Accepted Article

The potential usefulness of human leukocyte antigen typing for celiac disease screening: A systematic review and meta-analysis

Alicia Diaz-Redondo, José Miranda-Bautista, Javier García-Lledó, Javier Perez Gisbert, Luis Menchén

OR 3758 solo inglés

The potential usefulness of human leukocyte antigen typing for celiac disease screening: A systematic review and meta-analysis

Alicia Díaz-Redondo[1], José Miranda-Bautista[2], Javier García-Lledó[2], Javier P. Gisbert[3,4] and Luis Menchén [2,4,5]

Departments of [1]Preventive Medicine and [2]Digestive Diseases. Hospital General Universitario Gregorio Marañón - Instituto de Investigación Sanitaria Gregorio Marañón. Madrid, Spain. [3]Department of Digestive Diseases. Hospital Universitario de La Princesa. Instituto de Investigación Sanitaria Princesa (IP). Madrid, Spain. [4]Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD). Instituto de Salud Carlos III. [5]Department of Medicine. Universidad Complutense. Madrid, Spain

Correspondence: Luis Menchén. Sección de Gastroenterología, Servicio de Aparato Digestivo. Hospital General Universitario Gregorio Marañón. C/ Dr. Esquerdo 46. 28007 Madrid, Spain

e-mail: lamenchen.hgugm@salud.madrid.org

ABSTRACT

Background: The presence of specific Human Leukocyte Antigen-DQ2 and DQ8 seems to be necessary for celiac disease development, but the real contribution of its typing for screening is still uncertain. We aim to conduct a systematic review and meta-analysis of the diagnostic performance of Human Leukocyte Antigen typing tests for celiac disease screening.

Methods: Systematic review of published studies assessing accuracy of human leukocyte antigen DQ2 and DQ8 typing for the detection of celiac disease were selected. MEDLINE and EMBASE were searched from 1st January 2004 until 31st December 2013. Two independent researchers carried out selection and classification of studies, data extraction and analysis. Meta-analysis combining sensitivities, specificities and likelihood ratios of HLA-DQ2 and DQ8 for the diagnosis of celiac disease were carried out.

Results: 6 studies including 1303 individuals were finally evaluated. Pooled sensitivity was 98% (95% confidence interval: 97-99). Overall specificity was 45% (95% confidence interval: 41-48). Regarding specificity, studies were heterogeneous and a subgroup analysis was done according to the type of population included. Overall negative likelihood ratio was 0.05 (0.03-0.09).

Conclusions: Due to its great sensitivity and low negative likelihood ratio, Human Leukocyte Antigen-DQ2/DQ8 typing would be an appropriate test for ruling out celiac disease in the general population suffering related symptoms, and even more in at risk population.

Key words: Celiac disease. HLA. Systematic review. Meta-analysis.

2

INTRODUCTION

Celiac disease (CD) is an immune-based, chronic inflammatory disorder of the small bowel mucosa, which occurs in genetically predisposed individuals precipitated by the ingestion of gluten, the major storage protein of wheat and similar grains (1). Characteristic –although not specific– intestinal villous atrophy leads to the development of malabsorption of nutrients that varies depending of the affected segments. Clinical manifestations are diverse and often so subtle such as iron deficiency anaemia or metabolic bone disease (2); however it should be considered that CD is associated to an increase mortality with respect to healthy population (3), mainly related to a higher risk of development of enteropathy-associated T cell lymphoma in patients with long-standing, undertreated disease. The gluten free diet all lifelong has shown to be the unique and effective treatment for both clinically and histological improvement, in addition of reducing morbidity and mortality (4). Therefore, a prompt and accurate diagnosis is desirable in order to achieve disease remission and to decrease the risk of complications.

In recent years it has been demonstrated that CD is more prevalent that it was originally though, affecting up to 1% of population in western countries (5). Nevertheless, its reported prevalence varies depending, most importantly, on the diagnostic methods and criteria used to estimate it. This fact reflects not only the pitfalls of CD diagnosis nowadays, but also the wide spectrum of clinical manifestations of the disease, previously outlined: from asymptomatic patients diagnosed in a screening program for high-risk individuals, to patients with severe diarrhoea and malabsorption diagnosed in the early childhood (6,7). Therefore, it is important to advance knowledge of new tools to improve diagnosis of this disorder. In this sense, published data suggests that detection of at-risk HLA haplotypes could have a good effectiveness for the CD screening, especially due to its high negative predicted value (8). The pivotal role of certain HLA haplotypes in CD pathogenesis has been well characterized during the last decade, and the presence of specific HLA haplotypes –HLA-DQ2 and DQ8– seems to be necessary, although not sufficient, for the development of CD (9). HLA-DQ2 (mainly encoded by the allelic combination of DQA1*0501/DQB1*0201) and HLA-DQ8 (encoded by the allelic combination of

3

DQA1*0301/DQB1*0302) heterodimers are present, respectively, in 90-95% and 5-10% of the patients affected with CD (10,11). Therefore, CD diagnosis seems to be very unlikely in the absence of such alleles.

There is not a test to make the definitive diagnosis of CD, neither to exclude this illness; although intestinal biopsy is recognized as the gold standard, its results are different depending on the severity of the disease and the ability of the pathologist. Serologic tests have a high specificity and sensitivity but can oscillate in cases with mild intestinal damage and with the gluten intake (12,13). Moreover, serology seems to have lower sensitivity and specificity in adults (14). In 2004, National Institute of Health (NIH) published a systematic review about diagnosis and treatment of this illness (5); conclusions about application of determination of HLA-DQ2 and HLA-DQ8 as screening of CD were not specified. More recently, the American College of Gastroenterology published their clinical guidelines on the diagnosis and management of adulthood CD (15), recommending that HLA testing should not be used routinely in the initial diagnosis of the disorder, but it might be used to effectively rule it out in selected scenarios such as seronegative disease, cases with discrepant serology and histology, or evaluation of patients already on gluten-free diet. The aim of this study is to assess the accuracy of the determination of the presence of HLA-DQ2 and HLA-DQ8 as the first step in diagnosis of CD.

METHODS

Study identification and selection

Bibliographical searches were performed in MEDLINE and EMBASE electronic databases, according to the following search strategy: (hla OR hla-dq antigens OR hla antigens OR hladq OR histocompatibility antigens OR histocompatibility testing OR histocompatibility) OR ((leukocyte OR leukocytes OR leukocyte OR leucocytes) AND (antigen OR antigens)) AND (dq2 OR dq8 OR hladq2 OR hladq8 OR d2 OR d8) OR ((celiac OR celiacs OR celiac OR celiac OR gluten OR glutens OR glutenin OR glutenins OR gliadin OR gliadins OR celiac disease) AND (hla(w)antigen OR hla antigens OR hla-dq antigens)). Limits: English and Spanish; humans; from 2004 January the first to 2013 July the first. Inclusion criteria were: Any article in people both sex, any age and race,

4

and which allowed to construct a two by two table, extracting true positive, false positive, true negative and false negative values of HLA-DQ2 and/or HLA-DQ8 analysis as first step, or simultaneously with serology and regardless of serology result, for the diagnosis of CD from each study. We excluded duplicate articles, letters to the editor, editorials, case series, narrative reviews, systematic reviews and meta-analyses, besides those not related with the object of our study, after title and/or abstract reading. Moreover, studies were not included in case of the absence of intestinal biopsy for all the subjects. According to these criteria, two independent reviewers, reaching a consensus when discrepancies appear, did identification and selection of the studies. The selection process was documented according to PRISMA criteria (16).

Study quality

Study quality was assessed using QUADAS (Quality Assessment of Studies of Diagnostic Accuracy included in Systematic Reviews) checklist (17). It has 14 questions –each of them answered as "yes", "no" or "unclear"– evaluated separately, without summary score because of its difficult interpretation and moreover, because it would mean the loosing of information from each of them separately. Items try to show if patients of the study are representative of those to receive the test in practice, the accuracy and consistency of reference standard, the appropriateness of time between standard and the evaluated test, the interpretation of results, the explanation of withdrawals, and the synthesis and analysis of data in the selected studies. Both reviewers, reaching a consensus when discrepancies appeared, also made this assessment.

Data extraction

After a critical reading, according to the inclusion criteria, the following variables were extracted: Author, publication year, publication country, population characteristics, type of the study, prevalence (pre-test probability), recommendations about applications of the HLA tests; true positive, false positive, true negative and false negative values (input variables) to calculate our output variables. We contacted with the authors to get full data in case it needed. The same two independent reviewers, reaching a consensus when discrepancies appeared, also did data abstraction. For

5

further relevant studies we checked the reference list of identified trials.

Data synthesis

We calculated the sensitivity, specificity, positive and negative likelihood ratios (LRs) (output variables), and their corresponding 95% confidence intervals (95% CIs) for each study. LRs express how many times more likely a given test result would be expected in a patient with the disease, than in another without the target disease. LRs were used to estimate the post-test probability for the disease, so that can help in selecting the clinical usefulness of this diagnostic test. We can take positive LRs > 10 and negative LRs < 0.1 as providing convincing diagnostic evidence, while values > 5 and < 0.2 respectively as supporting moderate diagnostic evidence (18). To calculate LRs, if the event of one of the cells of the cross table contained a zero value; 0.5 points were added to all the cells. The heterogeneity of all indexes was statistically evaluated through a homogeneity test based on the χ^2 test (p value ≥ 0.1). The I^2 statistic was used to estimate the impact of heterogeneity on the results. In case of heterogeneity, subanalysis was conducted, depending on the population included in the study and its pre-test probability of suffering from CD. Sensitivities, specificities, and LRs of the individual studies were combined to assess their pooled indexes performing their corresponding meta-analyses, using a random effects model (DerSimonian and Laird). The analyses were carried out using the statistical software Meta-DiSc (version 1.4).

RESULTS

Study identification and selection

Based on described search strategy, 967 articles were initially identified. Figure 1 shows the flowchart with the results of the search and how studies were excluded till get the 6 studies included in the final analysis. Briefly, articles were screened at different steps, starting with the most sensitive strategy, and progressively increasing specificity in order to minimize relevant looses: First, and after screening studies following search strategy and limits, we excluded duplicate articles, letters to the editor, editorials, case series, narrative reviews, systematic reviews and meta-analyses, besides those not related with the object of our study, after title and/or abstract

6

reading (n = 853). Secondly, and after full-text review, including references, other studies were excluded: Those without information about HLA analysis in CD, or that studied other genes (n = 31). Finally, potentially appropriated studies were not included because the absence of intestinal biopsy for all the subjects, the use of HLA DQ not as first step for CD diagnosis, or because they did not allowed us to calculate our output variables (n = 77).

Data extraction and description of the included studies

In the final selection, a total of 1303 individuals were included in those selected articles, which used the typing of HLA-DQ2 or HLA-DQ8 haplotypes for the evaluation of CD as first step or regardless serology results, and containing information for calculating output variables (8,19-23). Table I summarizes these 6 cross-sectional studies; all of them were focused on European populations, being three of them carried out in Spain.

Qualitative analysis of included studies

Methodological quality of the included studies was evaluated through the QUADAS scale (Table II). The articles, all of them high quality cross-sectional studies, correspond to III evidence level and to B recommendation rank according to the Oxford Centre for Evidence-Based Medicine.

Quantitative analysis of output variables

In table III, we show the output variables that were obtained as a result of the data extraction from the included studies in this review. All of them had a very high negative predictive value, while positive predictive values were low and variable among them. Figure 2A reflects the high sensitivity (near 100%) that the test showed in all the articles. Specificity analysis revealed significant heterogeneity ($p < 0.001$; I^2 95.0%) (Fig. 2B); this fact led us to perform a subgroup analysis (see below). Regarding to negative LR, its values were homogenous (Fig. 2C), and very near to 0 in all the included studies. Its mean value of 0.05 (95% CI: 0.03-0.09) indicates that the probability of a negative result in a person without CD is 20 times higher than in a

7

patient with the disease (Fig. 2C).

Subanalysis: Specificity and positive LR analysis

Taking into account the heterogeneity of the studies about their specificity and positive LR, a subanalysis was performed according the population included in each study and the estimated prevalence of CD in it. The studies of Hadithi et al. (8) and Fernández-Bañares et al. (21) (subgroup 1) were grouped because their main population was symptomatic people (pre-test probability 1/56); on the other hand, Kapitany et al. (19), Karinen et al. (20), Santaolalla et al. (22) and Klapp et al. (23) (subgroup 2) analysed already diagnosed CD patients, first degree relatives or population at risk for CD (pre-test probability higher than 1/10). Subgroup specificity is shown in figure 3: subgroup 1 had a global specificity of 0.56 (95% CI: 0.52-0.61), and it was homogeneous (p = 0.33; I^2 0); subgroup 2 was heterogeneus (p < 0.001; I^2 91.0) about the specificity (0.29; 95% CI: 0.25-0.34). Regarding to positive LR (Fig. 4), it was 2.17 (95% CI: 1.84-2.55) for subgroup 1, indicating that the probability of a positive result corresponds to a person with CD is more than twice higher than to a person without the disease. Subgroup 2 was heterogeneous for the positive LR test (p = 0.001, I^2 81.6%).

DISCUSSION

Diagnosis of CD, especially in adult patients, still represents a challenge for clinicians. In adults, serology accuracy seems to be significantly lower than in children, a fact that may be related to the higher proportion of cases with lesser degrees of villous atrophy (13,24). The low sensitivity of serologic testing in patients with atypical clinical picture and less histologically severe forms of the disease makes mandatory the search of new diagnostic tools. The results reported herein suggest that HLA-DQ2/DQ8 typing would be an appropriate test for ruling out CD, both in the general population and in high-risk groups, due to its excellent sensitivity (near 100%) and low negative LR, with a global value of 0.05; this indicates that the probability of a negative result corresponds to a individual without CD is 20 times higher than to an affected patient.

It is already known that the presence of specific HLA haplotypes constitutes the

8

strongest genetic determinant of CD risk. It has been estimated that HLA-DQ2 and HLA-DQ8 heterodimers are present, respectively, in 90-95% and 5-10% of the patients affected with CD (10,11), so it is tempted to hypothesize that CD diagnosis is very unlikely in the absence of such alleles. On the other hand, and as previously outlined, sensitivity of serologic tests may be significantly lower in those cases with mild intestinal damage, in adult population, and in patients under gluten-free diet (12-14). In this sense, one of the strengths of the HLA-DQ2/DQ8 analysis lies on the absence of the bias treatment paradox, as its determination does not change before or after gluten-free diet; moreover, it is not an observer-dependent technique. In addition, the result of the test does not vary during lifetime, neither with disease severity. The HLA-DQ2/DQ8 determination has not been used in routine clinical practice due to its low specificity, high cost and low accessibility; however, as it is increasingly accessible, it becomes necessary to specifically evaluate its potential use in the clinical setting.

From the 967 selected articles in this systematic review, only six of the 83 references that were classified as potentially included, were finally included in the meta-analysis. Most of the excluded studies, including some case series, in this review had a high methodological quality, but did not accomplish the inclusion criteria and, consequently, they were not useful for the aim of our study; most of them did not use the index test (HLA determination) as the first step in the screening of CD. Other reasons to exclude studies were not using biopsy as the gold standard or biopsy only in those cases with a positive serologic test, making it impossible to obtain the data for calculate our output variables.

The 6 studies included in the present meta-analysis comprise more than 1,300 individuals. All of them were cross-sectional studies carried out in Europe (CD predominantly occurs in people of European origin) and published between 2006 and 2013. Two studies included paediatric patients (19,23), whereas the remaining four evaluated mainly adult population (8,20-22). Female population was slightly predominant in all except in the Santaolalla et al. study (22); the objective of this last study was not directly related with our endpoint, but it reported the variables needed for our study and it confirmed the high sensitivity and high negative predictive value of HLA testing. The other included studies agreed in recommending HLA typing to rule

out CD in their study populations. As all of the aforementioned studies are European, we can assume an adequate validity when extrapolating these results in Europe and countries with predominantly European ancestors. In this sense, those studies showing higher frequencies of HLA-DQ2 and DQ8 negative CD patients (up to 9%) have been performed in Latin-American populations (25,26). The absence of HLA-DQ2 and DQ8 should be, therefore, carefully interpreted in non-Caucasian populations in terms of CD screening.

The present meta-analysis –including studies that agree in the high negative predictive value of the HLA-DQ2 and HLA-DQ8 analysis– demonstrates a very low negative LR of this determination; it means that an individual with a negative result in HLA-DQ2 or DQ8 analysis is unlikely to present the disease. Values of included studies have demonstrated to be homogenous, and close to 0 in all the included studies, which holds up a convincing diagnostic evidence. As previously mentioned, its global value of 0.05 indicates that the probability of a negative result corresponds to an individual without CD is 20 times higher than to a person with the disease.

On the other hand, a high variability exists in the positive LR values and specificity data; this fact seems to be related to the heterogeneity of the analysed population in each study. Therefore, we performed a subgroup analysis in order to minimize the patient selection bias frequently found in systematic reviews. The subgroups were formed as follows: The first one included Hadithi et al. (8) and Fernández-Bañares et al. (21) studies, that analysed symptomatic patients –even though with non-specific symptoms, such as chronic watery diarrhoea, anaemia, weight loss or abdominal pain–. The results of the meta-analysis of these studies could be extrapolated to population with similar symptoms, and would benefit of the high sensitivity of the test with a good but not optimum positive LR of 2.17 –being more than twice as probable that a positive result in the test will correspond to a patient, anyway– and a better global negative LR of 0.04. Population analysed in the four studies which comprised the subgroup 2 is supposed to have higher probability of CD; Kapitany et al. (19) and Klapp et al. (23) analysed previous diagnosed CD patients, while Karinen et al. (20) and Santaolalla et al. (22) focus their studies in risky population, such as first-degree relatives of patients with CD and type I diabetes mellitus, which are known to share

10

genetic load. This subgroup resulted also heterogeneous with respect to the specificity; moreover, the results for the positive LR of the second subgroup remained heterogeneous. In this context, we must be very careful when interpreting their global results about specificity and positive LR.

Although additional studies are needed to dissipate the doubts concerning the natural history of the disease –specially in adult patients– and therefore the real importance of the available diagnostic tools in CD, we can conclude that, due to its great sensitivity and low negative LR, HLA-DQ2/DQ8 determination would be an appropriate test for ruling out CD in the general population suffering related symptoms, and even more in at risk population.

REFERENCES

1. Kagnoff MF. Overview and pathogenesis of celiac disease. Gastroenterology 2005;128:S10-8.

2. Green PH, Cellier C. Celiac disease. N Engl J Med 2007;357:1731-43.

3. Tio M, Cox MR, Eslick GD. Meta-analysis: Coeliac disease and the risk of all-cause mortality, any malignancy and lymphoid malignancy. Aliment Pharmacol Ther 2012;35:540-51.

4. Case S. The gluten-free diet: How to provide effective education and resources. Gastroenterology 2005;128:S128-34.

5. National Institutes of Health Consensus Development Conference Statement on Celiac Disease, June 28-30, 2004. Gastroenterology 2005;128:S1-9.

6. Fasano A. Celiac disease -how to handle a clinical chameleon. N Engl J Med 2003;348:2568-70.

7. Rampertab SD, Pooran N, Brar P, et al. Trends in the presentation of celiac disease. Am J Med 2006;119:355 e9-14.

8. Hadithi M, von Blomberg BM, Crusius JB, et al. Accuracy of serologic tests and HLA-DQ typing for diagnosing celiac disease. Ann Intern Med 2007;147:294-302.

9. Kagnoff MF. Celiac disease: Pathogenesis of a model immunogenetic

disease.   J   Clin   Invest   2007;117:41-9.

10.   Bevan S, Popat S, Braegger CP, et al. Contribution of the MHC region to the familial risk of coeliac disease. J Med Genet 1999;36:687-90.

11.   Liu E, Rewers M, Eisenbarth GS. Genetic testing: Who should do the testing and what is the role of genetic testing in the setting of celiac disease? Gastroenterology   2005;128:S33-7.

12.   Rostami K, Kerckhaert J, Tiemessen R, et al. Sensitivity of antiendomysium and antigliadin antibodies in untreated celiac disease: Disappointing in clinical practice. Am J Gastroenterol 1999;94:888-94.

13.   Abrams JA, Diamond B, Rotterdam H, et al. Seronegative celiac disease: Increased prevalence with lesser degrees of villous atrophy. Dig Dis Sci 2004;49:546-50.

14.   Hill ID. What are the sensitivity and specificity of serologic tests for celiac disease? Do sensitivity and specificity vary in different populations? Gastroenterology   2005;128:S25-32.

15.   Rubio-Tapia A, Hill ID, Kelly CP, et al. ACG clinical guidelines: Diagnosis and management of celiac disease. Am J Gastroenterol 2013;108:656-76; quiz 677.

16.   Urrutia G, Bonfill X. Declaracion PRISMA: una propuesta para mejorar la publicacion de revisiones sistematicas y metaanalisis. Med Clin (Barc) 2010;135:507-11.

17.   Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic   reviews.   BMC   Med   Res   Methodol   2003;3:25.

18.   Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine   Working   Group.   JAMA   1994;271:703-7.

19.   Kapitany A, Toth L, Tumpek J, et al. Diagnostic significance of HLA-DQ typing in patients with previous coeliac disease diagnosis based on histology alone.   Aliment   Pharmacol   Ther   2006;24:1395-402.

12

20.   Karinen H, Karkkainen P, Pihlajamaki J, et al. HLA genotyping is useful in the evaluation of the risk for coeliac disease in the 1st-degree relatives of patients with coeliac disease. Scand J Gastroenterol 2006;41:1299-304.

21.   Fernandez-Banares F, Esteve M, Salas A, et al. Systematic evaluation of the causes of chronic watery diarrhea with functional characteristics. Am J Gastroenterol 2007;102:2520-8.

22.   Santaolalla R, Fernandez-Banares F, Rodriguez R, et al. Diagnostic value of duodenal antitissue transglutaminase antibodies in gluten-sensitive enteropathy. Aliment Pharmacol Ther 2008;27:820-9.

23.   Klapp G, Masip E, Bolonio M, et al. Celiac disease: The new proposed ESPGHAN diagnostic criteria do work well in a selected population. J Pediatr Gastroenterol Nutr 2013;56:251-6.

24.   Tursi A, Brandimarte G, Giorgetti G, et al. Low prevalence of antigliadin and anti-endomysium antibodies in subclinical/silent celiac disease. Am J Gastroenterol 2001;96:1507-10.

25.   Parada A, Araya M, Perez-Bravo F, et al. Amerindian mtDNA haplogroups and celiac disease risk HLA haplotypes in mixed-blood Latin American patients. J Pediatr Gastroenterol Nutr 2011;53:429-34.

26.   Kotze LM, Nisihara R, Utiyama SR, et al. Absence of HLA-DQ2 and HLA-DQ8 does not exclude celiac disease in Brazilian patients. Rev Esp Enferm Dig 2014;106:561-2.

13

Fig. 1. Study selection.

Fig. 2. Global sensitivity, specificity and negative likelihood ratio.

Fig. 3. Subgroup specificity.

Fig. 4. Subgroup positive likelihood ratio.

Table I. Summary of included studies

Table II. Results of QUADAS scale

Table III. Results of data extraction of included studies