Research and Design on Knowledge Base of NCCN-Based Cancer CDSS

ZHAO Jiang-sheng 1, Huang Xi 2, Yang Lv-qing 1

1 Software School of Xiamen University, Xiamen, China
2 Xiamen Ricebird Software Technology Co. Ltd, Xiamen, China
zjs@xmu.edu.cn

Abstract. At present, China faces a severe situation in which the incidence of malignant tumors is rising. At the same time, the search engines obtain a large amount of useless or even false medical information, and the tensions between doctors and patients are worsening every day. In order to enable patients to quickly and easily obtain the most authoritative and effective treatment solution, we are in imminent need of the research and development of tumor clinical diagnosis decision-making system. This system is based on NCCN's latest clinical practice guidelines for malignant tumors, using the decision-making support system approach for research and development. This paper researches and analyzes the practice guidelines and gives the structure design for the knowledge base.

1. Introduction

According to statistics, the overall incidence of malignant tumors in China has been rising in recent years, reaching 250.28/100,000 in 2011. By 2014, it has risen to 276.16/100,000, and is still deteriorating. [1] It is obvious that the situation of the incidence of malignant tumors in China is very serious. This serious illness brings great pressure to the patients themselves and their families. It can be said to be a disaster for them physically, mentally and economically. However, once people suffer from such a major disease, they generally only use various search engines to obtain medical treatment or rescue information, which is the main approach and method at present. But being profit-driven, these search engines are full of false medical information, which makes it difficult for people to distinguish between true and false, and is thus deceptive. At the same time, due to the serious inequality of medical information between doctors and patients, the patients are not sure whether the treatment to be accepted is correct or not. Once the effect is not good, they are likely to doubt the correctness of the doctor's treatment. As a result, the mistrust between doctors and patients is deepening, the tensions between doctors and patients are tremendous, and disputes between doctors and patients are endless. The situation is especially prominent for major diseases such as tumors. It is the responsibility and mission of every conscientious doctor and technician to bring to patients the most authoritative and effective treatments in the world.

1.1 Introduction to NCCN
NCCN®, also known as the National Comprehensive Cancer Network, is a non-profit consortium composed of 23 world-leading American cancer centers. Its purpose is to wholeheartedly promote the quality and efficiency of medical management of cancer patients. Through the expertise and leadership of clinical talents in member institutions, NCCN develops clinical resources that provide invaluable information to decision makers in the health care system. [2]

Based on the latest cancer research results of the world, NCCN has developed 47 clinical guidelines for malignant tumors (hereinafter referred to as NCCN guidelines) that are freely implemented globally and kept up to date, and are recognized and followed by clinicians around the world. The NCCN guidelines ensure that clinicians have access to state-of-the-art and the most standardized prevention, examination, diagnosis, treatment, and other supportive therapies that cover approximately 97% of cancers to achieve the optimal outcomes [3].

The NCCN guidelines are very authoritative and are the most widely accepted diagnostic guide for oncologists worldwide, revered by them as a classic.

1.2. The official NCCN website in China

In 2015, NCCN set up its official website in China, with academician Sun Yan as the convener, organizing domestic professionals to translate and update the NCCN guidelines [2]. However, only a small part of the NCCN guidelines has been translated and the update is not timely, while the development of cancer research is rapid and the NCCN update speed is quite fast. Taking non-small cell lung cancer as an example, from January to September 2017, NCCN had updated 8 times according to the latest research results in this field, providing 8 versions of treatment guidelines for non-small cell lung cancer. Obviously, in this way it is difficult to meet the needs of domestic oncology medical workers.

1.3. The proposal of NCCN-based tumor clinical diagnosis decision-making system

The NCCN Guide is currently the world’s most authoritative guide for oncology diagnosis. The development of an authority-based NCCN guideline, both for physicians and for patients’ diagnostic decision-making systems, is thus urgent.

2. Introduction to clinical diagnosis decision system

2.1. Decision-making Support System

Decision-making Support System, DSS, is a system developed on the basis of management information system by deepening the application concept of it. [4] The basic structure of the decision-making system consists of a database management system, a model library management system, a method library management system, a knowledge base, and a problem processing system, as shown in Figure 1.
2.2. Clinical Decision Support System

On the other hand, Clinical Decision Support System, CDSS, refers to a computer system that supports clinical decision making, which fully utilizes available computer technology to improve and elevate decision-making efficiency for semi-structured or unstructured medical problems through human-computer interaction. [5]

The knowledge base-based CDSS generally consists of three parts, namely the knowledge base, the inference engine and the human-machine communication interface part, but only partially depends on the rules and connections including the compiled information. The IF-THEN rules are usually used to store and manage the knowledge. [5]

In this system, NCCN as an authoritative practice guide for clinical diagnosis and treatment of cancer can be regarded as a knowledge base for decision-making support.

3. Design of the System

3.1. Components of the system

The system consists of two parts: the foreground and the background. The background is mainly composed of knowledge base management, method library management, security module, system configuration and other modules. The foreground mainly includes functional modules such as homepage, personal data modification, and diagnostic management.

3.2. Basic structure of the knowledge base

The knowledge base is an essential element in supporting the decision-making process of the system and needs to be entered and maintained by professional medical personnel. According to the analysis of NCCN-related data (with cervical cancer data as the example only), the NCCN structural library can be divided into the following basic structures:

- Examination
- Discuss the results of the examination
- Do the next examination / do the treatment

Taking a few items randomly from the NCCN data for analysis, we can see the first CERV-1:
Figure 2. CERV-1

As shown in Figure CERV-1, it is obvious that the content described in the "WORKUP" stage is the examination, and the "CLINICAL STAGE" is the examination structure. According to the different examination results, there will be different treatment solutions or examination solutions, and then the system continues to analyze, with the most complicated "IA2, IB1" as an example.

After the examination of the results of this period, the NCCN knowledge base provides two decision-making options: CERV-2, (fertility sparing), and CERV-3 and CERV-4 (non-fertility sparing). Next, we select CERV-3 for analysis.

Figure 3. CERV-3

After entering CERV-3, an examination was actually carried out. That is, “with or without lymph-vascular space invasion”. If not, the decision tree provides the next examination request for a “cone biopsy”. After the examination is completed, according to the biopsy results, there are three cases: “negative margins and inoperable”, “negative margins and operable” and “positive margins for dysplasia or carcinoma”. According to these three cases, the decision tree gives a treatment or requires observation to continue.
After the treatment is completed as required, the decision tree moves the diagnostic process to CERV-10, i.e. “See Surveillance”.

3.3. Design of the knowledge base

Based on the above discussion, we find that the support library is actually a process of continuous decision making. And this decision making can be seen as a “finite state machine”. The basic structure of a finite state machine consists of a State and a Transition. By mapping it to the top, each examination/treatment solution can be considered as a status, and the examination result is treated as a condition for status transition. So, for patients or doctors, the use of this system equals a process that uses workflow.

The system regards decision making as a workflow, treats the solutions as steps of the workflow, and treats the possible result as a status transition set of the workflow or the decision link for references in the knowledge base.

3.3.1. Decision making. In this system, a complete diagnosis of a condition is called a decision making, so decision making is an important part of the knowledge base. Each decision making can consist of multiple sub-decision makings. Decision making is a tree structure that is suitable for workflows, while sub-decision makings are similar to sub-workflows. The data structure of the decision making is composed of ID, decision making name, decision making version, keyword, description, creator, creation time, editor and other fields.

3.3.2. Solution. Each node in the decision tree is called a solution, and each decision is made up of one or more entry solutions. The solution consists of the body of the solution and possible outcomes. The data structure consists of ID, solution name, solution type, specific solution description, version, level, keyword, description, creator, creation time, editor, edit time and other fields.

The keyword of the solution means that in other parts of the system, if the \\"#keyword#\" expression is used, the text is linked to the solution automatically. According to NCCN, the types of programs can be divided into "examination", "treatment", and "observation".

3.3.3. Possible results. Each solution has zero or one possible result. As long as there is a possible result, it means that the system can refine the solution, that a further examination/care/treatment
solution can be provided. The data structure consists of ID, solution ID [foreign key], result name, keyword, description, result details, and the solution ID pointed to by the result.

The result details describe how the user can determine which outcome he or she belongs to. For example, under what circumstances an index on an examination report belongs to the result. While keywords and descriptions are reserved fields.

3.3.4. Reference link. In order to make the system easy to understand and use for the user, we design the reference link for each node as the # and @ symbols that can be seen in Weibo.

4. Conclusion
The key to the NCCN-based oncology clinical diagnosis decision-making support system lies in the design of the knowledge base. Since the NCCN clinical practice guide contains 47 kinds(categories) of related diagnostic methods for malignant tumors, there are many differences in the structural expression of each type of method, causing notable difficulties to the designs of knowledge base and system. In the process, R&D personnel and professional doctors need to discuss with each other to conclude with a relatively common knowledge base data structure. In the next step, in order to make the system easy-to-use for the patient, and to find out the correct diagnosis according to the results of the existing examination, we will make the system give the possible examination results provided by the professional clinician team for each step node, and based on these results, determine its possible next step node. Due to its professionalism, authority and ease of use, the system can be used as a quick reference manual for patients and even professional doctors.

References
[1] Chinese Network for Industrial Information http://www.chyxx.com/industry/201704/511654.html
[2] Official Website of NCCN in China http://www.nccnchina.org/nccn-guidelines-china.aspx
[3] Official Website of NCCN in America https://www.nccn.org/professionals/physician_gls/f_guidelines.asp/site
[4] Baidu Encyclopedia https://baike.baidu.com/item/%E5%86%B3%E7%AD%96%E6%94%AF%E6%8C%81%E7%B3%BB%E7%BB%9F/85457?fr=aladdin
[5] Baidu Encyclopedia https://baike.baidu.com/item/CDSS/386665
[6] Clinical Decision Support System J 2018 Chinese Nursing Management
[7] Zhang Xiaolong and Luo Mingjian J 2005 Decision tree pruning algorithm with IF-THEN rules Computer Applications vol 25