of 0.61–0.80, 0.84–0.97, and 0.59–0.80).

As above, the 133 biopsies in this set were selected for inclusion based on signout diagnosis, without regard to whether the original pathologist had GI subspecialist training.

Table 1 Original and consensus diagnostic breakdown of colon biopsies

	Normal	Lymphocytic colitis	Collagenous colitis	Microscopic colitis
Overall cohort ($n = 133$)				
Original diagnoses	24 (18%)	47 (35%)	44 (33%)	18 (14%)
Final group consensus diagnosis	34 (26%)	57 (43%)	42 (32%)	n/a
Cases signed out by non-GI pathologists ($n = 71$)				
Original diagnoses	16 (23%)	24 (34%)	18 (25%)	13 (18%)
Final group consensus diagnoses	23 (32%)	29 (41%)	19 (27%)	n/a
Cases signed out by GI pathologists ($n = 62$)				
Original diagnoses	8 (13%)	23 (37%)	26 (42%)	5 (8%)
Final group consensus diagnoses	11 (18%)	28 (45%)	23 (37%)	n/a

There were 62 cases originally diagnosed by GI pathologists, and 71 diagnosed by non-GI pathologists. These two populations appeared potentially dissimilar in regard to the final consensus diagnoses rendered by the group; the former included 11/62 normal cases (18%), 28/62 LC cases (45%), and 23/62 CC cases (37%), while the latter included 23/71 normal cases (32%), 29/71 LC cases (41%), and 19/71 CC cases (27%). The group originally signed out by GI pathologists therefore had a larger proportion of cases deemed pathologic and a larger proportion of cases deemed LC, which had the lowest rate of agreement pre-consensus conference. However, this difference was not statistically significant ($P = 0.13$). Trichrome staining had been ordered in 7 cases (11%) originally signed out by GI pathologists (including the case that received three different independent diagnoses in this study) and 6 cases (8%) originally signed out by non-GI pathologists.

Among 58 cases originally signed out by non-GI pathologists, the original diagnosis and the group consensus diagnosis were in agreement in 49 ($\kappa = 0.77$, 95% CI = 0.63–0.91); among 57 cases originally signed out by GI pathologists, agreement was present in 44 ($\kappa = 0.63$, 95% CI = 0.46–0.80). Eighteen cases originally signed out as MC were not included in these comparisons, as this diagnosis was not utilized by the reviewing group, and an original diagnosis of MC therefore could not be faithfully matched to a diagnosis of LC or CC.

The statistical analysis was re-run with the LC and CC groups combined into one MC group; the individual calculated κ values all remained fairly similar and did not change in magnitude value (e.g., all κ values considered in the "substantial agreement" range stayed within that range). Using this "normal/MC" dichotomy, the cases originally signed out by GI pathologists were still not significantly different diagnostically from those originally signed out by non-GI pathologists ($P = 0.073$). Also using this dichotomy, among 71 cases originally signed out by non-GI pathologists, the original diagnosis and the group consensus diagnosis were in agreement in 64 ($\kappa = 0.76$, 95% CI = 0.59–0.92); among 62 cases originally signed out by GI pathologists, agreement was present in 53 ($\kappa = 0.44$, 95% CI = 0.14–0.75).

Discussion

The diagnosis of MC relies on a good biopsy and an accurate assessment of the histological alterations. Specific criteria for LC and CC have been endorsed, and many examples of these diseases exceed the minimum criteria. The diagnosis is therefore straightforward in many cases; however, difficulty may arise when cases approach or just barely exceed minimum criteria, or criteria are not strictly employed. Accordingly, Olesen et al. [14] reported that among a cohort of 1018 patients, 97 patients had MC and only 67 of them (69%) were correctly diagnosed at the time of original review using proper histologic criteria (though the qualifications of the original and reviewing pathologists were not specified). Bjørnbak et al. [1] reassessed biopsies from a number of patients with MC and incomplete MC and found that 15% of the patients with incomplete MC in fact met published criteria for MC.

Potentially less concerning is when a case resembles LC but has a somewhat thickened collagen table, suggesting the possibility of CC. This was also assessed in the study by Bjørnbak et al. [1], which found that 48% of CC cases had a mild increase in IELs and 24% of LC cases had a slightly thickened collagen layer (> 5 μm but < 10 μm). Additionally, Vigren et al. [15] retrospectively identified 664 patients with a diagnosis of CC or LC; upon reevaluation of the biopsies, three diagnoses were changed from CC to LC and six from LC to CC.

Other diagnoses can enter the histologic differential, such as acute self-limited colitis, inflammatory bowel disease, and medication-induced colitis, but these almost always create macroscopically visible colitis, which allows for distinction from MC when clinical history is available.

Our study found that, while group consensus allows for discussion of difficult cases and agreement over a final diagnosis, the individual group members agreed with the consensus opinion most of the time (with κ scores ranging from 0.69 to 0.91). Magnitude scoring is somewhat arbitrary with regard to the kappa statistic, but these κ scores would generally be considered "substantial" to "near-perfect" agreement [16].

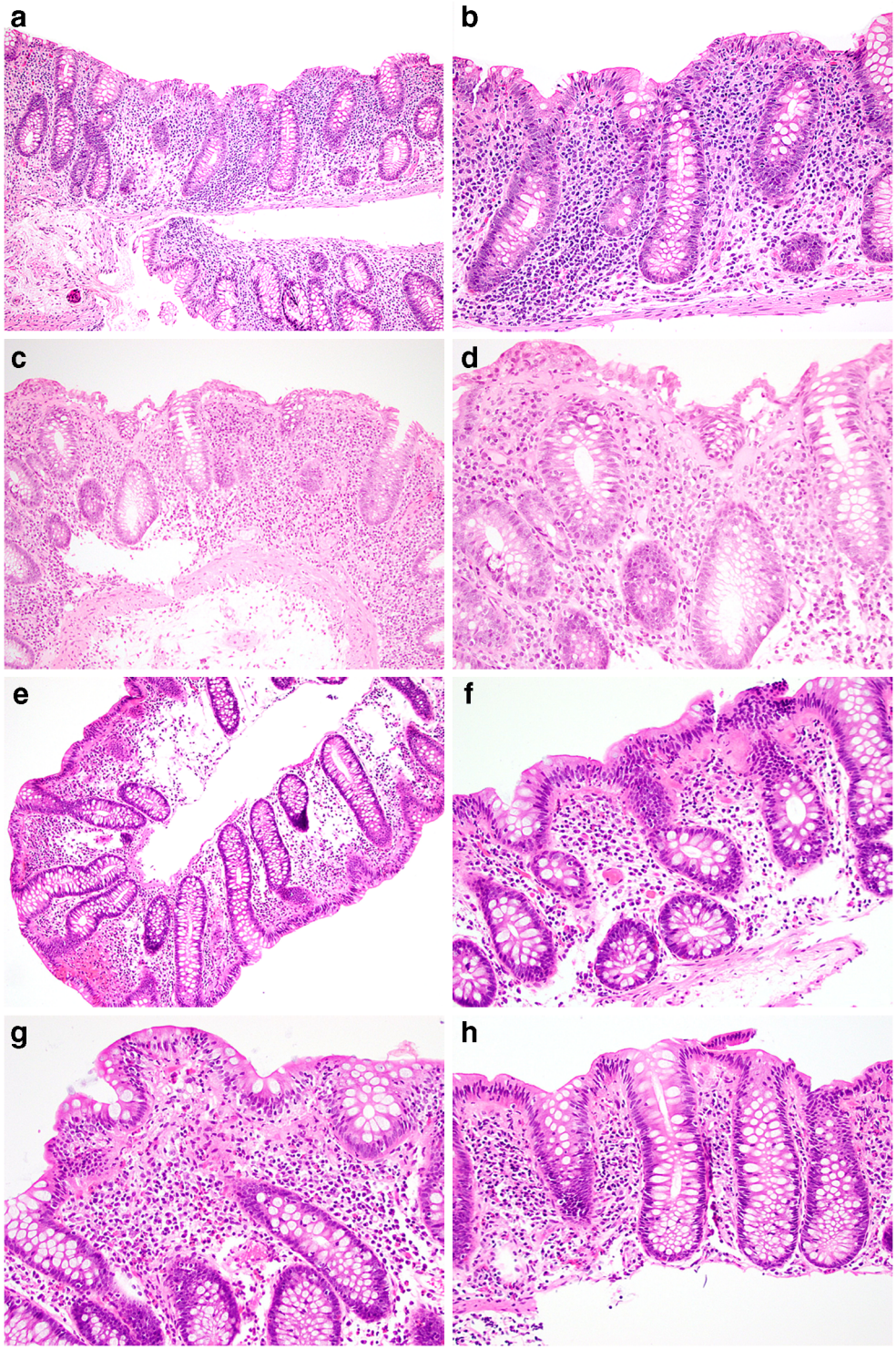

Fig. 1 **a**, **b** This colon biopsy was independently diagnosed as lymphocytic colitis by all three gastrointestinal pathologists (original magnification × 100 and × 200). **c**, **d** This colon biopsy was independently diagnosed as collagenous colitis by all three gastrointestinal pathologists (original magnification × 100 and × 200). **e**, **h** This colon biopsy was diagnosed differently by all three pathologists (i.e., it was called normal, lymphocytic colitis, and collagenous colitis) (original magnification × 100 and × 200). The final diagnosis after group consensus was collagenous colitis

Our study was not designed to determine overall accuracy of original or study-related diagnoses, given that a gold standard for LC and CC is difficult to establish, particularly in a setting such as ours where borderline cases were included; many patients in this study improved following treatment, but many also experienced resolution of symptoms without therapy, which can occur in MC.

Two previous studies have addressed the reproducibility of histopathological diagnosis of MC, with results similar to ours. Limsui et al. [12] had four GI pathologists twice review 90 colon biopsies, including 20 LC and 20 CC cases, and classify them into one of six categories. Overall interobserver agreement for the cases was 69% ($\kappa = 0.76$, 95% CI 0.69–0.83) and 70% ($\kappa = 0.71$, 95% CI 0.61–0.79), and

interobserver agreement for MC versus not MC was 91% (κ = 0.90, 95% CI 0.82–0.96) and 88% (κ = 0.83, 95% CI 0.73–0.92). The study also found good intraobserver agreement (κ = 0.89 for MC versus not MC), which we did not evaluate.

Fiehn et al. [13] had two GI pathologists and one pathology trainee twice review 125 biopsies (25 each of LC, CC, incomplete MC, inflammatory bowel disease, and normal). Interobserver agreement was 59% (κ = 0.64, 95% CI 0.59–0.69) and 67% (κ = 0.70, 95% CI 0.65–0.75) when allocating the cases into five diagnostic categories, and it improved when combining LC, CC, and incomplete MC into one category. Intraobserver agreement was generally also high (κ = 0.96, 0.88, and 0.92) when combining LC, CC, and incomplete MC. This study may have been affected by the participating pathologists knowing there were 25 cases for each diagnosis.

We note that GI-trained pathologists in particular usually attempt to utilize existing criteria in order to make a diagnosis of either LC or CC, rather than a less committal diagnosis of MC. For this reason, more cases were originally signed out as MC by non-GI pathologists than by GI pathologists. The use of MC as a top-line diagnosis may have also been influenced by the time period in which the cases were signed out, as 13 of the 18 "MC" cases (72%) were signed out in 2007, while the 2007 cases represented 69 of the 133 cases in the study (52%) (P = 0.078).

To our knowledge, MC has not previously been studied in the context of a subspecialty group consensus model, which we used to create a "final" diagnosis in cases without unanimous diagnostic agreement. Upon comparing originally rendered diagnoses with the final diagnosis determined in our study, cases signed out by non-GI pathologists agreed 84% of the time (κ = 0.77), while cases signed out by GI pathologists agreed only 77% of the time (κ = 0.63). This outcome may be partially explained by the fact that the two groups were somewhat dissimilar; after recording the final diagnoses and comparing the two groups, the cases signed out by non-GI pathologists had a higher proportion of normal cases and a lower proportion of LC cases, which may have potentially made the set easier. This is a limitation of our study that could not be controlled for, as the cases were selected prior to the establishment of the final diagnoses, and difficult cases (i.e., those that approached and perhaps barely exceeded criteria for LC or CC) were purposefully included. We also specifically chose not to utilize ancillary staining, in order to allow direct comparisons of the H&E appearance for each case. In routine practice, a trichrome histochemical stain or tenascin immunohistochemical stain can facilitate the diagnosis of CC [17–19], and immunohistochemical staining for CD3 and CD8 may potentially be of some value in the diagnosis of LC [20, 21].

In the past several years, SSSO has been adopted in many academic and community pathology practices [22]. It ostensibly provides superior patient care, as (for example) all GI pathology cases are signed out by pathologists who largely or only deal with GI material, making them more skilled at the subject matter. It may also facilitate group consensus conferences, wherein all the pathologists in one subspecialty convene to discuss difficult cases, settling upon a final opinion. Our study was not able to assert that GI pathologists more reproducibly diagnosed LC or CC than non-GI pathologists, perhaps due to differences in the two study populations or perhaps due to overall pathologist acumen (as subspecialty training can increase familiarity with material but may not replace some aspects of innate ability). It is also possible that, prior to SSSO at our institution, cases perceived as difficult were transferred from a non-GI pathologist to a GI pathologist, which may account for the higher rate of diagnosis of MC by GI pathologists. Additionally, our findings may indicate that gestalt is sometimes relied upon for the diagnosis of LC and CC, particularly since definitional criteria may still vary among pathologists, and/or that the diagnostic criteria for LC and CC are not sufficiently emphasized during GI fellowship training. Our study was not designed to investigate techniques that would improve interobserver agreement, but it suggests that a relatively low threshold for second/consensus opinion and immunohistochemical or special staining would be of diagnostic benefit.

Our study confirmed that pathologists generally have good interobserver agreement when diagnosing MC. It remains unclear whether GI pathologists are truly superior at these distinctions compared with non-GI pathologists, and the degree to which factors such as adherence to (variably defined) criteria, reliance on ancillary stains, and natural aptitude influence interobserver variability. Given that the three GI pathologists in our study often but not always agreed with the final group consensus diagnoses, we advocate showing difficult or unclear cases at a consensus conference; this may increase diagnostic accuracy and also may decrease the need for ancillary staining.

Contributions MS screened the slides, collected data, and co-wrote the manuscript. CLW, MGD, and ARH performed independent and consensus review of the slides. RSG designed the study, screened selected slides, analyzed the data, and co-wrote the manuscript.

Compliance with ethical standards This research was performed with appropriate Institutional Research Board approval. Human participants were not recruited. A waiver of HIPAA authorization was obtained.

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Bjornbak C, Engel PJ, Nielsen PL, Munck LK (2011) Microscopic colitis: clinical findings, topography and persistence of histopathological subgroups. Aliment Pharmacol Ther 34:1225–1234
2. Munch A, Aust D, Bohr J et al (2012) Microscopic colitis: current status, present and future challenges: statements of the European Microscopic Colitis Group. J Crohns Colitis 6:932–945

3. Nyhlin N, Bohr J, Eriksson S, Tysk C (2008) Microscopic colitis: a common and an easily overlooked cause of chronic diarrhoea. Eur J Intern Med 19:181–186
4. Tangri V, Chande N (2009) Microscopic colitis: an update. J Clin Gastroenterol 43:293–296
5. Fernández-Bañares F, Casanova MJ, Arguedas Y, Beltrán B, Busquets D, Fernández JM, Fernández-Salazar L, García-Planella E, Guagnozzi D, Lucendo AJ, Manceñido N, Marín-Jiménez I, Montoro M, Piqueras M, Robles V, Ruiz-Cerulla A, Gisbert JP, the Spanish Microscopic Colitis Group (SMCG) (2016) Current concepts on microscopic colitis: evidence-based statements and recommendations of the Spanish Microscopic Colitis Group. Aliment Pharmacol Ther 43:400–426
6. Lazenby AJ, Yardley JH, Giardiello FM, Jessurun J, Bayless TM (1989) Lymphocytic ("microscopic") colitis: a comparative histopathologic study with particular reference to collagenous colitis. Hum Pathol 20:18–28
7. Baert F, Wouters K, D'Haens G et al (1999) Lymphocytic colitis: a distinct clinical entity? A clinicopathological confrontation of lymphocytic and collagenous colitis. Gut 45:375–381
8. Langner C, Aust D, Ensari A, Villanacci V, Becheanu G, Miehlke S, Geboes K, Münch A, the Working Group of Digestive Diseases of the European Society of Pathology (ESP) and the European Microscopic Colitis Group (EMCG) (2015) Histology of microscopic colitis-review with a practical approach for pathologists. Histopathology 66:613–626
9. Magro F, Langner C, Driessen A, Ensari A, Geboes K, Mantzaris GJ, Villanacci V, Becheanu G, Nunes PB, Cathomas G, Fries W, Jouret-Mourin A, Mescoli C, de Petris G, Rubio CA, Shepherd NA, Vieth M, Eliakim R (2013) European consensus on the histopathology of inflammatory bowel disease. J Crohns Colitis 7:827–851
10. Geboes K (2008) Lymphocytic, collagenous and other microscopic colitides: pathology and the relationship with idiopathic inflammatory bowel diseases. Gastroenterol Clin Biol 32:689–694
11. Guagnozzi D, Landolfi S, Vicario M (2016) Towards a new paradigm of microscopic colitis: incomplete and variant forms. World J Gastroenterol 22:8459–8471
12. Limsui D, Pardi DS, Smyrk TC, Abraham SC, Lewis JT, Sanderson SO, Kammer PP, Dierkhising RA, Zinsmeister AR (2009) Observer variability in the histologic diagnosis of microscopic colitis. Inflamm Bowel Dis 15:35–38
13. Fiehn AM, Bjørnbak C, Warnecke M, Engel PJ, Munck LK (2013) Observer variability in the histopathologic diagnosis of microscopic colitis and subgroups. Hum Pathol 44:2461–2466
14. Olesen M, Eriksson S, Bohr J, Järnerot G, Tysk C (2004) Microscopic colitis: a common diarrheal disease. An epidemiological study in Orebro, Sweden, 1993-1998. Gut 53:346–350
15. Vigren L, Olesen M, Benoni C, Sjöberg K (2012) Are collagenous and lymphocytic colitis different aspects of the same disease? Scand J Gastroenterol 47:1448–1453
16. Milner DA Jr, Meserve EEK, Rinda Soong T, Mata DA (2017) Statistics for pathologists. Demos Medical, New York, pp 83–84
17. Fernandez-Banares F, Salas A, Esteve M (2008) Pitfalls and errors in the diagnosis of collagenous and lymphocytic colitis. J Crohns Colitis 2:343–347
18. Salas A, Fernandez-Banares F, Casalots J, Gonzalez C, Tarroch X, Forcada P, Gonzalez G (2003) Subepithelial myofibroblasts and tenascin expression in microscopic colitis. Histopathology 43:48–54
19. Anagnostopoulos I, Schuppan D, Riecken EO, Gross UM, Stein H (1999) Tenascin labelling in colorectal biopsies: a useful marker in the diagnosis of collagenous colitis. Histopathology 34:425–431
20. Mohamed N, Marais M, Bezuidenhout J (2011) Microscopic colitis as a missed cause of chronic diarrhea. World J Gastroenterol 17:1996–2002
21. Tysk C, Bohr J, Nyhlin N, Wickbom A, Eriksson S (2008) Diagnosis and management of microscopic colitis. World J Gastroenterol 14:7280–7288
22. Sarewitz SJ (2014) Subspecialization in community pathology practice. Arch Pathol Lab Med 138:871–872