

**VALIDATION OF MORPHOMETRIC ANALYSES OF SMALL-INTESTINAL BIOPSY
READOUTS IN CELIAC DISEASE**

Juha Taavela
Syventävien opintojen opinnäyte
Tampereen yliopisto
Lääketieteen yksikkö
Keliakiaryhmä
Lokakuu 2013

Tampereen yliopisto
Lääketieteen yksikkö
Keliakiaryhmä

TAAVELA JUHA: OHUTSUOLEN LIMAKALVON HISTOLOGISTEN SUUREIDEN
VALIDOINTI KELIAKIASSA

Kirjallinen työ, 11s.
Ohjaaja: Professori Markku Mäki

Lokakuu 2013

Avainsanat: Nukkalisäke/kuopake-suhde, leikkaussuunta, lymfosyytitiheys

Keliakian diagnostiikan kulmakivi on edelleen gluteenin aiheuttama limakalvovaurio ohutsuolessa. Vaurion aste arvioidaan yleensä käyttäen luokiteltuja asteikkoja, mutta nämä ovat usein liian epätarkkoja jättäen pieniä merkittäviä muutoksia huomiotta. Tämän vuoksi on tullut tarve kehittää validoituja jatkuvia muuttujia limakalvovaurion arviointiin.

Arvioimme standardoidun toimintamenettelymme toimivuutta valikoidussa 93 potilaan ryhmässä. Näytteet sisälsivät eri asteisia limakalvovaurioita ja ne arvioitiin käyttäen jatkuvia muuttujia: nukkalisäke/kuopake-suhdetta ja valkosolutiheyttä. Kaksi akkreditoitua tutkijaa teki mittaukset sokkoutetusti. Tutkijan ja tutkijoiden välinen luotettavuus nukkalisäke/kuopake-suhteessa ja lymfosyytitiheydessä tutkittiin Bland-Altmanin metodilla ja luokan sisäisellä korrelaatiolla.

Molemmat tutkijat tunnistivat yhtenevästi palat, joita ei saanut toimintamenettelymme mukaan mitata. Bland-Altmanin-metodi ja luokan sisäinen korrelaatio osoittivat luotettavia tuloksia ja keskihajonnaksi saatiin tutkijan välisessä analyysissä 0,159 ja tutkijoiden välillä 0,227. Keskihajonnaksi CD3-positiivisessa valkosolutiheydessä saatiin tutkijan sisäisessä analyysissä 16,5% ja tutkijoiden välisessä analyysissä 17,1%.

Käytettäessä standardoitua toimintamenettelyämme saavutettiin luotettavia ja toistettavia tuloksia jatkuville limakalvovaurion mittareille. Kliinisesti merkittäviksi muuttoksiksi asetettiin keskihajonnan mukaisesti nukkalisäke/kuopake-suhteessa 0.4 ja CD3-positiivisessa valkosolutiheydessä 30%.

SISÄLLYSLUETTELO

1. VALIDATION OF MORPHOMETRIC ANALYSES OF SMALL-INTESTINAL BIOPSY READOUTS IN CELIAC DISEASE	4
2. Abstract	4
3. Introduction	4
4. Material and Methods	5
4.1 Study design and small-bowel mucosal specimens	5
4.2 Biopsy readout standard operating procedure	5
4.3 Effect of tangentially cut sections on the evaluation of villus height-crypt depth ratio	5
4.4 Statistics	5
4.5. Ethics	6
5. Results	6
6. Discussion	6
7. Conclusions	9
8. References	10

Validation of Morphometric Analyses of Small-Intestinal Biopsy Readouts in Celiac Disease

Juha Taavela¹, Outi Koskinen¹, Heini Huhtala², Marja-Leena Lähdeaho¹, Alina Popp^{1,3}, Kaija Laurila¹, Pekka Collin⁴, Katri Kaukinen^{4,5,6}, Kalle Kurppa¹, Markku Mäki^{1*}

1 Tampere Center for Child Health Research, University of Tampere and Tampere University Hospital, Tampere, Finland, **2** Tampere School of Health Sciences, University of Tampere, Tampere, Finland, **3** Institute for Mother and Child Care "Alfred Rusescu" and University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania, **4** Department of Gastroenterology and Alimentary Tract Surgery, Tampere University Hospital, Tampere, Finland, **5** Department of Medicine, Seinäjoki Central Hospital, Seinäjoki, **6** School of Medicine, University of Tampere, Tampere, Finland

Abstract

Background: Assessment of the gluten-induced small-intestinal mucosal injury remains the cornerstone of celiac disease diagnosis. Usually the injury is evaluated using grouped classifications (e.g. Marsh groups), but this is often too imprecise and ignores minor but significant changes in the mucosa. Consequently, there is a need for validated continuous variables in everyday practice and in academic and pharmacological research.

Methods: We studied the performance of our standard operating procedure (SOP) on 93 selected biopsy specimens from adult celiac disease patients and non-celiac disease controls. The specimens, which comprised different grades of gluten-induced mucosal injury, were evaluated by morphometric measurements. Specimens with tangential cutting resulting from poorly oriented biopsies were included. Two accredited evaluators performed the measurements in blinded fashion. The intraobserver and interobserver variations for villus height and crypt depth ratio (VH:CrD) and densities of intraepithelial lymphocytes (IELs) were analyzed by the Bland-Altman method and intraclass correlation.

Results: Unevaluable biopsies according to our SOP were correctly identified. The intraobserver analysis of VH:CrD showed a mean difference of 0.087 with limits of agreement from -0.398 to 0.224 ; the standard deviation (SD) was 0.159. The mean difference in interobserver analysis was 0.070, limits of agreement -0.516 to 0.375 , and SD 0.227. The intraclass correlation coefficient in intraobserver variation was 0.983 and that in interobserver variation 0.978. CD3⁺ IEL density countings in the paraffin-embedded and frozen biopsies showed SDs of 17.1% and 16.5%; the intraclass correlation coefficients were 0.961 and 0.956, respectively.

Conclusions: Using our SOP, quantitative, reliable and reproducible morphometric results can be obtained on duodenal biopsy specimens with different grades of gluten-induced injury. Clinically significant changes were defined according to the error margins (2SD) of the analyses in VH:CrD as 0.4 and in CD3⁺-stained IELs as 30%.

Citation: Taavela J, Koskinen O, Huhtala H, Lähdeaho M-L, Popp A, et al. (2013) Validation of Morphometric Analyses of Small-Intestinal Biopsy Readouts in Celiac Disease. PLoS ONE 8(10): e76163. doi:10.1371/journal.pone.0076163

Editor: Karol Sestak, Tulane University, United States of America

Received: July 1, 2013; **Accepted:** August 18, 2013; **Published:** October 11, 2013

Copyright: © 2013 Taavela et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study is supported by the Competitive State Research Financing of the Expert Responsibility Area of Tampere University Hospital, the Foundation for Pediatric Research, The Finnish Medical Foundation, the Finnish Celiac Society, the Academy of Finland, the Sigrid Juselius Foundation and Elna Kaarina Savolainen's fund allocated for the development of cancer treatment. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: markku.maki@uta.fi

Introduction

In celiac disease the characteristic gluten-induced small-intestinal mucosal injury develops gradually [1,2]. The spectrum of mucosal changes contains two separate measurable parameters, inflammation reflected by intraepithelial lymphocytic infiltration and morphological damage which includes villous atrophy and crypt hyperplasia. The mucosal lesion is the gold standard for diagnosing celiac disease and is present in patients both with and without clinical symptoms or signs of the disease. Well-known grouped classifications in histological assessment are one described by Marsh and modified by Oberhuber and another produced by Corazza and Villanacci [1,3,4]. These classifications are practical

in clinical work, but allocation to specific groups may be challenging and minor histologic changes are easily missed [4]. Such small but significant changes can be caused even by relatively small amounts of gluten and are particularly important when the clinical effectiveness of pharmacological therapies is evaluated.

Also noteworthy is that only a morphologically healed mucosa, and not the disappearance of symptoms, is a prerequisite for the long-term well-being of a patient [5]. The key issue is thus whether a potential drug or vaccine is able to prevent or attenuate the gluten-induced mucosal damage [6,7]. Optimally, histological evaluation should be made with validated and reliable readout tools [4]. Unfortunately, recent studies have shown poor reproducibility when using the results of grouped classification as

the primary outcome. Moreover, evaluators have given discrepant results in borderline findings, which are diagnostically the most critical part of the deterioration [8–10]. An important reason for these diagnostic problems might be incorrect orientation of biopsy specimens, leading to tangential cuttings and faulty interpretations [11–14]. Hence, validated standard operating procedures (SOPs) are needed to ensure reliable and reproducible results.

We here evaluated the quantitative morphological (villus height-crypt depth ratio, VH:CrD) and inflammatory (density of intraepithelial lymphocytes, IEL) variables used in the assessment of different degrees of damage in small-intestinal mucosal biopsies. The aim was to provide a standardized methodology and cut-off values for significant gluten-induced changes in the small-intestinal mucosa to be employed in routine clinical practice, and in academic and pharmacological studies in celiac disease.

Materials and Methods

Study design and small-bowel mucosal specimens

The study was conducted in the University of Tampere and Tampere University Hospital. The material comprised altogether 93 small-intestinal mucosal specimens from 72 patients, which were obtained from the prospectively collected database and biobank maintained by our study group. Altogether 16 specimens were obtained from newly diagnosed untreated celiac disease patients, 44 from patients on a gluten-free diet, 16 from patients who underwent gluten challenge and 17 specimens from non-celiac disease controls. The mean age of the celiac patients was 57 years (range 15–81) and 67% were women. The mean age of the non-celiac controls was 50 years (range 21–73 and 59% of them were women. The small-bowel biopsies were selected (see later) to represent variable stages of mucosal injury ranging from completely normal histology to overt mucosal atrophy and crypt hyperplasia. Further, specimens with good ($n=81$) or poor orientation ($n=12$) were included.

The forceps biopsy specimens were formalin-fixed and embedded in paraffin and another set of biopsies from the same patient were embedded in optimal cutting temperature compound, snap-frozen and stored at -70°C until used. For morphometric analyses, paraffin-embedded standard 2- μm -thick sections were processed and stained with hematoxylin-eosin (HE) and additional sections from the same biopsy block were stained with monoclonal CD3-specific antibody (CD3, Ab-2; Thermo Scientific, Waltham, MA). The small-intestinal mucosal VH:CrD was evaluated from at least three separate VH:CrD units by measuring villi lengths (μm) and crypt depths (μm), and the result was given as the average of the ratios. IEL densities were counted under light microscopy both in HE-stained sections [15] and in sections where T lymphocytes were stained with monoclonal CD3-specific antibodies [16]. At least 300 epithelial cells were counted in a continuous length of the epithelium, and results were expressed as number of IELs per 100 epithelial cells.

Frozen 5- μm -thick sections were processed and CD3⁺ IELs stained with monoclonal antibody Leu-4 (Becton Dickinson, San Jose, CA). The IELs were then counted with a 100 \times flat-field light-microscope objective throughout the surface epithelium; at least 30 fields measuring 1.6 mm in epithelial length were counted and the IEL density was expressed as cells/mm of epithelium [17].

The selected paraffin-embedded biopsy blocks comprised a wide spectrum of different degrees of mucosal injury and VH:CrD and density of CD3⁺ IELs were evaluated in parallel by two readers (JT,OK). Also, one evaluator had prior to the present evaluation studied and given results on all of these specimens (JT). The specimens were chosen by an independent selector (MM) and all

measurements were made by the evaluators in blinded fashion without knowing the results obtained by the other evaluator or the clinical data of the patients. One evaluator (JT) studied and reported results on the VH:CrD measurements and all IEL countings, while the other (OK) evaluated only the VH:CrD and IELs from CD3⁺-stained paraffin specimens.

Biopsy readout standard operating procedure

Our study group has adopted the morphometric small-intestinal mucosal biopsy measurements, as already described by Kuitunen and co-workers in 1982 [13], and subsequently formed the currently used SOP for histological readouts. SOP includes teaching technicians to obtain correctly oriented cuttings of biopsy specimens for morphometric evaluation. A crucial step in the procedure is that an accredited evaluator, besides producing acceptable interobserver and intraobserver morphometric results, be able to identify cases with an inadequate specimen and/or poor biopsy orientation, where measurements of villus-crypt units are not viable. It is essential that only biopsies in which the plane of sectioning is perpendicular to the luminal surface be considered, as judged by the fact that the crypts of Lieberkuhn are cut longitudinally and not in cross sections [11]. In a case of poor orientation resulting in tangential cuttings the evaluator asks for recuttings until reliable morphological readouts can be obtained. In IEL measurements the results are independent of biopsy orientation and recutting of specimens is rarely needed.

Effect of tangentially cut sections on the evaluation of villus height-crypt depth ratio

A case example from routine clinics illustrating the effect of tangential sectioning and subsequent recutting on morphological biopsy readouts is presented. To highlight the importance of correct sectioning, another formalin-fixed and paraffin-embedded biopsy specimen was cut twice so that the plane of sectioning was both perpendicular and tangential to the luminal surface. These latter biopsy slices were HE-stained and grouped in Marsh-Oberhuber classes (0, 1, 2, 3a, 3b, and 3c) by five independent pathologists who were blinded to each other's results and to the double cutting of the same biopsy block.

Further, a computerized 3D model was created to evaluate the effect of different planes of sectioning on the morphological readouts of the mucosa. First, computerized blocks were made representing small-intestinal mucosa in three different stages of morphological injury and sections of these blocks then cut in such a way that the plane of sectioning was both perpendicular and tangential to the luminal surface. The computer modelling program used was Autodesk Inventor 2008 (Autodesk, San Rafael, CA).

Statistics

Intraobserver and interobserver variations were analysed by the Bland-Altman method, linear regression analyses and intraclass correlation coefficients (ICC) [18,19]. In the Bland-Altman method, the differences between two quantitative measurements are plotted against the averages of the two measurements, and the results reported as the mean difference between the two measurements and limits of agreement, which are defined as the mean difference plus and minus twice the standard deviation of the differences. Twice the standard deviation was used as margin of error and a change exceeding this would thus be considered clinically significant. In the Bland-Altman plot, the x axis shows the mean of the results of the two measurements and the y axis represents the absolute difference between the two measurements.

When there is an increase in the variability of the differences as the magnitude of the measurements increases, differences on the y axis are recommended to be expressed as percentages of the values on the axis (i.e. proportionally to the magnitude of the measurements) [20]. Inter-method agreement was assessed with ICC. Quantitative data were expressed as number of subjects (n), mean and ranges.

Ethics

The patient recruitment and sample collection were approved by the Ethics Committee of Tampere University Hospital and all patients gave written informed consent.

Results

All 12 preselected biopsy specimens with tangential cuttings rendering them unacceptable for VH:CrD analyses were correctly identified by both study evaluators. The villus heights in the remaining 81 specimens with proper orientation ranged from 1 μm to 595 μm and the crypt depths from 152 μm to 535 μm . The subsequent mean value in VH:CrD was 1.39 (range 0.01–3.23). In IEL countings, five HE-stained and five CD3⁺ paraffin specimens and seven frozen specimens were excluded for technical reasons resulting from insufficient amounts of epithelial cells. The mean density of IELs in HE-stained specimens was 32 (range 8–82) per 100 epithelial cells, in paraffin-embedded CD3⁺-stained specimens 38 (range 9–103) per 100 epithelial cells and in CD3⁺-stained frozen specimens 68 (range 12–190) cells/mm.

The Bland-Altman plots illustrating the agreement in intraobserver and interobserver analyses in VH:CrD are shown in Figures 1A and 1B and the corresponding numerical values in Bland-Altman analyses are given in Table 1. In both intraobserver and interobserver VH:CrD analyses, the mean differences in the two measurement series compared were below 0.1, ensuring that there was no marked systematic error in the study. Twice the standard deviations, which represent the error range of the measurements, were 0.318 in intraobserver and 0.454 in interobserver analyses. The intraobserver and interobserver linear regressions are shown in Figures 1C and 1D and the ICCs in Table 1.

Bland-Altman statistics for the IEL countings are shown in Table 2. In the intraobserver analyses of IEL countings, the twice standard deviations representing the error ranges were 34.2% in CD3⁺ paraffin, 33.0% in CD3⁺ frozen and 53.2% in HE stained specimens.

The ICCs presented showed excellent agreement in the CD3⁺ stainings and the agreement was good in the HE countings (Table 2). Additionally, the inter-method comparison between frozen and paraffin-embedded CD3⁺ stainings showed an ICC of 0.679 ($p < 0.001$). These IEL densities from different fixations and stainings are mutually convertible and the following equations can be obtained from the linear regression analyses: IEL densities of paraffin-embedded CD3⁺ stainings (CD3Paraf) can be converted to the conventional IEL densities in HE-stained paraffin specimens (HEParaf) by the equation $\text{HEParaf} = 0.619 * \text{CD3Paraf} + 8.061$, CD3⁺ frozen specimens (CD3Fro) to HE paraffin IEL density by the equation $\text{HEParaf} = 0.442 * \text{CD3Fro} + 8.843$ and IEL densities from paraffin CD3⁺ to frozen CD3⁺ staining by the equation $\text{CD3Fro} = 1.466 * \text{CD3Paraf} + 10.416$.

The effect of incorrect biopsy orientation on histological evaluation is shown in Figure 2. In routine clinics, the pathologist interpreted the specimen in Figure 2A as normal mucosa, the villous structures seen thus excluding celiac disease. However, due to a high clinical suspicion of disease, biopsy re-evaluation was

asked for and recuttings performed by reason of original tangential cutting as judged by biopsy cross-sectioning of the crypts. In these reoriented specimens (Figure 2B), crypt hyperplasia with villous atrophy compatible with celiac disease was obvious. To study further this source of diagnostic error, our selected block from another patient was evaluated by five independent pathologists. All graded the tangential sectioning showing only cross-sections of crypts but tall villi as representing Marsh 0–1 (Figure 2C). When the same block had been cut perpendicular to the luminal surface, the crypts were correspondingly cut longitudinally and the same five pathologists graded the specimen as Marsh 3a–3b (Figure 2D). In contrast, no marked differences can be seen in the IEL densities in tangential and perpendicular cuttings (Figure 3).

In Figure 4, the computerized 3D-model demonstrates the effect of differing cuttings, i.e. perpendicular and tangential, in the evaluation of small-intestinal mucosal biopsy specimens. The hallmark of tangential cuttings is circular cross-sectioning of the mucosal crypts. Poor biopsy orientation results in incorrect interpretation of VH:CrD and/or incorrect Marsh class.

Discussion

The present validation study demonstrated that quantitative variables which measure separately gluten-induced morphological (VH:CrD) and inflammatory (IEL density) changes are reliable and reproducible. Morphometric variables can be utilized for outcome measurements in both academic studies and forthcoming drug trials in celiac disease. By following our SOP, we were able to ensure strict limits of agreement in the Bland-Altman method and excellent intraobserver and interobserver ICCs in VH:CrD in biopsies with different grades of small-intestinal mucosal injury. It is obvious that the forceps biopsy specimens obtained by endoscopy are small and difficult to orientate correctly. The cornerstone of our SOP has thus been to teach the accredited evaluator when morphological variables are not amenable to measurement. In all biopsy samples the sectioning of the crypts should be inspected as a part of routine in order to avoid the reading of tangential cuttings, wherein circular cross-sections of crypts can be used as hallmark. In the present study, both accredited evaluators were able to correctly identify the specimens not suitable for subsequent morphological analyses. We believe that our methodology for biopsy orientation and measurement of morphology can be easily applied to any standard laboratory settings.

VH:CrD measurement is a sensitive tool in detecting minor changes in gluten consumption [8,21]. We have previously used a change of 0.5 in VH:CrD as clinically relevant in assessing the effect of a gluten-free diet in celiac disease patients and also during gluten challenges [22]. The present findings would argue that this value could be reduced, since the mean difference was below 0.1 and twice the standard deviation was only 0.318 in the intraobserver Bland-Altman analysis. A cautious new cut-off value of 0.4 could thus be assigned to represent a clinically relevant difference between measurements. In the interobserver analysis again, the limit needed to obtain 95% of subjects was 0.454, indicating a slight observer-dependency in VH:CrD measurements. We infer that small but significant gluten-induced mucosal architectural changes can be reliably measured by morphometry.

The IEL densities in immunohistochemical CD3 stainings here showed better limits of agreement in the intraobserver analyses than the corresponding measurements in HE-stained specimens. In contrast, the ICC between IEL measurements of the paraffin and frozen specimens was good, and either one can be used for reliable results. It should also be noted that the frozen sections are

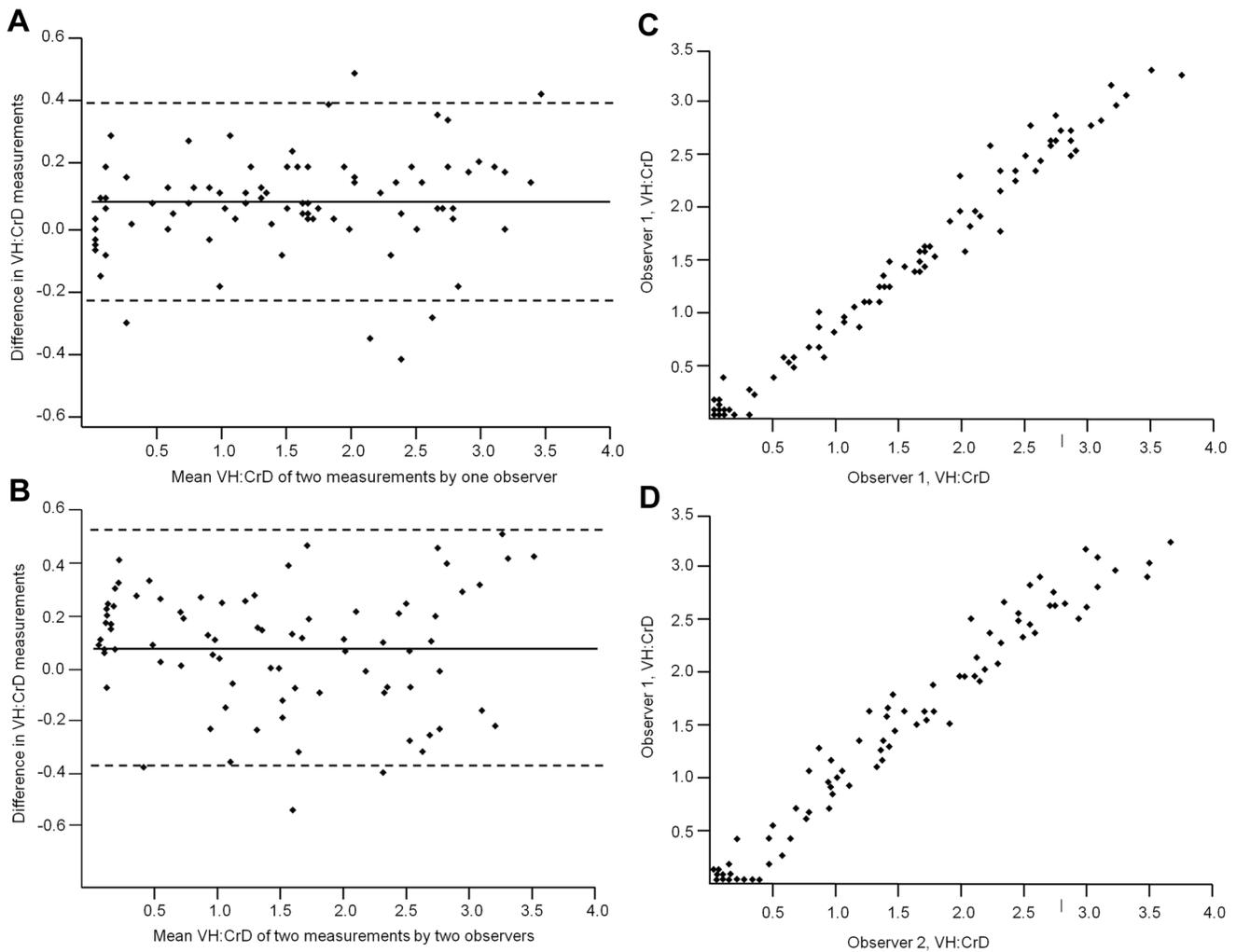


Figure 1. Graphical illustrations of the reliability and reproducibility of villus height crypt depth ratio. Panels A and B show the Bland-Altman plots of small-intestinal mucosal villus height crypt depth ratio and Panels C and D present the regression analyses for intraobserver and interobserver analyses, respectively. The solid lines in panels A and B indicate the mean difference between the measurements and the dashed lines correspond to the 95% limits of agreement.

doi:10.1371/journal.pone.0076163.g001

associated with better tissue preservation than the paraffin associated specimens. However, frozen samples are more burdensome due to specific technical requirements. We have used frozen samples in academic research in order to measure, in addition to CD3⁺ cells, the $\alpha\beta^+$ and $\gamma\delta^+$ T cell densities [16]. This also gives an opportunity for internal staining control, as the density of all T cells (CD3⁺) should be within the magnitude of the sum of $\alpha\beta^+$ and $\gamma\delta^+$ T cell densities. In the present study, twice the standard

deviation was in the intraobserver analysis in frozen specimens 33.0% and in paraffin specimens 34.2%. These findings agree well with previous studies in which a change of more than 30% in lymphocyte count has been considered clinically relevant [20]. The limits of agreement in the absolute IEL density values were wider in our study than those previously published, as we had substantially higher mean lymphocyte densities overall [23]. We thus provided values in percentages, as these are advocated when

Table 1. Bland-Altman statistics with absolute values and intraclass correlation coefficients (ICC) for analysing agreement and repeatability in small-bowel mucosal villus height crypt depth ratio (VH:CrD).

	Mean difference (95% CI)	Standard deviation	Limits of agreement	ICC
VH:CrD				
<i>Intraobserver</i>	0.087 (0.051 to 0.121)	0.159	−0.224 to 0.398	0.983
<i>Interobserver</i>	0.070 (0.020 to 0.121)	0.227	−0.375 to 0.516	0.978

CI, confidence interval.

doi:10.1371/journal.pone.0076163.t001

Table 2. Bland-Altman statistics with absolute and percentage values and intraclass correlation coefficients (ICC) for analysing agreement and repeatability in the density of intraepithelial lymphocytes (IELs) of paraffin CD3⁺, frozen CD3⁺ and hematoxylin-eosin (HE) stained small-bowel mucosal biopsy specimens.

	Mean difference (95% CI)	Standard deviation	Limits of agreement	ICC
Paraffin CD3+ IELs/100 enterocytes				
<i>Intraobserver</i>				
absolute values	1.9 (0.7 to 3.2)	5.9	-13.5 to 9.6	0.961
percentage values	5.1 (1.4 to 8.7)	17.1	-38.6 to 28.4	
<i>Interobserver</i>				
absolute values	6.9 (4.4 to 9.4)	11.6	-15.9 to 29.6	0.842
percentage values	24.1 (18.8 to 29.4)	24.7	-24.3 to 72.5	
Frozen CD3+ IELs/mm				
<i>Intraobserver</i>				
absolute values	2.0 (-0.3 to 4.3)	10.9	-19.4 to 23.4	0.956
percentage values	1.0 (-2.5 to 4.5)	16.5	-31.3 to 33.3	
HE IELs/100 enterocytes				
<i>Intraobserver</i>				
absolute values	0.7 (-2.6 to 1.3)	9.2	-17.4 to 18.7	0.854
percentage values	8.1 (-2.5 to 13.7)	26.6	-44.0 to 60.2	

CI, confidence interval.
doi:10.1371/journal.pone.0076163.t002

the variability in the measurements increases with higher IEL densities [19]. These percentage values would thus be more appropriate when assessing the effect of a gluten-free diet or novel drugs in celiac disease.

Soon after the discovery of the gluten-induced detrimental effect on the mucosal architecture in celiac disease it became clear that great care should be taken in orienting specimens in order to ensure representative vertical sections [11,24]. One of the main reasons for missing a celiac disease diagnosis is incorrect biopsy orientation, resulting in cross-sectioning of the crypts and thus loss of evidence of crypt hyperplasia [10,25]. In contrast, the IEL densities are mainly unaffected by the biopsy orientation as shown

here in Figure 3. The importance of having complete villus-crypt units and longitudinally cut crypts is highlighted by the exemplified biopsy cuttings seen in Figures 2 and 4. Further, when what are in reality low merged and convoluted villous ridges are cut tangentially the outcome is falsely normal-looking villi, as also seen in our computerized model (Figure 4). It is also possible that a normal mucosal architecture is falsely classified as celiac disease when the cutting is almost fully tangential giving the biopsy histological appearance of a “flat” lesion. Morphologically the appearance would be similar to a tangentially cut true “flat” lesion (Figure 4). False interpretation could occur especially if no cross-sections of villi are present. These examples further demonstrate

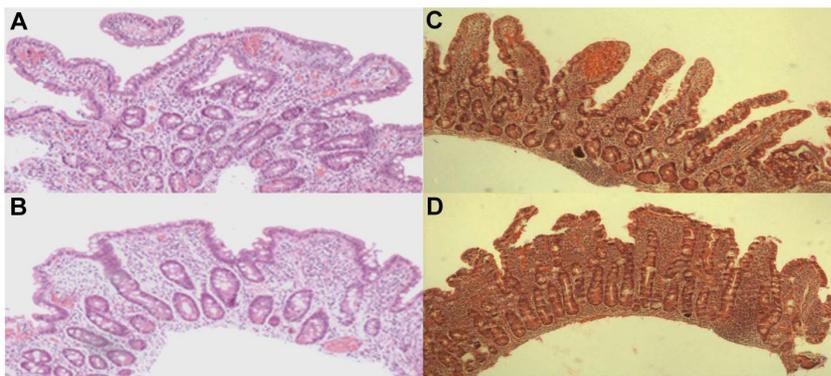


Figure 2. Two small-intestinal biopsy samples from routine clinics that show the importance of biopsy orientation in the interpretation of specimens. Sectionings A and C are cut tangential and B and D perpendicular to the luminal surface showing the effect of different orientation to the same small-intestinal mucosal biopsy block. The hallmark of tangential cutting is the cross-sectioning of the crypts while in correct vertical cutting the crypts are cut longitudinally. In routine clinics, the tangentially cut sectioning A was interpreted as normal, and on re-evaluation upon high clinical suspicion of celiac disease, the biopsy block was tilted and recut. The recut biopsy sample (B) reveals crypt hyperplasia and villous atrophy compatible with celiac disease. To further highlight this potential source of diagnostic error, five independent pathologists were asked to interpret another biopsy block with slices cut in different planes. All graded the tangential specimen C to be morphologically normal (Marsh 0–1) and the properly oriented specimen D to have villus atrophy and crypt hyperplasia (Marsh 3b or 3c).
doi:10.1371/journal.pone.0076163.g002

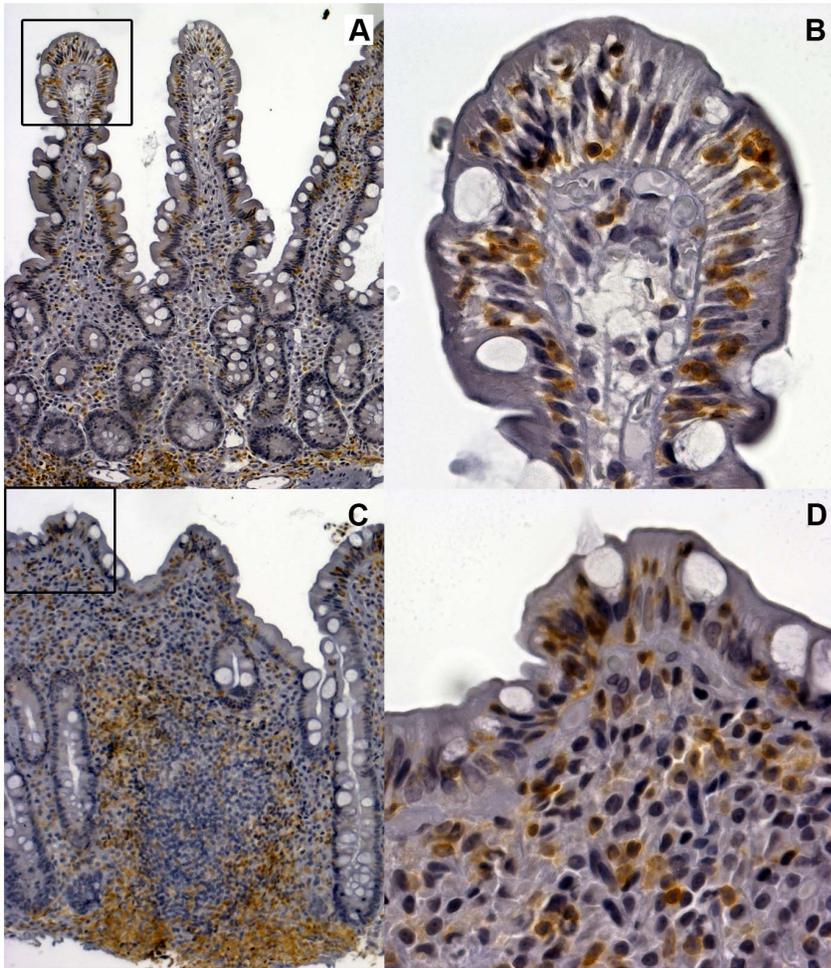


Figure 3. The intraepithelial lymphocyte densities are unaffected by the orientation of biopsy blocks. Panels A and C show tangential (A) and perpendicular (C) cuttings of paraffin embedded blocks stained for CD3⁺ T cells, i.e. the same biopsy block as in Figures 2C and 2D. The panels B and D present 40× magnified pictures of the mucosa from the places presented by the black rectangles in panels A and C. In these magnifications, the intraepithelial lymphocyte densities are 50 per 100 epithelial cells and 54 per 100 epithelial cells, respectively.
doi:10.1371/journal.pone.0076163.g003

that training of evaluators to correctly assess duodenal biopsy cuttings is essential in the diagnosis of celiac disease. One possibility to obtain good quality specimens is to orientate the biopsy on acetate cellulose filter before cutting [26].

There has recently been debate as to what should be the primary outcome when evaluating the effects of new treatments in celiac disease [6,27]. Although there is already an established mode of treatment, the gluten-free diet, there is an unmet need for novel pharmacological alternatives. Up to 60% of celiac disease patients may experience symptoms even on a strict gluten-free diet [28]. Also, mucosal healing may be incomplete with injury persisting in up to 60–80% of long-term treated patients in some series [29,30]. Such lingering mucosal damage predisposes to celiac disease-associated complications even if symptoms disappear on a gluten-free diet [5,31]. As a result, biopsy readouts should be favored in phase 2 clinical trials instead of self-perceived clinical outcomes. It is evident that if a drug is able to inhibit the gluten-induced mucosal injury similarly to a gluten-free diet, clinical symptoms, signs of malabsorption and serum celiac disease-specific antibodies will also disappear. In later phase 3 trials, well-validated patient-related and non-invasive biomarker outcomes can be

favored as endpoints and biopsies are needed only in subgroups of patients.

Conclusions

By following our SOP, excellent intraobserver and interobserver agreement in detecting small but significant changes in the small-intestinal mucosa can be achieved. The cut-off for clinically significant morphological duodenal mucosal healing or gluten-induced mucosal deterioration values as measured morphometrically (VH:CrD) was set as 0.4. A 30% or higher change in T-cell IEL densities, the marker of mucosal inflammation, can be regarded as clinically significant. The present results confirmed that proper orientation of biopsy specimens and recognition of incorrect tangential cuttings is crucial in ensuring reliable and reproducible histological results in celiac disease.

Acknowledgments

We thank Mr. Eero Heinonen for help in the creation of the 3D model with the Autodesk Inventor.

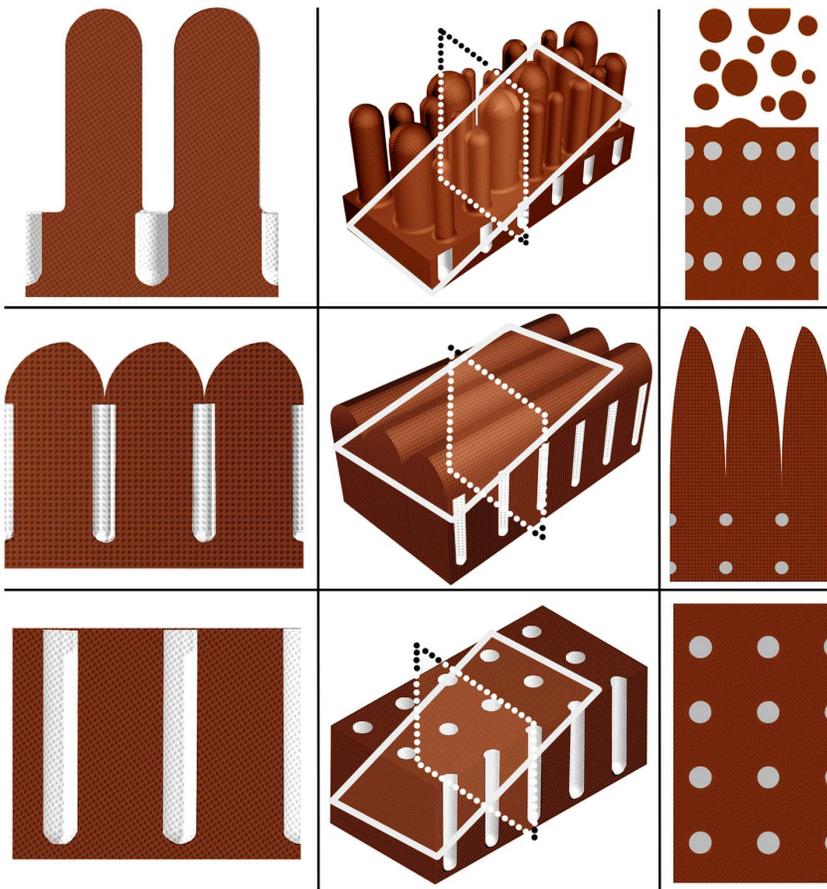


Figure 4. Computerized 3D-model demonstrates the effects of correct and incorrect planes of cutting on readout results. In the middle column are the biopsy blocks, in which the dashed and solid lines represent planes of sectioning. In the left column are sectionings cut perpendicular to the luminal surface and in the right tangentially cut sectionings. For example, in the middle row the computerized block shows merged and convoluted low villous ridges which in perpendicular cutting results in subtotal villous atrophy with deep crypts (left) but in tangential cutting in tall villi with only cross-sections of crypts (right).
doi:10.1371/journal.pone.0076163.g004

Author Contributions

Performed the experiments: JT OK. Analyzed the data: JT HH. Contributed reagents/materials/analysis tools: JT KL. Wrote the paper: JT. Conception and design of the work: JT OK HH ML AP KL PC KK

KK MM. Drafted the manuscript, revised it critically for important intellectual content and gave final approval of the version to be published: JT OK HH ML AP KL PC KK KK MM.

References

- Marsh MN. (1992) Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* 102: 330–354.
- Kurppa K, Collin P, Viljamaa M, Haimila K, Saavalainen P, et al. (2009) Diagnosing mild enteropathy celiac disease: a randomized, controlled clinical study. *Gastroenterology* 136: 816–823.
- Oberhuber G. (2000) Histopathology of celiac disease. *Biomed Pharmacother* 54: 368–372.
- Corazza GR, Villanacci V. (2005) Coeliac disease. *J Clin Pathol* 58: 573–574.
- Rubio-Tapia A, Rahim MW, Sec JA, Lahr BD, Wu TT, et al. (2010) Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *Am J Gastroenterol* 105: 1412–1420.
- Lähdeaho ML, Lindfors K, Airaksinen L, Kaukinen K, Mäki M. (2012) Recent advances in the development of new treatments for celiac disease. *Expert Opin Biol Ther* 12: 1589–1600.
- Sollid LM, Khosla C. (2011) Novel therapies for coeliac disease. *J Intern Med* 269: 604–613.
- Corazza GR, Villanacci V, Zambelli C, Milione M, Luinetti O, et al. (2007) Comparison of the interobserver reproducibility with different histologic criteria used in celiac disease. *Clin Gastroenterol Hepatol* 5: 838–843.
- Mubarak A, Nikkels P, Houwen R, Ten Kate F. (2011) Reproducibility of the histological diagnosis of celiac disease. *Scand J Gastroenterol* 46: 1065–1073.
- Arguelles-Grande C, Tennyson CA, Lewis SK, Green PH, Bhagat G. (2011) Variability in small bowel histopathology reporting between different pathology practice settings: impact on the diagnosis of coeliac disease. *J Clin Pathol* 65: 242–247.
- Risdon RA, Keeling JW. (1974) Quantitation of the histological changes found in small intestinal biopsy specimens from children with suspected coeliac disease. *Gut* 15: 9–18.
- Shidrawi RG, Przemioslo R, Davies DR, Tighe MR, Ciclitira PJ. (1994) Pitfalls in diagnosing coeliac disease. *J Clin Pathol* 47: 693–694.
- Thurlbeck WM, Benson JA Jr., Dudley HR Jr. (1960). The histopathologic changes of sprue and their significance. *Am J Clin Pathol* 34: 108–117.
- Collin P, Kaukinen K, Vogelsang H, Korponay-Szabó I, Sommer R, et al. (2005) Antiendomysial and antihuman recombinant tissue transglutaminase antibodies in the diagnosis of coeliac disease: a biopsy-proven European multicentre study. *Eur J Gastroenterol Hepatol* 17: 85–91.
- Kuitunen P, Kosnai I, Savilahti E. (1982) Morphometric study of the jejunal mucosa in various childhood enteropathies with special reference to intraepithelial lymphocytes. *J Pediatr Gastroenterol Nutr* 1: 525–531.
- Patey-Mariaud De Serre N, Cellier C, Jabri B, Delabesse E, Verkarre V, et al. (2000) Distinction between coeliac disease and refractory sprue: a simple immunohistochemical method. *Histopathol* 37: 70–77.
- Järvinen TT, Kaukinen K, Laurila K, Kyrönpalo S, Rasmussen M, et al. (2003) Intraepithelial lymphocytes in celiac disease. *Am J Gastroenterol* 98: 1332–1337.

18. Bland JM, Altman DG. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
19. Bartko JJ. (1966) The intraclass correlation coefficient as a measure of reliability. *Psychol Rep* 19: 3–11.
20. Pollock MA, Jefferson SG, Kane JW, Lomax K, MacKinnon G, et al. (1992) Method comparison - a different approach. *Ann Clin Biochem* 29: 556–560.
21. Lähdeaho M, Mäki M, Laurila K, Huhtala H, Kaukinen K. (2011) Small-bowel mucosal changes and antibody responses after low- and moderate-dose gluten challenge in celiac disease. *BMC Gastroenterol* 11:129.
22. Kaukinen K, Peräaho M, Collin P, Partanen J, Woolley N, et al. (2005) Small-bowel mucosal transglutaminase 2-specific IgA deposits in coeliac disease without villous atrophy: A prospective and randomized clinical study. *Scand J Gastroenterol* 40: 564–572.
23. Veress B, Franzén L, Bodin L, Borch K. (2004) Duodenal intraepithelial lymphocyte-count revisited. *Scand J Gastroenterol* 39: 138–144.
24. Shiner M. (1964) Problems in interpretation of intestinal mucosal biopsies. *JAMA* 188: 45–48.
25. Rostom A, Murray JA, Kagnoff MF. (2006) American Gastroenterological Association (AGA) institute technical review on the diagnosis and management of celiac disease. *Gastroenterology* 131: 1981–2002.
26. Ravelli A, Villanacci V. (2012) Tricks of the trade: How to avoid histological pitfalls in celiac disease. *Pathol Res Pract* 208: 197–202.
27. Taavela J, Kurppa K, Collin P, Lähdeaho ML, Salmi T, et al. (2013) Degree of damage to the small bowel and serum antibody titers correlate with clinical presentation of patients with celiac disease. *Clin Gastroenterol Hepatol* 11: 166–171.
28. Midhagen G, Hallert C. (2003) High rate of gastrointestinal symptoms in celiac patients living on a gluten-free diet: controlled study. *Am J Gastroenterol* 98: 2023–2026.
29. Bardella MT, Velio P, Cesana BM, Prampolini L, Casella G, et al. (2007) Coeliac disease: a histological follow-up study. *Histopathol* 50: 465–471.
30. Lee SK, Lo W, Memco L, Rotterdam H, Green PH. (2003) Duodenal histology in patients with celiac disease after treatment with a gluten-free diet. *Gastrointest Endosc* 57: 187–191
31. Kaukinen K, Peräaho M, Lindfors K, Partanen J, Woolley N, et al. (2007) Persistent small bowel mucosal villous atrophy without symptoms in coeliac disease. *Aliment Pharmacol Ther* 25: 1237–1245.