Using Patients Like My Patient for Clinical Decision Support: Institution-Specific Probability of Celiac Disease Diagnosis Using Simplified Near-Neighbor Classification

Brian H. Shirts, MD, PhD^{1,2}, Sterling T. Bennett, MD, MS^{1,3}, and Brian R. Jackson, MD, MS^{1,4}

¹Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT, USA; ²Department of Laboratory Medicine, University of Washington, Seattle, WA, USA; ³Department of Pathology, Intermountain Medical Center, Murray, UT, USA; ⁴ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT, USA.

BACKGROUND: Interpretation of a diagnostic test result requires knowing what proportion of patients with a “similar” result has the condition in question. This information is often not readily available from the medical literature, or may be based on different clinical populations that make it nonapplicable. In certain settings, where correlated screening parameters and diagnostic data are available in electronic medical records, a representation of diagnostic test performance on “patients like my patient” can be obtained.

OBJECTIVE: We sought to integrate patient demographic and physician practice information using a simplified nearest neighbor algorithm. We used this method to illustrate the relationship between tTG IgA test result and duodenal biopsy for celiac disease in a local diagnostic context.

PARTICIPANTS: We used a data set of 1,461 paired tissue transglutaminase (tTG) IgA and definitive duodenal biopsy results from Intermountain Healthcare with data on patient age and ordering physician specialty. This was split into a discovery set of 1,000 and a validation set of 461 paired results.

MAIN MEASURES: Accuracy of the local discovery data set in predicting probability of positive duodenal biopsy and confidence intervals around predicted probability in the test data compared to probabilities of positive biopsy implied from published logistic regression and from published sensitivity and specificity studies.

KEY RESULTS: The near-neighbor method could estimate probability of clinical outcomes with predictive performance equivalent to other methods while adjusting probability estimates and confidence intervals to fit specific clinical situations.

CONCLUSIONS: Data from clinical encounters obtained from electronic medical records can yield prediction estimates that are tailored to the individ-

ual patient, local population, and healthcare delivery processes. Local analysis of diagnostic probability may be more clinically meaningful than probabilities inferred from published studies. This local utility may come at the expense of external validity and generalizability.

KEY WORDS: personalized medicine; multifactorial analysis; gluten sensitive enteropathy; laboratory test information content; evidence based medicine; nearest neighbor.

J Gen Intern Med 28(12):1565–72

DOI: 10.1007/s11606-013-2443-z

© Society of General Internal Medicine 2013

INTRODUCTION

Interpretation of a diagnostic test result requires knowing what proportion of patients with similar results and pretest probabilities have the diagnosis in question. In principle, such information may be derived from test performance descriptions in the medical literature, but diagnostic likelihood ratios are often not readily available or may be based on different clinical populations, rendering them non-applicable. A summary of diagnostic test performance on “patients like my patient,” based on data available from local electronic medical records (EMRs), might be useful as decision support to both physicians and their patients in many situations.

Our hypothesis was that combining local demographic and physician practice data from EMRs with laboratory diagnostic information would yield post-test probabilities that are accurate and, in some cases, give more pertinent, patient-specific information than estimates available from literature. In addition, most multifactorial analyses are not intuitive to clinicians or patients, limiting their expediency. We identified nearest-neighbor analysis as a method that could facilitate multifactorial integration of diagnostics with clinical outcomes data. We describe a simplified near-neighbor analysis that yields probabilities intuitively understood by clinicians and patients and can be reasonably implemented by hospitals and laboratories with generally available statistics software.

Electronic supplementary material The online version of this article (doi:10.1007/s11606-013-2443-z) contains supplementary material, which is available to authorized users.

Received October 29, 2012

Revised March 13, 2013

Accepted March 20, 2013

Published online May 4, 2013

We present the example of celiac disease to illustrate the potential clinical utility near-neighbor analysis. Celiac disease is an example of a condition that is initially evaluated in the primary care setting, may be difficult to diagnose, affects a large number of patients, and can be treated effectively by generalists. Because treatment requires substantial lifestyle modifications, diagnostic accuracy is crucial. Although many general internists and pediatricians feel comfortable diagnosing uncomplicated cases, knowing when to refer patients to specialists for invasive diagnostic procedures is important. Near-neighbor analysis can be used as a tool to gather information about the performance of diagnostics in the hands of both experts and generalists to illustrate where additional value is gained with expert opinion.

The diagnosis of celiac disease incorporates clinical information with results from imperfect laboratory tests.^{1,2} Tissue transglutaminase IgA (tTG IgA) testing has become the favored initial test, often followed by duodenal biopsy to provide final confirmation of a presumptive diagnosis based on improvement of symptoms with diet modification.^{3–5} However, multiple diagnostic algorithms exist, many immunoassays are available for screening, correlation between pathologists in diagnostic biopsies is not perfect, and recent studies have questioned the need for biopsy before diet modification, particularly in pediatric populations with high tTG IgA levels.^{6–10} Not surprisingly, physician adherence to clinical guidelines is not uniform.³ Shared physician–patient decision-making may be important, particularly in light of varying individual thresholds for dietary modification.^{11,12} We used deidentified data from Intermountain Healthcare in a pilot study to evaluate whether near-neighbor analysis might improve decision support presented to clinicians along with tTG IgA test results.

METHODS

Development of Near-Neighbor Analysis Method Using Local Data

We sought a method that could report a likelihood ratio or the probability of a presumed diagnosis being true based on multiple clinical factors. Nearest-neighbor algorithms are one method to categorize cases using complex information by plotting a case in a graph where each factor is a separate dimension, and looking at the distance from the point defined by the case with other, already categorized, points in the factor space.^{13,14} For simplicity of illustration and calculation, we chose to use boxes or “hypercubes” centered on the patient’s clinical characteristics and laboratory values. Similar methods using hypercubes have been described by others.^{13,15} Briefly, a set of individuals with known diagnostic outcomes defines outcome distributions in a

multi-dimensional test space, with each diagnostic measure defining a separate dimension. Each individual’s values define a point in this multidimensional space. The area in close proximity to an unknown patient’s diagnostic point (near-neighbor space) contains previously identified cases (those with the disease of interest) and controls (those without the disease of interest) that are most similar to the test patient for the diagnostic measures (see Fig. 1). We used counts of cases and controls in the near-neighbor space to calculate the binomial probability and confidence intervals that the unknown patient is a case.

We provide our function in R statistical programming language (<http://www.r-project.org/>) to facilitate implementation of similar near-neighbor analysis by others (Supplemental appendix available online).

Study Sample

We retrospectively identified 3,207 individuals who had received any tTG IgA testing and a duodenal biopsy at one of the 17 hospitals or over 70 clinics associated with Intermountain Healthcare (Salt Lake City, UT) between Jan 2008 and Oct 2011. All tTG IgA testing

Figure 1. Illustration of the near-neighbor method for two-dimensional analysis to predict probability of duodenal biopsy using only tTG IgA and patient age in a discovery sample of 1,000 Intermountain Healthcare patients. Natural log-transformed tTG IgA values are plotted by age. The orange box corresponds to the near-neighbor space surrounding a 17-year-old patient with a tTG IgA of 3 units. Of the 161 points inside the orange box (some of which are plotted on top of others), six represent positive biopsies, indicating a 3.9 % chance this patient will have a positive biopsy (95 % CI [1.4 % to 8.2 %]). The green box corresponds to the near-neighbor space surrounding a 58-year-old patient with a tTG IgA of 55 units. Of the 16 points in this box, ten represent positive biopsies indicating a 63 % chance (95 % CI [35 % to 85 %]) this patient will have a positive biopsy. (Note that tTG IgA values of 0 were transformed to -1 , since $\ln(0)$ is undefined).

was performed at ARUP laboratories (Salt Lake City, UT) using the INOVA QUANTA Lite kits (INOVA Diagnostics, San Diego, CA) according to the manufacturer's specifications. A total of 1,681 tTG IgA results were paired with biopsy results more than 30 days distant and 1,526 tTG IgA/biopsy pairs were performed within 30 days of each other. In the 1,526 chronologically close pairs, 1,461 biopsy results could be classified as positive or negative for celiac disease and 65 were indeterminate or equivocal. Data on patient age was available for all 1,461 individuals. Ordering clinician specialty was available for the majority and "other" was coded for 60 patients where ordering clinician specialty was undefined. We subdivided these 1,461 tTG IgA/biopsy pairs into a discovery sample of 1,000 individuals and a validation sample of 461 individuals, with chronologically earlier pairs in the discovery sample as would happen if the algorithm was applied to clinical care. The study was approved by Intermountain Healthcare IRB (RMS#1024175).

Comparison of Local Data with Data from Meta-Analysis of tTG IgA Performance

In order to determine if test performance metrics derived from published literature accurately described local test performance, we calculated sensitivity, specificity, and positive and negative predictive values for our local tTG IgA performance and compared these values to the sensitivity and specificity reported in a recent meta-analysis.¹⁶

Determination of Potential Near-Neighbor Diagnostic Factors

There is no simple framework to determine optimal dimensionality for nearest-neighbor analysis, so for many clinical scenarios, some judgment may be necessary to identify potentially important local variables. Added dimensions decrease the number of near-neighbor observations for any patient increasing the width of confidence intervals around probability estimates (see Fig. 2).

We considered age and tTG IgA as continuous variables and sex, specialty of ordering physician, and setting (inpatient/outpatient) as categorical variables that might predict biopsy results. Although alternative methods exist to select variables, we used univariate logistic regression to evaluate each factor individually, with duodenal biopsy result as a binary dependent variable and a cutoff of $p < 0.05$ for nominal evidence that the variable might be useful in near-neighbor analysis. We used natural log transformations of tTG IgA, with $\ln(0)$ coded as -1 , which produced a distribution that was more uniform than untransformed tTG IgA (see Fig. 1).

Figure 2. Illustration of the effect of the size of the near-neighbor space on the probability of a positive biopsy as a function of tTG IgA. For large, middle, and small near-neighbor spaces, tTG IgA space length was 0.77, 0.53, and 0.26 $\ln(\text{tTG IgA})$ units, respectively. Near-neighbor analysis using very large and very small cubes is shown as *dotted black* and *dotted gray* lines, respectively. In near-neighbor analysis, smaller cube lengths created less smooth probability estimate curves and larger cube lengths decreased the predictive power at high tTG IgA levels. This figure illustrates that box size for clinically meaningful probabilities and adequate smoothness can be determined empirically.

Based on previous observations of simulated near-neighbor samples, we examined physician specialty looking for cohesive specialty groups with at least 200 observations in the discovery set. We found two sufficiently large physician groups, gastroenterologists ($n=417$) and pediatricians (including pediatric subspecialists, $n=270$). The remaining observations ($n=313$), predominantly family medicine ($n=111$) and general internists ($n=79$), were grouped with those without physician specialty recorded (48 in discovery set). A larger sample may have facilitated addition of other variables or further subdivision of physician specialty.

Estimating Clinically Meaningful Near-Neighbor Space

In near-neighbor analysis, smaller cube sides create less smooth probability estimate curves and large confidence intervals around estimates (Fig. 2). Conversely, larger cube sides produce less tailored probability estimates, but with smaller confidence intervals. At extremes, the largest cube size captures the entire data set and yields a maximally accurate estimate of the positivity rate for the entire population, and the smallest cube size captures only patients exactly like the patient in question, potentially producing inaccurate estimates for rare situations and no estimate for novel situations (see Fig. 2).

A clinically useful cube size will include enough patients to produce probability estimates that balance accuracy with precision. We found that optimal balance for the data available may vary by clinical situation and physician preference.

In our sample, increasing cube size decreased the predictive power, particularly at high tTG IgA levels. We experimented with varying cube side lengths using the discovery sample to identify a cube length large enough to produce a smooth curve without sacrificing predictive power (middle line, Fig. 2).

Comparison of Local Near-Neighbor Probability with Estimates from Published Sources

We compared probability estimates from near-neighbor analysis with probability estimates from two published sources: 1) the probability inferred from tTG IgA sensitivity and specificity,¹⁶ which is most commonly used by physicians; and 2) probability from a published logistic regression using tTG IgA to predict biopsy results with similar criteria to our study.¹⁷ We compared probability estimates by plotting these as a function of tTG IgA, the common variable in all methods. We then compared relative accuracy of near-neighbor analysis by using predicted probabilities as risk scores in ROC analysis performed with R ('colAUC function'). We also illustrated differences in probability predictions using Bland-Altman plots.

RESULTS

Comparison of Local Data with Data from Meta-Analysis of tTG IgA Performance

Sensitivity and specificity of tTG IgA at Intermountain Healthcare using the manufacturer's cutoff of 20 units were 0.63 (95 % CI [0.55–0.71]) and 0.93 (95 % CI [0.90–0.96]), respectively. Positive and negative predictive values were 0.78 (95 % CI [0.69–0.85]) and 0.87 (95 % CI [0.83–0.90]), respectively (see Table 1). From our 95 % confidence intervals, we infer these were significantly different than sensitivity (0.89) and specificity (0.98) reported in recent meta-analysis.¹⁶

Table 1. Contingency Table for Validation Data Set

Actual	Positive tTG IgA test	Negative tTG IgA test	Total
Positive biopsy	79	46	125
Negative biopsy	22	314	336
Total	101	360	461

Selection of Important Local Diagnostic Factors and Hypercube Side Length for Near-Neighbor Analysis

Of the five variables considered, tTG IgA, age, and physician specialty were identified as factors to potentially include in near-neighbor analysis ($p < 0.05$ in individual logistic regression). Gastroenterologists' patients appeared more likely to have positive biopsies regardless of tTG IgA and age, suggesting physician specialty may be important. It was apparent that pediatric specialty may be a surrogate for age, which would lead to a concern of confounding in many statistical analyses; however, in near-neighbor analysis this is not a concern so long as there are sufficient data to produce adequately precise probability estimates in all dimensions. If two or more variables create similar groups, one variable redundantly displays similar clustering seen in other dimensions, but this will not interfere with neighbor counts, as neighbors would stay close in all dimensions. Neither sex nor setting (inpatient vs. outpatient) was a significant predictor of biopsy results in initial analysis.

We chose a cube side length equivalent to drawing a box which goes ± 10 years from the patient's age and ± 0.54 from the patient's $\ln(\text{tTG IgA})$ result (illustrated by the orange and green boxes drawn in Fig. 1). This cube captures on average about 1/25th of the patients in the discovery sample for each validation patient (see Fig. 1 and middle line, Fig. 2).

Probability Estimates Using Near-Neighbor Analysis

With the discovery sample of 1,000 individuals, the near-neighbor analysis incorporating tTG IgA, age, and ordering physician specialty was able to predict the probability of a positive biopsy for the 461 individuals in the validation set with a median probability confidence interval width of 27 %.

Comparison of Local Near-Neighbor Probability with Estimates from Published Sources

Visual examination of scatter plots of near-neighbor probability compared with other estimates revealed several situations where age and physician specialty might modulate the probability of a positive biopsy at Intermountain Healthcare (Fig. 3). It is clear that estimates derived from sensitivity and specificity based on absolute tTG IgA cutoff don't facilitate the intermediate probabilities predicted by near-neighbor and logistic regression methods. In addition, there is a cluster of patients with relatively high probability of positive biopsy in near-neighbor analysis with low to moderate tTG IgA consisting of older patients, particularly those over 65 not seen by gastroenterologists. The patients with the lowest and highest probabilities of positive biopsy were very young patients, particularly those seen by

Figure 3. Comparison of probability estimates of positive duodenal biopsy from near-neighbor analysis with those derived from published sensitivity and specificity and from published logistic regression. **Black points**—near-neighbor probability estimates for 461 individuals in the validation data set using tTG IgA and age, with near-neighbor estimates derived from the Intermountain Healthcare discovery data set. **Red lines**—probability estimates inferred from sensitivity and specificity reported by meta-analysis of tTG IgA studies.¹⁶ **Blue line**—probability of positive biopsy using published logistic regression. **Orange vertical lines** show the tTG IgA manufacturer's cutoffs for “weak positive” and “moderate to strong positive.”

pediatricians and generalists (‘other’ specialty). ROC analysis showed that near-neighbor and logistic regression predictions had similar area under the curve (AUC=0.81 and AUC=0.82, respectively), with estimates produced using sensitivity and specificity performing slightly worse (AUC=0.78) (Fig. 4). Bland–Altman plots illustrate that despite similar ROC curves, there can be large differences in individual probability estimates generated by different methods (Fig. 5). Probability estimates derived from sensitivity and specificity incorrectly classify indeterminate results as either high or low probability (Fig. 5a), and estimates based on the published logistic regression appear to overestimate probability of positive biopsy for most patients (Fig. 5b).

DISCUSSION

We have shown that near-neighbor analysis can yield accurate probability estimates and can illustrate pertinent patient-specific and institution-specific information. In the example described, near-neighbor-derived probability estimates for positive duodenal biopsy based on institution-specific data were a better reflection of local clinical practice than tTG IgA performance metrics in published studies. Some differences between local and general

Figure 4. ROC analysis comparing relative accuracy of near-neighbor analysis by using predicted probabilities as risk scores.

probability estimates may be because published studies do not reflect uniform medical practice.¹⁸ For example, studies of celiac disease often exclude individuals already on a gluten-free diet, yet some patients begin diet modification before diagnostic testing or between screening and confirmatory tests.¹⁶ In addition, diagnostic studies rely on multiple duodenal biopsies, as villous atrophy can be patchy; yet clinical practice regarding the number of biopsies is not uniform.^{19,20} Consequently, local differences in practice may explain discrepancies between near-neighbor probability estimates and those derived from published studies. Furthermore, near-neighbor analysis reports continuous likelihood where most published studies use only binary cutoffs, so the greatest discrepancies are often near cutoff values. Whatever the sources of deviation, local estimates are intuitively more applicable to individual physicians and patients. A corollary of this observation is that the specific results generated by our analysis may not be generalizable to other institutions. This is not a limitation of the nearest-neighbor approach, but rather, a strength, consistent with our hypothesis that local analysis may be necessary to provide accurate estimates of the true information content of diagnostic testing.

The practice of using local information to improve diagnostic decision support is not a new idea,^{21,22} but has yet to be adopted widely. We have described simplified near-neighbor analysis as one method of multifactorial analysis that is intuitive and relatively easy to interpret. Although other methods of analysis could be used on local data, we found simplified near-neighbor analysis was exemplary for relatively straightforward health-system-specific analysis because it is clinically intuitive and can be implemented by laboratorians or information technology staff using data that already exist in

Figure 5. Bland-Altman analyses. a Near-neighbor probabilities are compared to binary probability estimates derived from reported sensitivity and specificity.¹⁶ b Near-neighbor probabilities are compared to estimates based on published logistic regression. Logistic regression appears to overestimate probability of positive biopsy for most patients.¹⁷

EMRs. Furthermore, near-neighbor analysis facilitates diagnostic decision support that captures idiosyncrasies of the local diagnostic environment. Decision support with these features would be particularly valuable for conditions are initially assessed and potentially managed in the primary care setting, yet may be diagnostically challenging and are frequently referred for specialist evaluation. We have tested near-neighbor analysis on

simulated data sets of prostate cancer diagnosis with favorable results (data not shown). Other potential targets include disorders of thyroid function, autoimmune diseases, thrombophilia risk, anemia workup, and many others.

An important limitation of near-neighbor analysis is that it requires relatively large data sets for precise probability estimates, particularly when many variables are considered; however, EMRs are designed to gather such data. Another limitation is that near-neighbor analysis may generate wide confidence intervals for some probability estimates, particularly at distribution extremes and with small data sets. Although wide confidence intervals may be a concern, there are frequently situations where pertinent published data are simply not available, and nearest-neighbor analysis may fill in the gaps. In these situations, the drawback of the diagnostic reality of wide confidence intervals pales in comparison to the false certainty implied by artificially high sensitivity and specificity derived from studies with very narrow, clinically unrealistic scope. We believe that the near-neighbor approach encourages appropriate physician flexibility to consider unmodeled mitigating factors, when appropriate. Similarly, for our analysis, we ignored indeterminate and equivocal biopsy results in order to facilitate binomial comparison with published literature, but nearest-neighbor analysis can theoretically accommodate these data.

Information that encapsulates pre-test and post-test probabilities may be necessary for optimal shared decision-making by giving physicians and patients a view of other “patients like me.” This concept is illustrated in the Text Box 1 with three examples from our data set. Local analysis could also lead to the identification of opportunities for clinical system improvement and quality control, such as a need for systematic duodenal biopsy performance to improve sensitivity.⁶ Data aggregation could be extended to inform physicians about local practice patterns by creating decision-support statements such as, “General internists referred patients for gastroenterology evaluation in 5 % of patients in the same age group with similar tTG IgA values,” or “For patients in the same age group with similar tTG IgA values referred for gastroenterology evaluation, biopsy was obtained in 70 % with a 94 % positivity rate.” Ideally, the availability of accurate information on local diagnostic test performance will lead to reduction in redundant or invasive procedures. For celiac disease, these include situations where high tTG IgA levels may justify preemptive initiation of dietary modifications without biopsy.^{11,12} Local near-neighbor analysis may be useful in many management situations where there is a high degree of practice variability between institutions and individual physicians. For example, near-neighbor analysis could aid in summarizing the variable utility of pre-

operative cardiac evaluations in relation to clinical outcomes.

Text Box 1.

- The near-neighbor space surrounding a 14-year-old patient with a tTG IgA of 4 units and seen by a pediatrician encompasses 33 similar patients, one of whom had a positive biopsy. An example test report that illustrates this information would state:
 - tTG IgA – negative.
 - Based on our experience with similar patients there is a 3% chance this patient will have a positive biopsy (CI [0% to 16%]).
- The near-neighbor space surrounding a 63-year-old patient with a tTG IgA of 11 units and seen by a gastroenterologist encompasses 16 similar patients, 5 of whom had a positive biopsy. An example test report would state:
 - tTG IgA – negative.
 - Based on our experience with similar patients there is a 31% chance this patient will have a positive duodenal biopsy (CI [11% to 59%]).
- The near-neighbor space surrounding a 23-year-old patient with a tTG IgA of 130 units and seen by a gastroenterologist encompasses 13 similar patients, 12 of whom had a positive biopsy. An example test report would state:
 - tTG IgA – moderate to strong positive.
 - Based on our experience with similar patients there is a 92% chance this patient will have a positive duodenal biopsy (CI [62% to 100%]).

EMRs lend themselves to this type of decision support application. EMR-based query tools have long allowed queries for similar cases, but these have been used for research and education more than clinical care.^{23,24} Near-neighbor analysis can be used as a paradigm for linking diagnostic and outcome data through a data-driven process that is similar to the way physicians naturally develop intuitions about the performance of a test in their hands. Unlike results of published studies, local nearest-neighbor probabilities and practice patterns are readily updatable as medical practice changes under the influence of clinical decision support, taking the expression “meaningful use” to a new level.

Acknowledgements: Funding for this study was provided by the ARUP Institute for Clinical and Experimental Pathology. An abstract based on preliminary data generated for this article was presented at the “Pathology Informatics 2011” meeting in Pittsburgh, PA.

Conflict of Interest: The authors declare that they do not have a conflict of interest.

Corresponding Author: Brian H. Shirts, MD, PhD; Department of Laboratory Medicine, University of Washington, Rm NW120, 1959 NE Pacific Street, Box 357110, Seattle, WA 98195, USA (e-mail: shirtsb@uw.edu).

REFERENCES

1. Leffler DA, Schuppan D. Update on serologic testing in celiac disease. *Am J Gastroenterol*. 2010;105:2520–4.
2. Giersiepen K, Leigemann M, Stuhldreher N, Ronfani L, Husby S, Koletzko S, Korponay-Szabo IR. Accuracy of diagnostic antibody tests for coeliac disease in children: summary of an evidence report. *J Pediatr Gastroenterol Nutr*. 2012;54:229–41.
3. Parakkal D, Du H, Semer R, Ehrenpreis ED, Guandalini S. Do gastroenterologists adhere to diagnostic and treatment guidelines for celiac disease? *J Clin Gastroenterol*. 2012;46:e12–20.
4. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, Troncone R, et al. European Society for Pediatric Gastroen-

- terology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr*. 2012;54:136–60.
5. **Rozenberg O, Lerner A, Pacht A, Grinberg M, Reginashvili D, Henig C, Barak M.** A novel algorithm for the diagnosis of celiac disease and a comprehensive review of celiac disease diagnostics. *Clin Rev Allergy Immunol*. 2012;42:331–41.
 6. **Arguelles-Grande C, Tennyson CA, Lewis SK, Green PH, Bhagat G.** Variability in small bowel histopathology reporting between different pathology practice settings: impact on the diagnosis of coeliac disease. *J Clin Pathol*. 2011;12:12.
 7. **Barker CC, Mitton C, Jevon G, Mock T.** Can tissue transglutaminase antibody titers replace small-bowel biopsy to diagnose celiac disease in select pediatric populations? *Pediatrics*. 2005;115:1341–6.
 8. **Hopper AD, Hadjivassiliou M, Hurlstone DP, Lobo AJ, McAlindon ME, Egner W, Wild G, et al.** What is the role of serologic testing in celiac disease? A prospective, biopsy-confirmed study with economic analysis. *Clin Gastroenterol Hepatol*. 2008;6:314–20.
 9. **Donaldson MR, Book LS, Leiferman KM, Zone JJ, Neuhausen SL.** Strongly positive tissue transglutaminase antibodies are associated with Marsh 3 histopathology in adult and pediatric celiac disease. *J Clin Gastroenterol*. 2008;42:256–60.
 10. **Vivas S, Ruiz de Morales JG, Riestra S, Arias L, Fuentes D, Alvarez N, Calleja S, et al.** Duodenal biopsy may be avoided when high transglutaminase antibody titers are present. *World J Gastroenterol*. 2009;15:4775–80.
 11. **Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnarsley P, Cording E, et al.** Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27:1361–7. Epub 2012 May 1323.
 12. **Legare F, Ratte S, Stacey D, Kryworuchko J, Gravel K, Graham ID, Turcotte S.** Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev*. 2010; CD006732.
 13. **Darrell T, Indyk P, Shakhnarovich G, eds.** *Nearest-Neighbor Methods in Learning and Vision Theory and Practice*. Cambridge: MIT Press; 2005.
 14. **Papadopoulos AN, Manolopoulos Y.** *Nearest Neighbor Search: A Database Perspective*. 1st ed. New York: Springer; 2005.
 15. **Millard BL, Niepel M, Menden MP, Muhlich JL, Sorger PK.** Adaptive informatics for multifactorial and high-content biological data. *Nat Methods*. 2011;8:487–93. Epub 2011 Apr 2024.
 16. **van der Windt DA, Jellema P, Mulder CJ, Kneepkens CM, van der Horst HE.** Diagnostic testing for celiac disease among patients with abdominal symptoms: a systematic review. *JAMA*. 2010;303:1738–46.
 17. **Toftedal P, Nielsen C, Madsen JT, Titlestad K, Husby S, Lillevang ST.** Positive predictive value of serological diagnostic measures in celiac disease. *Clin Chem Lab Med*. 2010;48:685–91.
 18. **Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES.** Regional variations in diagnostic practices. *N Engl J Med*. 2010;363:45–53. Epub 2010 May 2012.
 19. **Ravelli A, Villanacci V, Monfredini C, Martinazzi S, Grassi V, Manenti S.** How patchy is patchy villous atrophy?: distribution pattern of histological lesions in the duodenum of children with celiac disease. *Am J Gastroenterol*. 2010;105:2103–10. Epub 2010 Apr 2106.
 20. **Hopper AD, Cross SS, Sanders DS.** Patchy villous atrophy in adult patients with suspected gluten-sensitive enteropathy: is a multiple duodenal biopsy strategy appropriate? *Endoscopy*. 2008;40:219–24. Epub 2008 Dec 2005.
 21. **Osl M, Ohno-Machado L, Baumgartner C, Tilg B, Dreiseitl S.** Improving calibration of logistic regression models by local estimates. *AMIA Annu Symp Proc*. 2008:535–39.
 22. **Jiang X, Osl M, Kim J, Ohno-Machado L.** Calibrating predictive model estimates to support personalized medicine. *J Am Med Inform Assoc*. 2012;19:263–74.
 23. **Safran C, Porter D, Lightfoot J, Rury CD, Underhill LH, Bleich HL, Slack WV.** *ClinQuery: a system for online searching of data in a teaching hospital*. *Ann Intern Med*. 1989;111:751–6.
 24. **Safran C, Chute CG.** Exploration and exploitation of clinical databases. *Int J Biomed Comput*. 1995;39:151–6.