

## RESEARCH COMMUNICATION

# Unified Modeling Language (UML) for Hospital-based Cancer Registration Processes

Naomi Shiki<sup>1</sup>, Yuko Ohno<sup>1\*</sup>, Ayumi Fujii<sup>2</sup>, Taizo Murata<sup>2</sup>, Yasushi Matsumura<sup>2</sup>

## Abstract

**Objective:** Hospital-based cancer registry involves complex processing steps that span across multiple departments. In addition, management techniques and registration procedures differ depending on each medical facility. Establishing processes for hospital-based cancer registry requires clarifying specific functions and labor needed. In recent years, the business modeling technique, in which management evaluation is done by clearly spelling out processes and functions, has been applied to business process analysis. However, there are few analytical reports describing the applications of these concepts to medical-related work. In this study, we initially sought to model hospital-based cancer registration processes using the Unified Modeling Language (UML), to clarify functions. **Methods:** The object of this study was the cancer registry of Osaka University Hospital. We organized the hospital-based cancer registration processes based on interview and observational surveys, and produced an As-Is model using activity, use-case, and class diagrams. After drafting every UML model, it was fed-back to practitioners to check its validity and improved. **Results:** We were able to define the workflow for each department using activity diagrams. In addition, by using use-case diagrams we were able to classify each department within the hospital as a system, and thereby specify the core processes and staff that were responsible for each department. The class diagrams were effective in systematically organizing the information to be used for hospital-based cancer registries. Using UML modeling, hospital-based cancer registration processes were broadly classified into three separate processes, namely, registration tasks, quality control, and filing data. An additional 14 functions were also extracted. Many tasks take place within the hospital-based cancer registry office, but the process of providing information spans across multiple departments. Moreover, additional tasks were required in comparison to using a standardized system because the hospital-based cancer registration system was constructed with the pre-existing computer system in Osaka University Hospital. Difficulty of utilization of useful information for cancer registration processes was shown to increase the task workload. **Conclusion:** By using UML, we were able to clarify functions and extract the typical processes for a hospital-based cancer registry. Modeling can provide a basis of process analysis for establishment of efficient hospital-based cancer registration processes in each institute.

**Key Words:** Hospital-based cancer registry - unified modeling language (UML) - business modeling

*Asian Pacific J Cancer Prev*, 9, 789-796

## Introduction

Hospital-based cancer registries represent information sources for population-based cancer registries. These also provide a resource for evaluating treatment performance at relevant medical facilities, which is useful information to improve the quality of clinical care. Therefore establishing the processes for hospital-based cancer registry is an urgent issue. In Japan, hospital-based cancer registries were suggested as one of the functions of designated cancer care hospitals (Ministry of health and labor, 2008). As a result, there has been a steadily increasing number of medical facilities that implement hospital-based cancer registries. In order to prepare hospital-based cancer registries, Center for Cancer Control

and Information Services in the National Cancer Center has assumed a central role in producing standardized registration formats and in conducting practical cancer registration seminars (Nishimoto, 2008; Okamoto, 2008; Sobue, 2008).

There is, however, little systematic information available on work content, manpower, required time, management techniques, etc., which is necessary for successfully implementing cancer registries. Clarifying the functions and labor needed for a hospital-based cancer registry should be useful for the establishment of cancer registry at each medical facility.

In general, the work content of hospital-based cancer registry involves multiple departments such as clinical, pathology, and medical information. Therefore it is

<sup>1</sup>Dept. of Mathematical Health Science, Graduate School of Medicine, <sup>2</sup>Dept. of Medical Information Science, Osaka University Hospital, Osaka University, Osaka 565-0871, Japan \*For correspondence: ohno@sahs.med.osaka-u.ac.jp

necessary to organize these complex work processes. In Japan, standardization is inadequate (Nishimoto, 2008), and currently registration procedures at each medical facility differ greatly depending on the hospital-based cancer registration system in use and the progress status in converting medical records to electronic form. In order to clarify the needed functions and labor, it is also necessary to compare work processes between multiple departments and facilities. In order to clarify functions and to undertake comparative analyses, it is necessary to model the subject of investigation.

For visualizing hospital-based cancer registration processes, we adopted Unified Modeling Language (UML), which in recent years has received attention in process analysis with business modeling techniques. UML is a unified language that represents organized processes and systems as diagrams (Penker et al., 2000). Because of its generality, UML is applied not only to systems development but also to a variety of other fields; there have been many reports on the utility of UML in process analyses involving the clarification of work processes and functions (Knape et al., 2003; Kumarapeli et al., 2007). There are however, very few reports of using UML modeling for cancer registration processes (Lyalin and Williams, 2005), and no reports with respect to hospital-based cancer registration processes. In this study we sought to model hospital-based cancer registration processes using UML and clarify the various functions involved.

## Materials and Methods

For hospital-based cancer registries, a standardized registration format has been established. However, at medical facilities already conducting registrations, items and contents vary widely depending on the facility. The flow of processes involved in cancer registrations differs greatly depending on the medical record management system in use and the degree of progress in the conversion of medical records to an electronic format. Currently, these processes are introduced and implemented separately by each medical facility.

In view of the difficulty of establishing a "standardized" system for hospital-based cancer registries, in this study, we sought to understand the implementation of processes involved in cancer registry using a case study, and then to model hospital-based cancer registration processes using UML. In this study the subject case was the hospital-based cancer registry at Osaka University Hospital with 1076 beds (hereafter referred to as "the hospital"). The hospital is a specialty hospital and medical facility that endeavors to become a designated cancer care hospital. Along with the other hospitals in Japan, the hospital already has its own hospital information system and medical payment computer system.

### *Process of UML modeling*

We conducted an interview and observational survey on work content with those responsible for implementing hospital-based cancer registries. The survey was conducted in August, 2008. We modeled the obtained

information with UML and sought to specify the hospital-based cancer registration processes and the functions demanded in cancer registry at the hospital.

UML is an application built by unifying a model notational system based on object-oriented development techniques, and is used in NEEDS engineering and process analysis (Rumbaugh et al., 2005; OMG, 2008). There are 13 types of diagrams in UML. By representing organized processes and systems as diagrams, UML visualizes work contents and systems. In this investigation, we used use case, activity, and class diagrams to model the hospital-based cancer registration processes as the As Is model.

First, we produced activity diagrams of the hospital-based registration processes. An activity diagram is a so-called process flow chart. In an activity diagram, the processes performed by different organizations or systems are partitioned by borders called swim lanes. In addition, information and systems treated within processes are displayed as entities. Hospital-based cancer registration processes span across multiple departments and deal with various data types such as patient information. In conducting functional analyses of hospital-based cancer registration processes, it is necessary to organize complex processing steps. We therefore decided to first produce activity diagrams.

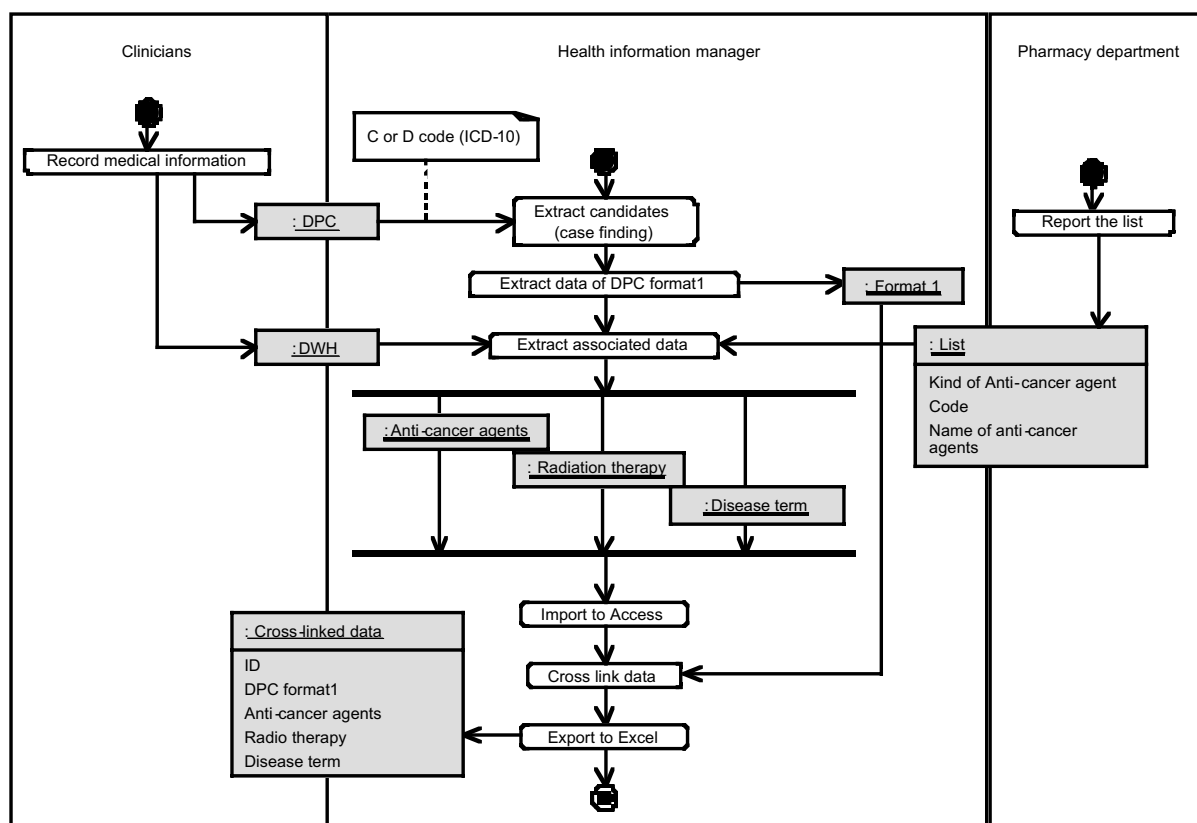
We then produced use case diagrams based on work contents clarified by the activity diagram. Use case diagrams provide a broad overview of the function and range of a system. Our goal with a use case diagram was to visualize the functions needed by the hospital-based cancer registry and the staff implementing registrations. We first extracted the work contents forming the core of hospital-based cancer registration as use cases. We then extracted the staff involved in hospital-based cancer registration processes as actors. With the extracted use cases and actors we produced a use case diagram.

Finally, using class diagrams we graphically represented the information utilized in hospital-based cancer registration processes, and the relationships that existed between the various types of information. Classes are abstractions of people or things that are handled within various processes. For hospital-based cancer registries, various types of data are employed. In this study, class diagrams were chosen in order to specify useable information and information needed for work processes. In describing the UML, we met successively with implementers of hospital-based cancer registrations, who provided feedback to ensure content accuracy.

## Results

### *Composition of the hospital-based cancer registry office*

The hospital-based cancer registration office was established in the medical information department. Two health information managers (one full-time, one part-time) primarily attend to registration tasks as cancer registrars. The full-time health information manager places priority on carrying out hospital-based cancer registration duties. The part-time health information manager attends to registration tasks in a helper capacity during times when hospital-based cancer registration duties are overloaded.



**Figure 1. Activity Diagram for Registration Tasks (Case-finding)**

Besides of the two registrars, another health information manager (not cancer registrar) takes charge of a part of duties.

#### System overview

The hospital is constructing and enacting an original hospital-based cancer registry using an electronic medical records system. This involves a system in which the health information managers perform the registration tasks and complete the registry via confirmation from the attending doctors. The subjects targeted for hospital-based cancer registry are "hospitalized patients diagnosed with cancer on or after January 1, 2007."

At the hospital, both of electronic and paper medical records are employed. Nearly all information registered in the electronic medical records, including patient information, prescription contents, tests, and surgery information, etc., is stored in the Data Ware House (DWH), where it can be searched for specific clinical information. There is also a registration system for diagnosis procedure combination (DPC; Japanese case mix classification), which is also searchable for DPC information. Pathology reports are kept independent of the hospital information system but the report database can be searched with the cooperation of the pathology department.

For the electronic medical records system, a form (created with Microsoft Word) and template (created with Template Master) for use in hospital-based cancer registrations have been prepared (Fujii et al., 2008). Moreover, while The TNM Classification of Malignant Tumors is an item for which registration errors are likely to occur, mistakes can be decreased by selecting stage from a number of choices made as small as possible due

to an input format using a hierarchical structure. In addition, information such as patient names can be obtained automatically from the electronic medical records. The registered data output is in the form of an XML file. This data is registered as is in the electronic medical records database and is stored in DWH, one patient one value one record, by overnight batch processing. Furthermore, information required for hospital-based cancer registration is extracted from the data stored in DWH, and after value code conversion, is stored in the hospital-based cancer registry database.

When filing data to the population-based cancer registry, the registered data extracted from the hospital-based cancer registry database are first read into HosCanR2.1 (Client server edition; free software of standardized registration system) to conduct quality control. After checking for errors with HosCan-R and correcting the mistakes, the data is then submitted in a specified format and posted to the population-based cancer registry.

#### UML modeling, Activity diagram

Figures 1-3 show the flow of hospital-based cancer registration processes using activity diagrams. From the organization of processes, we extracted broadly three processes; "registration tasks," "quality control," and "filing data to the population-based cancer registry" processes. In this study, "registration task" processes are shown in Figures 1 and 2, while "quality control" processes and "filing data to the population-based cancer registry" are shown in Figure 3.

In an activity diagram, each organization is partitioned by borders to organize the flow of and relationships

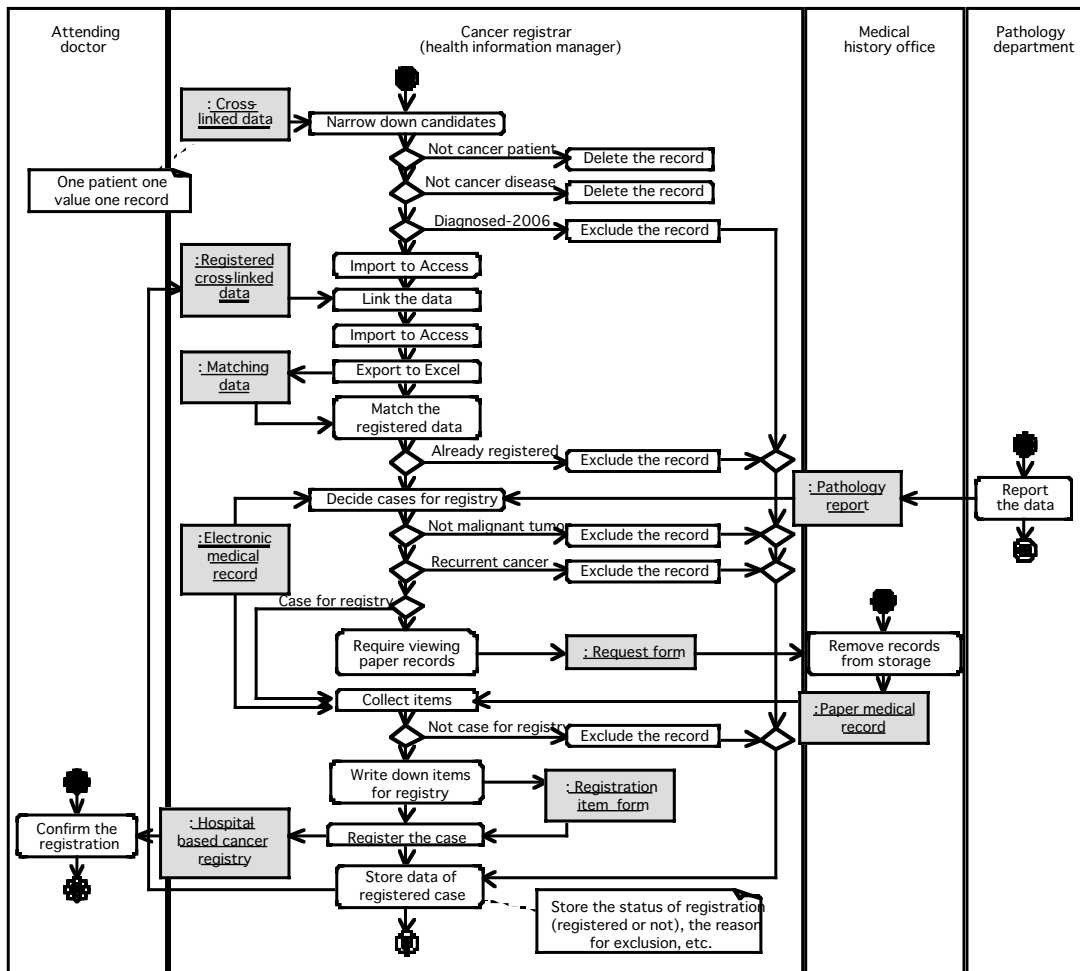


Figure 2. Activity Diagram for Registration Tasks (Narrowing Down Cases and Registration)

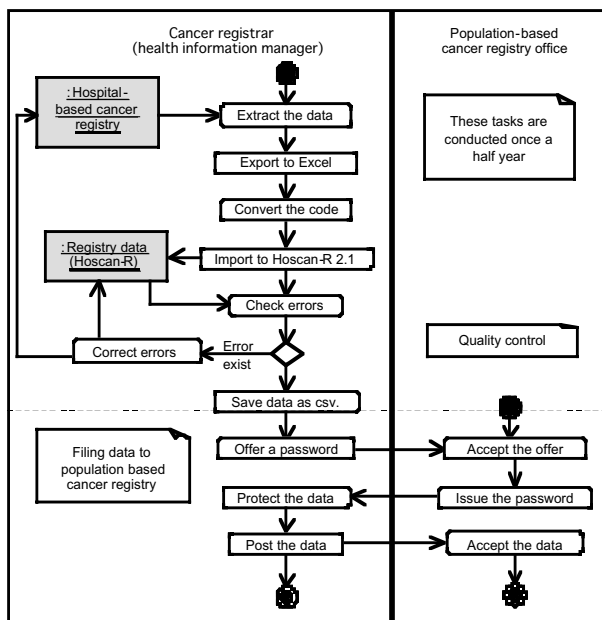


Figure 3. Activity Diagram for Quality Control and Filing Data to the Population-based Cancer Registry

between organizational processes. The data handled within work processes are shown as “entities,” and specify the exchange of information between organizations. The “registration tasks ” required task in the pharmacy department, pathology department and the medical history office. The work content therein primarily involved each

department providing data to the hospital-based cancer registry office.

The flow of the various processes is indicated by arrows, specific types of which include “fork nodes” and “join nodes.” This configuration shows that multiple processes are enacted in parallel and that subsequent tasks proceed only after the preceding ones are completed. For example in Figure 1, data extraction is performed on multiple occasions, but cross-linking tasks cannot proceed unless the extraction task has been completed. Also, when work content changes depending on conditions, these branching points are shown by a diamond-shape known as a guard conditions. For instance in Figure 2, when records exist in which the disease name overlaps for a patient in cross-linked data, the overlapping record is deleted. In activity diagrams for the hospital, there were many guard conditions present in portions of the registration task, which primarily accompanied a decision about whether or not a patient was considered case for registration. This judgment is a complicated task and requires a high level of expertise. Its complexity was also confirmed in the activity diagram.

In the hospital, an original registration system was used. Standardized registration systems such as HosCan-R proposed by the National Cancer Center, enable the simultaneous extraction, registration and management of data. However, in the case of original systems, tasks such as data extraction from hospital information and data cross-linking arise (Figure 1). It is also recognized that

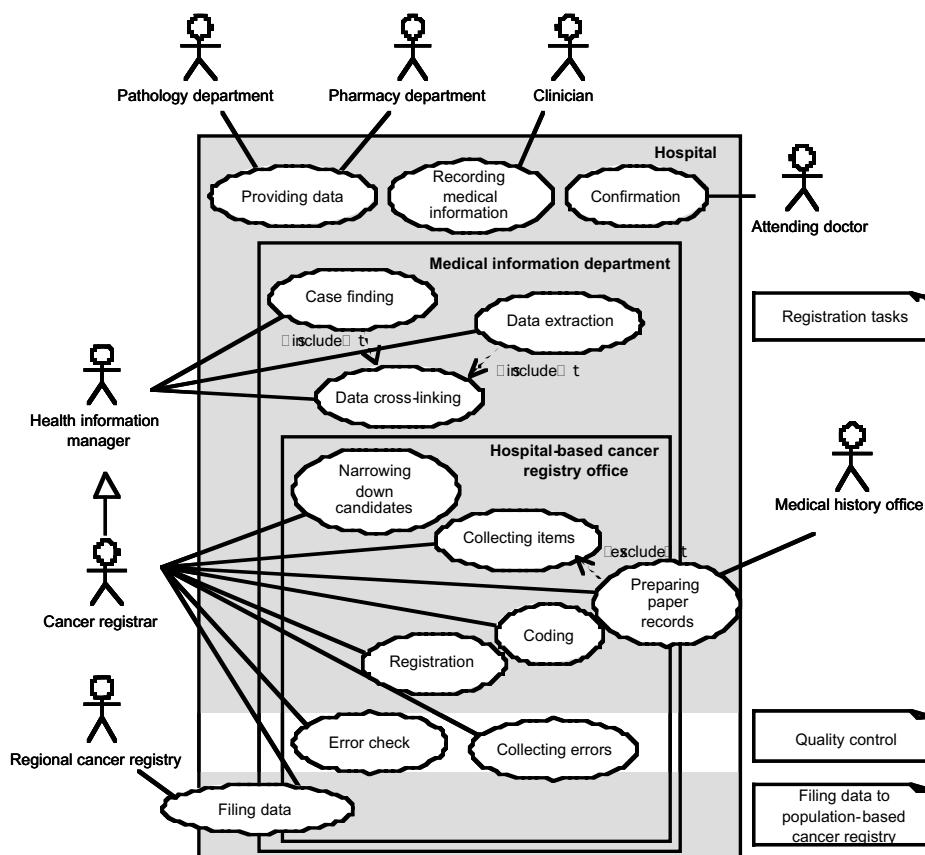


Figure 4. Use Case Diagram for a Hospital-based Cancer Registry

quality control needs to be performed separately (Figure 3). In addition, at the hospital, registration is done using electronic and paper medical records. Conditions requiring the viewing of paper medical records generated work tasks in the medical history office (Figure 2).

We note that follow-up of registered cases have not been implemented yet in the hospital, because the registry recently started.

*Use-case diagram*

Figure 4 shows the use case diagram for hospital-based cancer registration processes. On the basis of the activity diagram, we extracted 14 use cases. In addition to process implementers belonging to the hospital-based cancer registry office (health information managers as cancer registrars), we also extracted actors such as clinicians, pathology and pharmacy departments, and the medical history office. In addition, we extracted the population-based cancer registry office as an actor outside of the hospital.

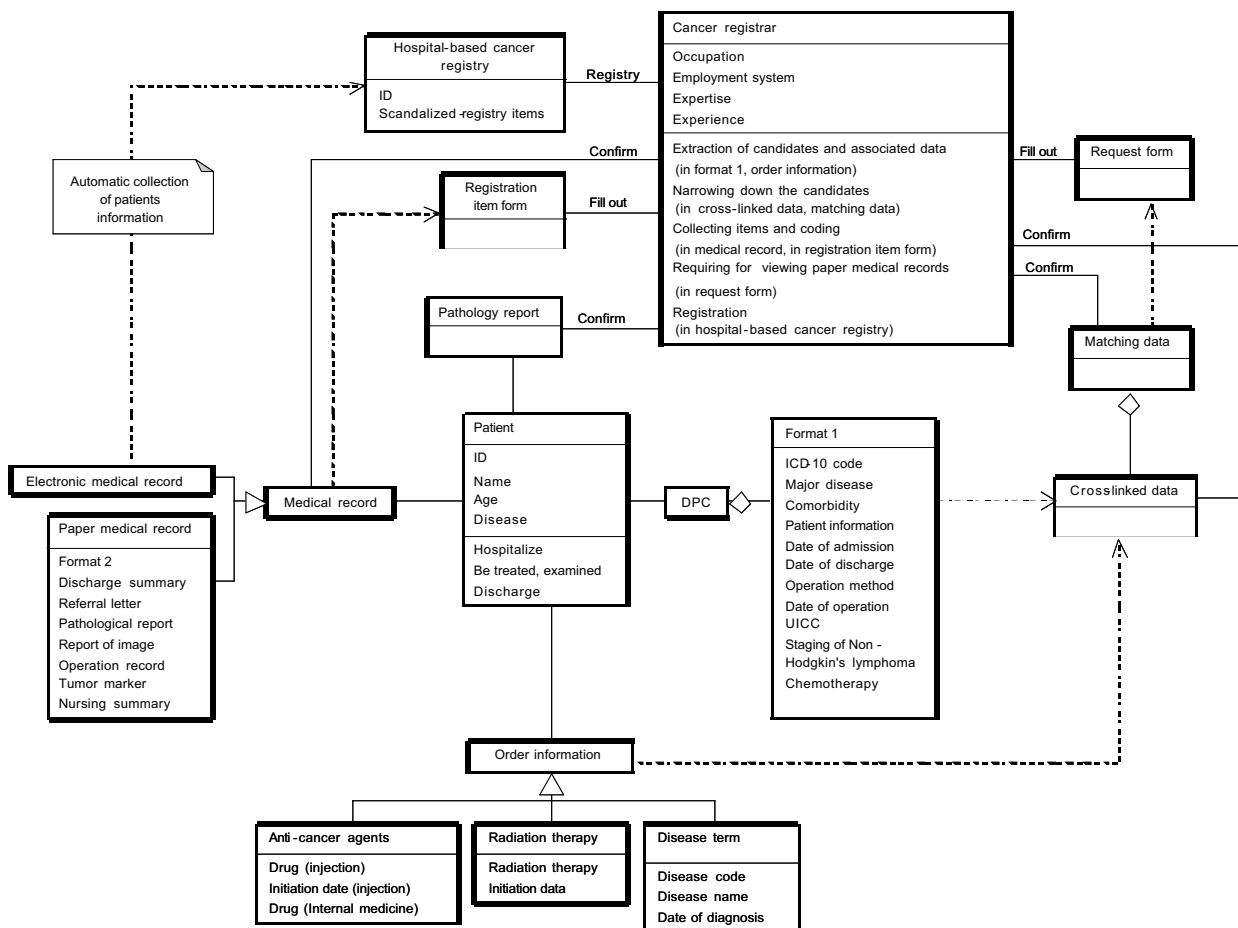
In a use case diagram, the relationships between use cases are specified. One of the relations present in use case diagrams is “include.” This is a relation in which a separate function is necessary when executing a particular function. For instance, “case-finding” and “data extraction” are essential when “data cross-linking”. In this case, “case-finding” and “data extraction” are shown to be in an “include” relationship with respect to “data cross-linking”. Separately, there is also an “exclude” relationship. This indicates a relationship in which under certain circumstances, a separate function is expanded whenever a particular function is executed. For example,

when performing the “collecting items”, “preparing paper records” is necessary only when examination of paper medical records is required. This sort of relationship is termed “exclude.”

The rectangular border in use case diagrams indicates the range of a system. In the present investigation we extracted the 3 layers of “hospital”, “medical information department”, and “hospital-based cancer registry office”, as systems. The use cases within each border represent the functions of each system. In addition, those responsible for enacting these functions are represented as the actors. The straight line connecting a use case and an actor indicates that the actor is responsible for enacting that function. For instance, the medical history office is responsible for “preparation of paper medical records” and is not in charge of any other use case. On the other hand, the cancer registrars are involved in many use cases. For hospital-based cancer registry, there is a need to decrease the clinician’s burden as much as possible. In the hospital, the clinicians were only involved in “recording medical information” and “confirmation,” which could be done during daily consultations.

*Class diagram*

Figure 5 shows the class diagram concerning information utilized in hospital-based cancer registrations. The definition of class includes workers and entities. Workers are staff who perform work tasks, and entities are passive objects operated by workers. In the present investigation, we extracted cancer registrars as workers. Their properties included occupation, employment system, expertise, and experience. These can be considered as



**Figure 5. Class Diagram for Information Utilized in Hospital-based Cancer Registration Tasks**

factors influencing the execution of hospital-based registration processes. In addition, the tasks shown in the activity and use case diagrams were suggested as operations. These can be considered the contents of the work performed by cancer registrars. Entities included DPC format 1, information for orders, clinical records, pathology reports, cross-linked data, matching data, request form (for viewing paper medical records), registration item form, hospital-based cancer registry, and other classes comprising these entities. The properties of each class are represented by the medical examination information possessed by each class (not shown in Figure 5).

A class diagram visualizes relationships between classes. Interclass relationships include “association,” “generalization,” “aggregation,” and “dependencies.” The present class diagram also confirmed a number of relationships. “Generalization” indicates a relation in which one of the related classes (superclass) is more general relative to the other (subclass). For example, information for orders is in a “generalization” relationship with information of anti-cancer agents, disease name, and radiation therapy. This relation is indicated by drawing a straight line from the subclass to the superclass and adding an open triangle marker at the end. Alternately, the “aggregation” interclass relationship indicates a relation in which a portion of the superclass is a subclass. For instance, matching data include data from cross-linked data, and so the relation between these classes is classified

as “aggregation.” In this case, a straight line is drawn from the subclass to the superclass and an open diamond marker is added at the end.

The class diagram shows that the process of extracting and narrowing down cases does not utilize pathological information, and that it is required to access the pathology report each time. This is because the pathology report is independent of the hospital information system. If, during the initial stages of registration tasks, the pathology data could be accessed and the data for cross-linked data could be linked, it is likely that the extracting and narrowing-down process could be simplified.

In deciding on cases for registration, medical records are examined and a decision is ultimately made on whether the patient will be registered. Then the necessary items are collected. During these processes, the informational resources particularly used as references include discharge summaries and letters of referral—sources for which clinical information is summarized. However, examination of the class diagram shows that letters of referral exist as a property of paper medical records. Viewing paper medical records requires filing a request for use, which can take considerable time.

**Discussion**

In the context of current demands for establishing processes for hospital-based cancer registry, it is necessary to clarify the functions and labor needed for these

processes. In general, hospital-based cancer registration consists of complex processing steps. In addition, the work contents differ depending on each medical facility. For these reasons, it is necessary to visualize and model the processes when analyzing registration procedures.

We therefore focused our attention on a business modeling technique using UML, a technique that has gained much attention in recent years. By employing UML, which uses graphic representation as the principle in which to model processes, it is possible to clarify process functions. Furthermore, using a unified language such as UML raises the prospect that such a model can be adapted to comparative analyses of processes conducted by multiple facilities.

By using activity diagrams, it was possible to represent the flow of registration processes, including the exchange of information between departments. This diagram seems appropriate for processes that involve many exchanges of data such as seen in hospital-based cancer registry. In addition, by organizing the results of interview surveys using this diagram, we were able to classify the hospital-based cancer registration processes broadly into three process types—"registration tasks," "quality control," and "filing data to the population-based cancer registry." These work contents can be viewed as functions that are necessary for the successful enactment of hospital-based cancer registry. Currently, it is claimed by Center of Cancer Control and Information Services in the National Cancer Center, that hospital-based cancer registration processes can be divided into "registration (including quality control)," "summarizing information", "providing information," and "follow-up" (Sobue et al., 2007). However, in the case of the hospital, "quality control" needs to be performed separately due to the existing surroundings of the system. In such cases, it is likely to be necessary to take up quality control as a core process separate from the "registration" process. By analyzing the current state of affairs using UML, it is possible to extract precisely what processes and functions are necessary, depending on the conditions at each medical facility.

Because implementation conditions differ at each medical facility, functional analyses appropriate for each condition may be necessary. Using activity diagrams, we were also able to represent changes in work content under a variety of conditions, and to extract particularly complex task areas. Activity diagrams could be used to simply and clearly model the flow of hospital-based cancer registration processes that span across multiple departments. We suggest these diagrams are effective for understanding the characteristics of the various processes.

Use case diagrams statically specify the relationships between the functions, the range and responsible implementers of a system. When applied to hospital-based cancer registry, we were able to view each of the hospital departments as a range of the system, and were able to define the work contents forming its core as use cases, and those responsible for the work as actors.

Hospital-based cancer registration processes were classified broadly into three processes using activity diagrams, and were divided further into 14 use cases in the hospital. In use case diagrams, it is possible to position

processes represented at border regions of a system as joint processes with other departments. In the case of the hospital, processes of "preparing of paper records" and "filing data" spanned across multiple departments. Another process outside of the hospital-based cancer registry office was "providing data", "recording medical information" and "confirmation."

The tasks of cancer registry require high expertise and workload. Thus, it is said that cancer registrars are needed from the view of reliability of data and efficiency in Japan (Nishimoto, 2008). In the hospital, cancer registrars mainly implement cancer registry so that clinicians' load is relatively small. To implement hospital-based cancer registry with the least possible burden on the clinic, it will likely be necessary to consider, for these processes that require coordination with other departments in particular, the efficient distribution of responsibilities and how they ought to operate ideally.

In a use case diagram, by defining the range of the system as the departments within the hospital, it is possible to extract processes that require coordination with other departments or that can be completed within a single department. Use case diagrams may be useful to considering the roles of each department in each task.

In class diagrams, we were able to summarize the information to be utilized and the relationships between the information. By considering cancer registrars as workers within the same class diagram, it is possible to simultaneously specify the factors involved in process execution, the information to be utilized, and the contents of the work.

In order to simplify the registration tasks, it is useful to have an information source in which the required information has been summarized. However, in the case of the hospital, it was necessary to view paper medical records. Moreover, while pathology report was useful in narrowing down cases and determining whether or not a patient was a case for registration, its use was limited. For hospital-based cancer registry, it has been recommended that multiple sources of information be used in order to prevent omissions and increase accuracy (Jensen et al., 1991). This suggests it is necessary to consider coordination with the hospital information system.

By considering those who implement the processes as workers, and the information handled as entities, class diagrams make it possible to see which processes require what information, as well as to identify what is lacking.

As seen above, the advantage of UML is its ability to simply and clearly specify complex processing steps by modeling on the basis of obtained information. In addition, UML can extract the functions demanded by a process, as well as identify problem areas (Knape et al., 2003; Kumarapeli et al., 2007). In investigating work issues, UML is extremely useful and applicable to hospital-based cancer registration processes. It may be possible to comparable analysis depending on facilities by modeling processes with UML, unified language. Moreover based on the modeled hospital-based cancer registration materials, it is possible that by conducting more detailed process analyses and quantitative task load surveys such

as time study, useful material may be provided in preparing hospital-based cancer registries in the coming years.

## Acknowledgements

This study was supported by a Health and Labor Sciences Research Grant for Clinical Cancer Research (H19-009) and a Grants-in-Aid for Scientific Research (B) (18310105).

## References

- Fujii A, Matsumura Y, Mineno T, et al (2008). On the construction and implementation of electronic medical records systems, University hospital information management conference suppl 2008, 138-140 (in Japanese).
- Jensen O.M, Parkin D.M, MacLennan R, Muir CS, Skeet RG (1991). Cancer Registration: Principals and Methods. Lyon, France, International Agency for Research on Cancer.
- Knappe T, Hederman L, Wade VP et al (2003). A UML Approach to Process Modelling of Clinical Practice Guideline for Enactment. *Stud Health Technol Inform*, **95**, 635-40.
- Kumarapeli P, De Lusignan S, Ellis T, Jones B (2007). Using Unified Modelling Language (UML) as a process-modelling technique for clinical-research process improvement. *Med Inform Internet Med*, **32**, 51-64.
- Lyalin D, Williams W (2005). Modeling cancer registration processes with an enhanced activity diagram. *Methods Inform Med*, **44**, 11-13.
- Ministry of Health and Labor, Department of health general affairs section, Office of promoting measures for cancer (2006). Guidelines for preparing cancer clinic coordinating center hospitals. (in Japanese) (<http://www.mhlw.go.jp/topics/2006/02/tp0201-2.html>. 2008.10.15 accessed)
- Nishimoto H (2008). Current status and future vision of hospital-based cancer registries in Japan. *Nippon Geka Gakkai Zasshi*, **109**, 101-4.
- Okamoto N (2008). A history of the cancer registration system in Japan. *Int J Clin Oncol*, **13**, 90-96.
- OMG. UNIFIED MODELING LANGUAGE. (<http://www.uml.org/>. 2008.10.15 accessed)
- Penker M, Eriksson H.E (2000). Business Modeling With UML: Business Patterns at Work. New York, John Wiley & Sons.
- Rumbaugh J, Jacobson I, Booch G (2005). The Unified Modeling Language Reference Manual. Second Edition. Addison-Wesley Pub(sd).
- Sobue T (2008). Current activities and future directions of the cancer registration system in Japan. *Int J Clin Oncol*, **13**, 97-101.
- Sobue T, Tsukuma H, Okamoto N, Ajiki W (2007). Handbook on Population-based Cancer Registration in Japan (fifth revision).