Development of a Pioneering Clinical Support System Utilizing Information Technology

Clinical Informatics and Genome Analysis

Doubun HAYASHI,¹ MD, Yasushi IMAI,² MD, Hiroyuki MORITA,² MD,
Hideo FUJITA,² MD, Koshiro MONZEN,¹ MD, Tomohiro HARADA,² MD,
Takefumi NOJIRI,² MD, Tadashi YAMAZAKI,³ MD, Tsutomu YAMAZAKI,³ MD,
and Ryozo NAGAI,² MD

SUMMARY

Nowadays, evidence-based medicine has entered the mainstream of clinical judgement and the human genome has been completely decoded. Even the concept of individually designed medicine, that is, tailor-made medicine, is now being discussed. Due to their complexity, however, management methods for clinical information have yet to be established. We have conducted a study on a universal technique which enables one to select or produce by employing information processing technology clinical findings from various clinical information generated in vast quantity in day-to-day clinical practice, and to share such information and/or the results of analysis between two or more institutions. In this study, clinically useful findings have been successfully obtained by systematizing actual clinical information and genomic information obtained by an appropriate collecting and management method of information with due consideration to ethical issues. We report here these medical achievements as well as technological ones which will play a role in propagating such medical achievements. (Jpn Heart J 2004; 45: 315-324)

Key words: RCN System, Database system, Clinical informatics, Data mining, Information technology, Genome analysis, Evidence-based medicine, Tailor-made medicine

AS its population ages, major diseases in Japan have shifted from acute types of disease such as infections to chronic types such as life-style related diseases. Meanwhile, the government has recently implemented in some institutions a Prospective Payment System in which the government reimburses remuneration at a predetermined amount and health care reform with the major purpose being reorganization of the health care system from an economical point of view is already in place. As a result, medical institutions are now required to safely and effi-

From the ¹ Department of Pharmacoepidemiology, ² Department of Cardiovascular Medicine, ³ Department of Clinical Bioinformatics, Graduate School of Medicine and Faculty of Medicine, University of Tokyo, Tokyo, Japan.

Address for correspondence: Doubun Hayashi, MD, Department of Pharmacoepidemiology, Graduate School of Medicine and Faculty of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.

Received for publication September 17, 2003.

Revised and accepted September 25, 2003.

ciently offer patients with chronic diseases more effective health care services than those currently available in order to cope with this reform.

However, in responding to social demands such as this, there are various issues which are difficult to solve completely only by precautions against human error or allocation of appropriate human resources. Therefore, the first step to solve such issues seems to be the preparation of fundamental information that is readily available prior to actual clinical treatment and the formation of a foundation upon which evidence-based medicine (EBM) is practiced. Since clinical research activity in Japan to evaluate clinical efficacy and safety is not up to the levels in Europe and the US, health care is mainly offered according to the clinical guidelines depending on the clinical data compiled in these countries. However, it must be noted that there are many differences arising from race, incidence of certain diseases, living environment, enzymatic activities for metabolizing drugs, and so on, and there is more and more need to collect fundamental clinical information and analyze clinical data obtained in Japanese subjects.

With the view of resolving these issues, much is expected recently in the realization of effective as well as efficient and safe health care brought about by systematizing clinical information. But in a majority of the cases, the attempt to select useful medical findings from clinical information is being made by individual clinical studies. It is considered that developing a technique to share such individual clinical information between health care institutions all over Japan widely accepted and practiced will in the end greatly contribute to the realization of a safe and at the same time effective health care system.

The objective of this study was to establish a universal technique with which to extract medical findings and to systematize diagnostic procedures by applying pioneering information technology (IT), with the aim of resolving the above-mentioned problems. In other words, we have constructed a clinical information management system with the following functions, in order to offer safe health-care services, to realize efficient and effective healthcare, to improve the level of healthcare in general and further to overcome economical problems, all by implementing IT in the management of fundamental clinical information:

- Having clinical information in an electronic format (database construction)
- Comprehensive data analysis with data mining as the main function
- Real-time network linking of clinical information

Further, since this system will be utilized to assist in making a diagnosis in a clinical setting based on the vast amount of fundamental data accumulated in real-time, we have named this system the "Real-time Clinical Navigator System" (hereinafter referred to as "RCN System").

METHODS

Clinical information database: In this study, we have converted to electronic data a vast amount of information obtainable in day-to-day clinical activity, that is, various types of clinical information such as events (“death”, “acute myocardial infarction”, “cardiac failure”, “stroke”, etc.), laboratory findings (“Tchol”, “HbA1C”, “heart rate”, etc.) or prescriptions (“drug for treatment of angina pectoris”, “antithrombotic drug”, “anticoagulant”, etc.) and constructed its database. In collecting such data, we proceeded through the informed consent procedure with each individual patient after we had obtained approval from the ethics committee of the University of Tokyo Faculty of Medicine. We only inputted clinical data and carried out genomic analysis for those patients who provided informed consent.

One of the major features of the RCN System interface is to store data in the form of a so-called chronological table, such as that used in historical science, with the passage of time in mind. There are the following two advantages in data storage methods according to the passage of time:

- It is possible to refer to the clinical information of a patient in time sequence.
- It is easy to establish the starting point in a prospective or retrospective investigation.

To be more specific, presenting a patient's clinical information in a time sequence such as a chronological table enables one to look at the patient's clinical history visually. Figure 1 shows the screen displaying a time sequence. Also, it is

Figure 1. Presentation in time sequence of a patient's clinical history.

designed so that one can refer to detailed data for each item in Figure 2. The presentation style which enables making a diagnosis on the spot such as this may save time in grasping the clinical history of a patient in an actual clinical situation and is highly useful. Further, in the case of an emergency, this could be an effective supportive tool for treating patients without delay.

In addition, Figure 3 shows an example of investigation of variation in total cholesterol. Here, we show an image of the investigation of the results of the first

Figure 2. Details of each clinical item.

Figure 3. Investigation of variation in total cholesterol.

laboratory blood test of total cholesterol (Tchol) in three patients and the variation in Tchol six months later in the same three patients. By arranging side by side the starting points of the three patients which are actually different, one can conduct epidemiological surveys such as prospective and transversal studies.

Also, it is possible to conduct an investigation following up a certain specific lesion in a coronary artery for progression, regression, restenosis, and so on as time passes. Further, we have developed a function to show the degree of stenosis in each lesion in time sequence, which enables one to readily investigate the progression, regression, and restenosis of coronary atherosclerosis in each lesion.

In this way, it is possible to investigate and analyze the data accumulated based on day-to-day clinical information in time sequence. Utilizing the RCN System, one can always peruse in real-time the results, similar to those obtained in a clinical research study, which would require enrollment of a great number of patients, a large amount of money, and follow-up over a long period of time.

Technique to extract clinical findings by data analysis: Based on the RCN System mentioned in the previous section, we are going to describe a method for extracting medical findings. In this study, we have developed and describe a comprehensive data analysis technique called data mining and an analytical system to extract, process, and present the data in accordance to the purpose.

Data mining is a technology which has become generally accepted and used in recent years and by setting up a simple criterion this technique can efficiently search for a rule satisfying the criterion in a vast amount of data^{1,2)} Table I shows an example of the results of analysis by data mining. Table I deals with “a patient with myocardial infarction (MI)” and shows premises when an evaluation standard with high degrees of certainty and support and a high odds ratio is selected. By utilizing such comprehensive data analysis, one can not only confirm clinical findings obtained based on a rule of thumb, but also extract findings completely unknown so far.

Figure 4 shows an example of a screen indicating the results in analytical function and automatic graph making function of the RCN System. One of the characteristics of the RCN System lies in its function enabling one to grasp trends in the data visually, including medical findings obtained by data mining. This is useful as a function to extract/process the data by the simple operation of selecting conditions, and present the results of the search graphically. Graphic presentation helps one to display trends in the data and the differences between groups and could greatly assist one in making clinical decisions.

Utilizing the technique to extract clinical findings by data mining and an analytical system contribute greatly to the establishment of a totally new diagnostic method or clinical guidelines.

Table I. Example of Data Analysis by Data Mining

Premise	Conclusion	Certainty	Support	Odds ratio
Gene X1 = A/C Gene Y1 = A/A Male	MI (+)	0.8163	0.1646	4.7281
Gene X2 = A/A Gene Y2 = A/A Familial IHD (-) Male	MI (+)	0.8438	0.1588	4.6703
Gene X3 = C/C Gene Y3 = defect Male	MI (+)	0.875	0.1657	6.3194

Certainty means a ratio of data securing the conclusion to those satisfying the premise. Support means a ratio of data satisfying both premises and conclusion to all of the data. Odds ratio means how many times more readily the data satisfying the premises can secure the conclusion when compared with the data which does not satisfy the premises. For example, in the rule in the first line of Table I a male patient with gene type X1 being A/C, hetero type, and gene type Y1 being A/A, homo type develops MI with a probability of 81.63 % (certainty: 0.8163) and such a rule occurs in 16.46 % of the total number of patients (support: 0.1646). Also, this shows that a male patient with gene X1 being A/C, hetero type and gene Y1 being A/A, homo type tends to develop MI 4.7281 times as frequently as those who have different gene type (odds ratio: 4.7281).

Figure 4. Example of graphic image of the analytical system of case data. Here, the patients are divided into two groups, one developing diabetes and the other no diabetes. The ratio of the patients who develop restenosis after recanalization of coronary artery is shown.

RESULTS

Biochemistry markers useful in diagnosis of coronary artery diseases without left ventricular dysfunction: In cases other than cardiac failure, we have studied atrial natriuretic peptide (ANP) and brain natriuretic peptide (BNP) which are known as markers indicating the severity of cardiac failure or prognosis of the patients. When limiting the cases to those with normal contraction ability of the left ventricle, we have obtained the finding showing that ANP and BNP are correlated with the severity of coronary artery diseases. As shown in Figures 5A and 5B, the higher the values of ANP and BNP, the greater the number of lesions in the coronary artery tends to be. We will continue to collect and analyze more data.

Factors regulating restenosis in coronary angioplasty: The ratio of patients with inadequately controlled diabetes who develop restenosis after coronary arterial intervention tends to be high compared with other patients.³⁾ Table II shows the relation between the existence of diabetes, HbA1C, and restenosis. In the case of patients with diabetes whose HbA1C values are higher than 7.0, we have obtained the result that, compared with other cases, they have a higher ratio of restenosis after recanalization of a coronary artery.

Further, in our data, when pioglitazone, a drug for the treatment of diabetes, is used, there was a tendency toward a lower rate of restenosis after coronary intervention treatment, compared with other drugs. Table III shows the relation between drugs for the treatment of diabetes and restenosis rate.

Figure 5. A: Correlation between plasma ANP level and the severity of coronary artery diseases. B: Correlation between plasma BNP level and the severity of coronary artery diseases.

Table II. Relation Between Existence of Diabetes, HbA1C, and Restenosis Rate

	With diabetes	Without diabetes
HbA1C 7.0 or more	39.2 (%)	27.5 (%)
HbA1C less than 7.0	29.6 (%)	-

Table III. Relation Between Drugs for the Treatment of Diabetes and Restenosis Rate

Drugs	Rate of restenosis
Sulfonylurea	33.6 (%)
Glucosidase inhibitor	29.7 (%)
Pioglitazone	25.0 (%)
Insulin	36.5 (%)

For three drugs except pioglitazone, restenosis after recanalization occurs at a rate of approximately 30% or higher, while in the cases where pioglitazone is prescribed, the rate is as low as 25%.

Table IV. Relation Between Ile823Met Polymorphism and Various Factors

	Ile/Ile	Ile/Met+Met/Met	Statistics
HDL-C	44.9 ± 11.5	49.0 ± 15.1	<i>P</i> = 0.04
BMI	23.5 ± 3.7	23.1 ± 3.3	NS
Tchol	180.9 ± 31.8	183.5 ± 36.4	NS
TG	125.7 ± 65.1	129.4 ± 87.9	NS
LDL-C	115.1 ± 32.7	114.9 ± 37.9	NS
Smoking	59 (61.5)	115 (68.0)	NS
Hypertension	65 (67.7)	120 (71.0)	NS
Diabetes	20 (20.8)	46.0 (27.2)	NS

It becomes clear that patients with Ile823Met including Met have an HDL concentration significantly higher than those without Met⁹⁾. No significant difference was observed in other factors such as obesity (BMI), total plasma cholesterol (Tchol), tryglycerides (TG), bad cholesterol (LDL-C), smoking, hypertension, and diabetes.

Ile823met and lipid concentration: It is reported that Ile823Met polymorphism of the ABCA1 gene has an effect on the concentration of good cholesterol (HDL-C).⁴⁾ Taking into consideration differences in race and eating habits, we have investigated what results would be obtained in Japanese people. Table IV shows the relation between Ile823Met polymorphism and various factors in patients receiving no administration of the drug for the treatment of hyperlipidemia.

Genetic markers for coronary artery diseases: We have selected approximately 50 genes which have something to do with arterial sclerosis and investigated the possible influence of polymorphism of such genes on the occurrence of coronary artery diseases, especially myocardial infarction. Analysis of the Gensini Score, known as an index for arterial sclerosis in coronary arteries, and background factors such as the presence of hypertension, obesity, and diabetes, and the existence of coronary arterial sclerosis lesions is ongoing and we are planning to report the results shortly.

DISCUSSION

The aims of this study were to contribute to clinical practice from the approach of having clinical information in electronic form by IT, describe a universal technique to extract clinical findings by constructing an RCN System, and demonstrate the effectiveness of such a system by describing the clinical findings obtained so far.

The clinical findings obtained so far include:

- Biochemical markers (ANP, BNP) having a correlation with the severity of coronary diseases in patients with normal cardiac function
- That the existence of diabetes and drugs for the treatment of diabetes influence the restenosis rate after recanalization of a coronary artery
- That the ABCA1 gene has an effect on lipid concentration, which is a risk factor for heart diseases
- Analysis of genetic polymorphism to find out specific sequence(s) which may have an influence on the onset of coronary diseases, myocardial infarction, and severity of coronary arterial sclerosis.

Based on these results, one can see that preparation of establishment of prompt extraction of useful clinical findings and useful diagnosing method is steadily carried out.

At this point in time, we are at the stage where the construction of a clinical database and extraction of clinical findings by data analysis by local health care institutions have been completed. From the view of spreading social technology, any development to be made from now on is extremely important. By cooperating with health care institutions all over Japan, the collection and analysis of more extensive clinical data or medical check up data will become possible. As a result, the sharing of clinical data among all health care institutions in Japan will become possible. By this, physicians can promptly grasp their patients' clinical history and background and treat them immediately and effectively. Patients can visit any health care institution in Japan and receive appropriate health care services without being subjected to duplication of testing.

Nationwide linking of databases accompanying the exchange of data does not present significant difficulty thanks to widespread internet and broadband circuits. While the challenge of implementing an effective security system required for the handling of clinical data should be met by all means, the sharing of clinical information data between health care institutions all over Japan should be feasible. One can expect that clinical findings and completely new diagnostic methods obtained by accumulation and analysis of vast amounts of clinical data by network linking with health care institutions throughout Japan will be shared by all of the health care institutions in Japan.

Also, up until now we have constructed the system concentrating on cardiovascular disease related items. Our next challenge is to widely promote the RCN System without emphasis on any particular disease or clinical department. Further, it is not necessary to limit the application of this system to health care institutions and it is possible to establish links to other industries, such as the pharmaceutical and insurance industries. For the pharmaceutical industry, information as to what type of patient receives what drug and whether there are any side effects would be valuable in the development of pharmaceutical products. For the insurance industry, information concerning the risk of onset of diseases subject to insurance payment as well as the risk of death could be used in setting insurance rates, etc.

In this study, we have sought to contribute to the establishment of clinical guidelines and realize the sharing of clinical information by "one case history per one patient" system. We hope that in the end our attempt will greatly contribute to the realization of a safe and at the same time precise and efficient health care system in Japan.

ACKNOWLEDGEMENT

A part of this study was performed under contract to the New Energy and Industrial Technology Development Organization (NEDO). The authors are grateful for the assistance received from many people during this study. The authors particularly thank people at Hitachi Corporation for providing us with valuable advice and instruction on system design, data analysis, data mining, etc.

REFERENCES

1. Fukuda T, Morimoto Y, Tokuyama G. "Data Mining Data Science Series 3" published by Kyoritsu Publishing, 1999. (in Japanese).
2. Usama M, F Gregory, P Shappiro, *et al.* Advances in Knowledge Discovery and Data Mining. The AAAI Press, 1996.
3. Eric Van Bel, Christophe Bateurs, Jean-Marc Lablanchele, *et al.* Restenosis rates in diabetic patients: a comparison of coronary stenting and balloon angioplasty in native coronary vessels. *Circulation* 1997; 96: 1454-60.
4. Bodzioch M, Orso E, Lackner KJ, *et al.* The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. *Nat Genet* 1999; 22: 347-51.
5. Harada T, Imai Y, Nojiri T, *et al.* A common Ile823 Met variant of ATP-binding cassette transporter A1 gene (ABCA1) alters high density lipoprotein cholesterol level in Japanese population. *Atherosclerosis* 2003; 169: 105-12.