The Development of an Expert System for Supporting the Diagnosis of Diffuse Interstitial Lung Diseases by High Resolution Computed Tomography[1]

Heon Han[a], Sung Hoon Chung[b] and Young Moon Chae[c]

[a]*Kangwon National University College of Medicine*
17-1 Hyoja-3-Dong, Chuncheon City, Kangwondo Province, 200-093, Korea
Tel: +82-33-258-2328, Fax: 82-33-257-4636, E-mail: hanheon@kangwon.ac.kr

[b]*Gyeongsang National University Hospital*
900 Gazwa-dong, Jinju City, South Kyongsang Province660-701, Korea
Tel: +82-55-750-8200, Fax: +82-55-758-1568, E-mail: shchung@nongae.gsnu.ac.kr

[c]*Yonsei University*
134 Shinchon-dong, Sodaemoon-gu, Seoul 120-749, Korea
Tel: +82-2-361-5048, Fax: +82-2-392-7734, E-mail: ymchae@yumc.yonsei.ac.kr

Abstract

The purpose of this study was to develop an expert system supporting the diagnosis of diffuse interstitial lung disease by high resolution computed tomography.

CLIPS (C language integrated production system) with rule-based reasoning was used to develop the system. Development of expert system had three stages knowledge acquisition, knowledge representation, and reasoning. Knowledge was obtained and integrated, from tables and figure legends of a representative textbook in the domain of this expert system, High-Resolution CT of the Lung, by Webb WR, Mueller NL, and Naidich DP. The acquired knowledge was analyzed to form a knowledge base. Overlapping knowledge was eliminated, similar pieces of knowledge were combined and professional terms were defined. The most important knowledge of findings was then selected for each disease. After groupings of combined findings were made, disease groups were analyzed sequentially to determine final *diagnoses.*

The system was based upon the input of 69 diseases, 185 findings, 73 conditions, 387 status, and 62 rules. The system was set up to determine the diagnoses of diseases from the combination of findings using forward reasoning. In an empirical trial, the system was applied to support the diagnosis of 40 cases of diffuse interstitial lung diseases. The performance of two doctors with support of the system was compared to that of another two doctors without support of the system. The two doctors with the support of the system made more accurate diagnoses than the doctors without the support of the system. The system is believed to be useful for the diagnosis of rare diseases and for cases with many possible differential diagnoses.

In conclusion, an expert system supporting the high resolution computed tomographic diagnosis of diffuse interstitial lung disease was developed and the system is thought to be useful for medical practice.

Key words:
Computers; Expert system; High resolution computed tomography; Diffuse interstitial lung disease

1. Introduction

Radiological diagnosis is composed of perception of abnormal findings, analysis of the abnormal findings, integration of knowledge according to the abnormal findings, and choice of most possible disease entity as a diagnosis and another possible disease entities as differential diagnoses.

High resolution computed tomography (HRCT) of diffuse interstitial lung diseases (DILD) is an important field in chest radiology, and has much information such as nodules, reticular opacities, and basal distribution, and has many disease entities, over one hundred in number, such as idiopathic pulmonary fibrosis, chronic eosinophilic pneumonia, and diffuse panbronchiolitis.

Analyzing and combining of HRCT findings leads radiologists to reach diagnoses and differential diagnoses. In the analysis of HRCT, many findings are used for diagnosis, and there are many differential diagnoses in some cases, and some cases are very rare in incidence. Therefore, diagnosis is sometimes difficult, and requires a wide and deep range of knowledge and experience. This is the reason an expert system is thought to be useful for this field.

2. Structure of DILD diagnosing system

2.1. Knowledge base

Knowledge is acquired from a representative textbook in this field, "High Resolution CT of the Lung" by R. Webb et al[1]. In the textbook, there are 17 tables explaining the diseases having a specific finding. For example, honeycombing is the finding of diseases such as sarcoidosis, IPF, silicosis, hypersensitivity pneumonitis, asbestosis, and histiocytosis-X. And there are 36 tables explaining the findings of a specific disease. For example, BOOP/COP shows findings such as patchy bilateral air-space consolidation, ground-glass opacity, subpleural and/or peribronchovascular distribution, bronchial wall thickening, dilatation in abnormal areas, and small nodular opacities, often peribronchiolar. And there are 99 figure legends. For example, lymphangitic carcinoma has findings of peribronchovascular interstitial thickening, interlobular septal thickening, subpleural nodules, subpleural interstitial thickening, and nodular thickening of major fissure. All the tables and figure legends are used as knowledge bases composed of 671 fragments of knowledge linking findings and diseases (Table 1). The knowledge fragments are simplified and standardized, and overlapping knowledge is eliminated. For example, "peripheral lung" and "peripheral distribution" have the same meaning, so the two are unified to one.

2.2. Inference engine

CLIPS 6.1[2] is used as an inference engine to develop the system.

2.3. Knowledge representation

Selection of some findings or finding groups is linked to a disease group, which is composed of many diseases containing the findings. As the next step, the disease group is required to choose to have or not to have another new finding. Consequently, final diagnosis and differential diagnosis is reached (Figure 1). At each step, expert knowledge, definition of terminology and example images can be supplied as reference.

Rules are made with forward reasoning as follows. T105 is a disease group composed of 7 diseases. P119 is a condition composed of one finding. "Adding of P119 to T105 results in T119" is a rule. T119 is a new disease group composed of 3 diseases out of T105. (Figure 2). It is a logic of "if, then".

The system is composed of 69 diseases, 185 findings, 73 conditions, 387 status, and 62 rules.

3. Empirical application of the system

To evaluate the usefulness of the system, 4 resident doctors are tested for 40 cases. Two doctors diagnosed the cases with support of the system, and another two doctors diagnosed the cases without support of the system, and the results were compared. Two doctors in 3rd grade showed no significant difference between them. But, for 4th grade residents, the difference was significant. A doctor with support of the system made correct answers in 9 cases more than a doctor without support of the system. The 9 cases were BOOP, lipoid pneumonia, PCP, BO, CMV, DPB, HP, and PAP. The diseases are somewhat rare diseases to diagnose or difficult diseases to diagnose because they have many differential diagnoses.

4. Discussion

This system is made from only one textbook with a limited knowledge base of HRCT information.

More extensive knowledge can also be obtained from other textbooks, many journals, and also clinical experience. Clinical information such as age, sex, symptom & duration, past history, and laboratory findings can be added to the knowledge base, in the future.

Test of four resident doctors is not sufficient to prove the usefulness of the system, but, diagnosis of rare disease is improved by the help of this system.

Authors think this system is now useful only for residents of the radiology department. But with modification of the system and strengthening of knowledge base, we can also make a more useful expert system not only for radiologists, but also for medical students or doctors other than radiologists.

This research provides only the basis for the complete system, and many tasks, such as knowledge supplementation, are still in need of development. But this research can be considered significant in the fact that it focused on 60 different diseases, which covers a rather larger target compared to earlier research, such as the experiment on diagnoses of five different liver diseases[3].

This system has the advantage of having an apparent target scope, a clear display of opinions, more advanced knowledge, and diagnosis through stage-oriented analysis. Furthermore, it is possible to select and add from medical knowledge only the information necessary for diagnosis, easy to display the relationship between the combination of opinions and diagnosis groups, and simple to make improvements in the system, such as modifications and implementation of knowledge. To further improve this system, new collaborative research by radiologists and knowledge engineers on knowledge expression or inferential methods is required to sufficiently increase the amount of knowledge so that it can easily be utilized in clinical practice.

5. Conclusion

In this paper, a medical expert system for diagnosing diffuse interstitial lung disease was developed with rule based reasoning. The system is composed of 69 diseases, 185 findings, 73 conditions, 387 status, and 62 rules, and can help doctors improving the diagnosis of diffuse interstitial lung disease. In the future, the system can be further improved very easily by strengthening the knowledge base.

References

1. Webb W.R., Mueller N.L. and Naidich D.P. 1996. *High-resolution CT of the Lung*. Second edition. Lippincott-Raven. Philadelphia. 41-269.

2.http://www.ghgcorp.com/clips/WhatIsCLIPS.html

3. Maclin P.S. and Dempsey J. 1992. Using an Artificial Neural Network to Diagnose Hepatic Masses. *J of Med Syst*. 16(5): 215-225.

Table 1. Example of knowledge acquisition from the tables explaining findings

Finding	Corresponding disease list

Finding	Corresponding disease list
Honeycombing	Sarcoidosis
Honeycombing	IPF
Honeycombing	Silicosis
Honeycombing	Hypersensitivity pneumonitis
Honeycombing	Asbestosis
Honeycombing	Histiocytosis X

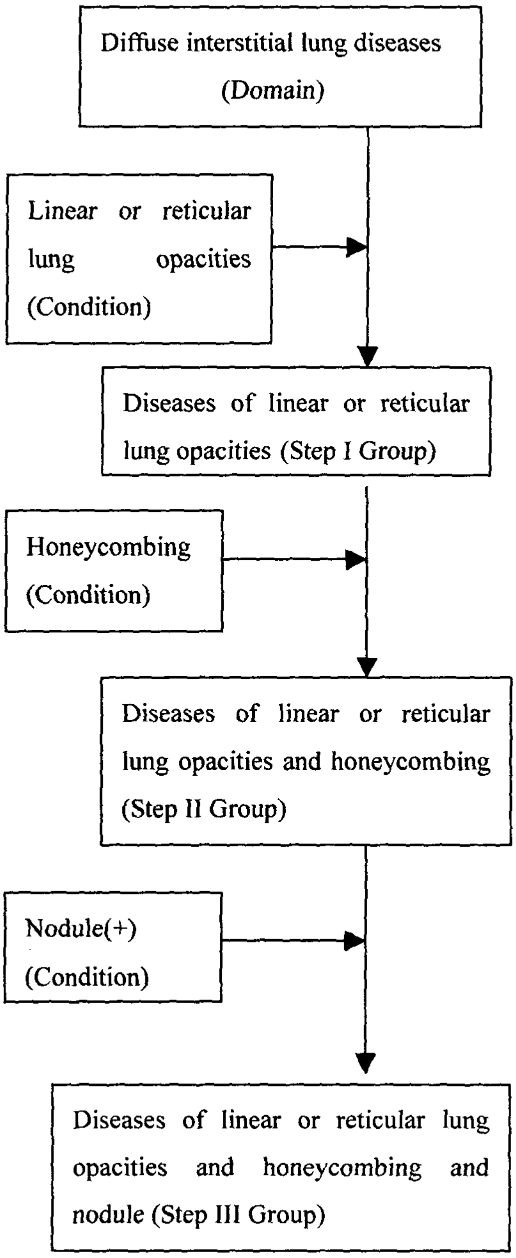

Figure 1. Example of hierarchy of forward reasoning in the system

105	IPF
T105	Asbestosis
T105	Lymphangitic carcinomatosis
T105	Pulmonary edema, interstitial
T105	Tuberculosis
T105	Bronchopneumonia
T105	Diffuse panbronchiolitis

P119	Tree-in-bud appearance

T105	P119	T119

T119	Tuberculosis
T119	Bronchopneumonia
T119	Diffuse panbronchiolitis

Figure 2. Example of rule making