

# The Development of Clinical Document Standards for Semantic Interoperability in China

Peng Yang, MD<sup>1</sup>, Feng Pan, MD<sup>1</sup>, Danhong Liu, PhD<sup>1</sup>, Yongyong Xu, PhD<sup>1</sup>, Yi Wan, PhD<sup>1</sup>, Haibo Tu, PhD<sup>1</sup>, Xuejun Tang, MD<sup>2</sup>, Jianping Hu, MD<sup>2</sup>

<sup>1</sup>Institute for Health Informatics, Fourth Military Medical University, Xian; <sup>2</sup>Center for Health Statistics and Information, Ministry of Health, Beijing, China

**Objectives:** This study is aimed at developing a set of data groups (DGs) to be employed as reusable building blocks for the construction of the eight most common clinical documents used in China's general hospitals in order to achieve their structural and semantic standardization. **Methods:** The *Diagnostics* knowledge framework, the related approaches taken from the Health Level Seven (HL7), the Integrating the Healthcare Enterprise (IHE), and the Healthcare Information Technology Standards Panel (HITSP) and 1,487 original clinical records were considered together to form the DG architecture and data sets. The internal structure, content, and semantics of each DG were then defined by mapping each DG data set to a corresponding Clinical Document Architecture data element and matching each DG data set to the metadata in the Chinese National Health Data Dictionary. By using the DGs as reusable building blocks, standardized structures and semantics regarding the clinical documents for semantic interoperability were able to be constructed. **Results:** Altogether, 5 header DGs, 48 section DGs, and 17 entry DGs were developed. Several issues regarding the DGs, including their internal structure, identifiers, data set names, definitions, length and format, data types, and value sets, were further defined. Standardized structures and semantics regarding the eight clinical documents were structured by the DGs. **Conclusions:** This approach of constructing clinical document standards using DGs is a feasible standard-driven solution useful in preparing documents possessing semantic interoperability among the disparate information systems in China. These standards need to be validated and refined through further study.

**Keywords:** Electronic Health Records, Standards, Health Level Seven, Information Systems

**Submitted:** November 18, 2011

**Revised:** December 13, 2011

**Accepted:** December 23, 2011

## Corresponding Authors

Danhong Liu, PhD and Yongyong Xu, PhD

Institute for Health Informatics, Fourth Military Medical University,  
Changle West Road 169, Xian 710032, PR China. Tel: +86-29-8477-  
2180, Fax: +86-29-8477-4858, E-mail: liudanh@fmmu.edu.cn,  
xuyongy@fmmu.edu.cn

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2011 The Korean Society of Medical Informatics

## 1. Introduction

In China, hospital information systems have been developed for more than 30 years and have gone through 4 application stages: single computer, department level, hospital-wide level and the current regional health information network level [1]. Because semantic interoperability was rarely considered in the development of the first three stages, the majority of encounter information, such as patient identifiers, demographic information, main patient problems, diagnoses, observations, medications, procedures, assessments, and expenditures could only be shared and exchanged within a specific hospital and could not be shared or exchanged be-

tween hospitals or external health institutions [2,3]. Therefore, developing standards for clinical information to promote semantic interoperability has become a priority in implementing the China National New Health Reform [4]. Thus far, 8 clinical documents, which are the most commonly used in China's general hospitals, have been medically identified with normalized contents and issued by the Ministry of Health (MOH) [5].

To be exchangeable, there must be standards for these documents to support semantic interoperability. Fortunately, several organizations have developed relevant standards, such as templates for the Continuity of Care Document (CCD) [6] in the Health Level Seven (HL7) [7], content modules [8,9] of Patient Care Coordination (PCC) in Integrating the Healthcare Enterprise (IHE) [10], and content modules of the Healthcare Information Technology Standards Panel (HITSP) [11], which were all based on the HL7 Clinical Document Architecture, Release Two (CDA R2) [12,13]. However, we cannot use them directly in our clinical documents because their development backgrounds and application conditions are different from those used in China. One difference is that the content of the templates and content modules do not suit our needs completely. Some information, such as medical expenses, administrative use, and quality assessment, are not present in the templates or content modules. Another difference is in the coding of value sets. Taking gender as an example, the codes in HL7 are "F = Female, M = Male, and UN = Undifferentiated", whereas the codes in China are "0 = unknown, 1 = male, 2 = female, and 9 = unaccounted." These codes have been widely used across China and have been a national standard (GB/T2261.1).

In this study, which is based on clinical record sheets in China's hospitals and references approaches from the HL7 CCD, the IHE PCC and the HITSP, we attempt to further develop a set of data groups (DGs) based on the CDA R2 as reusable building blocks to construct the 8 most common clinical documents in China's general hospitals. This will allow for structural and semantic standardization and promote interoperability.

## II. Methods

### 1. The Contents of Clinical Documents

The 8 most common clinical documents in China's general hospitals are: 1) the outpatient medical record summary; 2) the emergency medical record summary; 3) the inpatient medical record summary; 4) the basic medical synopsis (a brief summary of medical activities concerning the evolution of illness, including examining, diagnosing, and treat-

ing); 5) the inpatient outline (summary information during a hospital stay, which is usually as the first page of a paper-based medical record after discharge); 6) the discharge summary; 7) the referral summary; and 8) the labor and deliver record summary.

### 2. Chinese Health Data Dictionary

The Chinese National Health Data Dictionary (CNHDD) is a metadata repository that must comply with the standards for the construction of databases and health information systems. The metadata in the CNHDD was generalized and abstracted from various health information systems and legacy systems, with each metadata describing attributes of data identification, definitions, collection, usage guides, references and administration. At present, more than 1,500 metadata are available, and these metadata can be browsed by visiting the website described in [14]. In this research, we acquired standardized contents of each data item in the DGs by matching each data item with the metadata in the CNHDD.

### 3. Formulation Process of the DGs and Clinical Documents

1) Step 1: Development of the DGs' architecture and contents  
First, 1,487 original clinical record sheets from 14 representative general hospitals, including 4 hospitals with more than 2,000 beds, 6 hospitals with 1,000-2,000 beds and 4 hospitals with 500-1,000 beds across the country were collected. After merging the original sheets and removing redundant elements, 145 clinical record sheets were formed [15]. This dramatic reduction in the number of elements resulted from similar clinical procedures in most hospitals. Second, the framework of *Diagnostics* [16] knowledge and the approaches of the HL7 CCD, the IHE PCC and the HITSP for assembling templates and modules were considered together to propose the DGs' architecture. Third, the proposed DGs were used to construct 145 clinical record sheets as a pilot study to test their integrality. If the DGs could not completely build these sheets, the DGs were returned to a redefinition process. Lastly, data items within the DGs were identified by combining data items from the original sheets with the DG's architecture. Data items from the sheets were categorized and arranged in their related DGs, and data items having similar properties were abstracted. For example, data items *B-mode ultrasonography examination ID*, *X-ray examination ID*, *CT examination ID* and other examinations IDs were abstracted to two data items of *examination type* and *examination ID*. The *B-mode ultrasonography*, *X-ray*, *CT*, and other examinations became the codes in the value set for examina-

tion type after the abstraction.

## 2) Step 2: Definition of the structure, content and semantics of each DG

In this study, the HL7 CDA was chosen as our standard to represent the semantics of DGs and clinical documents for two reasons: 1) the HL7 CDA is a document markup standard that specifies the structure and semantics of a clinical document for the purpose of exchange [13], which suits our needs, and 2) the HL7 CDA has been chosen as the data exchange standard by MOH in the Technology Solution of Establishing Hospital Information Platform for Electronic Health Record (EHR) in China; thus, our standards should comply with the MOH standards [17].

By mapping each data item of a DG to the corresponding data element in the HL7 CDA, the structure of the DG was acquired. By matching each data items of the DG with the metadata in the CNHDD, a standardized description of the DG's items were obtained. Based on both above results, the DGs' semantics were defined. All the data items in the DGs have corresponding data elements in the CDA, and 90% of the data items were standardized directly by matching them with the CNHDD.

## 3) Step 3: Construction of each clinical document with the DGs

If one or more data item in each clinical document was

found in a DG, they were replaced by the DG. Thus, the contents of the clinical document were changed from being comprised of data items to being comprised of DGs, upon which the structure and semantics of the clinical document were finally produced.

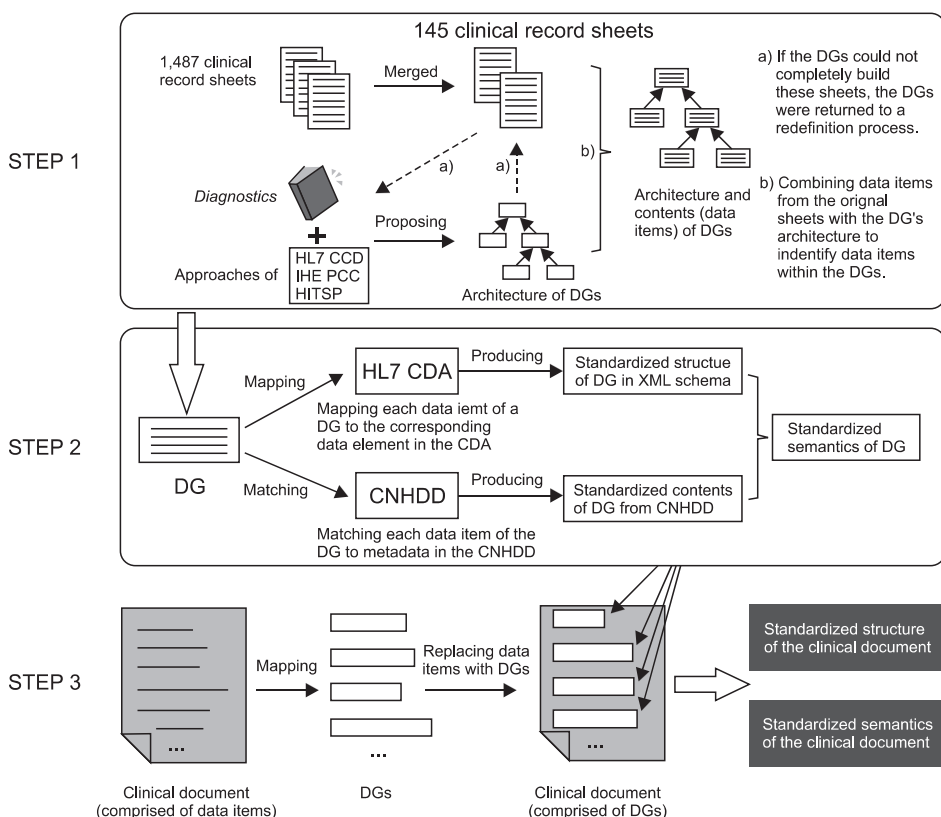
During the formulation process, 4 discussion meetings were held to discuss the integrity and rationality of the developed DGs and the accuracy and significance of the clinical documents structured by the DGs. Altogether, 25 people participated in the consultations, including MOH leaders, health information experts, senior physicians, surgeons and software development engineers. The formulation did not proceed to the next step unless results of the current step were approved by 95% of those consulted.

The formulation process of the DGs and clinical documents are shown in Figure 1.

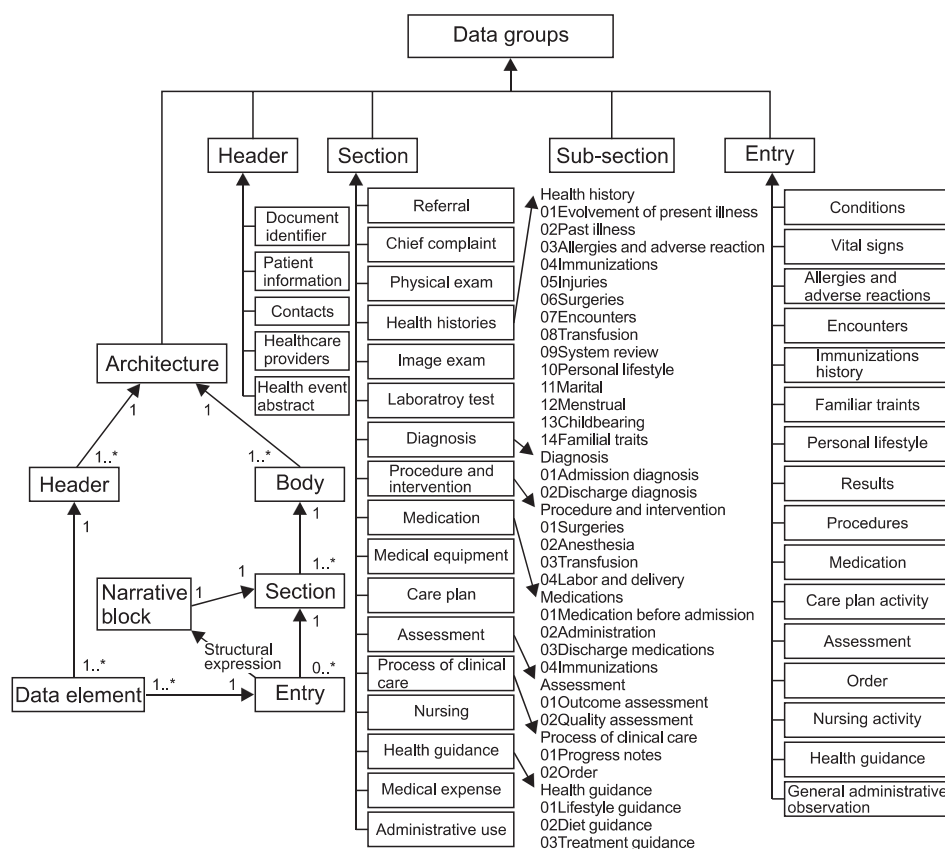
## III. Results

### 1. The Architecture and Contents of the DGs

Altogether, 5 header DGs and 65 body DGs, including 48 section DGs and 17 entry DGs, were proposed. The section DGs consisted of 17 section DGs and 31 sub-section DGs. Of the section DGs, *Health Histories*, *Diagnosis*, *Procedure and Intervention*, *Medications*, *Assessment*, *Process of Clinical Care* and *Health Guidance* all contained



**Figure 1.** The formulation process of the data groups (DGs) and clinical documents, which includes three steps: development of the DGs' architecture and contents, definition of the structure, content and semantics of each DG, and construction of each clinical document with DGs. HL7: Health Level Seven, IHE: Integrating the Healthcare Enterprise, HITSP: Healthcare Information Technology Standards Panel, CNHDD: Chinese National Health Data Dictionary, CDA: Clinical Document Architecture.



**Figure 2.** The architecture and contents of the data groups (DGs). Altogether, 5 header DGs, 48 section DGs and 17 entry DGs were proposed. Each DG contains one or more data items. A body consists of one or more section DGs. A section DG contains a single narrative block and zero or more entry DGs which represent the narrative block by structured data items.

sub-section DGs (Figure 2).

Each DG conveys specific information. A header DG conveys identification information for documents, patients and involved providers. A body DG comprised of relevant section DGs conveys clinical report information. A section DG contains a single narrative block and possible (zero or more) entry DGs representing narrative content by structured data items (Figure 2). Thus far, narrative blocks of most of the section DGs can be represented by entry DGs, except for the *Referral*, *Medical Equipment Use*, *System Review*, *Marital History*, *Menstrual History*, *Childbearing History* and *Progress Note* narrative blocks. More entry DGs will be developed to represent these unstructured section DGs in future studies.

## 2. The Internal Structure, Content and Semantics of DGs

### 1) The standardized structure of DGs

The contents of 5 header DGs were structured by 12 data elements in the HL7 CDA. Data elements of typeId, templateId, id, code, title, effectiveTime, confidentialityCode and author were combined to represent the DG *Document Identifier*, recordTarget represents *Patient Information*, participant represents Contacts, documentation of represents *Healthcare Providers* and component of represents *Health Event Abstract*.

The contents of 65 body DGs were structured by the ele-

ments within component. Each section DG has one or more templateIds specifying its identifier, a code specifying the type of narrative block with Logical Observation Identifiers Names and Codes (LOINC) [18], a text that describes the content of a narrative block, and possible entry DGs representing the narrative block of structured data items. When matching to LOINC, most narrative blocks have matching LOINC codes, especially for those related to laboratory tests. For a few narrative blocks that have complex contents and cannot be matched completely with a LOINC code, we split them into several simple parts that have specific LOINC codes to be matched and use several components to represent them accordingly.

Meanwhile, the data items of 17 entry DGs are represented by data elements of the CDA classes act, encounter, observation, organizer, procedure, substanceAdministration and supply. The 16 former entry DGs are used to describe information related to clinical activities, while the last entry DG, *General Administrative Observation*, is developed exclusively to describe information for hospital management, such as the length of hospital stay, the cure rate, and the death rate.

### 2) Standardized contents of the DGs from the CNHDD

Standardized metadata attributes of the data items in the DGs were acquired after matching each data item of a DG

Table 1. The semantics description of the entry DG allergies and adverse reactions

Data item		Definition	Length and format	Health Level 7 (HL7) clinical document architecture				
				Level <sup>a</sup>	Parent/element	Card.	Element's attribute	Data type
Allergy time				n + 1	Act/templateId	1..*	Root	II
								EHR.ETY.03
Allergy substance	03.01	The date this allergy symptom occurred	YYYY	n + 2	Observation/EffectiveTime	0..1		TS
	03.02	The code of the substance causing allergy	NN	n + 2	Observation/code	1..1		CE
	03.03	Name of materials that lead to allergy	A[A(30)]	n + 6	Observation/Participant /.../PlayingEntity/name	0..*		ST
Allergy symptom	03.04	Code of materials that lead to allergy	N[N(2)]	n + 6	PlayingEntity/code	0..*		CE
								VD.ETY.03.02 code of medicine causing allergy
								VD.ETY.03.03 code of food causing allergy
Severity	03.05	Code that describing the reaction symptom	NN	n + 4	Observation/value	1..1		CE
								VD.ETY.03.04 code of reaction symptom
Severity	03.06	Code identifying the severity of allergy	N	n + 4	Observation/value	1..1		CE
								VD.ETY.03.05 severity code

<sup>a</sup>n represents the level of class, "plus 1" means the level increases by 1. <sup>b</sup>Data groups (DGs) and value sets were all assigned identifiers. A DG was assigned with 8 or 10 letters. The first 3 letters was "EHR" and the 4th, 5th, and 6th was the DG's type containing three-letter (HRD, SEC or ETY). The last 2 or 4 letters were the sequence number. A value set was assigned an identifier with 9 letters. The first 2 letters are the "VD" and the 3rd -7th are the DG's type and its sequence number. The last 2 letters are the sequence number of this value set.



with corresponding data elements in the CNHDD. The matched data items are instances or specializations of the data elements in the CNHDD. For example, the data item *doctor's name* is an instance of the data element *name*, and the *date of allergy* is an instance of *date*. During the matching process, 95% of the data items have direct corresponding matches in the CNHDD, and 5% of the data items cannot be matched or are only mapped to codes in the value sets. Regarding these problems, the data items in the DGs are returned to the redefinition process, or the data elements and value sets in the CNHDD are added or adjusted after discussions with experts in developing and maintaining the CNHDD.

### 3) Standardized semantics of the DGs

Based on standardized structure and content, the semantics of each DG were acquired. For example, Table 1 shows the matched standardized metadata attributes of the entry DG *Allergies and Adverse Reactions* and its representation structured by the HL7 CDA. The values of attributes (including *definition*, *length and format*, *data type* and *value set*) for the data items are derived from the CNHDD. In line with the CNHDD, the contents of *Parent/element*, *card. (cardinality)*, *element's attribute and value* are defined and represented by the HL7 class of act and nested observation. Meanwhile, the relationships of *allergy substance*, *symptom* and *severity* are connected by the element entryRelationship, and their rela-

tionships are specified by MFST and SUBJ.

*Length and format* is described in the same manner as the descriptions used in METeOR [19]; *data type* is the HL7 Version 3 data type [6], and *value set* is the code collection for data item whose data type is CE. In addition, the value sets were standardized by referring to the ISO/IEC 11179-3 [20]. Their coded values are defined according to the sequence of a national standard (e.g., *sex code* from GB/T2261.1-2003), several code systems (e.g., *diagnosis code* from ICD-10), 8 clinical documents, collected clinical record sheets and the CNHDD.

Using instance data, an XML file of the DG can be produced. Figure 3 shows the XML file of the entry DG *Allergies and Adverse Reactions* with actual data (allergy substance -penicillin, allergy symptom-hives).

### 3. Semantics of the Clinical Documents Structured by DGs

One or more data item in each clinical document can be mapped to corresponding data items in the DG and then replaced with that DG. For example, the data item *provider's hospital name* and *provider's department name* were replaced with *Healthcare Providers* (EHR.HRD.04). *Type of laboratory test*, *name of laboratory test*, *value of laboratory test and measurement unit* were replaced with the section DG *Laboratory Test* (EHR.SEC.06). Finally 8 clinical documents were all structured by a number of DGs (Table 2), based on which standardized structures and semantics of the documents for

```

<entry>
  <act classCode="ACT" moodCode="EVN">
    <templateId root="EHR.ETY.03 ">
      <!-- **** allergy substance -penicillin **** -->
      <observation classCode="OBS" moodCode="EVN">
        <code code='01' displayName='drug allergy'
codeSystem="VD.ETY.03.01" codeSystemName='allergen code' />
        <participant typeCode="CSM">
          <participantRole classCode="MANU">
            <playingEntity classCode="MMAT">
              <name>Penicillin</name>
              <code code='02' displayName='Penicillin'
codeSystem="VD.ETY.03.02" codeSystemName='code of medicine causing allergy' />
            </playingEntity>
          </participantRole>
        </participant>
      <!-- **** allergy symptom - hives **** -->
      <entryRelationship typeCode="MFST" >
        <observation classCode="OBS" moodCode="EVN">
          <code xsi:type="CD" code="04" displayName="Hives"
codeSystem="VD.ETY.03.04" codeSystemName="code of reaction symptom" />
        </observation>
      </entryRelationship>
    </observation>
  </act>
</entry>

```

Figure 3. An XML instance of the entry data group (DG) *Allergies and Adverse Reactions*. The allergy substance-penicillin and allergy symptom-hives were represented by the element observation nested in the element act respectively. Their relationship was connected by the element entryRelationship.

Table 2. The clinical documents structured by a number of DGs

Document name (TemplateId)	S	Names of data groups <sup>a</sup>
Outpatient/emergency medical record summary (EHR.DOC.01/ EHR.DOC.02)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Healthcare Providers <sup>R</sup>
	B	Chief Complaint <sup>R</sup> , Physical Exam <sup>R2</sup> , Evolvement of Present Illness <sup>R</sup> , Past Illness <sup>R2</sup> , Allergies and Adverse Reactions <sup>R</sup> , Exam <sup>R2</sup> , Laboratory Test <sup>R2</sup> , Diagnosis <sup>R</sup> , Procedure and Intervention <sup>O</sup> , Medication <sup>O</sup> , Care Plan <sup>O</sup> , Health Guidance <sup>O</sup>
Inpatient medical record summary (EHR.DOC.03)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Contacts <sup>R2</sup> , Healthcare Providers <sup>R</sup> , Health Event Abstract <sup>O</sup>
	B	Referral <sup>O</sup> , Chief Complaint <sup>R</sup> , Physical Exam <sup>R2</sup> , Evolvement of Present Illness <sup>R</sup> , Past Illness <sup>R2</sup> , Allergies and Adverse Reactions <sup>R</sup> , Exam <sup>R</sup> , Laboratory Test <sup>R</sup> , Diagnosis <sup>R</sup> , Procedure and Intervention <sup>O</sup> , Medication <sup>R</sup> , Medical Equipment <sup>O</sup> , Care Plan <sup>O</sup> , Assessment <sup>R2</sup> , Process of Clinical Care <sup>O</sup> , Nursing <sup>R2</sup> , Health Guidance <sup>O</sup>
Basic medical synopsis (EHR.DOC.04)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Contacts <sup>R2</sup> , Health Event Abstract <sup>R</sup>
	B	Evolvement of Present Illness <sup>R</sup> , Past Illness <sup>R2</sup> , Allergies and Adverse Reactions <sup>R</sup> , Medication <sup>R</sup>
Inpatient outline (EHR.DOC.05)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Contacts <sup>R</sup> , Healthcare Providers <sup>R</sup> , Health Event Abstract <sup>O</sup>
	B	Evolvement of Present Illness <sup>R</sup> , Allergies and Adverse Reactions <sup>R</sup> , Laboratory Test <sup>R</sup> , Diagnosis <sup>R</sup> , Surgeries <sup>R2</sup> , Anesthesia <sup>R2</sup> , Transfusion <sup>R2</sup> , Assessment <sup>R</sup> , Medical Expense <sup>R</sup>
Discharge summary (EHR.DOC.06)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Healthcare Providers <sup>R</sup> , Health Event Abstract <sup>O</sup>
	B	Diagnosis <sup>R</sup> , Care Plan <sup>O</sup> , Order <sup>R</sup> , Health Guidance <sup>O</sup> , Lifestyle Guidance <sup>O</sup> , Diet Guidance <sup>O</sup> , Treatment Guidance <sup>O</sup>
Referral summary (EHR.DOC.07)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Contacts <sup>R2</sup> , Healthcare Providers <sup>R2</sup> , Health Event Abstract <sup>R2</sup>
	B	Referral <sup>R</sup> , Diagnosis <sup>R</sup> , Process of Clinical Care <sup>R</sup>
Labor and deliver record summary (EHR.DOC.08)	H	Document Identifier <sup>R</sup> , Patient Information <sup>R</sup> , Healthcare Providers <sup>R2</sup> , Health Event Abstract <sup>O</sup>
	B	Diagnosis <sup>R</sup> , Past Illness <sup>R2</sup> , Allergies and Adverse Reactions <sup>R</sup> , Menstrual History <sup>R</sup> , Childbearing History <sup>R</sup> , Exam <sup>O</sup> , Laboratory Test <sup>R</sup> , Diagnosis <sup>R</sup> , Labor and Delivery <sup>R</sup> , Medication <sup>R</sup> , Assessment <sup>R2</sup> , Process of clinical care <sup>O</sup>

S: structure, H: header; B: body, DGs: data groups, EHR: Electronic Health Record, HL: Health Level.

<sup>a</sup>Each DG was given the attribute of optionality (R, R2, or O). In data exchange, R means the contents of DGs must be sent R2 means if the content is known, it must be sent, and if it is not known, it need not be sent O means the DGs' content may be sent selectively according to the actual medical requirements.

semantic interoperability were produced in the HL7 XML format. For example, Figure 4 shows the detailed structure and semantics of *outpatient medical record summary document* structured by 3 header DGs and 12 section DGs in XML Schema.

## IV. Discussion

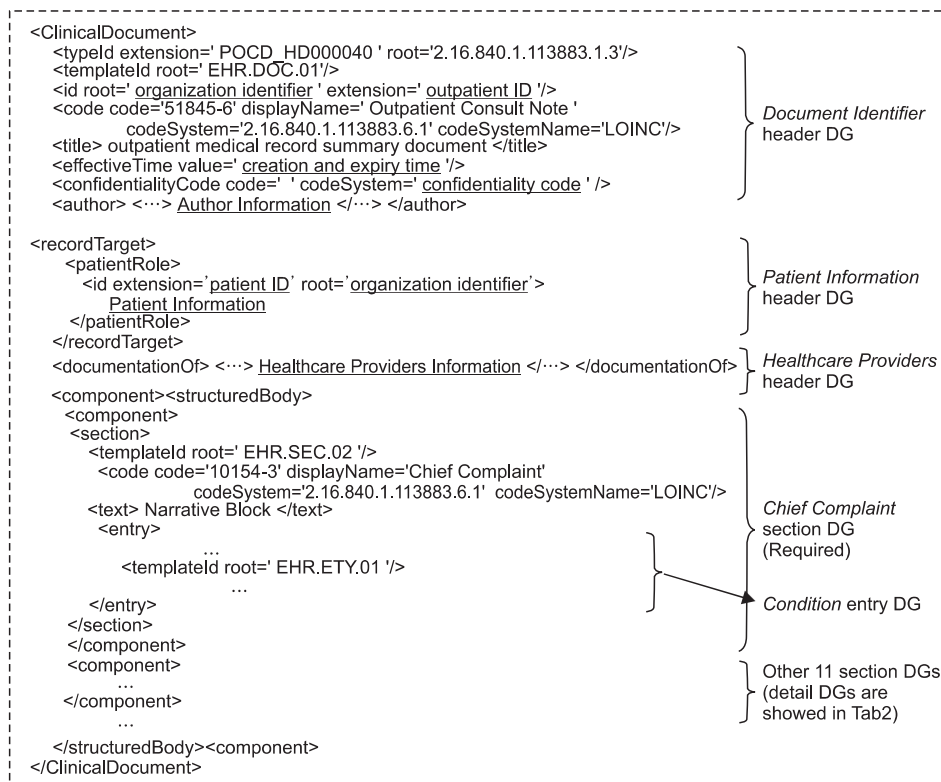
### 1. Localization of the HL7 Standards in Our Research

The methodology of building shareable clinical documents using the HL7 CDA is a recognized solution [21,22], yet we do not use it completely because the attributes of the data elements and codes of the value sets in the HL7 do not com-

pletely suit our needs. Therefore, we customized the HL7 CDA in two ways in our research. According to our business needs and on the condition that the architecture of the HL7 CDA remains unchanged, one way to customize the HL7 CDA was to adjust the attributes (e.g., cardinality, data type) of data items in the DGs, and the other method was to redefine codes of parts of value sets. Therefore, a contribution of our research is to promote the use of the HL7 standards in China.

### 2. Characteristics of DG-based Clinical Documents

Based on DGs, clinical documents have certain characteristics. First, clinical documents constructed by DGs will



be structured, enabling embedded information to be more complete and accurate [23,24]. Second, more than just these eight clinical documents can be built by flexibly reusing the DGs. The architecture of the DGs can stay stable merely by adding codes to value sets and adjusting the data items' attributes in the DGs when more documents need to be built. Third, DGs and data items that are irrelevant to clinical documents will be excluded by defining their attributes of *optionality* and *cardinality*, which can keep clinical documents clear and concise.

### 3. Differences between DGs and the Components of CCD, PCC and HITSP

Almost all the sections of the CCD, the PCC and the HITSP can be matched to corresponding section DGs except for two: the section describing medical care expenses and the section representing hospital management information. The CCD, the PCC and the HITSP use *Payers section* to specify organizations or individuals who may pay for a patient's healthcare, while we use *Medical Expense section* to describe actual expenditures that have been paid. We use *Administrative Use section* to describe the information used for hospital management (e.g., cure rate, death rate), whereas this section is absent in the sections of the CCD, the PCC, and the HITSP. Furthermore, entries also differ as a result of the differing sections. In conclusion, these differences come from

business variations among different cultural and language backgrounds.

## Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## Acknowledgements

This work was supported by the Research Grant (Grant No. 81102202; 81171427) from National Natural Science Foundation of China, by the Research Grant (Grant No. 2009JM4028) from Science Foundation of Shaanxi Province, and by the National Science and Technology Infrastructure Program from the Ministry of Science and Technology of China (Grant No. 2008BAI52B01).

## References

1. Information Steering Committee Office, The Ministry of Health. The white paper on China's hospital information systems [Internet]. Beijing, China: The Ministry of Health; 2008 [cited at 2011 Nov 15]. Available from: [http://www.chima.org.cn/pe/DataCenter/UploadFiles\\_8400/200812/20081219115545203.pdf](http://www.chima.org.cn/pe/DataCenter/UploadFiles_8400/200812/20081219115545203.pdf).



2. Liu D, Wang X, Pan F, Yang P, Xu Y, Tang X, Hu J, Rao K. Harmonization of health data at national level: a pilot study in China. *Int J Med Inform* 2010; 79: 450-458.
3. Liu D, Wang X, Pan F, Xu Y, Yang P, Rao K. Web-based infectious disease reporting using XML forms. *Int J Med Inform* 2008; 77: 630-640.
4. Ministry of Health of the People's Republic of China. The state council approved the final draft of the long-awaited healthcare reform [Internet]. Beijing, China: Ministry of Health of the People's Republic of China; c1999-2006 [cited at 2011 Dec 25]. Available from: <http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohbgt/s3582/200901/38889.htm>.
5. Ministry of Health of the People's Republic of China. Basic architecture and data standards of electronic health records in hospitals [Internet]. Beijing, China: Ministry of Health of the People's Republic of China; c1999-2006 [cited at 2011 Dec 25]. Available from: <http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohbgt/s6694/200908/42155.htm>.
6. HL 7 Wiki. Continuity of care document (CCD) [Internet]. HL 7 Wiki; 2011 [cited at 2011 Dec 23]. Available from: [http://wiki.hl7.org/index.php?title=Continuity\\_of\\_Care\\_Document\\_\(CCD\)](http://wiki.hl7.org/index.php?title=Continuity_of_Care_Document_(CCD)).
7. Health Level Seven (HL 7) International. HL 7 [Internet]. Ann Arbor, MI: HL 7 International; c2007-2011 [cited at 2011 Dec 22]. Available from: <http://www.hl7.org/>.
8. IHE Patient Care Coordination (PCC). Technical framework [Internet]. IHE International; c2011 [cited at 2011 Dec 20]. Available from: [http://www.ihe.net/technical\\_framework/index.cfm#pcc](http://www.ihe.net/technical_framework/index.cfm#pcc).
9. American National Standard Institute (ANSI). HITSP/C83: HITSP CDA content modules component [Internet]. New York, NY: ANSI; c2009 [cited at 2011 Dec 19]. Available from: [http://www.hitsp.org/ConstructSet\\_Details.aspx?&PrefixAlpha=4&PrefixNumeric=83](http://www.hitsp.org/ConstructSet_Details.aspx?&PrefixAlpha=4&PrefixNumeric=83).
10. IHE International. IHE: changing the way healthcare [Internet]. IHE International; c2011 [cited at 2011 Dec 19]. Available from: <http://www.ihe.net/>.
11. American National Standard Institute. HITSP: enabling healthcare interoperability [Internet]. New York, NY: American National Standard Institute; 2009 [cited at 2011 Dec 20]. Available from: <http://www.hitsp.org>.
12. Health Level Seven (HL7) International. Clinical document architecture [Internet]. Ann Arbor, MI: HL7 International; c2007-2011 [cited at 2011 Dec 25]. Available from: <http://www.hl7.org/implement/standards/cda.cfm>.
13. Dolin RH, Alschuler L, Boyer S, Beebe C, Behlen FM, Biron PV, Shabo Shvo A. HL7 clinical document architecture, release 2. *J Am Med Inform Assoc* 2006; 13: 30-39.
14. China national health data dictionary and metadata management system [Internet]. [cited at 2011 Dec 25]. Available from: <http://www.chiss.org.cn/>.
15. Tu H, Yu Y, Yang P, Tang X, Hu J, Rao K, Pan F, Xu Y, Liu D. Building clinical data groups for electronic medical record in China. *J Med Syst* 2010 Jul 14 [Epub]. <http://dx.doi.org/10.1007/s10916-010-9540-x>.
16. Chen W, Pan X. Diagnostics, version 7. Beijing: People Health Press; 2008.
17. Ministry of Health of the People's Republic of China. Technology solution of establishing hospital information platform for electronic health record in China [Internet]. Beijing, China: Ministry of Health of the People's Republic of China [cited at 2011 Dec 10]. Available from: <http://www.moh.gov.cn/publicfiles/business/htmlfiles/mohbgt/s6694/201103/51091.htm>.
18. McDonald C, Huff S, Mercer K, Hernandez, JA, Vreeman DJ. Logical observation identifiers names and codes (LOINC<sup>®</sup>) users' guide [Internet]. Indianapolis, IN: LOINC; 2011 [cited at 2011 Nov 17]. Available from: <http://loinc.org/downloads/files/LOINCManual.pdf>.
19. Australian Institute of Health and Welfare. Metadata online registry [Internet]. Australian Institute of Health and Welfare; [cited at 2011 Nov 16]. Available from: <http://meteor.aihw.gov.au/content/index.phtml/itemId/181162>.
20. International Organization for Standardization. ISO/IEC international standard, information technology-metadata registries (MDR). Part 3, registry meta model and basic attributes. Geneva: International Organization for Standardization; 2011.
21. Johnson SB, Bakken S, Dine D, Hyun S, Mendonca E, Morrison F, Bright T, Van Vleck T, Wrenn J, Stetson P. An electronic health record based on structured narrative. *J Am Med Inform Assoc* 2008; 15: 54-64.
22. Jian WS, Hsu CY, Hao TH, Wen HC, Hsu MH, Lee YL, Li YC, Chang P. Building a portable data and information interoperability infrastructure-framework for a standard Taiwan electronic medical record template. *Comput Methods Programs Biomed* 2007; 88: 102-111.
23. Sleszynski SL, Glonek T, Kuchera WA. Standardized medical record: a new outpatient osteopathic SOAP note form: validation of a standardized office form against physician's progress notes. *J Am Osteopath Assoc* 1999; 99: 516-529.
24. Aghili H, Mushlin RA, Williams RM, Rose JS. Progress notes model. *Proc AMIA Annu Fall Symp* 1997: 12-16.