Optimising the follow-up of adult coeliac disease with a clinical-based score to identify patients in need of a histological reassessment: a retrospective single centre study

Gaia Harder1†, Annalisa Schiepatti1‡, Federico Biagi2*, Federica Borrelli de Andreis1, Simona Agazzi1, Gian Marco Gabrielli1 and Catherine Klersy3

1Coeliac Centre/First Department of Internal Medicine, Fondazione IRCCS Policlinico San Matteo, University of Pavia, Pavia, Italy
2Istituti Clinici Scientifici Maugeri IRCCS, Gastroenterology Unit of Pavia Institute, University of Pavia, Pavia, Italy
3Clinical Epidemiology & Biometry, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

(Submitted 6 June 2019 – Final revision received 31 January 2020 – Accepted 9 February 2020 – First published online 20 February 2020)

Abstract

Follow-up modalities for adult coeliac patients remain controversial. Non-invasive markers to identify coeliac patients on a gluten-free diet (GFD) with persistence of villous atrophy (VA) are still lacking. We aim to develop a score to stratify coeliac patients on a GFD according to their risk of having persistent VA and to tailor follow-up modalities accordingly. The clinical notes of over 700 coeliac patients attending our unit (September 1999–November 2018) were retrospectively examined. A total of 273 patients on a GFD with a histological follow-up performed 12–24 months after diagnosis were selected. We developed a bivariable model based on diet adherence and clinical response evaluated by previously validated methods. A four-level score (0, 1-5, 3, 4) was obtained. Patients on a strict GFD and with good clinical conditions (score 4) have a very low risk of persistence of VA (2 (95 % CI 1, 5) %). Conversely, the risk is very high (46 (95 % CI 25, 68) %) in patients with poor adherence to a GFD and unsatisfactory clinical response (score 0-5). A score of 1-5 (poor GFD adherence and persistent well-being) is linked with a high risk (23 (95 % CI 15, 36) %). Risk is intermediate (6 (95 % CI 3, 10) %) in patients scoring 3 (strict GFD and no/partial clinical improvement). Three patients who developed complications belonged to this scenario. Patients at low risk of persistent VA can be followed-up non-invasively, whereas a biopsy should be repeated in those at high/very high risk. Case-by-case evaluation is needed in patients at intermediate risk. Studies on a larger sample size are required to confirm these data.

Key words: Coeliac disease: Gluten-free diet: Duodenal biopsy: Endomysial antibodies: Tissue transglutaminase

Coeliac disease (CD) is a lifelong gluten-dependent enteropathy characterised by an extremely heterogeneous clinical picture, high prevalence in the general population and increased morbidity and mortality4–5. Currently, a strict gluten-free diet (GFD) is the cornerstone for treatment of CD, resulting in complete clinical and histological recovery in the vast majority of patients1–5. Although rare, complications of CD may occur, leading to long-term reduced quality of life and poor outcomes7–10. The role of mucosal healing on a GFD in predicting poor outcomes in coeliac patients has been the focus of several papers that have provided conflicting results7–22. While some found that long-term outcomes and mortality in coeliac patients were not influenced by persistent villous atrophy (VA)13,22, the majority demonstrated that mucosal healing may play a role in preventing long-term complications of CD14,17,19–21. On the other hand, assessment of histological response requires a gastroscopy, a procedure that may expose the patients to additional risks, albeit minimal, and impacts on healthcare costs23. Follow-up of coeliac patients is a controversial issue. Although the need for regular follow-up is unquestionable, neither who should see these patients, nor the practical modalities to deliver follow-up care cost-effectively, have been established so far. Currently, the major international guidelines on CD do not recommend a routine re-biopsy strategy in all coeliac patients but are concordant in suggesting a second duodenal biopsy in

Abbreviations: CD, coeliac disease; GFD, gluten-free diet; TTA, tissue transglutaminase antibodies; VA, villous atrophy.

* Corresponding author: Professor Federico Biagi, email federico.biagi@icsmaugeri.it
† These authors contributed equally to this work.
‡ PhD course in Experimental Medicine, University of Pavia, Italy.
those with persisting symptoms despite being on a GFD\(^1\). Moreover, despite the availability of different non-invasive surrogate markers to assess GFD adherence\(^{24-26}\), there is still no widespread consensus on whether responsibility for coeliac patients' follow-up should be taken by dietitians, primary care physicians, gastroenterologists or tertiary referral centres\(^1\). Moreover, which follow-up modalities to employ are even more controversial. Although in clinical practice evaluation of adherence to a GFD can be made by clinical assessment, dietetic review and serology, demonstration of histological recovery on a GFD is also part of the routine follow-up in some centres\(^{19,27,28}\).

In our referral centre, since 1999, we have always offered routine follow-up duodenal biopsy to all coeliac patients regardless of their clinical picture at diagnosis, and their clinical and serological response to a GFD. A few years ago, to evaluate the appropriateness of this strategy, we analysed the results of these follow-up biopsies and found that, after 12–15 months on a GFD, VA was still present in 8% of coeliac patients\(^{28}\). Persistence of VA despite being on a GFD has also been shown by other authors\(^{14,16,19,27}\).

Since at present non-invasive markers to identify the minority of CD patients on a GFD with persistent VA are lacking, the aim of our present work is to develop a scoring system to stratify coeliac patients on a GFD according to their risk of having persistence of VA. This would enable us to propose a cost-effective modality for the follow-up of coeliac patients, based on referral to a histological reassessment only those coeliac patients at higher risk of persistence of VA, thus avoiding it in patients at lower risk.

### Patients and methods

#### Patients and study design

The clinical notes of all adult coeliac patients with a follow-up duodenal biopsy performed while on a GFD in our centre between September 1999 and November 2018 were retrospectively examined. We highlight the fact that we have been recommending a follow-up duodenal biopsy in all coeliac patients attending our unit regardless of their clinical or serological response to a GFD. In order to reduce data variability, only coeliac patients with a duodenal biopsy taken between 12 (so 2) and 24 (so 2) months after the start of a GFD were enrolled in the present study.

All patients had been previously diagnosed with CD, either in our centre or in other hospitals, on the basis of an upper gastrointestinal endoscopy with duodenal biopsies showing frank VA (at least Marsh 3a lesions or Corazza/Villanacci B) and positive IgA endomysial antibodies (EmA) and/or tissue transglutaminase antibodies (TTA)\(^1\). Patients diagnosed at other hospitals and referred to our centre attended our outpatient clinic to confirm the diagnosis of CD and provide them with gluten-free food certificates and regional health service tax exemption. For all the referred patients, review of histological slides confirmed the diagnosis of VA. Patients found to be affected by CD in childhood and those with potential CD (i.e. positive EmA/TTA and Marsh 0–2 lesions), seronegative or complicated form of CD were excluded\(^1\).

#### Biopsy samples and coeliac serology

To guarantee good quality samples in our unit, since 1999, we have been taking four duodenal biopsies in the second part of the duodenum during upper endoscopy performed under mild sedation. Biopsies were then oriented on cellulose nitrate paper, formalin-fixed and paraffin-embedded and cut perpendicularly with respect to the luminal surface. Immunohistochemistry for CD3 and haematoyxin–eosin staining were routinely performed. At our centre, a blood sample for EmA testing is taken on the same day on which patients undergo duodenal biopsy. EmA are tested by indirect immunofluorescence (dilution 1:5) on monkey oesophagus/jejunum sections (INOVA diagnostics). We specify that we do not routinely test for TTA. In our experience, both sensitivity and specificity of EmA and TTA are very similar and satisfactory\(^{29}\). Finally, we always measure total IgA, IgG and IgM to exclude IgA deficiency or hypogammaglobulinaemia.

#### Evaluation of gluten-free diet adherence

All patients have been correctly instructed by a skilled dietitian with expertise on CD on how to follow a GFD. Evaluation of the adherence to a GFD was made on the basis of a standard five-level score we previously developed and validated\(^{24}\). Briefly, patients scoring 3 and 4 were considered adherent to a GFD, while scores 0, 1 and 2 were considered unsatisfactory. Since this scoring system was developed in March 2008, patients seen before that date were evaluated by means of a structured dietetic interview performed by skilled dietitian\(^{24}\).

#### Criteria to assess histological and clinical response to a gluten-free diet

Histological response was considered satisfactory in all the patients with resolution of VA (Marsh grade <3 or Corazza/Villanacci A). On the contrary, patients with persistence of VA on follow-up duodenal histology (Marsh ≥3 or Corazza/Villanacci B)\(^1\) were scored as to 'with unsatisfactory histological response'. In order to evaluate the clinical response to a GFD, we adopted the same criteria previously described\(^{28}\). Briefly, patients were divided into two distinct categories on the basis of their clinical response to a GFD. Asymptomatic patients at diagnosis and patients with complete resolution of all symptoms and laboratory alterations were classified as 'complete recovery or persistent well-being'. On the other hand, patients without such a significant improvement were labelled as 'partial improvement or no significant improvement'.

#### Statistics and development of the score

Continuous variables that were normally distributed were summarised using mean values and standard deviations and those that were not-normally distributed as median and interquartile range. Categorical variables were summarised as counts and percentages. A two-sided \(P\) value of <0.05 was considered statistically significant. Logistic regression was used to assess the association of predefined variables with a persistent VA biopsy finding in the selected time window. OR and 95% CI were computed.
To develop the score, given the low number of events, no multivariable model could be fitted (according to the 1:10 predictors to events rule). Thus, a bivariable model including clinical and easy to measure variables was used for risk stratification. These were identified a priori. On the basis of our clinical experience and our previous works, we fixed GFD adherence and clinical response as predictors. The coefficients derived from the model were rounded to the nearest 0.5, inverted and summed-up for a predictive score calculation. Model discrimination was assessed by computing the area under the ROC curve and 95% CI. The probability of an unfavourable biopsy and its 95% CI were derived from the model, and it was plotted against the score for clinical use.

Sample size. The sample size was determined by feasibility. The minimally detectable effect sizes and the power for these comparisons are summarised in online Supplementary Table. We use two-sided α-levels of 2.5% (adjusting for two comparisons) and 5% and a power of 80% in the calculations. We use Stata 15.1 (StataCorp) for computation.

Ethics. Although all patients signed informed consent before the biopsies, they could not be asked to consent specifically for this study because of its retrospective nature. Since many patients were lost to follow-up, after verifying the good quality of the data, the data were all irreversibly anonymised. They all agreed to have their data used for scientific purposes, and none of them signed against it. The study was approved by the ethics committee of the Fondazione IRCCS Policlinico San Matteo according to the 1975 Declaration of Helsinki (6th revision, 2008).

Results

Since 1999, more than 700 adult coeliac patients have been diagnosed and/or followed-up in our centre. A total of 273 of them (203 female, mean age at diagnosis of CD 35.4 (so 12.0) years) received a thorough re-evaluation between 10 and 26 months and were therefore included in the present study. In accordance with the Oslo definitions(30), 141 patients were affected by a classical form of CD, 108 by a non-classical form and twenty-four with the asymptomatic one. The development of the scoring system was based on these 273 patients. Table 1 shows the collected clinical and biotic findings.

<table>
<thead>
<tr>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
</tr>
<tr>
<td>Clinical form of coeliac disease according to Oslo definitions(1)</td>
</tr>
<tr>
<td>Classical</td>
</tr>
<tr>
<td>Non-classical, asymptomatic</td>
</tr>
<tr>
<td>Gluten-free diet adherence†</td>
</tr>
<tr>
<td>Strict</td>
</tr>
<tr>
<td>Poor</td>
</tr>
<tr>
<td>Months to second biopsy</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>25th–75th percentiles</td>
</tr>
<tr>
<td>Clinical picture at time of follow-up*</td>
</tr>
<tr>
<td>Persistent well-being, complete recovery</td>
</tr>
<tr>
<td>Partial improvement, no significant improvement</td>
</tr>
<tr>
<td>EmA at time of follow-up*</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>Histological improvement at follow-up</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
</tbody>
</table>

EmA, endomysial antibodies. * Data were not available for all the patients. † Based on the gluten-free diet score developed in Pavia or on a dietary interview performed by expert personnel(28).

Table 1. Collected clinical and biotic findings (Numbers and percentages; median and 25th–75th percentiles)

Based on the gluten-free diet score developed in Pavia or on a dietary interview performed by expert personnel(28), the patients could be grouped according to their risk of having persistent VA into four clinical scenarios: very high risk of having persistent VA which means poor GFD adherence (−2·5), partial improvement/no significant clinical improvement (−1) and constant term (+4) leading to a final score of 0.5. The other risk scores can be calculated in the same way: high risk score (poor GFD adherence and persistent well-being/full clinical recovery, final result 1·5), intermediate risk (strict GFD adherence and partial/no clinical improvement, final result 3) and low risk (strict GFD adherence and persistent well-being/full clinical recovery, final result 4). In patients, who have been on a GFD for 10–26 months, the persistence of symptoms together with a poor adherence to a GFD is linked to a very high risk of persistence of VA (almost 50%). On the other hand, good adherence to a GFD and good clinical conditions at time of follow-up have a low risk of persistence of VA (2%). Table 3 is a ‘user-friendly’ version of our scoring system. It also provides our suggestions for follow-up of coeliac patients.

Finally, according to the aforementioned score, twenty-eight of the 263 patients (10.6%) followed-up in our centre belong to the very high and high risk groups, seventy patients (26.6%) to the intermediate risk group and the last 165 patients (62.7%) to the low risk one (ten patients could not be included because of some missing data). Three patients later developed...
complications, one abdominal B cell-lymphoma, one enteropathy-associated T-cell lymphoma and one refractory CD type 1. They all belonged to intermediate risk scenario.

Discussion

CD is a chronic condition that requires lifelong treatment with a GFD(1–4). CD is also characterised by increased mortality, mainly due to its complications, that may even occur many years after the initial diagnosis(5,7–10). Therefore, delivery of the best follow-up care is an unquestionable necessity for coeliac patients. Having said this, the most cost-effective approach and timing of follow-up for CD are still a matter of debate(1–4).

With the present study, we have provided a practical instrument for the decision-making process in the follow-up of coeliac patients. The score we obtained is extremely simple and can quickly be administered in everyday clinical practice, being based on clinical and dietary evaluation. For the first time, we have quantified and standardised here the direct role of these two parameters. A proposal for a strategy to optimise the follow-up of coeliac patients based on our score is shown in Table 3. Risk of persistent VA is ‘very high’ (≥46%) for patients in scenario 1 and ‘high’ (≥23%) in scenario 2. They should both be referred for gastroscopy with duodenal biopsy and complete routine laboratory tests to exclude malignant complications of CD(1–4,7–10). At the same time, strategies to improve GFD adherence are mandatory in these patients. On the contrary, patients in scenario 4 are at low risk (risk ≤2%) and they could be safely followed-up non-invasively by performing laboratory evaluation and dietary assessment only. Finally, the result that all patients who developed complications belonged to the ‘intermediate risk’ group for persistent VA is crucial. Based on our proposals, these patients should have been followed up case-by-case. In this group of patients at moderate risk (6%, scenario 3) deciding whether to perform a follow-up biopsy could be based on the clinical picture and age at diagnosis of CD, two parameters that predict risk of complications in coeliac patients(69). This means that patients at higher risk of complications, that is, those diagnosed after the age of 40 years and because of classical symptoms require a follow-up biopsy, whereas patients at low risk of complications, that is, those diagnosed aged <40 years old and presenting with non-classical symptoms or no symptoms could be followed-up non-invasively.

**Table 3.** Proposal for follow-up strategies based on different risks of persistent villous atrophy according to clinical response and gluten-free diet (GFD) adherence

<table>
<thead>
<tr>
<th>GFD adherence</th>
<th>Clinical response</th>
<th>Persistent well-being/complete recovery</th>
<th>Partial improvement/no significant improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strict</td>
<td>Score 4 – non-invasive follow-up</td>
<td>Score 1.5 – repeat biopsy</td>
<td>Score 0.5 – repeat biopsy</td>
</tr>
<tr>
<td>Poor</td>
<td>Score 3 – case-by-case evaluation based on age and clinical pattern at diagnosis</td>
<td>Score 0.5 – repeat biopsy</td>
<td>Score 0.5 – repeat biopsy</td>
</tr>
</tbody>
</table>

Dark grey, very high/high risk patients; light grey, intermediate risk; white, low risk.
Although we were not able to perform specifically a cost-effectiveness analysis and to internally validate our score, on the basis of our results, we can estimate that a follow-up duodenal biopsy may be avoided in more than 60% of patients and that it would be certainly necessary in only 10% of them. Another major strength of our study is that our score does not suffer from a selection bias because we did not refer patients for follow-up biopsy on the basis of their clinical picture, but we referred all of them indiscriminately. Moreover, GFD adherence and the clinical picture were evaluated with previously standardised and validated methods\(^{(24,28)}\). Limitations of the study include the low sample size and the single centre experience. Moreover, we have to underline that two different instruments for assessing adherence to the GFD were used for patients followed up before and after 2008.

A multivariable logistic model based on sex, clinical pattern of CD, serology at follow-up, GFD adherence and clinical response would have been the ideal instruments to foresee histological persistence of VA. However, given the low number of events, we were able to develop bivariable models only. We have to specify that in our model, we have not taken serology into consideration as a predictor to address follow-up of coeliac patients. Although normalisation of coeliac specific serology usually occurs on a GFD, it has been demonstrated that negative TTA cannot be used as a reliable marker of histological healing\(^{(31,32)}\). Our results confirm the need to focus our attention onto good dietary instruction in order to improve GFD adherence and on the investigation of non-invasive strategies to assess adherence to a GFD\(^{(24–26,33,34)}\). We think that these alternative methods to identify dietary lapses, including urinary gluten peptides, may be particularly useful for patients who do not require an invasive follow-up (scenario 4) and those in whom a case-by-case evaluation is needed (scenario 3)\(^{(35)}\).

The precise timing for histological reassessment in patients who do require a follow-up duodenal biopsy is another key issue. For the development of our score, we considered a time frame of 10–26 months since the diagnosis of CD. Therefore, our proposal may be that of evaluating histological recovery not earlier than 12–18 months since the beginning of a GFD.

Finally, we are aware that patients with mild symptoms or asymptomatic at time of diagnosis may be the most challenging to follow-up accurately, given the difficulty in the assessment of clinical response to a GFD. However, considering that risk of complications and mortality are not increased in these patients\(^{(3,10,35)}\), we think that strict GFD adherence is the main goal to achieve.

For the time being, follow-up of coeliac patients has been performed on the clinical experience of the different centres and there are no guidelines based on scientific studies. This study is the very first step towards the development of evidence-based criteria to tailor the follow-up of coeliac patients cost-effectively. Nonetheless, our score will need external validation and possibly a multicentre collaborative study, so as to consider more candidate predictors in a multivariate analysis.

Acknowledgements

A. S. is supported by an Investigational Fellowship by Collegio Ghislieri, Pavia, Italy (Assegno di ricerca annuale per giovani ricercatori).

F. B. is the guarantor of this article, and he is accepting full responsibility for the conduct of the study. He has had access to the data and has control of the decision to publish. A. S. and F. B. developed and organised the study; G. H., F. B. d. A., G. M. G. and S. A. collected the data; C. K. performed the statistical analysis; G. H., F. B. and A. S. wrote the manuscript. All authors approved the final version of the article, including the authorship list.

The authors declare that there are no conflicts of interest.

Supplementary material

For supplementary material referred to in this article, please visit https://doi.org/10.1017/S0007114520000537

References


