J Med Syst (2012) 36:723–736
DOI 10.1007/s10916-010-9540-x

ORIGINAL PAPER

Building Clinical Data Groups for Electronic Medical Record in China

Haibo Tu · Yingtao Yu · Peng Yang · Xuejun Tang · Jianping Hu · Keqin Rao · Feng Pan · Yongyong Xu · Danhong Liu

Received: 28 March 2010 / Accepted: 6 June 2010 / Published online: 14 July 2010

Abstract This article aims at building clinical data groups for Electronic Medical Records (EMR) in China. These data groups can be reused as basic information units in building the medical sheets of Electronic Medical Record Systems (EMRS) and serve as part of its implementation guideline. The results were based on medical sheets, the forms that are used in hospitals, which were collected from hospitals. To categorize the information in these sheets into data groups, we adopted the Health Level 7 Clinical Document Architecture Release 2 Model (HL7 CDA R2 Model). The regulations and legal documents concerning health informatics and related standards in China were implemented. A set of 75 data groups with 452 data elements was created. These data elements were atomic items that comprised the data groups. Medical sheet items contained clinical records information and could be described by standard data elements that exist in current health document protocols. These data groups match different units of the CDA model. Twelve data groups with 87 standardized data elements described EMR headers, and 63 data groups with 405 standardized data elements constituted the body. The later 63 data groups in fact formed the sections of the model. The data groups had two levels. Those at the first level contained both the second level data groups and the standardized data elements. The data groups were basically reusable information units that served as guidelines for building EMRS and that were used to rebuild a medical sheet and serve as templates for the clinical records. As a pilot study of health information standards in China, the development of EMR data groups combined international standards with Chinese national regulations and standards, and this was the most critical part of the research. The original medical sheets from hospitals contain first hand medical information, and some of their items reveal the data types characteristic of the Chinese socialist national health system. It is possible and critical to localize and stabilize the adopted international health standards through abstracting and categorizing those items for future sharing and for the implementation of EMRS in China.

Keywords Electronic medical record · HL7 CDA R2 · Data group · Data element

H. Tu · P. Yang · F. Pan · Y. Xu (✉) · D. Liu (✉)
Institute for Health Informatics,
Fourth Military Medical University,
Changle West Road 17#,
Xi'an 710032, China
e-mail: xuyongy@fmmu.edu.cn
e-mail: liudanh@fmmu.edu.cn

Y. Yu
Division of Medical Service, TangDu Hospital,
Fourth Military Medical University,
XinShi Road 1#,
Xi'an 710038, China

X. Tang · J. Hu · K. Rao
Center of Health Statistics and Information, Ministry of Health,
XiZhiMenWaiNan Road 1#,
Beijing 100044, China

Introduction

Electronic health records have been rarely adopted in U.S. hospitals due to a lack of interoperability [1]. The lack of interoperability in the health domain in China has prevented health informatization [2]. Many standards have been created to engender interoperability. Health Level 7 Clinical Document Architecture Release 2 (HL7 CDA R2) [3, 4] is currently one of most widely adopted clinical document

architectures model. The Continuity of Care Document (CCD) attempted to harmonize the Continuity of Care Record (CCR) [5, 6] with the CDA. The Integrating Healthcare Enterprise (IHE) suggested combining existing standards to improve interoperability [7]. The Taiwan electronic Medical record Template (TMT) was designed to achieve semantic interoperability in EHR exchanges nationally [8].

The Chinese Ministry of Health has proposed an overall approach to healthcare informatization for the Chinese mainland to support the national strategy of providing basic medical insurance to all Chinese residents. The approach was described as laying three foundations, building three level platforms, and improving profession application systems [9]. The three foundations will be reflected in the nationally used and standardized Electronic Health Record (EHR) for all residents, the standardized Electronic Medical Record (EMR) structure for all medical pavilions, and the national health information data dictionary. This research focused on the second foundation and was supported by a grant from the Ministry of Health. Whereas the goals and purposes of the second foundation are to build the basic EMR structure and its data element standard, this research proposes a basic EMR content structure to establish a guideline for the EMRS of all healthcare providers in China.

According to a survey on urban resident health conditions conducted in 2009, only 20% of Chinese citizens had regular physical examinations [10]. This means that most people learn about their health conditions only when visiting a doctor after becoming ill. Only then do medical institutes obtain their health information. Therefore, in most cases, medical institutes are positioned to collect and report the medical information of residents, and these medical institutes become the best source of the personal and health information of residents. The statistical data of the Ministry of Health shows that in 2007, most national- and provincial-level hospitals owned a local Hospital Information System (HIS), and about 38% of county-level hospitals had various levels of HIS.

The factor that most influences the variety in HIS levels is the absence of standards that could be adopted in the HIS software. *The Basic Function Specification of HIS* issued by the Ministry of Health in 2001 [11] is the only functional guideline for hospital data collection, and *The Common Data Elements of Health Records* issued in January 2009 by the Ministry of Health is not sufficient to support the requirements of defining the basic EMR structure and content.

As a component of HIS, an EMR system (EMRS) is designed to obtain the original medical data from the HIS database, create a medical summary of patient stay in a hospital, and respond to the information requirements from EHR platforms of different levels. Also, the EMRS could report the most recently updated information to the EHR platforms. Therefore, the EMR is the information source for the EHR of regional healthcare information platforms.

Our final aim was to propose a basic EMR content structure to clarify the types and formats of data needed by the regional EHR platforms. These data are to be obtained, managed and organized by the EMRS of healthcare providers. This content structure includes data groups and is independent of any specific EMRS. This structure will become the basis for all institutional EMR systems in China.

HL7 V3 Clinical Document Architecture (CDA) has provided an abstract and holistic information model to cover all forms of health information. However, the holistic and abstract attributes of the CDA created difficulty in building a particular EMR guideline in China. Therefore, we adopted the CDA model as the fundamental theory of this research.

Methodology

This study's method comprises four steps: (1) proposing a theoretical EMR content structure; (2) collecting original medical record sheets from hospitals and integrating these sheets into a basic set of medical sheets; (3) comparing the proposed basic EMR content structure to items of the basic set of medical sheets and creating basic data groups for the EMR structure; and (4) enlisting standard data elements for data groups and building the content structure for the EMR. Figure 1 shows a flowchart outlining this research.

The basic content structure of the EMR

We regarded every medical sheet as a complete clinical medical document. An EMR is the entire collection of medical documents for a patient. A basic EMR content structure was proposed to outline information to be included in the EMR and the manner in which they should be included. We used the HL7 CDA R2 model and medical diagnostics theory [12, 13]. The HL7 CDA R2 model offered the basic document structure that categorized information into the header and the body, while medical diagnostics theory provided the logic for organizing medical information.

To build an EMR content structure based on medical knowledge and the HL7 CDA R2 model, we invited medical experts to participate in relevant discussions and revising sessions. The components of the basic structure were described using a list of data groups matching the header and body as defined in the HL7 CDA R2 model.

A data group was defined as a composite data structure that aggregates relevant information elements. Occasionally, a data group could be considered as the aggregation of smaller data groups and data elements. Generally, the main

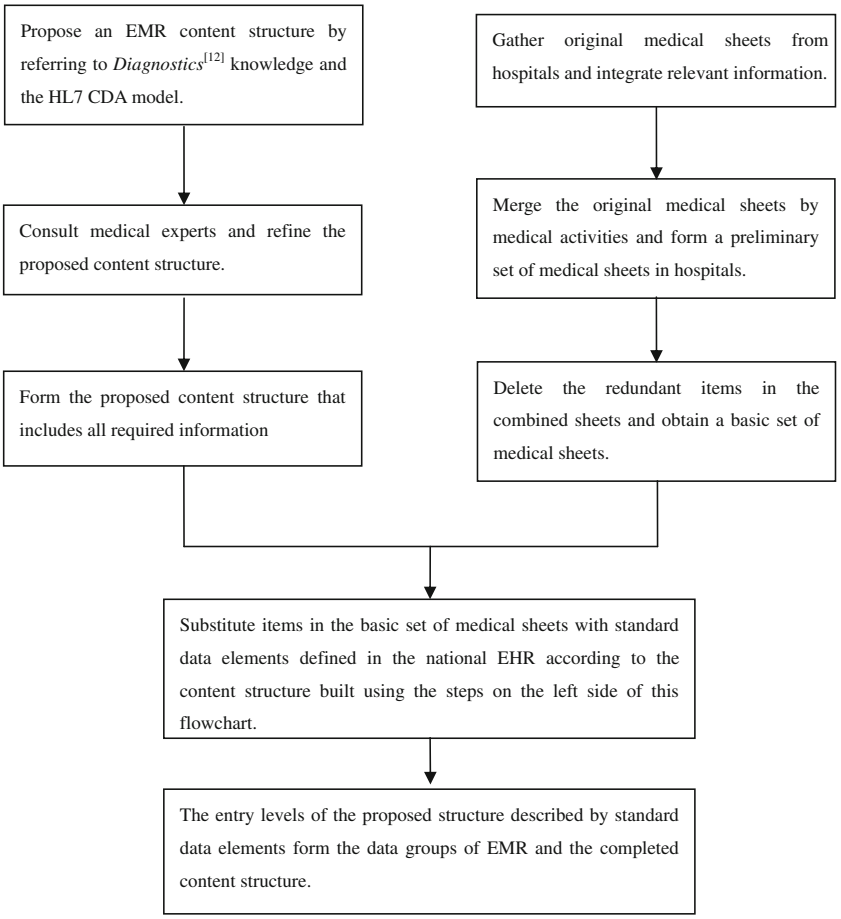

Fig. 1 Research steps

data elements composing data groups were essential for semantic interoperability in a specific context. Data groups organized the relevant data elements, and the values of these groups were assigned through the data elements. Examples of data groups include symptom, medication, surgery, and document ID.

A data element was the basic unit of data. Data elements were specified by atomic units of data that have precise meaning or semantics. The list of data groups in this step did not define the data elements included within the groups in detail. That will be done when comparing the groups and the set of medical record sheets. The listed groups were used to outline the content of the EMR.

The framework of diagnostics was helpful in creating the groups and in defining the relationships between them. The experts that participated in the discussions included senior physicians, surgeons, general practitioners, and specialists. It was particularly important to enlist the participation of the senior engineers who were familiar with HIS designing in these discussions. Two discussion sessions were held to discuss content structure. Expert opinions were used to improve the structure of the data groups and to identify their relationships. After this, the data groups were compared with the data elements included in the basic set of medical sheets.

Integrating the information collected from medical sheets used in hospitals

The collection of medical sheets from hospitals

Expert opinions and hospital medical sheets were collected simultaneously. The collection was conducted from January to March 2009. Seventeen hospitals in China were selected to conduct the collection. The hospitals selected had the following characteristics: (1) more than 50% were large-scale hospitals (large: more than 1,000 beds, medium: more than 500 beds, small: less than 500 beds); (2) all selected hospitals were 3A level; (3) at least two Chinese medicine hospitals of medium scale were included; and (4) at least one specialized hospital was included.

Merging and creating a preliminary set of medical sheets from hospitals

All medical sheets collected were to be categorized according to the medical activities targeted in the records. In categorizing, we created a list of medical activities and a list of medical profession fielding those activities. Overall consideration of the fields of profession and the proposed

content structure was helpful in creating the basic EMR content structure in step 2.4. Those original sheets that were categorized into the same activity, we merged their items on a single sheet. Then, based on the original sheets collected, we grouped the data on a single sheet. After cataloging and merging, a preliminary set of medical sheets was achieved.

Checking and integrating data elements in the medical sheets

Data elements in the preliminary set were combined from the same-purpose-sheets from different hospitals, and were not comprehensive. The sheets collected from specific hospitals did not necessarily contain information for default elements, such as the names and codes of the hospitals, hospital legal counsel, and the codes', authors, and the names of the data entry operators.

We regarded each medical sheet as a complete clinical medical document. Based on the HL7 CDA R2 model, each document was composed of a header and body. Elements that were essential to form the header of a complete EHR according to the *EHR Basic Architecture and Data Standard* issued by the Chinese Ministry of Health in May 2009 were defined as default elements. The data types of the elements defined in the *EHR Basic Architecture and Data Standard* were HL7 V3 data types.

First, each preliminary sheet was reviewed, and default data elements were added if necessary without changing the names and data types of these standard elements. Personal information items were added according to the regulations of the *Basic Dataset of Electronic Health Record* issued by the Ministry of Health in 2007 [14]. At this point, the preliminary set of medical sheets became the basic set, and default data elements were added and redundant data elements deleted.

Substitution of items in the basic set of medical sheets with standard data groups and data elements according to the content structure constructed using comparison and reconstruction

The data groups of the proposed EMR content structure formed in step 2.1 were compared with the basic medical sheets from step 2.2. The comparison had two aims: first, we intended to guarantee that the lists of data groups were exhaustive. Second, we intended to categorize the data elements of these sheets into groups to create the EMR basic data groups and EMR basic data elements.

After revising the basic medical sheets based on expert opinions, the data groups had the properties of the basic set of medical information. By assigning an ID code to each data group, the basic EMR data groups could be listed without including data elements. After careful cataloging, the integrated data elements of the basic medical sheets created in step 2.2.3 became the source of the EMR basic data elements.

The value ranges of data elements were set according to current national coding standards. For example, all national institutions in China were assigned a standard name and code in GB/T 17538-1998. The education of patients was coded in GB/T 4658-1984. After coding the data elements, the groups became real data groups that contained data elements.

Building basic EMR content templates and structures

Building basic EMR content templates

The basic medical sheets represented different medical activities, and records of similar activities were attributed to the same data group. After the elements of the data groups were revised on experts' opinion, we used the data groups to rebuild the basic medical sheets and created EMR templates. The templates were composed of the basic data groups that contained data elements. The data groups that were included in the template were decided based on the medical activities recorded. The cardinality values in the templates were defined in the process of rebuilding. That means the times that a data group could be reused in a template were defined in the cardinality value. The templates were called EMR templates and they could be used to guide what data must be included in EMRS. All templates formed the basic EMR templates.

Building basic EMR content structure

The basic templates, the basic data elements, and the basic data groups were organized together after thorough consideration of the fields of medical profession and the EMR data groups. The relationships between them were proposed.

Results

EMR content structure

The basic content structure of the EMR was based on diagnostics theory and the levels of the HL7 CDA R2 model. The model of the structure is shown in Fig. 2. The clinical document rectangle represents an EMR, which was the core of the figure. The content of the EMR was composed of data from 25 groups and some groups may include several levels of subgroups.

Most of the groups in the body came from medical knowledge, such as the main complaints of the patient, their present health problems, history of illness, family genetics, personal life and marital history, and surgical

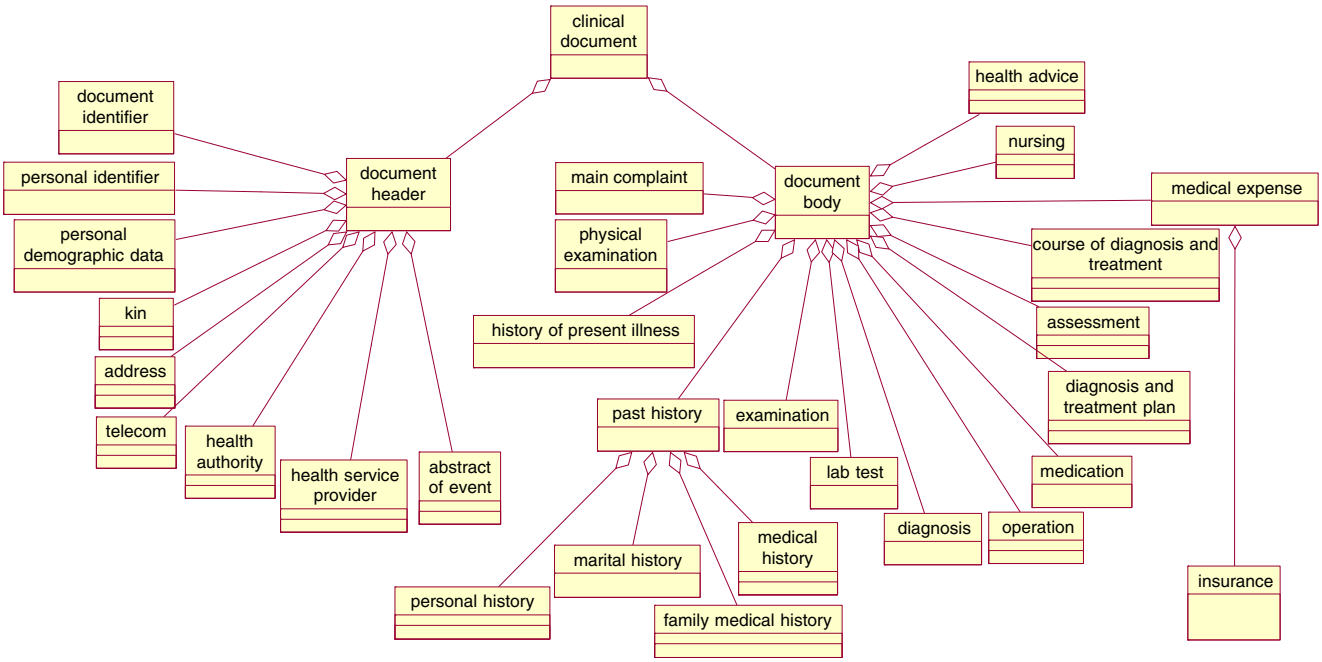

Fig. 2 Information levels of the EMR structure

history. The groups were at different levels in the EMR structure. For instance, personal history is part of 'past history', and the groups could be extended with subgroups when necessary.

According to the HL7 CDA R2 model, a clinical document consists of two main parts: a header and a body. The body of the EMR was composed of diagnostic groups, while the header of the EMR was composed of additional groups, such as demographic data and kin information.

To form the basic data groups, the depth of the levels should be sufficient to reach the leaf level of the structure, until subgroups that were only composed of data elements. The specific levels were confirmed after categorizing the medical sheets into these groups.

An EMR content structure with 25 data groups and 45 subgroups was proposed after the discussion sessions with medical experts. Parts of the groups are shown in Fig. 2.

Creating a basic set of medical sheets and basic data elements based on the collection of medical profession sheets of hospitals

All sheets used in hospitals were regarded as medical record documents that contain medical profession information. We collected and merged the sheets from selected hospitals. After merging, the list of medical fields and profession activities was generated and a basic set of hospital medical sheets was created. The basic data elements were acquired after merging different elements and eliminating redundant elements in the basic set of medical sheets.

Collecting medical record sheets

Hospitals in China were classified into three grades and four levels using the *Hospital Grade Management Script* issued by the Ministry of Health in 1989. Grade 3A is the top level of the highest grade. The evaluation criteria were based on hospital practice variety, mission, facilities, service quality, and management.

Adoption of a hospital information system was not listed in the criteria because the research was focused on the medical profession content that was independent of any specific system, while most selected hospitals have had HISs of different scales and scopes.

The collection of sheets in hospitals was conducted thoroughly, meaning that all sheets used in diagnosis and treatment were captured and collected. Sheets on patient insurance and identity information were also collected. For hospitals with computerized sheets, the sheets were printed and collected.

Seventeen 3A hospitals were selected to collect the sheets for medical profession. The selected hospitals were of different scales, and three Chinese medicine hospitals and one specialized hospital were included. Table 1 shows the characteristics of the hospitals selected.

The three Chinese medicine hospitals were also categorized as hospitals with less than 1,000 beds. In the collection of all 1,747 sheets, 260 were from the three Chinese medicine hospitals and 110 were from the specialized hospital.

Table 1 The accommodation capacity and hospital types of selected hospitals

Total selected hospital number	Total beds >2,000	Total beds (1,000–2,000)	Total beds (500–1,000)	Total beds <500	Number of general hospitals	Number of Chinese medicine hospitals	Number of specialized hospitals
17	4	6	6*	1	13	3*	1

The six hospitals (6) with beds number within 500 to 1,000 contains the three Chinese medicine hospitals (3*)

Merging and creating a preliminary set of medical record sheets in hospitals

One hundred and forty-five extractant sheets were acquired after the original 1,747 sheets were merged and redundant elements were eliminated. Taken together, all 145 sheets were taken as the basic set of medical sheets. The dramatic reduction in the number of medical sheets was due to hospitals across China providing mostly similar procedures. Hospitals in China adopted the *Basic Regulation of Patient Record* issued by the Ministry of Health as the rules for writing medical records. The regulation stipulated the scope and format of medical records. The scope was classified to many medical activities that included list of items. After merging and sorting the sheets, we categorized them to the activities, listed the medical fields and added the healthcare organization to integrate the information.

We generated 7 medical fields that contained 18 classes of first-level medical activity records and 62 second-level medical activity records that were the subgroups of the 18 first-level medical activity records. Table 2 shows the seven medical fields and the categorization of the 18 first level medical activity records.

A medical sheet represented a medical activity, and the categorization of the medical activity records depended on medical fields. Medical activity records were categorized to represent all medical sheets and were coded as EMRxxxx. By adding two digits at the end of the code, we could specify medical activity at the second level. Similar activity records that occurred on different occasions, such as outpatient or

Table 2 Medical activity records in different medical fields

Medical fields	Categorization of medical activity records
A medical summary	EMR0100 medical summary
B encounter medical records	EMR0200 outpatient/ambulant encounter medical records
	EMR0300 prescription
	EMR0601 encounter nursing records(nursing record)
	EMR0501 encounter procedure and treatment records(general)
	EMR0400 observation and laboratory test records
	EMR0700 informed consent information
C inpatient medical records	EMR0800 inpatient medical summary
	EMR0900 inpatient registry
	EMR1000 inpatient medical processing
	EMR1100 inpatient medical orders
	EMR0601 inpatient nursing records (nursing operation records)
	EMR0602 inpatient nursing records (nursing assessment and care plan)
	EMR0501 inpatient procedure and treatment records (general)
	EMR0502 inpatient procedure and treatment (labor records)
	EMR0400 observation and laboratory test records
	EMR1200 discharge records
	EMR1300 transfer records
	EMR0700 informed consent information
D health examination records	EMR1600 health examination record
E referral records	EMR1400 referral records
F legal medical evidence and report	EMR1700 legal medical evidence and report
G healthcare organization information	EMR1500 healthcare organization information

inpatient treatment, were categorized into one record. For example, nursing records for outpatient and inpatient treatment were found bearing the same record—EMR0601.

All 145 sheets were categorized into the above 62 catalogs of second level medical activity records according to the similarity of the data elements included. According to the first level category to which they belonged, their correspondent data groups would form the EMR basic template after comparison with the proposed EMR content structure.

Checking on the integrated data elements in the medical record sheets

The default elements that belonged to the header groups were obtained by referring to the *Basic Dataset of Electronic Health Record*. The *Basic Dataset of Electronic Health Record* was created to support the EHR information requirement. It was a component of the *EHR Basic Architecture and Data Standard*. Its data elements are consistent with the HL7 Data Type model. In addition, specific elements used exclusively in China were included in the basic dataset. Some of the elements had coding systems in China. For example, the coding of patent education was taken from the standard GB/T 4658-1984, the coding of patient occupation was taken from the standard GB/T 6565-1999, and years of working was categorized under Date type.

We used 46 data elements and 13 value domains to assign values to elements in the *Basic Dataset of Personal Information*. The 145 sheets were reviewed and their elements were compared with the dataset. The elements of the basic dataset were added into the sheets after a careful selection that filtered out default elements. The selection process was necessary because converting the sheets into a single complete document only required some elements in the basic dataset.

Substitute items in the basic set of medical sheets with standard data elements defined in the national EHR according to the proposed content structure

After comparison, some data elements were independent of the data groups. For example, ABO and Rh blood types are two inherent biological components that will not change in the life span of an individual. We classified blood type in group H.02.001, representing personal biological identifiers, as a subgroup of H.02, representing identifiers of target service. As a result, the data group H.02 was composed of data elements and sub-data groups.

Each data group had an ID code. The structure of the data group ID codes was designed similar to that for H.00.001, where H indicated that the group was a component of the header of a clinical document and the two-digit number after the first separator dot indicated the first level sequence. If there was a subgroup, another three-digit serial number was added after the second separator dot. The level of the sequence indicated the affiliation between data groups. For example, H.01.001 denoted the data group of personal biological identifiers, which was one of the components of the target service data groups. The initial letter could be H or S, where H represents “header” and S represents “section of the clinical document body”.

We added 5 subgroups when compared with the proposed EMR structure. The structure of the groups made it easy to add more subgroups by adding a serial number in the last three-figure number.

Seventy-five data groups were created according to the data elements in the sheet sets. Table 3 shows that subgroups could be added easily during medical development.

Data elements were defined in detail after being categorized into data groups. The cardinality of each element showed how many times it could be repeatedly used in one instance and the cardinality value was assigned according to the HL7 model. Meanwhile, we clarified their definitions and specified the data types and the formats of the data types. The values of the elements were set by coding where possible.

Finally, we obtained 451 data elements that served as the components of the data groups. Table 4 shows a part of the table of groups with the data elements, where group H.02 had two subgroups as shown in the table.

Some elements defined in the *EHR Basic Architecture and Data Standard* were defined using codes. Their codes were defined in the Code Value table with ‘CV’ followed by a series of numbers. For example, the element “ID-type code” was defined in the table as CV0100.03, in which the type of the ID code was coded and listed (see Appendix I).

The structure of the data groups could be graphically represented as shown in Fig. 3. In Fig. 3, the data groups had two levels, and one data group could be included in another data group. These were called sub-data groups. Data elements were composed the two levels of data groups and constrained by value domain.

Sub-data groups were composed of data elements alone, whereas data groups were composed of sub-data groups and data elements. The sub-data groups were used to assemble groups of relevant data that were used together.

Building the EMR basic content structure and the EMR basic content templates

Basic EMR content templates

Based on the categorization of medical activity records, seventeen templates were drafted to cover the medical

Table 3 Seventy-five data groups

No.	ID	Name	No.	ID	Name
1	**H.01**	**Document ID**	39	S.04.008	Systematic review
2	**H.02**	Identifiers of target service	40	S.04.009	Personal lifestyle
3	H.02.001	Personal biological identifiers	41	S.04.010	Marital
4	H.02.002	Personal risk identifiers	42	S.04.011	Menstrual
5	**H.03**	**Personal demographic data**	43	S.04.012	Childbearing
6	**H.04**	**Kin**	44	S.04.013	Family health
7	**H.05**	**Address**	45	S.04.014	Risk exposure
8	**H.06**	**Telecom**	**46**	**S.05**	**Examination**
9	**H.08**	**Healthcare organization**	47	S.05.001	Request
10	**H.09**	**Health service provider**	48	S.05.002	Report
11	**H.10**	**Health event summary**	49	S.05.003	Imaging report
12	**S.01**	**Main complaint(symptom/signs)**	**50**	**S.06**	**Laboratory test**
13	S.01.001	Symptom: fever	51	S.06.001	Request
14	S.01.002	Symptom: mucocutaneous hemorrhage	52	S.06.002	Report
15	S.01.003	Signs: abdominal pain	53	S.06.003	Sample
16	S.01.004	Signs: dropsy	**54**	**S.07**	**Diagnosis**
17	S.01.005	Symptom: allergic reaction/adverse reaction	**55**	**S.08**	**Operation/procedure**
18	**S.02**	**Physical examination**	56	S.08.001	Surgery
19	S.02.001	Physical examination: general condition	57	S.08.002	Anesthesia
20	S.02.002	Physical examination: skin	**58**	**S.09**	**Medication**
21	S.02.003	Physical examination: lymph nodes	59	S.09.001	Vaccination
22	S.02.004	Physical examination: head	60	S.09.002	Blood transfusion
23	S.02.005	Physical examination: neck	**61**	**S.10**	**Diagnosis and treatment plan**
24	S.02.006	Physical examination: chest	62	S.10.001	Patient alert
25	S.02.007	Physical examination: belly	63	S.10.002	Informed consent information
26	S.02.008	Physical examination: genital organs, anus, rectum	64	S.10.003	Clinical pathway
27	S.02.009	Physical examination: vertebral column,	65	S.10.004	Chinese medicine treatment
28	S.02.010	Physical examination: deformity	**66**	**S.11**	**Assessment**
29	**S.03**	**Present problem**	67	S.11.001	Therapeutic outcome
30	S.03.001	Infectious disease	68	S.11.002	Medical quality
31	**S.04**	**History of health and/or healthcare**	**69**	**S.12**	**Process of diagnosis and treatment**
32	S.04.001	Disease or trauma	70	S.12.001	Medical process
33	S.04.002	Surgery	71	S.12.002	Order
34	S.04.003	Diagnosis and treatment	**72**	**S.13**	**Medical expense**
35	S.04.004	Blood transfusion	73	S.13.001	Insurance
36	S.04.005	Immunization	74	S.14	Nursing
37	S.04.006	Allergy	75	S.15	Health advice
38	S.04.007	Medication			

sheets used in most hospitals. The relationship between the templates and medical activity records is shown in Table 5.

The templates consisted of data groups that could be used to express a specific sheet by assigning values to data elements. Table 6 is a part of the table showing templates with groups. Ten data groups were used to compose template MT01, Personal Health Summary, in Table 6, and nineteen data groups were used to compose template MT12, Inpatient Medical Process Records. The templates represented different medical activities, and their elements should be included in specific EMRS.

The template model and its implementation

Each template was composed of data groups, and its cardinality value was defined in its model (see Fig. 4).

Table 4 List of the selected data groups and elements

Data groups	Data elements' ID	Data elements	Cardinality	Definition	Data type of the data elements	Format of the data type	Value
H.02 identifiers of target service	HR01.01.002.01	ID-type code	1..1	The ID code type or documentary evidence type	S	N2	The CV0100.03 code of ID type (see Appendix I)
	HR01.01.002.02	ID-code	1..1	The ID code or documentary evidence	S	N..30	
	HR01.01.002.03	ID-valid date	0..1	The ID validity date	D	D8	
	HR01.01.002.04	ID-date of expiration	0..1	The ID expiration date	D	D8	
	HR01.01.002.05	ID-assigning authority	0..1	The name of the health authority assigning the ID	S	AN..70	
	HR02.01.001.01	name-referring to	0..1	To whom the name refers	S	A..20	The CV0100.01 code of name type.
	HR02.01.001.02	name-code of referring to	0..1	The code to whom the name refers	S	N2	The CV0100.02 code to whom the name refers.
	HR02.01.002	name	1..1	The name	S	A..30	
	HR42.01.012	patient type code	0..1	The patient type code.	S	N1	
H.02.001 personal biological identifiers	HR51.03.003	ABO blood type code	1..1	The ABO blood type code.	S	N1	The CV5103.02 ABO blood type code.
	HR51.03.004	RH blood type code	1..1	The RH blood type code.	S	N1	0-RH(−); 1-RH(+); 3-NA
H.02.002 personal risk identifiers	HR01.03.001.01	personal risk name	1..1	The personal risk name.	S	AN..20	
	HR01.03.001.02	personal risk code	1..1	The personal risk code.	S	N2	

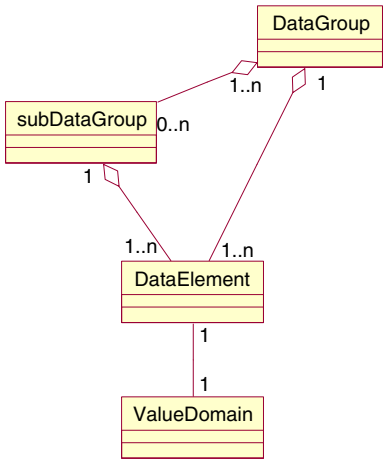

Fig. 3 The relationship of data group, data element and data value

According to its cardinality value, data group H.01 could be repeated multiple times in one personal health summary, in a similar manner in which each data element repeat time in each data group would be repeated according to its cardinality definition.

When applying a template to represent a specific medical sheet, data groups could be reused in this instance and data elements could be reused in this data group as well.

The basic EMR content structure

The basic data elements and data groups were the basic components of the EMR content structure. Figure 5 shows the relationship between the EMR, medical field, and data groups.

Discussion

The basic proposed content structure

Apart from the basic header-body structure as defined in the CDA model, the content system of diagnostics provided a basic framework for organizing clinical information. We proposed that content structure be regarded as a guideline for the collection of medical sheets in hospitals. The structure outlined the basic content of the EMR and directed the categorization of medical information. The discussion sessions with medical experts contributed to the reliability and rationality of the proposed structure. Other theory bases, excluding diagnostics, may also be adopted, but by choosing diagnostics, many actual descriptions of medical data could be borrowed to systematically organize clinical documents. Ultimately, it is easier to combine diagnostics with the CDA model.

Table 5 EMR templates

Template ID	EMR basic templates title	Catalog of medical activity records
MT01	Personal health summary	EMR0100 medical summary
MT02	Outpatient/emergency medical records	EMR020001 outpatient/emergency medical records
		EMR020002 records of temporary stay for observation
MT03	Prescription	EMR030001 prescription of Western medicine
		EMR030002 prescription of Chinese medicine
MT04	Nursing: processing records	EMR060101 general nursing records
		EMR060102 special care records
		EMR060103 surgical nursing records
		EMR060104 temperature records
		EMR060105 intake and output volume records
		EMR060106 inspection records
MT05	Nursing: nursing assessment and care plan	EMR060201 admission assessment records
		EMR060202 care plan records
		EMR060203 discharge assessment and guidance records
		EMR060204 disposable material using records
MT06	Procedure and treatment: general procedure and treatment	EMR050101 procedure records
		EMR050102 surgery records
		EMR050103 anesthesia records
		EMR050104 blood transfusion records
MT07	Procedure and treatment: labor records	EMR050201 lie-in records
		EMR050202 uterine-incision delivery records
		EMR050203 spontaneous labor records
MT08	Observation/lab test	EMR040001 observation records
		EMR040002 laboratory test records
MT09	Informed consent information	EMR070001 consent records for operation
		EMR070002 consent records for special examination and treatment
		EMR070003 consent records for special medicine or material
		EMR070004 consent records for blood transfusion
		EMR070005 records of critical illness
MT10	Inpatient medical summary	EMR080001inpatient medical summary
MT11	Inpatient registry	EMR090001 admission records
		EMR090002 discharge records during 24 h
		EMR090003 death records during 24 h
MT12	Inpatient medical process records	EMR100001 first medical process records
		EMR100002 general daily process records
		EMR100003 senior doctor investigation records
		EMR100004 complex cases discussion records
		EMR100005 handing-over records
		EMR100006 departmental transfer records
		EMR100007 episode summary records
		EMR100008 rescuing records
		EMR100009 consultation records
		EMR100010 pre-surgery records
		EMR100011 pre-surgery discussion records
		EMR100012 after-surgery records
		EMR100013 discharge summary records
		EMR100014 death records
		EMR100015 death discussion records
MT13	Inpatient medical orders	EMR110001 standing orders

Table 5 (continued)

Template ID	EMR basic templates title	Catalog of medical activity records
		EMR110002 stat orders
MT14	Discharge records	EMR120001 discharge records
MT15	Transfer records	EMR130001 transfer records
MT16	Referral records	EMR140001 referral records
MT17	Healthcare provider	EMR150001 healthcare provider records

Obtaining information from the medical sheets used in hospital

Collecting medical sheets from hospitals

China has more than 18,000 hospitals above the county-level classification, among which are over 800 3-A hospitals. The hospitals involved in this research were suitably representative of all hospitals in China. First, the geographical locations of these hospitals were distributed among nine provinces in northern and southern China. Also, four of these hospitals were very large-scale hospitals with over 2,000 beds. Second, China has a unitary medical system, and all hospitals operate with similar management concepts with

Table 6 Partial templates and relevant data groups

Template ID	EMR basic templates title	Composed data groups	Medical activity records
MT01	Personal health summary	H.01 document ID	EMR010001 patient identifier information
		H.02 identifiers of target service	EMR010002 patient health record
		H.02.001 personal biological identifiers	EMR010003 summary of health event
		H.02.002 personal risk identifiers	EMR010004 medical expense record
		H.03 personal demographic data	
		H.04 kin	
		H.05 address	
		H.06 telecom	
		H.07 insurance	
		H.10 health event summary	
		S.04.001 history of health and/or healthcare: disease or trauma	
		S.09 medication	
		S.13 medical expense	
MT12	Inpatient medical process records	H.01 document ID	EMR100001 first medical process records
		H.02 identifiers of subject of care	EMR100002 general daily process records
		H.03 personal demographic data	EMR100003 senior doctor investigation records
		H.08 healthcare organization	EMR100004 complex cases discussion records
		H.09 health service provider	EMR100005 handing-over records
		H.10 health event summary	EMR100006 department transfer records
		S.01 main complaint (symptom/signs)	EMR100007 episode summary records
		S.02 physical examination	EMR100008 rescuing records
		S.03 present problem	EMR100009 consultation records
		S.04 history of health and/or healthcare	EMR100010 pre-surgery records
		S.05 examination	EMR100011 pre-surgery discussion records
		S.06 laboratory test	EMR100012 after-surgery records
		S.07 diagnosis	EMR100013 discharge summary records
		S.08 operation/procedure	EMR100014 death records
		S.09 medication	EMR100015 death discussion records
		S.10 diagnosis and treatment plan	
		S.11 assessment	
		S.12 process of diagnosis and treatment	
		S.15 health advice	

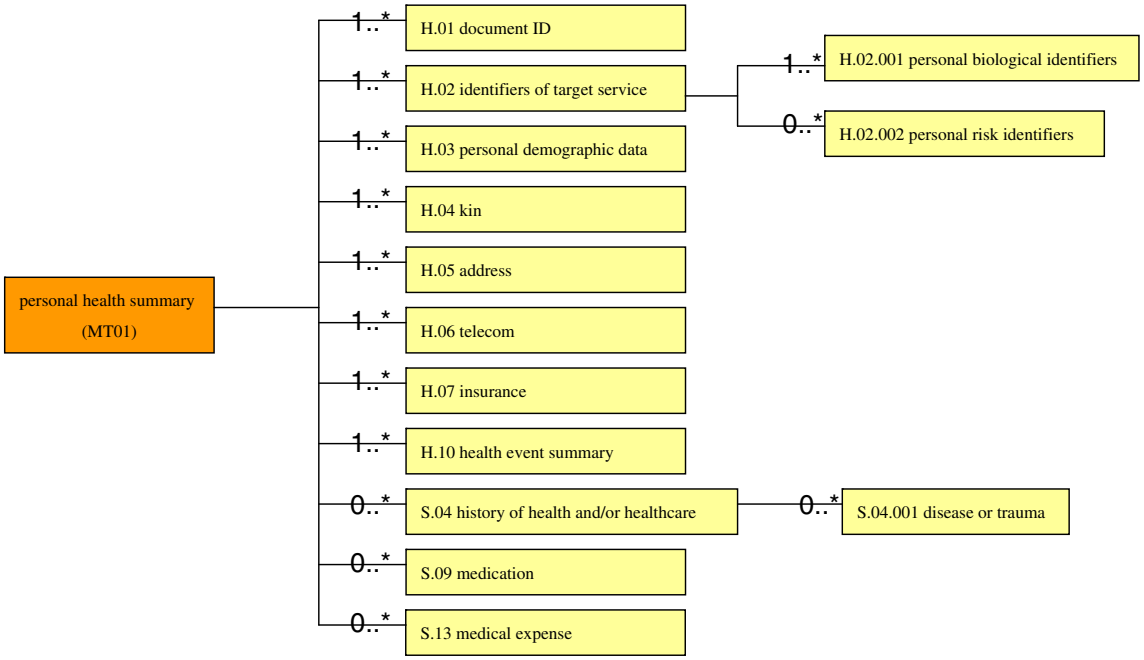

Fig. 4 The model of medical template MT01

little variety found among them. Third, the selected 17 hospitals are digital hospitals as certified by the Ministry of Health in December 2008, and the type of medical information that their medical sheets represented might be better presented in hospitals that use more information technology.

Merging and creating a preliminary set of medical sheets in hospitals

Our collection was one of several surveys in Chinese hospitals collecting medical profession sheets. We collected as many sheets as possible through the survey and found

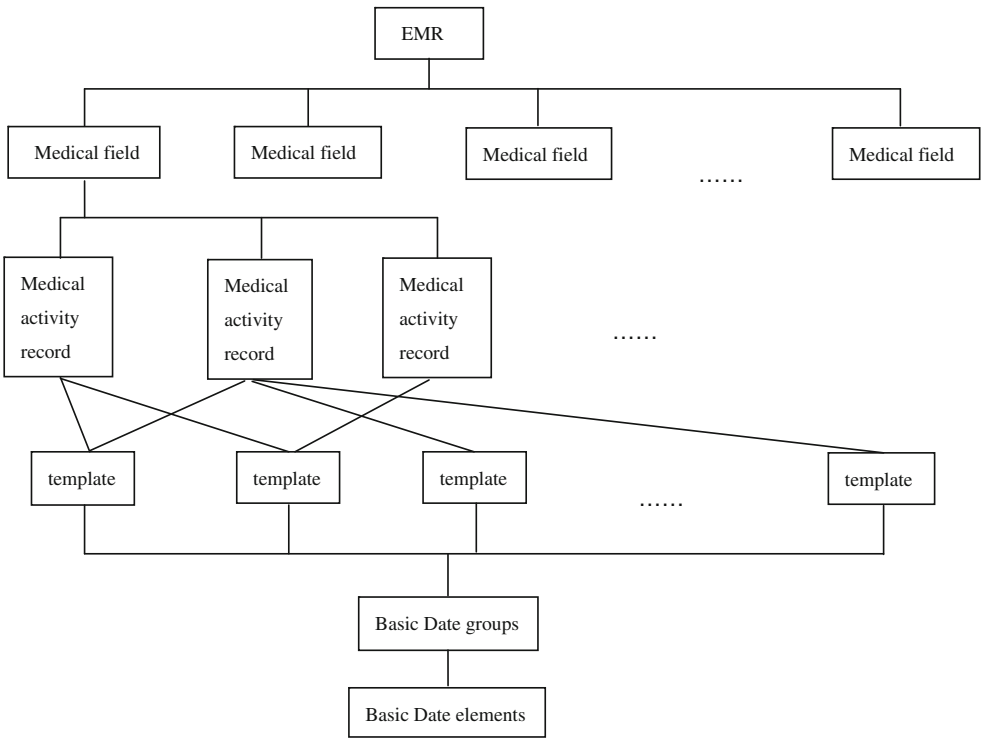

Fig. 5 The relationships among the EMR, medical field, and data groups

three characteristics of the medical profession in hospitals: (1) the medical profession bears great similarities in most Chinese hospitals, (2) Chinese medicine hospitals have adopted Western medicine to a great extent, and (3) most hospitals did not consider information exchange among different healthcare providers.

The set of medical sheets resulting from this research reflected information cataloging in hospitals in general. The set that we collected will also provide a useful reference for future EMR research and development.

Checking on the integration of data elements

Because most hospitals did not consider information exchange among different healthcare centers, almost all sheets lacked unified identification. The ID number of different sheets in one hospital may be generated by different systems or in different ways from the method used in a different hospital. The process of complementation elements for medical sheets was constrained by the national data element standard, the Basic Dataset of Person Information, assuring the interoperability of information among local EMRS and regional EHR platforms [15].

The regional EHR platform was a national EHR architecture composed of three levels: national, provincial and regional or city levels. The building of a regional EHR platform has begun in several regions with the goal of sharing personal health information regionally and (ultimately) nationally. The EMR were data sources for the regional EHRS; therefore, consistency with EHR standards was a basic requirement for EMR standards.

Comparison of proposed EMR content structure with the basic medical sheets

The proposed EMR content structure influenced the categorization of items on medical sheets. This process verified the capacity of the content structure and to what extent the vessel could cover all items from those original data sheets from the hospitals. The matching process might lead to the addition of data groups or data elements. Eventually, this research sought to build a universal structure composed of data groups, and the data groups could append data elements when necessary. When data groups became the basic information units of the EMR content structure, their values were assigned by acquiring data elements.

The relationship between EMR templates and data groups

Data groups were used to rebuild the real medical record sheets and form the templates of the activity sheet. The advantage of the templates was their practicality. By assigning values to data elements and data groups, the abstract EMR structure became concrete and could be used as a guideline in building EMRS.

Future research will focus on building standards and rules for EMRS in hospitals. The standards will be applied to the EHR architecture at national, provincial, and regional levels. This will also facilitate cross-institutional sharing of medical information.

Conclusion

We proposed an EMR basic content structure with a basic set of medical record sheets, a basic set of data groups with data elements, and finally a basic set of templates. The proposed structure covers most information that hospitals collect regarding their patients, and the universality of structure and concise table format allow for its implementation in specific EMRS. The EMR basic content structure will help to promote the standardization of the EMR system and aid in building local EHR platforms in China.

Acknowledgments This research was supported by National Science and Technology Support Project Grant (No. 2008BAI52B01) from the Ministry of Science and Technology of the People's Republic of China. Special thanks are extended to the hospitals that participated in the collection of the sheets, providing valuable material. We wish to acknowledge the support of Dr. Haijing Tu and editors of American Journal Experts for their assistance in the writing.

Appendix I

Table 7 Table of CV0100.03: ID-type codes

Code	Definition	Description
01	Residential ID card	default
02	Travel passes for Taiwan residents to enter or leave the mainland	
03	Hong Kong/Macao identity card	
04	Household register certificate	
05	Passport	
06	Military officer card	
07	Military staff card	
08	Soldiers card	
09	Driver's license	
10	Handicap certificate	
11	Medical insurance card	
12	Birth certificate	
13	Family id	
14	Inpatient medical record	
15	In-patient bed no.	
16	Outpatient medical record	
91	Death certificate	
99	Other id	

References

1. Jha, Ashish K., DesRoches, Catherine M., Campbell, Eric G., Donelan, Karen, Rao, Sowmya R., Ferris, Timothy G., Shields, Alexandra, Rosenbaum, Sara, and Blumenthal, David, Use of Electronic Health Records in U.S. Hospitals. *The New England Journal of Medicine* 360:1628–1638, 2009.
2. Zhang, Yuhai, Xu, Yongyong, Shang, Lei, and Rao, Keqin, An investigation into health informatics and related standards in China. *International Journal of Medical Informatics* 76:614–620, 2007.
3. HL7 Health Level Seven. Available on World Wide Web at: http://www.hl7.org/, Last date visited:2010-01-12.
4. Dolin, R. H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F. M., Biron, P. V., and Shabo, S. A., HL7 Clinical Document Architecture, Release 2. *J. Am. Med. Inform. Assoc.* 13:30–39, 2006.
5. ASTM Committee E31 on Healthcare Informatics develops standards. Available on World Wide Web at: http://www.astm.org/cgi-bin/SoftCart.exe/COMMIT/COMMITTEE/E31.htm?L+mystore+fdxl9438, Last date visited: 2010-01-12.
6. Kibbe, D. C., Phillips, J., and Green, L. A., The continuity of care record. *Am. Fam. Physician* 70:1220–1223, 2004.
7. IHE IT Infrastructure Technical Framework, Volume 2a (ITI TF-2a), Transactions Part A—Sections 3.1–3.28, Revision 6.0, Final Text, August 10, 2009. Available on World Wide Web at: http://www.ihe.net/Technical_Framework/upload/IHE_ITI_TF_6-0_Vol2a_FT_2009-08-10-2.pdf. Last date visited: 2010-01-12.
8. Jian, Wen-Shan, Hsu, Chien-Yeh, Hao, Te-Hui, Wen, Hsyien-Chia, Hsu, Min-Huei, Lee, Yen-Liang, Li, Yu-Chuan, and Chang, Polun, Building a portable data and information interoperability infrastructure—framework for a standard Taiwan Electronic Medical Record Template. *Computer Methods and Programs in Biomedicine* 88:102–111, 2007.
9. Interviewing Keqin Rao, the director of statistic department of Ministry of Health. 2009. Available on the World Wide Web at : http://www.h-ceo.com/doc/2009/04/15749.shtml. Last date visited: 2010-1-12(Chinese).
10. Use of Hospital Information System in China. Available on World Wide Web at: http://www.cio360.net/Page/1806/InfoID/294896/SourceId/13718/PubDate/2009-04-16/Default.aspx. Last date visited: 2010-01-12(Chinese).
11. The HIS functional guideline. Available on World Wide Web at: http://wsj.maoming.gov.cn/new_page_1.htm. Last date visited: 2010-01-12(Chinese).
12. Chen, Wenbin, and Pan, Xianglin, *Diagnostics. Version 7*. People Health Press, BeiJing, pp. 6–551, 2008.
13. Medical Document Rule, Ministry of Health. Available on World Wide Web at: http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohyzs/s3585/200904/40299.htm. Last date visited: 2010-01-15.
14. The basic Architecture and Data Standard of EHR, Ministry of Health. Available on World Wide Web at: http://www.chima.org.cn/pe/Article/ShowArticle.asp?ArticleID=763. Last date visited: 2010-01-12(Chinese).
15. The Draft of Construction Guidelines for Health Information Platform basing on EHR, Ministry of Health. Available on World Wide Web at: http://www.chima.org.cn/pe/Article/ShowArticle.asp?ArticleID=764. Last date visited: 2010-01-12(Chinese).