

The medical data is organized into lists that are based on the MML structure and that include disease name information, laboratory information, and progress reports. The fifth column in the list indicates the medical institution where the patient was treated. "Maiko Hospital" in this column indicates a Kyoto area medical institution and data which was uploaded to the Kyoto EHR system in CDA rel.1 (MML3.0) format. On the other hand the data of "Haniwa hospital" was uploaded to Miyazaki EHR system in MML2.3 format. In this way, this super site is able to merge and display data from EHR systems in different regions.

Observations

In this study, we considered, and verified by testing, a mechanism for integrating local EHR systems and providing medical cooperation that spans multiple regions. We constructed a super site (Super Dolphin) with data mapping functions for the purposes of matching patient IDs from different regions using an upper-level directory structure, and of compatibility between different data structures. Japan Medical Network Association was established as the operating body for operation of this super site. Although the upper-level directory structure is simple, it is highly universal and is expected to provide large benefits for medical cooperation between regions within a country and with regions where a unique national-level ID cannot be used. In this study, we also succeeded in mapping between different data structures. As described before, in this case conversions were only performed in one-to-one combinations (MML2.3 and CDA rel.1(MML3.0)). Naturally in order to use this mechanism to handle a broad range of clinical data in a versatile manner, the development of a correspondence table for the various standards will be necessary. In this case, determining how to coordinate the different levels of detail in the information may be a larger problem than mapping. However, the structures of the minimum necessary data that needs to be recorded for medical purposes do not differ greatly, and we believe that there will be no serious problems. In the future, we intend to increase the number of local EHR systems which participate in this super site and verify the effectiveness of this fundamental infrastructure, working towards achieving a national-level EHR.

References

1. Canada Health Infoway <https://www.infoway-inforoute.ca/about-infoway/news/open-letter-to-canadians/canadas-ehr-progress>, last accessed Jan 16 2011
2. NHS online http://www.nhs.uk/attachments/NHS_Reprint_P1.pdf, last accessed Feb 16 2011
3. Examples of EHR in Japan (Japanese) [http://www.wam.go.jp/wamappl/bb13GS40.nsf/0/976ec0d3117a45cf4925708a001e8e62/\\$FILE/siryu2-1.pdf](http://www.wam.go.jp/wamappl/bb13GS40.nsf/0/976ec0d3117a45cf4925708a001e8e62/$FILE/siryu2-1.pdf), last accessed Feb 16 2011
4. The number of Americans who moved declined sharply last year http://www.cleveland.com/business/index.ssf/2009/04/the_number_of_americans_who_mo.html, last accessed Nov 2 2010
5. Statistics Bureau and the Director-General for Policy Planning and Statistical Research and Training Institute, Ministry of Internal Affairs and Communications of Japan. <http://www.stat.go.jp/english/index.htm>, last accessed Feb 16 2011
6. Takada, A., Guo, J., Tanaka, K., Sato, J., Suzuki, M., Suenaga, T., Kikuchi, K., Araki, K., and Yoshihara, H., Dolphin project—Cooperative regional clinical system centered on clinical information center. *J. Med. Syst.* 29(4):391–400, 2005.
7. Halamka, J. D., Making the most of federal health information technology regulations. *Health Aff.* 29(4):596–600, 2010.
8. Halamka, et al., Health care IT collaboration in Massachusetts: The experience of creating regional connectivity. *J. Am. Med. Inform. Assoc.* 12(6):596–601, 2005.
9. 2009 MIC communication use trend survey. http://www.soumu.go.jp/main_content/000064217.pdf, last accessed Feb 20 2011
10. Eichelberg, M., et al., A survey and analysis of electronic healthcare record standards. *ACM Comput. Surv.* 37(4):277–315, 2005.
11. Araki, K., Ohashi, K., Yamazaki, S., Hirose, Y., Yamashita, Y., Yamamoto, R., Minagawa, K., Sakamoto, N., and Yoshihara, H., Medical markup language (MML) for XML-based hospital information interchange. *J. Med. Syst.* 24(3):195–211, 2000.
12. Guo, J., Takada, A., Tanaka, K., Sato, J., Suzuki, M., Suzuki, T., Nakashima, Y., Araki, K., and Yoshihara, H., The development of MML (Medical Markup Language) version 3.0 as a medical document exchange format for HL7 messages. *J. Med. Syst.* 28(6):523–533, 2004.
13. Health Level 7 (HL7). <http://www.hl7.org/>, last accessed Nov 2 2010
14. Yoshihara, H., Applications of electronic medical records to local cooperative health care—Cooperative health care on the national level in Japan by linking EHR systems. *J. Jpn. Med. Assoc.* 135(9):1989–1992, 2006.
15. Maldonado, J. A., et al., LinkEHR-Ed: A multi-reference model archetype editor based on formal semantics. *Int. J. Med. Inform.* 78(8):559–570, 2009.
16. ASTERIA. <http://www.infoteria.com/en/>, last accessed Nov 2 2010
17. Rhapsody. <http://www.orionhealth.com/products/rhapsody/>, last accessed Aug 12 2010
18. Yoshihara, H., Minagawa, K., and Tanaka, H., The concept of Global Medical History Database (GMHD) and Its prototype. *Jpn. J. Med. Inf.* 17(3):237–240, 1997.
19. NPO Japan Medical Network Association <http://www.ehr.or.jp/>, last accessed Feb 20 2011
20. Higo-Med Project. <http://133.95.89.5/dolphin/>, last accessed Feb 20 2011
21. Haniwa net. <http://www.haniwa-net.jp/>, last accessed Feb 20 2011
22. HOT project. <http://www.ocean.shinagawa.tokyo.jp/hot/>, last accessed Feb 20 2011
23. Maiko-Net. <http://www.e-maiko.net/>, last accessed Nov 2 2010
24. Japan Gigabit Network. <http://www.jgn.nict.go.jp/english/index.html>, last accessed Nov 2 2010

Development of Fundamental Infrastructure for Nationwide EHR in Japan

Tadamasa Takemura · Kenji Araki · Kenji Arita ·
Toshiaki Suzuki · Kazuya Okamoto · Naoto Kume ·
Tomohiro Kuroda · Akira Takada · Hiroyuki Yoshihara

Received: 4 November 2010 / Accepted: 15 March 2011
© Springer Science+Business Media, LLC 2011

Abstract The movement of create medical information systems that is now taking place involves both progress in EMR (Electronic Medical Records)—computerization of records at hospitals and clinics, and also in EHR (Electronic Health Records) in which information is shared with individual regions. However, the geographical coming and going of people in modern society is extremely active. Naturally the places these people move to are not necessarily within the same region. For this reason, even if the basic unit for the health care supply system is in practical terms limited to the local level, if services are restricted to only one region, many persons may be unable to receive the benefits of health care cooperation. In this study, we constructed a mechanism for a medical cooperation system which links the EHR systems of individual regions and is able to create a one-patient, one-record system on the national level. In this paper, we will provide a report of this mechanism and of the 4-year operational trial.

Keywords Nation level EHR · Dolphin Project · XML data mapping · Directory service

Introduction

The movement of create medical information systems that is now taking place involves both progress in EMR (Electronic Medical Records) computerization of records at hospitals and clinics, and also in EHR (Electronic Health Records) in which information is shared with individual regions. A variety of trials have been carried out worldwide for this purpose, primarily in developed countries, and informatics is also receiving attention as an effective means of improving the efficiency of medical services in newly industrialized and developing countries as well. For example under the leadership of the state, Canada and England have invested at least 1.6 billion U.S. dollars [1] and 20 billion U.S. dollars [2] respectively. It is said that the United States will invest 20 billion U.S. dollars in switching to electronic medical documents. For EHR, many successful examples in sharing medical information within regions continue to be reported from around the world, and several EHR projects have been carried out in Japan as well [3].

However the geographical coming and going of people in modern society is extremely active. In the United States, 35 million people change their place of residence each year [4], and it is said that on average a person in Japan moves 5 times in his or her life [5]. Naturally the places these people move to are not necessarily within the same region. For this reason, even if the basic unit for the health care supply system is in practical terms limited to the local level, if services are restricted to only one region, many persons may be unable to receive the benefits of health care

T. Takemura (✉) · K. Okamoto · N. Kume · T. Kuroda ·
H. Yoshihara
Department of Medical Informatics, Kyoto University Hospital,
Kyoto, Japan
e-mail: takemura@kuhp.kyoto-u.ac.jp

K. Araki · K. Arita
Department of Medical Informatics, University of Miyazaki
Hospital,
Miyazaki, Japan

T. Suzuki
Japan DynaSystems Inc.,
Nagoya, Japan

A. Takada
National Cerebral and Cardiovascular Center,
Suita, Japan

cooperation. In this study, we constructed the actual fundamental infrastructure needed for local medical EHR, and carried out EHR projects in several regions of Japan [6], and during this study we found that it will be impossible to gain a comprehensive grasp of patient medical information at the national level only because no solution to the problems caused by movement of patients. Of course, we cannot use national identified number in Japan, so it make the solving problems more difficult.

Therefore when considering the future development of EHR, a mechanism for consolidating local-level medical information on the national level, as well as functions for data compatibility and other purposes, will be needed.

In this study, we constructed a mechanism for a medical cooperation system which links the EHR systems of individual regions and is able to create a one-patient, one-record system on the national level. In this paper, we will provide a report of this mechanism and of the 4-year operational trial.

Methods

As we discuss a mechanism for medical cooperation between regions, we will first describe the current conditions of inter-regional medical cooperation in Japan.

Local-level EHR

Many regions in the world have created EHR systems for managing patient medical data within that region, and many projects have been launched using these systems as hubs for coordinated health care and the provision of medical record [1, 7, 8]. For these purposes, it is necessary to ensure safe routes of information between the medical institutions and the system, and also to create a mechanism that allows patients to safely view medical data via the internet, which penetration rate is 78.0% in Japan [9]. The formulation and operation of an open standard for exchanging medical data from a wide variety of medical records are also important [10]. In Japan as well, there are many EHR systems operating in individual regions. In these cases, the systems are operated in a way that makes best use of the unique characteristics of each region. Data exchange is accomplished in a variety of ways, including direct connections to the hospital information systems of large scale hospitals, and exchange using MML (Medical Markup Language) [11, 12] or HL7 (Health Level 7) [13]. Because it is the local governments which are directly faced with a need for health care cooperation in the region, in many cases the systems are operated under the leadership of the local governments, and currently it is difficult to carry out activities that span multiple regions.

Construction of a mechanism for wide-area medical cooperation

As described above, attempts to integrate local EHR systems and carry out services over a wide area face a number of problems. One is data-level integration. Although some believe that collecting data using a single unified format is sufficient, this approach is not practical when one considers the current conditions in which many independent local EHR systems are operating, using various formats. The solution is data conversion (mapping) on the content level between different data structures.

Another problem is fragmentation among EHRs because of lack of national level patient's identification. It is thought that this problem can be resolved by assigning an internal upper-level ID at upper-level sites in place of the unique IDs used on the local level, and to assign the local IDs to these upper-level IDs (essentially assigning them to an upper-level directory structure) [14]. Following is a description of data mapping and the upper-level directory structure.

Data mapping

Absorbing differences in data structures can be accomplished by constructing a mechanism for XML (eXtensible Markup Language) data mapping. Figure 1 shows a concept diagram of XML data mapping. A document has a format showed in the left-hand side while another document has a format showed in the right-hand side. For example, the document on the left-hand side in Fig. 1 defines the patient ID as <ID>, while the document on the right-hand side defines it as <SocNum>. If these two are considered equivalent, they can be mapped so that they can be converted back and forth. In the same way, <given name> and <first name> is another example that is often seen. If the XML label and the data indicated by that label have the same code system, they can also be mapped. For example, if <disease> in the left-hand document contains an ICD-10 code, then it can be converted to the <ICD code> in the right-hand document.

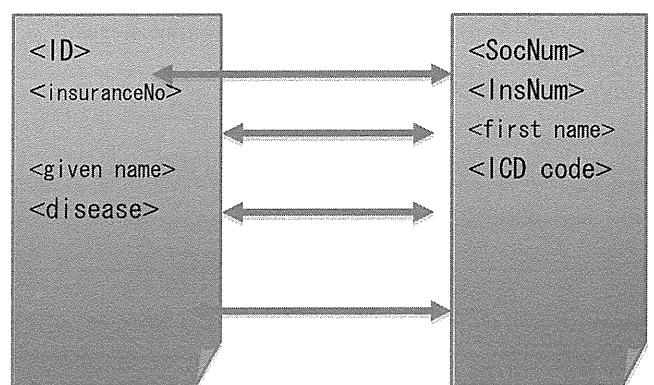


Fig. 1 XML data mapping

There have been reports of cases in which XML data mapping is used for bidirectional data conversion between EMR and EHR [15], and commercially products such as Asteria from the Infoteria Corporation (Japan) [16] and Rhapsody™ from Orion Health (New Zealand) [17] have been marketed as middleware intended for medical use. Use of these sorts of products makes data compatibility possible.

Upper-level directory structure

On the national level, if it is possible to issue and use a unique patient ID to each citizen on the national level, then such IDs can be used. However in many countries including Japan, use of these IDs in EHR is difficult. If a person is issued different patient IDs by multiple local EHR systems, it is necessary to understand that these different patient IDs actually indicate the same person. For this purpose, when a certain local EHR system issues a patient ID, an authorized organization can issue an upper-level patient ID for the national level, and can manage the links between patient IDs in multiple local EHR systems [18]. Using this mechanism, when a search for user data is performed using any local

Fig. 2 Concept diagram of upper-level directory structure

EHR system patient ID, it is possible to send a search request to other local EHR systems by means of the upper-level patient ID and return complete and integrated search results. Figure 2 shows a concept diagram of this process.

Actual system

In this study, we constructed a nationwide-capable EHR directory service (super site), with named “Super Dolphin” that includes the XML data mapping and upper-level directory structure described earlier, and verified that it is possible to link multiple local EHR systems together. Specifically, the subjects were two regions of Japan (Miyazaki and Kyoto) where EHR systems are actually operating. These two local EHR systems were connected to the super site that we constructed, and this super site was given the name “Super Dolphin”. The NPO Japan Medical Network Association, which was established since 2005 to implement nationwide EHR manages this super site [19].

Table 1 is an overview of the two local EHR systems which were the subjects of this test.

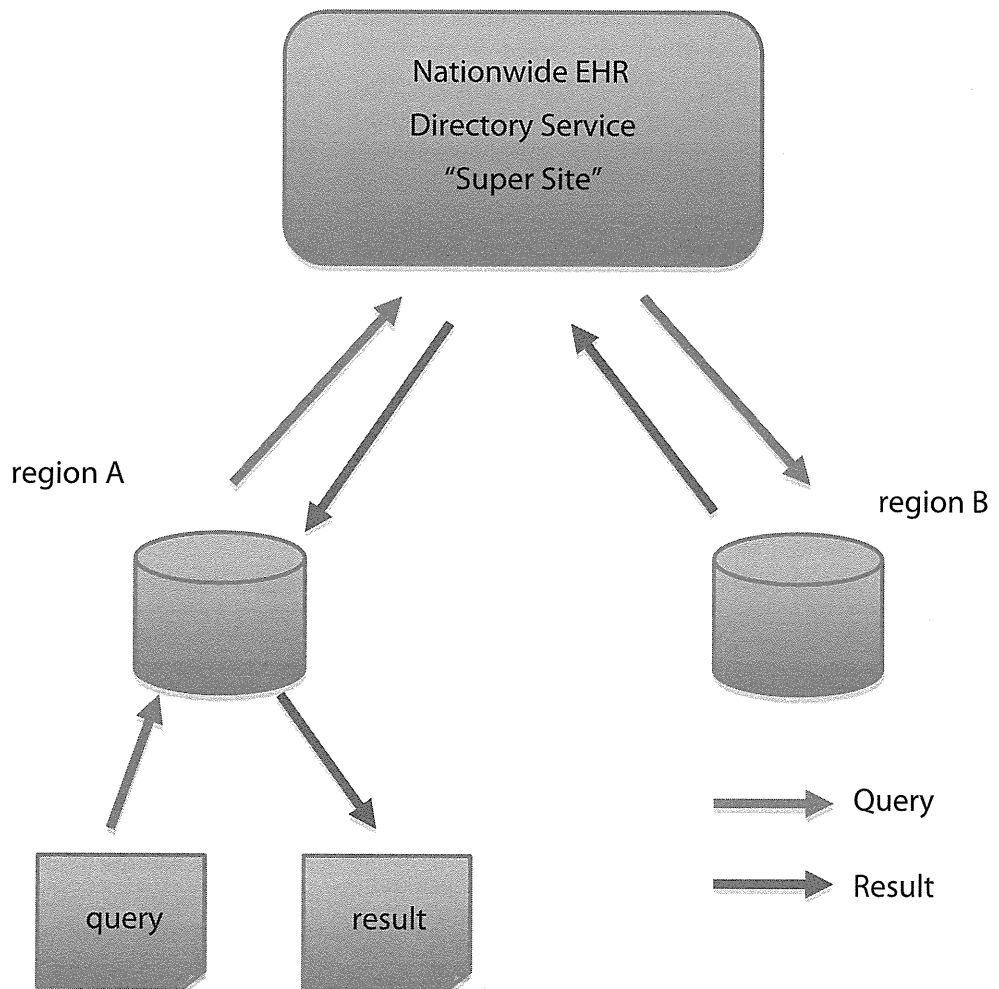


Table 1 Test conditions

-
- 1) Local EHR systems and using data formats
 Miyazaki EHR system (haniwa): Using MML2.3
 Kyoto EHR system (maiko): Using MML3.0 (CDA* rel.1 compliant)
- 2) Upper-level site: Super Dolphin
- 3) Paths: The two local EHR systems and Super Dolphin were connected by Japan Gigabit Network version2 (JGN2**)
 The internet is used for the communications route from the medical institutions to the center server
-

CDA* clinical document architecture

JGN2** is research and development testbed network operated by the National Institute of Information and Communications Technology (NiCT) in Japan

Both the Miyazaki and Kyoto EHR systems are EHR systems that were constructed within the framework of the Dolphin Project [2]. The Dolphin Project was proposed by Yoshihara et al. in 1997 [11], and took its first step toward becoming reality in 2000 as a R&D project of the Ministry of Economy, Trade and Industry, Japan. Subsequently, experimental EHR services were launched in two regions, Kumamoto [20] and Miyazaki [21], in December 2001 and remain in use today. Later, full-scale projects aimed at providing practical services were launched in Tokyo [22], Kyoto [23], and other major cities.

The framework of the Dolphin Project involves integrated management of the medical data stored in the EHR system central server under a certain level of security. This allows medical practitioners to centrally view the medical data of patients who have concluded treatment agreements, and allows coordinated medical care. Patients can also view their own medical data (electronic record disclosure) and can enter symptoms and other information into their own records. The central server is connected to clinics, hospitals, laboratory test services, pharmacies, home nursing-care stations, and other facilities, which can send information such as past histories, laboratory results, letters of introduction, and discharge summaries. This information is all integrated and stored for each patient. In addition to sharing of local treatment data, this information is also used as a backup for the record data of each medical institution. In the Dolphin Project, the data of each medical institution is sent to the central server using MML, HL7, or other data format and is stored by the server in a database. A web interface is provided to the patients and medical practitioners. At present, each region is currently operating an original system utilizing the above basic design but making use of the local characteristics. The scale of each local project is as shown below (Table 2).

In this study, each patient is issued a unique patient ID in the local EHR system where person wants to receive service. Using this ID, the patient is able to view patient's own medical information within the region. When a patient wants to view his/her own medical information from another region, by linking the patient IDs from multiple

local EHR systems, Super Dolphin allows medical information from different regions to be viewed.

When a search for medical information is performed on the Miyazaki or Kyoto system, first a query is sent to the database of that local EHR system using the patient ID as the key. At the same time, the local EHR system sends a query to Super Dolphin to check whether or not that patient ID is linked with patient IDs in other regions. As a result of this query, if the patient ID is found to be linked to an EHR system patient ID in another system, Super Dolphin uses this link information to request a search. The obtained data is converted to the data structure used by the data center which sent the request, and displayed. Communications between each EHR system and the super site utilize a local area network that uses the JGN2 network (Japan Giga Network version2) [24] provided jointly by the Ministry of Internal Affairs and Communications (MIC) and by NiCT. For local EHR systems and users, communication uses SSL with security functions utilizing Certification Authorities. The overall configuration is shown in Fig. 3.

Results

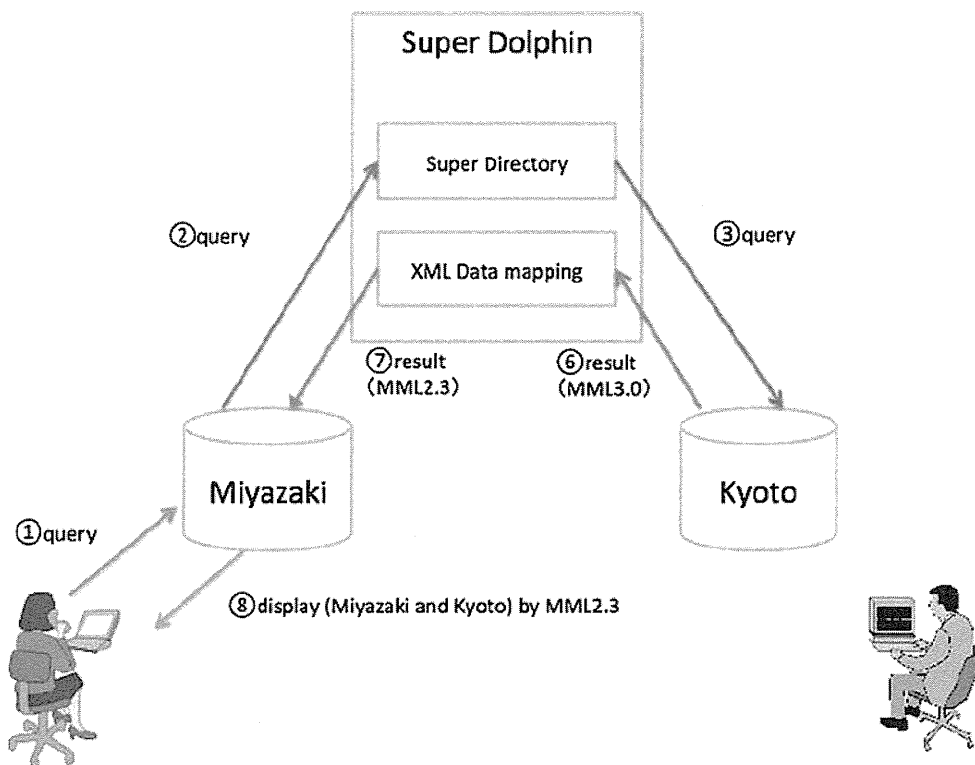
Figure 4 shows an example of the results from display of patient medical data.

Table 2 The scale of each local project

	Miyazaki	Kyoto
Registered patients	1078	1,100
Registered medical institutions	84	5
Registered physicians	478	2,000
Monthly views (physicians)	185	100
Monthly views (patients)	60	2,000
No. of documents sent (text)	1,600,000	7,000,000
No. of documents sent (images)	85,000	86,000
Year started	2002	2007

Measurement date is 30 Oct 2010

Fig. 3 Overall super dolphin configuration



The screenshot shows a web application interface for 'DMS(Dolphin MML Kaisei Server) Patient Frame'. It displays a patient's profile and a table of medical documents. The table has columns for document type, name, date, author, institution, and author category. Callouts indicate that some data is 'derived from Kyoto' and other is 'derived from Miyazaki'.

選択	文書名	作成日	作成者	作成施設	作成資格
<input type="checkbox"/>	患者基本情報	2005年03月02日	医8801	まいご病院	その他の医療従事者
<input type="checkbox"/>	患者基本情報	2005年03月02日	医8801	まいご病院	その他
<input type="checkbox"/>	病名	2005年03月01日	医8801	まいご病院	医師
<input type="checkbox"/>	病名	2005年03月01日	医8801	まいご病院	医師
<input type="checkbox"/>	プログレスノート	2005年03月01日	医8801	まいご病院	医師
<input type="checkbox"/>		2005年01月24日	医8801	まいご病院	医師
<input type="checkbox"/>	記録	2002年01月29日	増嶋 次郎	ほにわ病院 宮医大分院	医師
<input type="checkbox"/>	記録	2002年01月29日	増嶋 次郎	ほにわ病院 宮医大分院	医師

(a): check box,(b): medical document name, (c): document date, (d): document author name, (e): medical institution name, (f): author category(doctor, nurse, etc)

Fig. 4 Patient medical data

The medical data is organized into lists that are based on the MML structure and that include disease name information, laboratory information, and progress reports. The fifth column in the list indicates the medical institution where the patient was treated. “Maiko Hospital” in this column indicates a Kyoto area medical institution and data which was uploaded to the Kyoto EHR system in CDA rel.1 (MML3.0) format. On the other hand the data of “Haniwa hospital” was uploaded to Miyazaki EHR system in MML2.3 format. In this way, this super site is able to merge and display data from EHR systems in different regions.

Observations

In this study, we considered, and verified by testing, a mechanism for integrating local EHR systems and providing medical cooperation that spans multiple regions. We constructed a super site (Super Dolphin) with data mapping functions for the purposes of matching patient IDs from different regions using an upper-level directory structure, and of compatibility between different data structures. Japan Medical Network Association was established as the operating body for operation of this super site. Although the upper-level directory structure is simple, it is highly universal and is expected to provide large benefits for medical cooperation between regions within a country and with regions where a unique national-level ID cannot be used. In this study, we also succeeded in mapping between different data structures. As described before, in this case conversions were only performed in one-to-one combinations (MML2.3 and CDA rel.1 (MML3.0)). Naturally in order to use this mechanism to handle a broad range of clinical data in a versatile manner, the development of a correspondence table for the various standards will be necessary. In this case, determining how to coordinate the different levels of detail in the information may be a larger problem than mapping. However, the structures of the minimum necessary data that needs to be recorded for medical purposes do not differ greatly, and we believe that there will be no serious problems. In the future, we intend to increase the number of local EHR systems which participate in this super site and verify the effectiveness of this fundamental infrastructure, working towards achieving a national-level EHR.

References

1. Canada Health Infoway <https://www.infoway-inforoute.ca/about-infoway/news/open-letter-to-canadians/canadas-ehr-progress>, last accessed Jan 16 2011
2. NHS online http://www.nhs.uk/attachments/NHS_Reprint_P1.pdf, last accessed Feb 16 2011
3. Examples of EHR in Japan (Japanese) [http://www.wam.go.jp/wamappl/bb13GS40.nsf/0/976ec0d3117a45cf4925708a001e8e62/\\$FILE/siryou2-1.pdf](http://www.wam.go.jp/wamappl/bb13GS40.nsf/0/976ec0d3117a45cf4925708a001e8e62/$FILE/siryou2-1.pdf), last accessed Feb 16 2011
4. The number of Americans who moved declined sharply last year http://www.cleveland.com/business/index.ssf/2009/04/the_number_of_americans_who_mo.html, last accessed Nov 2 2010
5. Statistics Bureau and the Director-General for Policy Planning and Statistical Research and Training Institute, Ministry of Internal Affairs and Communications of Japan. <http://www.stat.go.jp/english/index.htm>, last accessed Feb 16 2011
6. Takada, A., Guo, J., Tanaka, K., Sato, J., Suzuki, M., Suenaga, T., Kikuchi, K., Araki, K., and Yoshihara, H., Dolphin project—Cooperative regional clinical system centered on clinical information center. *J. Med. Syst.* 29(4):391–400, 2005.
7. Halamka, J. D., Making the most of federal health information technology regulations. *Health Aff.* 29(4):596–600, 2010.
8. Halamka, et al., Health care IT collaboration in Massachusetts: The experience of creating regional connectivity. *J. Am. Med. Assoc.* 296(6):596–601, 2005.
9. 2009 MIC communication use trend survey. http://www.soumu.go.jp/main_content/000064217.pdf, last accessed Feb 20 2011
10. Eichelberg, M., et al., A survey and analysis of electronic healthcare record standards. *ACM Comput. Surv.* 37(4):277–315, 2005.
11. Araki, K., Ohashi, K., Yamazaki, S., Hirose, Y., Yamashita, Y., Yamamoto, R., Minagawa, K., Sakamoto, N., and Yoshihara, H., Medical markup language (MML) for XML-based hospital information interchange. *J. Med. Syst.* 24(3):195–211, 2000.
12. Guo, J., Takada, A., Tanaka, K., Sato, J., Suzuki, M., Suzuki, T., Nakashima, Y., Araki, K., and Yoshihara, H., The development of MML (Medical Markup Language) version 3.0 as a medical document exchange format for HL7 messages. *J. Med. Syst.* 28(6):523–533, 2004.
13. Health Level 7 (HL7). <http://www.hl7.org/>, last accessed Nov 2 2010
14. Yoshihara, H., Applications of electronic medical records to local cooperative health care—Cooperative health care on the national level in Japan by linking EHR systems. *J. Jpn. Med. Assoc.* 135(9):1989–1992, 2006.
15. Maldonado, J. A., et al., LinkEHR-Ed: A multi-reference model archetype editor based on formal semantics. *Int. J. Med. Inform.* 78(8):559–570, 2009.
16. ASTERIA. <http://www.infoteria.com/en/>, last accessed Nov 2 2010
17. Rhapsody. <http://www.orionhealth.com/products/rhapsody/>, last accessed Aug 12 2010
18. Yoshihara, H., Minagawa, K., and Tanaka, H., The concept of Global Medical History Database (GMHD) and Its prototype. *Jpn. J. Med. Inf.* 17(3):237–240, 1997.
19. NPO Japan Medical Network Association <http://www.ehr.or.jp/>, last accessed Feb 20 2011
20. Higo-Med Project. <http://133.95.89.5/dolphin/>, last accessed Feb 20 2011
21. Haniwa net. <http://www.haniwa-net.jp/>, last accessed Feb 20 2011
22. HOT project. <http://www.ocean.shinagawa.tokyo.jp/hot/>, last accessed Feb 20 2011
23. Maiko-Net. <http://www.e-maiko.net/>, last accessed Nov 2 2010
24. Japan Gigabit Network. <http://www.jgn.nict.go.jp/english/index.html>, last accessed Nov 2 2010

複数の機械学習手法を用いた退院時サマリからの 自動 DPC コーディング

岡本 和也*・内山 俊郎**・竹村 匡正*・足立 貴行**
糸 直人*・黒田 知宏*・内山 匡**・吉原 博幸*

Automatic DPC Code Selection from Discharge Summaries Using Several Machine Learning Methods

Kazuya OKAMOTO,* Toshio UCHIYAMA,** Tadamasa TAKEMURA,* Takayuki ADACHI,**
Naoto KUME,* Tomohiro KURODA,* Tadasu UCHIYAMA,** Hiroyuki YOSHIHARA*

Abstract A DPC code expresses a primary disease, a complication, and procedures, etc. In 2010, 1334 hospitals use DPC codes for calculations of medical fees. Since, in the hospitals, the medical fee of each case is calculated based on one DPC code, each case must be classified into one DPC code. However, the classification is difficult in some cases because patients sometimes have various conditions. Therefore, automatic DPC code selections using machine learning are being studied. Suzuki et al. evaluated automatic DPC code selections from discharge summaries using a vector space method. However, there are general machine learning methods except for the vector space method. Hence, we must evaluate other machine learning methods exhaustively for improvement of accuracy of automatic DPC code selections. Therefore, we evaluated automatic DPC code selections from discharge summaries using naïve Bayes method, SVM, concept base method, and another vector space method which is different from the vector space model used by Suzuki et al. We considered these machine learning methods as general ones. We also focus on characteristics of each machine learning methods on automatic DPC code selections and we utilize a method which combines some machine learning methods. First, the combining method estimates confidences of the machine learning methods bases on classification scores that the machine learning methods regard as classification evidence. Next, the combining method adopts the method whose confidence is highest. We compared accuracy of the methods using discharge summaries created in 2008 fiscal year in Kyoto University Hospital. As a result, SVM classified 72.2% of the cases into correct DPC codes though the vector space model utilized by Suzuki et al. classified 64.8% into correct DPC codes. Moreover the combining method classified 76.1% into correct DPC codes. In conclusion, we achieved significant improvement.

Keywords : DPC code selection, machine learning, discharge summary.

1. はじめに

2010年度現在、全国にある一般病床の約半数の病床にあたる1334施設が診断群分類（DPC）包括評価を用いた

入院医療費の定額支払い制度の対象病院となっている。これらの対象病院では、診療報酬請求のために、対象となる各入院患者に対して主傷病に加えて副傷病や処置等を表すコードであるDPCコードを1つ付与する必要がある。しかし、多様な容態を示す患者を1つのコードに割り当てることは困難である。実際、京都大学医学部附属病院では2008年4月から2009年3月の間に退院した延べ17993人の患者のデータのうち延べ17555人（97.6%）の患者データを診療情報管理士が監査し、1957人（10.9%）のDPCコードを医師の協力の下に変更を行い、結果としてDPCコードに基づく診療報酬の請求額が2億6500万円の増加となる変更が生じた[1]。そのようなDPCコーディング

生体医工学シンポジウム2010発表（2010年9月，札幌）
2010年7月30日受付，2010年10月1日改訂
Received July 30, 2010; revised October 1, 2010.

*京都大学医学部附属病院医療情報部
Department of Medical Informatics, Kyoto University
Hospital

**日本電信電話株式会社サイバースリユーション研究所
NTT Cyber Solutions Laboratories, NTT Corporation

の困難さのため、診療文書を解析対象とした機械学習による自動 DPC コーディングの可能性の検討が行われている。これまでに、Suzuki らは千葉大学医学部附属病院の退院時サマリを解析データとして、ベクトル空間法を用いた自動 DPC コーディングが試み、人手で付与された DPC コードとの一致率を調査した[2]。そして、DPC コード全 14 桁が一致するデータの割合は 56.5%、主傷病を表す DPC コード上位 6 桁が一致するデータの割合は 80.7% となったことが示された。

しかし、ベクトル空間法以外にも一般的な機械学習手法は複数存在することから、自動 DPC コーディングの精度向上のために他の機械学習手法を網羅的に検討する必要がある。

そのため、我々は同じく退院時サマリを解析データとし、機械学習手法としてナイーブ・ベイズ法、サポートベクターマシン (SVM)、概念ベース法、さらに、既存研究で用いられたベクトル空間法と異なるベクトル空間法 (平均ベクトル空間法) を用いて自動 DPC コーディングを行ない、それらの精度の比較を行う。

さらに、我々は自動 DPC コーディングにおいて各々の機械学習手法に特徴があることに注目し、各機械学習手法が分類の根拠としているそれぞれの分類スコアから分類結果に対する各機械学習手法の確信度を推定し、確信度の最も高い結果を選択する方法、すなわち、確信度の最大化に基づいて複数の機械学習手法を統合する手法を試みる。

2. 機械学習による自動 DPC コーディング

DPC コードは主傷病・副傷病・手術・処置・重症度等によって定められる全 14 桁のコードであり、図 1 に示すような構造になっている。

但し、上位の桁と下位の桁には依存関係がある。例えば、1, 2 桁目は主要診断群 (MDC: Major Diagnostic Category) を表しており、神経系疾患や眼科系疾患など 18 分類が用いられており、主要診断群を表す上位 2 桁と 3-6 桁目の上位 6 桁で傷病名を表している。なお、x は区別がない場合に用いられる。

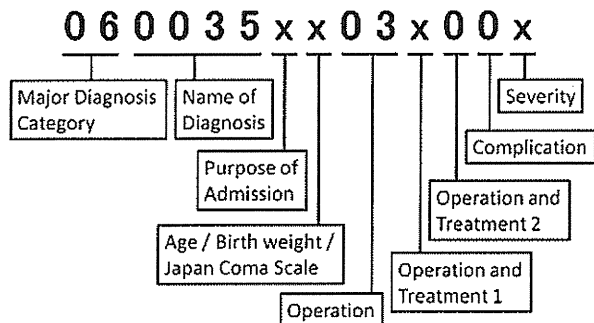


図 1 DPC コードの構造

Fig. 1 The structure of DPC codes.

そして、DPC を用いた入院医療費の定額支払い制度において、対象病院では入院患者に対し DPC コードを 1 つ付与し、その DPC コードを医療費請求に利用している。但し、DPC コードの種類は 1 万種類を超える種類が存在するのに対し、DPC を用いた入院医療費の定額支払い制度において使用されている DPC コードは限られる。(本研究で用いた解析データが作成された 2008 年 4 月から 2009 年 3 月時点では 1572 種類であった。) そのため、使用されている DPC コードに当てはまらない症例は出来高で支払われている。この DPC コードの付与は入院患者の症状、あるいは、入院患者に対して行った手術・処置といった各種医療情報をもとに行われる。

そして、自動 DPC コーディングとは、各症例に関連する医療情報を解析データとし、各 DPC コードをクラスとして自動クラス分類を行うことに等しい。

また、DPC コードは上位 6 桁で傷病名を表しているため、先行研究においては、上位 6 桁の傷病名をクラスとして自動クラス分類することにも重要性があると考えられ、DPC コード上位 6 桁の一致率が調べられている [2, 3]。(本研究で用いた解析データが作成された 2008 年 4 月から 2009 年 3 月時点では上位 6 桁に対応する傷病名の種類は 452 種類であった。)

以下では、自動 DPC コーディングの解析データとして用いる退院時サマリについて述べ、続けて、各種機械学習手法について述べる。

2.1 退院時サマリ

退院時サマリは、患者の退院後に患者の入院時の情報を要約したもので、主訴、傷病名、既往歴、経過記録、退院時所見、備考などが記載されている。京都大学医学部附属病院では退院時サマリの項目は定められているが、内容に関しては自由記載方式を取っており、各項目に含まれるデータは自然言語形式のデータである。退院時サマリは患者の退院後に患者の症例を要約して記載されるため、解析データとして有用であると考えられており、これまでも様々な研究において利用されてきた [4, 5]。

2.2 機械学習手法

機械学習手法を用いたクラス分類では、あらかじめ定められたクラスに自動的にデータの振り分けを行う。各機械学習手法は、正しいクラスの情報が付与された教師データを機械的に学習し、学習結果に基づいて未分類のテストセットをクラス分類する。

現在、機械学習手法として有用であると考えられている手法として、ベクトル空間法 [2, 6-8]、ナイーブ・ベイズ法 [9]、SVM [10]、概念ベース法 [11-13] が挙げられる。

また、複数の機械学習手法を統合することで、精度を高める試みがなされている [14-20]。中でも内山らはロジスティック回帰分析による確信度を用いて複数の異なるモデルの機械学習手法を統合する手法を提案し、提案手法の有

効性を示した[19, 20].

以下では本研究で用いる各種機械学習手法について、また、確信度を用いた複数の機械学習手法の統合手法について概説する。

2・2・1 ベクトル空間法 ベクトル空間法は各教師データと各テストデータをベクトル空間に配置し、各テストデータと教師データのベクトル空間における相対的な位置からクラス分類を行う手法である[6-8].

ベクトル空間の各軸にはデータの要素が対応する。本研究で解析データとして用いる退院時サマリは自然言語に基づくデータであり、各単語が各要素に対応すると考える。

さらに、データをベクトル空間に配置するために、各要素の値を定める必要がある。値を定める際に、頻度と重要度の積を利用することが一般的であり[21-23]、要素 e の重要度 $idf(e)$ として以下の式で表されたものを使用することが多い[8, 24].

$$idf(e) = \log \frac{N}{df(e)}, \quad (1)$$

ここで、 N は全データ数、 $df(e)$ は要素 e が含まれるデータ数である。本研究でも各要素の値として頻度と重要度の積を利用する。

ベクトル空間法では、まず、ベクトル空間の軸と要素の単位に従い教師データをベクトル化し、ベクトルを $L2$ ノルムが 1 になるように正規化した上でベクトル空間に配置する。次に、各テストデータも教師データと同様にベクトル空間に配置する。そして、ベクトル空間に配置したテストデータのクラス分類を行う。このクラス分類を行う際に、各教師データのベクトルと比較してクラス分類を行う手法（逐一ベクトル空間法）と平均ベクトルと比較してクラス分類を行う手法（平均ベクトル空間法）がある。

逐一ベクトル空間法では、ベクトル空間に配置したテストデータと各教師データの 2 つのベクトルの内積を求め、最も内積が大きくなる教師データのクラスにテストデータを分類する。なお、Suzuki らはこの各教師データのベクトルと比較してクラス分類を行うベクトル空間法を利用した[2].

平均ベクトル空間法では、同じクラスの教師データを集めて平均のベクトルを求めて正規化し、各クラスの平均ベクトルとする。そして、ベクトル空間に配置したテストデータのベクトルと各平均ベクトルの内積を求め、最も内積が大きくなる平均ベクトルのクラスにテストデータを分類する。

2・2・2 ナイブ・ベイズ法 ナイブ・ベイズ法は要素の頻度を単純に計測することで、クラス分類モデルを得る方法である[9]。本研究においては、多項モデル (multinomial model) のナイブ・ベイズ法に重要度を反映できるように変形する。

ナイブ・ベイズ法ではデータ D がクラス C_k に属する

事後確率 $P(C_k|D)$ を考え、事後確率が最も高くなるクラスにデータを分類する。ここで、事後確率 $P(C_k|D)$ はベイズ則により、

$$P(C_k|D) = \frac{P(C_k)P(D|C_k)}{P(D)}, \quad (2)$$

と書ける。 D はクラス分類を行う際に与えられるので、 $P(D)$ は一定である。なお、本研究では事前確率 $P(C_k)$ を全てのクラスに対して同じと設定する。そのため、尤度 $P(D|C_k)$ が最も大きくなるクラスを選択すればよいということになる。

そして、データ D の構成要素を e_1, \dots, e_n 、各構成要素に対応する頻度を f_1, \dots, f_n とすると、 $D = \{(e_i, f_i) | i = 1, \dots, n\}$ と表すことができる。さらに、データに含まれる要素に関して条件付き独立を仮定すると、対数尤度 $\log P(D|C_k)$ は

$$\log P(D|C_k) = \sum_{i=1}^n f_i \log P(e_i|C_k) + c, \quad (3)$$

となる[9]。ここで、 $P(e_i|C_k)$ はクラス C_k のデータにおいて要素 e_i が出現する確率を指し、 c はクラス C_k によらない値を持つ変数である。そのため、各クラスの尤度を比較するためには右辺の第 1 項を用いればよく、 $P(e_i|C_k)$ は実際にクラス C_k に属するデータ中の要素数をカウントすることで求めることができる。但し、要素数のカウントにおいては、ゼロ頻度問題[25]を回避するために、Good-Turing 推定法[26-28]を用いて平滑化を行う。

そして、本研究では式(3)の右辺第 1 項の f_i の代わりに頻度と重要度 $idf(e_i)$ をかけた t_i を用いる。但し、 $\sum_{i=1}^n t_i = 1$ となるように正規化を行う。

上記のように $P(e_i|C_k)$ を求めることでクラスごとの尤度 $P(D|C_k)$ の大小関係を求め、最も尤度 $P(D|C_k)$ が大きいクラスへ分類を行う。

2・2・3 サポートベクターマシン サポートベクターマシン (SVM) はカーネルという関数を用いて写像されたベクトル空間に配置された教師データを「うまく」分離する超平面を見つけ、自動分類させるテストセットをベクトル空間に配置して見つけた超平面に従い分類を行う 2 値分類器である[10].

本研究ではカーネルに与えるベクトルとして、各単語をベクトルの軸とし、頻度を値とするベクトルを用いる。そして、カーネルとして線形カーネルを用いる。また、SVM を用いて多クラスの分類を行う場合には、複数の SVM を組み合わせ、適切な指標により 1 つの分類を与える必要がある。本研究では、1 つのクラスを正のクラス、残りの全てのクラスを負のクラスとした SVM をクラスの数だけ生成し、各 SVM で各テストデータを分類し、超平面からの距離が最も正の方向に離れているクラスに分類を行う方法[29]を用いる。

2・2・4 概念ベース法 ベクトル空間法では各軸に単

語が対応するのに対し、概念ベース法では、各軸に意味属性が対応する。そして、各単語は意味属性を軸としたベクトル空間上のベクトルとして捉えられる[11-13]。そのため、概念ベース法では単語の意味的類似性をベクトル間の距離として定量的に評価できるため、単語の一致・不一致に関わらず、概念的な内容に基づいて分類できる長所（汎化能力の高さ）を持つ。

単語に対応するベクトルを作成する方法はいくつか知られているが、本研究では単語・意味属性間共起に基づく方法を用いる[13]。また、必要となる意味属性として、先行研究[13]においても使用されている日本語語彙大系[30]における一般名詞意味体系の意味属性を用いる。

さらに、医療に特化した概念ベースを生成するため、医療文書を用いて概念ベースの生成を行う。まず、医療文書の中に出現する単語と意味属性とが、一文内において共起した回数を要素とする共起行列 X を生成する。医療文書の中に出現する単語の種類が n 個の場合、日本語語彙大系[30]における一般名詞意味体系の意味属性の数が 2715 個であるため、 X は $n \times 2715$ の行列となる。さらに、行列 X を特異値分解により 3 つの行列 $X = U\Sigma V$ に分解する。ここで、 U 、 V はユニタリ行列で、 Σ は対角要素が特異値（固有値）の対角行列である。この左特異値行列 U の各行が単語の特徴（概念）を表していると考えられる。そして、 Σ の特異値の大きい方から順に d 番目までの特異値に対応する列ベクトルを U から取り出し、 $n \times d$ 列の部分行列にする。（本研究では $d = 2000$ とする。）最後に、この部分行列の各行ベクトルの L2 ノルムを 1 に正規化し、概念ベース B とする。

ここで、概念ベースを用いて、データのベクトル化を行う。まず、概念ベースを用いて教師データの各単語をベクトル化する。この際に、単語 w の重要度 $idf(w)$ を掛け合わせて利用する。そして、全ベクトルを足し合わせて教師データのベクトルとし、L2 ノルムが 1 になるように正規化した上でベクトル空間に配置する。

次に、各テストデータも教師データと同様にベクトル空間に配置する。そして、ベクトル空間に配置したテストデータのクラス分類を行う。平均ベクトル空間法と同様に、クラスごとに正規化された平均ベクトル作成し、内積を求めて最も大きい内積となる平均ベクトルのクラスに分類することで行う。

2.2.5 確信度を用いた組み合わせクラス分類器 クラス分類器を統合することで、新たなクラス分類器とする試みがなされており[14-20]、内山らはロジスティック回帰分析による確信度を用い、複数の異なるモデルの機械学習手法の統合に適用し、有効性を示した[19, 20]。そこで、本研究では内山らと同様に、確信度を用いた異なるモデルの組み合わせクラス分類器を用いる。

まず、ロジスティック回帰分析を利用した確信度の推定

に関して述べる。確信度は各機械学習手法の結果の正しさを示す指標である。確信度の推定には、各機械学習手法が最も正しいらしいと推定したクラスの分類スコア、2 番目に正しいらしいと推定したクラスの分類スコアが用いられる。ここで、分類スコアとはベクトル空間法や概念ベース法における内積、ナイーブ・ベイズ法における尤度、SVM における超平面からの距離を表している。具体的に、確信度 P は以下の 2 式から求められる[18]。

$$P = \frac{\exp Z}{1 + \exp Z}, \quad (4)$$

$$Z = \beta_0 + \beta_1 L_1 + \beta_2 L_2, \quad (5)$$

ここで、 β_0 、 β_1 、 β_2 はモデルパラメータ、 L_1 、 L_2 はそれぞれ最も正しいらしいと推定したクラスの分類スコア、2 番目に正しいらしいと推定したクラスの分類スコアである。そして、ロジスティック回帰分析により、モデルパラメータ β_0 、 β_1 、 β_2 を導出する。このモデルパラメータ β_0 、 β_1 、 β_2 の導出は機械学習手法ごとに、また、各機械学習手法が最も正しいらしいと推定したクラスごとに行われ、最も正しいらしいと推定したクラスの正誤情報、最も正しいらしいと推定したクラスの分類スコア、2 番目に正しいらしいと推定したクラスの分類スコアからなる教師データが用いられる。

次に、確信度を用いた組み合わせクラス分類器[19]について述べる。i) 複数の異なるモデルの機械学習手法でテストデータのクラス分類を行う。ii) 各機械学習手法が最も正しいらしいと推定したクラス、各機械学習手法が最も正しいらしいと推定したクラスの分類スコア、および、2 番目に正しいらしいと推定したクラスの分類スコアから各機械学習手法の確信度を求める。iii) 最も確信度が大きい機械学習手法が最も正しいらしいと推定したクラスにテストデータを分類する。

3. 実験 1

本節では、退院時サマリを解析データとして利用し、各機械学習手法を用いた自動 DPC コーディングを行い、評価する。

3.1 解析データと方法

本研究では、2006 年 4 月から 2009 年 3 月に京都大学医学部附属病院において作成された退院時サマリを用いて実験を行う。そこで、実際に解析の対象となる退院時サマリを以下のように選択した。まず、2008 年 4 月から 2009 年 3 月に作成された退院時サマリを DPC コードに紐付けを行い、13189 件の退院時サマリを選択した。（この際に紐付けられた DPC コードは診療情報管理士によって修正されたものである。なお、出現する DPC コードの種類は 1117 種類であった。）さらに、各 DPC コードがそれぞれ 20 件以上の退院時サマリに対応するように 8869 件の退院時サマリを選択した。このとき、DPC コードの種類は

156 種類であった。また、本研究では各退院時サマリから主訴、傷病名、既往歴、経過記録、退院時所見、備考を抜き出し解析データとして利用する。

これらの解析データを 2 節で示したベクトル空間法、ナイーブ・ベイズ法、線形カーネルを利用した SVM、概念ベース法を用いて、自動 DPC コーディングを行う。Suzuki らの結果[2]と比較するため、ベクトル空間法は逐一ベクトル空間法と平均ベクトル空間法の両方を用いる。そして、各単語 w の重要度 $idf(w)$ を算出するための医療文書として、また、概念ベースを作成するための医療文書として、2006 年 4 月から 2009 年 3 月に京都大学医学部附属病院において作成された退院時サマリ 60402 件をコーパスとして用いる。

なお、評価は 10 分割交差により、人手で DPC コードを割り当てた結果と一致する割合を調べる。さらに、DPC コーディング機能を実応用する際には、必ずしも 1 つの DPC コードのみを出力する必要はなく、医師に DPC コードの候補を提示するという選択肢もあると考えられる。そこで、各機械学習手法の分類スコアが高い上位 5 つの DPC コードと実際の DPC コードとの比較も行う。

3.2 結果

各機械学習手法を利用した DPC コーディングの結果を表 1 において示す。ここで、Precision の列では最も分類スコアが高かった DPC コード (クラス) が人手で与えた DPC コードと一致した割合を表わしており、Prec. of top 5 の列では分類スコアが高かった上位 5 つの DPC コードに人手で与えた DPC コードが含まれる割合を表している。そして、14 digits の列では各機械学習手法が分類した DPC コードが実際の DPC コードと全 14 桁一致している割合を表わしており、6 digits の列では各機械学習手法が分類した DPC コードと実際の DPC コードの上位 6 桁が一致している割合を表している。また、各精度の右に示している数値は 10 分割交差ごとの標準偏差である。なお、ppVS, aveVS はそれぞれ逐一ベクトル空間法、平均ベクトル空間法を表し、NB はナイーブ・ベイズ法、CB は概念ベース法を表している。

3.3 考察

3.2 の結果から、逐一ベクトル空間法に対する他の手法の有効性が認められる。特に、DPC コード全 14 桁を評価対象とした場合、SVM の 1 位精度が高く、DPC コード上位 6 桁を評価対象とした場合、概念ベース法の 1 位精度が高い。なお、概念ベース法を用いた場合、人手で与えた DPC コード全 14 桁が 5 位以内に含まれる割合 (5 位以内精度) が 95% となり、非常に高い精度となることがわかった。

ここで、本実験で用いたデータ量が各機械学習手法間の比較を行うに十分であったかどうか判断するために、符号検定を用いて評価を行った。具体的には逐一ベクトル空間法と他の各機械学習手法を対象とし、1 位精度の結果を用いて符号検定を行った。表 2 に符号検定の結果を示す。表 2 において、1 列目の Comparative method の列では逐一ベクトル空間法と比較する機械学習手法を表しており、ppVS の列では逐一ベクトル空間法が正解しており比較手法が誤っているテストデータの数を表わしており、逆に、3 列目の Comparative method の列では逐一ベクトル空間法が誤っており比較手法が正解しているテストデータの数を表わしている。また、p-value は両手法に差がないと仮説を立てた場合の確率を表している。いずれも、p-value が 0.01 より小さく、1% の有意水準で見た時、仮説は棄却され、統計的に有意であるといえる。そのため、本実験で用いたデータ量は十分であったと考えられる。

実験 1 において、精度の高かった機械学習手法に関する考察であるが、SVM は一般的にも精度の高い機械分類手法であり、その有用性が本実験においても示されたものと考えられる。また、概念ベース法に関しては各単語を概念により抽象化することによる単語の一致・不一致に関わらない汎化能力の高さが DPC コード上位 6 桁、あるいは、5 位以内精度といった解空間を拡張させた場合において有効に働いたものと考えられる。

さらに、各機械学習手法の DPC コードごとの 1 位精度を見ていくと、機械学習手法ごとに特徴があることがわかる。例として、表 3 に本研究で扱った 156 種類の DPC

表 1 各機械学習手法を利用した DPC コーディングの結果
Table 1 The results of DPC code selection using several machine learning methods.

Method	Precision (%)		Prec. of top 5 (%)	
	14 digits	6 digits	14 digits	6 digits
ppVS	64.8 ± 1.2	79.2 ± 1.2	91.6 ± 1.0	95.8 ± 0.7
aveVS	68.0 ± 2.1	82.5 ± 2.3	91.7 ± 2.0	95.8 ± 1.5
NB	70.4 ± 2.3	85.0 ± 1.5	93.1 ± 1.5	96.5 ± 1.4
SVM	72.2 ± 2.6	83.3 ± 2.0	88.3 ± 2.7	94.6 ± 1.8
CB	69.0 ± 2.0	85.2 ± 1.9	95.0 ± 1.2	97.5 ± 0.7

表 2 逐一ベクトル空間法と他の機械学習手法を対象とした符号検定結果

Table 2 The results of sign tests between the point-by-point vector space model and the other machine learning methods.

Comparative method	ppVS	Comparative method	p-value
aveVS	1475	1759	3.0×10^{-7}
NB	1273	1765	2.2×10^{-19}
SVM	1183	1839	4.0×10^{-33}
CB	1412	1777	5.1×10^{-11}

表3 DPCコードごとの1位精度(%)

Table 3 The precision of code selection for some DPC codes (%).

DPC code	ppVS	aveVS	NB	SVM	CB
010060x099x00x	30.8	33.3	55.1	34.6	30.8
050070xx97x0xx	37.9	41.4	51.7	31.0	62.1
070570xx010xxx	72.0	88.0	76.0	84.0	76.0
100260xx99x00x	66.7	55.6	74.1	74.1	81.5
140010x199x0xx	70.1	69.0	80.5	93.1	67.8

コードのうち一部のDPCコードごとの1位精度を示す。

表3の結果から、機械学習手法ごとに適したDPCコードがあることが見て取れる。この手法ごとの差異は、例えば、ベクトル空間法やナイーブ・ベイズ法が各単語を独立したものとして網羅的に扱うことができること、SVMがある分類を特徴付ける単語を発見することを得意としていること、概念ベース法が単語の意味的類似性を定量的に評価することができる汎化能力の高さを持つことなどに起因していると考えられる。

機械学習手法ごとに適したDPCコードがあることから、複数の機械学習手法を適切に組み合わせることで精度を向上させることができる可能性がある。但し、逐一ベクトル空間法と平均ベクトル空間法は似た特徴を持っているため、精度の良い平均ベクトル法を用いて他の機械学習手法と組み合わせるのが適当であると考えられる。

4. 実験 2

本節では3・3項の考察に従い、確信度を用いて平均ベクトル空間法、ナイーブ・ベイズ法、SVM、概念ベース法を組み合わせる新たなクラス分類器を構築し、退院時サマリの自動DPCコーディングを行い、実際のDPCコードとの比較を行う。

4.1 解析データと方法

解析データは3節と同じ2008年4月から2009年3月に作成された8869件の退院時サマリを用いる。そして、クラス分類には2・2・5の確信度を用いた組み合わせクラス分類器を用いる。組み合わせる機械学習手法は平均ベクトル空間法、ナイーブ・ベイズ法、SVM、概念ベース法である。

なお、確信度を用いた組み合わせクラス分類器のモデルパラメータ β_0 、 β_1 、 β_2 を算出し、機械学習手法を選択するには10分割交差を用いる。

4.2 結果

複数の機械学習手法を組み合わせるクラス分類器の結果を表4に示す。

但し、全ての組み合わせについて掲載することは困難であるため、3種類、及び、4種類の機械学習手法についての組み合わせに関する結果のみを掲載した。

表4 複数の機械学習手法を組み合わせたクラス分類器のDPCコーディングの結果

Table 4 The results of DPC code selection by utilizing the method which combines several machine learning methods.

Combined method	Precision (%)	
	14 digits	6 digits
aveVS+NB+SVM	75.8±2.5	87.6±2.1
aveVS+NB+CB	74.0±2.3	87.4±2.1
aveVS+SVM+CB	75.9±2.5	87.4±1.7
NB+SVM+CB	76.0±2.5	87.9±1.8
aveVS+NB+SVM+CB	76.1±2.6	87.9±1.9

なお、確信度を用いた組み合わせクラス分類器は最も確からしい1位のDPCコードのみを算出する分類器であるため、5位以内精度は考慮していない。

DPCコード全14桁を評価対象とした場合、精度が最も高くなるのは本実験で用いた4つの機械学習手法全てを組み合わせるクラス分類器であり、上位6桁を評価対象とした場合は精度が最も高くなるのはナイーブ・ベイズ法、SVM、概念ベース法を組み合わせる場合であった。

4.3 考察

逐一ベクトル空間法と比較し、本実験で用いた4つの機械学習手法を組み合わせるクラス分類器はDPCコード全14桁、上位6桁を評価対象とした場合に、それぞれ11.3%、8.6%もの大幅な精度の向上が見られた。

また、各機械学習手法を単一で用いた時、DPCコード全14桁、あるいは、上位6桁を評価対象とした場合に、最も精度が高かった機械学習手法はそれぞれSVMと概念ベース法であった。そして、これらの結果と比較すると、組み合わせるクラス分類器の精度がそれぞれ3.9%、2.6%向上しており、確信度を用いた組み合わせるクラス分類器の有効性が示された。

5. 全体の考察

実験1では、既存手法として用いられた逐一ベクトル空間法に比べ、他の一般的な機械学習手法の精度が優れていることが示された。また、実験2では複数の機械学習手法を確信度を用いて組み合わせることで精度の高いクラス分類器を構築できることが示された。

DPC包括評価は米国において用いられているDiagnosis-Related Group (DRG) systemをもとにして導入されたものであり、DRG systemもDPC包括評価と同様に各症例を1つの症例群に割り当て、症例群ごとにあらかじめ定められた診療報酬を決定するものである。そのため、DRG systemにおける診療文書からの症例群決定においても本研究と同様の結果が得られる可能性があると考えられる。

また、今後の課題としていくつかの解決すべき問題点がある。1つ目は、そもそも教師データとして用いた DPC コードが正しい DPC コードであるという保証がないという点である。実際、京都大学医学部附属病院では 2008 年度に医師が付与した DPC コードを診療情報管理士が確認することが 10.9% の DPC コードの変更に繋がっている [1]。このことから、DPC コーディングは決して簡単な作業ではなく、多くの付与間違いが発生していると考えることが自然である。そのため、今後、本来どの DPC コードを付与すべきだったのか解析を行うシステムの構築を行う必要があると考える。そして、DPC コードの付与間違いによる本実験への影響であるが、本実験では診療情報管理士により修正が行われた DPC コードを用いているため、DPC コードの付与間違いは非常に少数となっていると考えられ、本実験の各精度が逆転する程の影響はないものと推測される。また、より正確な DPC コードが付与されれば、誤った教師データを学習することが減ることで、より精度の高い自動 DPC コーディングが実現できると考えられる。

問題点の 2 つ目は、本研究で扱った DPC コードの種類が全 DPC コードの 1 割程であるということである。これは機械学習手法を利用する際に、ある程度の教師データが必要となることが原因である。但し、本研究で扱った DPC コードに対応する症例は全症例の 7 割を占めており、今後症例数を増やすことで扱える DPC コードの種類が増え、症例の割合も上昇すると考えている。しかし、一部の DPC コードに関しては、将来的にも十分な数が揃わないことも考えられる。その際には一部の DPC コードに関してはルールを作り、機械学習手法と併用して利用して行く必要があると考えられる。

6. ま と め

本研究では退院時サマリーを解析データとして複数の機械学習手法を用いて DPC コーディングを行った。結果として、先行研究において利用された逐一ベクトル空間法では DPC コード全 14 桁の 1 位精度が 64.8% であったのに対し、複数の機械学習手法を確信度によって組み合わせクラス分類を行うことで 76.1% と 11.3% の精度の向上を達成した。また、診療文書をもとに概念ベースを作成した概念ベース法を用いると 5 位精度が 95% と非常に高い精度となることがわかった。

謝辞 本研究の成果の一部は、NTT からの受託研究によるもの、および、科研費 22700257 の助成を受けたものである。

文 献

1. 大井美香子, 田村寛, 福嶋実, 金子悦代: 1 年間の DPC コー

- ディング適正化作業の検討. 第 35 回日本診療情報管理学会学術大会予稿集, 静岡, 2009, p. 82.
2. Suzuki T, Yokoi H, Fujita S, Takabayashi K: Automatic DPC code selection from electronic medical records: text mining trial of discharge summary. *Methods Inf Med.* **47** (6): 541-548, 2008.
 3. 西貴士, 魚澤正克: データマイニングの手法を用いた診療行為からの DPC 病名決定モデル, 第 29 回医療情報学連合大会予稿集 (CD-ROM), 広島, 2009, 2-C-4-6.
 4. 小野大樹, 高林克己, 鈴木隆弘, 横井英人, 井宮淳, 里村洋一: テキストマイニングによる退院サマリー自動分類の試み. *医療情報学.* **24**(1): 35-44, 2004.
 5. 岡本和也, 竹村匡正, 黒田知宏, 長瀬啓介, 吉原博幸: 文脈に基づく類似診療文書検索システム. *生体医工学.* **44** (1): 199-206, 2006.
 6. Salton G, Buckley C: Term-weighting approaches in automatic text retrieval. *Inf Process Manage.* **24** (5): 513-523, 1988.
 7. 石井健一郎, 上田修功, 前田英作, 村瀬洋: わかりやすいパターン認識. オーム社, 東京, 1998.
 8. 徳永健伸, 辻井潤一: 情報検索と言語処理. 東京大学出版会, 東京, 1999.
 9. McCallum A, Nigam K: Comparison of Event Models for Naive Bayes Text Classification. *Proc of AAAI-98 Workshop on Learning for Text Categorization*, 1998, pp. 41-48.
 10. Boser BE, Guyon IM, Vapnik VN: A Training Algorithm for Optimal Margin Classifiers. *Proc of the 5th Annual ACM Workshop on Computational Learning Theory*, 1992, pp. 144-152.
 11. Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R: Indexing by latent semantic analysis. *J Am Soc Inf Sci.* **41**: 391-407, 1990.
 12. Schutze H: Automatic Word Sense Discrimination. *Comput Ling.* **24**(1): 97-123, 1998.
 13. 別所克人, 内山俊郎, 内山匡, 片岡良治, 奥雅博: 単語・意味属性間共起に基づくコーパス概念ベースの生成方式. *情報処理学会論文誌.* **49**(12): 3997-4006, 2008.
 14. 藤野昭典, 上田修功, 斉藤和巳: 最大エントロピー原理に基づく付加情報の効果的な利用によるテキスト分類. *情報処理学会論文誌.* **47**(10): 2929-2937, 2006.
 15. 工藤拓, 松本裕治: 半構造化テキストの分類のためのブースティングアルゴリズム. *情報処理学会論文誌.* **45** (9): 2146-2156, 2004.
 16. Schapire RE, Singer Y: BoosTexter: A Boosting-based System for Text Categorization. *Machine Learning.* **39** (2-3): 135-168, 2000.
 17. Schapire RE, Singer Y: Improved boosting algorithm using confidence-rated predictions. *Mach Learn.* **37** (3): 297-336, 1999.
 18. 高橋和子: 素性選択によるアンサンブル学習に関する一考察. *情報処理学会自然言語処理研究会 SIGNL-188*, 2008, pp. 23-28.
 19. 内山俊郎, 別所克人, 内山匡, 奥雅博: 確信度推定を用いた複数分類器の結合. *人工知能学会第 84 回知能ベースシステム研究会予稿集*, 2009, pp. 33-38.
 20. 内山俊郎, 岡本和也, 竹村匡正, 足立貴行, 桑直人, 黒田知宏, 内山匡, 吉原博幸: 医療分野に特化した概念ベースを用いた DPC コーディング. *信学技報 PRMU2010*, 2010, pp. 157-162.
 21. Salton G, Yang CS: On the specification of term values in automatic indexing. *J Doc.* **29**(4): 351-372, 1973.

22. Salton G: A Theory of Indexing. Society for Industrial and Applied Mathematics, Philadelphia, 1975.
23. Salton G, Yang CS, Yu CT: A theory of term importance in automatic text analysis. J ASIS. 26(1): 33-44, 1975.
24. 北研二, 津田和彦, 獅々堀正幹: 情報検索アルゴリズム. 共立出版, 東京, 2002.
25. 北研二: 確率的言語モデル. 東京大学出版, 東京, 1999.
26. Good IJ: The population frequencies of species and the estimation of population parameters. Biometrika. 40(3-4): 237-264, 1953.
27. Church KW, Gale WA: A comparison of the enhanced good-turing and deleted estimation methods for estimating probabilities of english bigrams. Comput Speech Lang. 5:19-54, 1991.
28. Chen SF, Goodman J: An Empirical Study of Smoothing Techniques for Language Modeling. Technical Report at Harvard University, TR-10-98, 1998.
29. Hamel LH: Knowledge Discovery with Support Vector Machines. Wiley-InterScience, Massachusetts, 2009.
30. 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系. 岩波書店, 東京, 1997.

・ 岡本 和也 (オカモト カズヤ)

平 16 京大・工・情報・卒. 平 21 同大・情報・了. 平 19 日本学術振興会特別研究員 (DC2). 平 21 京大・情報・特定研究員. 同年同大・病院・特定助教. 現在に至る. 医療情報学, アルゴリズムの研究に従事. 電子情報通信学会論文賞・猪瀬賞等受賞. 博士 (情報学).



日本医療情報学会, 日本医療・病院管理学会, 電子情報通信学会の会員.

内山 俊郎 (ウチヤマ トシオ)

昭 62 東工大・工・電気電子・卒. 平 1 同大・工・修士了. 同年(株)NTT データ入社. 平 3-5 南カリフォルニア大客員研究員. 平 11-17 通信・放送機構研究員, 特別研究員. 平 18 より日本電信電話(株)サイバースリソリューション研究所所属. Web データマイニング, 分光色再現の研究に従事. 博士 (工学). 電子情報通信学会, 情報処理学会, IEEE の会員.



竹村 匡正 (タケムラ タダマサ)

平 10 阪大・医・保健・卒. 平 15 同大・医・了. 同年京大・医・助手. 平成 17 同大・病院・運営企画室助手 (兼任). 平 19 同大・医・講師. 現在に至る. 病院経営, 医療情報学, 特に医療における自然言語処理の研究等に従事. 博士 (保健学).



日本医療情報学会, 日本医療・病院管理学会, 日本医療マネジメント学会等の会員.

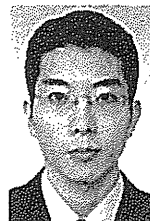
足立 貴行 (アダチ タカユキ)

平 7 九工大・情報・卒. 平 9 同大・情報・修士了. 同年 NTT に入社. 平 17-20 NTT レゾナント, 平 21 より NTT サイバースリソリューション研究所所属. テキストマイニングに関する研究開発に従事. 情報処理学会の会員.



桑 直人 (クメ ナオト)

平 18 京大・情報・了. 同年日本学術振興会特別研究員 (PD). 平 19 京大・医・助教. 平 22 ABB Corporate Research 客員研究員 (兼任). 現在に至る. 医用 VR, 医療情報学等の研究に従事. 博士 (情報学). 日本 VR 学会, 日本医療情報学会, 日本 VR 医学会, 情報処理学会の会員.



黒田 知宏 (クロダ トモヒロ)

平 6 京大・工・情報・卒. 平 10 奈院大・情報・了. 同助手. 京大・医・講師. オウル大・理・情報・客員教授. 阪大・基礎工・准教授を経て, H21 京大・医・准教授, 病院・医療情報部副部長. 現在に至る. 仮想・強調現実感, 福祉情報学, 医療情報学等の研究に従事. 日本 ME 学会阿部賞等受賞. 博士 (工学). 日本生体医工学会, IEEE, 日本 VR 学会, 日本医療情報学会等の会員.



内山 匡 (ウチヤマ タダス)

昭 60 名大・理・物理・卒. 昭 62 同大・理・修士了. 同年 NTT に入社. 平 10-13 NTT コミュニケーションズ, 平 16-18 NTT レゾナントにてポータルサービスの開発等に従事. 平 19 より NTT サイバースリソリューション研究所主幹研究員. ポータルサービスシステムの研究開発に従事. 情報処理学会, 電子情報通信学会, 日本応用数理学会の会員.



吉原 博幸 (ヨシハラ ヒロユキ)

昭 55 宮崎医大・医・卒. 昭 59 同大・医・了. 平 5 同講師. 平 7 同教授. 平 10 ハーバード大学, MIT・客員准教授. 平 12 熊大・医・教授. 平 14 大連医科大学・客員教授. 平 15 京大・医・教授, 病院・医療情報部部長. 現在に至る. 医療情報学, 外科学の研究等に従事. 医学博士.



日本生体医工学会, 日本医療情報学会, 日本外科学会, 日本医療・病院管理学会等の会員.

研究資料

インシデントレポート・医事管理データによる転倒・転落に
起因する追加的医療費算出の試み

江上 廣 一¹⁾ 廣瀬 昌 博^{1,2)} 竹村 匡 正³⁾ 岡本 和 也³⁾
 津田 佳 彦¹⁾ 大濱 京 子¹⁾ 本田 順 一¹⁾ 島 弘 志¹⁾
 今中 雄 一⁴⁾ 吉原 博 幸³⁾

入院患者の転倒・転落に起因する検査や治療は、原疾患の診療とは異なり、本来不要で、その追加的医療費は保険者である行政や病院管理者等にとっては看過できず、医療経済学上、正確な把握が必要である。本院における2007年～2009年度の3年間のインシデントレポート7,717件について、転倒・転落件数は影響度レベル別にレベル2が824件、レベル3aが298件およびレベル3bが46件であった。そのうち、あらたな医療費（追加的医療費）が確認できた件数（費用判明率）および1件あたりの平均追加的医療費は、レベル2が205件（24.9%）および10,070±7,934円、レベル3aが186件（62.4%）および12,859±15,772円、レベル3bが34件（73.9%）および226,723±281,065円で、レベル2・3a間（ $p=0.027$ ）、3a・3b間（ $p<0.001$ ）でも有意差を認めた。この結果から、年間あたりの追加的医療費を推計すると、各レベル別にレベル2が2,744千円、レベル3aが1,281千円およびレベル3bが3,476千円で、総額約7,000千円と推計された。したがって、医療安全管理の立場から、効率的な転倒・転落防止策として、3b事例を防止することが重要であることが示唆された。

キーワード：追加的医療費、転倒・転落、影響度レベル、インシデントレポートシステム、医事データ、臨床研修病院

I. 目 的

わが国における国レベルでの医療安全対策の一つとして、2002年10月厚生労働省がすべての病院および有床診療所の管理者に対して、医療の安全を確保する体制の整備を義務付けた¹⁾。そのうち、医療機関内における事故報告等の医療に係る安全の確保を目的とした改善のための方策を講ずることとは、

院内におけるインシデント・アクシデントの報告制度を指すもので、一般にはインシデントレポートシステムと呼ばれ、収集されるレポートはインシデント・アクシデントの区別なくインシデントレポートと総称されている。この制度は、類似のインシデント・アクシデント発生防止に有効であることが知られているが、その一方で医療従事者の負担の割にフィードバックが少ないなどの指摘もされている²⁾。

一方、2003年4月特定機能病院に包括的医療費支払制度であるDPC（Diagnosis, Procedure, and Combination）が特定機能病院に導入されるなど、

¹⁾ 聖マリア病院 医療の質管理本部

²⁾ 島根大学医学部附属病院 病院医学教育センター

³⁾ 京都大学医学部附属病院 医療情報部

⁴⁾ 京都大学大学院医学研究科 医療経済分野

医療費適正化政策のもと、医療施設には医療の安全確保と医療の質の向上とともに効率的で効果的な病院運営が求められている³⁾。

さて、収集されるインシデントレポートのうち、影響度（傷害）レベル⁴⁾によればレベル2以上の事例は、発生したインシデント・アクシデントによって検査や治療を実施することになり、その検査や治療は、その当該患者には本来不要で、あらたな医療費は病院が一部負担する場合も少なくない。また、わが国の医療費適正化政策の観点からも原疾患とは異なる医療費の発生は望ましくない。一方、医療制度が異なり、安易な比較は禁物であるが、米国では65歳以上を対象とした医療扶助制度 Medicare は、合理的に防止できる（reasonably prevented）事例については、それが院内で発生した場合、Medicare は支払いに応じないという。そのなかには、部位間違い手術、遺物遺残や重症褥瘡、転倒による傷害が含まれている⁵⁾。

とくに転倒・転落による事例はインシデントレポートの2割前後を占め⁶⁾、件数的にも医療経営上、看過できない状況である。しかも、わが国では超高齢者社会にあつて、入院患者の高齢化も著しい状況では、病院管理者ばかりでなく、行政にとってもその追加的医療費を把握しておく必要がある。しかしながら、海外における研究は有害事象に費やす医療費の推計に関するものが多く⁷⁻⁹⁾、また、国内の研究では対象とする転倒・転落事例の件数が少なく¹⁰⁻¹²⁾、相当数の転倒・転落事例に関する追加的医療費を検討したものはなく、しかも、影響度レベルで検討した研究はない。

そこで、本研究は、インシデントレポートシステムで報告される事例のうち、転倒・転落に起因する、あらたな追加的医療費について医事データを用いて影響度レベル別に算出し、適切で的確な医療安全活動に資することを目的とした。

II. 方 法

本研究は、聖マリア病院医の倫理委員会の承認を得て実施されている。また、患者やインシデント・アクシデントに関与した職員らの個人情報については、匿名化するなど、細心の注意を払って研究を遂行した。

＜病院概要＞

聖マリア病院（以下、本院）は、福岡県久留米市にあり、一般病床1,188床、療養型病床100床、精神病床60床を含めた1,354床を擁する、救命救急センターを兼備した急性期病院である。年間の入院患者数は約320,000人、外来患者は約370,000人であり、救急車搬入台数は年間約10,000台、救急搬入患者は約60,000名に達する。

＜医療の質管理本部およびインシデントレポートシステム＞

1999年の医療事故を契機に医療安全管理委員会が設置され、医療安全管理本部は2003年に設置された。同本部は2005年に発展的に拡充し、医師を本部長として、専任医師、専任看護師、専任薬剤師各1名ずつを含め、計7名で構成され、種々の医療安全活動を展開しており、2011年4月にはその名称が医療の質管理本部に変更された。院内で発生したインシデント・アクシデントを収集するインシデントレポートシステムは、1999年看護部から始まり、2006年1月には、オンラインシステムを導入し、現在に至っている。

インシデントレポートの入力には、1. 事象レベル、2. 特に報告を求める事例、3. 報告者と発生日時、4. 発生場所、5. 患者情報、6. 発見者と関連職種、7. 当事者基本、および8. 当事者勤務状況の順に入力し、最終的には1枚のシートで閲覧できる（図1）。そのうち、事象レベルは、国立大学附属病院の影響度レベルや小林らの報告と同様、レベル1からレベル5まで傷害の内容はまったく同じである^{4,10)}。また、特に報告を求める事象は3b以上の事象を指し重要事例としており、これがアクシデント事例に、レベル2および3a事例はインシデント事例に相当する。

＜対象および方法＞

2007年4月1日～2010年3月31日までの3年間にインシデントレポートシステム Safe Master[®]（アステム社、福岡）により報告された転倒・転落事例を対象とし、以下の手順で調査した。

- 1) 報告された事例のうち、医療費の発生する、影響度レベル2以上の事例を収集する。
- 2) そのうち、インシデントレポートシステムによるインシデント報告（図1）に記載された【発生時の対応】をもとに検査オーダー等を確認し、原疾患の医療費とは異なるものを抽

インシデント報告書		出力日 2011/03/05 10:27:40
インシデント定義名	転倒・転落用	
インシデントNo	0010095	
報告日	2009年10月07日	
1. 事象レベル	3a	
2. 発生日時	2009年10月06日 23時30分	
3. 発生場所	病室	
4. 報告者	*****	
5. 発見者	同職業者	
6. 当事者	【敬語】 当事者無し (患者本人による) *****	
7. 患者基本情報	*****	
8. 発生場面	【机上の場面】 排せ行动中 【きっかけ】 その他：尿器を取ろうとして	
9. 発生内容	【状況】 自立歩行中 【区分】 転倒	
10. 発生要因	【患者側】 歩行器 降階剤 利尿剤 尿意・便意が頻回 トイレ介助が必要 歩行できないという過信 (歩行訓練開始) 1週間以上 (直前のスコア) 11 (一つ前の評価日) 1週間以内 (一つ前のスコア) 11 【環境】 他衣服の不具合： 【設備】 (ベッド側) 3本指 (ベッドの種類) ハイ・ローベッド (起床センサー) 対象外の患者だった 【ケア提供者側】 急変・重症による多忙 ナースコール頻回 薬剤服用後の観察が不十分 リスクは少しあった その他医療側の要因：ope後入院の患者が病室直後にて対応中	
11. 事前に取りられていた防止策/受傷軽減策	【防止策】 3本指 床ずれ誘導 他配慮：衣服の考慮	
12. 障害内容	【外傷等】 軽微受傷あり	
13. 検査・処置対応	【診察】 11 外科 【診察日時】 2009年10月07日 00時10分 【処置】 湿布、消毒、軟膏処置等	
14. 患者・家族への説明・反応	【説明】 説明する予定	
15. 事前考察	【生命の危険度評価】 生命の危険度は低い 1/万 【発生頻度】 月1回以上 【考察】 注意すれば防げた事例 【医師への報告】 2009年10月06日 23時40分 (医師) 当直医 【管理者への報告】 2009年10月06日 23時43分 (管理者) 管理部長	
16. 発生状況要旨 (記述)	【発生状況要旨】 23:30前後入院の患者さんをope室より入院させない(尿測定・嘔気への対応)への説明を行っていた時、退出勤務者が物音がした為訪室、ベッド下に座りこんでいる所発見、本人どこを打ったか不明、右前腕に沁みあり、呼吸変化無し、打撲痕、頭部に問題無し。	
17. 発生要因 (記述)	【発生要因】 机上昇しリハビリ中、日中歩行器で1人で歩行出来ているが就寝前にベッドリハ内服しておりふらつき有り。	
18. 発生時の対応 (記述)	【発生時の対応】 Dr診察しびらん部の消毒後、一せ保護、おしり・意識レベルの観察	
19. 今後の対策 (記述)	【今後の対策】 ナースを呼ぶように指導、排せ誘導を行う	
医療安全管理者	2009/10/08 安全管理本部長：本田 洋 2009/10/07 安全管理：江上 洋 2009/10/07 安全管理：江上 洋	

図1 インシデントレポートシステムにおけるインシデント報告書(例)

出する。

- 3) 医事システム会計カード(レセプトデータ: NEC社製PC-IBARS)と検査オーダーを照合することで、診療の有無を確認する。
- 4) 照合された検査や治療などの診療項目について医事システム会計カードから、医療費を収集する。

調査方法を具体的に示すと、次のようである。例えば、図2は、本研究のうち、某診療科で発生した転倒・転落事例で、当該月の11日に転倒が発生し、その後、整形外科医師がオーダーした診療行為等の一覧表である。本院では、入院患者に転倒・転落が生じた場合、整形外科医師がこれに関する診療行為についてオーダーすることとなっている。

図2に示される左端の欄には各診療行為の項目が並び、当該月の1日から31日までの診療行為の有無が示されている。すなわち、「0:算定なし」は当該の診療行為が実施されていないことを示し、「1:

算定あり」は、その診療行為が実施されたことを示している。したがって、この一覧表によって、転倒・転落により、原疾患による医療費とあらたに生じた追加的医療費を区別し、算出することが可能である。

以上のように本研究は、インシデントレポートシステムで報告される転倒・転落事例をもとに、電子カルテにより転倒・転落の発生を確認し、医事データを用いて転倒・転落が発生したことに起因する追加的医療行為を同定し、当該行為の医療費を推計する手法を用いたretrospective case studyである。

<本研究における影響度レベル>

院内で発生する事象については、一般には、インシデントおよびアクシデントとして把握されている施設が多い。それらについて、小林らが明確に定義している¹⁰⁾。国立大学附属病院安全管理協議会では、平成14年10月開催の第1回国立大学医学部附属病院医療安全管理協議会において、報告のあったすべての事象を「インシデント」とし、インシデントの

転倒・転落発生日 原疾患による診療行為

あらたな診療行為

科	項目名称	使用量	点数	10日	11日	12日	13日	14日	15日	16日	17日	18日	19日	20日	21日	22日	23日
16	骨盤単純	0	210	0	1	0	0	0	0	0	0	0	0	0	0	0	0
16	大腿骨単純	0	223	0	1	0	0	0	0	0	0	0	0	0	0	0	0
16	股関節単純	0	210	0	1	0	0	0	0	0	0	0	0	0	0	0	0
16	単純CT撮影(マルチスライスCT)	0	770	0	1	0	0	0	0	0	0	0	0	0	0	0	0
14	関節内骨折観血的手術(股)	0	17800	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	閉鎖循環式全身麻酔5	0	6100	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	局所麻酔	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	酸素	150	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	セボフレン	20	617	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	スタドール注1mg O. 1%1・	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	キシロカイン注射液「0. 5%」エビレナミン	20	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	アナペイン注7. 5mg/・ 0. 75%20・	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	サリンヘス 6%500・	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	フレゾフォル1%注 200mg20・	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	ネオンネジンコーワ注1mg O. 1%1・	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	ディオバン錠80mg	0. 5	21	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	アーナスト錠2. 5mg	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	アムロジンOD錠5mg	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	朝食後30分	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	コニール錠4	2	33	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	リスモダンR錠150mg	2	33	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	朝・夕食後30分	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	ウルソ錠100mg	3	11	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	重カマ「ヨシダ」	1. 5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
14	ガスモチン散 1%	1. 5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0

0:算定なし 1:算定あり

図2 診療行為別算定表

うち、①「医療側に過失があり」、②「患者に一定程度の障害があり」、③「①と②間に因果関係がある」ものを「医療事故」とし、アクシデントや医療過誤も同義とすると定めた経緯がある。また、その後、国立大学附属病院における医療上の事故等の公表に関する指針の別表に示されているように、「傷害の継続性」と「傷害の程度」を加味した影響度分類が示されており、本院においても同様に「影

響度レベル」として取り扱っている⁴⁾(表1)。

したがって、表1から影響度レベル2以上の事象については、検査、処置や治療など何らかの診療行為が実施されており、原疾患とは異なるあらたなコストが発生する。

表1 影響度レベル

レベル	傷害の継続性	傷害の程度	傷害の内容
レベル0	未実施(未然に発見)		エラーや医薬品・医療用具の不具合が見られたが、患者様には実施されなかった
レベル1	なし		患者への実害はなかった(何らかの影響を与えた可能性は否定できない)
レベル2	一過性	軽度	処置や治療は行わなかった(患者観察の強化、バイタルサインの軽度変化、安全確認のための検査などの必要性は生じた)
レベル3a		中等度	簡単な処置や治療を要した(消毒、湿布、皮膚の縫合、鎮痛剤の投与など)
レベル3b		高度	濃厚な処置や治療を要した(バイタルサインの高度変化、人口呼吸器の装着、手術、入院日数の延長、外来患者の入院、骨折など)
レベル4a	永続的	軽度-中等度	永続的な障害や後遺症が残ったが、有意な機能障害や美容上の問題は伴わない
レベル4b		中等度-高度	永続的な障害や後遺症が残り、有意な機能障害や美容上の問題を伴う
レベル5	死亡		死亡(原疾患の自然経過によるものを除く)

4) より引用

III. 結 果

A. 病院基本データ (表2)

表2は、2007年度から2009年度までのインシデント・アクシデントレポート件数および平均在院日数と入院患者数から入院患者延べ数が求められ、転倒・転落の報告数を転倒・転落数と同じと仮定すると、

転倒・転落の発生率

$$= \text{転倒・転落数} \div (\text{入院患者数} \times \text{平均在院日数}) \times 1,000$$

$$= \text{転倒・転落数} \div \text{入院患者延べ数} \times 1,000$$

ここで、単位は1,000患者・日 (patient-dayもしくはoccupied bed)である。これは、在院日数が長くなると転倒・転落の危険性が高くなることを考慮に入れたclinical indicatorであり、本院では1.81~1.90で3年間を通じて1.87で、年度による統計学的有意差は認めなかった。

B. 影響度レベル別報告件数と追加的医療費 (表3)

表3は、影響度レベル別追加的医療費を示してい

表2 病院概要

	2007年度	2008年度	2009年度	全体
平均在院日数(日)	20.6	20.8	21.4	21.0
入院患者数(人)	15,549	15,076	15,112	45,737
入院患者延べ数(人・日)	321,286	312,717	324,372	958,375
インシデント・アクシデント全数	2,760	2,559	2,398	7,717
転倒・転落報告件数	610	565	615	1,790
転倒・転落発生率(1,000患者・日)	1.90*	1.81*	1.90*	1.87

$$\begin{aligned} \text{転倒・転落の発生率} &= \text{転倒・転落数} \div (\text{入院患者数} \times \text{平均在院日数}) \times 1,000 \\ &= \text{転倒・転落数} \div \text{入院患者延べ数} \times 1,000 \end{aligned}$$

*: 年度による統計学的有意差はない

表3 影響度レベル別追加的医療費 (医療費の単位は千円)

年度	影響度レベル	2	3a	3b	合計
2007	Ne	275	113	19	610
	iNe	77	89	11	177
	iCe	736	1,076	2,755	4,568
	iCe/iNe	9.6±7.6	13.1±19.5	250.1±353.8	26.3±103.3**
	Ne×iCe/iNe	2,640	1,480	4,752	16,043 (8,872)
2008	Ne	264	89	8	565
	iNe	46	41	7	94
	iCe	449	500	1,002	1,951
	iCe/iNe	10.7±9.7	12.2±13.1	143.1±184.5	21.2±59.4**
	Ne×iCe/iNe	2,825	1,089	1,145	11,978 (5,059)
2009	Ne	285	96	19	615
	iNe	82	56	16	154
	iC	841	728	3,952	5,521
	iCe/iNe	10.3±7.2	13.0±10.1	247.0±269.5	35.9±111.3**
	Ne×iCe/iNe	2,936	1,248	4,693	22,079 (8,877)
全年度	ΣNe	824	298	46	1,790
	ΣiNe	205	186	34	425
	ΣiCe	2,027	2,304	7,709	12,040
	ΣiCe/ΣiNe	10.1±7.9*	12.9±15.8*	226.7±281.1*	28.6±98.5
	ΣNe×ΣiCe/ΣiNe	8,322	3,844	10,428	51,194 (22,594)

Ne: 追加的医療費の発生したと考えられる件数, iNe: 追加的医療費が確認できた件数, iCe: 確認できた追加的医療費。

()内はレベル別の総和を示す。* レベル2・3aおよび3a・3b間で有意差を認めた。** 年度による有意差は認めない。