Development of an Open Metadata Schema for Clinical Pathway (openCP) in China

Wei Xu1; Yanxin Zhu1; Yibing Geng2
1Department of Medical Informatics, Second Military Medical University, Shanghai, China; 2Shanghai Banpo Network Technologies LTD., Shanghai, China

Keywords
Clinical pathway, metadata, archetype

Summary
Background: By the end of the first quarter of 2017, 1212 standard Clinical Pathways (CPs) had been developed and implemented in China since 2009, however, Chinese standard CPs are still paper-based CPs.
Objectives: To optimize the data structure and content of Chinese standard CPs based on Chinese health standards by two-level modeling approach and archetype method, and develop an Open Metadata Schema for Clinical Pathway (openCP) in China.
Methods: We chose 43 diseases covering 7 departments of internal medicine as samples and source materials including Respiratory, Gastroenterology, Neurology, Cardiology, Nephrology, Endocrinology, and Hematology, and we followed 4 steps to develop openCP: identify reference model of openCP; develop archetype models; standardize archetype terms and metadata; and evaluate its functionality.
Results: We have built hierarchical reference model and archetypes of openCP, created and standardized archetype terms and metadata, and evaluate the functionality of openCP. The qualitative comments from 5 doctors were inspiring.
Discussion and Conclusions: Medical staff may easily build localized CPs in openCP which can be integrated into EHR systems and shared among different hospitals and clinics.

Correspondence to:
Wei Xu
Department of Medical Informatics
Second Military Medical University
800, Xiangyan Road, Yangpu District
200433 Shanghai, China
E-mail: zhuizhui@smmu.edu.cn; xuzhuizhui@hotmail.com

1. Introduction
Clinical pathway (CP) is an important information entity in evidence-based medicine information system architecture, which can provide high quality healthcare through well-organized processes based on best clinical evidence. CP may have many other terms such as "care map", "care plan", "care pathway", "critical pathway", and "care path", and they are all used to describe "standardized, evidence-based multidisciplinary management plans, which identify an appropriate sequence of clinical interventions, timeframes, milestones and expected outcomes for a homogenous patient group" [1]. A well-designed CP is an important management tool to provide high quality and cost-efficient health care [2].

By the end of the first quarter of 2017, 1212 standard CPs in 30 specialties had been developed in China and implemented in more than 7000 hospitals since a trial version of the Guiding Principle of CP was issued in October 2009 [3]. CP standard procedures and CP checklists are the main documents of Chinese standard CPs. A CP standard procedure is a text describing care workflows, while a CP checklist (the following CP means CP checklist) is a two-dimensional table with both care topics (vertical) and timeframe (horizontal) options. The care topics are categorized into Main Care, Main Orders, Main Nursing, and Variance [4], and fine-grained items of specific clinical activities or process are listed under each care topic. By implementing standard CPs for more than 8 years, many hospitals in China have improved health care management, standard medication, health resource integration and patient satisfaction.

A CP needs to be both standardized and localized. On the one hand, a standard and structured CP is easy for computer operation and makes healthcare consistent. On the other hand, a localized CP makes clinical staff possible to follow. Furthermore, a standard and structured CP facilitates creation, revision, use and interoperability of other systems.

However, Chinese standard CPs are still paper-based CPs. In reviewing and analyzing 1212 standard CPs, we found that although standard CPs can basically meet the requirements of most hospitals, they have following deficiencies. (1) Standard CPs are not well-structured. They have a two-level structure and discreet items to represent clinical topics and activities, however, the categories of care topics are ambiguous and can be misleading. For example, Main Care in care topics includes most of care processes or steps. There may be quite a few specific items which make medical staffs not easy to locate key information; (2) standard CPs are not well-standardized. Fine-grained items under each care topic are expressed in text, and...
words and expressions are inconsistent when representing the same meaning among different diseases; and (3) there are no standard variance codes which can facilitate statistic analysis [4]. These deficiencies hinder the efficient development and use of standard CPs for different diseases and make it difficult to adapt to computer-based or even web-based environment.

It is important to computerize CPs to increase the efficiency of paper-based CP and promote the secondary use of data, and schema model of CPs is the basis for implementing electronic CPs [5]. We intended to develop an Open Metadata Schema for Clinical Pathway (openCP) in China based on Chinese health information standards by using two-level modeling approach and archetype method so that electronic CPs can be easily built for displaying, recording, ordering, editing and calculating.

2. Methods
2.1 Background

We chose 43 diseases covering 7 departments of internal medicine as samples and source materials including Respiratory, Gastroenterology, Neurology, Cardiology, Nephrology, Endocrinology, and Hematology. All these CPs hold discreet items similar to data items in electronic health records (EHRs) to some extent. One of the differences is that CP items focus more on clinical processes and actions, while EHR items focus more on clinical recording.

One of the key functions of electronic CPs is that it can be integrated with EHR and avoid redundant data entries [6, 7]. Therefore, we designed openCP based on structure and content of Chinese EHR Standard. Chinese EHR Standard is a trial version standard for EHR information model and datasets, and is composed of Data Structure, Clinical Information Model, Data Group and Data Element Standard, and Basic Template and Data Set Standard [8]. Data Structure includes four hierarchical levels of document, section, data group, and data element from the top-level container to the bottom-level element. Data Group and Data Element Standard has 76 nested two-level data groups, 465 data elements and 76 value domain code sets.

We referenced data structure and content of Chinese EHR Standard by using the two-level modeling approach and archetype method to develop openCP for the following two reasons.

Two-level modeling approach is composed of reference model as lower level and archetypes as upper level model, and it is successfully used in openEHR architecture [9]. Archetypes, analogous to building blocks, can be reusable, composable and hierarchically organized, and they are responsible for essential semantic constraints at the knowledge level [10, 11]. Originally introduced to denote domain concept model in EHR, they are now widely used in many other medical information fields such as supporting data recording in clinical trials and representing definition language of clinical practice guidelines [12, 13]. Since data structures and most of the data elements in clinical pathways are consistent with those for EHR and clinical research, we believe archetypes may also be efficient to create openCP. Furthermore, openCP designed by using archetypes is independent of any information model or system. Archetypes enable defining, adding, or revising of clinical content by medical experts without the need for technical understanding [14], and it can be easily integrated into different medical information models or systems.

Chinese EHR Standard is also built by two-level modeling with Data Structure and Clinical Information Model corresponding to the openEHR Reference Model while Data Group and Data Element Standard corresponding to openEHR Archetypes [8]. In our previous work, we successfully built an Open Metadata Schema for Prospective Clinical Research (openPCR) based on Chinese EHR Standard by using archetype method [15]. Development of openCP through the consistent way can facilitate compatibility and integration with EHR and clinical research models.

We followed 4 steps to develop openCP: identify reference model of openCP; develop archetype models; standardize archetype terms and metadata; and evaluate its functionality.

2.2 Step 1: Identify Reference Model of openCP

Chinese standard CPs include 2 parts: the data category of Basic Information and time-varying data categories of Main Care, Main Orders, Main Nursing, Variances, and Signature which can be repeated according to the timeline. Only Main Orders can be further divided into Long-term Orders, Short-term Orders and Discharge Orders. Under each category is listed discreet textual data items (Figure 1).

There is an ontological gap between data categories and data items. The data categories such as Main Care, Main Orders, or Main Nursing are too broad concepts to be searched; the discreet data items under these data categories are difficult to be browsed and located. Data categories of Basic Information, Main Care, Main Orders and Main Nursing in Chinese standard CPs can be further divided into more subordinate ones which can well organize data items.

We developed the reference model of openCP according to Data Structure of Chinese EHR Standard with the hierarchical level of document, section, data group, and data element. Specific archetypes can be created based on the reference model.

2.3 Step 2: Develop Archetype Models of openCP

Data items in Chinese standard CPs are variously expressed in textual words or sentences. Some of them with the same meaning may be expressed differently in different diseases, or even within one disease. For example, a data item expressed as "observe drug adverse reaction" was found in clinical pathway of community-acquired pneumonia in Chinese standard CPs, while a data item of the same meaning expressed as "observe drug reaction and side effects" in clinical pathway of chronic obstructive pulmonary disease. Some of data items may present more than one concepts within one data item. For example, "determine therapy strategy and perform the empirical anti-infection treatment". We extracted all the data items from 43 diseases of internal medicine, and dissected data items with more than one concepts into discreet items.
Data items of the same meaning were considered identical.

We then reorganized the structure of Chinese standard CPs, added middle levels between data categories and data items and built the hierarchical archetype models of openCP. Since Chinese EHR Standard has Data Group and Data Element Standard as its archetypes, we created openCP archetypes against Data Group and Data Element Standard. That is if there are existing archetypes (i.e., data groups or data elements) in Chinese EHR Standard, we adopted it directly; if not, we created new ones. We also searched the openEHR Clinical Knowledge Manager (http://www.openehr.org/ckm/) to check the corresponding archetypes. If a suitable archetype was available, we mapped it to the corresponding openCP archetype.

2.4 Step 3: Standardize Archetype Terms and Metadata

Terms of archetypes including data groups and data elements we newly created were standardized according to Chinese EHR Standard, NHDD (National Health Data Dictionary of China) and WS/T 303-2009 (Rules for Data Element Standardization for Health Information).

NHDD is National Health Data Dictionary of China (also called Meta Data Repository), and it provides formal healthcare data elements and their metadata and promote data sharing and exchanging [16]. It holds 278 data elements and 85 value codes. Each formal data element has 8 attributes including Chinese Name, Identifier, English Name, Definition, Object Class, Property, Presentation, and Format. For example, Data Element “Food Composition” has attributes of Chinese Name “膳食结构”, Identifier “PAT010104_100011”, English Name “Food Composition”, etc.

WS/T 303-2009 refers to Information Technology – Metadata Registries (MDR) (ISO/IEC 11179–3:2004). It specifies not only naming rules for data elements, but also basic metadata attributes describing data elements [17]. A name of a data element can be given according to one object class term, one property term, one representation term, and (or) qualifier terms by semantic and syntactical rules. For example, the name of a data element expressed as “diet component code” was specified by class term “diet”, property term “component”, and presentation term “code”.

There are two kinds of data items we extracted and dissected in Chinese standard CPs. One is generic that can be shared among Chinese standard CPs of different diseases and different departments, and the other is dedicated to CPs of specific diseases. As to the generic data items, we adopted the names of data elements from Chinese EHR Standard or NHDD to represent data items of the same meaning in Chinese standard CPs. If there were no corresponding data elements, we created new names for the data items according to WS/T 303-2009. These standardized data items became core data elements. As to the specific data items, we divided them into 2 categories and standardized them respectively. One can be transformed into core data elements plus specific details. The other was standardized as department-specific or disease-specific data elements.

Metadata attributes of archetypes were grouped into six categories: Identification, Definition, Relation, Representation, Administration, and Usage guides [15]. In relation, we created Mapping attribute to represent connections of archetype terms with those in other standards.

2.5 Step 4: Evaluate the Functionality of openCP

Following the idea of openEHR Archetype Editor, we developed a test openCP platform and created a couple of clinical pathways to test the usability and functionality of openCP. We evaluated openCP by adopting standard functions of electronic CPs [6]. Wakamiya and Yamauchi classified the standard functions into 6 categories and 17 functions. OpenCP is a metadata scheme focusing on data structure and content, therefore we excluded those functions that depend on system designs (1 category and 10 functions). We created 5 evaluation factors based on the included 7 functions (Table 1).

Factor 3 and 4 can be tested by doctors through the openCP platform. Based on the two evaluation factors, we created 4 questions including: Does openCP have clear and concise structure? Does openCP have clear and concise content? Is openCP...
archetype easy to create or revise? And is data items in a CP easy to add or delete?

We reviewed electronic CP system in a tertiary hospital in Shanghai, and invited 5 doctors, who used electronic CPs in that hospital every day and were familiar with them, to test a CP created by openCP, answer the 4 questions, and talk about their feelings towards the usability.

3. Results

We have built hierarchical reference model and archetypes of openCP based on Chinese EHR Standard, created and standardized archetype terms and metadata, and evaluate the functionality of openCP. The qualitative comments from 5 doctors were inspiring.

3.1 Results of Step 1

We developed reference model of openCP based on Data Structure of Chinese EHR Standard, and we also keep data categories of Chinese standard CPs as sections (Table 2).

Corresponding to data structure of Chinese EHR Standard, openCP reference model has four hierarchical levels including Document, Section, Data Group, and Data Element.

Table 2 Reference model of openCP based on Chinese EHR Standard and Chinese standard CP.

<table>
<thead>
<tr>
<th>Document</th>
<th>Section</th>
<th>Data group</th>
<th>Data element</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>Data category</td>
<td>Data group</td>
<td>Data item</td>
</tr>
</tbody>
</table>

3.2 Results of Step 2

3.2.1 Top-level Archetype Model of openCP

At the top level, the Document contains the root object, Clinical Pathway, and the Sections include data categories of Basic Information, Main Care, Main Orders, and Main Nursing in original Chinese standard CPs (Figure 2).

3.2.2 Archetype Model of Basic Information Section

Basic Information Section presents the state or situation of a patient, which includes demographic information, patient ID, diagnoses, implementation time, etc. Data groups of Service object identifier, Demographics, Event summary, Diagnoses and Procedures in Chinese EHR Standard contain the corresponding identification information [18] and can be reused in Basic Information Section (Figure 3). We got 5 data groups and 12 core data elements in Basic Information Section.

3.2.3 Archetype Model of Main Care Section

Main Care Section contains a series of care plans, observations, actions, and assessments. Most of data items have corresponding ones in Chinese EHR Standard. We adopted directly 9 Data Groups in Chinese EHR Standard (Figure 4) including Physical exam, Specific exam, Lab data, Procedures, Medications, Care/Treatment plans, Assessments, Encounters/Episodes notes, and Health guidance. We combined Data Groups of Physical exam, Specific exam, and Lab data into one Data Group: Diagnosis tests, and combined Data Groups of Procedures and Medications into Interventions. We also created 4 new data groups: Discharge plans, Ward rounds, Consultations and Transfers, and we got 10 data groups in Main Care Section.

We further divided Data Groups of Diagnosis tests, Interventions and Care/Treatment plans into sub-data groups, for those data elements have great difference (Figure 5). We developed 93 core data elements that can be shared among different CPs.
3.2.4 Archetype Model of Main Orders Section

A clinical order is an action-oriented instruction by an authorized practitioner describing an intervention in the health of a specific patient [19]. Since orders formulate medical instructions, while activities implement instructions, they share many common data items. Therefore, Section of Main Orders can reuse data groups and elements in Section of Main Care. In Main Orders Section, we kept Long-term Orders, Short-term Orders and Discharge Orders as sub-sections. In addition to reused data groups of Diagnosis tests, Interventions and Health guidance in Main Care Section, we added data groups of Nursing, Diet and Patient activities (▶Figure 6).

We developed 22 generic data elements in data groups of Diet, Nursing and Patient activities, such as data elements of High-quality-protein diet, Low-phosphorus diet, Low-purine diet, etc.

3.2.5 Archetype Model of Main Nursing Section

Main Nursing Section involves care or nursing activity conducted by nurses. Nurses are involved in clinical care and implement clinical orders of doctors, so Main Nursing Section reuses most of the data groups and data elements in Main Care Section and Main Orders Section such as data groups of Interventions, Health guidance, Nursing, Diet, Patient activities, etc. Although different sections may have subordinate data groups and data elements, these data groups or data elements do not strictly belong to a specific superordinate section. All the data groups and data elements can be reused by any section if necessary.

Data group of Variance is an independent data group for CP in openCP which has 2 data elements including Variance reason and Variance code.

3.3 Results of Step 3

We created standardized archetype terms of data groups and data elements according to Chinese EHR Standard, NHDD and WS/T 303-2009. As to specific data items

<table>
<thead>
<tr>
<th>Data group ID</th>
<th>Data group name</th>
<th>Data group name</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.02</td>
<td>Physical exam</td>
<td>Diagnosis tests</td>
</tr>
<tr>
<td>S.05</td>
<td>Specific exam</td>
<td>Interventions</td>
</tr>
<tr>
<td>S.06</td>
<td>Lab data</td>
<td>Care/Treatment plans</td>
</tr>
<tr>
<td>S.08</td>
<td>Procedures</td>
<td>Assessments</td>
</tr>
<tr>
<td>S.09</td>
<td>Medications</td>
<td>Encounters/Episodes notes</td>
</tr>
<tr>
<td>S.10</td>
<td>Care/Treatment plans</td>
<td>Diagnosis tests</td>
</tr>
<tr>
<td>S.11</td>
<td>Assessments</td>
<td>Interventions</td>
</tr>
<tr>
<td>S.12</td>
<td>Encounters/Episodes notes</td>
<td>Interventions</td>
</tr>
<tr>
<td>S.15</td>
<td>Health guidance</td>
<td>Health guidance</td>
</tr>
</tbody>
</table>

Figure 3 High-level archetype of Basic Information Section.

<table>
<thead>
<tr>
<th>Data group ID</th>
<th>Data group name</th>
<th>Data group name</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.02</td>
<td>Service object identifier</td>
<td>Service object identifier</td>
</tr>
<tr>
<td>H.03</td>
<td>Demographics</td>
<td>Demographics</td>
</tr>
<tr>
<td>H.10</td>
<td>Event summary</td>
<td>Event summary</td>
</tr>
<tr>
<td>S.07</td>
<td>Diagnoses</td>
<td>Diagnoses</td>
</tr>
<tr>
<td>S.08</td>
<td>Procedures</td>
<td>Procedures</td>
</tr>
</tbody>
</table>

Figure 4 High-level archetype of Main Care Section.
in Chinese standard CPs, we transformed them into core data elements with specific details, or with linkage to internal or external standards or rules. For example, in Chinese standard CP of Sustained Ventricular Tachycardia, Data Item of “Transfer to implantable cardioverter defibrillator (ICD) surgery process, if patient has indication of ICD implantation after surgery of EPS and RFCA” can be transformed to core data element of Postoperative disposition plus specific details of “Transfer to (ICD) surgery process, if patient has indication of ICD implantation after surgery of EPS and RFCA” (Figure 7).

Since another category of specific data items in Chinese standard CPs is unsuitable to transform into core data elements with details, we transformed them into specific data elements by using Chinese EHR Standard or NHDD, or created new data elements according to WS/T 303-2009. For example, in Chinese standard CP of Gastroduodenal Ulcer, Data Item of “Observe patients with abdominal symptoms and signs, stool characters, and any complications” can be transformed into specific data elements of Abdomen, Stools, and Complications under sub-data group of Observation (Figure 8).

Each archetype term was described by six metadata categories of Identification, Definition, Relation, Representation, Administration, and Usage guides. In Relation, metadata attribute of Mapping can be created to link to other standards, such as openEHR archetypes or Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). For example, Data Element “Adverse reaction” in Sub-Data Group “Observation” under Data Group “Diagnosis tests” can be mapped to Entry “Adverse reaction risk” under Entry “Evaluation” in openEHR Archetypes, or be mapped to Concept “Adverse reaction” (ConceptID:281647001) in SNOMED CT (Figure 9).

### 3.4 Results of Step 4

According to a hospital information member, all the tertiary hospitals in Shanghai have implemented CP system since 2012, and the usage rates of CPs were evaluated every year. We visited a tertiary hospital and found that its electronic CPs were created by computerizing Chinese standard CPs. Electronic CPs have been used mainly by attendings and residents.

We then invited 3 attendings and 2 residents who used electronic CPs every day to try our test openCP platform, and talk about their feelings. The answers to the 4 questions were positive (Table 3).

The feelings towards using the test openCP platform were encouraging. All the interviewers felt the CP created by the platform is concise and clear, easy to browse and locate, and convenient to use, and 3 of them said it will definitely save time and promote effectiveness.

We evaluated openCP using the 5 evaluation factors. As to Factor 1 and 2, the design mechanism guarantees the integration of openCP with those EHR system adopting Chinese EHR Standard as well as other EHR and CPOE systems. OpenCP was developed based on data structure and content of Chinese EHR Standard.
thermore, the metadata attribute of Mapping connected archetypes to the correspondences in other health information systems or standards.

As to Factor 3, we created openCP with 4-level data structure and all the data items were standardized according to Chinese EHR Standard or NHDD. One interviewer compared the openCP platform with electronic CPs he used “Many data items in electronic CPs in our hospital are long sentences, and it is difficult to read and create. Different doctors often create the same meaning with different expressions. Data items in the openCP platform are concise and clear, and the detailed information familiar to doctors is hidden. I am familiar with and easy to follow or locate these data items”. Another interviewer expressed her views of the platform “The structure is following the data structure of EHR, and I am familiar with it”.

Factor 4 has 2 meanings: easy creation or revision of archetype terms, and easy addition or deletion of data items in a CP. Most of the archetypes in openCP will be created by professionals, however, medical experts also have the possibility to create the archetypes to meet their own needs. But one interviewer felt unsure whether it is convenient to create an archetype term with many metadata attributes. All the interviewers felt it is easy to add or delete data items in a CP. One interviewer said: “I can easily browse the hierarchical trees, enter into the data group, find and add the data element I need”.

Regarding to Factor 5, all the CPs should have variance to record the deviation from the pathway [20]. We created data elements of Variance reason and Variance code, with Variance reason recording change, deviation, or dropout of the care and Variance code calculating statistics for evaluation of the implementation of CPs.

4. Discussion and Conclusions

More than 1212 Chinese standard CPs have been issued successively since 2009, which improved quality of health care and narrowed its difference among different areas, and promoted the satisfaction of patients during implementation. However, Chinese standard CPs are still paper-based
CPs, and are not well-structured and standardized.

To our knowledge, although there are extensive researches to computerize CPs [21, 22, 23], no study has focused on modeling or standardizing paper-based Chinese standard CPs. Consistent with data structure of Chinese EHR Standard, we have built archetype models of openCP and standardized archetype terms and metadata, and evaluate the functionality of openCP. Basic Information Section included 5 data groups holding common information which does not change with time. 14 data groups were created and reused in Main Care Section, Main Orders Section and Main Nursing Section.

It is not easy to standardize CPs for that although they can be developed based on national standard level and shared by different tertiary hospitals and clinics, they should also be localized from much coarse-grained to fine-grained models to meet various clinical needs. We found that two-level modeling and archetypes were efficient ways to develop CP models, for they provide such mechanism to facilitate the development. Archetypes of single unit can be created respectively and combined to develop more complex units. OpenCP has generic and relatively small standard information modules, and different hospitals or departments can choose archetypes of data groups or data elements from generic information modules and add customized specific data groups or data elements.

Electronic CP is a major component of clinical information system, and it can always be integrated into EHRs [6, 24, 25] for various reasons. Every phase of CPs has close connections with specific elements in EHR [6]. CPs from practice data in the EHR may influence practice review and shared decision making [26]. CPs and EHRs might be intelligently combined for clinical research purposes [27], and CP variance can be identified by comparing EHR against CP [28]. Since CPs are management plans for clinical care, they are following clinical workflows and in accordance with clinical records to a great extent. The consistency of data structures between CPs and EHRs can facilitate the later integration of CPs into EHRs. By adopting Data Structure and most Data Groups and Data Elements of Chinese EHR Standard, CPs created by openCP can be mapped to and integrated into Chinese EHR Standard based EHR system model, and furthermore it can be mapped to our previously developed openPCR platform, which makes secondary use of clinical data possible.

We have not only develop openCP based on Chinese health related standards, but also considered interoperability with other international standards. OpenCP archetype terms could be linked to openEHR archetypes and SNOMED CT concepts by archetype metadata attribute of Mapping. SNOMED CT is a comprehensive clinical health terminology as a common global language for health terms [29]. Chinese version of SNOMED International was translated more than 20 years ago [30]. Although its electronic version is seldom used, we can create the mapping for future needs.

The functionality of openCP was evaluated and the results were inspiring. All the interviewers felt the CP created by the test openCP platform clear and concise, and convenient to browse and locate information.

The research has some limitations. This time we created openCP by using Chinese standard CPs of internal medicine, and next time we will add generic data groups and data elements for other departments and complete the openCP step by step.

We have only developed a test platform for creating a couple of CPs, the test platform is far from complete, and has not been used in real clinical environment. Although we invited doctors to test the platform and got the positive results, the interview sample size is small. In the future, we will cooperate with hospitals, test the efficiency and efficacy of CPs made by openCP, and improve its usability and functionality.

We developed openCP platform and generic CP information model based on Chinese health standards, and structured and standardized Chinese standard CPs. Medical staff may easily build localized CPs in openCP which can be integrated into EHR systems and shared among different hospitals and clinics.

References