Article:
Shiha, M.G. orcid.org/0000-0002-2713-8355, Nandi, N., Raju, S.A. et al. (7 more authors) (2024) Accuracy of the no-biopsy approach for the diagnosis of coeliac disease in adults: a systematic review and meta-analysis. Gastroenterology, 166 (4). pp. 620-630. ISSN 0016-5085

https://doi.org/10.1053/j.gastro.2023.12.023

Accuracy of the no-biopsy approach for the diagnosis of coeliac disease in adults: a systematic review and meta-analysis

Mohamed G. Shiha, Nicoletta Nandi, Suneil A. Raju, Graeme Wild, Simon S. Cross, Prashant Singh, Luca Elli, Govind K. Makharia, David S. Sanders, Hugo A. Penny

PII: S0016-5085(23)05688-3
DOI: https://doi.org/10.1053/j.gastro.2023.12.023
Reference: YGAST 66066

To appear in: *Gastroenterology*
Accepted Date: 19 December 2023

Please cite this article as: Shiha MG, Nandi N, Raju SA, Wild G, Cross SS, Singh P, Elli L, Makharia GK, Sanders DS, Penny HA, Accuracy of the no-biopsy approach for the diagnosis of coeliac disease in adults: a systematic review and meta-analysis, *Gastroenterology* (2024), doi: https://doi.org/10.1053/j.gastro.2023.12.023.

Systematic Review and Meta-analysis

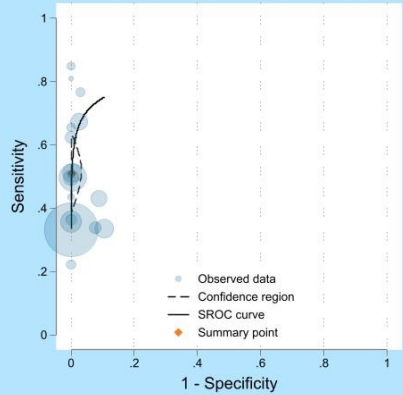
Accuracy of the no-biopsy approach in adults with suspected coeliac disease

tTG ≥ 10xULN *vs* **Duodenal biopsy Marsh ≥2**

18 Studies

15 Countries

12,103 Participants



51%
Sensitivity

100%
Specificity

98%
PPV

Selected adults with IgA-tTG ≥10×ULN and a moderate to high pre-test probability of coeliac disease could be diagnosed based on serology

Gastroenterology

Accuracy of the no-biopsy approach for the diagnosis of coeliac disease in adults: a systematic review and meta-analysis

Mohamed G Shiha[1,2], Nicoletta Nandi[1,3], Suneil A Raju[1,2], Graeme Wild[4], Simon S Cross[5], Prashant Singh[6], Luca Elli[3], Govind K Makharia[7], David S Sanders[1,2], Hugo A Penny[1,2]

[1]Academic Unit of Gastroenterology, Sheffield Teaching Hospitals, Sheffield, UK

[2]Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, UK

[3]Center for Prevention and Diagnosis of Celiac Disease, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[4]Department of Immunology, Sheffield Teaching Hospitals, Sheffield, UK

[5]Department of Histopathology, Sheffield Teaching Hospitals, Sheffield, UK

[6]Division of Gastroenterology, Department of Medicine, University of Michigan, Michigan, USA

[7]Department of Gastroenterology and Human Nutrition, All India Institute of Medical Sciences, New Delhi, India

Corresponding author: Dr Mohamed Shiha, Academic Unit of Gastroenterology, Royal Hallamshire Hospital, Sheffield S10 2JF, UK; Mohamed.shiha1@nhs.net

Conflict of interest:

Mohamed G Shiha: Received speaker honorarium from Thermo Fisher.

Nicoletta Nandi: No conflict of interest to declare.

Suneil A Raju: No conflict of interest to declare.

Graeme Wild: No conflict of interest to declare.

Simon S Cross: No conflict of interest to declare.

Prashant Singh: No conflict of interest to declare.

Luca Elli: No conflict of interest to declare.

Govind K Makharia: No conflict of interest to declare.

David S Sanders: No conflict of interest to declare.

Hugo A Penny: Received speaker honorarium from Thermo Fisher.

Authors contribution

Conception: MGS, DSS, HAP. Literature search and data extraction: MGS, NN. Statistical analysis: MGS. Risk of bias assessment: MGS, RAS. Initial drafting of the manuscript: MGS. Data interpretation, critical revision of the manuscript and final approval of the submitted version: all authors.

Funding

HAP funded by a Clinical Lecturers grant (CL-2021-04-002) from the NIHR, PS is supported by K23DK129327 grant from NIDDK

Data Availability

Data used in this meta-analysis are publicly available.

Abstract

Background & Aims

Current international guidelines recommend duodenal biopsies to confirm the diagnosis of coeliac disease in adult patients. However, growing evidence suggests that IgA anti-tissue transglutaminase (tTg) antibody levels ≥10 times the upper limit of normal (ULN) can accurately predict coeliac disease, eliminating the need for biopsy. We performed a systematic review and meta-analysis to evaluate the accuracy of the no-biopsy approach to confirm the diagnosis of coeliac disease in adults.

Methods

We systematically searched MEDLINE, EMBASE, Cochrane Library and Web of Science from January 1998 to October 2023 for studies reporting the sensitivity and specificity of IgA-tTG ≥10×ULN against duodenal biopsies (Marsh grade ≥2) in adults with suspected coeliac disease. We used a bivariate random-effects model to calculate the summary estimates of sensitivity, specificity, positive and negative likelihood ratios. The positive and negative likelihood ratios were used to calculate the positive predictive value (PPV) of the no-biopsy approach across different pre-test probabilities of coeliac disease. The methodological quality of the included studies was evaluated using the QUADAS-2 tool. This study was registered with PROSPERO, number CRD42023398812.

Results

A total of 18 studies comprising 12,103 participants from 15 countries were included. The pooled prevalence of biopsy-proven coeliac disease in the included studies was 62% (95% CI, 40% - 83%).

The proportion of patients with IgA-tTG ≥10×ULN was 32% (95% CI, 24% - 40%). The summary sensitivity of IgA-tTG ≥10×ULN was 51% (95% CI, 42% - 60%), and the summary specificity was 100% (95% CI, 98% - 100%). The area under the summary receiver operating characteristic curve was 0.83 (95% CI, 0.77 – 0.89). The PPV of the no-biopsy approach to identify patients with coeliac disease was 65%, 88%, 95%, and 99% if coeliac disease prevalence was 1%, 4%, 10% and 40%, respectively. Between-study heterogeneity was moderate (I^2 =30.3%), and additional sensitivity analyses did not significantly alter our findings. Only one study had a low risk of bias across all domains.

Conclusion

The results of this meta-analysis suggest that selected adult patients with IgA-tTG ≥10×ULN and a moderate to high pre-test probability of coeliac disease could be diagnosed without undergoing invasive endoscopy and duodenal biopsy.

Keywords:

Adult; Biopsy; Celiac Disease; Humans; Immunoglobulin A; Transglutaminases

What you need to know

Background and context

The diagnosis of coeliac disease in adults currently involves a two-step process, starting with the detection of tissue transglutaminase (tTG) antibodies and/or serum endomysial antibodies (EMA), followed by a confirmatory endoscopy and duodenal biopsy. Due to the increased accuracy of serological tests, paediatric guidelines adopted a no-biopsy approach, whereby children with IgA-tTG levels ≥10 times the upper limit of normal (ULN) and positive EMA can be diagnosed with coeliac disease without biopsy. However, applying this no-biopsy approach to diagnose adult patients with coeliac disease is highly controversial.

New findings

In a meta-analysis of 18 studies with >12,000 adult participants, we found that IgA-tTG levels ≥10×ULN are highly indicative of coeliac disease in adult patients referred to secondary care with a 100% specificity and a positive predictive value of 98%. The predictive value of the no-biopsy approach varies according to the prevalence of coeliac disease in the studied population.

Limitations

All studies were conducted in secondary and tertiary care settings, and results may not be generalizable to primary care.

Impact

The no-biopsy approach could lead to a shorter time to diagnosis, increased patient satisfaction and reduced healthcare costs.

Lay summary

The diagnosis of coeliac disease involves blood tests, to detect elevated antibodies triggered by eating gluten, and endoscopy with biopsy from the small intestine to prove the damage to the intestinal lining. We found that when the levels of the antibodies are very high, damage to the intestinal lining is almost certain, and endoscopy may not be required in all cases.

Introduction

Coeliac disease is a common autoimmune disorder characterised by an immunological response to dietary gluten in genetically susceptible individuals[1]. Although it is estimated that coeliac disease affects nearly 60 million people worldwide, the majority of patients remain undiagnosed, misdiagnosed or experience significant diagnostic delays[2]. Undiagnosed coeliac disease is associated with significant morbidity, reduced quality of life, and serious long-term complications such as increased risks of osteoporosis, cardiovascular diseases, and cancers[1,3]. Currently, the diagnosis of coeliac disease in adults is based on a combination of serological testing followed by endoscopy and duodenal biopsy to confirm the diagnosis[4]. However, this approach is invasive, expensive, and often associated with long waiting times, which can delay diagnosis and treatment.

In recent years, there have been significant advancements in the diagnostic accuracy of serological tests for coeliac disease. These have led to a step change in the paediatric guidelines, whereby children with IgA anti-tissue transglutaminase (tTG) antibody levels ≥10 times the upper limit of normal (ULN) along with positive endomysium antibodies (EMA) can be diagnosed with coeliac disease without a confirmatory duodenal biopsy[5]. A subsequent prospective multicentre study confirmed the reliability of the no-biopsy approach to diagnose coeliac disease in children with a positive predictive value (PPV) of >99%[6].

Despite the evidence supporting the no-biopsy approach in children, applying the same approach to adults remained controversial. Thus, current international guidelines still recommend duodenal biopsies to confirm the diagnosis of coeliac disease in adults [7–9]. The aim

of this systematic review and meta-analysis was to evaluate the accuracy of the no-biopsy approach in adult patients with suspected coeliac disease.

Methods

Registration of review protocol

This study was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for diagnostic test accuracy (PRISMA-DTA) guidelines (Supplementary materials) [10], based on a *priori* registered protocol (PROSPERO; CRD42023398812).

Search strategy and study selection

We systematically searched MEDLINE, EMBASE, Cochrane Library and Web of Science for relevant studies from January 1998 to the 2nd of April 2023 to identify studies evaluating the diagnostic performance of IgA-tTG ≥10×ULN compared with duodenal biopsies in adult patients (age ≥16 years) with suspected coeliac disease. We restricted the literature search to start from 1998 following the publication of a landmark study by Dieterich et al., which defined how coeliac disease is diagnosed in children and adults using IgA-tTG [11]. There were no language restrictions. The literature search was repeated on the 3rd of October 2023 with a refined search strategy to ensure that no relevant studies have been missed. Two reviewers (MGS and NN) independently screened the titles and abstracts of all citations against the inclusion criteria. The full-text articles of all potentially relevant studies were retrieved and further evaluated in more detail using standardised forms. We also manually searched the bibliographies of the relevant reviews and included studies for any additional eligible studies. The full search strategy is shown in the supplementary materials.

We included studies that met the following criteria: (1) included adult patients (age ≥16 years) at risk of coeliac disease (2) reported IgA-tTg cut-off levels of ≥10×ULN (3) coeliac disease diagnosed

based on a Marsh ≥2 lesions on duodenal biopsy (4) Published in full-text articles. We excluded studies that included only paediatric patients, conference abstracts, case reports, reviews, editorials, and practice guidelines. Studies with insufficient information to create 2x2 contingency tables for the diagnostic accuracy of IgA-tTG ≥10×ULN were also excluded.

Data extraction

Two reviewers (MGS & NN) extracted the data from eligible studies using a standardized excel spreadsheet. The following data were extracted, where available: study country, study design, study period, inclusion criteria, participants' number and characteristics, the prevalence of coeliac disease, number of true positives, false positives, false negatives, and true negatives. Disagreements between reviewers were resolved by consensus.

Risk of bias and quality assessment

The risk of bias assessment was independently assessed by four reviewers using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS- 2) tool based on the following domains: patient selection, index test, reference standard, and flow and timing[12]. Studies that did not explicitly state whether consecutive or random sampling was made were judged as having a high risk of bias in the patient selection domain of the QUADAS-2 tool. The index test domain of the QUADAS-2 tool was judged as unclear, if the authors did not provide sufficient details of the IgA-tTG assay used. The reference standard domain of the QUADAS-2 tool was judged as having a high risk of bias if the authors did not explicitly state whether duodenal biopsy was interpreted without knowledge of the IgA-tTG results. Finally, the flow and timing domain of the QUADAS-2 tool was judged to have an unclear risk of bias if the authors did not report the exact time interval

between IgA-tTG and duodenal biopsy. The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) criteria was used to assess the quality of evidence[13]. Disagreements between reviewers were resolved by consensus.

Study outcomes

The primary outcome of the meta-analysis was the diagnostic accuracy of IgA-tTG ≥10×ULN in identifying patients with coeliac disease, compared with intestinal biopsy as the reference standard.

Data synthesis and statistical analysis

The prevalence of coeliac disease and proportion of patients with IgA-tTG ≥10×ULN in the included studies were pooled and estimated with 95% CI using a random effects model. We used 2x2 tables to calculate the summary estimates of sensitivity, specificity, positive and negative likelihood ratio of IgA-tTG ≥10×ULN using a bivariate random effects model. Summary estimates of the sensitivity and specificity of the included studies were presented in forest plots. A summary receiver operating characteristic (SROC) curve was constructed and the area under the curve was calculated[14].

The unconditional positive and negative predictive values were assessed based on a uniform prior distribution of coeliac disease. However, the prevalence of coeliac disease varies according the studied population, and it is estimated to be approximately 1% of the general population[15], 4% of patients with irritable bowel syndrome type symptoms[16], and 10% of people with a family history of coeliac disease[17]. Therefore, we used these common pre-test probabilities of coeliac

disease to estimate the post-test probabilities if the test is positive or negative, using Fagan's nomograms[18].

We assessed heterogeneity by visual inspection of the forest plot, bivariate box plot and using Cochran Q χ^2 test and the I^2 statistics [19]. To identify potential outliers and estimate the influence of individual studies, we used Cook's distance (Cook's D). Additionally, we evaluated the risk of publication bias using Deek's funnel plot asymmetry test[20]. A p-value of < 0.05 was considered statistically significant. All statistical analyses were performed with Stata version 17 (StataCorp, College Station, Texas, USA), using the "metaprop", "midas" and "metadta" commands.

Results

Study selection and characteristics

The search strategy identified 17,576 citations from four electronic databases, of which 82 articles appeared to be relevant and eligible for full-text screening (Figure 1). A total of 18 studies comprising 12,103 participants from 15 countries met the criteria for inclusion in the meta-analysis [21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38]. The main characteristics of the included studies are summarized in Table 1.

All the studies were conducted in secondary and tertiary care settings and excluded patients with known coeliac disease or on a gluten-free diet. All studies included adult patients with suspected coeliac disease who underwent serology and duodenal biopsy. The inclusion and exclusion criteria of each study are summarised in the supplementary material. Three studies included repeated measurements of IgA-tTG assays across different commercial kits [22,28,36]. The PPV of IgA-tTG ≥10×ULN to identify patients with coeliac disease was similar across the different IgA-tTG assays in all three studies. In our primary analyses, we included the Celikey IgA assay results as the shared assay between the three studies and reported the results of the other assays in separate sensitivity analyses. Three studies were published as letters [23,31,32], including Sugai et al. which was a post-hoc analysis of an earlier study [23,39]. We have decided to include these studies in our primary analysis and conducted a sensitivity analysis excluding them to estimate their influence on the results. The prevalence of biopsy-proven coeliac disease in the included studies was 62% (95% CI, 40% - 83%) with a high heterogeneity between studies (I^2=99.9%) (Supplementary Figure 1). The proportion of patients with IgA-tTG ≥10×ULN was 32% (95% CI, 24% - 40%) with a high heterogeneity between studies (I^2=99.3%) (Supplementary Figure 2).

Diagnostic performance of the no-biopsy approach

The summary sensitivity of IgA-tTG ≥10×ULN was 51% (95% CI, 42% - 60%), and the summary specificity was 100% (95% CI, 98% - 100%) for the diagnosis of coeliac disease (Figure 2). The positive and negative likelihood ratios were 183.42 (95% CI, 30.1 – 1114.6) and 0.49 (95% CI 0.34 – 0.59), respectively. The diagnostic odds ratio was 373 (95% CI, 60 – 2314). The area under the summary receiver operating characteristic curve was 0.83 (95% CI, 0.77 – 0.89) (Figure 3). The unconditional PPV was 98% (95% CI, 96% - 99%), and the unconditional NPV was 62% (95% CI, 61% - 63%) (Figure 4).

The PPV of the no-biopsy approach to identify patients with coeliac disease was 65%, 88%, 95%, and 99% if coeliac disease prevalence was 1%, 4%, 10% and 40%, respectively (Figure 5). The prevalence of 40% represents the lower confidence interval of the pooled prevalence from the included studies. The PPV and NPV of the no-biopsy approach across different coeliac disease prevalences are shown in Supplementary Figure 3. The diagnostic accuracy results and downstream consequences of testing four hypothetical adult cohorts with different pre-test probabilities of coeliac disease are presented in absolute terms per 1,000 patients tested in supplementary Figure 4[40].

Heterogeneity assessment and sensitivity analyses

Between-study heterogeneity for sensitivity was high (I^2=92.3%), while there was low heterogeneity for specificity (I^2=1.5%). The generalized between-study heterogeneity was moderate (I^2=30.3%). The bivariate box plot showed that most studies clustered within the median distribution and 95% confidence bound of the data points, with only two outliers [27,29] (Figure 6). Further influence analysis using Cook's distance confirmed that both outlier studies

had a significant influence on the results (Supplementary Figure 5). Excluding both outlier studies did not significantly alter the results with a summary sensitivity of 49% (95% CI, 42% – 57%) and a summary specificity of 99% (95% CI, 98% - 100%) (Supplementary Figure 6).

Sub-group analysis of 13 studies reporting Marsh 3 lesions on duodenal biopsies yielded similar diagnostic performance of IgA-tTG ≥10×ULN with a summary sensitivity of 51% (95% CI, 40% - 62%) and summary specificity of 100% (95% CI, 98% - 100%) (Supplementary Figures 7 - 8). Excluding the 3 studies published in letters did not significantly alter the results with a summary sensitivity of 54% (95% CI, 44% - 63%) and a summary specificity of 100% (95% CI, 98% - 100%) (Supplementary Figure 9)[23,31,32]. Furthermore, there were no significant differences in the summary sensitivity and specificity between retrospective and prospective studies (Supplementary Figure 10). The results were also not significantly altered after sensitivity analyses using the different assays in Oyaert et al. [22] (Supplementary Figure 11), Ylönen et al. [28] (Supplementary Figures 12 - 14), Castelijn et al. [36] (Supplementary Figure 15). There was no evidence of Deek's funnel plot asymmetry to suggest publication bias (p=0.05) (Supplementary Figure 16).

Risk of bias and quality assessment

The outcomes of the methodological quality assessment of the included studies using the QUADAS-2 tool are summarized in the supplementary material. There was only one study with a low risk of bias in all domains[25]. However, there were no concerns regarding applicability as all the studies reflected real life clinical practice. The overall certainty of evidence was downgraded to moderate due to serious risks of bias (Supplementary Table 1).

Discussion

This is the first systematic review and meta-analysis to evaluate the accuracy of the no-biopsy approach for the diagnosis of coeliac disease in adults. A total of 18 studies with 12,103 participants from 15 countries were included in this meta-analysis. Summary data showed that IgA-tTG ≥10×ULN has an overall sensitivity of 51% (95% CI, 42% - 60%) and an overall specificity of 100% (95% CI, 98% - 100%) for detecting coeliac disease. The PPV of IgA-tTG ≥10×ULN to identify patients with coeliac disease was 98% (95% CI, 96% - 99%). However, this high predictive value varied according to the pre-test probability of coeliac disease in the studied population. We provided PPV estimates of IgA-tTG ≥10×ULN for common pre-test probabilities of coeliac disease to aid clinicians and patients in reaching an informed decision on a no-biopsy diagnosis based on the best available evidence.

The results of this study demonstrate that the no-biopsy approach, that has been incorporated in paediatric practice to diagnose coeliac disease for over a decade, can be safely extrapolated to selected adult patients in secondary care settings. This has significant implications for clinical practice by reducing the diagnostic delays, risks and healthcare costs associated with endoscopy. In a recent study, we estimated that the cost of diagnosis in adults could be reduced by over 75% if endoscopy and biopsy were avoided[41].

Despite the consistent evidence supporting the no-biopsy approach in diagnosing adult patients with coeliac disease, there have been some concerns regarding its applicability. One potential concern with the relying on serology testing alone is the possibility of false-positive diagnosis of coeliac disease[9]. This could lead to unnecessary dietary restriction and negative effects on patients' quality of life. Although our results did not show that the PPV of IgA-tTG

≥10×ULN to identify patients with coeliac disease was 100%, it is important to note that no diagnostic test for coeliac disease is 100% accurate even duodenal biopsy which is considered the gold standard. Studies have shown that adherence to recommended biopsy guidelines occurs in only 40% of cases [42,43], indicating that the diagnosis could be missed despite duodenal sampling. Furthermore, interpretation of histopathological changes can be subjective and substantial interobserver variability exist between different pathologists [6]. Therefore, the results interpreted as "false-positive" serology could have been false-negative histology [27]. The no-biopsy diagnosis of coeliac disease in patients with IgA ≥10×ULN could mitigate the risk of potential false-negative histology results. This is particularly relevant in cases where the histopathological findings are not diagnostic for coeliac disease due to inadequate sampling.

The lack of standardization of IgA-tTG assays across different laboratories is another concern [44]. However, studies directly comparing different IgA-tTG assays showed that a cut-off level ≥10×ULN had a consistent PPV for coeliac disease close to 100% [28,36]. This is in line with our results showing high diagnostic performance of IgA-tTG ≥10×ULN across different commercial kits, laboratories, and countries. Yet, local validation of this pathway is recommended to ensure the accuracy and applicability of the no-biopsy approach. Concerns have also been raised regarding the possibility of missing concurrent pathology in patients avoiding endoscopy and biopsy. While recent evidence suggests that patients with coeliac disease, including older patients, had no significant co-pathology that would have been missed if they avoided endoscopy [38,45,46], the decision to avoid endoscopy should be made on a case-by-case basis. Factors such as the patient's age, co-morbidities, risk factors and preferences should all be considered when making the decision of a no-biopsy diagnosis.

A crucial aspect of the successful implementation of the no-biopsy approach is that it should not be interpreted as a "no-referral" approach. Despite current guidelines mandating referral for biopsy in all patients with positive coeliac serology, reports from the UK, Israel and the USA showed that almost a third of patients were never referred from primary care [34,45,47,48].Therefore, a close collaboration and dialogue between primary and secondary care is necessary to implement the no-biopsy approach safely, and to promote adherence to the serology-biopsy guidelines. This would avoid over-diagnosis of coeliac disease in primary care which could have detrimental effects on patients' quality of life [49]. Importantly, it should be stressed that endoscopy would still be required for patients with <10 fold elevation of IgA-tTG, patients with red flag signs or symptoms and for those who wish to have a confirmatory biopsy before adhering to a lifelong gluten-free diet [50]. The development of clear clinical guidelines, educational initiatives and local diagnostic pathways would ensure that clinicians are well-informed and capable of appropriately assessing the pre-test probability of coeliac disease in different clinical settings.

The current European paediatric guidelines recommend that children with IgA-tTG ≥10×ULN require a positive EMA test in a separate blood sample before confirming the diagnosis of coeliac disease[5]. The same approach has been adopted in the Finnish guidelines for the diagnosis of adult coeliac disease as well as in the interim guidance issued by the British society of Gastroenterology during the COVID-19 pandemic [51,52]. However, our results suggest the possibility of re-evaluating the necessity of confirmatory EMA testing as IgA-tTG ≥10×ULN alone has an excellent predictive power for coeliac disease. EMA testing requires indirect immunofluorescence, which is costly, labour intensive and subject to inter-observer variability.

Consequently, many clinical laboratories have stopped performing EMA tests and their availability has progressively decreased over time [22]. Therefore, including EMA testing in the no-biopsy diagnostic pathway may hinder its implementation without having a clear added value.

This meta-analysis has important strengths. First, we conducted a comprehensive systematic literature search following a *priori* registered protocol and a pre-defined inclusion and exclusion criteria. Second, we performed extensive sensitivity analyses to explore causes of heterogeneity and to assess the robustness of our results. Third, we used the validated QUADS-2 tool to assess the risk of bias and applicability concerns in the included studies. Fourth, all the included studies used serology as the index test and Marsh ≥2 on duodenal biopsies as the reference standard. Restricting the analysis to only those evaluating the predictive value of IgA-tTG ≥10×ULN for Marsh 3 lesions did not alter the results, adding to the validity of our findings.

Our study also has some limitations that should be considered when interpreting the results. All the included studies were performed in secondary and tertiary care settings with a pooled prevalence of coeliac disease of 62%, which is higher than expected in clinical practice. The sensitivity and specificity of any diagnostic test can be influenced by the prevalence of disease in the studied population due to many clinical mechanisms (distorted patient spectrum, referral filter, or reader expectation) or artefactual mechanisms (distorted inclusion of participants including limited-challenge phenomenon, verification bias, or reference standard misclassification). We have adjusted for this by using the likelihood ratios to calculate the predictive values of the no-biopsy approach across different pre-test probabilities of coeliac disease. The results showed that the no-biopsy approach may have a limited utility in primary care, where the pre-test probability of coeliac disease is lower than 10%. Another limitation is

the retrospective nature of most of the included studies. However, the high predictive value of the no-biopsy approach remained consistent in prospective studies and across different countries and commercial IgA-tTG assays. Finally, only one study had a low risk of bias across all domains as most studies had selection bias, unclear time intervals between serology and histology which may have introduced misclassification bias, and lack of blinding to serology results which may have influenced the pathologists' interpretation of the histological findings. To avoid these potential sources of bias, future prospective studies should adhere to pre-defined protocols and report results according to the Standards for Reporting of Diagnostic Accuracy Studies [53].

Future research should focus on evaluating the diagnostic accuracy of the no-biopsy approach in primary care settings and in patients with a low-pretest probability of coeliac disease. It will also be important to assess the accuracy of lower thresholds of IgA-tTG to predict villous atrophy, and the added value of confirmatory testing with EMA. Additionally, given that most studies were conducted in Europe, further studies are needed to determine the generalizability of our findings to regions and countries with limited data on the accuracy of the no-biopsy approach, such as the USA. Importantly, studies exploring patients' preferences, the cost-effectiveness, and the regulatory aspects of implementing the no-biopsy approach are needed to determine its acceptability, feasibility and impact in clinical practice.

In conclusion, our meta-analysis of 18 studies including more than 12,000 participants provides evidence that IgA-tTG levels ≥10×ULN are highly indicative of coeliac disease in adult patients referred to secondary care. Close collaboration between primary and secondary care,

and shared decision making between clinicians and patients will be critical in implementing this no-biopsy approach.

References

1. Lebwohl B, Sanders DS, Green PHR. Coeliac disease. Lancet 2018;391:70–81. Available at: http://dx.doi.org/10.1016/S0140-6736(17)31796-8.

2. Makharia GK, Chauhan A, Singh P, et al. Review article: Epidemiology of coeliac disease. Aliment Pharmacol Ther 2022;56:S3–S17.

3. Kaukinen K. Updates on systemic consequences of coeliac disease. Nat Rev Gastroenterol Hepatol 2021;18:87–88.

4. Raiteri A, Granito A, Giamperoli A, et al. Current guidelines for the management of celiac disease: A systematic review with comparative analysis. World J Gastroenterol 2022;28:154–175.

5. Husby S, Koletzko S, Korponay-Szabó I, et al. European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. J Pediatr Gastroenterol Nutr 2020;70:141–156.

6. Werkstetter KJ, Korponay-Szabó IR, Popp A, et al. Accuracy in Diagnosis of Celiac Disease Without Biopsies in Clinical Practice. Gastroenterology 2017;153:924–935.

7. Ludvigsson JF, Bai JC, Biagi F, et al. Diagnosis and management of adult coeliac disease: Guidelines from the British society of gastroenterology. Gut 2014;63:1210–1228.

8. 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 Eur

Gastroenterol J 2019;7:583–613.

9. Rubio-Tapia A, Hill ID, Semrad C, et al. American College of Gastroenterology Guidelines Update: Diagnosis and Management of Celiac Disease. Am J Gastroenterol 2023;118:59–76. Available at: https://journals.lww.com/10.14309/ajg.0000000000002075.

10. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 2021;372.

11. Dieterich W, Laag E, Schöpper H, et al. Autoantibodies to tissue transglutaminase as predictors of celiac disease. Gastroenterology 1998;115:1317–1321. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0016508598700071.

12. Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med 2011;155:529. Available at: http://annals.org/article.aspx?doi=10.7326/0003-4819-155-8-201110180-00009.

13. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. Chinese J Evidence-Based Med 2009;9:8–11.

14. Takwoingi Y, Dendukuri N, Schiller I, et al. Chapter 11: undertaking meta-analysis. Cochrane Handb Syst Rev diagnostic test accuracy version 2022;2.

15. Singh P, Arora A, Strand TA, et al. Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2018;16:823-836.e2.

16. Irvine AJ, Chey WD, Ford AC. Screening for Celiac Disease in Irritable Bowel Syndrome: An

Updated Systematic Review and Meta-analysis. Am J Gastroenterol 2017;112:65–76.

17. Singh P, Arora S, Lal S, et al. Risk of celiac disease in the first- and second-degree relatives of patients with celiac disease: A systematic review and meta-analysis. Am J Gastroenterol 2015;110:1539–1548. Available at: http://dx.doi.org/10.1038/ajg.2015.296.

18. Letter FTJ. Nomogram for Bayes theorem. N Engl J Med 1975;293:257.

19. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–1558. Available at: https://onlinelibrary.wiley.com/doi/10.1002/sim.1186.

20. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005;58:882–893. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0895435605001423.

21. Hill PG, Holmes GKT. Coeliac disease: A biopsy is not always necessary for diagnosis. Aliment Pharmacol Ther 2008;27:572–577.

22. Oyaert M, Vermeersch P, Hertogh G De, et al. Combining antibody tests and taking into account antibody levels improves serologic diagnosis of celiac disease. Clin Chem Lab Med 2015;53:1537–1546.

23. Sugai E, Hwang HJ, Vázquez H, et al. Should ESPGHAN guidelines for serologic diagnosis of celiac disease be used in adults? A prospective analysis in an adult patient cohort with

high pretest probability. Am J Gastroenterol 2015;110:1504–1505.

24. Tola M Di, Marino M, Goetze S, et al. Identification of a serum transglutaminase threshold value for the noninvasive diagnosis of symptomatic adult celiac disease patients: a retrospective study. J Gastroenterol 2016;51:1031–1039.

25. Previtali G, Licini L, D'Antiga L, et al. Celiac Disease Diagnosis Without Biopsy: Is a 10× ULN Antitransglutaminase Result Suitable for a Chemiluminescence Method? J Pediatr Gastroenterol Nutr 2018;66:645–650.

26. Gülseren YD, Adiloğlu AK, Yücel M, et al. Comparison of non-invasive tests with invasive tests in the diagnosis of celiac disease. J Clin Lab Anal 2019;33:1–7.

27. Fuchs V, Kurppa K, Huhtala H, et al. Serology-based criteria for adult coeliac disease have excellent accuracy across the range of pre-test probabilities. Aliment Pharmacol Ther 2019;49:277–284.

28. Ylönen V, Lindfors K, Repo M, et al. Non-Biopsy Serology-Based Diagnosis of Celiac Disease in Adults Is Accurate with Different Commercial Kits and Pre-Test Probabilities. Nutrients 2020;12:2736. Available at: https://www.mdpi.com/2072-6643/12/9/2736.

29. Sinha SK, Berry N, Muktesh G, et al. Utility of narrow band imaging in predicting histology in celiac disease. Indian J Gastroenterol 2020;39:370–376.

30. Penny HA, Raju SA, Lau MS, et al. Accuracy of a no-biopsy approach for the diagnosis of coeliac disease across different adult cohorts. Gut 2021;70:876–883. Available at:

https://gut.bmj.com/lookup/doi/10.1136/gutjnl-2020-320913.

31. Paul SP, Lau WYS, Khan ZH, et al. Letter: no-biopsy pathway for diagnosing adult coeliac disease. Aliment Pharmacol Ther 2021;53:357–358.

32. Tashtoush LB, Bosanko NC, Broad SR, et al. Letter: the BSG COVID-19 interim coeliac disease guidance no-biopsy approach is safe in adults. Aliment Pharmacol Ther 2021;54:1090–1092.

33. Baykan AR, Cerrah S, Ciftel S, et al. A No-Biopsy Approach for the Diagnosis of Celiac Disease in Adults: Can It Be Real? Cureus 2022;14:3–8. Available at: https://www.cureus.com/articles/103368-a-no-biopsy-approach-for-the-diagnosis-of-celiac-disease-in-adults-can-it-be-real.

34. Johnston RD, Chan YJ, Mubashar T, et al. No-biopsy pathway following the interim BSG guidance reliably diagnoses adult coeliac disease. Frontline Gastroenterol 2022;13:73–76.

35. Beig J, Rostami K, Hayman DTS, et al. Is duodenal biopsy always necessary for the diagnosis of coeliac disease in adult patients with high anti-tissue transglutaminase (TTG) antibody titres? Frontline Gastroenterol 2022;13:287–294.

36. Castelijn DAR, Mulder AHL, Pol P Van Der, et al. Multicenter study to compare the diagnostic performance of CLIA vs. FEIA transglutaminase IgA assays for the diagnosis of celiac disease. Clin Chem Lab Med 2023.

37. Deane C, O'Connor E, O'Donovan H, et al. The strategic use of biopsy in the diagnosis of coeliac disease in adults. Dig Liver Dis 2023. Available at: https://linkinghub.elsevier.com/retrieve/pii/S1590865823007739.

38. Ciacci C, Bai JC, Holmes G, et al. Serum anti-tissue transglutaminase IgA and prediction of duodenal villous atrophy in adults with suspected coeliac disease without IgA deficiency (Bi . A . CeD): a multicentre , prospective cohort study. Lancet Gastroenterol Hepatol 2023;1253:1–10. Available at: http://dx.doi.org/10.1016/S2468-1253(23)00205-4.

39. Sugai E, Moreno ML, Hwang HJ, et al. Celiac disease serology in patients with different pretest probabilities: Is biopsy avoidable? World J Gastroenterol 2010;16:3144–3152.

40. Whiting P, Leeflang M, Salis I de, et al. Guidance was developed on how to write a plain language summary for diagnostic test accuracy reviews. J Clin Epidemiol 2018;103:112–119. Available at: https://doi.org/10.1016/j.jclinepi.2018.07.008.

41. Shiha MG, Nandi N, Hutchinson AJ, et al. Cost-benefits and environmental impact of the no-biopsy approach for the diagnosis of coeliac disease in adults. Frontline Gastroenterol 2023:flgastro-2023-102494. Available at: https://fg.bmj.com/lookup/doi/10.1136/flgastro-2023-102494.

42. Lebwohl B, Kapel RC, Neugut AI, et al. Adherence to biopsy guidelines increases celiac disease diagnosis. Gastrointest Endosc 2011;74:103–109. Available at: http://dx.doi.org/10.1016/j.gie.2011.03.1236.

43. Taylor MA, Blanshard RJ, Naylor G, et al. Do gastroenterologists have medical inertia

towards coeliac disease? A UK multicentre secondary care study. BMJ Open Gastroenterol 2021;8.

44. Beltran L, Koenig M, Egner W, et al. High-titre circulating tissue transglutaminase-2 antibodies predict small bowel villous atrophy, but decision cut-off limits must be locally validated. Clin Exp Immunol 2014;176:190–198.

45. Hoyle A, Gillett P, Gillett HR, et al. No-biopsy strategy for coeliac disease is applicable in adult patients: a 'real-world' Scottish experience. Frontline Gastroenterol 2022:flgastro-2022-102254. Available at: https://fg.bmj.com/lookup/doi/10.1136/flgastro-2022-102254.

46. Stefanolo JP, Zingone F, Gizzi C, et al. Upper gastrointestinal endoscopic findings in celiac disease at diagnosis: A multicenter international retrospective study. World J Gastroenterol 2022;28:6157–6167. Available at: https://www.wjgnet.com/1007-9327/full/v28/i43/6157.htm.

47. Guz-Mark A, Feldman BS, Ghilai A, et al. High rates of serology testing for coeliac disease, and low rates of endoscopy in serologically positive children and adults in Israel: lessons from a large real-world database. Eur J Gastroenterol Hepatol 2020;32:329–334. Available at: https://journals.lww.com/10.1097/MEG.0000000000001613.

48. Joelson AM, Geller MG, Zylberberg HM, et al. Numbers and Features of Patients With a Diagnosis of Celiac Disease Without Duodenal Biopsy, Based on a National Survey. Clin Gastroenterol Hepatol 2019;17:1089-1097.e2. Available at:

https://doi.org/10.1016/j.cgh.2018.09.006.

49. Biagi F, Bianchi PI, Campanella J, et al. The impact of misdiagnosing celiac disease at a referral centre. Can J Gastroenterol 2009;23:543–545.

50. Shiha MG, Penny HA, Sanders DS. Is There a Need to Undertake Conventional Gastroscopy and Biopsy When Making the Diagnosis of Coeliac Disease in Adults? J Clin Gastroenterol 2023;57:139–142. Available at: https://journals.lww.com/10.1097/MCG.0000000000001806.

51. Working group set up by the Finnish Medical Society Duodecim and the Finnish Gastroenterology Society. Celiac disease. Current care guidelines. 2018. Available at: https://www.kaypahoito.fi/hoi08001 [Accessed September 12, 2023].

52. Penny HA, SD GH, Gillett P. BSG Interim Guidance: COVID-19 spcific non-biopsy protocol those with suspected coeliac disease. 2020. Available at: https://www.bsg.org.uk/covid-19-advice/covid-19-specific-non-biopsy-protocol-guidance-for-those-with-suspected-coeliac-disease/ [Accessed September 12, 2023].

53. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. BMJ Open 2016;6:1–17.

Tables

Table 1 – Study characteristics

Author, year (Ref)	Country	Study design	Total participants	Patients with coeliac disease	IgA-tTG assay
Hill et al., 2008[21]	UK	Single centre, retrospective	146	139	Celikey® ELISA (Phadia, Freiburg, Germany)
Oyaert et al., 2015[22]	Belgium	Single centre, prospective	662	90	EliA Celikey ® IgA (Thermo Fisher, Uppsala, Sweden) & QUANTA Flash® (Inova Diagnostics, San Diego, USA)
Sugai et al., 2015[23]	Argentina	Dual-centre, Prospective	161	63	QUANTA Lite™ (Inova Diagnostic, San Diego, CA, USA)
Di Tola et al., 2016[24]	Italy	Single centre, retrospective	671	633	QUANTA Lite™ (Inova Diagnostic, San Diego, CA, USA
Previtali et al., 2018[25]	Italy	Single centre, retrospective	549	199	QUANTA Flash® (Inova Diagnostics, San Diego, USA)
Gülseren et al., 2019[26]	Turkey	Single centre, prospective	21	39	SIEMENS BNProSpec device and Siemens serum IgA kit (Siemens, Munich, Germany)

Fuchs et al. 2019[27]	Finland	Multicentre, retrospective	5500	274	Celikey® ELISA & QUANTA Flash®
Ylönen et al. 2020[28]	Finland	Multicentre, retrospective	836	207	Multiple assays*
Sinha et al., 2020[29]	India	Single centre, prospective	122	112	Celikey IgA Immunoassay (Thermo Fischer, Waltham, MA, USA)
Penny et al., 2021[30]	International[ǂ]	Multicentre, prospective and retrospective cohorts	1417	861	Multiple assays**
Paul et al., 2021[31]	UK	Single centre, retrospective	101	89	Not specified
Tashtoush et al. 2021[32]	UK	Single centre, retrospective	479	388	Not specified
Baykan et al., 2022[33]	Turkey	Single centre, retrospective	269	77	ELISA kit (Orgentec, Mainz, Germany) and an Alisei QS (SEAC Group, Italy)
Johnston et al., 2022[34]	UK	Single centre, retrospective	265	213	Orgentec IgA TTG ELISA (Orgentec Diagnostika, Mainz, Germany) and QUANTA Flash tTG IgA assay (Inova-Werfen, San Diego, USA)
Beig et al., 2022[35]	New Zealand	Single centre, retrospective	144	77	BioRad Autoimmune EIA

					Anti- TTG IgA immunoassay
Castelijn et al., 2023[36]	Netherland	Multicentre, retrospective	89	85	Multiple assays***
Deane et al., 2023[37]	Ireland	Single centre, retrospective	217	184	EliA Celikey IgA assay (Thermo Scientific, Uppsala, Sweden)
Ciacci et al., 2023[38]	International[ǂ]	Multicentre, prospective	636	359	Multiple assays****

[ǂ]UK, USA, Argentina, Iran, Netherland, Italy, Romania, Turkey

[♯]Italy, UK, Spain, Netherland, Romania, Israel, New Zealand, Argentina, India

*Celikey (Phadia, Freiburg, Germany), Orgentec (ORG 540A, Orgentec Diagnostika, Mainz, Germany), Inova (QUANTA Lite h-tTG, Inova Diagnostics, San Diego, CA, USA), and Eurospital (Eu-tTG, Trieste, Italy)

**ARUP Laboratories (Utah, USA), QuantaLite (Inova Diagnostics, San Diego, California), Eu- tTG (Eurospital, Italy), Euroimmune (Luebeck, Germany) and Celikey ELISA (Thermo Fisher, Freiburg, Germany)

***EliA™ Celikey® IgA FEIA (Phadia AB, Thermo Fisher Scientific, Uppsala, Sweden), QUANTA Flash® h-tTG IgA CLIA (Werfen/Inova Diagnostics) and Anti-Tissue Transglutaminase ChLIA (IgA) (EUROIMMUN Medizinische Labordiagnostika AG, Lübeck, Germany)

****QUANTAFLASH, QUANTA Lite ELISA R-tTG IgA, QUANTA Lite ELISA h-tTG IgA, BioPlex 2200 system, Phadia, Diamicron, Multiplex CytoBead CeliAK, Eurospital, IDS Diagnostica, AESKULISA tTg-A, Orgentec Diagnostika, IDS iSYS laboratories, Diasorin Liaison XL, Invitrogen

Figure legends

Figure 1 – PRISMA flow diagram of study selection

Figure 2 – Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease

Figure 3 – A summary receiver operating characteristic plot of the study estimates of IgA-tTG ≥10×ULN sensitivity and specificity

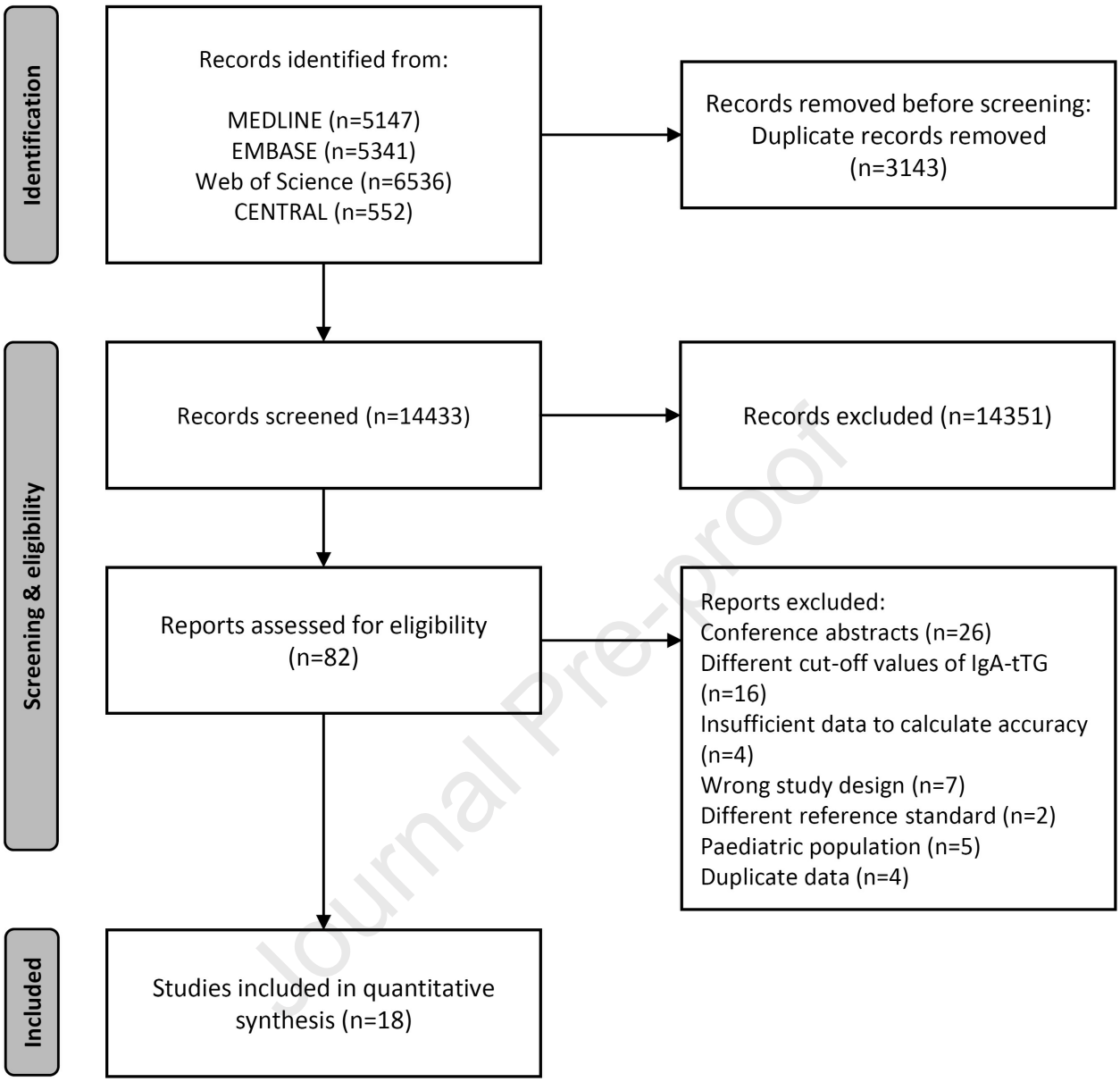
Figure 4 – Probability modifying plot showing the unconditional positive and negative predictive values of IgA-tTG ≥10×ULN to identify patients with coeliac disease

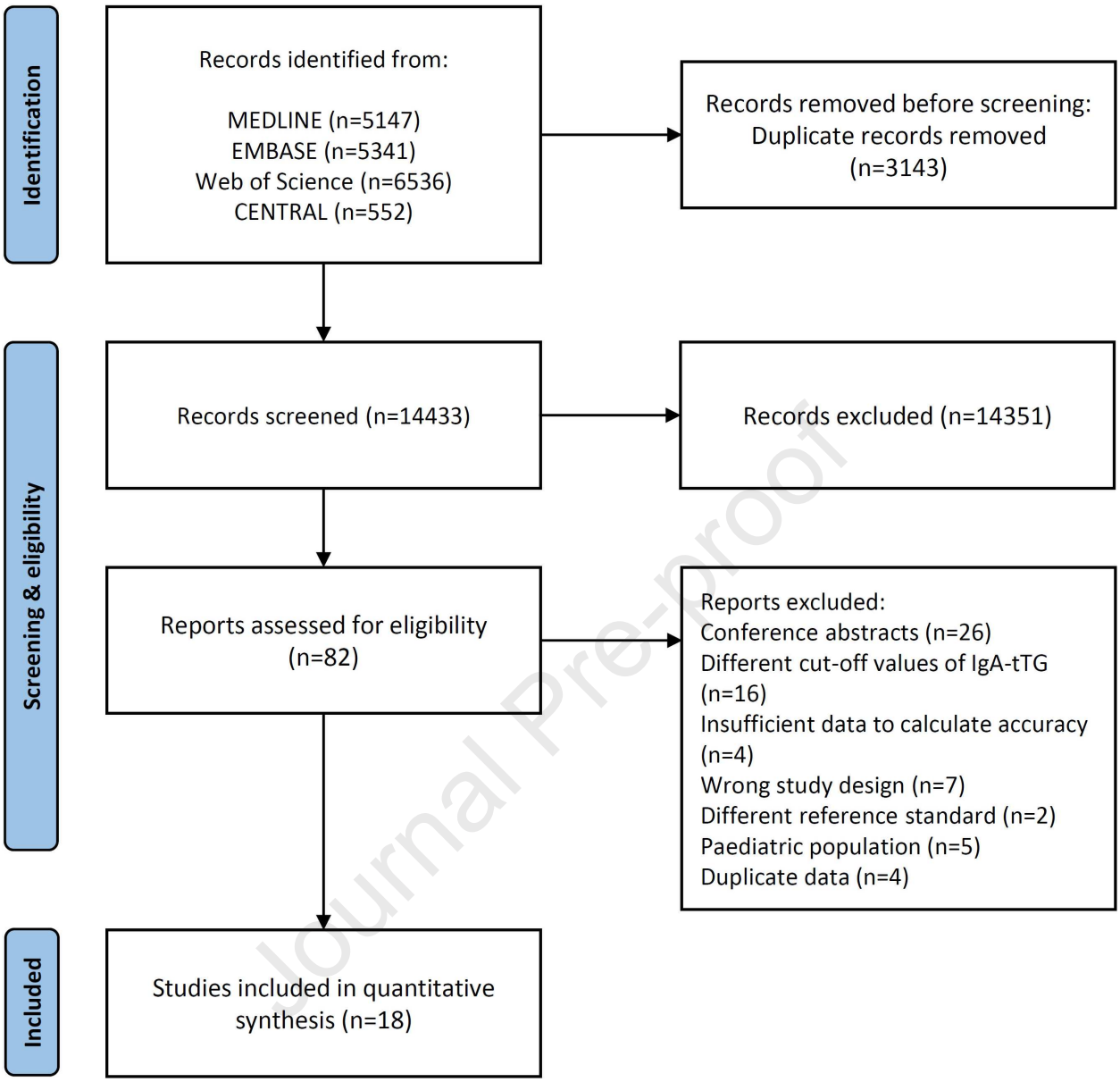
Figure 5 – Fagan's nomograms showing the positive (solid line) and negative (dash line) predictive values of IgA-tTG ≥10×ULN if the pre-test probability of coeliac disease is 1% (A), 4% (B), 10% (C), and 40% (D)

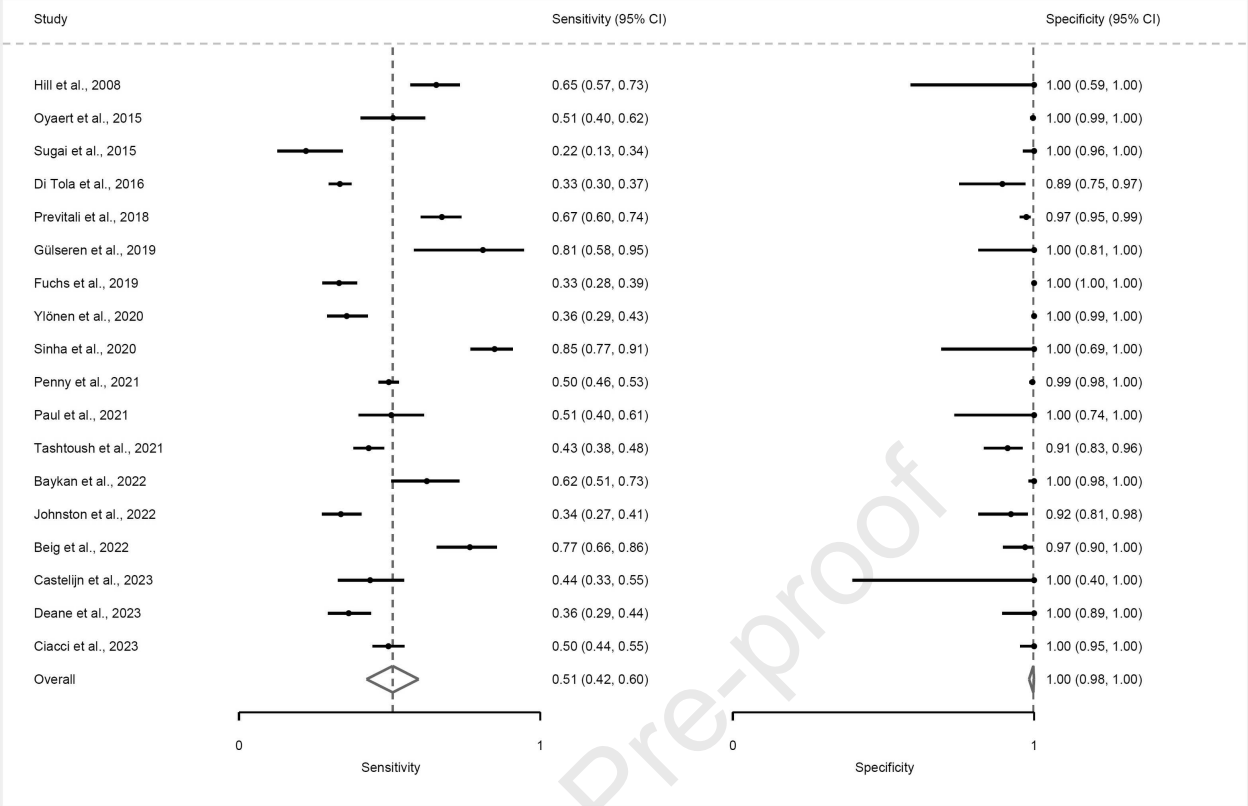
Figure 6 – A bivariate box plot of a random effects modeling of the sensitivity and specificity, with the inner oval representing the median distribution of the data points and the outer oval representing the 95% confidence bound

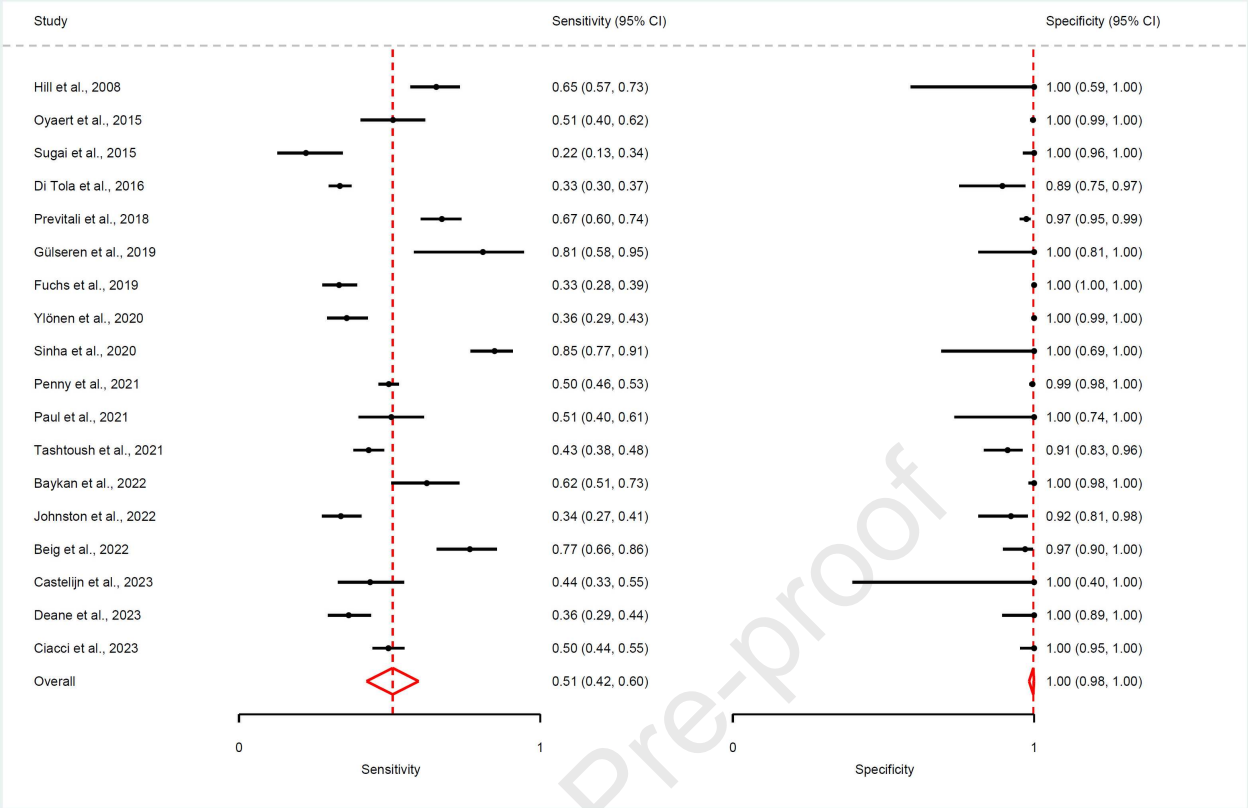
Acknowledgements

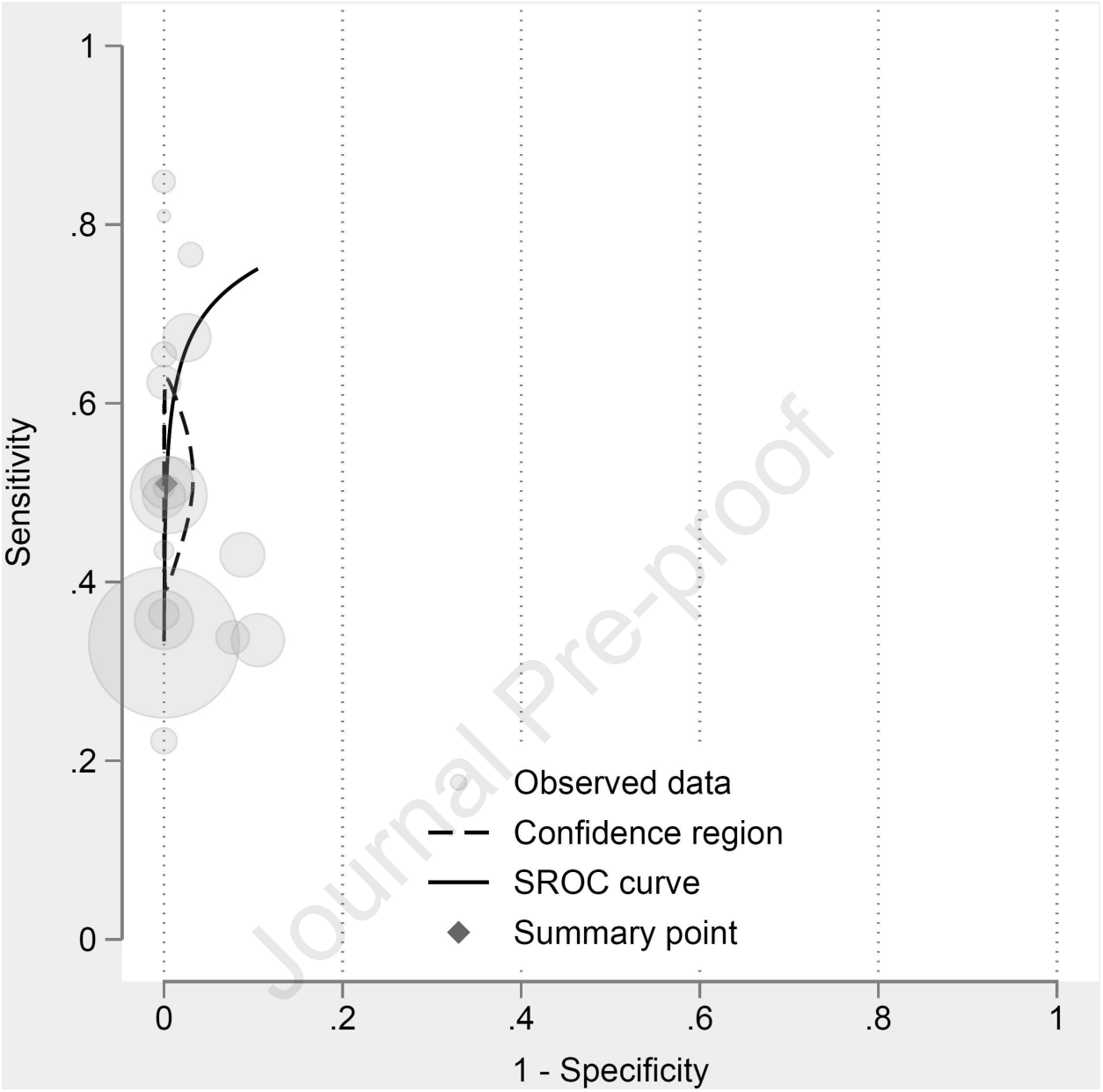
We are grateful to Dr Xavier Bossuyt, Dr Junaid Beig and Dr Daan A.R. Castelijn for providing data from their studies, and to Mr Daniel Froste from Sheffield Teaching Hospitals Library for his advice on the updated search strategy.

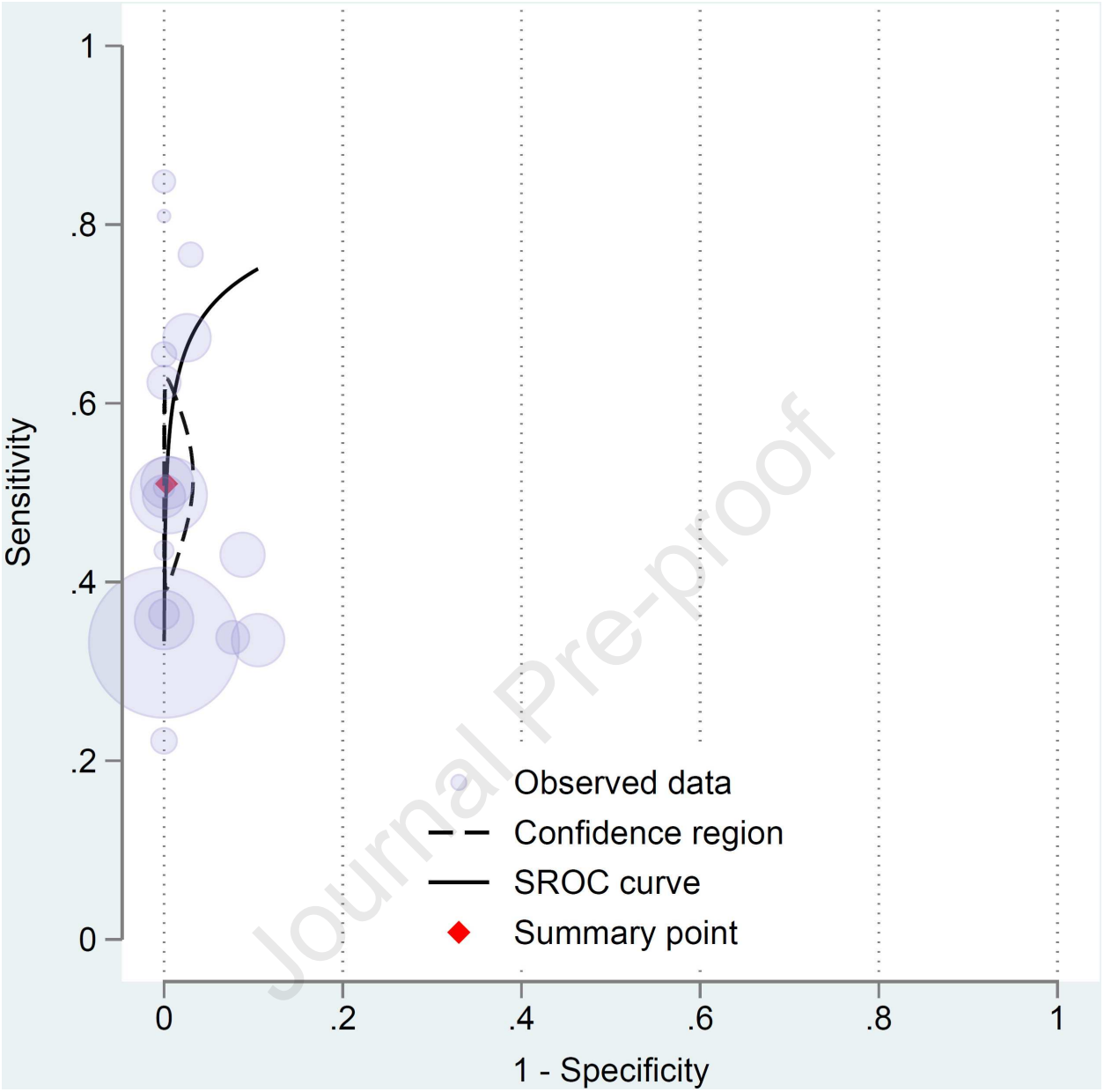


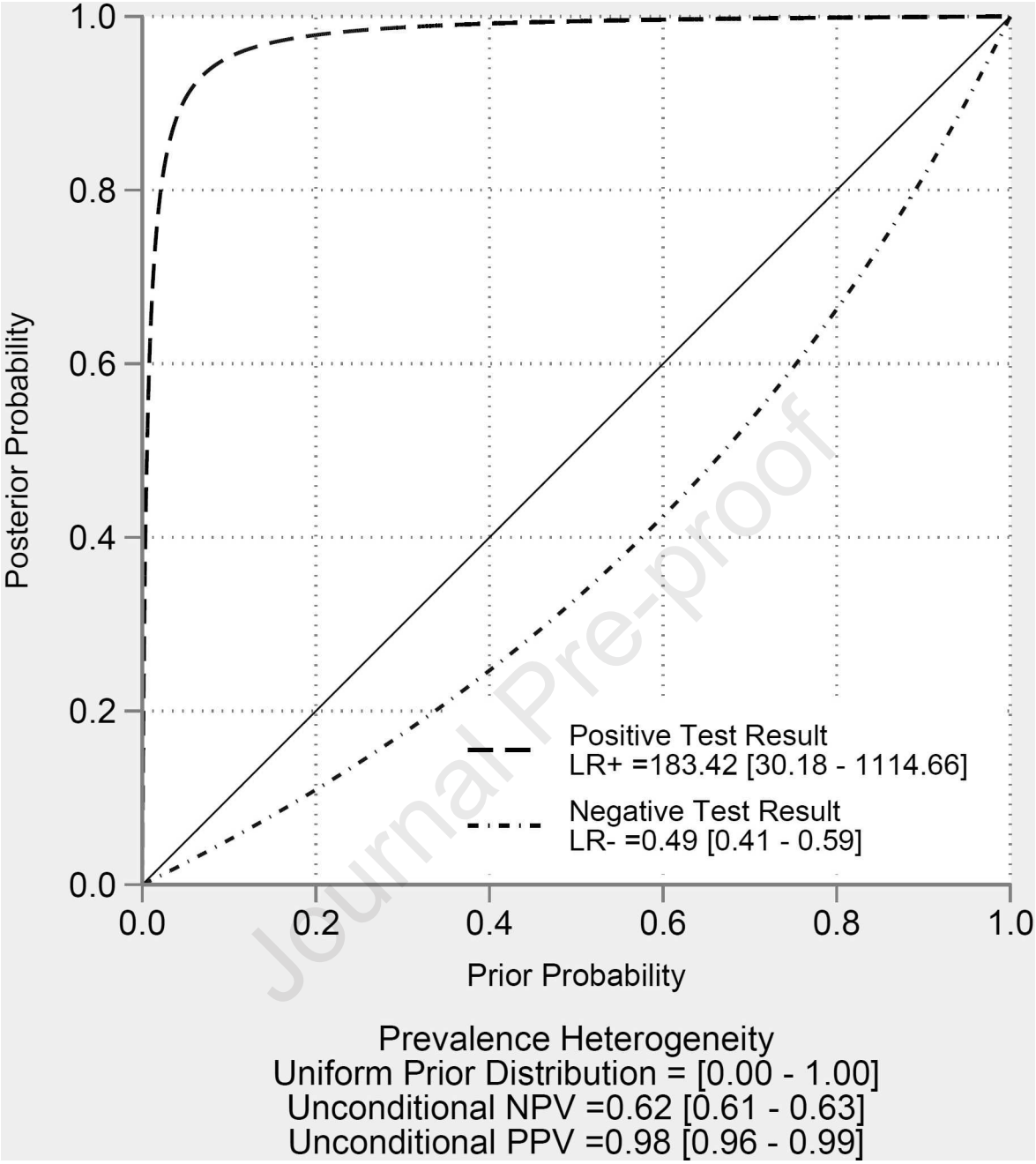


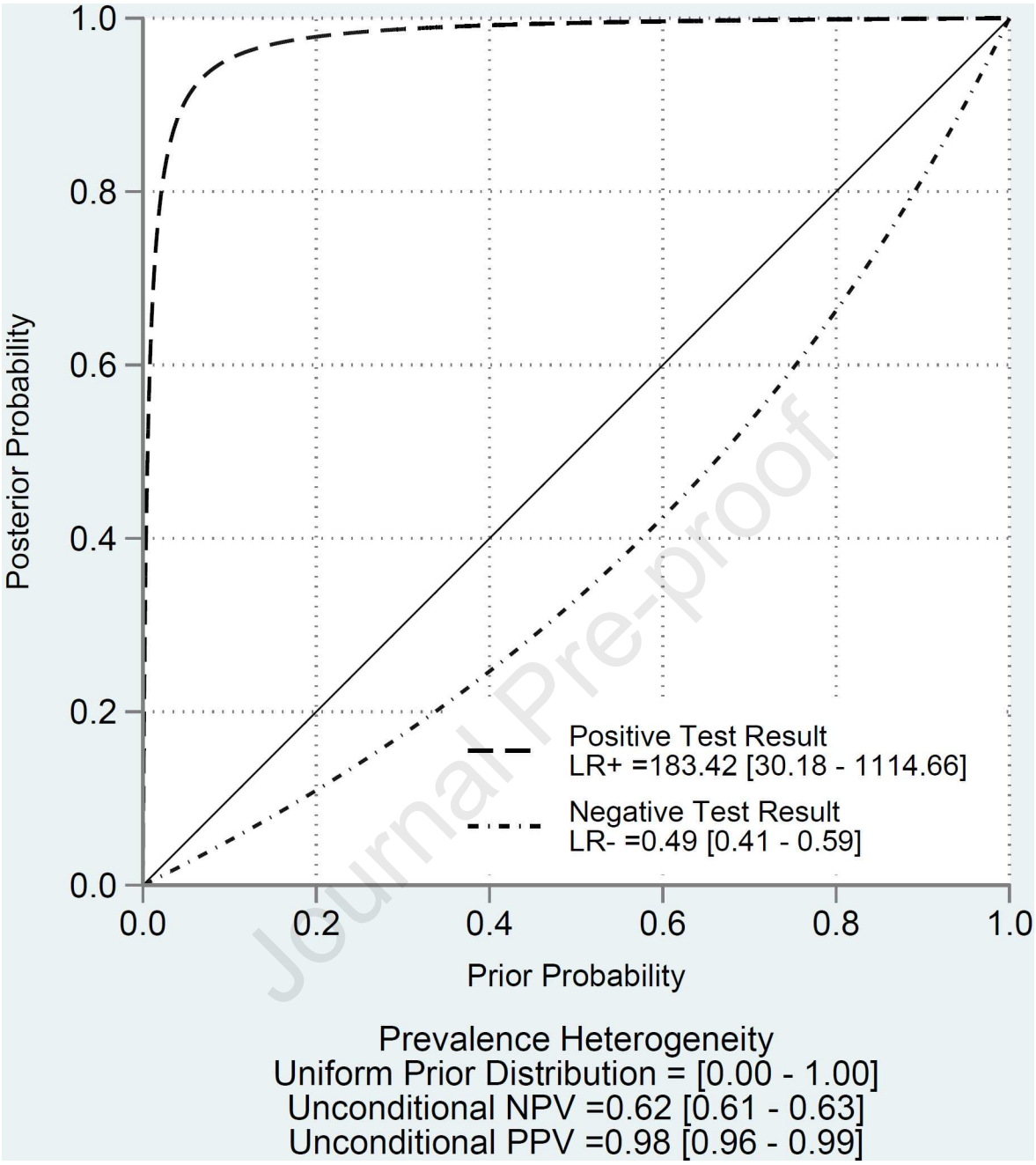


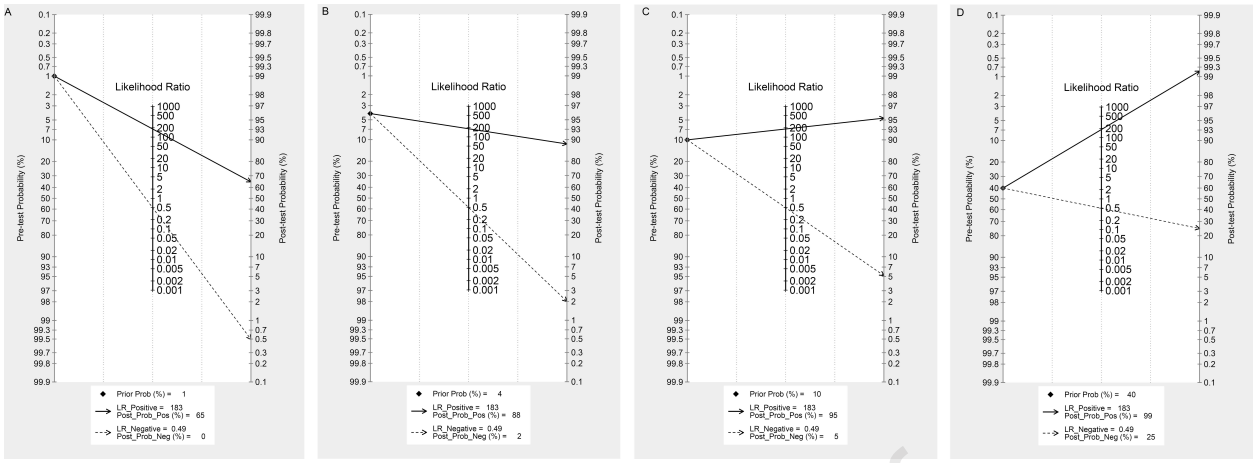


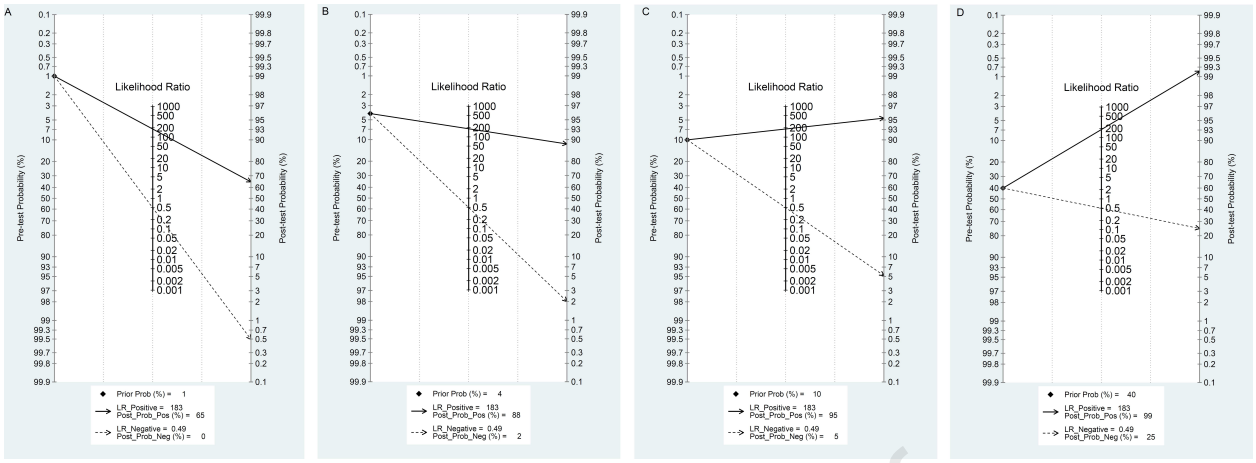


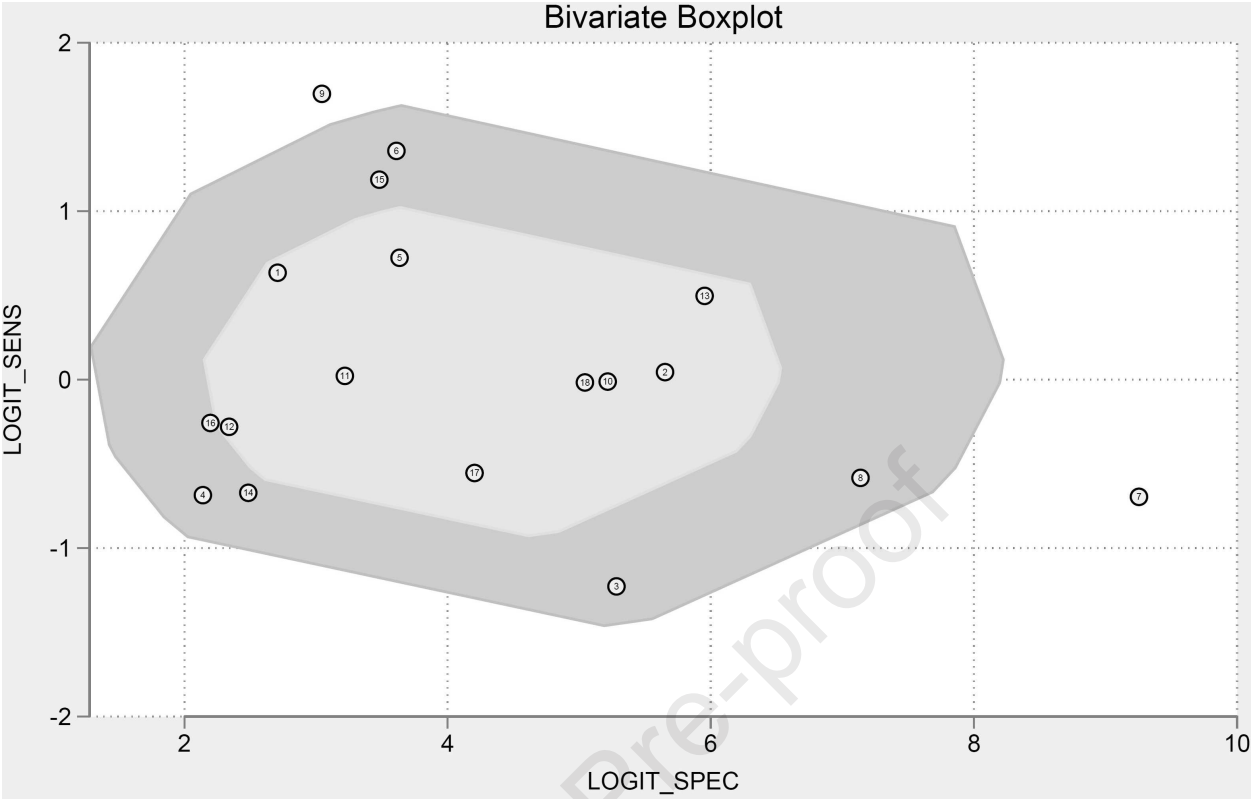


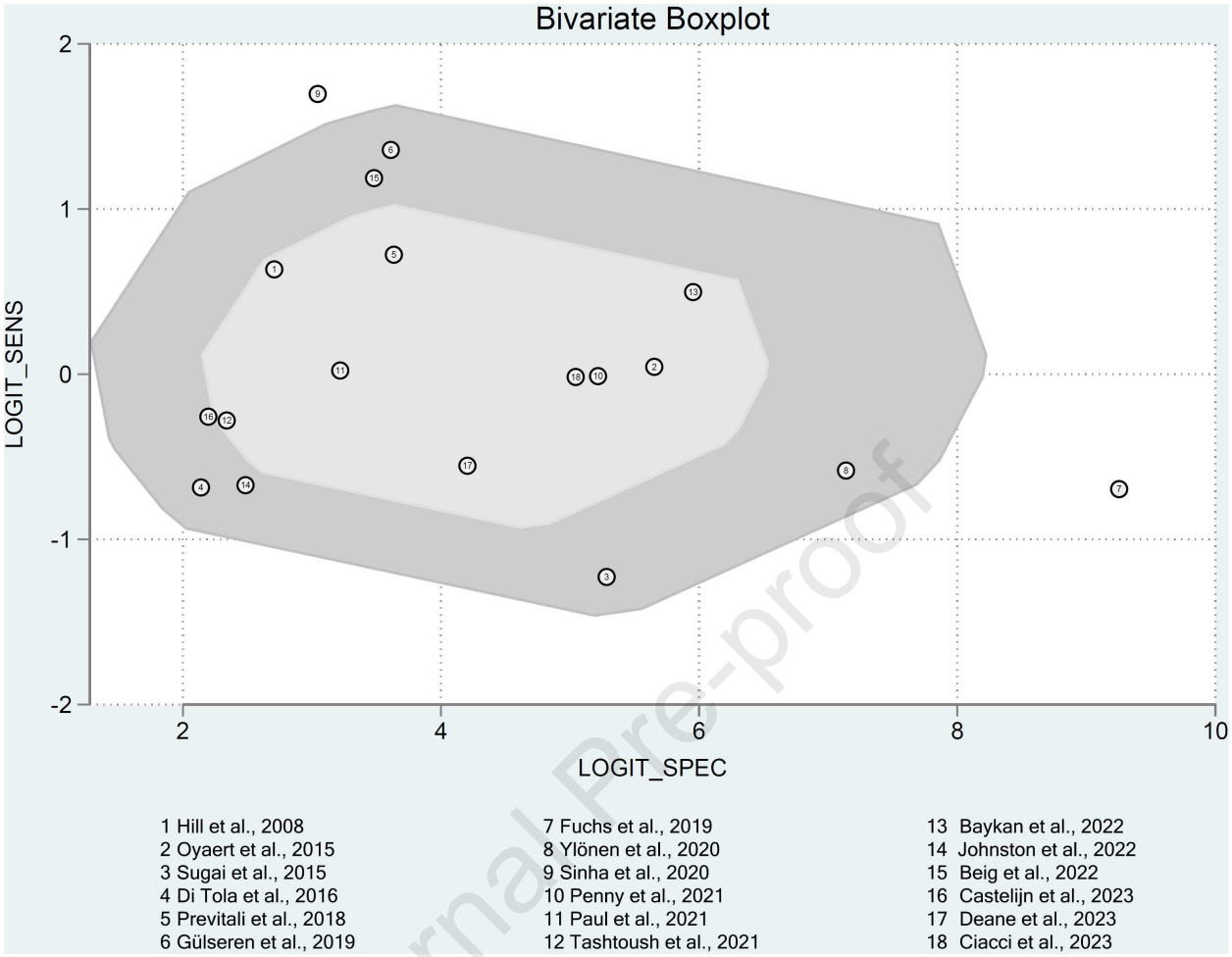












Supplementary Material

PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	3 - 4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	7
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	7
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	7 - 8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	9
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	9
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	S8 – S10
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	9 - 10
Data collection	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and	10

process		any processes for obtaining and confirming data from investigators.	
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	10
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	10 - 11
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	11 - 12
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	11 - 12

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	11 – 12
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	11 - 12
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	13
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	13 & S30 – S79
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	15 & S27 – S79

Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	14 & 32
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	14 & 32
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	14 - 15
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	16
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	19 - 20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	20
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	1

Adapted From: McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.

Search strategy

MEDLINE

1	Celiac Disease/	21983
2	((coeliac or celiac) adj4 disease).tw.	20539
3	Glutens/	7419
4	(gluten adj3 (intoleran* or sensitiv* hypersensi*)).tw,kw.	632
5	1 or 2 or 3 or 4	30597
6	Serologic Tests/	21742
7	serolog*.tw.	133586
8	Autoantibodies/	76218
9	immunoglobulin a/	35695
10	Transglutaminases/	7190
11	((anti-human or antihuman or anti human) adj4 transglutaminase).tw,kw.	33
12	((tissue or anti-tissue or antitissue or anti tissue) adj4 transglutaminase).tw,kw.	3474
13	(iga adj4 transglutaminase).tw,kw.	932
14	(tTg or anti-tTg or anti-httg).tw,kw.	2417
15	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14	256295
16	5 and 15	5942
17	limit 16 to yr="1998 -Current"	5147

EMBASE

1	Celiac Disease/	36614

2	((coeliac or celiac) adj4 disease).tw.	30645
3	Glutens/	10338
4	(gluten adj3 (intoleran* or sensitiv* hypersensi*)).tw,kw.	884
5	1 or 2 or 3 or 4	43688
6	Serologic Tests/	91532
7	serolog*.tw.	174000
8	Autoantibodies/	84246
9	immunoglobulin a/	72322
10	Transglutaminases/	8538
11	((anti-human or antihuman or anti human) adj4 transglutaminase).tw,kw.	56
12	((tissue or anti-tissue or antitissue or anti tissue) adj4 transglutaminase).tw,kw.	5525
13	(iga adj4 transglutaminase).tw,kw.	1912
14	(tTg or anti-tTg or anti-httg).tw,kw.	4302
15	6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14	358893
16	5 and 15	11390
17	limit 16 to yr="1998 -Current"	10753
18	limit 17 to "remove medline records"	5341

CENTRAL

ID	Search	Hits
#1	celiac disease with Publication Year from 1998 to 2023, in Trials (Word variations have been searched)	1113
#2	serology with Publication Year from 1998 to 2023, in Trials (Word variations have been searched)	4342

#3 transglutaminase with Publication Year from 1998 to 2023, in Trials (Word variations have been searched) 186

#4 #2 OR #3 4484

#5 #1 AND 4 552

Web of Science

(TS=("celiac disease" OR "coeliac disease" OR "gluten-sensitive enteropathy" OR gluten) AND TS=("serologic tests" OR "serological tests" OR "antibodies" OR "autoantibodies" OR "tissue transglutaminase" OR "tTG" OR "anti-tissue transglutaminase" OR "anti-tTG"))

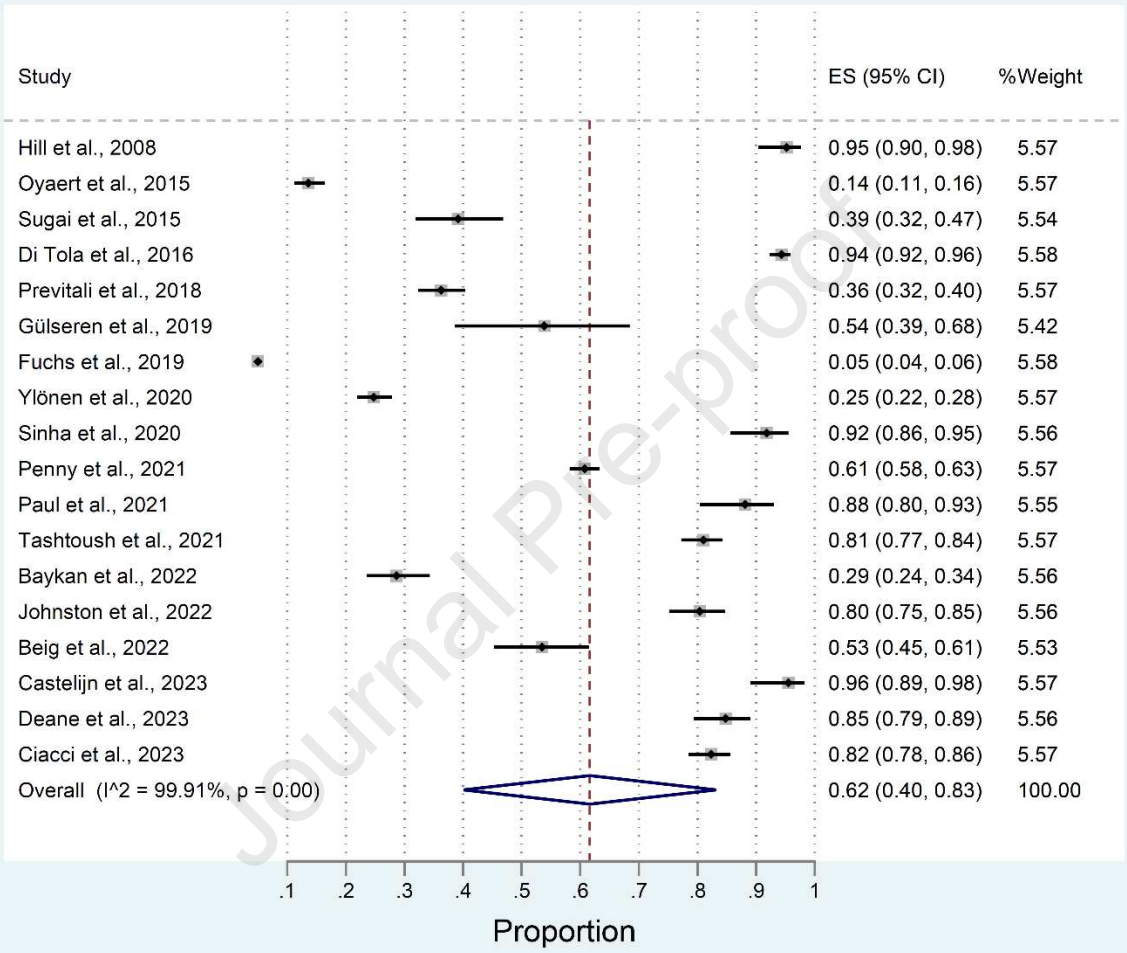
Supplementary Tables

Supplementary Table 1 – Certainty of evidence and grade of recommendation

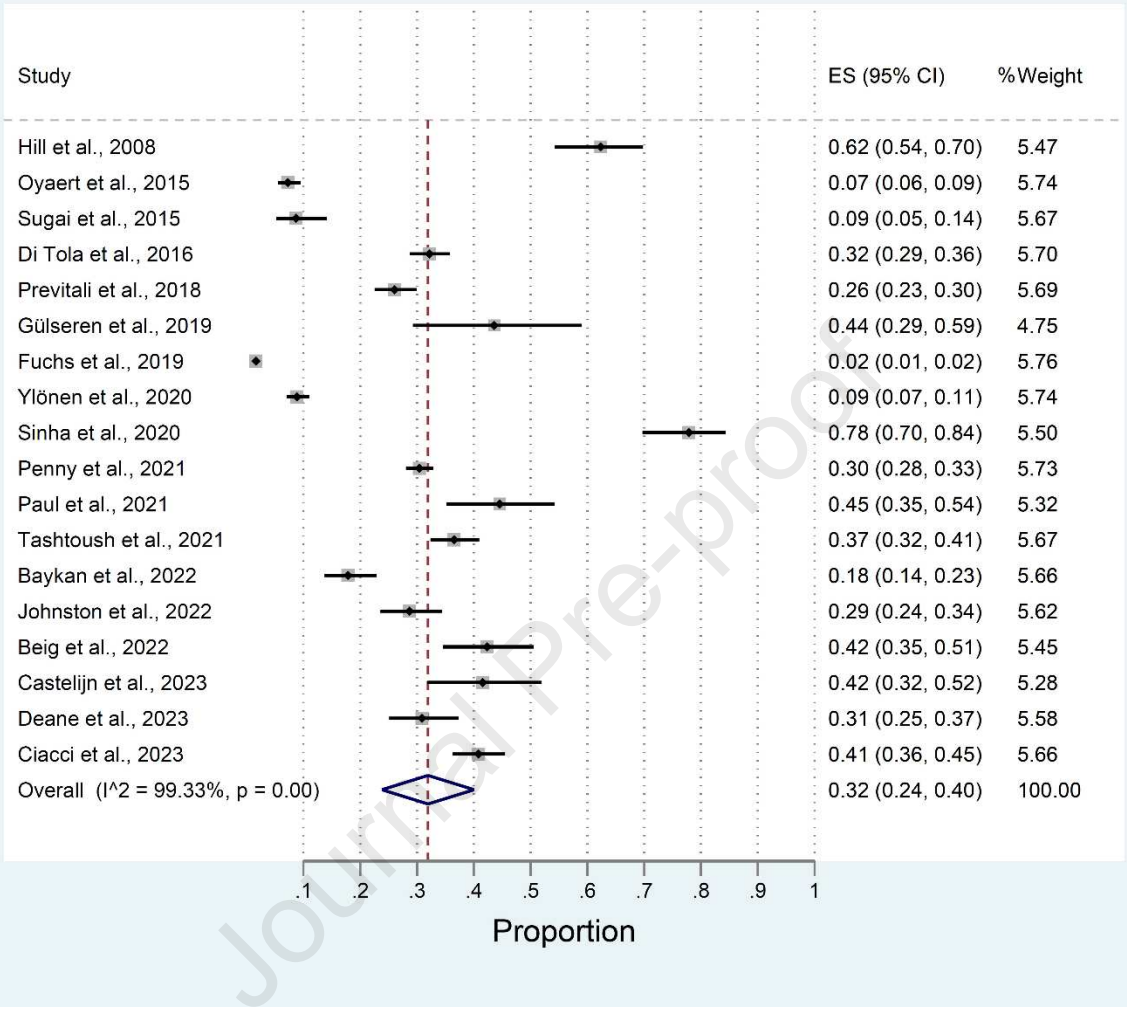
Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy certainty of evidence
			Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	
True positives (patients with Coeliac disease)	18 studies 4071 patients	Observational diagnostic test accuracy studies	serious	not serious	not serious	not serious	none	⊕⊕⊕◯ Moderate
False negatives (patients incorrectly classified as not having Coeliac disease)								
True negatives (patients without Coeliac disease)	18 studies 8032 patients							
False positives (patients incorrectly classified as having Coeliac disease)								

Supplementary Figures

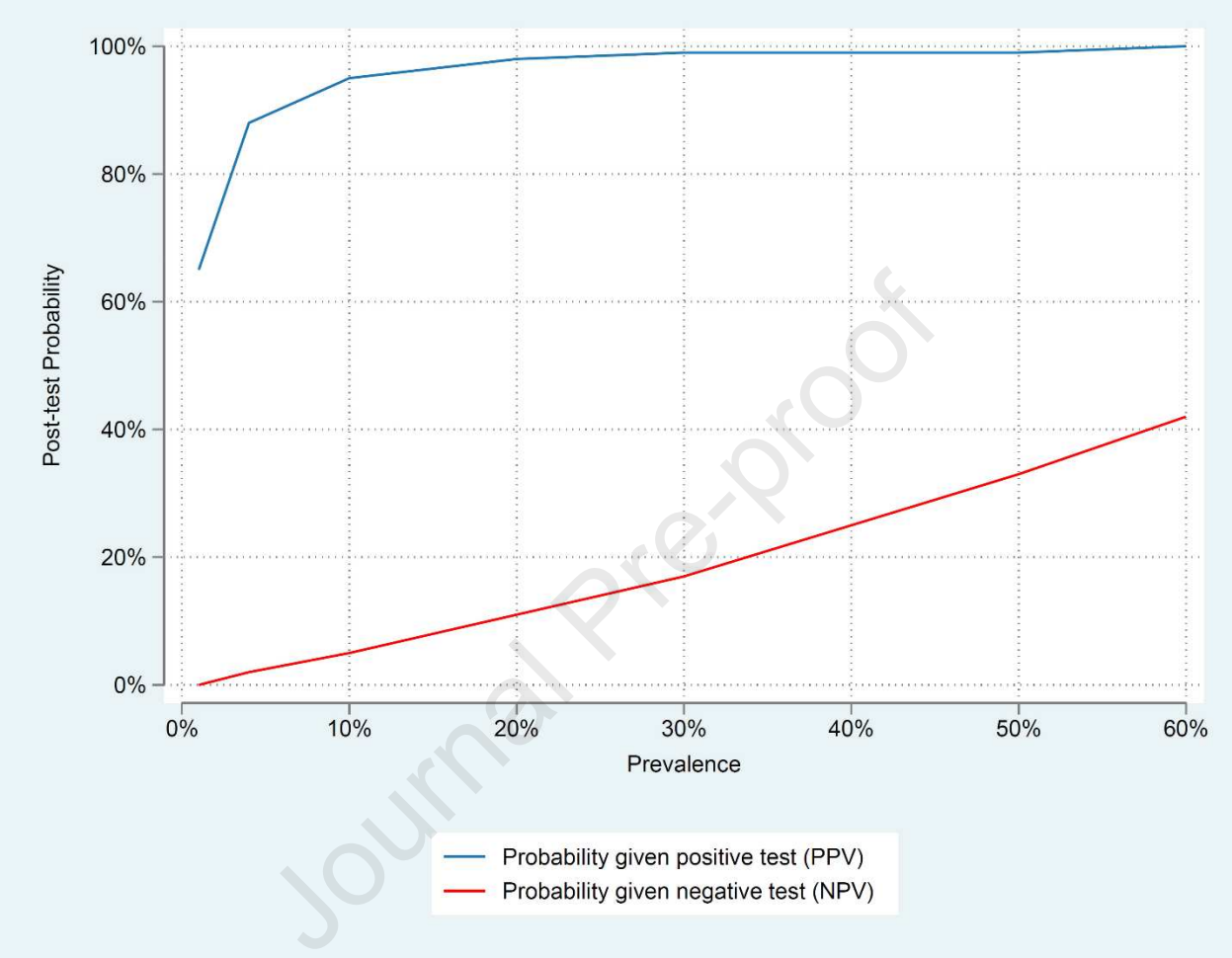
Supplementary Figure 1 – Prevalence of biopsy-proven coeliac disease in the included studies



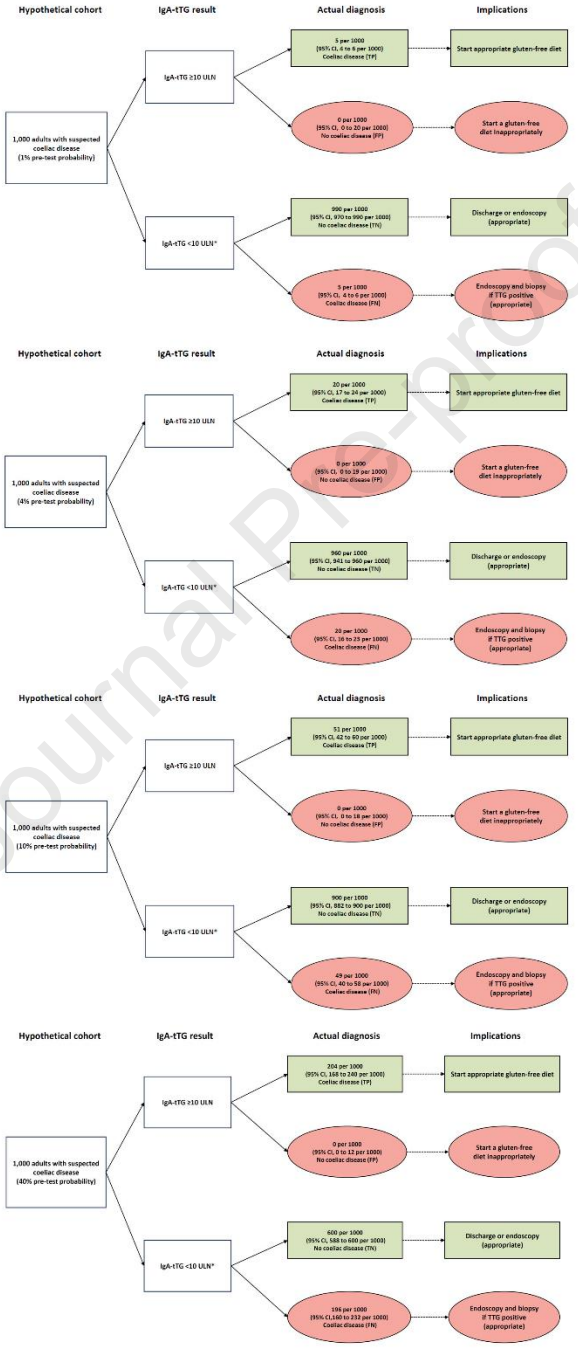
Supplementary Figure 2 – Proportion of patients with IgA-tTG ≥10×ULN in the included studies



Supplementary Figure 3 – Line graph showing the positive and negative predictive values of the no-biopsy approach across different coeliac disease prevalences

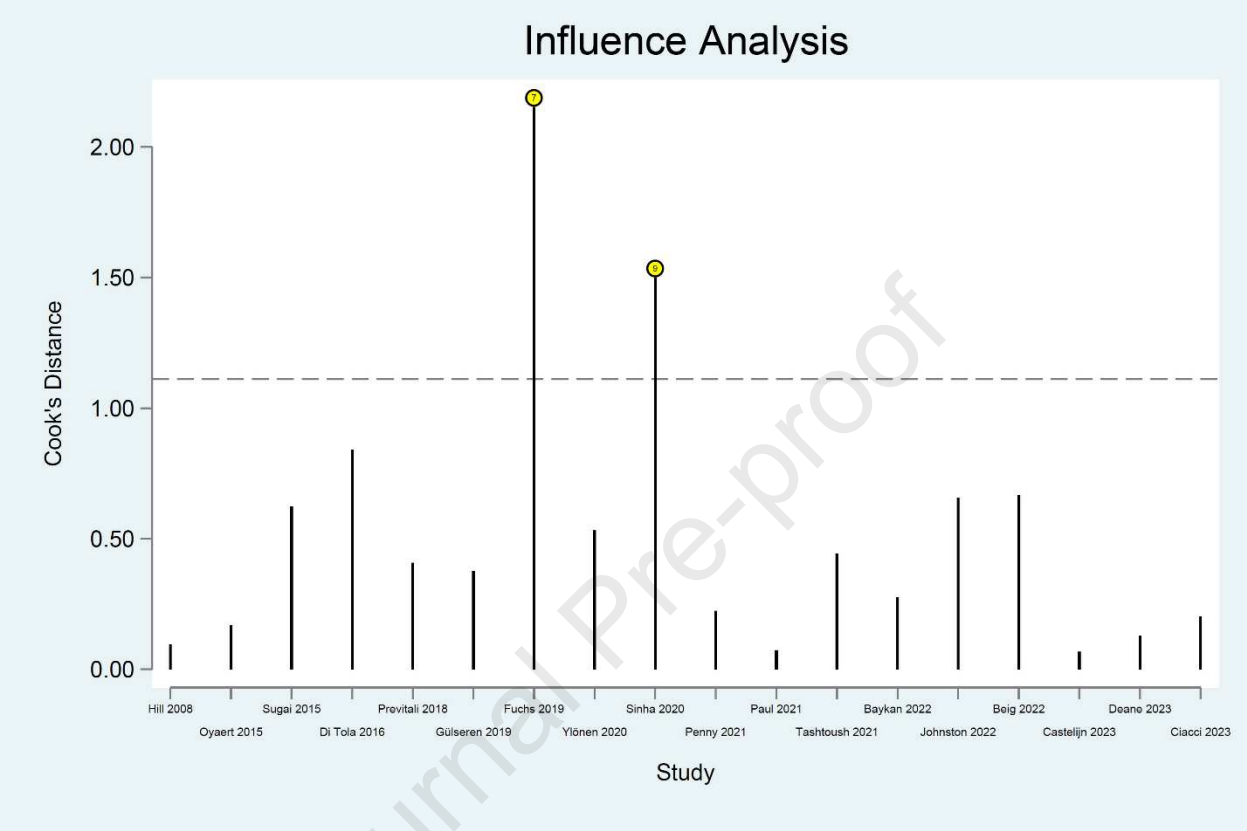


Supplementary Figure 4 – Test consequence graphic showing the results in absolute terms per 1000 patients tested with the downstream consequences of TP, TN, FP, FN in four hypothetical cohorts of adult patients with different pre-test probabilities of coeliac disease

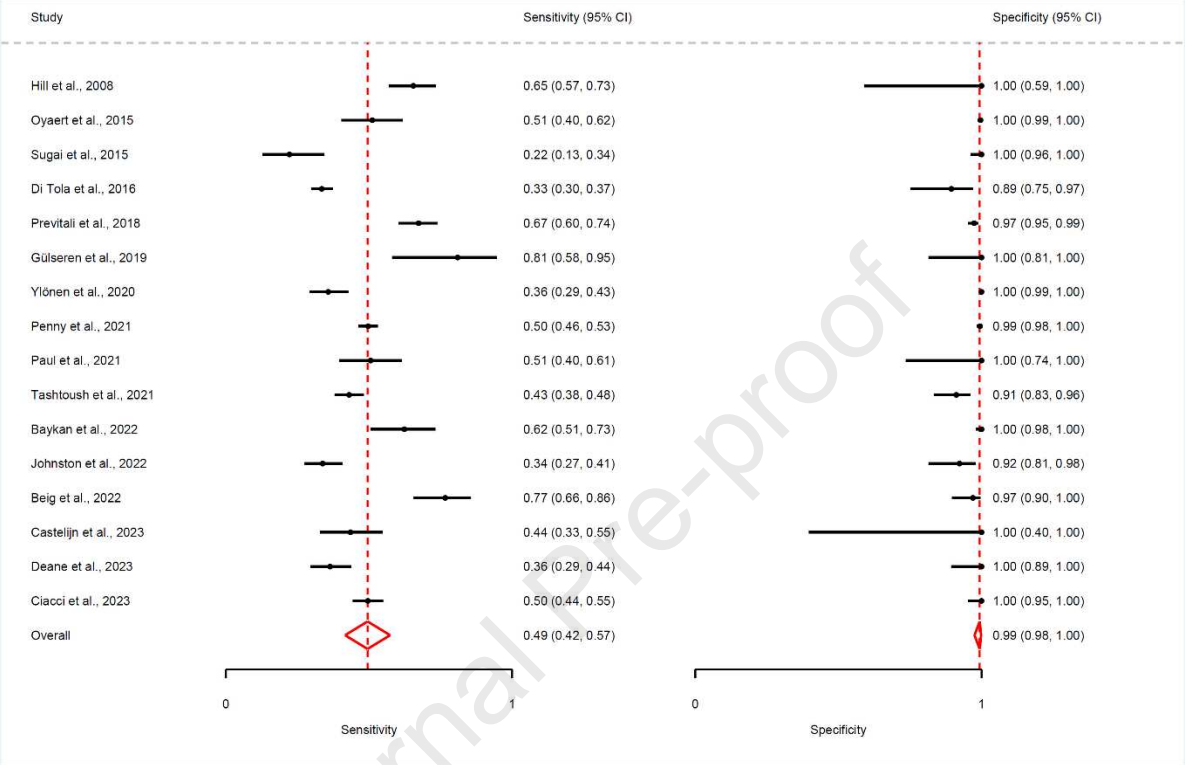


TP: true positive – test is positive (indicates coeliac disease) and patient has coeliac disease
FP: false positive – test is positive (indicates coeliac disease) but patient does not have coeliac disease
TP: true negative – test is negative (indicates coeliac disease not present) and patient does not have coeliac disease
FN: false negative – test is negative (indicates coeliac disease not present) but patient has coeliac disease
*patients with IgA-tTG <10 ULN could either have positive or negative serology which will determine whether endoscopy is indicated or not

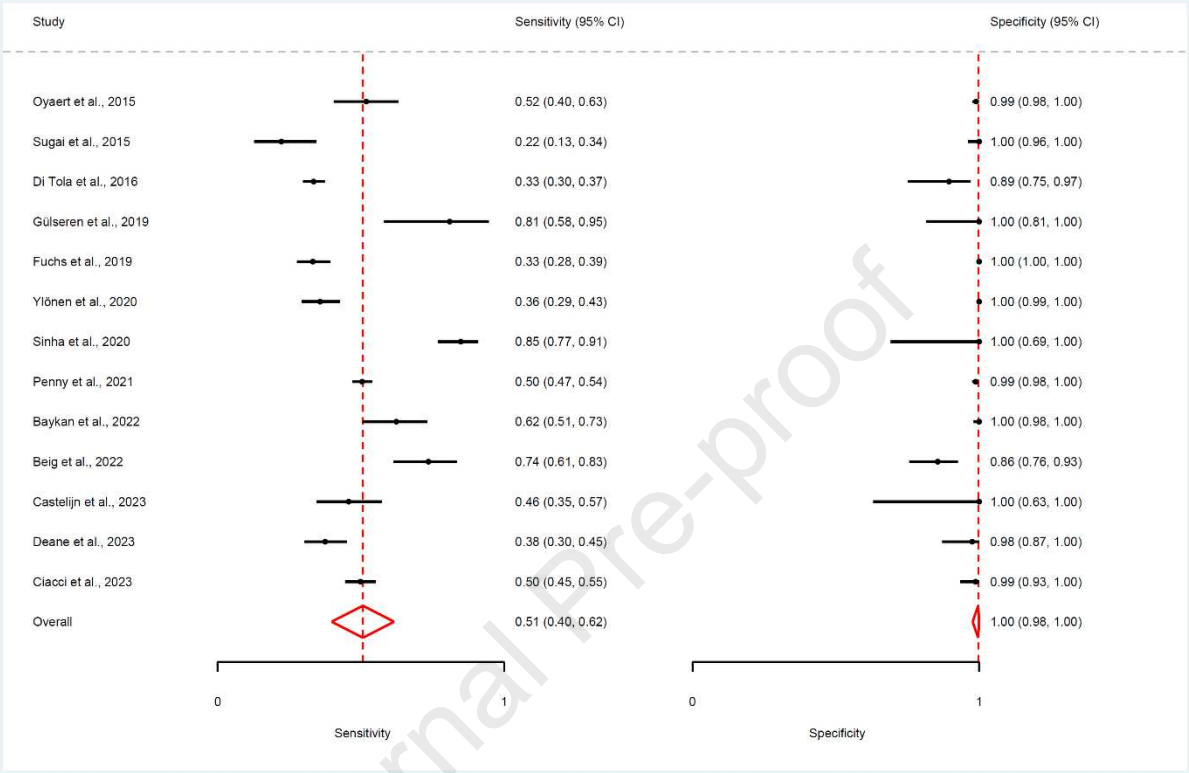
Supplementary Figure 5 – Cook's distance spike plot for detecting highly influential observations



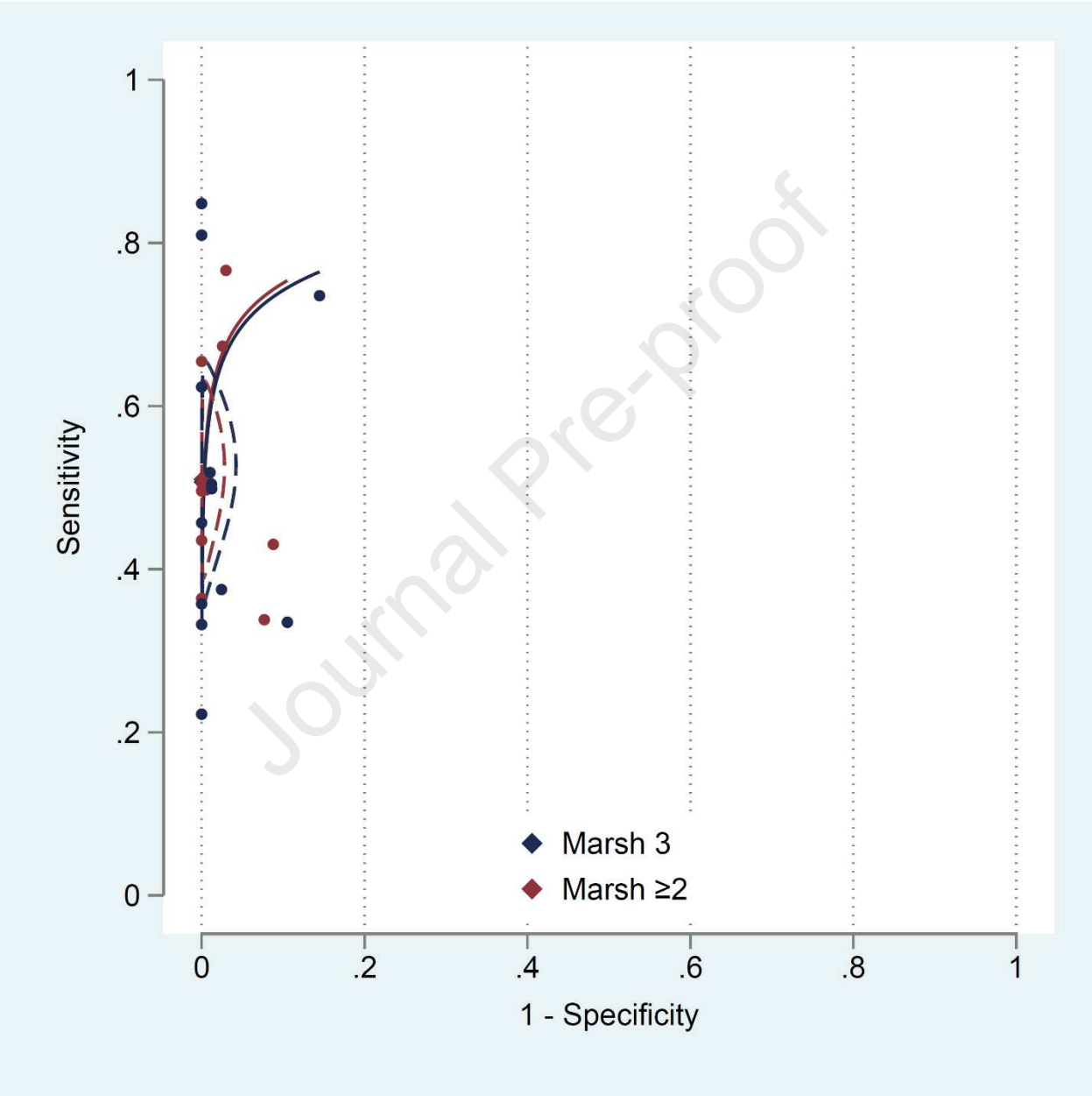
Supplementary Figure 6 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease after excluding outlier studies



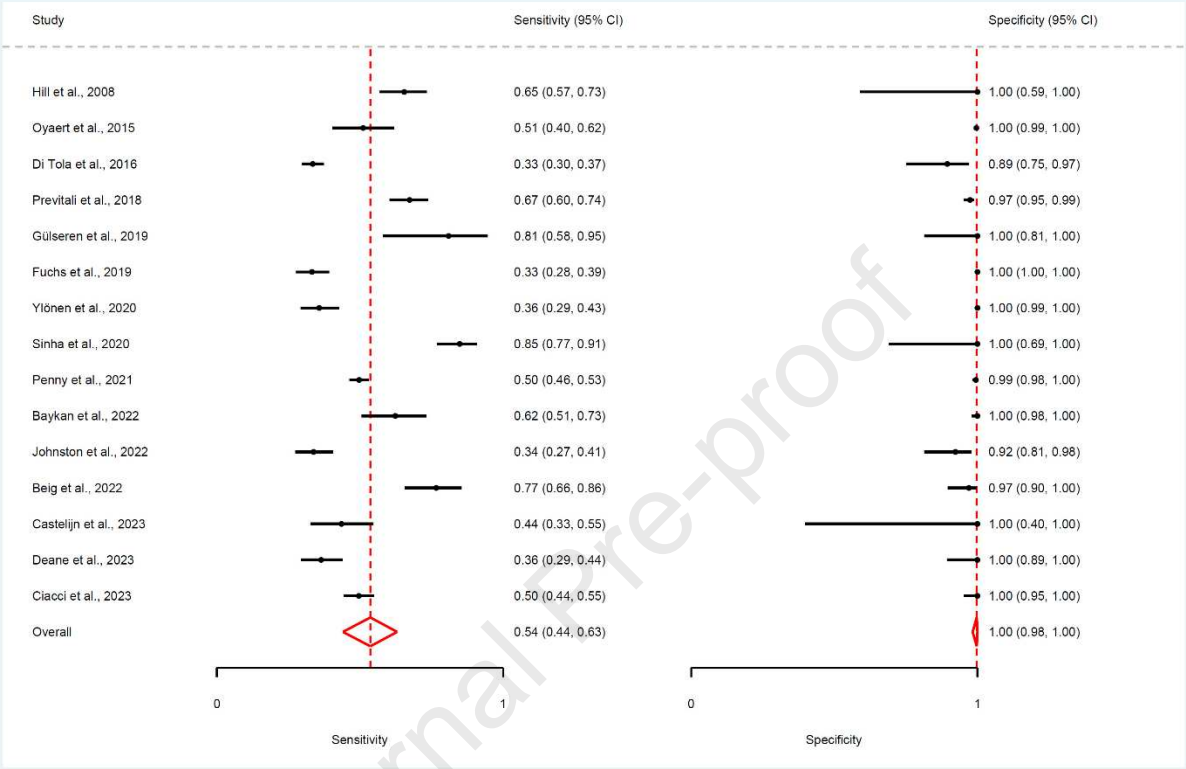
Supplementary Figure 7 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with Marsh 3 lesions



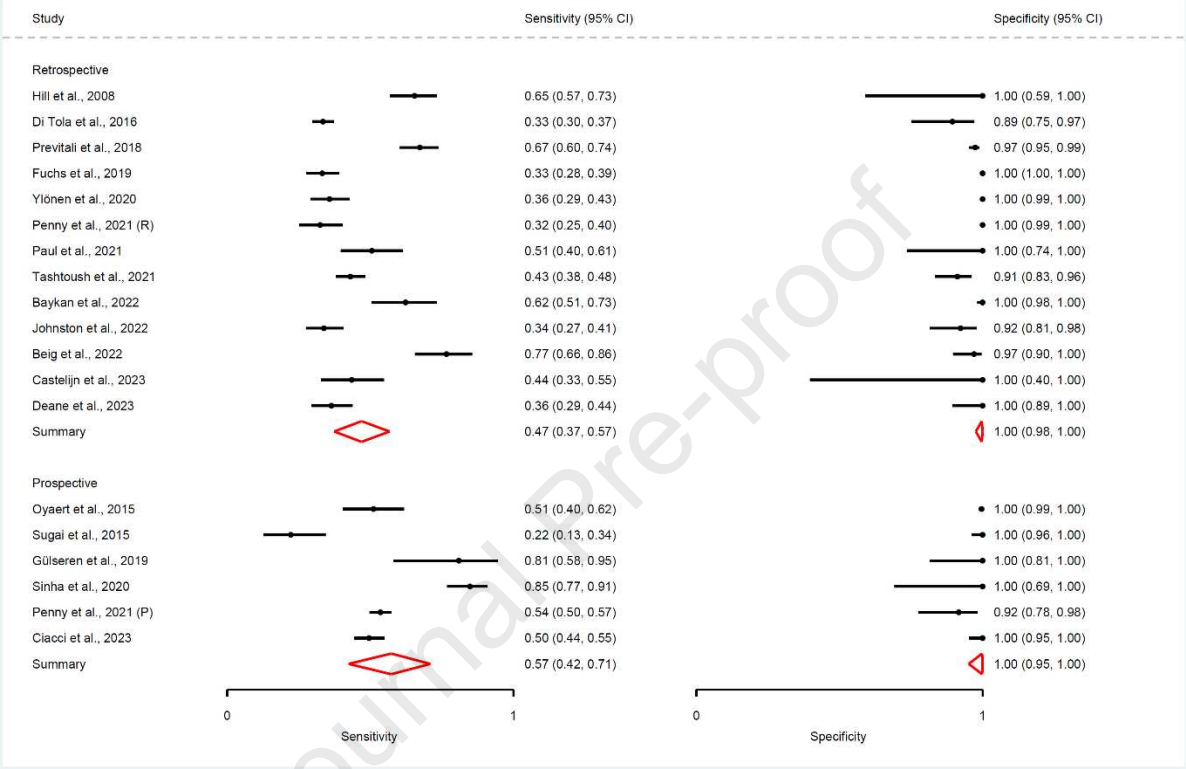
Supplementary Figure 8 - A summary receiver operating characteristic plot of the study estimates of IgA-tTG ≥10×ULN sensitivity and specificity according to Marsh 3 and Marsh ≥2 lesions



Supplementary Figure 9 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease excluding studies published in letters

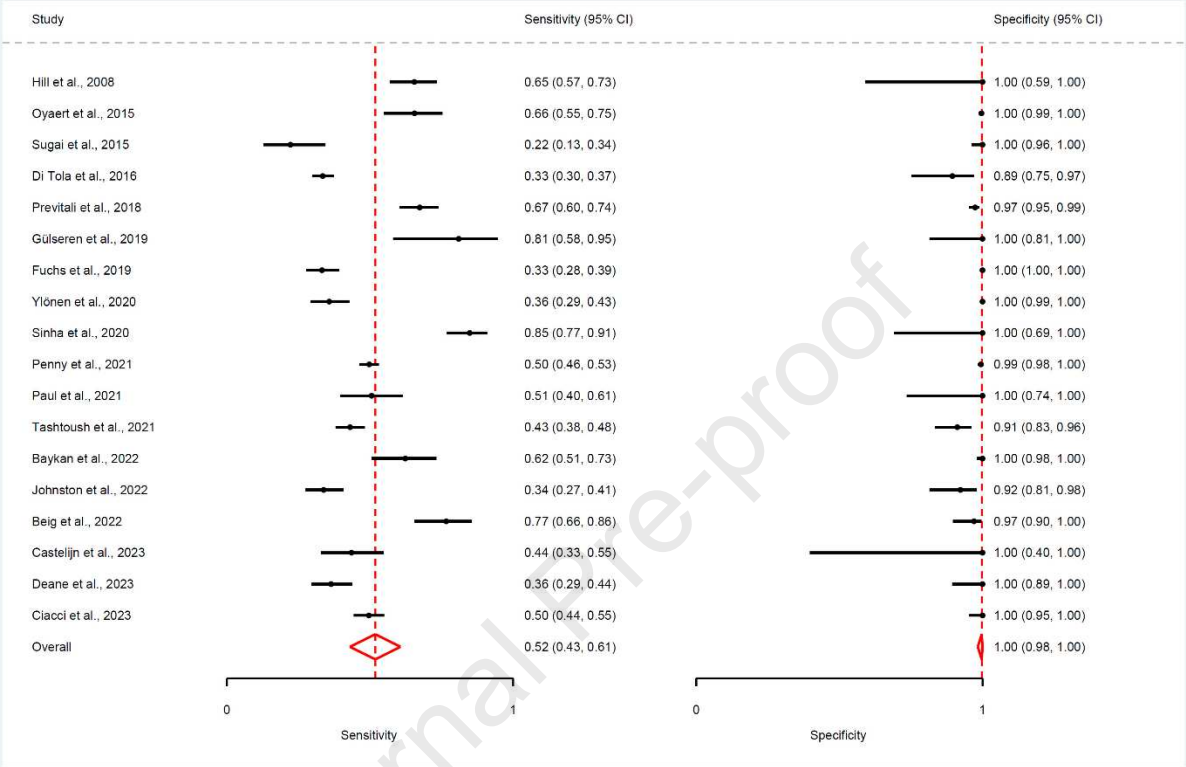


Supplementary Figure 10 – Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease sub-grouped according to study design (Retrospective vs prospective)

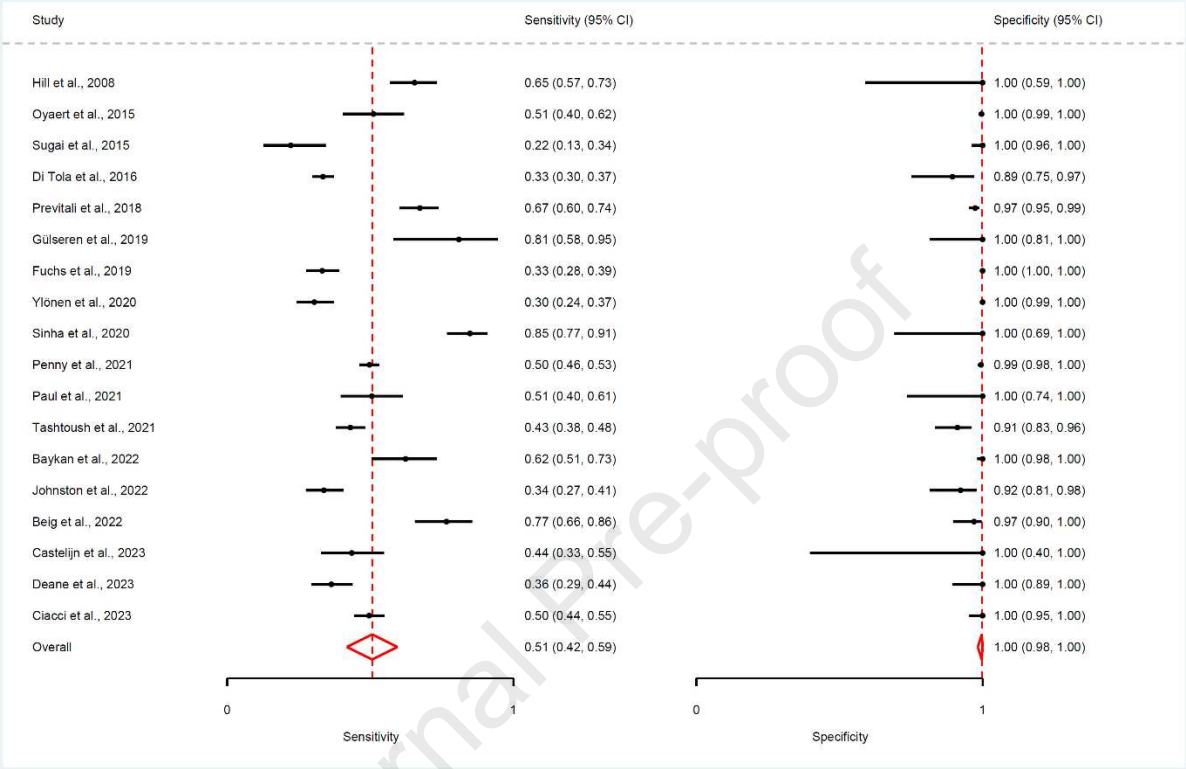


Penny et al., 2021 included one prospective cohort (P) and two retrospective cohorts (R)

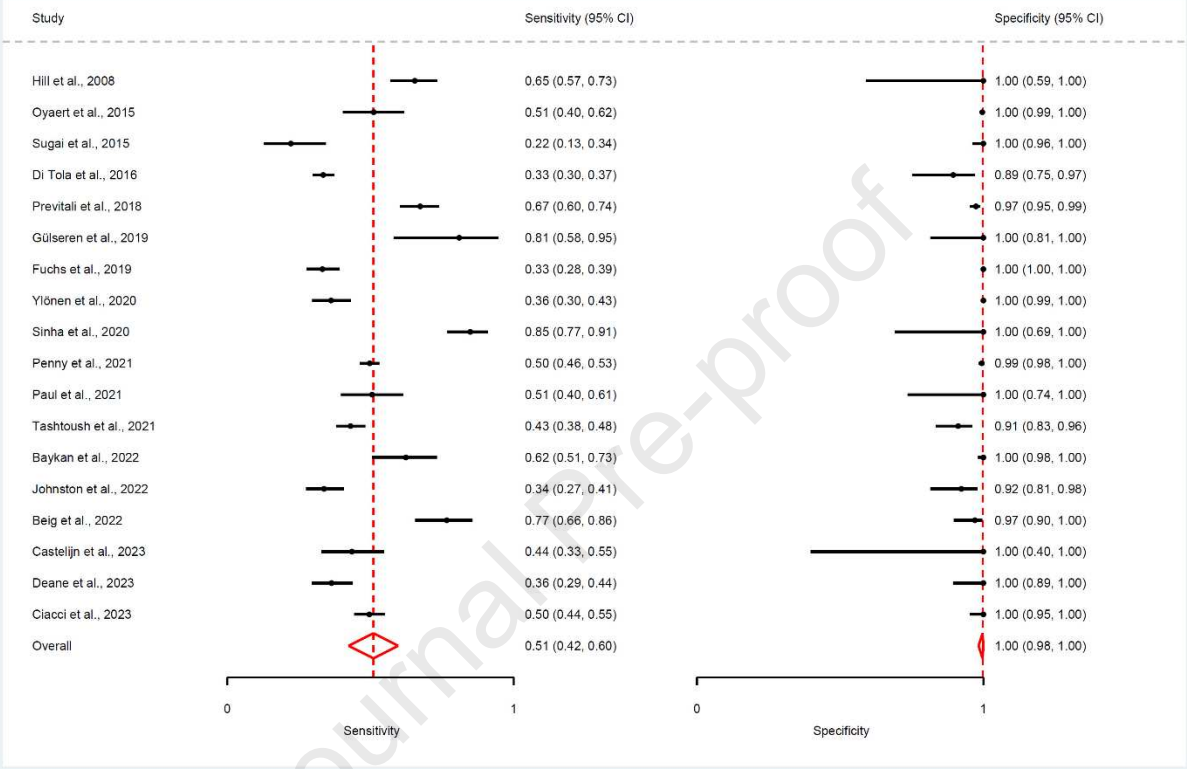
Supplementary Figure 11 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease using the Qunta Flash from Oyaert et al.



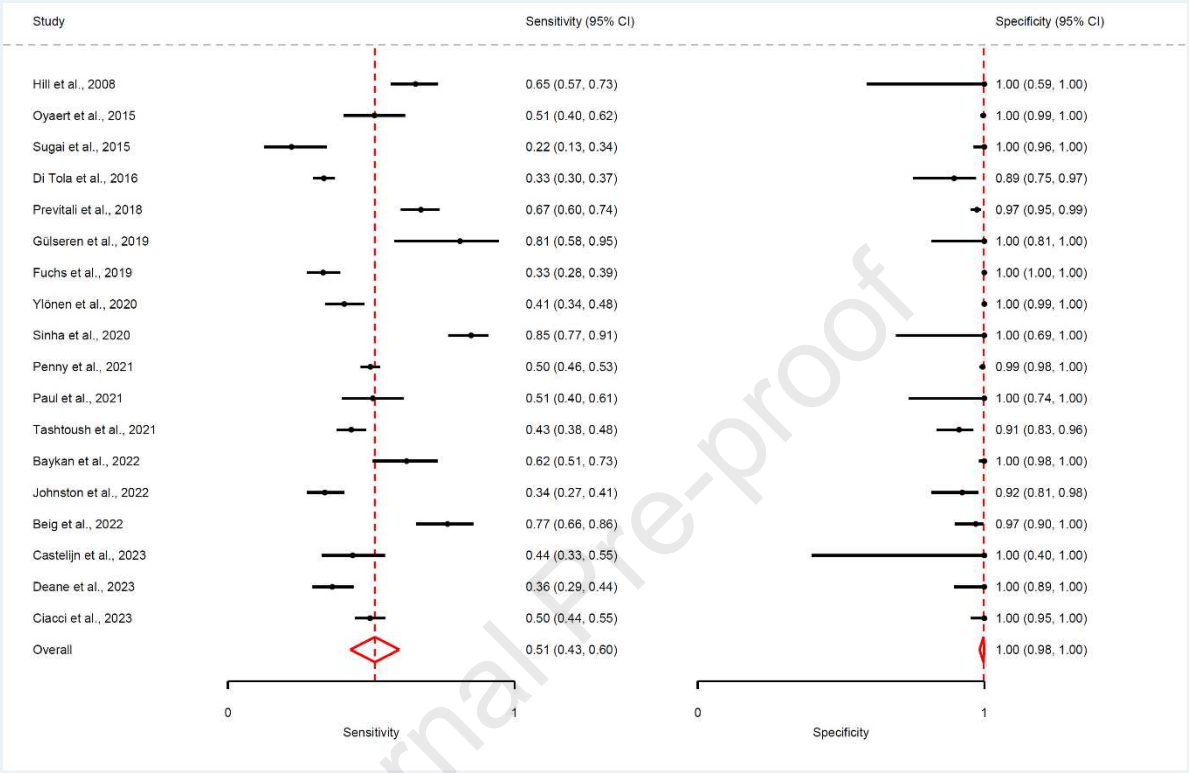
Supplementary Figure 12 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease using Orgentec assay from Ylönen et al.



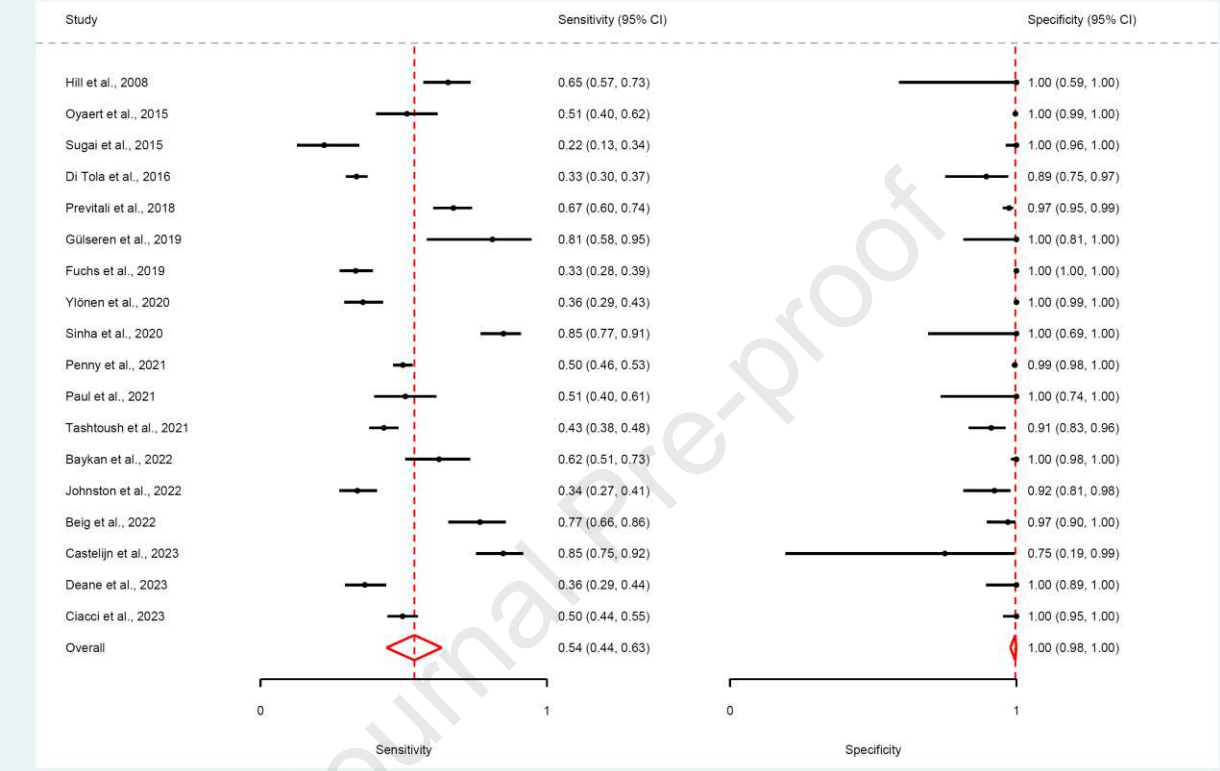
Supplementary Figure 13 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease using Inova Quanta Lite assay from Ylönen et al.



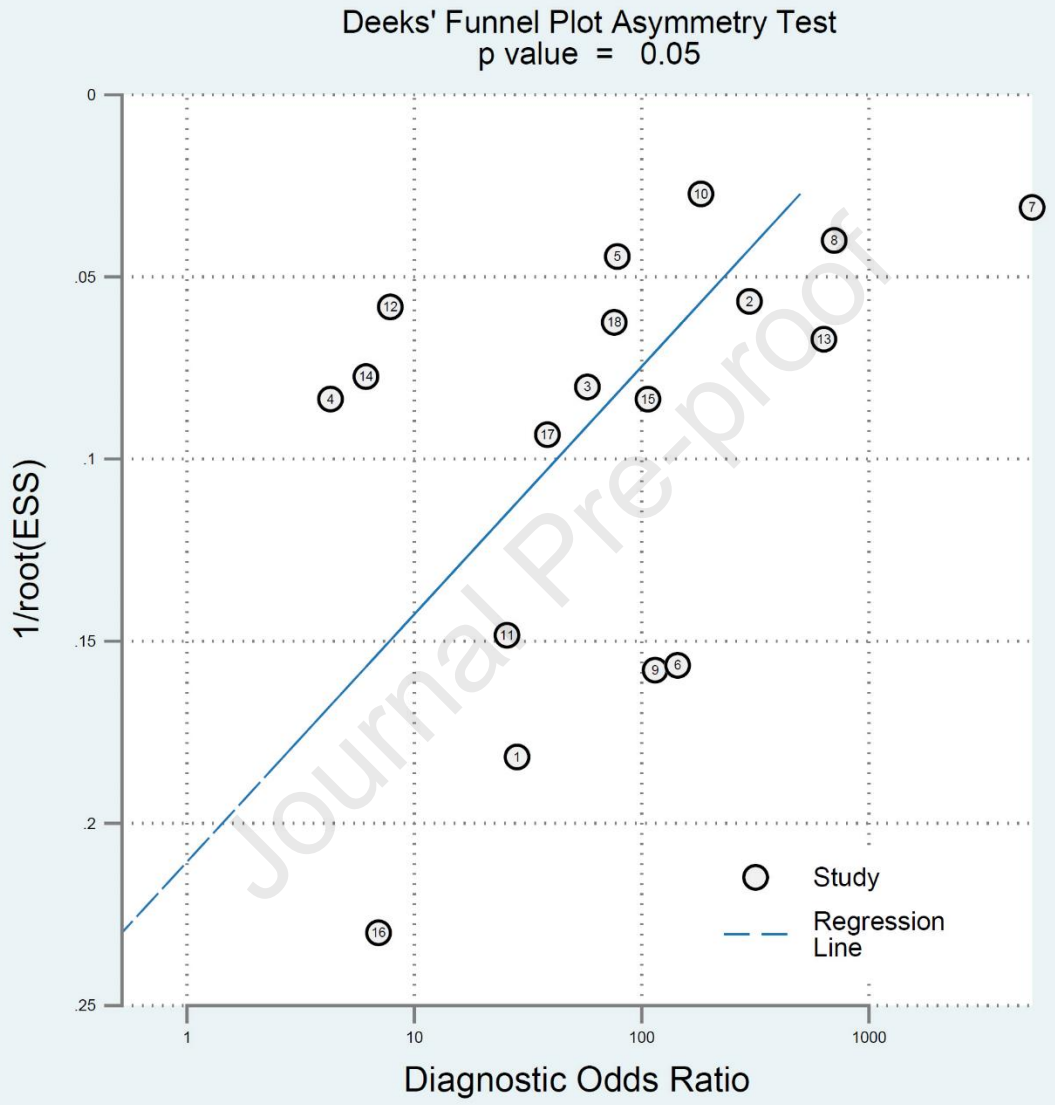
Supplementary Figure 14 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease using Eurospital assay from Ylönen et al.



Supplementary Figure 15 - Forest plot of summary sensitivity and specificity of IgA-tTG ≥10×ULN to identify patients with coeliac disease using the Qunta Flash or Euroimmun assays from Castelijn et al.



Supplementary Figure 16 – Deek's funnel plot of the included studies with superimposed regression line



1 Hill et al., 2008
2 Oyaert et al., 2015
3 Sugai et al., 2015
4 Di Tola et al., 2016
5 Previtali et al., 2018
6 Gülseren et al., 2019
7 Fuchs et al., 2019
8 Ylönen et al., 2020
9 Sinha et al., 2020
10 Penny et al., 2021
11 Paul et al., 2021
12 Tashtoush et al., 2021
13 Baykan et al., 2022
14 Johnston et al., 2022
15 Beig et al., 2022
16 Castelijn et al., 2023
17 Deane et al., 2023
18 Ciacci et al., 2023

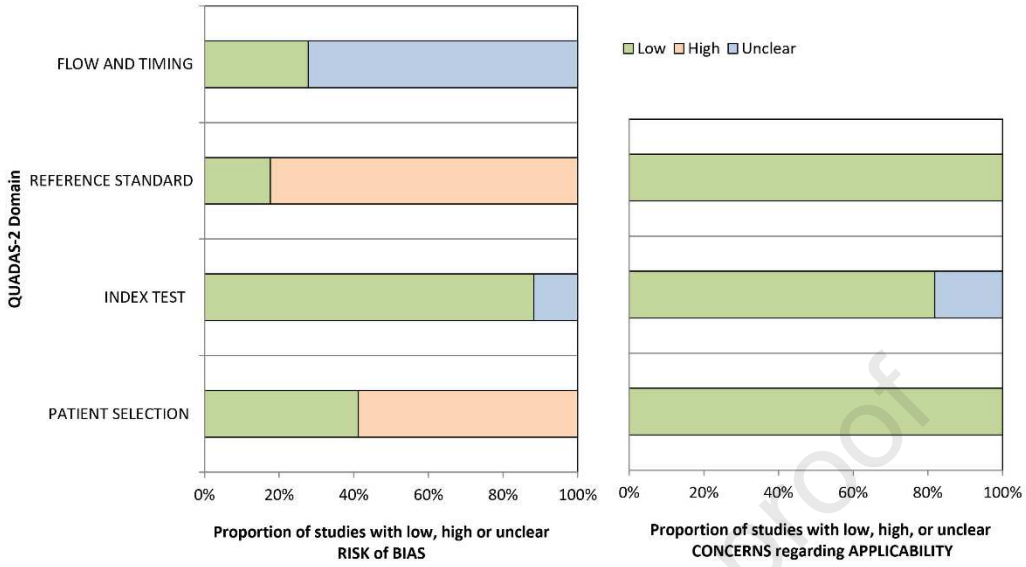
Risk of bias assessment using the QUADAS-2 tool

	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING
Hill et al., 2008	High	Low	High	Low
Oyaert et al., 2015	Low	Low	High	Unclear
Sugai et al., 2015	Low	Low	High	Unclear
Di Tola et al., 2016	Low	Low	Low	Unclear
Previtali et al., 2018	Low	Low	Low	Low
Gülseren et al., 2019	High	Low	High	Unclear
Fuchs et al., 2019	Low	Low	High	Unclear
Ylönen et al., 2020	Low	Low	High	Unclear
Sinha et al., 2020	Low	Low	High	Unclear
Penny et al., 2021	High	Low	High	Low
Paul et al., 2021	High	Unclear	High	Unclear
Tashtoush et al., 2021	High	Unclear	High	Unclear
Baykan et al., 2022	Low	Low	High	Low
Johnston et al., 2022	High	Low	High	Unclear
Beig et al., 2022	High	Low	High	Low
Castelijn et al., 2023	Low	Low	High	Unclear
Deane et al., 2023	High	Low	Low	Unclear
Ciacci et al., 2023	Low	Low	High	Unclear

Applicability concerns assessment using the QUADAS-2 tool

	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Hill et al., 2008	Low	Low	Low
Oyaert et al., 2015	Low	Low	Low
Sugai et al., 2015	Low	Low	Low
Di Tola et al., 2016	Low	Low	Low
Previtali et al., 2018	Low	Low	Low
Gülseren et al., 2019	Low	Low	Low
Fuchs et al., 2019	Low	Low	Low
Ylönen et al., 2020	Low	Low	Low
Sinha et al., 2020	Low	Low	Low
Penny et al., 2021	Low	Low	Low
Paul et al., 2021	Low	Unclear	Low
Tashtoush et al., 2021	Low	Unclear	Low
Baykan et al., 2022	Low	Low	Low
Johnston et al., 2022	Low	Low	Low
Beig et al., 2022	Low	Low	Low
Castelijn et al., 2023	Low	Low	Low
Deane et al., 2023	Low	Low	Low
Ciacci et al., 2023	Low	Low	Low

Summary of risk of bias and applicability concerns assessment using the QUADAS-2 tool



Detailed risk of bias and applicability assessment of the included studies using the QUADAS-2 tool

ID: **1** Author: Hill et al. Year: 2008 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Adult patients (>15 years old) presenting from April 2002 to December 2003 with Raised IgA-tTG and duodenal biopsy performed.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

A total of 146 patients were included (43 males, 103 females) with IgA-tTG >10 U/ml; median male age 48 years and median female age 42 years.

Do the included patients and setting match the question? Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Transglutaminase antibody was measured with human recombinant tTG as antigen using the

Celikey kit (Phadia GmbH,Freiburg, Germany). ULN values provided.

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ fromt he review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. Marsh > or equal 2 considered diagnostic for coeliac disease. No information on site/ number of duodenal biopsy sampling.

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Two patients were excluded because of the long interval between serology and biopsy

Describe the time interval and any interventions between index test(s) and reference standard:

Variable. Data provided in the study.

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Low

ID: **2** Author: Oyaert et al. Year: 2015 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

A total of 98 consecutive adults (older than 16 years) and 58 consecutive children (45 between 2 and 15 years old and 13 ≤ 2 years old) diagnosed with coeliac disease and 974 consecutive patients diagnosed as not having CD (disease controls) aged three months or older. All recruited patients underwent duodenal biopsy. Serum was collected at the time of endoscopy in all patients. Excluded patients: patients on GFD and/or previously diagnosed with coeliac disease

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	No
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

98 adults with coeliac disease were included (28 males/70 females) and 564 without coeliac disease (data regarding diagnostic accuracy in adults obtained after contacting the corresponding author)

Do the included patients and setting match the question? Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Blood samples were tested for TGA with 2 assay from 2 different manufacturers (EliA Celikey ® IgA - Thermo Fisher

and QUANTA FLASH ® IgA tTG IgA Inova). ULN values provided.

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy with Marsh grade 3 was considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of Unclear
the index test?

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability: Low
reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from
the 2x2 table (refer to flow diagram):

Excluded from study: GFD and previous Dx CD.

Excluded from 2x2: none.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **3** Author: Suagi et al. Year: 2015 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Prospective (post-hoc analysis). Cross-sectional. Adults >18 years old. Symptomatic. Post-hoc analysis of data collected from a previous prospective study, including adult patients (>18 years old) with potential intestinal disorders. Serology results including IgA-tTG and duodenal histology from 161 adults with were analysed.

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

A total of 161 patients were included, of whom 63 were diagnosed with coeliac disease.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Serology was performed using the QUANTA LiteTM assay, h-tTG IgA (Inova Diagnostic Inc., San Diego, CA, USA). ULN reported.

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted::

Duodenal biopsy. >3x biopsies from D2. Marsh grade 3 considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Excluded patients with prior serology, a previous diagnosis of CD, prior treatment with a gluten-free diet, or a former diagnosis of dermatitis herpetiformis.

Describe the time interval and any interventions between index test(s) and reference standard:

Serology was performed at the time of endoscopy.

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive a reference standard? Yes

Did patients receive the same reference standard? Yes

Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Low

ID: **4** Author: Di Tola et al. Year: 2016 Reviewer:

Consensus

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Consecutive adult patients (>18 years old) referred to the Gastroenterology Unit of Polyclinic Umberto from February 2009 to October 2014 were evaluated in this retrospective study. All patients were on a gluten-containing diet, evaluated clinically, subjected to blood collection for serology and in case of discordant results they also performed HLA typing.

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

Consecutive adult patients [149 male/ 522 female (ratio 1:3.5), median age 34, range 18–65 years]

Do the included patients and setting match the question? Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

QUANTA Lite R h-tTG IgA assay was used (INOVA Diagnostics, San Diego, CA). ULN levels reported.

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias? Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsies with Marsh grade 3 or 4 were diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability:

Low

reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

All patients were included, and all received both the index test and reference standard.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **5** Author: Previtali et al. Year: 2018 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Of 2565 consecutive patients on whom duodenal biopsy was performed at Papa Giovanni XXIII

Hospital from July 2012 to September 2016, 857 were included. All the patients who had undergone serological testing for coeliac disease analysed within +/- 3 months of duodenal biopsy and before the start of a gluten-free diet.

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

The study included 857 adults and children with suspected coeliac disease/

Do the included patients and setting match the question? Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

IgA-tTG levels were measured using the QUANTA Flash® h-tTG IgA assay (inova diagnostics)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. Multiple biopsies from D2. Marsh 2 or above were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

The study excluded patients who were unable to obtain clinical information and/or final diagnosis; IgA deficiency; dermatitis herpetiformis, or on a gluten-free diet at the time of assessment.

Excluded from 2x2: none.

Describe the time interval and any interventions between index test(s) and reference standard:

Serology was performed within 3 months of endoscopy, while patients on a gluten-containing diet.

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Low

ID: **6** Author: Gulseren et al. Year: 2019 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Adult patients (>17 years old) referred to a gastroenterology clinic because of the gastrointestinal symptoms and extraintestinal symptoms that could be associated with coeliac disease. Patients with malignancy, pregnant women, and Crohn's disease were excluded.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	No
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Thirty-nine patients with suspected coeliac disease. They all underwent duodenal sampling and serology, including IgA-tTG.

Do the included patients and setting match the question?Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Immo Diagnostics Enzyme Immunoassay Kits (Immo Diagnostics, Buffalo, NY, USA) were used to determine IgA-tTG. ULN reported.

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ fromt he review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsies with Marsh Grade 3 were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability: Low

reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

Excluded patients with known malignancy, Crohn's disease and pregnant women.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **7** Author: Fuchs et al. Year: 2019 Reviewer:

Consensus

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective evaluation of previous prospective studies. Different groups of patients with different pre-test probabilities. the high-risk group were patients referred for suspected coeliac disease based on clinical symptoms, the moderate risk were at-risk family members recruited from previously diagnosed individuals, and the low-risk group were patients selected randomly among a group of patients not presenting with possible coeliac disease-related symptoms.

Was a consecutive or random sample of patients enrolled? Yes

Was a case-control design avoided? Yes

Did the study avoid inappropriate exclusions?: Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

Three different cohorts with different pre-test probabilities of coeliac disease (low/moderate/high). Patients with known coeliac disease, dermatitis herpetiformis or on a gluten-free diet were excluded.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Cohorts 1 & 3 - EliA Celikey - ELISA

Cohort 2 - QUANTA Lite - ELISA

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

DOMAIN 3: REFERENCE STANDARD

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. Multiple biopsies from D2. Marsh 2 or above was considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability: Low

reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

Excluded patients on a gluten-free diet prior to biopsy, incomplete tests, refusal for biopsy, and death.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **8** Author: Ylonen et al. Year: 2020 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

This retrospective study was conducted at the Celiac Disease Research Center, Tampere University and Tampere University Hospital. The patients were collected from 836 adults, who were further categorised into two sub-cohorts based on assumed pre-test probability for celiac disease.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Unclear

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Adult patients (>18 years old) with suspected coeliac disease were included. Excluded patients with known coeliac disease or a on a gluten-free diet.

Do the included patients and setting match the question?Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

QUANTA Lite; Celikey; Orgentec; Eurospital - all ELISA.

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsies (Marsh 3 lesions). Some of the patients with inconclusive histology received the diagnosis in a re-biopsy after one year ("Gluten challenge"). In a subset, the diagnosis was set on the basis of special investigations and clinical, serological, and histological responses to the gluten-free diet. At least 4x duodenal biopsies were sampled.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

Excluded from study: none.

Excluded from 2x2: Marsh grade <3 in one aspect of analysis

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **9** Author: Sinha et al. Year: 2020 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

A single-centre prospective observational study included consecutive patients with suspected coeliac disease defined as a history suggestive of malabsorption and positive coeliac serology (IgA tTG-Ab ≥ 10 EliA U/mL)

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes

Did the study avoid inappropriate exclusions?: Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

Included 122 patients with suspected coeliac disease, median age 27 years (range 16 – 35) with a male-to-female ratio of 1:1.26. Patients with chronic kidney disease, cardiac disease, GI bleeding and pregnancy were excluded.

Do the included patients and setting match the question?Concerns regarding applicability:

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Celikey - ELISA.

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

If a threshold was used, was it pre-specified? Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

DOMAIN 3: REFERENCE STANDARD

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. 4x D2 biopsies. Marsh 3 lesions were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of Unclear
the index test?

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability: Low
reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from
the 2x2 table (refer to flow diagram):

Patients with chronic illnesses (chronic kidney disease, cirrhosis, congestive cardiac failure), active overt GI bleeding, clinically significant coagulopathy and who were pregnant/lactating and were unwilling for endoscopy were excluded from the study.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID:**10** Author: Penny et al. Year: 2021 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Three cohorts. Cross-sectional. Adults >16 years old.

1. Prospective. Consecutive. Reviewed in CD clinic with TTG to biopsy 6 weeks or less, no previous coeliac disease diagnosis, IgA competent, naivety to GFD.

2. Retrospective of prospective. Referred for UGI endoscopy due to symptoms.

3. Retrospective. Positive TTG and appropriate duodenal biopsies were available.

Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Cohort 1: 740 patients were recruited from CD specialistic clinic at RHH (Sheffield, UK).

Cohort 2: 778 patients were included.

Cohort 3: multicentre, international, all tertiary centres for CD with 145 patients included.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Cohort 1 - Aeskulisa & Celikey

Cohort 2 - Celikey

Cohort 3 - Multiple

All ELISAs.

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

DOMAIN 3: REFERENCE STANDARD

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. Marsh 3 lesions were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	No

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability: Low

reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

Excluded patients with IgA deficiency; already consuming a gluten-free diet; previous diagnosis of coeliac disease; IBD; incomplete testing or missing information; positive EMA test in primary care. No patients were excluded from cohort 1.

Describe the time interval and any interventions between index test(s) and reference standard:

serology and duodenal biopsy performed within six weeks from each other in the first and second cohorts.

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Low

ID:**11** Author: Paul et al. Year: 2021 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective cross-sectional study of patients with suspected coeliac disease (Positive IgA-tTG) from January 2013 to June 2020. Published as a letter.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	No

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Excluded children (<16 years old) and those who did not undergo endoscopy.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Unclear.

Were the index test results interpreted without knowledge of the results of the	Unclear

reference standard?

If a threshold was used, was it pre-specified? Unclear

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Unclear

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Unclear

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. Unclear on site/ sampling. Marsh 2 or more were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

Excluded from study: age <16 years old; not referred for biopsy; no biopsy done.

Excluded from 2x2: none.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID:**12** Author:Tashtoush et al. Year: 2021 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective analysis of patients with positive IgA-tTG who underwent endoscopy and duodenal biopsies. Reported in a letter.

Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Included adult patients with positive serology who underwent endoscopy and biopsy. Excluded children and those who were not referred or did not tolerate endoscopy.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Unclear.

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear
If a threshold was used, was it pre-specified?	Unclear

Could the conduct or interpretation of the index test have introduced bias?Risk of bias:

Unclear

B. APPLICABILITY:

Is there concern that the index test, its conduct, or **Concerns regarding applicability:**

Unclear

interpretation differ from the review question

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. Marsh grade 2 or more were considered diagnostic for coeliac disease. No details on site/ sampling of biopsies.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Excluded from the study: patients <16 years old; didn't want/ tolerate/ not suitable for endoscopy.

Excluded from 2x2: none

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **13** Author: Baykan et al. Year: 2022 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective, cross-sectional included patients who presented to the Gastroenterology Clinic of the Erzurum Regional Training and Research Hospital between January 1, 2018, and January 1, 2019, and underwent a serological test for coeliac disease and were also sampled for coeliac disease on endoscopy were included.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Recruitment among patients attending the outpatient clinic: 269 (165 Females). Serology for coeliac was performed in all patients as well as OGD with duodenal sampling. Patients younger than 18 years of age, those with more than one month of duration between celiac serology and endoscopy, those with a previous diagnosis of CD, those on a gluten-free diet, and those with selective IgA deficiency were excluded from the study.

Do the included patients and setting match the question?Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Orgentec - ELISA

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy; at least four from D1/D2. Full results for Marsh classification reported.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of Unclear
the index test?

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability: Low
reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from
the 2x2 table (refer to flow diagram):

Excluded from study: Patients younger than 18 years of age, those with more than one month of duration between celiac serology and endoscopy, those with a previous diagnosis of CD, those on a gluten-free diet, and those with selective IgA deficiency were excluded from the study. Excluded from 2x2: none.

Describe the time interval and any interventions between index test(s) and reference standard:

One month

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Low

ID: **14** Author: Johnston et al. Year: 2022 Reviewer:

Consensus

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective analysis of patients with positive IgA-tTG who had endoscopy and duodenal biopsies between 2012 and 2019.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Included 265 adult patients who had positive serology and underwent endoscopy with duodenal biopsies. Children and those who did not have endoscopy were excluded.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Orgentec for some - ELISA

QUANTA Flash - chemiluminescense

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. At least 5 from D1/D2. Marsh 2 or more were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition?	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the Concerns regarding applicability:

Low

reference standard does not match the review question?

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from

the 2x2 table (refer to flow diagram):

Excluded from the study: no biopsy, didn't want/ tolerate endoscopy.

Excluded from 2x2: none.

Describe the time interval and any interventions between index test(s) and reference standard:

Variable. Median three months (Range 1-13 months).

Was there an appropriate interval between index test and reference standard?	No
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** High

ID: **15** Author: Beig et al. Year: 2022 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective cross-sectional analysis of a laboratory database. Included patients with positive serology from January 2013 to December 2018 who also underwent endoscopy and duodenal biopsies while on a gluten-containing diet.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Included patients with available serology and biopsy results. Patients with duodenal biopsies performed before serology, children (<16 years old) and those who did not undergo endoscopy or already diagnosed with coeliac disease were excluded.

Do the included patients and setting match the question?Concerns regarding applicability:

Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Bio-Rad EIA - ELISA.

Were the index test results interpreted without knowledge of the results of the	Yes

reference standard?

If a threshold was used, was it pre-specified? Yes

Could the conduct or interpretation of the index test have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsies; median 5 biopsies. Marsh 1-3 lesions considered for coeliac disease diagnosis (full results for Marsh classification were provided by the corresponding author).

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? No

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who

were excluded from

the 2x2 table (refer to flow diagram):

Excluded from study: patients <16 years old, previous diagnosis of coeliac disease, and those with incomplete information

Excluded from 2x2: none

Describe the time interval and any interventions between index test(s) and reference standard:

Variable time interval between serology and biopsy.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **16** Author: Castelijn et al. Year: 2023 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

Retrospective study including adults and paediatric patients who underwent serological testing for coeliac disease. The corresponding author provided data from adult patients who underwent serology and duodenal biopsies. Controls were not included in accuracy analyses.

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	No
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Describe included patients:

Included adult patients who underwent serology and endoscopy with duodenal biopsies. Patients who did not undergo endoscopy and biopsy were excluded.

Do the included patients and setting match the question?Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Celikey - ELISA

QUANTA Flash & ChLIA - chemoluninescence

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Full data for Marsh classification was provided by the corresponding author.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? Concerns regarding applicability: Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Adult patients who did not have endoscopy and biopsies were excluded from accuracy analyses.

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	No
Did patients receive the same reference standard?	No
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID:**17** Author: Deane et al. Year: 2023 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

A single-centre, retrospective study included patients with suspected coeliac disease who underwent serology testing and duodenal biopsies from 2014 to 2019.

Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias? **Risk of bias:** High

B. APPLICABILITY:

Describe included patients:

Included patients with positive coeliac serology who underwent endoscopy and biopsy. Patients with IgA deficiency, on a gluten-free diet, <16 years old or had previous biopsies at other centres were excluded. The study also excluded those with negative serology and biopsy results

and those with negative serology and positive biopsy results (those were included in our meta-analysis)

Do the included patients and setting match the question?Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

EliA Celikey IgA assay on the Phadia250 Immunoassay analyser (Thermo Scientific, Phadia AB, Uppsala, Sweden) was used.

Were the index test results interpreted without knowledge of the results of the reference standard?	No
If a threshold was used, was it pre-specified?	Yes

Could the conduct or interpretation of the index test have introduced bias?Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question **Concerns regarding applicability:** Low

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

All patients had D2-3 biopsies taken, and D1 at the endoscopist's discretion. Not specified how many D2 biopsies. Histological evaluation was performed by an expert pathologist blinded to serological results. Marsh classification was used. Coeliac disease was diagnosed in the presence of Marsh 3 lesion.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Yes

Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk of bias:** Low

B. APPLICABILITY:

Is there concern that the target condition as defined by the reference standard does not match the review question? **Concerns regarding applicability:** Low

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

325 eligible patients

Excluded: 101 as they fitted exclusion criteria, 7 for IgA def, 29 serology and biopsy neg, 24 serology neg and biopsy pos.

164 has positive serology.

67 had TGA > x 10 ULN

Describe the time interval and any interventions between index test(s) and reference standard:

not defined

Was there an appropriate interval between index test and reference standard? Unclear

Did all patients receive a reference standard?	No
Did patients receive the same reference standard?	Unclear
Were all patients included in the analysis?	No

Could the patient flow have introduced bias? **Risk of bias:** Unclear

ID: **18** Author: Ciacci et al. Year: 2023 Reviewer: **Consensus**

DOMAIN 1: PATIENT SELECTION

A. RISK OF BIAS

Describe methods of patient selection:

In this prospective international study, consecutive patients with suspected coeliac disease in 14 tertiary centres were recruited from January 2018 to December 2020.

Was a consecutive or random sample of patients enrolled?	Yes
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?:	Yes

Could the selection of patients have introduced bias?

Risk of bias: Low

B. APPLICABILITY:

Describe included patients:

Included patients with suspected coeliac disease who underwent serology and duodenal biopsies. A total of 79 patients were excluded due to duplicate recording (n=6), absence or withdrawal of consent (n=6), IgA deficiency (n=3), previous coeliac disease diagnosis (n=2), treatment with a gluten-free diet (n=5), absence of local data on serum tTG-IgA (n=56), or unreadable histological images (n=1)

Do the included patients and setting match the question?

Concerns regarding applicability: Low

DOMAIN 2: INDEX TEST

A. RISK OF BIAS

Describe the index test and how it was conducted and interpreted

Multiple assays used in the different study centres.

Were the index test results interpreted without knowledge of the results of the reference standard? Yes

If a threshold was used, was it pre-specified? Yes

Could the conduct or interpretation of the index test have introduced bias? Risk of bias: Low

B. APPLICABILITY:

Is there concern that the index test, its conduct, or interpretation differ from the review question

Concerns regarding applicability: Low

DOMAIN 3: REFERENCE STANDARD

A. RISK OF BIAS

Describe the reference standard and how it was conducted and interpreted:

Duodenal biopsy. 6 biopsies from D1/D2. Marsh 3 lesions were considered diagnostic for coeliac disease.

Is the reference standard likely to correctly classify the target condition? Yes

Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias?	**Risk of bias:** High
Is there concern that the target condition as defined by the reference standard does not match the review question?	**Concerns regarding applicability:** Low

B. APPLICABILITY:

DOMAIN 4: FLOW AND TIMING

A. RISK OF BIAS

Describe any patients who did not receive the index test(s) and or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Excluded from the study: IgA deficiency, previous coeliac disease diagnosis, treatment with a gluten-restricted diet, a current or previous diagnosis of cancer, absence of local data on serum tTG-IgA, absence or withdrawal of written informed consent for study participation, unreadable duodenal histology, or duodenal villous atrophy with negative serum tTG-IgA test.

Excluded from 2x2: none (as did on both local biopsy results and after central evaluation).

Describe the time interval and any interventions between index test(s) and reference standard:

Unclear.

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive a reference standard?	Yes
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes

Could the patient flow have introduced bias? **Risk of bias:** Unclear

What you need to know

Background and context

The diagnosis of coeliac disease in adults currently involves a two-step process, starting with the detection of tissue transglutaminase (tTG) antibodies and/or serum endomysial antibodies (EMA), followed by a confirmatory endoscopy and duodenal biopsy. Due to the increased accuracy of serological tests, paediatric guidelines adopted a no-biopsy approach, whereby children with IgA-tTG levels ≥10 times the upper limit of normal (ULN) and positive EMA can be diagnosed with coeliac disease without biopsy. However, applying this no-biopsy approach to diagnose adult patients with coeliac disease is highly controversial.

New findings

In a meta-analysis of 18 studies with >12,000 adult participants, we found that IgA-tTG levels ≥10×ULN are highly indicative of coeliac disease in adult patients referred to secondary care with a 100% specificity and a positive predictive value of 98%. The predictive value of the no-biopsy approach varies according to the prevalence of coeliac disease in the studied population.

Limitations

All studies were conducted in secondary and tertiary care settings, and results may not be generalizable to primary care.

Impact

The no-biopsy approach could lead to a shorter time to diagnosis, increased patient satisfaction and reduced healthcare costs.

Lay summary

The diagnosis of coeliac disease involves blood tests, to detect elevated antibodies triggered by eating gluten, and endoscopy with biopsy from the small intestine to prove the damage to the intestinal lining. We found that when the levels of the antibodies are very high, damage to the intestinal lining is almost certain, and endoscopy may not be required in all cases.