

## Case series

# What is the most reproducible histopathological classification for celiac disease? study of the interobserver variability in a serie of 69 cases



**Amal Douida<sup>1,&</sup>, Laila Tahiri<sup>1,2</sup>, Mohamed Berraho<sup>3</sup>, Hida Moustapha<sup>4</sup>, Mounia Elyousfi<sup>5</sup>, Nawal Hammam<sup>1,2</sup>, Hind El Fatemi<sup>1,2</sup>, Laila Chbani<sup>1,2</sup>**

<sup>1</sup>Department of Anatomy Pathology, University Hospital Hassan II, Fez, Morocco, <sup>2</sup>Biomedical and Translational Research Laboratory, Medical School of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco, <sup>3</sup>Laboratory of Epidemiology, Clinical Research and Community Health, Faculty of Medicine and Pharmacy, Sidi Mohammed Ben Abdellah University, Fez, Morocco, <sup>4</sup>Department of Pediatric, University hospital Hassan II, Fez, Morocco, <sup>5</sup>Department of Gastroenterology and Hepatology, University Hospital Hassan II, Fez, Morocco

<sup>&</sup>Corresponding author: Amal Douida, Department of Anatomy Pathology, University Hospital Hassan II, Fez, Morocco

Received: 29 Feb 2020 - Accepted: 14 Apr 2020 - Published: 04 May 2020

Domain: Gastroenterology

Keywords: Celiac disease, Marsh–Oberhuber classification, Cozarra classification, interobserver variability

### Abstract

The Marsh classification, as modified by Oberhuber *et al.* based on 6 stage grading, has been accepted as a worldwide standard, but the considerable number of diagnostic categories involved makes it prone to a low interobserver and intraobserver agreement, and given the therapeutic issues, a more reproducible histological classification is essential in order to better classify the villous atrophy. A more simplified classification, which is based on 3 villous morphologies and an intraepithelial lymphocyte count of >25/100 enterocytes, has been proposed by Corazza. The aim of the present study was to evaluate the interobserver agreement in classifying celiac disease lesions according to both the Marsh–Oberhuber classification and new grading system proposed by Corazza. Sixty-nine cases of villous atrophy were selected for the study. The slides of duodenal biopsy specimens were reread by two pathologists, who were not given any clinical or biological information, and had to evaluate them according to both grading systems. The kappa statistic was used to assess agreement between the two pathologists. Overall mean kappa values were 0.48 (moderate) for the Marsh–Oberhuber classification versus 0.61 (good) for the Corazza classification system. The main finding of the study was a high interobserver variability between the two pathologists when using the Marsh–Oberhuber system of grading compared to the Cozarra classification.

Case series | Volume 3, Article 2, 04 May 2020 | 10.11604/pamj-cm.2020.3.2.22064

Available online at: <https://www.clinical-medicine.panafrican-med-journal.com/content/article/3/2/full>

© Amal Douida et al PAMJ - Clinical Medicine (ISSN: 2707-2797). This is an Open Access article distributed under the terms of the Creative Commons Attribution International 4.0 License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



## Introduction

---

Celiac disease (CD) is an enteropathy caused by an immune reaction triggered by dietary gluten. Its diagnosis is based on a bundle of clinical, biological, endoscopic and histological arguments. The classification used in practice is that of Marsh-Oberhuber [1,2], but a new, more simplified classification has been proposed by Corazza [3]. So, what is the most reproducible histopathological classification for celiac disease?

## Methods

---

Sixty-nine cases of villous atrophy of various etiologies were selected for the study. At least 3 biopsies were obtained from the duodenum of each patient and properly oriented. Biopsies were fixed in formalin for at least 6 hours and at least 4 sections levels were performed for each biopsy and stained with hematoxylin-eosin. The slides of duodenal biopsy specimens were reread by two pathologists who were not given any clinical or biological information. An immunohistochemical staining with CD3 was performed to recognize cases of latent celiac disease when serology and clinical data are not conclusive or when the histological diagnosis remains equivocal. Each pathologist graded villous atrophy according to Marsh-Oberhuber [1,2] and Corazza [3] classification. The Marsh-Oberhuber classification [1,2], was based on a 6-stage grading, namely stage 1, characterized by normal mucosal architecture with an increased number of intraepithelial lymphocytes (IEL); stage 2 hyperplastic lesions, characterized by an increase in crypt depth without villous flattening; stage 3 type 3a, 3b, and 3c destructive lesion, characterized by mild, marked, and complete villous flattening, respectively; and stage 4 hypoplastic lesions, characterized by villous atrophy with normal crypt height and IEL count. The new proposed grading system [3] classified the CD lesions into

non-atrophic (grade A) and atrophic (grade B). Grade A (Figure 1) was characterized by the isolated increase of IELs ( $>25/100$  enterocytes), whereas grade B (Figure 2) was split into B1, in which the villous/crypt ratio is less than 3/1, with still detectable villi, and B2 (Figure 3), in which the villi are no longer detectable. A comparison between the Marsh-Oberhuber and the new grading criteria is shown in the diagram below (Figure 4). For statistical methods, the Cohen kappa statistic was used to assess the agreement between the two pathologists. The strength of agreement as regards the kappa values was evaluated according to Landis and Koch [4], revised by Altman [5], as follows: poor,  $<0.20$ ; fair,  $0.21-0.40$ ; moderate,  $0.41-0.60$ ; good,  $0.61-0.80$ ; and very good,  $0.81-1.00$ .

## Results

---

Sixty-nine slides with duodenal biopsies for villous atrophy of various etiologies were reread by two pathologists who were not given any clinical or biological information. Each pathologist graded villous atrophy according to Marsh-Oberhuber and Corazza classification. The results were subsequently compared using the Cohen kappa statistic. Among the 69 patients taking part in the study, 47 were female and 22 were male (sex ratio=2, 2). The mean age was 15 years (range, 1-61 years). 75% were positive for IgA transglutaminase antibodies and 32.5% for IgG transglutaminase antibodies. The indications for biopsy were dominated by failure to thrive (90%), diarrhea (67.5%), and abdominal distension (47.5%). An immunohistochemical staining with CD3 was performed in only 30% of cases to recognize cases of latent celiac disease when serology and clinical data are not conclusive and when the histological diagnosis remains equivocal.

**According to Marsh-Oberhuber staging:** the first pathologist classified 2 cases in stage I, 3 cases in stage IIIa, 20 cases in stage IIIb, 37 cases in stage IIIc and 7 cases in stage 4. The second pathologist classified 1 case in stage 0, 2 cases in stage I, one case in stage 2, 2 cases in stage IIIa, 25 cases in stage IIIb, 26 cases in stage IIIc and 12 cases in stage 4. The distribution of cases according to Marsh-Oberhuber classification for the 2 pathologists is shown in the Figure 5. The overall agreement kappa was 0.48 which means moderate match rate between the two pathologists. We have evaluated the partial kappa values according to each type of lesion (Table 1). A very good agreement (kappa=1) was found for stage I lesions, whereas the agreement for 3a lesions was moderate (kappa=0.53). Fair agreement was observed for type 3b (kappa=0.35) and poor agreement for types IIIc and 4 (kappa=0.19). Finally, the agreement for type 2 was very low because the first pathologist did not classify any case in this stage.

**According to the New (Corazza) staging:** the first pathologist classified 2 cases in grade A, 23 cases in grade B1 and 44 cases in grade B2. The second pathologist classified 4 cases in grade A, 27 cases in grade B1 and 38 cases in grade B2. The distribution of cases according to the new staging classification for the 2 pathologists is shown in the Figure 6. The overall agreement kappa was 0.61 which means good match rate between the two pathologists. When considering partial kappa values according to each type of lesion (Table 2), a fair agreement was found for grade A (kappa=0.29), a good agreement for type B1 (0.72), and a moderate agreement for B2 (0.49). We noted that the most of the disagreements were found on Marsh-Oberhuber system of grading compared with the new grading system.

## Discussion

---

Celiac disease is a gluten-dependent enteropathy characterized by an increasing in prevalence [6], in morbidity [7], and in mortality rate [8]. The diagnosis requires a combined clinical and histopathology approach with sometimes a complementary immunohistochemical study. The recommended first-line test is serology with immunoglobulin A (IgA) tissue transglutaminase and IgA endomysial antibodies. These serological tests show high levels of sensitivity and specificity, but biopsy techniques have evolved, with routine use of upper endoscopy and biopsy, histopathologists playing a key role in diagnosis [9]. During histological interpretation a particular care should be paid to the number of biopsies [10], correct handling and processing [11], and method of grading CD mucosal abnormalities. To classify villous atrophy, we can use the classification of Marsh or that of Corazza. There is no doubt that the Marsh classification [1], as modified by Oberhuber *et al.* [2] has been accepted as a worldwide standard, but the considerable number of diagnostic categories involved makes it prone to a low interobserver and intraobserver agreement, and given the therapeutic issues, a more reproducible histological classification is essential in order to better classify this pathology. The aim of the present study was to evaluate the interobserver agreement in classifying CD lesions according to both the Marsh-Oberhuber classification and new grading system proposed by Corazza [3]. Two experienced pathologists, who work in centers in which at least 40 new pediatric and adult CD cases are diagnosed every year, evaluated 69 duodenal biopsy specimens by both methods. All evaluations were performed in a blind and independent of any clinical or biological information. The main finding of the study was a high interobserver variability between the two pathologists when using the Marsh-Oberhuber system of grading compared to the classification of Corazza.

The overall agreement kappa was 0.48 (moderate) for Marsh-Oberhuber and 0.61 (good) for Corazza classification. The results of our study are similar to those of other studies, notably those of Corazza [12], which showed that the overall mean kappa value was 0.35 (fair) for the "Marsh-Oberhuber" classification and 0.55 (moderate) for the Corazza classification. When considering the type of lesions, for the Marsh-Oberhuber classification, agreement was good for the stage I lesions, moderate for stage 3a, fair for stage 3b and poor for stages IIIc and 4. In this setting, in the Corazza study [12], agreement was good for the most extreme conditions, like normal mucosa and total villous atrophy (3c), and poor for the other conditions (1, 2, 3a, and 3b). For the Corazza staging, in our study, when considering the type of lesions, a fair agreement was found for grade A, a good agreement for type B1, and a moderate agreement for B2. In the Corazza study [12], a fair agreement was found for grade A, a moderate agreement for normal specimens and type B1, and a good agreement for B2. This result confirmed that when an increase of IELs is suspected, CD3 staining represents a mandatory adjunctive technique to their counting, because it allows a more precise evaluation [13]. Noted that, the upper limit of IELs is 25 per 100 enterocytes in the normal duodenal mucosa [14]. Since the emergence of the new classification proposed by Corazza [3], very few studies have been carried out on this subject. Our study, have shown that this classification is more reproducible than the classification of Marsh (2).

## Conclusion

---

The prospective evaluation of 69 cases of duodenal biopsies by the two pathologists showed that the simplified grading proposed by Corazza [3], is more reliable than the Marsh-Oberhuber [1,2] classification system used worldwide. It is known that the greater the number of diagnostic categories of

a method, the lower its diagnostic reproducibility [15]. So probably that the use of this more simplified grading system will allow a better diagnosis and facilitate the relationship between pathologists and clinicians.

## What is known about the topic

- Marsh classification has been accepted as a worldwide standard and it is the most used by pathologists;
- Considering the considerable number of diagnostic categories involved makes it prone to a low interobserver and intraobserver agreement;
- Very few studies have been carried out on this subject, and they have shown that the classification proposed by Corazza is more reproducible than the classification of Marsh.

## What this study adds

- Considering the reduced number of studies carried out on this subject, our study constitutes an addition;
- These results will push pathologists to use a more simplified classification to avoid disagreement;
- Our study has shown that the simplified grading proposed by Corazza, is more reliable than the Marsh-Oberhuber classification.

## Competing interests

---

The authors declare no competing of interest.

## Authors' contributions

---

Amal Douida drafted the manuscript. All authors read and approved the final manuscript.

## Tables and figures

---

**Table 1:** agreement between the two pathologists according to each type of lesion using the Marsh-Oberhuber classification

**Table 2:** agreement between the two pathologists according to each type of lesion using the Corazza Classification

**Figure 1:** (HESx200) normal mucosal architecture with increase of IELs (>25/100 enterocytes)

**Figure 2:** (HESx100) villous/crypt ratio is less than 3/1

**Figure 3:** (HESx100) complete villous atrophy

**Figure 4:** comparison between the Marsh-Oberhuber and the new grading system (Corazza)

**Figure 5:** distribution of cases according to Marsh-Oberhuber classification for the two pathologists.

**Figure 6:** distribution of cases according to Corazza classification for the two pathologists

## References

---

1. Marsh MN. Gluten, major histocompatibility complex and the small intestine: a molecular and immunobiologic approach to the spectrum of gluten sensitivity (celiac sprue). *Gastroenterology*.1992 Jan;102(1):330-54. **PubMed | Google Scholar**
2. Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease time for a standardized report scheme for pathologists. *Eur J Gastroenterol Hepatol*. 1999 Oct;11(10):1185-94. **PubMed | Google Scholar**
3. Corrao G, Corazza GR, Andreani ML, Torchio P, Valentini RA, Galatola G *et al*. Coeliac disease: some considerations on the histological diagnosis. *J Clin Pathol*. 2005 Jun;58(6):573-4. **PubMed | Google Scholar**
4. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar;33(1):159-74. **PubMed | Google Scholar**
5. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall.1991.
6. Maki M, Mustalahti K, Kokkonen J, Kulmala P, Haapalahti M, Karttunen T *et al*. Prevalence of celiac disease among children in Finland. *N Engl J Med*. 2003 Jun 19;348(25):2517-24. **PubMed | Google Scholar**
7. Collin P, Maki M. Associated disorders in coeliac disease: clinical aspects. *Scand J Gastroenterol*. 1994 Sep;29(9):769-75. **PubMed | Google Scholar**
8. Corrao G, Corazza GR, Bagnardi V, Brusco G, Ciacci C, Cottone M *et al*. Mortality in patients with coeliac disease and their relatives: a cohort study. *Lancet*. 2001 Aug 4;358(9279):356-61. **PubMed | Google Scholar**
9. Losowsky MS. A history of coeliac disease. *Dig Dis*. 2008;26(2):112-20. **PubMed | Google Scholar**
10. Mee AS, Burke M, Vallon AG, Newman J, Cotton PB. Small bowel biopsy for malabsorption: comparison of the diagnostic adequacy of endoscopic forceps and capsule biopsy specimens. *Br Med J*. 1985 Sep 21;291(6498):769-72. **PubMed | Google Scholar**

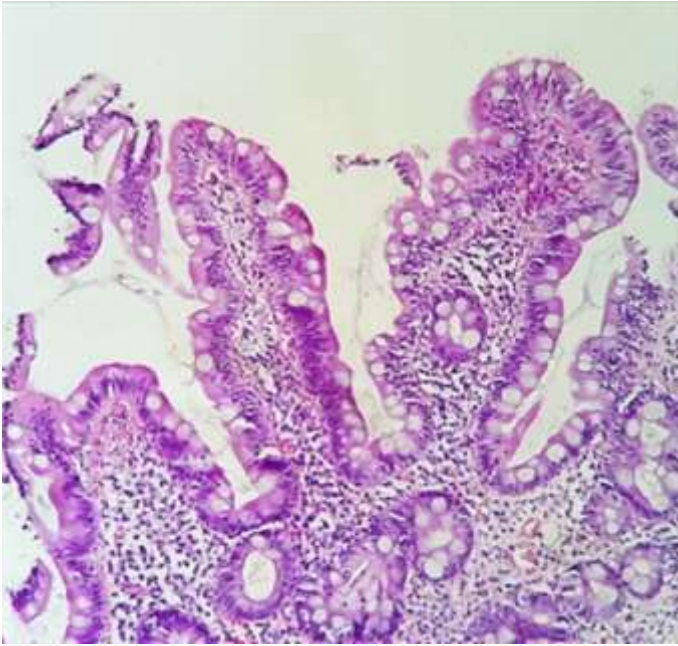
11. Perera DR, Weinstein WM, Rubin CE. Symposium on pathology of the gastrointestinal tract-Part II Small intestinal biopsy. *Hum Pathol.* 1975 Mar;6(2):157-217. **PubMed | Google Scholar**
12. Corazza GR, Villanacci V, Zambelli C, Milione M, Luinetti O, Vindigni C *et al.* Comparison of the interobserver reproducibility with different histologic criteria used in celiac disease. *Clin Gastroenterol Hepatol.* 2007 Jul;5(7):838-43. **PubMed | Google Scholar**
13. Veress B, Franzen L, Bodin L, Borch K. Duodenal intraepithelial lymphocyte-count revisited. *Scand J Gastroenterol.* 2004 Feb;39(2):138-44. **PubMed | Google Scholar**
14. Ensari A. Gluten-sensitive enteropathy (celiac disease): controversies in diagnosis and classification. *Arch Pathol Lab Med.* 2010 Jun;134(6):826-36. **PubMed | Google Scholar**
15. Koran LM. The reliability of clinical methods, data and judgments (first of two parts). *N Engl J Med.* 1975 Sep 25;293(13):642-6. **PubMed | Google Scholar**

**Table 1:** agreement between the two pathologists according to each type of lesion using the Marsh-Oberhuber classification

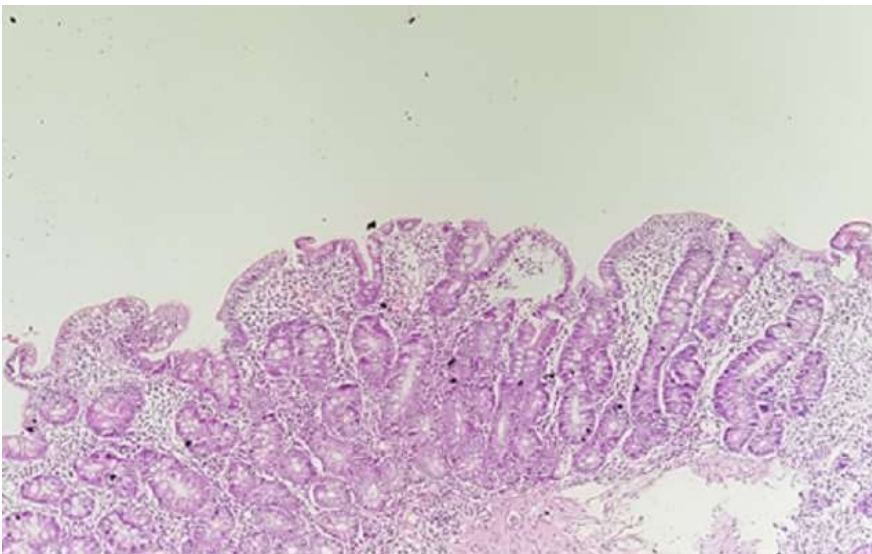
Type of lesion	Kappa value
1	1
2	0.04
3a	0.53
3b	0.35
3c	0.19
4	0.19

**Table 2:** agreement between the two pathologists according to each type of lesion using the Corazza classification

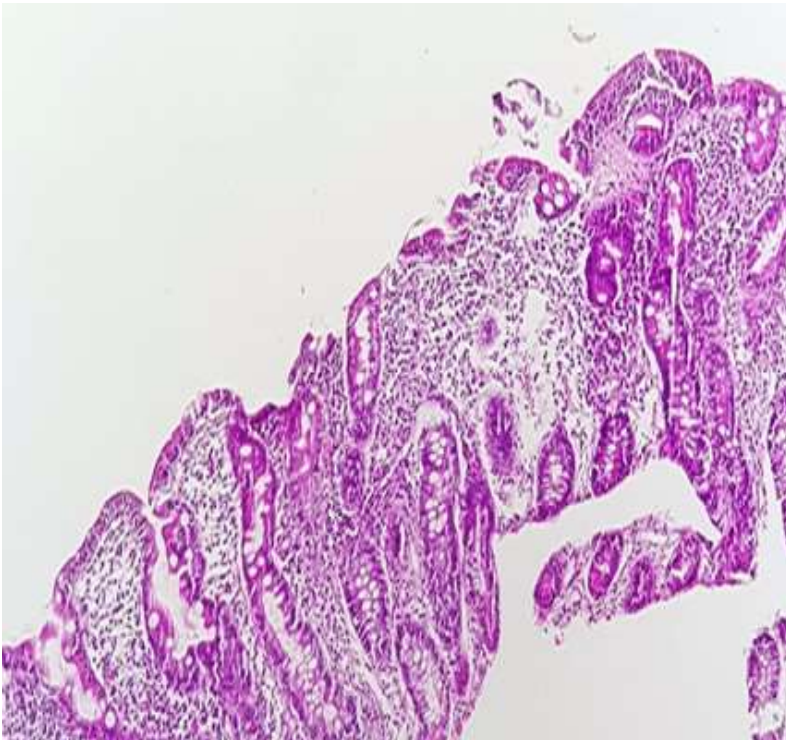
Type of lesion	Kappa value
A	0.29
B1	0.72
B2	0.49



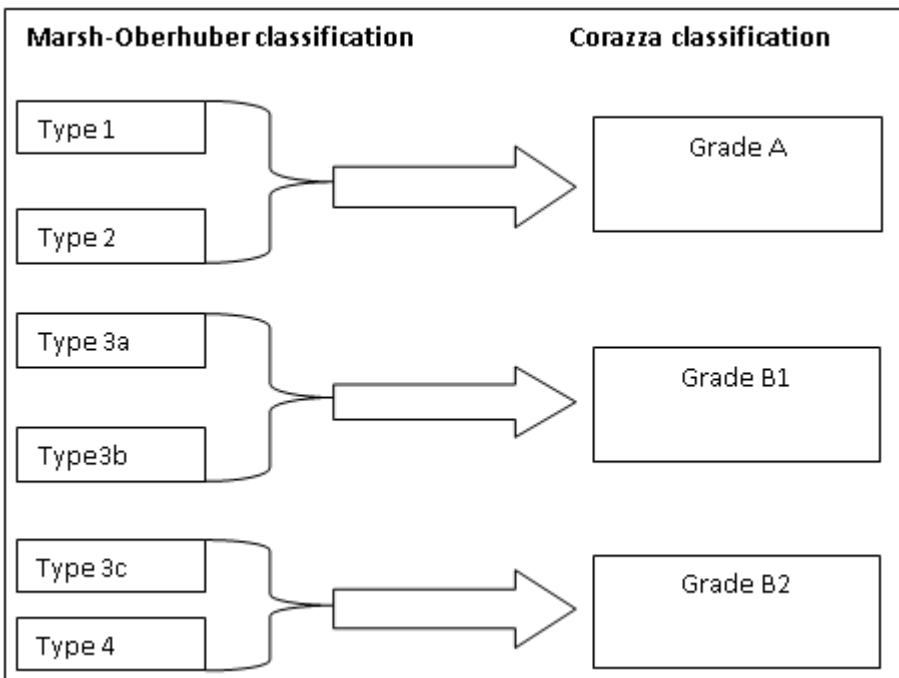
**Figure 1:** (HESx200) normal mucosal architecture with increase of IELs (>25/100 enterocytes)



**Figure 2:** (HESx100) villous/crypt ratio is less than 3/1

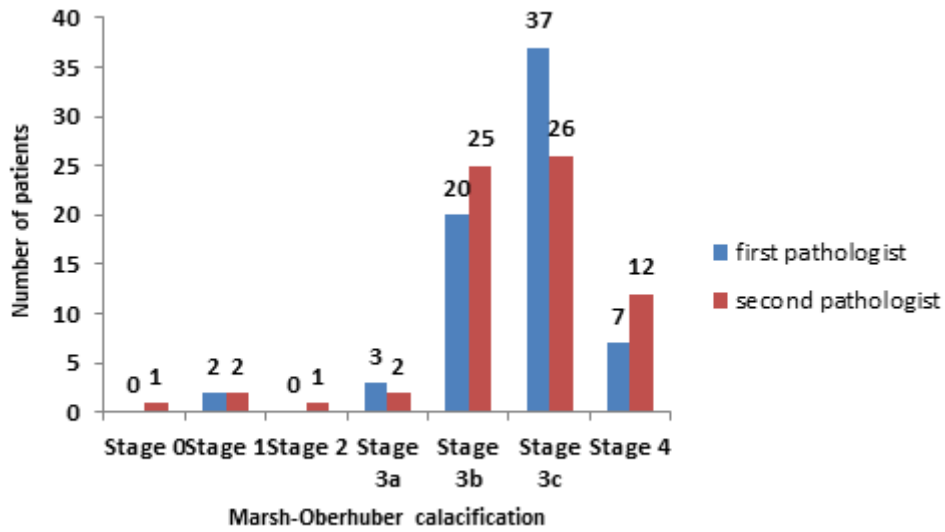


**Figure 3:** (HESx100) complete villous atrophy

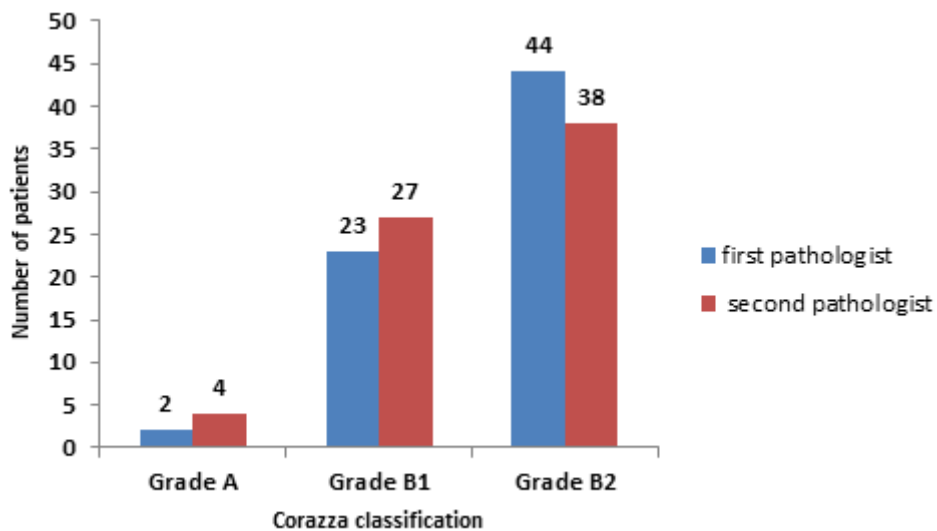


**Figure 4:** comparison between the Marsh-Oberhuber and the new grading system (Corazza)





**Figure 5:** distribution of cases according to Marsh-Oberhuber classification for the two pathologists



**Figure 6:** distribution of cases according to Corazza classification for the two pathologists