

# Patient Centered Event Representation for the Treatment of Multifactorial Diseases: Current Progress and Challenges

Yihan DENG<sup>a,1</sup>, Jan GAEBEL<sup>a</sup> and Kerstin DENECKE<sup>b</sup>

<sup>a</sup>*Innovation Center Computer Assisted Surgery, University of Leipzig*

<sup>b</sup>*Department of Medical Informatics, Bern University of Applied Sciences*

**Abstract.** Events such as clinical interventions, adverse drug events are one of the basic semantic units in the clinical workflow and are the foundation of the pathway representation. Current research has mainly concentrated on event recognition through concept mapping using from medical ontologies (UMLS, SNOMED CT) and gene relation detection in the biological context. However, the analysis of the patient status and the interaction between the patient status and a context event is still at the primary stage. In order to realize an efficient personalized treatment design and pathway planning, the correlation between a patient status and different types of clinical events should be analyzed. In this paper, we will provide a summary of the current research progress in clinical event detection in the biomedical domain and compare two approaches of event acquisition: an event schema produced using a guideline-based method and an expert-based annotation. We will apply the approaches to generate a structured annotation corpus and a special case of an event schema based on the complication classification and risk management in treatment of laryngeal cancer.

**Keywords.** Clinical event, patient status, clinical guideline, knowledge management

## 1. Introduction

Clinical narratives reflect the changes of the patient status in the entire diagnostic and therapeutic workflow. They record the patient status at different time points and also which types of the clinical events occurred, e.g. surgical interventions or pharmaceutical interventions. A better understanding of the interactions between the patient status and the corresponding clinical events could be realized by automatic event detection from clinical narratives. With the proceeding of the treatment process, the amount of the clinical narratives for one patient grows rapidly, especially for those complex disease involved with several clinical specialties. A formal representation of clinical events in the clinical domain is a crucial step for the follow up event detection and correlation analysis. Although a variety of event definitions can be found in the current research progress and shared tasks, none of them have provide a clear definition of an event referring to the patient status and its relations to clinical events. The deployment of event detection in the treatment planning has also not been discussed so

---

<sup>1</sup> Corresponding Author. Yihan.deng@medizin.uni-leipzig.de

far. In this paper, we will propose an integrated patient centered event schema based on our clinical studies and expert based annotation. A special case for the complication modeling in the laryngeal cancer treatment will be introduced. The distance between the current technology and practical usage will be discussed.

Clinical event extraction has made considerable progress over the last decade. Two subdomains have sparked the particular attentions of the research community. The first subdomain is related to the extraction of information on medication and adverse drug events. As the event in the medication domain is relatively standardized, the drug name together with its dosage, precondition and outcomes constitute the general form of the drug event. The adverse drug event has then normally the form of “object has a relation to an object”. The relation indicates the cause and effect between the drug and diseases or drug and the corresponding effects. The experimental system MedEx [1] has achieved a reliable performance in the medication extraction from discharge summaries with 93.2% accuracy.

The other subdomain is focusing on the extraction of the facts and relations and event detection in clinical narratives. Within the UMLS, concepts are assigned to semantic groups (anatomy, chemical and drugs, devices, disorders, geographic areas, living beings, objects, phenomena, physiology, procedures), which provide suitable foundations for the event detection [2, 3, 4]. From 2006 to 2014, plenty of evaluation and development challenges have been organized by the i2b2 shared task [5]. The challenges include not only the extraction of basic facts such as concept, assertion, but also the high level semantics such as relation recognition, temporal detection, smoking status recognition, obesity identification.

None of the existing event representations have considered the relation between the patient status and reflections on the one hand, and different types of biomedical events on the other hand. Further, the cooperative treatment for complex diseases requires also a reliable mutual understanding and determined agreement towards the patient status between caregivers from different specialties. In this paper, we will compare the event modeling for the treatment of multifactorial diseases (guideline and studies based) and the patient centered event modeling in a general domain (expert based). The current progress and the challenges in patient event modelling will be discussed.

## 2. Methods

The information on the patient status and his feedbacks are recorded in the electronic patient records, while general information on the treatment is available in relevant clinical guidelines and studies. In order to obtain the event representation based on these two data sources, two types of approaches are employed: 1) expert based annotation and 2) guideline and study-based event modeling (see figure 1).

1) The **expert based annotation** is a type of data driven approach: as a pre-stage for the follow-up automatic method, the patterns and principle of the event representation are recognized and obtained through manual annotation. The aim of the annotation in clinical narratives is the determination of relations between observed patient status and clinical treatments. During the annotation process, at least three clinical experts should work independently to conduct the annotation in clinical narratives. The patient records from controlled patient cohort with the same multiple diseases (same combination of ICD codes) and similar treatments (same combination

of procedure code) are used as annotation corpus. The recorded patient status and the corresponding events are annotated. After the annotation, the event and correlation between clinical event and patient status are determined based on annotation adjudication and expert based voting. The result with high inter annotator agreement will be accepted.

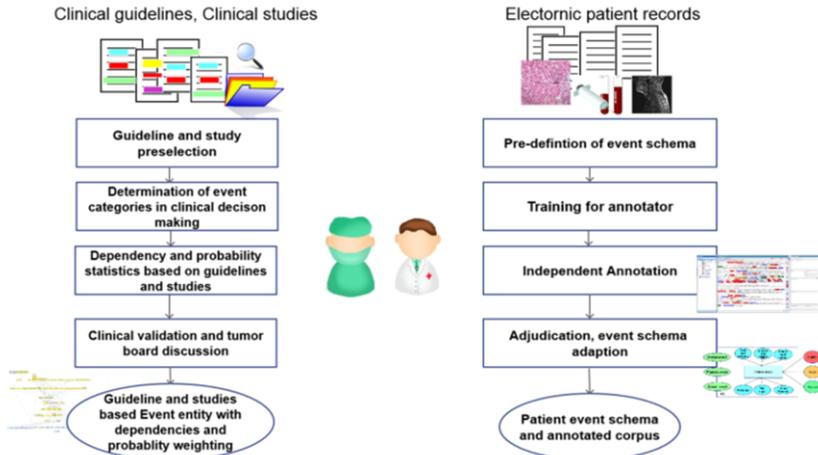


Figure 1. Two approaches for the generation of an event representation

2) The data sources of **guideline and study-based modeling** are the most recent research literatures and evaluated reports for clinical studies. The knowledge curator selects the important entities referring to treatment decisions from the aforementioned data sources whereas the entities are categorized into topic groups. Then, the dependencies between the entities are determined regarding the entire clinical treatment planning. At last, a retrospective study based on the decision and clinical outcomes from the local hospital are used as case experiences to adapt the dependence graph.

### 3. Results

With the aforementioned two approaches, a general patient centered event schema with the corresponding corpus is established. Further, a representation of adverse event with fine-grained hierarchy in the domain of laryngeal cancer treatment has been created.

#### 3.1. Expert based annotation

Three annotators (two medical backgrounds and one pharmaceutical background) have taken part in the task. 300 ICU nurse letters from MIMIC II<sup>2</sup> data base have been provided to the annotator. The patient status events are those expressions which represent the status of the patient. It is observed by the nurse on duty and documented as feedback and reaction of the patient after the treatment. The clinical event refers to the clinical treatment, surgical intervention, therapy planning, complications and organizational event (transfer, readmission) in the treatment pathway. Pharma event is

<sup>2</sup> <https://physionet.org/mimic2/>

related to usage of the medications, drugs. The last event type is social connection, the visiting of relatives or feedback from the social connection or even symptoms of the relatives can also reflect the status of the patient. According to the annotation, the two medical experts have reached 79.8% inter annotator agreement, while the pharma annotator has recognized the largest amount of medication events. The three types of events may have a positive or negative influence on the patient status. In total, 7080 patient status event and 2040 clinical events, 1380 pharma events and 535, social events, 6039 influence relations between patient status events and other events have been found in the corpus. The schema and the annotated corpus provided one general patient centered event representation. The simple influence pattern (positive and negative) between patient statuses can be extended with fine-grained interaction rules. The corpus is the foundation of further experiment for learning based event recognition.

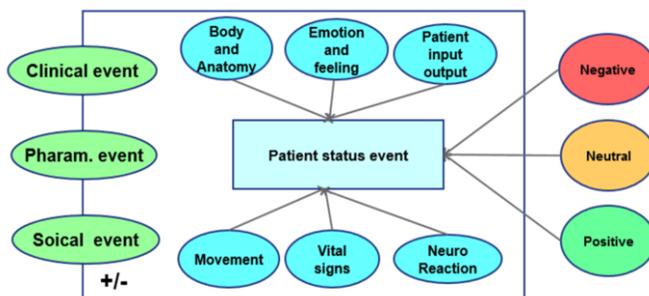


Figure 2 Event schema: Result of the expert based event annotation

### 3.2. Guideline and clinical studies based event modeling.

Within the clinical event, there is one special type of event called complication. It is one kind of unwanted, critical event happening during the treatment. It may threaten patient safety and lead to the failure of the entire treatment plan. We used this specific event type to exploit the guideline and study-based approach for generating an event representation. More than 65 clinical studies and guidelines relevant for head and neck cancer treatments were analyzed by knowledge curators who generated a hierarchical event representation. 6 superordinate categories of the treatment for the laryngeal carcinoma together with 146 entities of subtype for complications were summarized (see figure 3). The risk of all the subtypes of the complication in the surgical operation has been calculated according to the integrated statistics in the clinical studies.

## 4. Discussion

Event definition in the clinical domain is a complex problem, which relates to the semantic dependencies and temporal features. The current effectiveness of the automatic event recognition is still far from optimal for the usage in the clinical practice. The two event modeling studies for clinical event in this paper have shown the distance between data driven method and the practical clinical usage in one subdomain of the treatment of a multifactorial disease. The data driven method may start with a limited amount of data and results in a low complexity patient oriented event

representation, while the event modeling based on guidelines can be extremely sophisticated with full of hierarchical abstractions and specialized terminology. Even for one subdomain, high computing capacity is required to complete the complex modeling task. Hence, a suitable clinical domain with reasonable complexity should be chosen for the experiment in consideration of the computing resources.

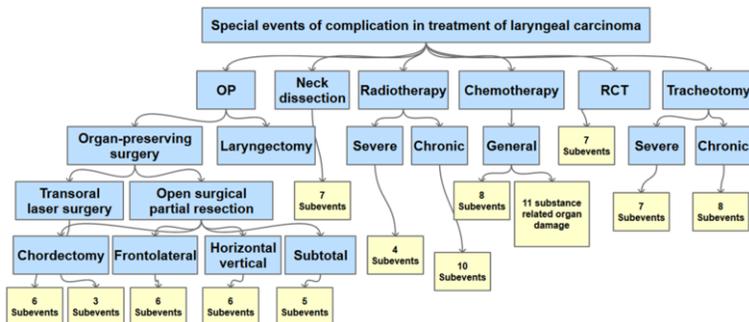


Figure 3. Adverse event in the treatment of laryngeal carcinoma

With patient centered event modeling the clinical knowledge is still required for the follow up predication task. The guideline based event representation must be enhanced with the feedback of the patient and increase the patient influencing factors in the entire treatment decision making process. Technically, the guideline and study based modeling could be extremely complex, while the patient event representation can be realized with current standard technology as independent application logic. In order to balance the differences of complexity between these two types of representation, a suitable form of the combination of two types of representations should be developed, e.g., the application logics based on patient centered event tasks can be packed into a separate multi-agent system. Each agent inquires the necessary guideline based event and dependencies as a snapshot from the guideline event modeling, the concrete predication logics will be outsourced in each independent agent system. The capacity bottleneck can then be avoided. The temporal features of the patient centered event can be defined as discrete observation time points e.g., 24, 48 or 72 hours after clinical interventions. As a next step, the special event of complication based on clinical guidelines will be enriched with patient status event from clinical narratives to facilitate the patient specific treatment decision making.

## References

- [1] H. Xu, SP. Stenner, S. Doan, KB. Johnson, LR. Waitman, JC. Denny. MedEx: a medication information extraction system for clinical narratives. *J Am Med Inform Assoc.* 2010 Jan-Feb; **17**(1):19-24.
- [2] W. Sun, A. Rumshisky, O. Uzuner. Evaluating temporal relations in clinical text: 2012 i2b2 Challenge. *J Am Med Inform Assoc.* 2013 Sep-Oct; **20**(5):806-13.
- [3] Ö. Uzuner, BR. South, S. Shen, SL. DuVall. 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *J Am Med Inform Assoc.* 2011 Sep-Oct; **18**(5):552-6
- [4] O. Bodenreider, and A. McCray. Exploring semantic groups through visual approaches. *Journal of Biomedical Informatics*, 2013. **36**:414–432
- [5] A. Stubbs, C. Kotfila, H. Xu, Ö. Uzuner. Identifying risk factors for heart disease over time: Overview of 2014 i2b2/UTHealth shared task Track 2. *J Biomed Inform.* 2015 Dec; **58** Suppl:S67-77.