



OPEN ACCESS

Original research

# Persistent villous atrophy predicts development of complications and mortality in adult patients with coeliac disease: a multicentre longitudinal cohort study and development of a score to identify high-risk patients

Annalisa Schieppati ,<sup>1,2</sup> Stiliano Maimaris,<sup>1,2</sup> Suneil A Raju,<sup>3</sup> Olivia L Green,<sup>3</sup> Giulia Mantica,<sup>1</sup> Amelie Therrien,<sup>4</sup> David Flores-Marin,<sup>4</sup> Justin Linden,<sup>4</sup> Fernando Fernández-Bañares ,<sup>5,6</sup> Maria Esteve,<sup>5,6</sup> Daniel Leffler,<sup>4</sup> Federico Biagi,<sup>1,2</sup> David S Sanders<sup>3</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2023-329751>).

For numbered affiliations see end of article.

## Correspondence to

Dr Annalisa Schieppati, Gastroenterology Unit of Pavia Institute, Istituti Clinici Scientifici Maugeri, IRCCS, Via Salvatore Maugeri 10, 27100, Pavia, Lombardia, Italy; [annalisa.schieppati01@universitadipavia.it](mailto:annalisa.schieppati01@universitadipavia.it)

Received 19 February 2023

Accepted 1 June 2023

Published Online First

26 June 2023



► <https://dx.doi.org/10.1136/gutjnl-2023-329751>



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Schieppati A, Maimaris S, Raju SA, et al. *Gut* 2023;**72**:2095–2102.

## ABSTRACT

**Objective** Persistent villous atrophy (pVA) in coeliac disease (CD) despite a gluten-free diet (GFD) has unclear meaning. We aimed to (i) study the relationship between pVA and long-term outcomes and (ii) develop a score to identify patients at risk of pVA.

**Design** This is a multicentre retrospective-prospective study consisting of a study cohort (cohort 1) and an external validation cohort (cohort 2) of patients with biopsy-proven CD diagnosed between 2000 and 2021. Cohort 1 was used to (i) compare long-term outcomes between patients with and without pVA (Marsh  $\geq 3a$ ) at follow-up biopsy and (ii) to develop a score to evaluate the risk of pVA, which was validated in cohort 2.

**Results** Of 2211 patients, 694 (31%) underwent follow-up duodenal biopsy and were included in the study cohort (491F, 44 $\pm$ 16 years). 157/694 (23%) had pVA. Risk of complications (HR 9.53, 95% CI 4.77 to 19.04,  $p<0.001$ ) and mortality (HR 2.93, 95% CI 1.43 to 6.02,  $p<0.01$ ) were increased in patients with pVA. A 5-point score was developed and externally validated (receiver operating characteristic area under the curve 0.78, 95% CI 0.68 to 0.89) to stratify patients by risk of pVA: low (0–1 points, 5% pVA), intermediate (2 points, 16% pVA) and high (3–5 points, 73% pVA). Predictors for pVA used in the score were age at diagnosis  $\geq 45$  years (OR 2.01, 95% CI 1.21 to 3.34,  $p<0.01$ ), classical pattern of CD (OR 2.14, 95% CI 1.28 to 3.58,  $p<0.01$ ), lack of clinical response to GFD (OR 2.40, 95% CI 1.43 to 4.01,  $p<0.001$ ) and poor GFD adherence (OR 48.9, 95% CI 26.1 to 91.8,  $p<0.001$ ).

**Conclusions** Risk of complications and mortality were increased in patients with pVA. We developed a score to identify patients at risk of pVA and in need of histological reassessment and closer follow-up.

## INTRODUCTION

Coeliac disease (CD) is a lifelong immune-mediated enteropathy due to the ingestion of gluten in genetically predisposed individuals.<sup>1–3</sup> CD is characterised by a high prevalence in the general population,

## WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Persistent villous atrophy in patients with coeliac disease despite a gluten-free diet is a common but still poorly defined clinical scenario.
- ⇒ Persistent villous atrophy in coeliac disease has been associated with a significantly increased risk of lymphoproliferative disorders and other comorbidities such as osteoporotic hip fractures, but not with an increase in mortality.
- ⇒ The clinical phenotype of patients with coeliac disease at risk of persistent villous atrophy is undefined.

## WHAT THIS STUDY ADDS

- ⇒ Based on a multicentre international study, risk of complications and mortality were increased in patients with coeliac disease with persistent villous atrophy.
- ⇒ A 5-point score to stratify patients according to their risk of persistent villous atrophy was developed and validated based on age at diagnosis  $\geq 45$  years, classical pattern of coeliac disease, lack of clinical response to a gluten-free diet and poor gluten-free diet adherence.
- ⇒ Up to 40% of patients with persistent villous atrophy do not have ongoing symptoms.
- ⇒ Our score allows clinicians to identify these patients at higher risk even in the absence of ongoing symptoms.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ We propose a personalised and cost-effective approach for the follow-up of adult patients with coeliac disease: follow-up biopsy in patients at high risk (score  $\geq 3$ ); avoidance of follow-up biopsy in those at low risk (score 0–1); case-by-case evaluation in those at intermediate risk (score 2).

heterogeneous clinical picture and increased morbidity and mortality.<sup>1–7</sup> A lifelong gluten-free diet (GFD) is the mainstay for treatment in CD, leading to improvement of symptoms and intestinal lesions and preventing poor long-term outcomes associated with CD.<sup>1–3 5 8 9</sup>

However, villous atrophy (VA) can persist in some patients despite a GFD, usually due to poor dietary adherence or slow responsiveness to a GFD or hypersensitivity to gluten and, less commonly, premalignant/malignant complications of CD.<sup>10–18</sup> Current literature provides contrasting data on the relationship between persistent VA (pVA) and long-term outcomes in patients with CD. pVA in CD has been associated with a significantly increased risk of lymphoproliferative disorders<sup>19</sup> and other comorbidities such as osteoporotic hip fractures, but not with an increased mortality.<sup>7 20–23</sup> Furthermore, few studies have evaluated which factors may be associated with pVA in CD,<sup>16 24</sup> so the clinical phenotype of patients at higher risk of pVA is still poorly defined.

Although follow-up duodenal biopsy is not routinely performed in all patients at most centres, and despite the growing interest for non-invasive tools to predict mucosal healing, duodenal histology remains the gold standard for evaluating the response to a GFD. Major guidelines suggest that histological reassessment is needed when clinical response to a GFD is unsatisfactory,<sup>1–3 10 25</sup> although pVA has also been shown to frequently occur in asymptomatic patients on a GFD.<sup>14 18</sup> Currently, a non-invasive tool to identify patients at high risk of having pVA and thus possibly in need of a follow-up duodenal biopsy is not available. Therefore, we aimed to:

1. Evaluate whether pVA was associated with an increased risk of complications and mortality.
2. Study clinical predictors of pVA, and develop and validate a score to identify patients with CD at high risk of pVA.

## PATIENTS AND METHODS

### Study design and setting

This is a multicentre longitudinal study (partly retrospective and partly prospective) on adult patients with CD (age at diagnosis  $\geq 18$  years) followed-up at the participating centres. The study consisted of two separate cohorts of patients, as following: cohort 1, the study cohort, was used to (i) retrospectively evaluate whether a relationship existed between pVA and poor long-term outcomes (risk of complications and mortality) and (ii) identify clinical predictors of pVA and develop a score to identify patient at risk of having pVA; cohort 2, the validation cohort, was used to externally validate the predictive score for pVA on a separate real-world multicentre cohort.

### Study population

#### Cohort 1

This was a retrospective cohort including all the patients with CD directly diagnosed at three major referral centres (Sheffield, UK; Pavia, Italy; Boston, USA) between 2000 and 2020 and followed-up until February 2020, just before the COVID-19 pandemic outbreak. This multicentre dataset has data on >2200 patients with CD. We have previously published a longitudinal study specifically on seronegative CD using datasets from the Sheffield (UK), Harvard (USA) and Pavia (Italy) centres.<sup>26</sup> However, we specify that for the purpose of the present study only patients with CD directly diagnosed in one of the three centres were included. This means that patients diagnosed elsewhere and referred to one of these centres for confirming the diagnosis, or subsequent follow-up were excluded in order to

avoid a selection bias. Data about follow-up including development of complications and mortality were collected in each centre, as previously described.<sup>26</sup> This cohort was used (i) to investigate the relationship between persistence of VA and development of malignant complications of CD and mortality and (ii) to identify predictors of pVA and develop a score to identify patients at higher risk of pVA. Patients who did not undergo follow-up duodenal biopsy were excluded from the study cohort.

#### Cohort 2

This was a separate real-world multicentre cohort of patients with CD used for external validation of the score to predict risk of pVA. This cohort included further patients enrolled prospectively who underwent follow-up biopsy from March 2020 to October 2022 at the three centres that provided data for cohort 1 (Sheffield, UK; Pavia, Italy; Beth Israel Deaconess Medical Center-Harvard, USA). To increase the generalisability of our results, we also enrolled patients from a fourth centre (Terrassa, Spain) who underwent follow-up duodenal biopsy from 2015 to August 2022. Moreover, unlike for cohort 1, where only directly diagnosed patients were enrolled, in cohort 2 we also included referred patients in addition to patients who were directly diagnosed in each participating centre.

### Diagnostic criteria for CD

For both cohort 1 and cohort 2, diagnosis of CD was made in accordance with widely accepted criteria,<sup>1–3 26 27</sup> as follows. Conventional seropositive CD was diagnosed on the basis of VA (Marsh  $\geq 3a$ ) on correctly oriented duodenal biopsies taken from second duodenal portion + positive IgA tissue transglutaminase antibodies (TTA)/IgA endomysial antibodies (EmA)/IgA-IgG deamidated gliadin peptide antibodies (DGP) in the absence of IgA deficiency. In patients with total IgA deficiency (total serum IgA <8 mg/dL), CD was diagnosed based on VA and positive IgG TTA/EmA/DGP responding clinically and histologically to a GFD. Seronegative CD was diagnosed based on VA, negative TTA/EmA/DGP responding clinically and histologically to a GFD in the absence of other known causes of VA. For both seronegative CD and CD + IgA deficiency, HLA-DQ2 and DQ8 typing was necessary.

Patients with potential CD (ie, architecturally normal duodenal mucosa and positive coeliac serology)<sup>1–3</sup> were excluded from the study.

### Follow-up and mortality study

Data on follow-up and mortality for cohort 1 were retrieved from the dataset of the study we previously published on long-term outcomes in seronegative CD.<sup>26</sup> Clinical response to a GFD was defined as improvement or disappearance of symptoms and biochemical abnormalities present at baseline. Patients with evidence of VA (Marsh  $\geq 3a$ ) despite a GFD on follow-up duodenal biopsy were considered as having pVA. Evaluation of GFD adherence was based on dietetic interview or validated questionnaires (Coeliac Dietary Adherence Test/Pavia score).<sup>28 29</sup> Data about development of complications, death and cause of death were obtained from a large and regularly updated local multidataset at the Royal Hallamshire Hospital, Sheffield, UK and from a large registry of the Coeliac Centre, Beth Israel Deaconess Medical Center, Boston, USA, as previously described.<sup>26</sup> For patients belonging to the Italian cohort data were obtained from clinical notes, and for those not attending the clinic in the 6 months prior to data collection, information

about health status were obtained by means of phone contact or through the Local Council Services.<sup>26 30</sup>

Causes of death, when available, were considered related to CD when they included premalignant/malignant complications of CD,<sup>30 31</sup> and unrelated to CD in all the other cases. We considered as complications of CD malignant and premalignant conditions arising in the abdomen, including refractory CD type 1 and type 2, ulcerative jejuno-ileitis, abdominal lymphomas and small bowel carcinomas, as previously described.<sup>30 31</sup> In particular, persistence of malabsorption symptoms and VA despite a strict GFD for at least 12 months in the absence of lymphoma, malignancies and non-coeliac enteropathies allowed the diagnosis of refractory CD (RCD).<sup>1–3 30–32</sup> Distinction between type 1 and type 2 RCD was made on the basis of presence of an aberrant immune-phenotype of intraepithelial lymphocytes assessed by means of flow cytometry, immunohistochemistry and molecular genetics for  $\gamma$ -T-cell receptor clonality.<sup>1–3 30–32</sup> Diagnosis of malignancies (primary lymphomas and carcinomas of the small bowel) were based on histopathology.<sup>29–32</sup>

### Data collection

The following data were collected for all the patients: age at diagnosis of CD, sex, clinical pattern of CD at diagnosis,<sup>33</sup> type of CD (conventional seropositive CD, CD+IgA deficiency, seronegative CD), degree of adherence to a GFD, clinical response to a GFD, results of follow-up duodenal biopsy, time from diagnosis at which biopsy was performed, duration of follow-up, development of complications of CD and/or death and cause of death when available.

### STATISTICAL ANALYSIS

Statistical analysis was performed using R V.4.1.2 (R Core Team (2022). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>). Categorical variables were summarised as total counts and percentages, and compared among groups with Fisher's exact test. Continuous variables were summarised as mean and SD or median and IQR, for skewed data. Continuous variables were compared among groups using Student's t-test or the Mann-Whitney U test, as appropriate. Patients with missing data were excluded from relevant analyses and no imputation of data was performed.

Complication-free survival and overall survival were compared among patients with and without pVA by means of Kaplan-Meier curves and the log-rank test. Multivariate Cox regression models for complication-free survival and overall survival were fitted according to recovery/persistence of VA with major factors known as predictors<sup>5 16 24 30</sup> of developing malignant complications and mortality in CD considered in the analysis as covariates. Analyses were stratified by diagnosing centre. HRs and 95% CIs were calculated. In addition to preplanned analyses, a post hoc Cox regression analysis adjusted for age at diagnosis, sex, presenting pattern of CD and stratified by diagnosing centre was also conducted to further investigate for differences in mortality between patients who underwent follow-up biopsy and those who did not.

Logistic regression analysis was performed to identify independent predictors of pVA. Variables with  $p < 0.10$  on univariate analysis or which were considered clinically relevant on the basis of data previously reported in the literature were included in the multivariate analysis.<sup>5 16 24 30</sup> Adjusted ORs and 95% CIs were calculated. Model discriminatory ability was evaluated by calculating receiver operating characteristic area under the

curve (ROC-AUC) and 95% CI. Internal validation of the model was performed by means of 10-fold cross-validation. Based on predictors identified by logistic regression analysis for pVA, a weighted 5-point score for predicting risk of pVA was developed. Score discriminatory ability was evaluated by calculating ROC-AUC. Youden's index was used to identify the best cut-off score for predicting persistence of VA.

External validation of both the logistic regression model and the 5-point score for predicting persistence of VA were performed by assessing their discriminatory ability on the validation cohort by evaluating their respective ROC-AUC and 95% CI. Sensitivity analyses were performed to investigate for potential biases arising from the inclusion of patients with seronegative CD. A two-sided  $p$  value  $< 0.05$  was considered statistically significant for all analyses.

## RESULTS

### Clinical and demographic characteristics of retrospective cohort 1

Over a 20-year period, 2182 patients (1533 F; mean age at diagnosis  $42 \pm 16$  years) were directly diagnosed with CD in three major centres (2084 conventional seropositive CD, 59 seronegative CD, 39 CD+IgA deficiency), as previously described.<sup>26</sup> The flow chart in online supplemental figure 1 shows patients included and excluded from the study. Six-hundred and ninety-four patients with CD (491 F, mean age at diagnosis  $44 \pm 16$  years) underwent follow-up duodenal biopsy after a median time on a GFD of 32 months (IQR 15–61). The vast majority of these 694 patients (83.6%) were adherent to a GFD. At time of follow-up, duodenal biopsy symptoms present at diagnosis had improved in 65.9% of patients, whereas 34.1% still had ongoing symptoms; 157/694 (22.6%) had persistence of VA on follow-up biopsy, whereas the remaining 537 patients (77.4%) had resolution of VA. In 33/157 patients (21%) with pVA, complications of CD were diagnosed. None of the remaining 124 patients (79%) developed complications, and pVA in these patients was attributed to poor GFD adherence and/or slow responsiveness. Table 1 shows comparison of clinical characteristics between patients with and without pVA on follow-up duodenal biopsy.

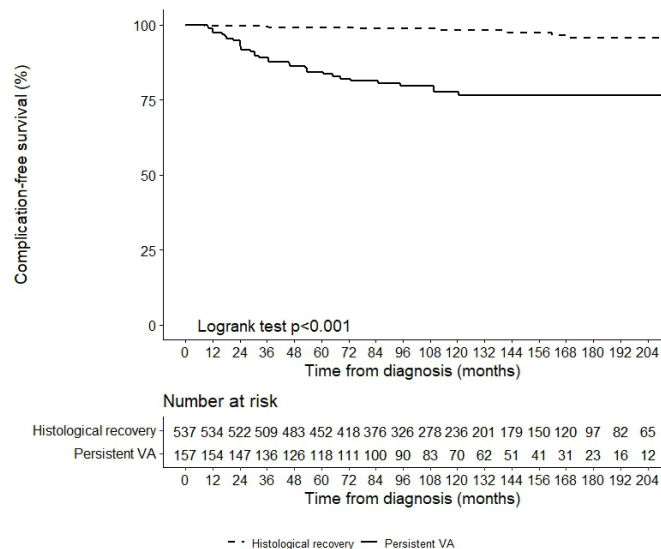
**Table 1** Comparison of clinical and demographic characteristics between patients with persistent VA and mucosal healing on follow-up duodenal biopsy

	Patients with persistent VA (n=157)	Patients with mucosal healing (n=537)	P value*
Age at diagnosis (years, mean $\pm$ SD)	47 $\pm$ 17	43 $\pm$ 15	<0.01
Gender, F	103 (65.6%)	388 (72.3%)	0.11
Classical pattern of CD	88 (56.1%)	263 (49.0%)	0.12
Clinical response to GFD	63 (40.1%)	394 (73.4%)	<0.001
Good GFD adherence	60 (38.2%)	520 (96.8%)	<0.001
Family history of CD	17 (10.9%)	74 (13.9%)	0.42
Associated autoimmune diseases	34 (21.7%)	133 (24.8%)	0.46
Seronegative CD†	20 (12.7%)	54 (10.1%)	0.38
Months on a GFD at follow-up duodenal biopsy (median, IQR)	30 (14–60)	32 (16–63)	0.32

\* $P < 0.05$  for statistical significance.

†Includes both patients with true seronegative CD and patients with CD associated with total IgA deficiency.

CD, coeliac disease; F, female; GFD, gluten-free diet; IQR, Interquartile range; SD, Standard Deviation; VA, villous atrophy.



**Figure 1** Kaplan-Meier curves comparing complication-free survival in patients with persistent villous atrophy (VA) (continuous line) and with histological recovery (dotted line) on follow-up duodenal biopsy. Five-year complication-free survival was 84.3% in patients with persistent VA on follow-up duodenal biopsy and 99.4% in those with mucosal healing. Ten-year complication-free survival was 77.7% in patients with persistent VA on follow-up duodenal biopsy and 98.4% in those with mucosal healing.

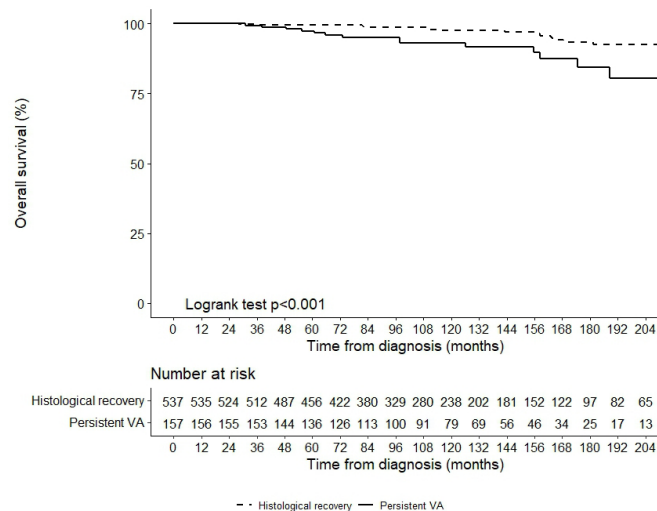
Age at diagnosis  $\geq 45$  years ( $p < 0.001$ ), lack of clinical response to a GFD ( $p < 0.001$ ) and poor GFD adherence ( $p < 0.001$ ) were associated with pVA at follow-up biopsy, while no significant association was found with sex or classical presentation of CD. Duration of a GFD before follow-up biopsy did not differ significantly between the two groups.

### Long-term outcomes

Median duration of follow-up in the study cohort was 110 months (IQR 74–160). Overall, during follow-up, 44 patients (6.3%) developed a complication and 32 patients died (4.6%). Complications of CD included type 1 refractory CD (25 patients), type 2 refractory CD (9 patients), enteropathy-associated T-cell lymphoma (4 patients), oesophageal cancer (3 patients), small bowel carcinoma, B-cell lymphoma and coeliac crisis (1 patient each). Causes of death were related to complications of CD in 10 patients, unrelated to CD in 12 patients and finally in 10 patients cause of death could not be ascertained.

### Complication-free survival and overall survival according to persistence of villous atrophy

Kaplan-Meier curves in [figure 1](#) and [figure 2](#) show complication-free survival and overall survival, respectively, which were both significantly lower in patients with pVA compared with those with mucosal recovery (both  $p < 0.001$ ). [Table 2](#) shows results of multivariate Cox models for complication-free survival and overall survival, according to age at diagnosis, pattern of clinical presentation of CD, sex and pVA/mucosal recovery. After stratification by centre and adjustment for age at diagnosis, pattern of clinical presentation of CD and sex, pVA was found to be independently predictive of both developing complications (HR 9.53, 95% CI 4.77 to 19.04,  $p < 0.001$ ) and mortality (HR 2.93, 95% CI 1.43 to 6.02,  $p < 0.01$ ).



**Figure 2** Kaplan-Meier curves comparing overall survival in patients with persistent villous atrophy (VA) (continuous line) and with histological recovery (dotted line) on follow-up duodenal biopsy. Five-year overall survival was 97.3% in patients with persistent VA on follow-up duodenal biopsy and 99.4% in those with mucosal healing. Ten-year complication-free survival was 93.1% in patients with persistent VA on follow-up duodenal biopsy and 97.5% in those with mucosal healing.

### Comparison between patients with and without follow-up duodenal biopsy

Patients who underwent follow-up duodenal biopsy were slightly older at diagnosis (mean  $44 \pm 16$  vs  $41 \pm 16$  years,  $p < 0.001$ ) and more frequently had a classical pattern of CD at diagnosis (50.6% vs 40.3%,  $p < 0.01$ ) than those who did not. There were no differences in gender. Patients who underwent follow-up duodenal biopsy developed complications of CD much more frequently than patients who did not, as shown in online supplemental figure 2. As shown in online supplemental figure 3, overall survival was slightly higher in patients who underwent follow-up biopsy ( $p < 0.05$ ). However, this difference was no longer significant (HR 0.70, 95% CI 0.44 to 1.12,  $p = 0.13$ ) at a post hoc multivariate Cox regression analysis adjusted for age at diagnosis, sex, presenting pattern of CD and stratified by diagnosing centre. Among patients who did not undergo follow-up duodenal biopsy, only 5 patients developed complications of CD, while 58 patients died. Causes of death were related to CD in

**Table 2** Results of multivariate Cox regression analysis investigating role of persistent VA in the development of complications and mortality

	HR (95% CI)	P value*
Development of complications		
Persistent VA at follow-up duodenal biopsy	9.53 (4.77 to 19.04)	<0.001
Age at diagnosis $\geq 45$ years	2.17 (1.10 to 4.29)	0.03
Classical pattern of CD	3.03 (1.52 to 6.02)	<0.01
Male sex	1.21 (0.65 to 2.23)	0.55
Mortality		
Persistent VA at follow-up duodenal biopsy	2.93 (1.43 to 6.02)	<0.01
Age at diagnosis $\geq 45$ years	9.45 (2.81 to 31.80)	<0.001
Classical pattern of CD	1.16 (0.57 to 2.36)	0.68
Male sex	1.24 (0.59 to 2.62)	0.57

\*P<0.05 for statistical significance.

CD, coeliac disease; VA, villous atrophy.

**Table 3** Results of multivariate logistic regression analysis investigating predictors of persistent villous atrophy at follow-up duodenal biopsy

	Adjusted OR (95% CI)	P value*
Age at diagnosis $\geq 45$ years	1.99 (1.21 to 3.27)	<0.01
Classical pattern of CD	2.18 (1.31 to 3.64)	<0.01
Lack of clinical response to GFD	2.32 (1.42 to 3.79)	<0.001
Poor GFD adherence	49.30 (26.30 to 92.20)	<0.001

\*P<0.05 for statistical significance.  
CD, coeliac disease; GFD, gluten-free diet.

4 patients, unrelated to CD in 42 (12 due to unrelated cancers, 9 due to cardiovascular diseases, 21 due to other heterogeneous causes) and unknown in 12 patients. Complications of CD in this group of patients included three B-cell lymphomas, one enteropathy-associated T-cell lymphoma and one small bowel carcinoma. Causes of death related to CD were significantly less frequent in patients who did not undergo follow-up duodenal biopsy than those who did ( $p<0.01$ ). Length of follow-up was also significantly longer among patients who underwent follow-up biopsy than those who did not (median 110 months (IQR 74–161) vs 60 (IQR 24–110),  $p<0.001$ ).

### Predictors of persistent VA

Multivariate logistic regression analysis (table 3) identified predictors of pVA, both at time of diagnosis and during follow-up. Classical presentation of CD at diagnosis ( $p<0.01$ ) and age at diagnosis  $\geq 45$  years ( $p<0.01$ ) as well as lack of clinical response to a GFD ( $p<0.001$ ) and poor GFD adherence ( $p<0.001$ ) were independent predictors of pVA at follow-up duodenal biopsy. Model discriminatory ability was very good (ROC-AUC 0.88, 95%CI 0.85 to 0.92). At internal validation

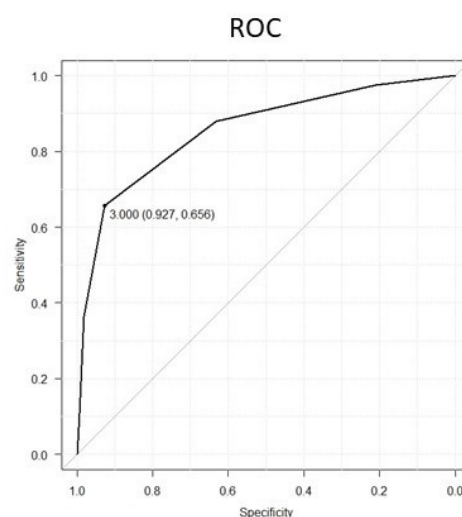
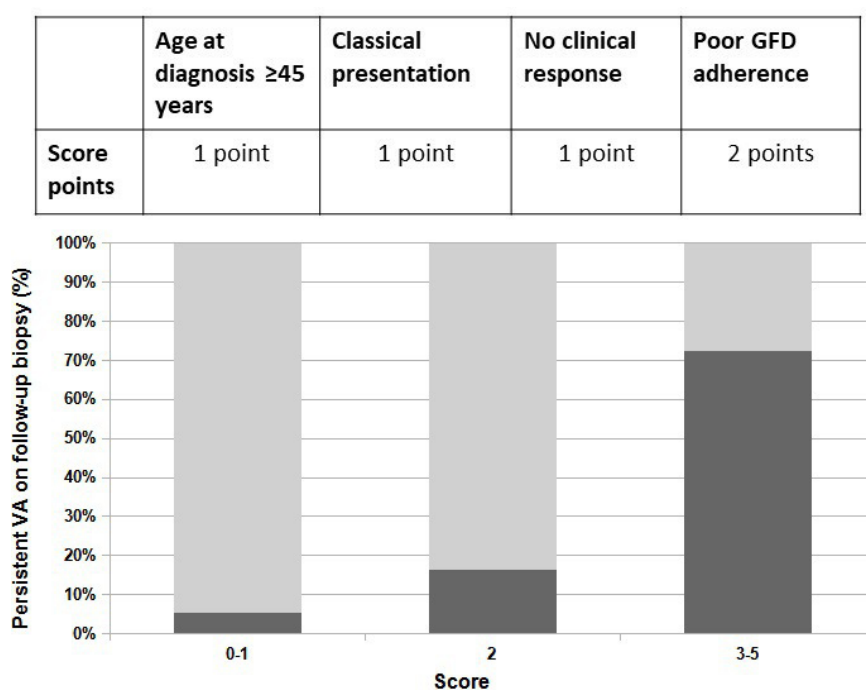
model, discriminatory ability remained very good after 10-fold cross-validation (ROC-AUC 0.86, 95% CI 0.83 to 0.89).

### Five-point score for predicting persistence of VA

Based on the logistic regression analysis results, we developed a weighted 5-point score (shown in figure 3) to stratify patients according to their risk of pVA. Patients could be categorised as low risk (score 0–1, 5% pVA), intermediate risk (score 2, 16% pVA) or high risk (score  $\geq 3$ , 73% pVA). Youden's index identified a score of 3 as the best cut-off point for identifying patients with and without pVA. On ROC analysis, the score showed very good predictive ability for pVA (ROC-AUC 0.86, 95% CI 0.82 to 0.89).

### External validation of the score on cohort 2

The validation cohort consisted of 144 patients with CD (113 F, mean age at diagnosis  $40 \pm 14$  years) who underwent follow-up duodenal biopsy after a median of 40 months (IQR 27–77) from diagnosis. Online supplemental table 1 shows a comparison of clinical and demographic features between cohort 1 and cohort 2. Twenty-six patients had persistence of VA (18.1%) while the remaining 118 (81.9%) showed resolution of VA, similarly to results of cohort 1 (pVA in 18.1% vs 22.6%,  $p=0.27$ ). Complications occurred in 5 patients (three type 1 refractory CD, one small bowel adenocarcinoma, one ampullary adenocarcinoma). Of these, the three patients with type 1 refractory CD had pVA while the patients with small bowel adenocarcinoma and ampullary adenocarcinoma did not have pVA. No patients died in the validation cohort. The logistic regression model to predict pVA maintained very good discriminatory ability (ROC-AUC 0.77, 95%CI 0.66 to 0.88) on the validation cohort. Likewise, the 5-point predictive score performed well in discriminating between patients with and without persistence of VA on the validation cohort (ROC-AUC 0.78, 95% CI 0.68 to 0.89).



**Figure 3** A 5-point clinical score to stratify patients according to their risk of having persistent VA at follow-up duodenal biopsy. A score of  $\geq 3$  identifies patients at high risk of persistent VA. The score is calculated by summing the points obtained for each item. Dark grey: patients with persistent VA; light grey: patients with mucosal healing. GFD, gluten-free diet; ROC, receiver operating characteristic; VA, villous atrophy.

### Sensitivity analyses

To evaluate for a potential bias in our results due to the inclusion of patients with seronegative CD, for which histological response to a GFD is crucial for diagnosis,<sup>26</sup> we repeated all analyses while excluding patients with seronegative CD. Results of all analyses remained substantially unchanged (data not shown).

### DISCUSSION

This multicentre study has identified the clinical phenotype of adult patients with CD at risk of having pVA despite a GFD, and has shown that persistence of VA predicted development of malignant complications and mortality in patients with CD. We have also developed and validated a score based on clinical data available at time of diagnosis of CD and during follow-up to identify patients at high risk of pVA, who therefore need follow-up biopsy. On the contrary, patients with a low score rarely had pVA, so follow-up biopsy can be avoided in these patients. Those with an intermediate risk should be evaluated on a case-by-case basis.

Although we cannot estimate the true prevalence of pVA in the present study, as not all the patients underwent follow-up biopsy, our results show that pVA despite a GFD is a common clinical scenario, occurring in roughly one-fifth of patients in a multicentre retrospective cohort spanning >20 years. Similar figures for pVA despite a GFD were also found in the separate external validation cohort. Previous estimates on the prevalence of pVA in the literature have shown variable results, with between 10% and 50% of patients with CD showing pVA even in the absence of symptoms while on a GFD.<sup>10–18</sup>

pVA was due to malignant complications of CD in 20% of cases, and more commonly due to poor GFD adherence (80%). Overall, the prevalence of complications in the present study was higher than previously reported, although comparisons are difficult to make due to differences in clinical and demographic aspects, and duration of follow-up.<sup>30–34</sup> It should be noted that roughly half of the complications in our cohort were due to refractory CD type 1, which is characterised by the best prognosis among the complications of CD, and for which it has recently been suggested that hidden gluten exposure may play a role.<sup>18,35</sup> Unfortunately, in this regard, we cannot provide data on measurement of minimal traces of gluten given the retrospective nature of our study. In the present study, follow-up biopsy was performed after a median time of 32 months (IQR 15–61), and there was no relationship between time to follow-up biopsy and persistence of VA, suggesting there was enough time for mucosal healing to occur. In this regard, previous studies have shown higher rates of incomplete mucosal recovery when follow-up biopsies were performed within 1–2 years from diagnosis.<sup>12,16,36</sup>

The result that pVA is associated with an increased risk of complications and mortality fills an important gap in the natural history of CD. In fact, this completes a pathogenetic sequence of events, in which persistent inflammation and tissue damage lead to poor long-term outcomes, as is known to occur in other chronic diseases such as inflammatory bowel disease.<sup>37,38</sup> Previous population-based studies showed an association between pVA and risk of lymphoproliferative disorders and osteoporotic hip fractures, but although they seemed to suggest an increased mortality, this last result was not statistically significant.<sup>7,12,13,16,19–23</sup> This difference in results may be related to differences in the study designs and populations under investigation and it could be hypothesised that in our cohort patients who underwent a follow-up duodenal biopsy may have been those with more severe disease (older at diagnosis and

with severe clinical symptoms despite a GFD), thus warranting a stricter follow-up. However, mortality in patients who underwent follow-up biopsy in our cohort was similar to those who did not, as confirmed at multivariate analysis. Despite this, causes of death related to CD were uncommon in patients without a follow-up duodenal biopsy. More precisely, the most common cause of death in patients who did not undergo follow-up biopsy were cancers unrelated to CD and cardiovascular disease. Unfortunately, a systematic evaluation of cardiovascular and other risk factors in these patients cannot be provided, as this was beyond the main aim of the study. Nevertheless, it may be that patients who underwent follow-up biopsy were more strictly followed up and this may have improved outcomes. However, we did not collect data on number of medical consultations to confirm this hypothesis. This result is in line with a previous Finnish study showing that patients with CD with more severe disease repeated a duodenal biopsy more commonly after 1 year since diagnosis, but overall repeating a duodenal biopsy did not impact on long-term outcomes.<sup>13</sup>

The association of pVA with other risk factors such as classical symptoms and advanced age at diagnosis ( $\geq 45$  years) delineates a specific subgroup of patients with a more severe disease phenotype, which may impact on the overall mortality of patients with CD and drive the slightly increase in mortality found in CD compared with the general population.<sup>7,28</sup> Nevertheless, poor GFD adherence was the most significant predictor of pVA in all analyses. This is not surprising and confirms that poor GFD adherence is a risk factor for poor outcomes in CD, including increased mortality.<sup>5</sup> This underlines the importance of strictly following a GFD for patients with CD. A personalised approach to implement and maintain GFD adherence, also over the long-term, should be part of the standard of care for patients with CD.<sup>9</sup> Lack of access to dietitians and personalised care still represents a major barrier to optimal care of patients with CD.

Our results have important implications for the clinical management of adult patients with CD, as they demonstrate that obtaining deep mucosal healing significantly impact disease progression and long-term outcomes. Noteworthy, the identification of a subgroup of patients at higher risk of pVA and poor long-term outcomes introduces the concept of a cost-effective and personalised follow-up of patients with CD based on specific disease phenotypes. This aspect is even more important if we consider that modalities for the follow-up of adult patients with CD are not standardised. Follow-up duodenal biopsy is usually not offered routinely to all patients at most centres and the decision on whether or not to repeat a duodenal biopsy is usually taken on a clinical basis, particularly when there is persistence of symptoms despite adherence to a GFD.<sup>1–3,10,39,40</sup>

Identification of potential clinical targets in the follow-up of patients with CD is a crucial requirement for clinicians. Based on our results, this includes implementing strategies for obtaining and maintaining over time strict adherence to a GFD,<sup>9</sup> control of persistent symptoms despite a GFD<sup>11,35,40</sup> and obtaining deep mucosal healing. A substantial degree of overlap between these aspects is likely to exist. Our results also confirm the findings of a prior study on smaller sample size that GFD adherence and clinical response to a GFD can be used to stratify patients who may need to repeat duodenal biopsy.<sup>41</sup> However, by identifying patients at risk of pVA our score allows clinicians to discriminate between patients likely to have underlying pVA and those with persistent symptoms likely unrelated to histological damage. Moreover, up to 40% of patients with pVA did not have ongoing symptoms. Our score also allows clinicians to identify these patients, who may nevertheless be at higher risk. Patients with

confirmed pVA may also be in need of more in-depth investigation, including capsule endoscopy and abdominal imaging to exclude ongoing complications and require closer clinical and dietetic follow-up to improve outcomes.

This study has some limitations, which possibly include lack of uniformity in methods used for diagnosis, evaluation of GFD adherence and follow-up modalities across different centres over >20 years. First, histological evaluation of duodenal architecture may have differed between centres, as some studies have suggested.<sup>42,43</sup> Although analyses were stratified by centre, it was not logistically feasible to share pathology material across centres to completely exclude potential biases in histological interpretation. Second, follow-up duodenal biopsies were not routinely performed in all patients, but only in those where it was clinically warranted as part of clinical care, thus not allowing a precise estimation of the prevalence of pVA. However, this reflects the real-world clinical practice and does not significantly limit the generalisability of our results. Thirdly, the inclusion of patients with seronegative CD in the present study may be of concern as histological recovery is one of the criteria for diagnosis of seronegative CD. However, we believe their inclusion increases the generalisability of our results, which also remained unchanged when analyses were repeated excluding seronegative patients. Finally, we could not evaluate the role of serology in predicting persistence of VA because we did not have standardised and centralised serological assessment due to the retrospective nature of cohort 1, with some participating centres mainly performing EmA while others mainly using TTA and/or DGP. Moreover, a recent meta-analysis showed low sensitivity of serology for identifying pVA.<sup>44</sup>

In conclusion, we have shown that patients with pVA have an increased risk of complications and mortality. Our score allows early identification of patients with CD at high risk of pVA who may require targeted interventions and personalised follow-up modalities to contrast poor long-term outcomes.

#### Author affiliations

<sup>1</sup>Department of Internal Medicine and Therapeutics, University of Pavia, Pavia, Italy

<sup>2</sup>Gastroenterology Unit of Pavia Institute, Istituti Clinici Scientifici Maugeri IRCCS, Pavia, Italy

<sup>3</sup>Academic Unit of Gastroenterology, Royal Hallamshire Hospital, Sheffield, UK

<sup>4</sup>Celiac Center, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA

<sup>5</sup>Department of Gastroenterology, Hospital Universitari Mutua Terrassa, Terrassa, Spain

<sup>6</sup>Centro de Investigación biomédica en red de enfermedades hepáticas y digestivas (CIBEREHD), Madrid, Spain

**Acknowledgements** This work was partially supported by the Ricerca Corrente funding scheme of the Italian Ministry of Health.

**Contributors** AS and DSS planned the study. AS, SM, SAR, GM, AT, DF-M, JL, FB, ME, DL, FB and DSS took care of the patients and collected the data. SM performed the statistical analysis. AS and SM interpreted the data and drafted the manuscript. All the authors critically revised and accepted the final version of the manuscript. Guarantor: Prof David S Sanders.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** DSS is the recipient of an educational grant from Dr Schär. AS and DL have received honoraria from Takeda Pharmaceuticals.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** Patients underwent clinical tests and assessments as part of their routine care. The Sheffield UK database was approved by the Yorkshire and the Humber Sheffield Research Ethics Committee, under registration number 14/YH/1216 as part of the Coeliac Disease Research Database. The database is used to identify efficiently and comprehensively patients eligible for a specific healthcare

intervention in order to help recruitment into trials, and for using routine clinical data to study the course of disease and effectiveness of healthcare used in daily coeliac clinical practice. The study was conducted in accordance with the Declaration of Helsinki (6th revision, 2008). The study protocol was approved by the Ethics Committee of IRCCS Pavia, ICS Maugeri, Pavia, Italy (protocol number CE2381 approved on 14th January 2020-amended on 05th April 2022), the BIDMC Ethical Board (2019P000927) and the Mutua-Terrassa Ethical Board (Code: EO/1011; approved on 25th March 2010; reviewed for the present project on 6th May 2022).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Annalisa Schieppatti <http://orcid.org/0000-0002-8493-7698>

Fernando Fernández-Bañares <http://orcid.org/0000-0002-1489-504X>

#### REFERENCES

- Ludvigsson JF, Bai JC, Biagi F, et al. British society of Gastroenterology. diagnosis and management of adult celiac disease: guidelines from the British society of Gastroenterology. *Gut* 2014;63:1210–28.
- Rubio-Tapia A, Hill ID, Kelly CP, et al. ACG clinical guidelines: diagnosis and management of celiac disease. *Am J Gastroenterol* 2013;108:656–76.
- Al-Toma A, Volta U, Auricchio R, et al. European society for the study of Coeliac disease (Esscd) guideline for coeliac disease and other gluten-related disorders. *United European Gastroenterol J* 2019;7:583–613.
- Singh P, Arora A, Strand TA, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clinical Gastroenterology and Hepatology* 2018;16:823–836.
- Corrao G, Corazza GR, Bagnardi V, et al. Mortality in patients with coeliac disease and their relatives: a cohort study. *Lancet* 2001;358:356–61.
- Biagi F, Corazza GR. Mortality in celiac disease. *Nat Rev Gastroenterol Hepatol* 2010;7:158–62.
- Lebwohl B, Green PHR, Söderling J, et al. Association between celiac disease and mortality risk in a Swedish population. *JAMA* 2020;323:1277–85.
- Holmes GK, Prior P, Lane MR, et al. Malignancy in coeliac disease - effect of a gluten free diet. *Gut* 1989;30:333–8.
- Schieppatti A, Maimaris S, Nicolardi ML, et al. Determinants and trends of adherence to a gluten-free diet in adult celiac patients on a long-term follow-up (2000–2020). *Clin Gastroenterol Hepatol* 2022;20:e741–9.
- Biagi F, Vattiato C, Agazzi S, et al. A second Duodenal biopsy is necessary in the follow-up of adult coeliac patients. *Ann Med* 2014;46:430–3.
- Leffler DA, Dennis M, Hyett B, et al. Etiologies and predictors of diagnosis in Nonresponsive celiac disease. *Clin Gastroenterol Hepatol* 2007;5:445–50.
- Pekki H, Kurppa K, Mäki M, et al. Predictors and significance of incomplete Mucosal recovery in celiac disease after 1 year on a gluten-free diet. *Am J Gastroenterol* 2015;110:1078–85.
- Pekki H, Kurppa K, Mäki M, et al. Performing routine follow-up biopsy 1 year after diagnosis does not affect long-term outcomes in coeliac disease. *Aliment Pharmacol Ther* 2017;45:1459–68.
- Kaukinen K, Peräaho M, Lindfors K, et al. Persistent small bowel Mucosal Villous atrophy without symptoms in coeliac disease. *Aliment Pharmacol Ther* 2007;25:1237–45.
- Lanzini A, Lanzarotto F, Villanacci V, et al. Complete recovery of intestinal mucosa occurs very rarely in adult coeliac patients despite adherence to gluten-free diet. *Aliment Pharmacol Ther* 2009;29:1299–308.
- Rubio-Tapia A, Rahim MW, See JA, et al. Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *Am J Gastroenterol* 2010;105:1412–20.
- Bardella MT, Velio P, Cesana BM, et al. Coeliac disease: a Histological follow-up study. *Histopathology* 2007;50:465–71.

- 18 Fernández-Bañares F, Beltrán B, Salas A, *et al.* Persistent Villous atrophy in de novo adult patients with celiac disease and strict control of gluten-free diet adherence: A multicenter prospective study (CADER study). *Am J Gastroenterol* 2021;116:1036–43.
- 19 Lebwohl B, Granath F, Ekbom A, *et al.* Mucosal healing and risk for lymphoproliferative malignancy in celiac disease: a population-based cohort study. *Ann Intern Med* 2013;159:169–75.
- 20 Lebwohl B, Michaëlsson K, Green PHR, *et al.* Persistent Mucosal damage and risk of fracture in celiac disease. *J Clin Endocrinol Metab* 2014;99:609–16.
- 21 Lebwohl B, Granath F, Ekbom A, *et al.* Mucosal healing and mortality in coeliac disease. *Aliment Pharmacol Ther* 2013;37:332–9.
- 22 Lebwohl B, Stephansson O, Green PHR, *et al.* Mucosal healing in patients with celiac disease and outcomes of pregnancy: a nationwide population-based study. *Clinical Gastroenterology and Hepatology* 2015;13:1111–1117.
- 23 Emilsson L, Lebwohl B, Green PH, *et al.* Mucosal healing and the risk of serious infections in patients with celiac disease. *United European Gastroenterol J* 2018;6:55–62.
- 24 Lebwohl B, Murray JA, Rubio-Tapia A, *et al.* Predictors of persistent Villous atrophy in coeliac disease: a population-based study. *Aliment Pharmacol Ther* 2014;39:488–95. 10.1111/apt.12621 Available: <http://doi.wiley.com/10.1111/apt.12621>.
- 25 Rej A, Elli L, Sanders DS. Persisting Villous atrophy and adherence in celiac disease: What does the patient want? What should a clinician advise? *Am J Gastroenterol* 2021;116:946–8.
- 26 Schieppatti A, Rej A, Maimaris S, *et al.* Clinical classification and long-term outcomes of Seronegative coeliac disease: a 20-year Multicentre follow-up study. *Aliment Pharmacol Ther* 2021;54:1278–89.
- 27 Schieppatti A, Sanders DS, Baiardi P, *et al.* Nomenclature and diagnosis of Seronegative coeliac disease and chronic non-coeliac Enteropathies in adults: the Paris consensus. *Gut* 2022;71:2218–25.
- 28 Biagi F, Andrealli A, Bianchi PI, *et al.* A gluten-free diet score to evaluate dietary compliance in patients with coeliac disease. *Br J Nutr* 2009;102:882–7.
- 29 Leffler DA, Dennis M, Edwards George JB, *et al.* A simple validated gluten-free diet adherence survey for adults with celiac disease. *Clinical Gastroenterology and Hepatology* 2009;7:530–536.
- 30 Biagi F, Schieppatti A, Maiorano G, *et al.* Risk of complications in coeliac patients depends on age at diagnosis and type of clinical presentation. *Dig Liver Dis* 2018;50:549–52.
- 31 Cellier C, Delabesse E, Helmer C, *et al.* Refractory sprue, celiac disease, and Enteropathy-associated T-cell lymphoma. *The Lancet* 2000;356:203–8.
- 32 Malamut G, Afchain P, Verkarre V, *et al.* Presentation and long-term follow-up of refractory celiac disease: comparison of type I with type II. *Gastroenterology* 2009;136:81–90.
- 33 Ludvigsson JF, Leffler DA, Bai JC, *et al.* The Oslo definitions for coeliac disease and related terms. *Gut* 2013;62:43–52.
- 34 Biagi F, Gobbi P, Marchese A, *et al.* Low incidence but poor prognosis of complicated coeliac disease: a retrospective Multicentre study. *Dig Liver Dis* 2014;46:227–30.
- 35 Penny HA, Rej A, Baggus EMR, *et al.* Non-responsive and refractory coeliac disease: experience from the NHS England national centre. *Nutrients* 2022;14:2776.
- 36 Galli G, Esposito G, Lahner E, *et al.* Histological recovery and gluten-free diet adherence: a prospective 1-year follow-up study of adult patients with coeliac disease. *Aliment Pharmacol Ther* 2014;40:639–47.
- 37 Cosnes J, Gower-Rousseau C, Seksik P, *et al.* Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* 2011;140:1785–94.
- 38 Colombel J-F, D'haens G, Lee W-J, *et al.* Outcomes and strategies to support a treat-to-target approach in inflammatory bowel disease: A systematic review. *J Crohns Colitis* 2020;14:254–66.
- 39 Ciacci C, Cirillo M, Cavallaro R, *et al.* Long-term follow-up of celiac adults on gluten-free diet: prevalence and correlates of intestinal damage. *Digestion* 2002;66:178–85.
- 40 Schieppatti A, Maimaris S, Lusetti F, *et al.* High prevalence of functional gastrointestinal disorders in celiac patients with persistent symptoms on a gluten-free diet: A 20-year follow-up study. *Dig Dis Sci* November 18, 2022.
- 41 Harder G, Schieppatti A, Biagi F, *et al.* 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. *Br J Nutr* 2020;123:1159–64.
- 42 Picarelli A, Borghini R, Donato G, *et al.* Weaknesses of histological analysis in celiac disease diagnosis: new possible scenarios. *Scand J Gastroenterol* 2014;49:1318–24.
- 43 Mubarak A, Nikkels P, Houwen R, *et al.* Reproducibility of the histological diagnosis of celiac disease. *Scand J Gastroenterol* 2011;46:1065–73.
- 44 Silvester JA, Kurada S, Szewajcer A, *et al.* Tests for serum Transglutaminase and endomysial antibodies do not detect most patients with celiac disease and persistent Villous atrophy on gluten-free diets: a meta-analysis. *Gastroenterology* 2017;153:689–701.