Rationale and objectives: Evaluation of Crohn’s disease by computed tomography enterography, magnetic resonance imaging enterography and colonoscopy is essential for disease monitoring. The aim of this study is to evaluate this exams acuity.

Materials and methods: Patients with histological diagnosis of Crohn’s disease who underwent computed tomography enterography, magnetic resonance imaging enterography and colonoscopy in the period of January 1st, 2009 and July 31st, 2016 and the realization of these exams did not exceed a time interval of 6 months was included. Sensitivity, specificity, positive and negative predictive values (PPV, NPV), Cohen’s kappa (K), agreement and disagreement were calculated.

Results: Comparing computed tomography enterography and magnetic resonance imaging enterography with colonoscopy, there was an agreement of 85.7% and a disagreement of 14.3% in Crohn’s disease overall detection, for both exams. Computed tomography enterography and colonoscopy showed greater agreement in abscesses and lumen reduction/dilatation detection (C = 95.2%) and magnetic resonance imaging enterography and colonoscopy in abscesses detection (C = 92.9%). Comparing magnetic resonance imaging enterography and computed tomography enterography, greater agreement was observed in detection of lumen reduction/dilatation (C = 96%). K showed considerable agreement in detection of mesenteric lymph nodes, fistulas, mural inflammation and thickening. The sensitivity, specificity, PPV and NPV were respectively set at 94.12% (95% CI 71.31–99.85), 50% (95% CI 6.76–93.24), 88.89% (95% CI 65.29–98.62) and 66.67% (95% CI 9.43–99.16) for CTE and 90.62% (95% CI 80.70–96.48), 33.33% (95% CI 4.33–77.72), 93.55% (95% CI 84.30–98.21) and 25% (95% CI 3.19–65.09) for MRIE.
Conclusions: Although computed tomography enterography presents better sensitivity and specificity than magnetic resonance imaging enterography, both present high agreement values in detection of characteristic Crohn’s disease findings, therefore the selection of the best test to monitor Crohn’s disease should take into account aspects such as age, tolerability, disease phenotype and resources availability.

Enterografia por tomografia computadorizada ou por ressonância magnética na doença de Crohn - qual escolher?

RESUMO

Fundamentação e objetivos: A avaliação da doença de Crohn por enterografia por tomografia computadorizada, enterografia por ressonância magnética e colonoscopia é essencial para o monitoramento da doença. Este estudo teve como objetivo avaliar a acuidade desses exames.

Materiais e métodos: O estudo incluiu pacientes com diagnóstico histológico de doença de Crohn submetidos à enterografia por tomografia computadorizada, enterografia por ressonância magnética e colonoscopia no período entre 1º de janeiro de 2009 e 31 de julho de 2016; os exames foram realizados em um intervalo de máximo de seis meses. Calculou-se a sensibilidade, especificidade, valores preditivos positivo e negativo (VPP, VPN), Kappa (K) de Cohen, concordância e discordância.

Resultados: Ao comparar enterografia por tomografia computadorizada e enterografia por ressonância magnética com colonoscopia, observou-se uma concordância de 85,7% e discordância de 14,3% na detecção global da doença de Crohn para ambos os exames. A enterografia por tomografia computadorizada e a colonoscopia mostraram maior concordância nos abscessos e na detecção da redução da luz (C = 95,2%) e enterografia por ressonância magnética e colonoscopia, na detecção de abscessos (C = 92,9%). Ao comparar a enterografia por ressonância magnética e a enterografia por tomografia computadorizada, observou-se maior concordância na detecção da redução/dilatação do lúmen (C = 96%). Os valores de K mostraram concordância considerável na detecção de linfonodos mesentéricos, fistulas, inflamação e espessamento mural. A sensibilidade, especificidade, VPP e VPN foram, respectivamente, 94,12% (IC 95%: 71,31 ± 99,85), 50% (IC 95%: 6,76 ± 93,24), 88,89% (IC 95%: 65,29 ± 98,62) e 66,67% (IC 95%: 9,43 ± 99,16) para ETC e 90,62% (IC 95%: 80,70 ± 96,48), 33,33% (IC 95%: 4,33 ± 77,72), 93,55% (IC 95%: 84,30 ± 98,21) e 25% (IC 95%: 3,19 ± 65,09) para enterografia por ressonância magnética.

Conclusões: Embora a enterografia por tomografia computadorizada apresente melhor sensibilidade e especificidade que a enterografia por ressonância magnética, ambas apresentam altos valores de concordância na detecção dos achados característicos da doença de Crohn. Assim, a seleção do teste mais adequado para monitorar a doença de Crohn deve levar em consideração aspectos como idade, tolerabilidade, fenótipo da doença e disponibilidade de recursos.

Introduction

Inflammatory bowel diseases encompasses a group of pathologies characterized by an inappropriate immune response to the endogenous intestinal microbial flora, with or without some auto-immunity component. With regard, specifically, to Crohn’s disease (CD), its incidence is of 0.1–16 cases per 100,000 inhabitants worldwide. In Europe, it is estimated that 12.7 cases of CD are diagnosed per 100,000 inhabitants per year. CD is defined as a chronic and idiopathic inflammatory process that can affect any segment of the gastrointestinal tract, from the mouth to the anus. The diagnosis of this disease is based on a set of data resulting from the clinical history, physical examination and complementary diagnostic tests, such as endoscopic, radiological, laboratory and histological examinations. Undoubtedly, colonoscopy has proven to be the elective exam for the diagnosis of CD because it allows a complete visualization of the entire colon, ileocecal valve and terminal ileum, which are effectively the most commonly affected anatomical regions.
With regard to investigation of small bowel disease, for decades, barium studies, such as enteroclysis and intestinal transit, were considered gold standard tests, with great impact on the diagnosis, evaluation of their anatomical distribution and the detection of fistulas, abscesses and signs of active phase or exacerbation of the disease. However, given the evolution that has been developed in the field of imaging, namely in terms of the improvement of imaging techniques, enterography, either by computed tomography (CTE) or magnetic resonance imaging (MRIE), has been supplanting techniques, previously used in the study of small intestine pathologies. Effectively, CTE has the advantages of allowing monitoring to evaluate the response to therapy. This latter aspect takes even more emphasis on young patients who will certainly need more frequent imaging. 

A retrospective study was performed. The inclusion criteria were: patients with histological diagnosis of CD who underwent CTE, MRIE and colonoscopy in the period of January 1st, 2009 and July 31st, 2016 at Braga Hospital and the performance of CTE, MRIE and colonoscopy did not exceed a time interval of 6 months. Patients with an incomplete clinical process in which it was not possible to collect data were excluded from the study.

The collected data were organized in an Excel (Microsoft Office 2010) database, and the Statistical Package for Social Sciences (SPSS) version 24.0 was used.

A descriptive analysis of the variables under study was performed to determine the means and frequencies. The sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) of CTE and MRIE were calculated in the detection of some specific CD findings, such as fistulas, abscesses, strictures, ICV alterations and mural stratification, when compared to colonoscopy results. For this purpose, the MedCalc online tool was used (available at http://www.medcalc.org/calc/diagnostic_test.php).

The values of agreement and disagreement between the three exams were calculated through the following formulas: $A = \frac{C^2}{P} \times 100$ and $D = \frac{D^2}{P^2} \times 100$.

For the agreement evaluation, the Cohen's kappa value was also calculated using the Vassar Stats online tool (available at http://vassarstats.net/kappa.html). According to the literature consulted, Cohen's kappa values between 0.01 and 0.20 show a poor agreement, between 0.21 and 0.40 a considerable agreement, between 0.41 and 0.60 a moderate agreement, between 0.61 and 0.80 a good agreement and between 0.81 and 1 an excellent agreement.

For all tests, it was assumed a significance of 0.05 and a confidence interval (CI) of 95%, calculated by the following formula: $P = 1.96 \times \sqrt{\frac{P(1-P)}{n}}$.

All the patients of the Hospital Braga provided written informed consent for surgery. The present study was also conducted in accordance with the Declaration of Helsinki (1964) and later versions and Good Clinical Practice guidelines and with the approval of the Hospital de Braga Ethics Committee (HB 32/2013) and by Ethics Subcommittee for Life and Health Sciences (SECVS 102/2016).

### Results

The study sample consisted of 159 patients, of which 49.7% were male ($n=79$) and 50.3% female ($n=80$). They are aged between 14 and 71 years, with a mean of 36.64 years. The mean age at diagnosis was 27.57 years, with the youngest age of diagnosis being 10 years and the oldest 57 years.

The most frequently observed CD phenotype was the inflammatory of ileum with a frequency of 41.5% ($n=66$), followed by the stricture of the ileum, 6% ($n=36$), inflammatory ileocolic, 17% ($n=8$), penetrating ileal CD, 2.5% ($n=4$), penetrating ileocolic CD and inflammatory colonic CD, 1.3% ($n=2$), and penetrating colic CD, 0.6% ($n=1$).

### Comparative analysis between CTE and colonoscopy findings

Of the 159 patients, only 21 patients had performed CTE and colonoscopy within the period of 6 months. Of these 21 patients, 40.9% were male ($n=9$) and 59.1% female ($n=12$). The most frequently observed phenotypes were stricture of the ileum and inflammatory of the ileum, 40.9% ($n=8$) and 22.7% ($n=5$), respectively.
The two exams presented an agreement of 85.7% (n = 18) for the detection of characteristic signs of CD, or for their exclusion and a disagreement of 14.3% (n = 3) (Table 1). The presence or absence of abscesses and lumen reduction presented the highest agreement between the two exams (95.2%) while the highest value of disagreement was found in the detection of ileocecal valve alterations (52.4%).

Comparative analysis between MRIE and colonoscopy findings

Of the 159 patients, only 70 had performed MRIE and colonoscopy in a period of 6 months, of which 37 were male (52.9%) and 33 were female (47.1%). The most frequently observed phenotype was inflammatory of the ileum, 35.7% (n = 25), followed by strictureing of the ileum, 24.3% (n = 17).

The two exams presented an agreement of 87.7% (n = 60) and a disagreement of 14.3% (n = 10) in the overall detection or exclusion of CD (Table 2). They showed greater agreement in the detection or exclusion of abscesses (κ = 0.929) and a greater disagreement in the detection of ICV alterations (D = 30%) (p = 0.078).

Comparative analysis between CTE and MRIE findings

Of the 159 patients, only 25 performed CTE and MRIE within the period of 6 months. Of these 25 patients, 40% were male (n = 10) and 60% female (n = 15). The earliest age at diagnosis was 14 years and the oldest 51 years (M = 31.56). Concerning the CD phenotype, the most common was strictureing of the ileum, 40% (n = 10), followed by inflammatory of the ileum, 24% (n = 6). Table 3 presents agreement, disagreement results and Cohen’s Kappa values between CTE and MRIE.

Analysis of sensitivity, specificity, positive predictive value and negative predictive value of CTE and MRIE

The CTE presents slightly higher values of sensitivity and specificity than MRIE, as can be seen in Table 4. On the other hand, MRIE presents higher positive predictive value in the detection of CD.

Discussion

CD is an inflammatory bowel disease of unknown etiology, whose incidence has increased over the last years, in all age groups. Colonoscopy is currently considered the gold standard for the diagnosis and monitoring of CD. However, this examination has some disadvantages such as patient intolerance or refusal and the fact that it does not allow the visualization of the majority of the small intestine, the most frequently affected segment of the digestive tract. CTE and MRIE, high resolution imaging techniques, have supplanted techniques previously used in the diagnostic

---

**Table 1 – Relation between CTE and colonoscopy in the detection or exclusion of CD characteristic findings.**

<table>
<thead>
<tr>
<th>CTE</th>
<th>Colonoscopy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Detect</td>
<td>Non-detect</td>
</tr>
<tr>
<td>Detect</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Non-detect</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>4</td>
</tr>
</tbody>
</table>

CTE, computed tomography enterography.

**Table 2 – Relation between MRIE and colonoscopy in the detection or exclusion of CD characteristic findings.**

<table>
<thead>
<tr>
<th></th>
<th>Colonoscopy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Detect</td>
<td>Non-detect</td>
</tr>
<tr>
<td>Detect</td>
<td>58</td>
<td>4</td>
</tr>
<tr>
<td>MRE</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Non-detect</td>
<td>64</td>
<td>6</td>
</tr>
</tbody>
</table>

MRE, magnetic resonance enterography.

**Table 3 – Agreement and disagreement results between CTE and MRIE and Cohen’s kappa values.**

<table>
<thead>
<tr>
<th></th>
<th>CTE = 1 MRE = 1</th>
<th>CTE = 0 MRE = 0</th>
<th>Total</th>
<th>CTE = 1 MRE = 0</th>
<th>CTE = 0 MRE = 1</th>
<th>Total</th>
<th>Cohen's kappa (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fistulas</td>
<td>8 (2)</td>
<td>76 (19)</td>
<td>84 (21)</td>
<td>8 (2)</td>
<td>8 (2)</td>
<td>16 (4)</td>
<td>0.40 (0.90–0.94)</td>
</tr>
<tr>
<td>Abscesses</td>
<td>0 (0)</td>
<td>92 (23)</td>
<td>92 (23)</td>
<td>0 (0)</td>
<td>8 (2)</td>
<td>8 (2)</td>
<td>NC</td>
</tr>
<tr>
<td>MLN</td>
<td>16 (4)</td>
<td>48 (12)</td>
<td>64 (16)</td>
<td>28 (7)</td>
<td>8 (2)</td>
<td>36 (9)</td>
<td>0.23 (0.60–0.63)</td>
</tr>
<tr>
<td>Mural inflammation</td>
<td>8 (2)</td>
<td>68 (17)</td>
<td>76 (19)</td>
<td>0 (0)</td>
<td>24 (6)</td>
<td>24 (6)</td>
<td>0.31 (0.70–0.79)</td>
</tr>
<tr>
<td>Mural thickening</td>
<td>72 (18)</td>
<td>8 (2)</td>
<td>80 (20)</td>
<td>4 (1)</td>
<td>16 (4)</td>
<td>20 (5)</td>
<td>0.34 (0.80–0.86)</td>
</tr>
<tr>
<td>ICV alterations</td>
<td>4 (1)</td>
<td>72 (18)</td>
<td>76 (19)</td>
<td>0 (0)</td>
<td>24 (6)</td>
<td>24 (6)</td>
<td>0.19 (0.76–0.76)</td>
</tr>
<tr>
<td>Strictures</td>
<td>0 (0)</td>
<td>84 (21)</td>
<td>84 (21)</td>
<td>4 (1)</td>
<td>12 (3)</td>
<td>16 (4)</td>
<td>NC</td>
</tr>
<tr>
<td>Lumen reduction</td>
<td>0 (0)</td>
<td>96 (24)</td>
<td>96 (24)</td>
<td>0 (0)</td>
<td>4 (1)</td>
<td>4 (1)</td>
<td>NC</td>
</tr>
<tr>
<td>Lumen dilation</td>
<td>8 (2)</td>
<td>72 (18)</td>
<td>80 (20)</td>
<td>4 (1)</td>
<td>16 (4)</td>
<td>20 (5)</td>
<td>0.34 (0.80–0.86)</td>
</tr>
<tr>
<td>MFD</td>
<td>28 (7)</td>
<td>44 (11)</td>
<td>72 (18)</td>
<td>20 (5)</td>
<td>8 (2)</td>
<td>28 (7)</td>
<td>0.43 (0.08–0.79)</td>
</tr>
<tr>
<td>Comb sign</td>
<td>16 (4)</td>
<td>68 (17)</td>
<td>84 (21)</td>
<td>4 (1)</td>
<td>12 (3)</td>
<td>16 (4)</td>
<td>0.57 (0.17–0.96)</td>
</tr>
</tbody>
</table>

CTE, computed tomography enterography; ICV, ileocecal valve; MLN, mesenteric lymph nodes; MFD, mesenteric fat densification; MRE, magnetic resonance enterography; NC, not calculable.
approach of small bowel pathology, since they allow a complete visualization of the intestine, without overlapping of loops, and the detection of extra-luminal findings characteristic of CD.

In this study, the acuity of these two exams was evaluated when compared to the colonoscopy, for the detection of typical CD findings.

Comparing CTE and MRIE respectively with colonoscopy, in the detection or exclusion of characteristic imaging findings of CD, an agreement of 85.7% and a disagreement of 14.3% were obtained. Pei-You et al., in a retrospective study with 60 patients, obtained lower values of agreement and disagreement, both for CTE and colonoscopy (80% and 20%, respectively), and for MRIE and colonoscopy (75% and 25%, respectively).

The observed disagreement can be explained by the fact that CD is manifested by the alternation between periods of exacerbation and remission of the disease as well as the time lag of 6 months between the two exams, although Amitai et al., rejected the existence of any impact of this time interval on the accuracy of a comparative analysis between the two examinations. Another factor that may be implicated in this difference is the presentation of CD in the chronic and acute phase. In the exacerbations of the disease, the signs typically found are: Comb sign, wall thickening (>3 mm), ulcers, fistulas, abscesses and mucosal erythema. Other signs such as densification of mesenteric fat and mesenteric lymph nodes may be present in both phases of the disease. It should be noted that some of the findings mentioned above are only detectable by one of the exams being studied, such as Comb sign and wall thickening that are only identified by MRIE and CTE. Finally, given that the disease manifests itself in such a varied way, not all patients have the same imaging findings, and therefore may only present some that are detectable only by one of the exams under study.

Comparing the CTE and colonoscopy, a higher disagreement was found in the detection or exclusion of changes in ICV (52.4%), as verified in the comparison of results between MRIE and colonoscopy (30%). These values of disagreement are expected since colonoscopy allows a direct visualization of the ICV, which can lead to more frequently detect involvement of this segment.

The agreement between CTE and MRIE was evaluated for 11 common signs of CD. We found higher agreement in the detection or exclusion of abscesses (92%) and lumen reduction and dilation (96%), while the lower agreement was found in the detection or exclusion of mesenteric lymph nodes (64%). The detection or exclusion of abscesses, changes in the ICV, lumen dilation and strictures was better by MRIE, whereas the detection or exclusion of mesenteric lymph nodes and mesenteric fat densification was better by CTE. These results are different from those found by Amitai et al. that compared ten signs of CD in 42 patients with histological diagnosis of the disease who underwent CTE and MRIE in a period of 6 months. Amitai et al. found higher agreement in the detection or exclusion of mural thickening (90.48%) and strictures (85.71%) and a lower agreement in the detection or exclusion of lymph nodes (54.76%) and lumen dilation (66.67%).

Concerning the coefficient of agreement, K, the highest values were 0.57 (95% CI 0.17–0.96) in the detection of the Comb sign, 0.43 (95% CI 0.08–0.79) in the detection of densification of mesenteric fat, revealing, according to Viera et al., a moderate agreement. Other findings showed considerable agreement, for example the detection of fistulas whose K was 0.40 (95% CI 0.0–0.94) and the detection of mural thickening and lumen dilation with a K of 0.34 (95% CI 0–0.86).

Regarding to the sensitivity and specificity of the CTE the values found were set at 94.12% and 50.00%, respectively. These results are similar to those found by Cakmakci et al., who obtained values of sensitivity and specificity of 100% and 53.9%, respectively. Although it is a very accessible test in the hospital setting and presents sensitivity for the detection of characteristic findings of CD close to 100%, CTE has some limitations, such as the use of ionizing radiation, which is of great importance in young patients and in need of frequent monitoring of the disease. In turn, MRIE presented a sensitivity and specificity of 90.62% and 33.33%, respectively. Also, these results were similar to those found by Cakmakci et al., which obtained sensitivity values of 88.9% and specificity of 38.4%.

The sample size and the retrospective nature of the study are limitations when comparing the acuity of CTE and MRIE with colonoscopy in the monitoring of CD. The temporal window between the three exams may also be a limiting factor in the study, since with CD has a dynamic course in time, the imaging findings may also be different over time. Another limitation of the study may be that the analysis of the exams was not always done by the same professional, thus constituting a bias of the observer.

### Table 4 – Characterization of CTE and MRIE regarding sensitivity, specificity, positive predictive value and negative predictive value.

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTE</td>
<td>94.12 (95% CI 71.31–99.85)</td>
<td>50.00 (95% CI 6.76–93.24)</td>
<td>88.89 (95% CI 65.29–98.62)</td>
<td>66.67 (95% CI 9.43–99.16)</td>
</tr>
<tr>
<td>MRE</td>
<td>90.62 (95% CI 80.70–96.48)</td>
<td>33.33 (95% CI 4.33–77.72)</td>
<td>93.55 (95% CI 84.30–98.21)</td>
<td>25.00 (95% CI 3.19–65.09)</td>
</tr>
</tbody>
</table>

CTE, computed tomography enterography; MRIE, magnetic resonance enterography; NPV, negative predictive value; PPV, positive predictive value.

### Conclusion

Early detection of characteristic signs of CD in active phase is essential for the institution of appropriate treatment in order to avoid the progression of the disease. The imaging evaluation of patients can be done using different techniques such as colonoscopy, CTE and MRIE, so it is pertinent to determine the acuity of these exams in the monitoring of patients with CD.

This study document that CTE presents better acuity than the MRIE in the detection of CD, when compared with the
colonoscopy, since it presents a better sensitivity and specificity (94.12% and 50.00%, respectively). It was also shown that the two exams show high agreement when evaluated for their ability to detect changes caused by CD, and that MRE was better at detecting abscess, strictures, lumen dilation and alterations of ICV, whereas CTE was better at detecting mesenteric lymph nodes and densification of the mesenteric fat.

Although CTE showed greater sensitivity and specificity (94.12% and 50.00%, respectively) than MRE (90.62% and 33.33%, respectively), the choice of the adequate exam should be weighted considering several factors such as the age of the patient, their tolerability, the CD phenotype and the availability of hospital resources.

Conflicts of interest

The authors declare no conflicts of interest.

REFERENCES