Celiac Disease: An Assessment of Subjective Variation and Diagnostic Reproducibility of the Various Classification Systems

By Manas Madan, Sanjay Piplani, Manisha Sharma, Tejinder Singh Bhasin, Mridu Manjari, Harjot Kaur, Jasmeet Kaur & Saumil Garg

Guru Ramdas Institute of Medical Sciences & Research, Amritsar, Punjab, India

Abstract-Introduction: Celiac disease (CD) is a chronic immune mediated disorder occurring in genetically predisposed individuals with intolerance to gluten, particularly its protein gliadin. The histological examination still remains the gold standard for its diagnosis. Marsh-Oberhuber classification is very widely used by pathologists for the diagnosis of CD and is valid under optimal clinical conditions. However, due to the presence of greater diagnostic categories, it lends itself to greater subjective variability and lower interobserver and intraobserver agreement and hence lower reproducibility of the diagnosis. Recently, Corazza and Villanacci introduced a classification that reduces the number of categories and the interobserver variation. This study was undertaken to observe the reproducibility of the Marsh-Oberhuber classification in comparison to the newer Corazza and Villanacci classification and determine the intra and interobserver variation in both the classifications.

Keywords: celiac disease, gluten, histopathology.

GJMR-C Classification : NLMC Code: WD 175

Strictly as per the compliance and regulations of:
Celiac Disease: An Assessment of Subjective Variation and Diagnostic Reproducibility of the Various Classification Systems

Manas Madan a, Sanjay Piplani a, Manisha Sharma a, Tejinder Singh Bhasin a, Mridu Manjari x, Harjot Kaur b, Jasmeet Kaur x & Saumil Garg y

Abstract

Introduction: Celiac disease (CD) is a chronic immune mediated disorder occurring in genetically predisposed individuals with intolerance to gluten, particularly its protein gliadin. The histological examination still remains the gold standard for its diagnosis. Marsh-Oberhuber classification is very widely used by pathologists for the diagnosis of CD and is valid under optimal clinical conditions. However, due to the presence of greater diagnostic categories, it lends itself to greater subjective variability and lower interobserver and intraobserver agreement and hence lower reproducibility of the diagnosis. Recently, Corazza and Villanacci introduced a classification that reduces the number of categories and the interobserver variation. This study was undertaken to observe the reproducibility of the Marsh-Oberhuber classification in comparison to the newer Corazza and Villanacci classification and determine the intra and interobserver variation in both the classifications.

Materials And Methods: The present study was a retrospective one and comprised of 86 patients who were already diagnosed as CD according to Marsh Oberhuber classification at Sri Guru Ramdass Institute of Medical Sciences and research, Amritsar, Punjab. The slides were retrieved from the archives and reexamined independently by two pathologists and re classified according to Marsh Oberhuber classification without either of them knowing the initial diagnosis. The slides were then shuffled and again classified according to Corazza and Villanacci classification by the same two pathologists. Then the initial diagnosis reported as per the Marsh Oberhuber classification was also noted. The intraobserver and interobserver variation among the two classification systems was then determined.

Conclusion: There is immense histological variation in CD and the spectrum is increasing along with the number of tests involved in its diagnosis. Histopathology is considered as the gold standard in its diagnosis along with the clinical history and serological findings. The classification systems are also ever evolving each with its merits and demerits. The modified Marsh classification system although efficacious and widely used lends itself to a greater subjective variation due to the large number of categories involved. The new classification system proposed by Corazza and Villanacci simplifies the above classification, reduces the number of categories and hence greater diagnostic reproducibility. Our study further corroborates this fact although it is limited by small sample size. More studies should be undertaken with a larger sample size to determine its validity, accuracy and reproducibility.

Keywords: celiac disease, gluten, histopathology.

I. Introduction

The term celiac was first used in the first century AD by the physician Celsius when he used the term Celiac for a diarrhea like disease. The understanding of Celiac disease (CD), also known as gluten induced enteropathy has come a long way since with regards to its etiology, pathogenesis and the various modalities of diagnosis. Now we are clear that this disease is a chronic immune mediated disorder occurring in genetically predisposed individuals with intolerance to gluten, particularly its protein gliadin. This elicits an abnormal immune mediated response characterized by chronic inflammation of small intestinal villi and associated with progressive disappearance of intestinal villi.(1,2)

The histological examination remains the gold standard for its diagnosis. (1,3,4) The diagnosis is based on biopsy showing the presence of characteristic histological changes in duodenum and jejunum that improve after gluten free diet. (2,3)

Histological abnormalities characteristic of CD were described in 1954 by Paulley. Marsh in 1990 classified the various histologic patterns seen in CD which were further modified by Oberhuber in 1999. This classified the histology into 5 categories (Type 0-4).(2,3,4)

Type 0: Preinfiltrative, Normal small intestinal architecture, < 30 Intraepithelial lymphocytes (IEL)/100 enterocytes.

Type I: Infiltrative type, normal villous:crypt ratio >3:1, > 30 IEL enterocytes.

Type II: Infiltrative hyperplastic: Normal villi, Crypt hyperplasia, increased IELs.

Type III: Destructive CD further subdivided into 3 sub categories.

Type IIIa: Mild villous atrophy, villi:crypt ratio <3:1, increased IELs.

Type IIIb: Marked villous atrophy, villi:crypt ratio <1:1, increased IELs.
Type IIIc: Total villous atrophy (flat mucosa), increased IELs.
Type IV: Atrophic type (hypoplastic)

The above classification is very widely used by pathologists for the diagnosis of CD and is valid under optimal clinical conditions. However, due to the presence of greater diagnostic categories, it lends itself to greater subjective variability and lower interobserver and intraobserver agreement and hence lower reproducibility of the diagnosis. (1,3,5)

Recently, Corazza and Villanacci modified the above classification. This newer classification reduces the number of categories. Type 1 and 2 have been clubbed into Grade A, 3a and 3b into Grade B1, 3c into grade B2. Type 4 category of Marsh Oberhuber has been deleted. This classification system further simplifies the criteria and reduces the number of categories and hence the interobserver variation. (1,2,3,5)

Table 1

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Type A (Non Atrophic)</th>
<th>Type B1 (Atrophic)</th>
<th>Type B2 (Atrophic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraepithelial Lymphocytosis</td>
<td>Present</td>
<td>Present</td>
<td>Present</td>
</tr>
<tr>
<td>Villi</td>
<td>Normal</td>
<td>Still detectable</td>
<td>Undetectable</td>
</tr>
<tr>
<td>Marsh Oberhuber Equivalent</td>
<td>Type 1 and 2</td>
<td>Type 3a and 3b</td>
<td>Type 3c</td>
</tr>
</tbody>
</table>

This study was undertaken to observe the reproducibility of the Marsh-Oberhuber classification in comparison to the newer Corazza and Villanacci classification and determine the intra and interobserver variation in both the classifications.

II. MATERIALS AND METHODS

The aim of the study was to observe the reproducibility of the classification systems in patients of CD and to assess the interobserver and intraobserver variation among these.

The present study was a retrospective one and comprised of 86 patients who were already diagnosed as CD according to Marsh Oberhuber classification at Sri Guru Ramdass Institute of Medical Sciences and research, Amritsar, Punjab.

The slides were retrieved from the archives and reexamined independently by two pathologists and reclassified according to Marsh Oberhuber classification without either of them knowing the initial diagnosis. The slides were then shuffled and again classified according to Corazza and Villanacci classification by the same two pathologists. Then the initial diagnosis reported as per the Marsh Oberhuber classification was also noted.

The intraobserver variation (among each of the two pathologists) was then noted among the two diagnosis (initial diagnosis and the diagnosis made after reexamination, both according to Marsh Oberhuber classification)

The interobserver variation was then determined among the two pathologists for the diagnosis made after reexamination according to Marsh Oberhuber classification

Also, the interobserver variation was determined among the two pathologists for the diagnosis made after reexamination according to Corazza and Villanacci classification.

III. RESULTS

86 patients were included in this study group.

Histological Examination: The histology was classified first according Marsh Oberhuber and then according to Corazza staging.

Results of initially reported diagnosis: The initial diagnoses for the 86 cases according to Marsh Oberhuber classification were as follows:

Table 2

<table>
<thead>
<tr>
<th>Category</th>
<th>Type I</th>
<th>Type II</th>
<th>Type IIIa</th>
<th>Type IIIb</th>
<th>Type IIIc</th>
<th>Type IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>18</td>
<td>03</td>
<td>13</td>
<td>17</td>
<td>35</td>
<td>00</td>
</tr>
</tbody>
</table>

When reclassified according to the same classification, following were the results of both the pathologists.

Table 3

<table>
<thead>
<tr>
<th>Category</th>
<th>Type I</th>
<th>Type II</th>
<th>Type IIIa</th>
<th>Type IIIb</th>
<th>Type IIIc</th>
<th>Type IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>17</td>
<td>02</td>
<td>16</td>
<td>14</td>
<td>37</td>
<td>00</td>
</tr>
</tbody>
</table>
Thus, there was a significant intraobserver and interobserver difference in categories type IIIa and IIIb of Marsh-Oberhuber classification whereas the difference was much less in the categories types I and IIIc. No case was diagnosed as CD type IV in all the three instances.

The results of both pathologists when classified according to the Corazza and Villanacci classification were as follows.

### Table 4

<table>
<thead>
<tr>
<th>Category</th>
<th>Type I</th>
<th>Type II</th>
<th>Type IIIa</th>
<th>Type IIIb</th>
<th>Type IIIc</th>
<th>Type IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>18</td>
<td>04</td>
<td>12</td>
<td>18</td>
<td>34</td>
<td>00</td>
</tr>
</tbody>
</table>

### Table 5

**Pathologist 1:**

<table>
<thead>
<tr>
<th>Category</th>
<th>Type A</th>
<th>Type B1</th>
<th>Type B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>21</td>
<td>30</td>
<td>35</td>
</tr>
</tbody>
</table>

**Pathologist 2:**

<table>
<thead>
<tr>
<th>Category</th>
<th>Type A</th>
<th>Type B1</th>
<th>Type B2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>22</td>
<td>30</td>
<td>34</td>
</tr>
</tbody>
</table>

Thus, much lesser interobserver variation was found when CD was classified according to Corazza and Villanacci classification.

### IV. Discussion

This study was undertaken in 86 already diagnosed cases of CD according to Marsh Oberhuber classification which were then reexamined by two pathologists independently and reclassified according to Marsh Oberhuber and Corazza Villanacci classification to assess the intraobserver and interobserver variation among the two classification systems.

CD is a highly variable disease histologically and can exhibit many microscopic patterns. Although histopathology is considered as the gold standard for its diagnosis, the correct diagnosis of CD depends on a combination of clinical features, serology and histopathological features to give a presumptive diagnosis of CD. The final diagnosis rests on the improvement of the symptoms/serological values/biopsy findings after gluten free diet.(2,3,5,6,7)

Due to a variety of histological patterns, many classification systems have been proposed in the past to categorize the various patterns that this disease exhibits. Initially proposed by Marsh and then modified by Oberhuber, the modified Marsh classification system has been widely used for the classification of CD. This system is no doubt efficacious and is valid under optimal clinical conditions.(2,3,4,6,7,8)

However, there are concerns about its validity and efficacy in daily clinical practice and with respect to an individual’s clinical presentation. Due to the large number of diagnostic categories, there tends to be lower intraobserver and interobserver agreement therefore leading to a lower reproducibility of the diagnosis.(1,2,3,4,5,6,8,9) The same was found in our study where there was both intraobserver and interobserver variation when CD was classified according to this classification. This variation was negligible in type I (Corazza type A) (Fig 1), and IIIc categories whereas it was much more pronounced in type IIIa and IIIb categories. This could be due to the fact that recognition of lesser degrees of villous abnormalities lends itself to a greater intraobserver and interobserver variability because of subjective differences in the recognition of these changes. The new classification system by Corazza groups these two categories into a single one (Type B1) (Table 2) (FIG 2).

Due to the reduction of the categories and hence a consequent reduction in the subjective variation (in seeing whether the villi are mildly atrophic or markedly atrophic but not yet completely flat), there tends to be better agreement among the various pathologists.(1,2,3,4,5,6,8,9) Our study further corroborated this as there was significantly improved intraobserver and interobserver agreement in type B1 category of Corazza when independently examined by two pathologists. (Table 2,3,4,5,6)
The type 2 hyperplastic lesions category according to modified marsh classification has been omitted in the new Corazza classification due to the doubts in its usefulness and efficacy. The new patients would already be diagnosed by increased IELs and hence this category does not impart any useful information about the disease although it definitely represents a spectrum of the histological stage.\(^\text{(2,3,5)}\) In our study, a very minor number of cases were classified into this category. (Table 2)

Type 3c category of Marsh classification has been re designated as B2 and these are the very commonly encountered lesions in CD. This category lends itself to very less subjective variation as the histological changes are very marked and striking leading to a greater diagnostic reproducibility.\(^\text{(2,3,4,5,7,8,9,10)}\) In our study too, very good intraobserver and interobserver agreement as found in this category. (Table 2,3,4,5,6) (FIG 3)
The last category of Marsh classification (type IV) has been omitted as it is virtually never seen in practice and has alos been made obsolete by a recent finding of an aberrant IEL clone characteristically seen in refractory sprue, ulcerative jejunoileitis and enteropathy type intestinal T cell lymphoma. No case was reported as type IV in our study too. (Table 2)

V. Conclusion

There is immense histological variation in CD and the spectrum is increasing along with the number of tests involved in its diagnosis. Histopathology is considered as the gold standard in its diagnosis along with the clinical history and serological findings. The classification systems are also ever evolving each with its merits and demerits. The modified Marsh classification system although efficacious and widely used lends itself to a greater subjective variation due to the large number of categories involved. The new classification system proposed by Corazza and Villanacci simplifies the above classification, reduces the number of categories leading to more intraobserver and interobserver agreement and hence greater diagnostic reproducibility. Our study further corroborates this fact although it is limited by small sample size. More studies should be undertaken with a larger sample size to determine its validity, accuracy and reproducibility.

REFERENCES Références Referencias

9. Corazza GR, Bonvicini F, Frazzoni M, gatto M, Gasbarrini G. Observer variation in assessment of...