

Intelligent Diagnostic Assistant for Complicated Skin Diseases through C5's Algorithm

Fatemeh Rangraz Jeddi¹,
Masoud Arabfard², Zahra Arab
Kermany³

¹Faculty of Paramedical Sciences, Kashan
University of Medical Science, Kashan, Iran

²Tehran University, Kish International Campus,
Islamic Republic of Iran

³Health Information Technology Department,
Kashan University of Medical Science, Kashan,
Iran

Corresponding author: Masoud Arabfard, PhD student,
Tehran University, Kish International Campus, Islamic
Republic of Iran, ORCID ID: <http://orcid.org/0000-0003-3955-613X>. E-mail: arabfard@ut.ac.ir, ORCID ID: <http://orcid.org/0000-0003-3955-613X>

doi: 10.5455/aim.2017.25.182-186

ACTA INFORM MED. 2017 SEP; 25(3): 182-186

Received: Jul 27, 2017 • Accepted: Sep 04, 2017

© 2017 Fatemeh Rangraz Jeddi, Masoud Arabfard,
Zahra Arab Kermany

This is an Open Access article distributed under the
terms of the Creative Commons Attribution Non-
Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Introduction: Intelligent Diagnostic Assistant can be used for complicated diagnosis of skin diseases, which are among the most common causes of disability. The aim of this study was to design and implement a computerized intelligent diagnostic assistant for complicated skin diseases through C5's Algorithm.

Method: An applied-developmental study was done in 2015. Knowledge base was developed based on interviews with dermatologists through questionnaires and checklists. Knowledge representation was obtained from the train data in the database using Excel Microsoft Office. Clementine Software and C5's Algorithms were applied to draw the decision tree. Analysis of test accuracy was performed based on rules extracted using inference chains. The rules extracted from the decision tree were entered into the CLIPS programming environment and the intelligent diagnostic assistant was designed then. **Results:** The rules were defined using forward chaining inference technique and were entered into Clips programming environment as RULE. The accuracy and error rates obtained in the training phase from the decision tree were 99.56% and 0.44%, respectively. The accuracy of the decision tree was 98% and the error was 2% in the test phase. **Conclusion:** Intelligent diagnostic assistant can be used as a reliable system with high accuracy, sensitivity, specificity, and agreement.

Keywords: Knowledge representation, Computer assisted decision making, Dermatology, Expert system, C5's Algorithm.

1. INTRODUCTION

The risk of skin diseases have been reported to be up to 80% (1). According to the World Health Organization's report in 2013, the disease burden in 15 categories of skin diseases was 3-7 million (2). Utility criteria that reported the disease burden to be in the range of zero to one have been reported to be 0.640 to 1.000 for skin diseases such as pemphigus. This rate was 0.650 for HIV and 0.740 for kidney diseases (1). Skin diseases create additional problems because of their special features such as physical deformity that necessitate early detection and treatment of these patients, but the diagnosis of skin diseases is a complicated task. In such a case, widespread knowledge of skin diseases is inevitable for diagnosis and treatment of skin diseases (3-5), which will help dermatologists in diagnosis, and can also be used in rural and remote areas where access to dermatologists is not always possible. In particular, diagnostic intelligent assistant are able to

simulate and model problem-solving capabilities in a specific field. This is possible because these systems obtain knowledge from an expert and make decisions based on artificial intelligence (6). The knowledge base contains all specialized knowledge that is related to the problem and inference engine acts in a way that is similar to human reasoning (7). Inference techniques can be backward (deny or prove a pre-specified target) or forward (reach a goal that is not already known). Different methods of representing expert knowledge systems include triad of O-A-V, rules, semantic networks, and frames (7, 8).

2. AIM

The main objective of this study is to design and implement an intelligent diagnostic assistant for complicated skin diseases through C5 algorithm based on expert knowledge using the rule-based system and forward chaining inference techniques.

Disease	No.	Sign	Common symptoms with other diseases
BCC	1	History of prolonged exposure to the sun	MM
BCC	2	History of tanning	MM
BCC	3	History of frequent sunburn	MM
BCC	4	History of phototherapy	MM
BCC	5	History of radiotherapy	MM
BCC	6	History of exposure to chemical substances	MM
BCC	7	Papules or nodules with prominent telangiectasia and pearl	
BCC	8	Ulcerated or bleeding lesion	MM
BCC	9	Pigmented lesion	MM
BCC	10	Erythematous Plaque or macula	
BCC	11	Sclerotic plaque	
BCC	12	History of genetic diseases such as Xeroderma Pigmentosa	MM
BCC	13	Patient's complaint of the skin ulcer	MM
BCC	14	History of the change in color or increase in the size of the mole	MM
BCC	15	Black or brown or nodules with irregular margins and multiple color	MM
BCC	16	Hemorrhagic nodule or tumor	MM
BCC	17	Brown-black plaque with irregular color and border on the face and sun damage skin	MM
BCC	18	Lymphadenopathy	MM, Scabies
PV	19	Irritation and pain of mucosal	LP
PV	20	Dysphagia or pain on swallowing food	
PV	21	Nasal bleeding	
PV	22	Odynophagia	
PV	23	Red Eyes	
PV	24	Skin sores and Blister	LP
PV	25	Erosions of the mucosal membranes	LP
PV	26	Skin blister	LP
PV	27	Skin ulcer and erosions	LP, Scabies
PV	28	Wet lesion	scabies
PV	29	Involvement of the skin and mucosa	LP
PV	30	Local involvement	LP
PV	31	Occurrence on the head and neck and seborrheic areas	
PV	32	Painful oral erosions with subsequent involvement of the entire body	LP
PV	33	Loose and fragile blisters on the normal skin without erythema	
PV	34	Nikolsky sign	
PV	35	Oral lesions	LP
PV	36	Vesicle	scabies
PV	37	Infected lesions	scabies
PV	38	Blister	LP
PV	39	Mucosal lesions(reticular white lines-ulceration-erosions)	LP
PV	40	Post inflammatory pigmentation	LP
Scabies	41	History of recent travel	
Scabies	42	History of imprisonment	
Scabies	43	History of itching in other members of the family	
Scabies	44	Itching	LP
Scabies	45	Itching worsens at night	
Scabies	46	Itching worsens with heat and hot baths	
Scabies	47	Symmetric lesions	LP
Scabies	48	Presence of skin barrow	
Scabies	49	Papules	LP
Scabies	50	Nodules	LP
Scabies	51	Pustule	
Scabies	52	Ecematous lesion	LP
Scabies	53	Contact dermatitis	
Scabies	54	Excoriated lesions	LP

Scabies	55	Barrow with undulating lines 1-10 mm long in the area between the fingers and wrist	
Scabies	56	Find vesicle between the fingers, wrist and hand sides	
Scabies	57	Itchy lesions	LP
LP	58	Distribution of the lesions	
LP	59	Appearance of skin lesions(itchy and purple's papules-polygonal-flat)	
LP	60	Wickham strae	
LP	61	Scaring hair loss	
LP	62	Nail changes	
MM	63	Family history of atypical(dysplastic) moles or melanoma	
MM	64	History of multiple moles	
MM	65	Family history of melanoma	
MM	66	Pigmented bands of nails	
MM	67	Hutchinson's sign	

Table 1. A detailed list of specific and common symptoms of the selected disease

3. METHOD

This applied-developmental study was done in Razi Dermatology hospital, affiliated to TUMS, at 2015. The research was conducted in two phases, which include designing knowledge Base based on dermatologists' knowledge as experts and implemented as follows:

3.1. Designing Knowledge Base

The Acquisition of knowledge was done by Delphi technique and by holding several sessions and interviews with the experts using checklists of signs and symptoms that were developed based on International Classification of Diseases 10 revision (ICD-10) and semiology book of dermatology. One coordinator organized the meetings and received dermatologists' opinions and suggestions and consulted with other professionals until a consensus was reached. A total of 106 symptoms were determined for disease diagnosis.

3.2. Extracting and Representing Knowledge

Drawing decision tree from the trained data

First 5 class diseases consideration, and for every class, it was collect 60 medical records patient for the selected skin diseases that contains attribute of disease and patients. Then it was collect 150 medical records of the patients without have the selected skin diseases. So, there was 450 medical record for training the decision tree model as true class label.

Overall the sum of samples are 450 medical records for 6 class contains 5 disease and 1 no disease. Data of 60 medical records of the patients for each disease (a total of 300 medical records), and 150 medical records of the patients without the selected skin diseases were entered into the Excel Microsoft Office database. Knowledge representation was obtained from the train data in the database and decision tree was drawn using Clementine Software and C5 Algorithm. Accuracy and error of the drawn decision tree was calculated.

Testing and final analysis of the accuracy of decision tree

To test the validity of the decision tree, data of 20 medical records of the patients for each of the selected diseases (a total of 100 medical records) and 50 medical records of the patients without the selected skin diseases were entered into the database and accuracy of the drawn decision tree was calculated by measuring the accuracy and error rates. Accuracy higher than 90% was considered as the precision value of the decision tree.

Extracting Rules from the Decision Tree

Rules of If/Then/Else were extracted using forward chaining inference technique from the drawn decision tree.

The Implementation of a Diagnostic Intelligent Assistant

Extracted rules from the decision tree were entered into CLIPS programming environment as RULE (Diagram 1).

4. RESULTS

According to dermatologist's view, diseases including pemphigus vulgaris (PV), basal cell carcinoma (BCC), lichen planus (LP), malignant melanoma (MM), and scabies require recognition by a diagnostic intelligent assisted.

4.1. Accuracy and Error Rates Obtained in the Training

No.	Diseases	No. signs
1	Basal cell carcinoma	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18
2	Pemphigus vulgaris	19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40
3	Scabies	18,27,28,36,37,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57
4	Lichen planus	19,24,25,26,27,29,30,32,35,38,39,40,44,47,49,50,52,54,57,58,59,60,61,62
5	Malignant melanoma	1,2,3,4,5,6,8,9,12,13,14,15,16,17,18,63,64,65,66,67

Table 2. Brief listing of specific and common symptoms of the selected diseases

Scabies N=60	PV N=60	MM N=60	LP N=60	Except with selected diseases N=150	BCC N=60	No. Diagnosis
0	0	0	0	0	60	Basal cell carcinoma
0	0	0	0	149	1	Except with selected diseases
0	1	0	59	0	0	Lichen planus
0	0	60	0	0	0	Malignant melanoma
0	60	0	0	0	0	Pemphigus vulgaris
60	0	0	0	0	0	scabies

Table 3. Analysis of Decision Tree in Training Phase. Accuracy: %99.56 wrong: %0.44

Scabies N=20	PV N=20	MM N=20	LP N=20	Except with selected diseases N=50	BCC N=20	No. Diagnosis
0	0	0	0	0	20	Basal cell carcinoma
0	0	0	0	49	1	Except with selected diseases
1	0	0	19	0	0	Lichen planus
0	0	20	0	0	0	Malignant melanoma
0	20	0	0	0	0	Pemphigus vulgaris
19	0	0	0	1	0	scabies

Table 4. Analysis of Decision Tree in Testing Phase. Accuracy: %98 wrong: %2

Test Phase

In order to prepare data for drawing the decision tree, common symptoms for diagnosing the selected diseases were extracted (Table 1). A summarized list of specific and common symptoms of the selected diseases (Table 2) was developed. The accuracy and error rates obtained in the training phase from the decision tree (Figure 1) were 99.56% and 0.44%, respectively (Table 3). The accuracy of the decision tree was 98% and the error was 2% in the test phase (Table 4). Rules were defined using forward chaining inference technique and were entered into Clips programming environment as RULE.

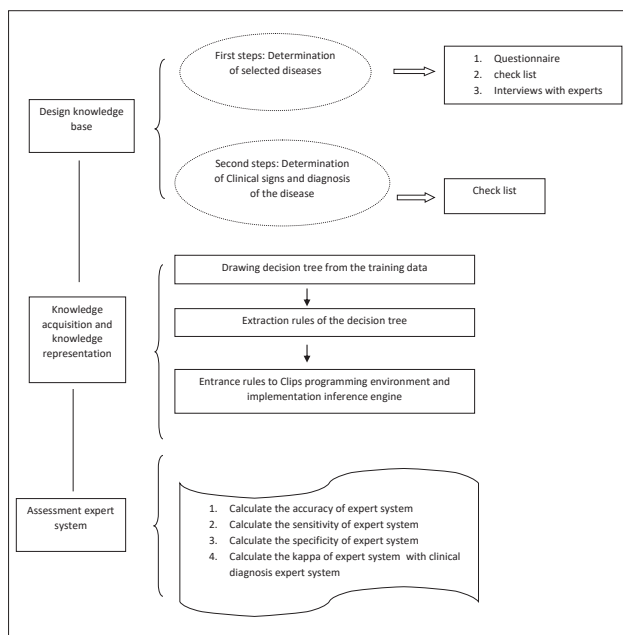


Diagram 1. Process Design, Implementation and Evaluation of Intelligent Diagnostic Assistant

5. DISCUSSION

The purpose of this research was to design and implement an intelligent diagnostic assistant for diagnosis of complicated skin diseases.

Strength of this study was preparation of a preliminary checklist for skin diseases using the International Classification of Diseases 10 Revision (ICD-10). Another strong point of our study was identification of diseases according to dermatologists' preferences.

Analysis of the data showed that in the training phase, the drawn decision tree diagnosed diseases with accuracy of 99.56% and error of 0.44%. Analysis in test phase showed accuracy of 98% and error of 2%.

Expert systems can be used to assist in the diagnosis of diseases (9). One study had suggested more than 60 rules as "Then-If" using the decision trees for 9 selected diseases (3). Another study had extracted rules from the decision tree by designing an expert system that was used to diagnose neurological diseases (10). Other studies had extracted rules by a two step decision tree of training and testing (11-16) or had used the decision tree to extract the rules that were used to design neural networks and estimate the risk of preeclampsia with an accuracy of 83.6% in the training phase and 93.8% in the testing phase (17). The methods used in these studies are in line with the methodology used in the current study. However, the accuracy of the drawn decision tree in the current study appeared to be 99.56% in the training phase and 98% in the testing phase.

Seto and colleagues designed a rules-based system in their study and used 105 rules for knowledge representation. Training and testing of the decision tree were not used in this research (18). Semantic network was used in the phase of knowledge representation which revealed causal relationships between variables and symptoms (19). These studies do not match with the findings of the current study. In this study we used a semantic network to obtain causal relationships between variables and symptoms. Although the se-

semantic network is a way of representing knowledge in intelligent system, the rules extracted from the decision tree is much easier to analyze due to the simplicity and ease of comprehension by the specialists, and is therefore superior to semantic network. Given that, expert systems are to be used in disease diagnosis, paying attention to the training and testing phases and evaluating the decision tree accuracy is essential. It is recommended that emphasis be placed on the training and testing phases and evaluation of the decision tree accuracy at the phase of knowledge representation as an inseparable part implementation of these systems.

Following implementation of the intelligent diagnostic assistant, and entering of the extracted rules from the decision tree to the intelligent diagnostic assistant, rule-based method was also used and the rules of If/Then/Else were extracted from the decision tree and were entered into the system programming (CLIPS) as RULE using forward chaining inference technique.

Doniz et al. used rule-based technique to enter rules of If/Then/Else as rule in the phase of knowledge representation to the proposed intelligent diagnostic assistant. Inference technique used in this study was forward (20). In some intelligent diagnostic assistants, rules of If/Then/Else, CLIPS programming environment and forward chaining inference technique (21, 23) or C programming language (24, 25) were used. The language used in both studies is consistent with our study. In one study. After extraction of "If/Then/Else" rules, the decision tree was drawn. By using forward chaining inference techniques, the selected rules were burned to reach a final diagnosis (26). CLIPS programming environment was also used (3, 10-12, 27) which is consistent with the current study.

Phase logic, decision trees and forward chaining inference techniques were used to determine the severity of pre-eclampsia (28). MATLAB programming environment was also applied in the intelligent diagnostic assistant for pneumonia diagnosis (29). Taking into consideration the fact that the MATLAB development environment is an interpretive language and that its execution speed is much slower than the compiled languages, it does not work so well in complex structures. In addition, its cost is high while the operating programs written in MATLAB are common and bring about essential problems (39). Therefore, in order to fix these defects, C programming language was used in the current study.

Usually, researchers use backward chaining inference technique in their studies (10) and C sharp language (30). Although some studies used the decision tree and the C language, the difference between them and the current study is the small number of tested data and the use of backward chaining in the inference engine (10, 30). Phase logic or neural network was used in designing the systems (15, 19) that is inconsistent with the current study. Given that all conditions of the current study were binary, it was not possible to use phase logic.

Some studies have used neural network in the phase of knowledge representation (17) that is not in consistence with the current study. The major problem of neural networks is that the user cannot understand the method of inference and conclusion of the system, learning system is dependent on the collection system of train data and the amount as well as the number of training courses. If train data is incomplete and

inadequate or the number of training courses is too many or too few, the system cannot have a good decision. The CLIPS program was usually used to design the intelligent diagnostic assistant due to its high flexibility, expandability and low cost (20).

6. CONCLUSION

Although artificial intelligence has great potential for improving medical decision-making, successful implementation of these systems requires special attention to details including using a method that is appropriate to the system target during the phase of knowledge representation. Paying attention to the details prepares the necessary ground for the use of these systems, like making specialists and physicians aware of the advantages of the intelligent diagnostic assistant in their professional area. Implementing intelligent diagnostic assistant with a user-friendly interface to increase physician's willingness is also essential to make effective use of these systems.

- **Clinical Relevance Statement:** The result of our study demonstrate that the intelligent diagnostic assistant can be used to determine the via Clementine Software and C5 Algorithm. Paying attention to details prepares the necessary ground for the use of these systems. Physicians should be aware of the advantages of the intelligent diagnostic assistant.
- **Acknowledgements:** This paper is the result of a master thesis in health information technology. Vice Chancellor for Research of University of Medical Sciences is highly appreciated for his financial support in the conduction of this study (Project No 9395).
- **Conflict of interest:** None declared.
- **Author's contribution:** Arabfard and Rangraz Jeddi conceived of the presented idea. Arab Kermany developed the theory and performed the computations. Rangraz Jeddi and Arab Kermany verified the analytical methods. Rangraz Jeddi and Arabfard encouraged to investigate and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

REFERENCES

1. Khatami A, Zartab H. Dermatology: A public health perspective. *Dermatology and Cosmetic* 2011; 2(4): 245-53.
2. Who methods and data sources for global burden of diseases estimates, Department of Health Statistics and Information Systems, November 2013: Available at http://www.who.int/healthinfo/global_health_estimates/en/.
3. Naser SSA, Akilla AN. A proposed Expert system for Skin Diseases Diagnosis. *Journal of Applied Sciences Research*. 2008; 4(12): 1682-93.
4. Maselena A, Hasan M. Skin diseases expert system using Dempster-shafer theory. *J Intelligent Systems and Applications*. 2012; 5: 38-44.
5. Rangraz Jeddi F, Abazari F, Moravveji A, Nadjafi M. Evaluating the ability of hospital information systems to establish evidence-based medicine in Iran. *Journal of medical systems*. 2013 Apr 1; 37(2): 1-7.
6. Naser SSA, Ola AZA. An expert system for diagnosing eye disease using clips. *Journal of Theoretical and Applied Information Technology*. 2008; 4(10).
7. Durkin J. *Expert Systems Design and Development*. Prentice Hall, Englewood Cliffs, N.J. 1994.
8. Jafari B, Azar A. Fuzzy decision tree; new modern opinion in strategy collection. *Public management investigation*. 2014; 7(19): 25-9.

9. Rangraz Jeddi F, Arabfard M, Arab Kermany Z, Gilasi H. The Diagnostic Value of Skin Disease Diagnosis Expert System. *Acta Inform Med.* 2016 Feb; 24(1): 30.
10. Al-Hajji A. A Rule-Based Expert System for Diagnosis and Symptom of Neurological Disorders "Neurologist Expert System (NES)". In *Proceedings of the 1st Taibah University International Conference on Computing and Information Technology (ICCIT)*. 2012: 12-14.
11. Alotaiby FA, Foda SG. Chemical Threats Diagnosis Expert System (CTDES). In *Advanced Materials Research*. 2012; 396-8; 1001-7.
12. Hota HS. Diagnosis of breast cancer using intelligent techniques. *International Journal of Emerging Science and Engineering (IJESE)*. 2013; 1(3): 45-53.
13. Matamoros A, Torrealba A, Rivas F, González S, Sánchez R, Molina L, Rivas-Echeverría C. Expert system for the pre-eclampsia prevention program. In: *Proceedings of the 4th WSEAS International Conference on Computational Intelligence, Man-machine Systems and Cybernetics (MSC'05) Miami USA*. 2005: 146-9.
14. Exarchos TP, Tsipouras MG, Exarchos CP, Papaloukas C, Fotiadis DI, Michalis LK. A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree. *Artificial Intelligence in Medicine*. 2007; 40(3):187-200.
15. Ali Keles A, Ayturk Keles B. ESTDD: Expert system for thyroid diseases diagnosis. *Expert systems with Application*. 2008; 34(1): 242-6.
16. Akbarzadeh TMR, Moshtagh-Khorasani M. A hierarchical fuzzy rule-based approach to aphasia diagnosis. *Journal of Biomedical Informatics*. 2007; 40(5): 465-75.
17. Neocleous CK, Anastasopoulos P, Nikolaidis KH, Schizas CN, Neocleous KC. Neural networks to estimate the risk for pre-eclampsia occurrence. In: *Neural Network. IJCNN International Joint Conference on IEEE*. 2009; 2221-5.
18. Seto E, Leonard K, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Developing healthcare rule-based expert systems: case study of a heart failure telemonitoring system. *International journal of medical informatics*. 2012; 81(8): 556-65.
19. Zolnoori, Maryam, Zarandi MHF. Designing fuzzy expert system for diagnosing and evaluating childhood asthma. Department of information technology management, Tarbiat Modares University. Master of Science in Information Technology Management. 2007: 199pp [In Persian].
20. Karagiannis A, Dounis T, Chalastras P, Tiropanis, Papachristos D. Design of Expert System for Search Allergy and Selection of the Skin Tests using CLIPS. *International Journal of Information and Communication Engineering*. 2007: 3(6).
21. Akter M, Uddin MS, Haque A. Diagnosis and management of diabetes mellitus through a knowledge-based system. In: *13th International Conference on Biomedical Engineering*. 2009: 1000-3.
22. Koutsoj C, Hatzilygeroudis I. FESMI: A Fuzzy Expert System for Diagnosis and Treatment of Male Impotence. In: *Knowledge-Based Intelligent Information and Engineering Systems*. Springer Berlin Heidelberg. 2004: 1106-13.
23. Santosh Kumar Patra P, Dipti S, Indrajit M. An expert system for diagnosis of human diseases. *International Journal of Computer Applications*. 2010; 13(1).
24. Guvenir HA, Emeksizb N. An expert system for the differential diagnosis of erythematous-squamous diseases. *Expert Systems with Applications*. 2000; 18(1): 43-9.
25. Garibaldi JM, Westgate JA, Ifeachor EC, Greene KR. The development and implementation of an expert system for the analysis of umbilical cord blood. *Artificial Intelligence in Medicine*. 1997; 10(2): 129-44.
26. Mailafiya I, Isiaka F. Expert System for Diagnosis of Hepatitis B. *West African Journal of Industrial and Academic Research* 2013; 4(1): 52-61.
27. Ghazanfari M, Kazemi Z. EDU-FES. A fuzzy expert system for advising students. *International Journal of Industrial Engineering and Production Management [Thesis]*. 1380; 12(3): 117-135 [In Persian].
28. Salari R, Developing of fuzzy expert system for detecting of preeclampsia [Thesis]. Tehran. Tehran University of medical Science. 2013 [In Persian].
29. Akramiane Arani A. Fuzzy expert system for the diagnosis of pneumonia. [Thesis]. Tehran. Tehran University of medical Science. 2013 [In Persian].
30. Kitporntheranunt M, Wiriyasuttiwong W. Development of a medical expert system for the diagnosis of ectopic pregnancy. *J Med Assoc Thai*. 2010; 93(2): 43-9.