Others Section

Original Article

Diagnosis of Celiac Disease using Fuzzy Logic Probabilistic System in North-Indian Patients

SUNNY THUKRAL[1], JATINDER SINGH BAL[2]

(cc) BY-NC-ND

ABSTRACT

Introduction: Fuzzy logic is gaining popularity with predictive healthcare applications in the field of medicine. It is consistently being used in recognising various chronic diseases with optimum accuracy. Fuzzy logic is also supposed to be a beneficial tool to diagnose celiac disease with approximate reasoning.

Aim: To design a fuzzy logic inference system to predict celiac disease based on individual symptoms in North-Indian patients.

Materials and Methods: The study was conducted in Punjab, Northern India, on 700 individuals using a questionnaire stratagem. SPSS software was used for the analysis of celiac data. The fuzzy logic system was implemented in Python software with 6 inputs and 1 fuzzy output parameter to determine the probability level of celiac disease.

Results: The fuzzy logic system produced 98.8% accuracy, 98.5% sensitivity and 98.39% specificity with minimal error rate. The prevalence of celiac disease obtained a 1:19 ratio from the survey in Punjab with 134 celiac patients.

Conclusion: Fuzzy logic yields better results in terms of sensitivity and specificity in comparison with other existing disease probabilistic systems. Hence, the fuzzy system is a fruitful tool for celiac disease disclosure without any painful testing strategy.

INTRODUCTION

Celiac disease is a type of chronic disease that endures in those carrying halotype *DQ2/DQ8* in the body [1]. The initial representation of celiac disease was given by Gee in 1888 [2]. After ingestion of gluten, an afflicted individual observes painful symptoms that appear instantly. The clinical features are abdominal bloating, anaemia, fatigue, abdominal pain, vomiting, skin rashes, diarrhea, failure to thrive, delayed growth, vitamin levels, anxiety, depression, etc., [3]. The manifestations differ from patient to patient and depend upon ppm (parts per million) gluten intake. During the world war, when wheat grain lessens in specific countries, it leads to a decline in certain celiac symptoms [4,5]. So, the notion of gluten-less products was prefaced to grab the celiac disease [6,7]. European Society for Paediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) guidelines [6] assures the tTG-IgA test for celiac diagnosis and if any ambiguous consequence, then proceeds with biopsy procedure for the confirmation of celiac disease. Gut damage is commonly seen in celiac patients with biopsy testing. An individual is also said to be celiac if the HLA *DQ2/DQ8* genes is present in the body with gastrointestinal symptoms.

The concept of fuzzy logic was given by Zadeh LA to deal with the problem of intermediate values or non-linear problems in 1965 [8]. A fuzzy system is a type of inference system which-takes input in the form of crisp data and evaluates them by comparing rules stored in the fuzzy database [9]. Mamdani and Sugeno fuzzy types of models can be used for producing a fuzzy knowledge base with different types of the membership function. Triangular and trapezoidal member functions are ordinarily used functions for the formation of fuzzy rules. Subsequently, the process of de-fuzzification produces an output value as the single probable value which is a prediction in terms of health applications [10]. Fuzzy logic is consistently used in recognising chronic diseases such as dental, cholera, liver diagnosis, viral infection, lung diseases, kidney diseases, etc., effectively and accurately. Alongside this, a lot of research has been done on the celiac disease with the help of clinical testing procedures in different countries. Researchers did exceptionally well to accomplish

submissions in diagnosing celiac disease with multiple clinical examinations as Tissue transglutaminase (tTG), Endomysial Antibody (EMA), Biopsy and Genetic Testing procedure [11-17].

So, this study was aimed to design a fuzzy logic inference system to predict celiac disease based on individual symptoms in North-Indian patients. The proposed system will abandon a fruitful consequence for entities and physicians for celiac disease disclosure in few seconds without any painful testing strategy.

MATERIALS AND METHODS

A wheat allergy Camp was organised on 7th April 2019, at DAV College, Amritsar, Punjab, India with the permission of Principal, Dr. Rajesh Kumar and under the guidance of three Gastrointestinal Physicians. The survey was conducted using a questionnaire on 700 individuals, using a random sampling technique. The questionnaire consisted of all the common gastrointestinal symptoms often seen by celiac patients. tTG clinical testing was done for the confirmation of celiac disease under the guidance of three gastrointestinal physicians [6]. The Body Mass Index (BMI) evaluation was also done, along with relevant family and medical history.

Sample size calculation: The sample size considered was based on the area of the population. The population of Amritsar is 11.3 lacs given in Census report published in 2011. So, Cochrane's sample size calculation was:

$x = ((Z^2 * pop(1-pop))/er^2)/ (1+(Z^2 * pop(1-pop))/er^2 \text{ No.})$, where,

x, is the total sampling unit to be taken,

Z, is the confidence level of up to 95% (Z=1.96),

Pop, is the sample population with 0.5 approximation

Er, is the error rate with 0.05 approximation

No. is the total population of the area to be examined.

$x = ((1.96^2 * 0.5(1-0.5))/0.05^2)/(1+(1.96^2 * 0.5(1-0.5))/0.05^2 * 1132383)$

$x = 384.029$

It requires only approx. 400 records for the dataset, but 700 instances

were collected to develop a fuzzy system with optimisation.

Inclusion criteria: Participants who were willing to be a part of this survey under the guidance of three gastrointestinal physicians.

Exclusion criteria: Participants with other existing chronic diseases were excluded from the study.

All the participants were subjected to general physical examination such as height, weight, BMI (as weight in kg, height in meter^2) and the same was recorded in the questionnaire. For BMI, individuals were divided into groups as Underweight (17.0-18.4 kg/m^2), Healthy Normal (18.5-24.9 kg/m^2), Overweight (25.0-29.9 kg/m^2), and Obese (≥30 kg/m^2).

The detail of celiac patients with values as never, mild, moderate, and severe were presented in the questionnaire for abdominal pain, abdominal bloating, anaemia, diarrhea, acidity, etc.

The questionnaire also consisted of a question on awareness concerning the celiac disease with three formal inputs as Aware, Not Aware, and Somewhat Aware.

STATISTICAL ANALYSIS

The data was analysed using Statistical Package for the Social Sciences (SPSS) software with principal component analysis and factor analysis for diminishing uncommon symptoms. For the implementation of a fuzzy celiac inference system, Python software was used to get the single probabilistic value with multiple predefined modules. To check the performance of the system, 700 records to be fed up to the fuzzy system to obtain sensitivity and specificity.

RESULTS

From the questionnaire data, the primary objective was to classify the most common symptoms among all in celiac patients to formulate a fuzzy logic predictive system. A total of 700 individuals who participated in the survey, with 134 celiac patients were confirmed. The mean age was found to be 21.72, with a standard deviation of 10.685.

[Table/Fig-1] shows symptomatic values given by celiac patients. The

Symptoms (Existing patients)	Never	Mild	Moderate	Severe
Abdominal pain	0	12	25	97
Abdominal bloating	40	66	12	17
Anaemia	1	4	57	72
Diarrhea	3	8	24	99
Acidity	66	58	8	2
Headache	60	66	7	1
Bone or joint pain	76	45	8	5
Defective tooth enamel	81	40	10	3
Vomiting	3	27	43	61
Fatigue (Weakness)	3	25	48	58
Constipation	77	46	7	4
Irritability	59	48	18	9
Anxiety	60	61	7	6
Skin rash	77	28	15	14
Weight loss	0	8	126	
Delayed growth	16	48	38	32

[Table/Fig-1]: Analysis of celiac symptoms with attributes.

most common symptoms were abdominal pain, diarrhea, anaemia, vomiting, and weight loss reported by celiac patients from the survey.

The 0-18 age-group attained a maximum number of 104 celiac patients from overall 134 celiac patients. It registers the early intake of wheat and breastfeeding, a vital role for the ingredients in the diet. Few celiac patients were also confirmed in other age-groups, as depicted in [Table/Fig-2].

		Age group (years)				
		0-18	18-35	35-60	60-100	Total
Celiac patient	No	172	326	64	4	566
	Yes	104	10	15	5	134
Total		276	336	79	9	700

[Table/Fig-2]: Celiac patient record with age-groups.

It is also noteworthy to focus on celiac patient's gender-wise records. From the survey data, 75 female celiac patients were found against 59 males, as shown in [Table/Fig-3].

		Celiac patient		
		No	Yes	Total
Gender	Female	232	75	307
	Male	334	59	393
Total		566	134	700

[Table/Fig-3]: Celiac patients with gender records.

Finally, the survey settled with the awareness level of celiac disease in Punjab, represented in [Table/Fig-4]. Majority (46.71%) of the patients were not aware of celiac disease.

|		Age group (years)				
		0-18	18-35	35-60	60-100	Total
Awareness	Not aware	111	168	47	1	327
	Somewhat aware	31	84	3	0	118
	Aware	134	84	29	8	255
Total		276	336	79	9	700

[Table/Fig-4]: Awareness with age-groups.

Based on these symptomatic conclusions, the idea of developing a fuzzy logic system was proposed to diagnose celiac disease using intelligent rather than a clinical approach. The development of a fuzzy logic system to detect celiac disease was implemented in PyCharm by importing various existing modules for evaluation and fuzzy graphical notation. The collections of symptoms were analysed from the SPSS tool to find out the most common symptoms achieved from the survey. The selection of vital parameters defined in the proposed system with six input parameters as anaemia, diarrhea, abdominal pain, vomiting, BMI, weight loss with a relevant range. The evaluation of all the input parameters defined using the triangular membership function for the fuzzy inference system. Crisp inputs were devised based on symptomatic nature using fuzzy values that differ according to the range. The fuzzy database constitutes of fuzzification process with a fuzzy if-then rule-based environment. Every input in the system is in the fuzzy act as antecedent and its outcome in the form of the consequent.

The scenario renders information with crisp values in numerical notation input to the fuzzy system as abdominal pain, anaemia and diarrhea to 10 points, vomiting to 7.5 with BMI 20.8, weight loss as 3.8 to the fuzzy logic probabilistic system as presented in [Table/Fig-5]. The input variables that have been practiced in the order, executed using a triangular membership function. It consists of multiple crisp values for each fuzzy variable, as abdominal pain with crisp values range; Mild (0-4), Moderate (1-9) and Severe (6-10). The entire range of severity of disease with symptoms is stewarded from 0-10 using triangular evaluation.

The outcome was 95.53873% according to the input given by an individual portrayed in [Table/Fig-6] with recommendations. The outcome signifies an individual with a strong chance of celiac disease based on the output. The output from the fuzzy system using the de-fuzzification technique accompanies the ESPGHAN guidelines by weightage given to every symptom [6].

It also prescribes the strong chance of celiac disease to avoid gluten in the diet and for the final confirmation must proceed with the tTG-IgA test.
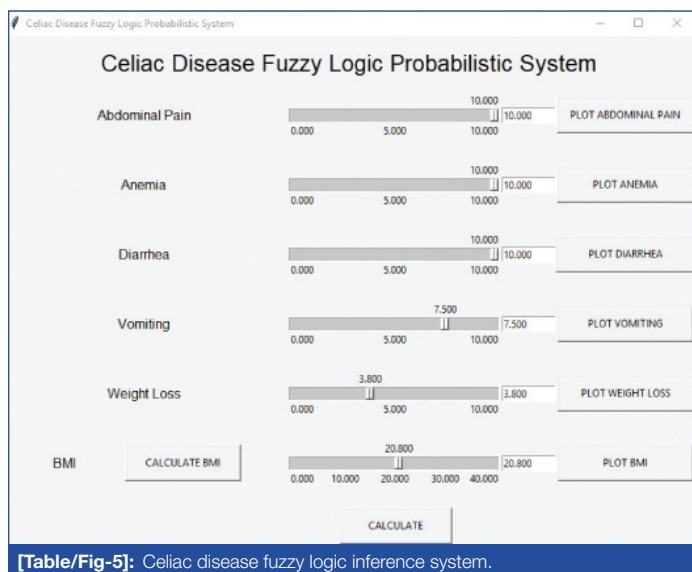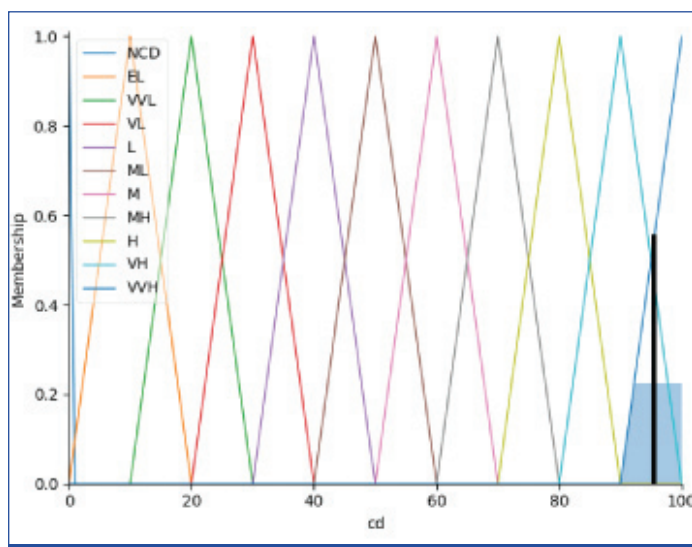

[Table/Fig-5]: Celiac disease fuzzy logic inference system.
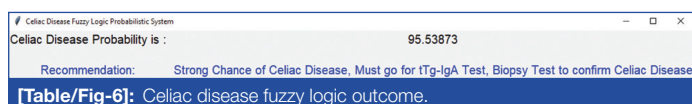
[Table/Fig-6]: Celiac disease fuzzy logic outcome.

The fuzzy graph outline in [Table/Fig-7] betokens fuzzy rules matched at the maximum with the entire fuzzy database. The process of de-fuzzification based on the area of maximum using different AND, OR combinations. Through the fuzzy logic system, it is possible to predict celiac disease with symptomatic information.

[Table/Fig-7]: Celiac disease fuzzy logic graph.
NCD: Non-celiac disease; EL: Extremely-low; VVL: Very-very-low; VL: Very-low; L: Low; ML: Medium-low; M: Medium; MH: Medium-high; H: High; VH: Very-high and VVH: Very-very-high

To check the accuracy of the fuzzy logic proposed system, every record obtained from the survey feed up to the system. The fuzzy logic system produced accuracy 98.8%, sensitivity 98.5% and specificity 98.39% with minimal error rate as depicted in [Table/Fig-8]. The outcome performance leads and helps physicians to use this proposed system to avoid clinical testing confirmation in comparison with existing fuzzy systems.

	Actual celiac disease (+)	Actual celic disease (-)
Predicted celiac disease (+)	132 (TP)	6 (FP)
Predicted celiac disease (-)	2 (FN)	560 (TN)

[Table/Fig-8]: Confusion matrix for performance evaluation.
TP: True positive; FP: False positive; FN: False negative; TN: True negative

DISCUSSION

Fuzzy logic is consistently used in recognising chronic diseases effectively and accurately. Various researchers deal with similar predictive systems with identical diseases using a fuzzy logic approach. Adeli M and Zarabadipour H fetched samples from the University of California, Irvine, (UCI) Machine Learning Repository and obtained 80% accuracy in detecting chronic diseases using the pattern recognition routine with fuzzy logic [18]. Parewe AM et al., admitted the hybrid fuzzy logic system with dental disease detection, achieved 82% accuracy [19]. Uduak A and Mfon M offered cholera diagnosis using fuzzy logic in 2013 with Matrix Laboratory (MATLAB) tool to prophesy the disease [20]. Liver diagnosis has been reported using fuzzy logic by Hashmi A and Khan MS, with three diverse types of liver diseases been realised by the fuzzy system [21]. Kaur R et al., proposed viral infection recognition using fuzzy logic with six different parameters to automate the system [22]. The symptomatic study has been characterised by Manikandan T et al., with the prediction of lung diseases using fuzzy logic with 95% accuracy of the system [23]. Recently in 2018, Zarandi MH and Abdolkarimzadeh M aimed a fuzzy expert system for diagnosing kidney disease in Iran on 400 specimens to differentiate chronic kidney or non-chronic disease patients having 80% accuracy [24]. So, there is a necessity to compose such intelligent systems to investigate diseases based on symptoms [25-27].

Limitation(s)

The limitation of the fuzzy system is that it is based on six input parameters with a specific range of every sign. Alongside this, the total number of instances for a more accurate system can be formulated with data from other geographical locations to get more accurate awareness and prevalence levels of celiac disease.

CONCLUSION(S)

The fuzzy logic system produced 98.8% accuracy, 98.5% sensitivity and 98.39% specificity with minimal error rate. The prevalence of celiac disease obtained a 1:19 ratio from the survey in Punjab with 134 celiac patients. The collected set of data from the survey elucidates that the female ratio carrying celiac disease is much more as compared to males. Few parameters like abdominal pain and varied BMI having vital inputs for the system as it resembles with all present celiac patients. As multiple expert systems accessible everywhere in the world for diagnosis of various diseases, fuzzy logic will be satisfactory to diagnose celiac disease. Due to sharpness in every specialty of medicine, fuzzy logic is associated with formulating intelligent machines for the determination of diseases. In this proposed system, it is ostensible that celiac disease can be foretold with symptoms. In the future, the system can be expanded with more than six parameters to predict celiac disease with a more precise nature as symptoms might vary in sophisticated individuals.

REFERENCES

[1] Kupper C. Dietary guidelines and implementation for celiac disease. Gastroenterology. 2005;128(4 SUPPL.1). https://doi.org/10.1053/j.gastro.2005.02.024.
[2] Gee SJ. On the coeliac affection. St. Bartholomew's Hosp Rep. 1888;24:17-20.
[3] Dowd B, Walker-Smith J. Samuel Gee, Aretaeus, and the coeliac affection. Br Med J. 1974;2(5909):45-47.
[4] Dicke WK, Weijers HA, Vande H. Coeliac disease II. The presence in wheat of a factor having a deleterious effect in cases of coeliac disease. Acta Paediatr. 1953;42(1):34-42. https://doi.org/10.1111/j.1651-2227.1953.tb05563.x.
[5] Marsh MN. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). Gastroenterology. 1992;102(1):330-54.
[6] Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr. 2012;54(1):136-60. doi:10.1097/MPG.0b013e31821a23d0.
[7] Murry JA, Watson T, Clearnnan B, Mitros F. Effect of a gluten-free diet on gastrointestinal symptoms in celiac disease. Am J ClinNutr. 2204;79(4):669-73. doi:10.10973/ajcn/79.4.669.

[8]   Zadeh LA. Fuzzy Sets. Elsevier, Information and Control. 1965;8(3):338-53. doi:10.1016/S0019-9958(65)90241-X.

[9]   Sadegh-Zadeh K. Fuzzy health, illness, and disease. Journal of Medicine and Philosophy. 2000;25:605-38.

[10] Mamdani EH, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies. 1975;7(1):01-13.

[11] Gulseren YD, Adiloglu AK, Yucel M, Dag Z, Filik L, Eyerci N, et al. Comparison of non-invasive tests with invasive tests in the diagnosis of celiac disease. J Clin Lab Anal. 2019;33:e22722. https://doi.org/10.1002/jcla.22722. PMID:30461063.

[12] Senapati S, Gutierrez-Achury J, Sood A, Midha V, Szperl A, Franke L, et al. Evaluation of European coeliac disease risk variants in a north Indian population. European Journal of Human Genetics. 2015;23(4):530-35. https://doi.org/10.1038/ejhg.2014.137.

[13] Digiacomo DV, Tennyson CA, Green PH, Demmer RT. Prevalence of gluten-free diet adherence among individuals without celiac disease in the USA: Results from the continuous national health and nutrition examination survey 2009-2010. Scandinavian Journal of Gastroenterology. 2013;48(8):921-25. https://doi.org/10.3109/00365521.2013.809598.

[14] Ford RPK. The gluten syndrome: A neurological disease. Medical Hypotheses. 2009;73(3):438-40. https://doi.org/10.1016/j.mehy.2009.03.037.

[15] Sood A, Midha V, Sood N, Avasthi G, Sehgal A. Prevalence of celiac disease among school children in Punjab, North India. Journal of Gastroenterology and Hepatology (Australia). 2006;21(10):1622-25. https://doi.org/10.1111/j.1440-1746.2006.04281.x.

[16] Mäki M, Mustalahti K, Kokkonen J, Kulmala P, Laurila K, Knip M, et al. Prevalence of celiac disease among children in Finland. New England Journal of Medicine. 2003;348(25):2517-24. https://doi.org/10.1056/nejmoa021687.

[17] Korponay-Szabó IR, Szabados K, Pusztai J, Uhrin K, Nemes E, Imre A. Population screening for coeliac disease in primary care by district nurses using a rapid antibody test: Diagnostic accuracy and feasibility study. British Medical Journal. 2007;335(7632):1244-47. https://doi.org/10.1136/bmj.39405.472975.80.

[18] Adeli M, Zarabadipour H. Automatic disease diagnosis systems using pattern recognition based genetic algorithm and neural networks. International Journal of Physical Sciences. 2011;6(25):6076-81. doi:10.5897/IJPS11.1075.

[19] Parewe AM, Mahmudy WF, Ramdhani F, Anggodo Y. Dental disease detection using hybrid fuzzy logic and evolution strategies. Journal of Telecommunications, Electronic and Computer Engineering. 2016;10(10):27-33.

[20] Uduak A, Mfon M. Proposed fuzzy framework for cholera diagnosis and monitoring. International Journal of Computer Applications. 2013;82(17):01-10. doi:10.5120/14252-1626.

[21] Hashmi A, Khan MS. Diagnosis blood test for liver disease using fuzzy logic. International Journal of Sciences: Basic and Applied Research. 2015;20(1):151-83.

[22] Kaur R, Kaur S, Rehani V. Fuzzy based automation system for predicting viral infections. International Journal for Innovative Research in Multidisciplinary Field. 2016;2(11):426-34.

[23] Manikandan T, Bharathi N, Satish M, Asokan V. Hybrid neuro fuzzy system for prediction of lung diseases based on the observed symptom values. Journal of Chemical and Pharmaceutical Sciences. 2017;8:69-76.

[24] Zarandi MH, Abdolkarimzadeh M. Fuzzy rule based expert system to diagnose chronic kidney disease. Advances in Intelligent Systems and Computing, Springer International Publishing. 2018;648:323-27. doi:10.1007/978-3-319-67137-6-37.

[25] Thukral S, Rana V. Versatility of fuzzy Logic in chronic diseases: A review. Medical Hypotheses. Elsevier. 2019;122:150-56. doi:10.1016/j.mehy.2018.11.017.

[26] Thukral S, Bal JS. Medical applications on fuzzy logic inference system: A review. Int J. Advanced Networking and Applications. 2019;10(4):3944-50. doi.org/10.35444/ijana.2019.10046.

[27] Bascunan KA, Vespa MC, Araya M. Celiac disease: Understanding the gluten-free diet. Eur J Nutr. 2017;56(2):449-59. doi:10.1007/s00394-016-1238-5.

PARTICULARS OF CONTRIBUTORS:
1.   Assistant Professor, Department of Computer Science, DAV College, Amritsar, Punjab, India.
2.   Vice-Chancellor, Department of Computer Science, Sant Baba Bhag Singh University, Jalandhar, Punjab, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:
Mr. Sunny Thukral,
DAV College, Amritsar, Punjab, India.
E-mail: sunnythukral@davcollegeasr.org