Building Evidence Graph for Clinical Decision Support

Jing Mei, Wen Sun, Jing Li, Haifeng Liu, Xiang Li, Yiqin Yu, Guotong Xie

IBM Research, Beijing, China
{meijing, sunwenbj, jingli, liuhf, lixiang, yuyiqin, xieguot}@cn.ibm.com

Abstract. Evidence-based medicine intends to optimize clinical decision making by using evidence. Semantic query answering could help to find the most relevant evidence. However, at point of care, it still lacks time for clinicians to do literature study. In this poster, we propose to build an evidence graph for clinical decision support, in which an evidence ontology is defined with extension of SWRL rules. On top of this graph, we do evidence query and evidence fusion to generate the ranking list of decision options. Our prototype implementation of the evidence graph demonstrates its assistance to decision making, by combining a variety of knowledge-driven and data-driven decision services.

Keywords. Evidence-based Medicine, Clinical Decision Support, Knowledge Graph

1 Introduction

Providing clinical decision support needs evidence [1]. The evidence could be a risk model from cohort studies (such as the Framingham cardiovascular disease 10-year risk model), or a rule model from clinical guidelines (such as the NICE guideline for management of type 2 diabetes), or a similarity analysis from electronic health records (such as the distribution of stroke treatment in similar patients), etc. However, at point of care, given a patient, what are the most relevant evidence to provide a personalized clinical decision support for him/her? To address this problem, there are two big challenges.

First is how to find the most relevant evidence. A general solution is a search engine, which requires (1) the evidence representation, (2) the evidence extraction, (3) the evidence indexing. Considering that evidence has semantics (e.g. disease hierarchy, drug interaction), the basic keyword-based search is not enough, and we propose to build an evidence graph for answering semantic queries.

Second is how to use the found evidence to provide a personalized clinical decision support. The state of the art solution is to rank the evidence search results for human reading, however, at point of care, clinicians rarely have time to do literature study. For instance, at point of care, given a patient with non-rheumatic atrial fibrillation, the clinician wants to know the patient's risk score of stroke, instead of reading papers – even if the most relevant papers such as the CHA2DS2-VASc study have been retrieved. Therefore, our evidence graph aims to making evidence machine understandable and processable, for giving the ranked decision options (instead of the ranked documents).

2 Evidence graph

Table 1 shows an ontology definition of the evidence graph. Referring to the PICO (Population, Intervention, Comparison, Outcome) framework which has been well established to formulate clinical questions [2], we defined the P,I,C,O four classes, in addition to another four classes of Evidence, Feature, Source, Terminology.

Table 1. An Evidence Ontology

Class	Description
Source	This class represents evidence sources from external files, web pages or data sets (e.g. the NICE guideline pdf file, the Framingham Heart Study web page, the CHARLS data repository).
Terminology	This class represents code systems (e.g. disease codes in ICD 10, lab test codes in LOINC).
Population	This class represents the population studied in the evidence (e.g. Chinese retired residents, or overweight diabetes patients).
Intervention	This class represents the intervention studied in the evidence (e.g. insulin therapy, or life style intervention).
Comparison	This class represents the comparison studied in the evidence (e.g. placebo, or an alternative intervention)
Outcome	This class represents the outcome studied in the evidence (e.g. death, or hospitalization)
Feature	This class represents features defined by logical and/or arithmetic expressions of terminologies (e.g. the feature of elder is defined as: age>=60, the feature of male is defined as: gender=male).

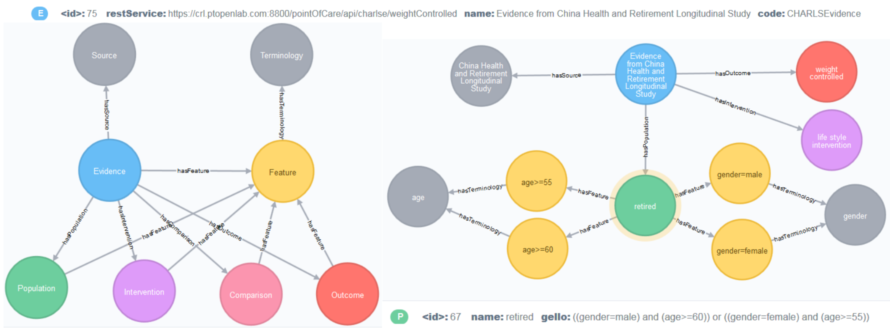

Fig. 1. Left hand side: classes and properties in the evidence ontology.
Right hand side: an evidence instance from the CHARLS data source.

In Figure 1, the left hand side is the 11 properties of 8 classes, such as hasSource (from Evidence to Source), hasPICO (from Evidence to P,I,C,O), hasFeature (from P,I,C,O to Feature), hasTerminology (from Feature to Terminology). The right hand side illustrates an evidence instance, which has data source from the Chinese Health and Retirement Longitudinal Study (CHARLS[1]). It has a population instance, namely retried, with 4 feature instances, i.e., gender=male, gender=female, age>=55, age>=60, followed by two terminology instances, namely gender and age. The retried population has an expression: ((gender=male) and (age>=60)) or ((gender=female) and (age>=55)). Besides, this evidence instance has attributes of id, name, code and restService, etc., where the restService is the URI of its decision service implementation.

Here, we remark two key points in the evidence graph. One is the evidence semantic query answering, and the other is the evidence fusion for decision option ranking. Actually, the semantic part is nature for our evidence graph, which is defined by an ontology with extension of SWRL rules. Below is a SWRL sample for evidence inference.

Evidence(?e1) & hasPopulation(?e1, ?p1) & hasFeature(?p1, ?f1) & hasTerminology(?f1, ?t1) & Evidence(?e2) & hasPopulation(?e2, ?p2) & hasFeature(?p2, ?f2) & hasTerminology(?f2, ?t2) & subTermOf(?t1, ?t2) => subEvidenceOf(?e1, ?e2)

From the ICD 10 code system, we have *subTermOf(Stroke, CardiovascularDisease)*, which implies that, given a patient with diagnosis of non-rheumatic atrial fibrillation (AF), the risk models of stroke and other cardiovascular diseases (CVD) could be all applicable for him/her.

The fusion part is novel, via calling the restService URI of each retrieved evidence. Taking the above AF patient as an example, both the CHA2DS2-VASc evidence (for stroke risk model) and the Framingham evidence (for CVD risk model) have been retrieved. The CHA2DS2-VASc evidence service takes input features of hypertension, diabetes, congestive heart failure, age >= 75, pre stroke or transient ischemic attach or thromboembolism, and outputs the stroke risk score from 0 to 6. The Framingham evidence service takes input features of diabetes, smoking, age, systolic blood pressure, total cholesterol, HDL cholesterol, and outputs the CVD risk score from 0 to 1. These different evidence services will be activated (most possibly by input of different features), and generate scores of different decision options. Finally, we will do a decision fusion of evidence services, which could leverage fusion algorithms such as majority voting, weighted average, and meta-classification [3]. In this example, the stroke risk socre is 3, the CVD risk score is 0.75, and after decision fusion, the stroke evidence would be diffused to the CVD evidence, getting the final CVD risk score of 0.86.

3 Conclusion and Outlook

An evidence graph for clinical decision support is appealing and challenging. This poster presents an early phase, and we have manually built an evidence graph, including 4 sources from Framingham, CHA2DS2-VASc, NICE, and CHARLS. Based on that, we implemented the Decision Advisor with integration of a care management system (namely Curam[2]), and Figure 2 is the screenshot of a patient's care plan. At point of care, suppose that a patient Cao Ping comes, and 3 objectives (a.k.a. outcomes) are selected. By semantic query answering on the evidence graph, we retrieve out the evi-

---

1 http://charls.ccer.edu.cn/en
2 http://www-01.ibm.com/software/info/curam/

dences applicable for Cao Ping. After posting the patient data to each evidence restService URI, 5 life-style intervention options are respectively scored by different evidence services, and the decision fusion finally gives the ranked list. Social activity has been first recommended to Cao Ping, with the highest rank score.

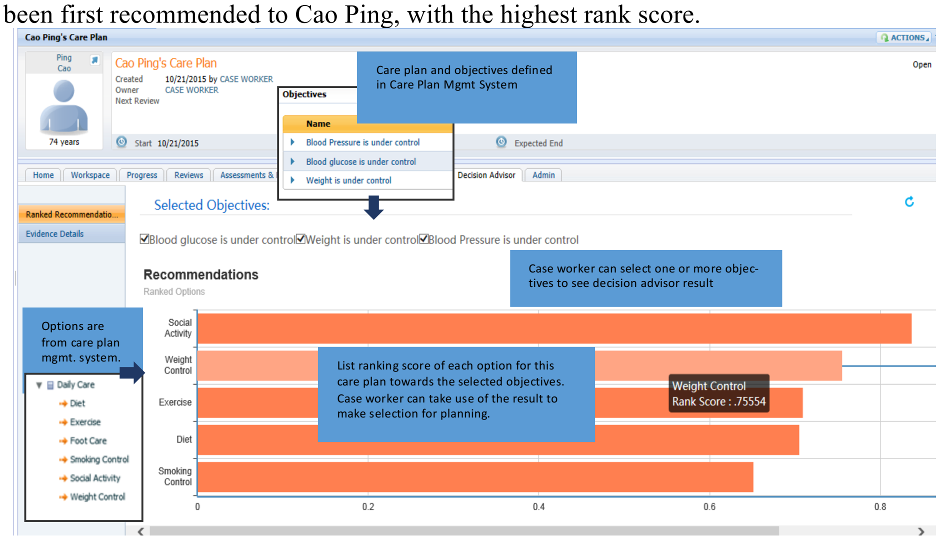

Fig. 2. Decision Advisor based on Evidence Graph

In future work, we are planning to release the evidence graph of large scale. First of all, we will leverage Bio2RDF [4] which has published 11 billion triples across 35 bioinformatics datasets. In particular, PubMed could serve as the evidence source, from which we expect to extract PICO elements. Besides, MeSH (Medical Subject Headings) could contribute to the terminology class, and NDC (National Drug Code) could contribute to the intervention class. A harder issue is the service generation, e.g. how to extract the risk factors from the evidence, and implement the risk calculator with publishment of the restService URI. Last but not least, we realize that evidence itself has the quality level and confidence score, while text mining and data mining technologies further exacerbate the uncertainty. Decision fusion paves the way to combine various evidences, however it still needs more investigation, special for evidence conflicts.

Reference:

1. Craig JC, Irwig LM, Stockler MR. Evidence-based medicine: useful tools for decision making. Med J Aust. 2001;174(5):248-53.
2. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. AMIA Annual Symposium Proceedings. 2006:359-63.
3. Mei J, Liu H, Li X, Xie G, Yu Y. A Decision Fusion Framework for Treatment Recommendation Systems. Studies in Health Technology and Informatics. 2015;216:300-4.
4. Belleau F, Nolin MA, Tourigny N, Rigault P, Morissette J. Bio2RDF: towards a mashup to build bioinformatics knowledge systems. J Biomed Inform. 2008 Oct;41(5):706-16.