

Chapter 4

Methodology

This chapter deals with a detailed methodology adopted for the design and development of a Customized Decision Support System for Breast and Cervix Cancer.

4.1 Overview

This research was conducted in order to design and develop a clinical decision support system for breast and cervix cancer. In order to reach the research goal, the researcher opted to review the literature to obtain the information regarding the existing clinical decision support system in cancer care with specific reference to breast and cervix cancer care. The researcher has also identified the present level of Knowledge and utilization pattern of cancer information system in practice among the oncologists. The perception and expectation of the oncologists about decision support system is also obtained and this forms the platform for the designing, testing and adopting a more successful clinical decision support system for breast and cervix cancer patients.

4.2 Research Design

A descriptive method of research was used for the study. John W Creswell [100] describes the descriptive method of research where the researcher gathered information based on the present existing condition. The method is used to define the nature of the situation, as it exists at the time of study and to explore the causes of a particular a phenomenon. In this study, the descriptive method was employed to identify the knowledge, utilization pattern, perception and expectation of respondents using a tested tool by recruiting and selecting an oncologist during the time of research. The aim of this research is to design and develop a customized decision support system for breast and cervix cancer where this method helped to obtain first hand data for deriving the rationale and sound conclusions and recommendations.

The following objectives were drawn in order to achieve the research goal:

Objective.1. To identify the various clinical decision support systems available in practice.

Objective.2. Understand the present level of knowledge among Oncologists with regard to information systems.

Objective.3. Assess the utilization of existing information system among oncologist.

Objective.4. Identify the Oncologists opinion on clinical decision support system.

Objective.5. Understand the Oncologist expectation from the clinical decision support system

Objective.6. Design and Develop a Customized Decision Support System for Breast and Cervix Cancer.

4.3 Methodology as per Objectives:

The methodology adopted for the achieving each objectives was:

4.3.1 Methodology for Objective 1:

Objective 1. To identify The Various Clinical Decision Support Systems Available In Practice.

An extensive literature search was conducted using the research database such as PubMed [101], Google Scholar [102], MedlinePlus [103], ProQuest [104], ACM Digital Library [105], IEEEExplore [106], Open-J-Gate [107], EBSCO [108] etc to identify the existing decision support system in oncology in general and with specific reference to breast and cervix cancer care.

It is evident from the literature survey that several systems have been developed in the field of clinical decision support system such as CASNET, PIP, ABEL, Leeds Abdominal Pain System, EON System [109], AAPHelp [110], Internist-I [111], Mycin [112], DxPlain [113], Isabel Healthcare [114], GDSI [115], PROMEDAS [116], Caduceus [43], DxPlain [45], Diagnosis Pro [46], Visual DX [47], RODIA [48], HELP [50], CASNET [51], ABEL [52], QMR [53] etc. but, there are only a few available in the domain of breast and cervical cancer care. Works of interest in this case [78]- [99] have been reviewed and limitations of the systems have been noted down and addressed during the design, development and implementation of a new customized system.

4.3.2 Methodology for Objectives 2, 3, 4 & 5

Objective 2. Understand the present level of knowledge among Oncologists with regard to information systems.

Objective 3. Assess the utilization of existing information system among oncologist.

Objective 4. Identify the Oncologists opinion on clinical decision support system.

Objective 5. Understand the Oncologist expectation from the clinical decision support system

4.3.2.1 Study Design

A survey was carried out in 12 cancer hospitals and research centers of southern India. The Oncologists were asked to rate their opinion on CDSS and suggest their expectations from the CDSS.

4.3.2.2 Study Setting

The Oncologists were selected from the different cancer hospitals and research centers of southern India.

4.3.2.3 Sample Size

In order to determine the level of perception and expectation from CDSS, 100 Oncologists were asked to participate in the study.

4.3.2.4 Inclusion Criteria

To obtain pertinent information, certain inclusion criteria were drawn. The Oncologists with MD and 5 years of working experience in oncology set-up qualified for the study. This qualification ensured that the participants understood the role of information support services in their practice and were able to give appropriate feedback in this regard. The participants who were not willing to participate in the study were excluded.

A purposive sampling method was adopted where the three states i.e. Karnataka, Kerala and Tamil-Nadu were included in the study. In these three states, 12 cancer hospitals and research centers were considered for selecting the respondents based on the availability and convenience and also as per the qualification of the Oncologists.

4.3.2.5 Data Collection Tool

In order to collect the data from the respondents, a validated and pre tested questionnaire was used. The questionnaire includes identification details with (Refer. Annexure.1) 5 major sections namely:

Section.1. Knowledge of Information System:

This Section consists of 7 items in Relation to the Knowledge of Cancer Information System in Terms of Computers, Peripheral Devices, Database Management System, Handling Electronic Images and Graphics, Various Technologies Involved Such as Electronic Health Records, Computerized Providers Order Entry, Picture Archiving And Communication System. The Oncologists were asked to respond by marking their response in a scale from 5 to 1 (**Excellent to Below Average**)

Section.2. Utilization of Existing Information System

A personal observation was done using a checklist consisting of 23 items to identify whether the existing information system of the hospital supports the Oncologists in documenting the patient information, decision making, patient care, research & education and reporting. **The observation was marked in terms of YES/NO.**

Section.3. Opinion on Decision Support System

To understand the Oncologists perception about the decision support system 8 parameters were included in this section where the Oncologists were asked

to mark whether the implementation of decision support system will increase their productivity, create hassles, do jobs better than people, decrease hospital cost, less work for people, upgrade the job functions of non-medical personnel, reduce the doctors control over medical practice. The Oncologists were asked to respond by marking their response in a scale from 5 to 1 (**Strongly Agree to Strongly Disagree**).

Section.4. Expectation from Decision Support System

This section intended to gather the Oncologists expectation from decision support system using 16 parameters in terms of documentation, patient care, clinical decision making, knowledge management, statistical analysis and evidence based practice. The Oncologists were asked to respond by marking their response in a scale from 5 to 1 (**Strongly Agree to Strongly Disagree**).

Section.5. Suggestions and Recommendations

This is an open end section where the Oncologists were expected to mark their recommendations in terms of knowledge base, patient information and case base.

4.3.2.6 Methods of Data Collection

The questionnaire was administered to the Oncologists after taking consent from the concerned hospital authorities. The respondents were first briefed about the research and its objectives and the purpose of survey. An informed consent was obtained from the Oncologists for being the part of the study where the data are collected by distributing the questionnaire among those who responded.

The existing workflow has also been observed to determine the total time for capturing and disseminating the patient information to the oncologists and also the issues related to the accessibility of clinical knowledge. The oncologists

were also asked to list out the areas of improvement in terms of information support service in their practice.

4.3.2.7 Statistical Analysis

Statistical Package of Social Sciences (SPSS) 16.0 version was used to analyze the data. The results of the survey are represented in terms of percentage and mean score. The Chi square test was done and $p < 0.05$ was considered significant. The average time for capturing and dissemination of patient information to the end users was also calculated to understand the issues related to the access and availability of information and clinical knowledge and to suggest a better CDSS for improving the information support service process.

4.3.3 Methodology for Objective 6:

Objective 6. Design and Develop a Customized Decision Support System for Breast and Cervix Cancer.

Based on the literature survey the limitations of existing system were identified and addressed during the designing phase of the decision support system “ONCOSYS”.

It has been decided to include five important modules (Figure 4.1) into the system to completely assist the oncologist in quality decision making during patient care. The modules and the features incorporated are:

4.3.3.1 Patient database module

The patient database module is being incorporated into the system to support the clinicians in recording, retrieving and updating patient’s information. AJCC 6th edition cancer staging system (Annexure 3 & 4) and ICD-Oncology-3rd version (Annexure 5) is also incorporated into the module to allow the

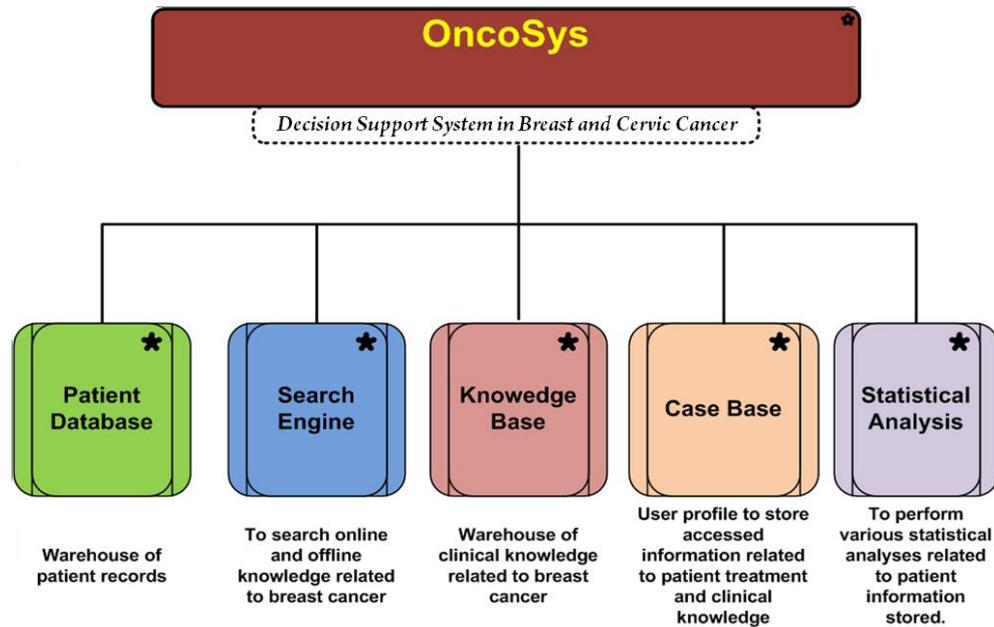


Figure 4.1: Modules of Clinical Decision Support System

clinicians in cancer staging and recording of topographical and morphological diagnosis correctly. The content of this module are decided based on the feedback received from the oncologist and literature review.

The following fields were included to capture the patient information:

- Identification, Demographic and Diagnostic Information

- Registration Number
- Date of First Diagnosis
- Patient's Name
- Age
- Gender
- Date of Birth
- Father's / Husband Name

- Address (Local and Permanent)
 - Basis of Diagnosis & Date
 - Details of Tumor (Primary & Secondary)
 - Topography and Morphology with ICD-Oncology Codes
- Details of Socio-economic Status, Family Income, Occupation**
- Details of Income
 - Dependent Family Member
 - Occupation
- Co-Morbid Condition - Details of Stages**
- Staging System
 - Clinical Stage UICC
 - TNM Stage AJCC
 - Pathological Stages
 - R Classification
 - Stage Group
 - Assessment of Staging
- Details of Cancer Directed Treatment**
- Details of Treatment given prior to reporting institution
 - Surgery
 - Chemotherapy

- Radiotherapy
- Hormonal Therapy
- Details of Treatment at reporting institution
 - Surgery
 - Chemotherapy
 - Radiotherapy
 - Hormonal Therapy
- Performance status before treatment
- Response to the treatment
- Complication during treatment
- Performance status after the treatment

- Follow-up Details

- Date
- Basis of Diagnosis
- Details of treatment and outcome
- Second primary
- Treatment
- Outcome
- Vital status
- Method of last Follow-up

4.3.3.2 Search engine module

A search engine module is included to allow the user to interact with the system through user interface for the access of offline and online knowledge of the respective domain i.e. breast and cervical cancer. The search criteria for the instant access of domain knowledge defined during the creation of search engine are such as:

- **Years of evidence** i.e. the search engine will assist the user for offline and online access of latest evidence available related to breast and cervical cancer. The system will sort out and display the recent evidence available as per the required search.

- **Reviewed article** i.e. after getting access to the clinical knowledge, the user can rank the page with a scale range from average (1) to excellent (5). The scores marked by different users will be automatically calculated and the average of the same will be considered. Next time when a new user interacts with the system and seeks for clinical knowledge related to breast cancer, the related URL of the article/page will be created and shown based on the rank marked by the other users. Clicking on respective URL will direct the user to the respective page.

4.3.3.3 Knowledge Base Module

Knowledge Base Module is included to support the user to keep updating the offline and online knowledge of breast cancer. Knowledge base creation process was conducted in five phases:

Phase 1 included the gathering of knowledge related to respective domain i.e. breast and cervical cancer. The sources of information looked for

during the preparation of the model were textbooks, journals, best practice guidelines, meta-analysis, systematic review and websites.

Phase 2 included the authentication of gathered knowledge. The authentications were conducted with the discussion of domain experts involved in cancer care. The criteria included for the authentication were knowledge source and contents.

Phase 3 included the creation of data warehouse for the storage of authenticated knowledge. The warehouse contained the detailed information related to breast and cervical cancer treatment.

Phases 4 & 5 included the specification of rules/key terms and algorithm for the required search.

The knowledge base is a well organized dataset where each document has unique identity and metadata. The metadata is a file provided for each dataset in the database. The metadata furnishes extensive general and medical characteristics of the dataset. The metadata provides identification, description, content, purpose, status, accessibility, creator, publisher, data quality, condition, spatial data organization, spatial reference and attribute description, distribution and metadata reference information.

Knowledge base is a the separate directory which resides in the server and documents are stored in the file format. Information can be further organized by grouping files together into logical stacks. Search folder empowers the user to organize files independent of their location on disk by specific set of criteria, formulated as search queries. Navigation and filter allow quick sorting and organization of files on the server machine.

OncoSys knowledge base will scan a set of documents written in a natural

language. It will be modeled according to the document set for predictive classification purpose and will populate the database with search index with metadata. This will also have the features to dynamically update the clinical knowledge. Creation of knowledge base uses database along with the file system. Only authorized users are allowed to access the database.

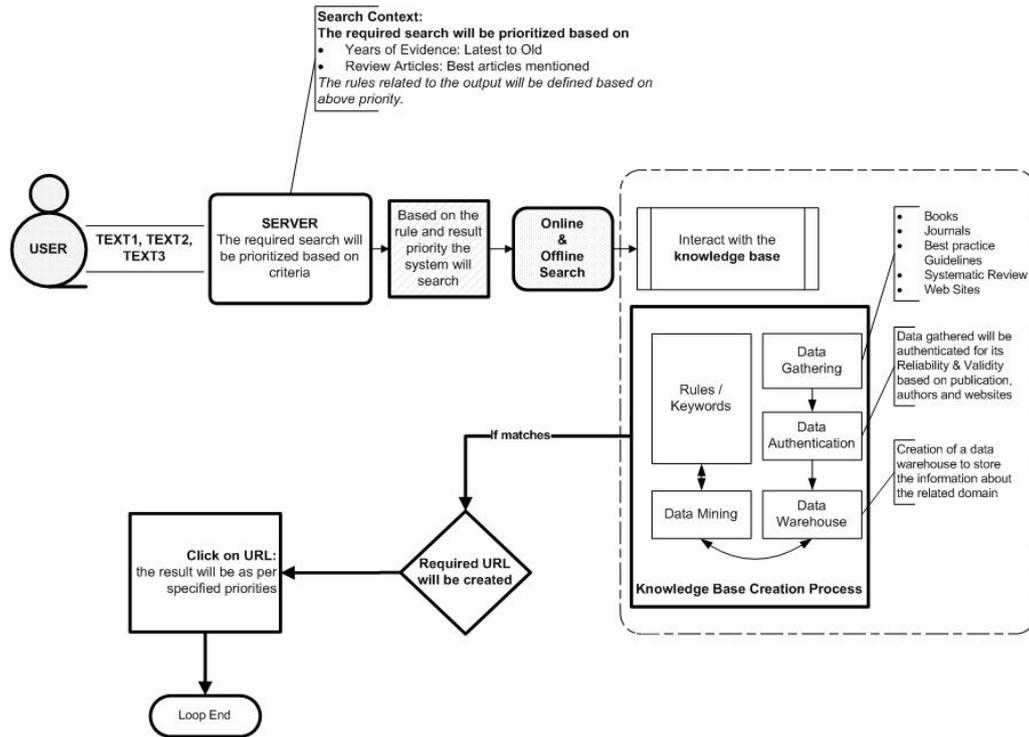


Figure 4.2: Model of Knowledge Base and Search Engine

Figure 4.2 represents the workflow model of knowledge base and search engine. The workflow is done in such a manner that when the user seeking the clinical knowledge related to breast and cervical cancer has to assign the keyword in the search text box under the search engine module. Inference engine of the system will act as an intermediary between the user and the system. It will accept and forward the request of the user in terms of keyword to the system for the required search. Once the system receives the command request, it will start searching the clinical knowledge based on rules/criteria specified

during the designing phase of search engine (section 4.3.3.2 of methodology). The system will first go for both offline and online search and if the request matches with the available knowledge (created during knowledge creation process 4.3.3.3) into the system the search result will be shown on user desktop screen.

SQL stored procedure are used to enhance the efficiency of the database access using **TSQL**. **XHTML** is used for client side scripting to make interface work dynamic and speedup the working of the system.

The search algorithm is designed using text analysis to extract the best document and list them below where the user has to click the document to view it. The text analytics is a set of linguistic, lexical, pattern recognition, extraction, tagging/structuring, visualization and predictive method. This can also be used to retrieve best practice. These techniques and processes are used to manage the domain knowledge, practices, and medications that are presented in the knowledge base in terms of text that are impermeable to automated processing.

4.3.3.4 Case Base Module

Case Base Module is included to assist the user in automatically storing the patient information and knowledge access by the users in their profile under the module. The stored information can be retrieved and used by login in to this module with proper authentication. The information provided by this module can be used by the user for the follow-up events as well as in the provision of continuity of patient care.

4.3.3.5 Statistical Module

Statistical Module is also included into the system because it is considered to be an important element for quality decision making in cancer control and prevention program and also for reporting to various governmental and non-governmental agencies. This module is found to be missing in the existing CDSS in practice. The following statistical analyses are incorporated into the module:

- Total number of patients based on Gender, Marital Status and Religion
- Total number of patients based on their geographical area
- Number and proportion of cancer by broad age group(0-14,15-34,35-16,65+)
- Number and proportion of cancer by five year age interval
- Trends in total number of cancer registered
- Number and relative proportion of cancer based on different method of diagnosis
- Number and relative proportion of microscopic diagnosis across different types of microscopic diagnosis
- Number and Relative proportion according to broad group of treatment
- Total number of patients treated
- Total number of treatment procedure performed and procedures/patient ratio
- Number and Relative proportion according to the treatment given

- Number and relative proportion based on ICO-Oncology Codes both Topography and Morphology

4.4 Security Issues built in OncoSys

Protection of health information should be a primary concern during the development and implementation of any IT application in a hospital setup. These issues have been addressed and measures taken during the design, development and implementation of OncoSys.

User authentication is achieved at database scope through User name and Password. If the user requests a page that requires authenticated access and user has not previously logged onto the site, then the user is redirected to a configure log-in page. The log-in page prompts the user to supply credential - typically a user name and password. These credentials are then passed to the server and validated against a user store, such as **SQL server database**. After the user credentials are authenticated, the user is redirected to the originally requested page.

To guarantee health information of the patient stored during registration of the patient **Transactions** are implemented. This ensures atomicity, consistency and durability of records.

Form Authentication Module is introduced to gain form level authentication and to avoid unauthorized user accessing the system. Form authentication uses an authentication ticket that is created when a user logs in to a site and then it tracks the user throughout the site. However, **ASP.NET version 2.0** supports form authentication, which results in the ticket being passed in query strings.

SHA1 (Secure Hashing Algorithm) is implemented for secure storage

of user data. It encrypts the data and sends throughout the network to prevent sniffing. Sniffing is observing the packets passing by a network and stealing data from a network usually in the form of password and ID name.

There is a possibility that an attacker can pass a malicious SQL command into the database. To avoid such SQL injection attack at the database side, precompiled stored procedures are used.

Weak input validation is common vulnerability that could allow the application to get affected with numerous injections attack. To avoid the injection attacks the form level data validation is incorporated into the system.

4.5 Architecture

OncoSys is a 3-layer application with User Interface, Business Logic layer and Data layer. Data access layer is incorporated using ADO.NET to communicate with database. The system requires the following to run the application:

4.5.1 User Interface

The external users are physicians and administrators of the hospital. The oncologist can access patient information and domain knowledge and also the statistics related to patient visit and treatment. The administrators will have an account to monitor the entire system in relation to access and use of the system by the end users.

4.5.2 Hardware Interface

The interface for accessing the system is personal computers of physicians and administrator. The PCs may be laptop or desktop connected to local server. Apart from these no other hardware is required.

4.5.3 Software Interface

MS SQL 2005, a database server to store the information. The Web Server i.e. IIS to respond to the users request from their browsers like Mozilla Firefox, Internet Explorer etc to access the system and its content. End users operating system can be any version of Windows, Linux or any other which supports TCP/IP protocols.

4.5.4 Communication Interface

The communication interfaces is a local area network connected to local hosting server.

4.6 User Acceptance Test

In the process of implementing OncoSys, a user acceptance test was conducted among the 16 expert Oncologists. The test was intended to determine the acceptance of OncoSys among the end users and also to identify their further requirement. A validated user acceptance test questionnaire was used to collect the response from the Oncologists. The questionnaire consists of 16 questions related to the user interface, 5 questions about the decision making interface and 10 related to the documentation and statistical interface of OnsoSys (Refer. Annexure.2). The OncoSys was demonstrated to the Oncologists where the related features and functionalities were explained to them. The questionnaire was then probed to the Oncologists and asked them to mark their response in a scale from Strongly Agree (5) to Strongly Disagree (1). The SPSS version 16.0 was used to analyze the collected data. The result of the test are represented in term of percentage and mean score.

4.7 Ethical Consideration

This research was not possible without the feedback and suggestions of the Oncologist. To ensure the respondents safety certain ethical issues were addressed during the research process, such as privacy and confidentiality of information supplied by them. In order to secure the consent of the selected participants, the researcher relayed all important details of the research including its aim, objectives and significance. These explanations helped the respondents to understand their contribution in completing and achieving the goal of the research. During the process of data collection, the respondents were also advised that they could express their unwillingness to be a part of this research. The investigator had taken all the concerned not to force participants to be the part of the research. The investigator had assured the respondents not to disclose their identification details in the research but the relevant data for answering the research questions.