Original Paper

Improving the Efficacy of the Data Entry Process for Clinical Research With a Natural Language Processing–Driven Medical Information Extraction System: Quantitative Field Research

Jiang Han[1,2*], MPH; Ken Chen[3*], BS; Lei Fang[3], MS; Shaodian Zhang[3,4], PhD; Fei Wang[3,5], PhD; Handong Ma[3,6], MS; Liebin Zhao[2,7], MS; Shijian Liu[1,2], PhD

[1]Pediatric Translational Medicine Institute, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[2]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[3]Synyi Research, Shanghai, China
[4]APEX Data and Knowledge Management Lab, Shanghai Jiao Tong University, Shanghai, China
[5]Department of Healthcare Policy and Research, Weill Cornell Medicine, New York, NY, United States
[6]Department of computer science, Shanghai Jiao Tong University, Shanghai, China
[7]Child Health Advocacy Institute, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China
*these authors contributed equally

Corresponding Author:
Shijian Liu, PhD
Pediatric Translational Medicine Institute
Shanghai Children's Medical Center
Shanghai Jiao Tong University School of Medicine
1678 Dongfang Road, Pudong New Area
Shanghai,
China
Phone: 86 86 21 38625637
Fax: 86 86 21 38625637
Email: arrow64@163.com

Abstract

Background: The growing interest in observational trials using patient data from electronic medical records poses challenges to both efficiency and quality of clinical data collection and management. Even with the help of electronic data capture systems and electronic case report forms (eCRFs), the manual data entry process followed by chart review is still time consuming.

Objective: To facilitate the data entry process, we developed a natural language processing–driven medical information extraction system (NLP-MIES) based on the i2b2 reference standard. We aimed to evaluate whether the NLP-MIES–based eCRF application could improve the accuracy and efficiency of the data entry process.

Methods: We conducted a randomized and controlled field experiment, and 24 eligible participants were recruited (12 for the manual group and 12 for NLP-MIES–supported group). We simulated the real-world eCRF completion process using our system and compared the performance of data entry on two research topics, pediatric congenital heart disease and pneumonia.

Results: For the congenital heart disease condition, the NLP-MIES–supported group increased accuracy by 15% (95% CI 4%-120%, *P*=.03) and reduced elapsed time by 33% (95% CI 22%-42%, *P*<.001) compared with the manual group. For the pneumonia condition, the NLP-MIES–supported group increased accuracy by 18% (95% CI 6%-32%, *P*=.008) and reduced elapsed time by 31% (95% CI 19%-41%, *P*<.001).

Conclusions: Our system could improve both the accuracy and efficiency of the data entry process.

(JMIR Med Inform 2019;7(3):e13331) doi:10.2196/13331

KEYWORDS

electronic data capture; electric medical records; case report form; natural language processing; field research

Introduction

According to ClinicalTrials.gov [1], the number of clinical trials worldwide has increased exponentially in recent years. Clinicians and researchers use evidence from interventional and observational trials to determine the effectiveness of treatments or interventions. Interventional trials, such as randomized controlled trials, compare the efficacy of interventions under relatively ideal cohorts to get unbiased estimates of effects. However, reality is far more complicated, and these ideal cohorts limit generalizability of results obtained to broader patient populations and settings. Moreover, due to high expenses and the short research cycle, interventional trials could hardly provide evaluations of effectiveness and safety for large populations and long-term follow-ups. As supplements, many observational trials, such as retrospective cohort studies, cross-sectional studies, and real-world evidence studies, use patient historical data collected at the point of care to compare effectiveness and safety of treatments in clinical practice settings in nonexperimental ways. Such observational trials usually have larger cohort sizes and longer follow-up periods. Growing interest in using these approaches poses new challenges to effective and efficient collection of patient electronic medical records (EMRs).

Manual data entry based on paper-and-pen case report forms (CRFs) followed by chart review is the conventional way of clinical trial data collection. With the development of health care information technology, electronic data capture (EDC) systems, which accelerate the data collection process and assure data quality with real-time data entry, review, analysis, and verification [2], emerge as a timely solution that is in high demand. Driven by the prevalent use of EDC systems, CRFs gradually transitioned from paper to electronic forms [3]. Many studies have suggested that data entry using electronic CRF (eCRF) applications of EDC systems could achieve higher efficiency and accuracy at a lower cost than the conventional paper-and-pen approach [2,4-8]. However, neither EDC nor eCRF fundamentally changed the essential ways of how the data are collected. Especially for observational trials using patient data, researchers still need to manually transcribe the data one by one from EMRs. The data entry process takes time and becomes a significant efficiency bottleneck.

The 2018 guidance from the US Food and Drug Administration [9] emphasized the importance of interoperability between electronic health records (EHRs) and EDCs. It also promoted the idea of secondary use of source data at the time of care to prepopulate eCRFs without specific user efforts. The guidance focused more on the use of structured data, such as demographics, vital signs, and laboratory data, but little on the use of unstructured clinical narratives, which account for about 80% of the patient care information [10]. To achieve data interoperability for these unstructured narratives, many EDC systems created predesigned patient information templates including standardized documentation or forms for coded data entry in lieu of free text documentation to structuralize the medical records [11,12]. Clinicians record patient information under the guidance of these templates, and at the same time the system stored the coded data from templates for future analysis. Patient information templates can help data collection for research and patient care, integrate EDC and EMRs, and automatically prepopulate the eCRF. However, limitations of the templates were obvious. For clinicians, the one-size-fits-all templates restricted freedom of expression. For researchers, the predesigned data elements limited usability of the data in different research topics.

The development of natural language processing (NLP) technologies provides new potential for better secondary use of free unstructured EMR data. Informatics for integrating biology and the bedside (i2b2) has posed NLP challenges to extract information, including clinical finding, test, treatment, medication, clinical event, and time information, from clinical notes and discharge summaries [13-16] and promoted a series of commercial medical applications focusing on post hoc structuralization of medical records [17-19]. Nonetheless, as one of the main topics on secondary use of patient EMR, unstructured data collection based on NLP technology has not been well studied.

In order to fill in this gap, we developed an NLP-driven medical information extraction system (NLP-MIES) based on i2b2 reference standards for concept extraction, assertion, and relation classification. After manually constructing eCRFs and binding data elements using concepts from the Systematized Nomenclature of Medicine–Clinical Terms (SNOMED-CT) or the radiology-specific ontology (RadLex) developed by the Radiological Society of North America, our system can scan clinical notes and image diagnostic reports, find related medical concepts, and automatically prepopulate data elements with associated values. To further compare the accuracy and efficiency between manual data entry and NLP technology–supported data entry, we conducted a randomized and controlled field experiment. We created a mock-up eCRF application that enables users to review medical records and enter, modify, and verify the data prepopulated by NLP-MIES. We recruited clinicians and researchers to use the application to finish a certain amount of simply designed eCRFs in the limited time. Based on these designs, we simulated a real-world eCRF filling process and aimed to quantitatively evaluate how NLP technologies could improve efficacy of data collection of clinical research and identify potential problems that are not neglectable in future NLP-driven EDC design.

Methods

Natural Language Processing–Driven Medical Information Extraction System

We leveraged the methods developed for the 2010 i2b2/Veterans Affairs (VA) challenge as the primary reference for Chinese medical NLP machine learning practices in NLP-MIES, which includes Chinese word segmentation, named entity recognition, assertion classification, and relation extraction [14,20-22]. On the basis of the predefined entities (medical problems, tests, treatments) and relation types (medical problems and treatments, medical problems and tests, medical problems and other medical problems) from the 2010 i2b2/VA challenge, in order to extract more information from medical records, we added four new entities (body structure, observable, qualifier, value) and four new types of relations (body structures and observables, medical problems and observables, observables and qualifiers, observables and values). After preprocessing by an associated value dimension algorithm [23], entities from medical texts can be rearranged according to their relations. We then adopted an improved longest common subsequence algorithm to map these aligned entities and relations into Chinese SNOMED-CT and RadLex concepts and synonyms [24]. Figure 1 shows the overall workflow of NLP-MIES.

Electronic Clinical Research Form

We constructed simple eCRFs for two disease conditions (pediatric congenital heart disease and pneumonia) to evaluate the efficacy of NLP-MIES. To make the eCRFs closer to the real ones, we invited clinical researchers from the departments of pediatric cardiothoracic surgery and pediatric respiratory medicine to help design the eCRFs. The types of CRF data elements include true-false (participant judges whether a certain condition or medical problem exists, doesn’t exist, or is not mentioned in a certain case and chooses the button accordingly—for example, patient had a disturbance of consciousness: true, false, or not mentioned); multiple choice (participant should click the button corresponding to one or more conditions or medical problems associated with a certain patient—for example, which of the following are the chief complaints of the patient: cardiac murmur, cyanosis, or dyspnea); and fill-in-the-blank (participant should enter the value for each data element—for example, the lesion size of ventricular septal defect is ___ cm). Figures 2 and 3 show examples of eCRF design.

Figure 1. Workflow of the natural language processing–driven medical information extraction system. EMR: electronic medical record; NLP: natural language processing; eCRF: electronic case report form.

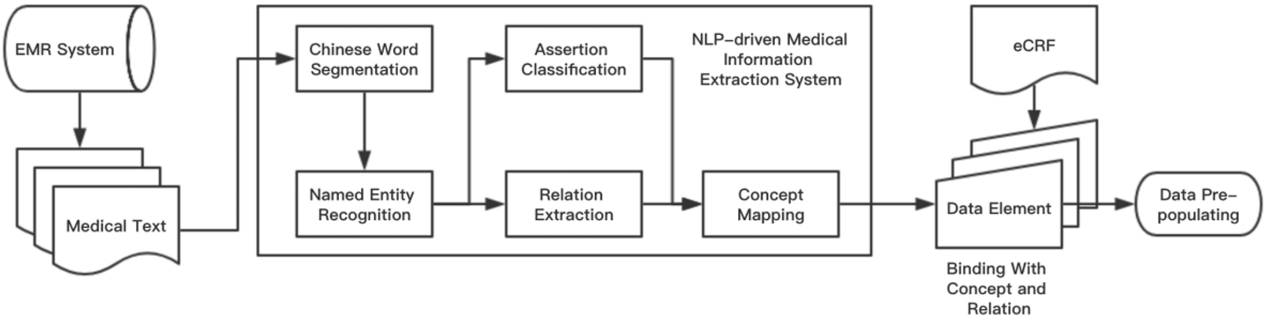
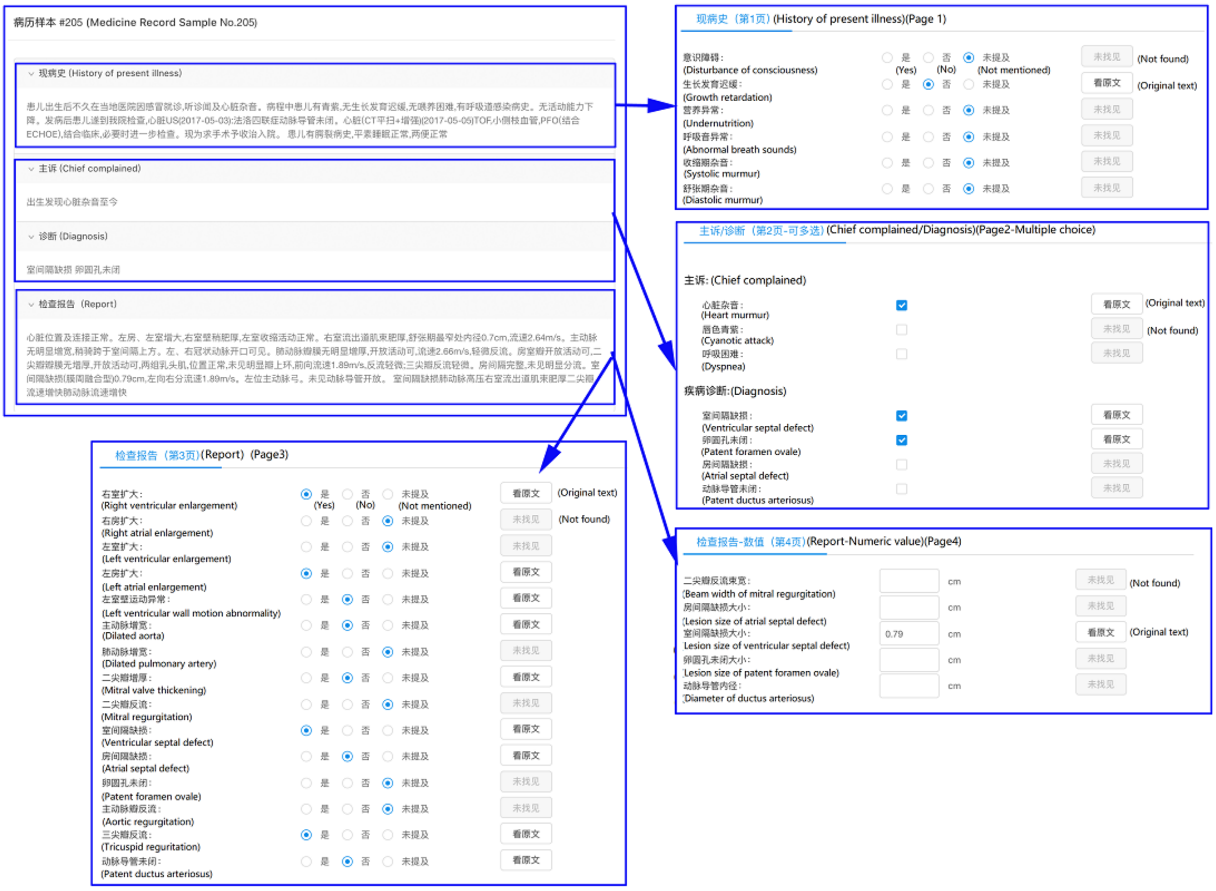


Figure 2. Electronic case report form design for congenital heart disease.



病历样本 #205 (Medicine Record Sample No.205)

现病史 (History of present illness)

患儿出生后不久在当地医院因感冒就诊,听诊闻及心脏杂音。病程中患儿有青紫,无生长发育迟缓,无喂养困难,有呼吸道感染病史。无活动能力下降。发病后患儿遂到我院检查,心脏US(2017-05-03):法洛四联症动脉导管未闭。心脏(CT平扫+增强)(2017-05-05)TOF,小侧枝血管,PFO(结合ECHOE),结合临床,必要时进一步检查。现为求手术予收治入院。 患儿有腭裂病史,平素睡眠正常,两便正常

主诉 (Chief complained)

出生发现心脏杂音至今

诊断 (Diagnosis)

室间隔缺损 卵圆孔未闭

检查报告 (Report)

心脏位置及连接正常。左房、左室增大,右室壁稍肥厚,左室收缩活动正常。右室流出道肌束肥厚,舒张期最窄处内径0.7cm,流速2.64m/s。主动脉无明显增宽,稍骑跨于室间隔上方。左、右冠状动脉开口可见。肺动脉瓣膜无明显增厚,开放活动可,流速2.66m/s,轻微反流。房室瓣开放活动可,二尖瓣瓣膜无增厚,开放活动可,两组乳头肌,位置正常,未见明显瓣上环,前向流速1.89m/s,反流轻微;三尖瓣反流轻微。房间隔完整,未见明显分流。室间隔缺损(膜周融合型)0.79cm,左向右分流速1.89m/s。左位主动脉弓。未见动脉导管开放。 室间隔缺损肺动脉高压右室流出道肌束肥厚二尖瓣流速增快肺动脉流速增快

现病史（第1页）(History of present illness)(Page 1)

意识障碍： (Disturbance of consciousness)	○ 是 (Yes)	○ 否 (No)	● 未提及 (Not mentioned)	未找见	(Not found)
生长发育迟缓： (Growth retardation)	○ 是	● 否	○ 未提及	看原文	(Original text)
营养异常： (Undernutrition)	○ 是	○ 否	● 未提及	未找见	
呼吸音异常： (Abnormal breath sounds)	○ 是	○ 否	● 未提及	未找见	
收缩期杂音： (Systolic murmur)	○ 是	○ 否	● 未提及	未找见	
舒张期杂音： (Diastolic murmur)	○ 是	○ 否	● 未提及	未找见	

主诉/诊断（第2页-可多选）(Chief complained/Diagnosis)(Page2-Multiple choice)

主诉: (Chief complained)

心脏杂音： (Heart murmur)	☑	看原文	(Original text)
唇色青紫： (Cyanotic attack)	☐	未找见	(Not found)
呼吸困难： (Dyspnea)	☐	未找见	

疾病诊断:(Diagnosis)

室间隔缺损： (Ventricular septal defect)	☑	看原文
卵圆孔未闭： (Patent foramen ovale)	☑	看原文
房间隔缺损： (Atrial septal defect)	☐	未找见
动脉导管未闭： (Patent ductus arteriosus)	☐	未找见

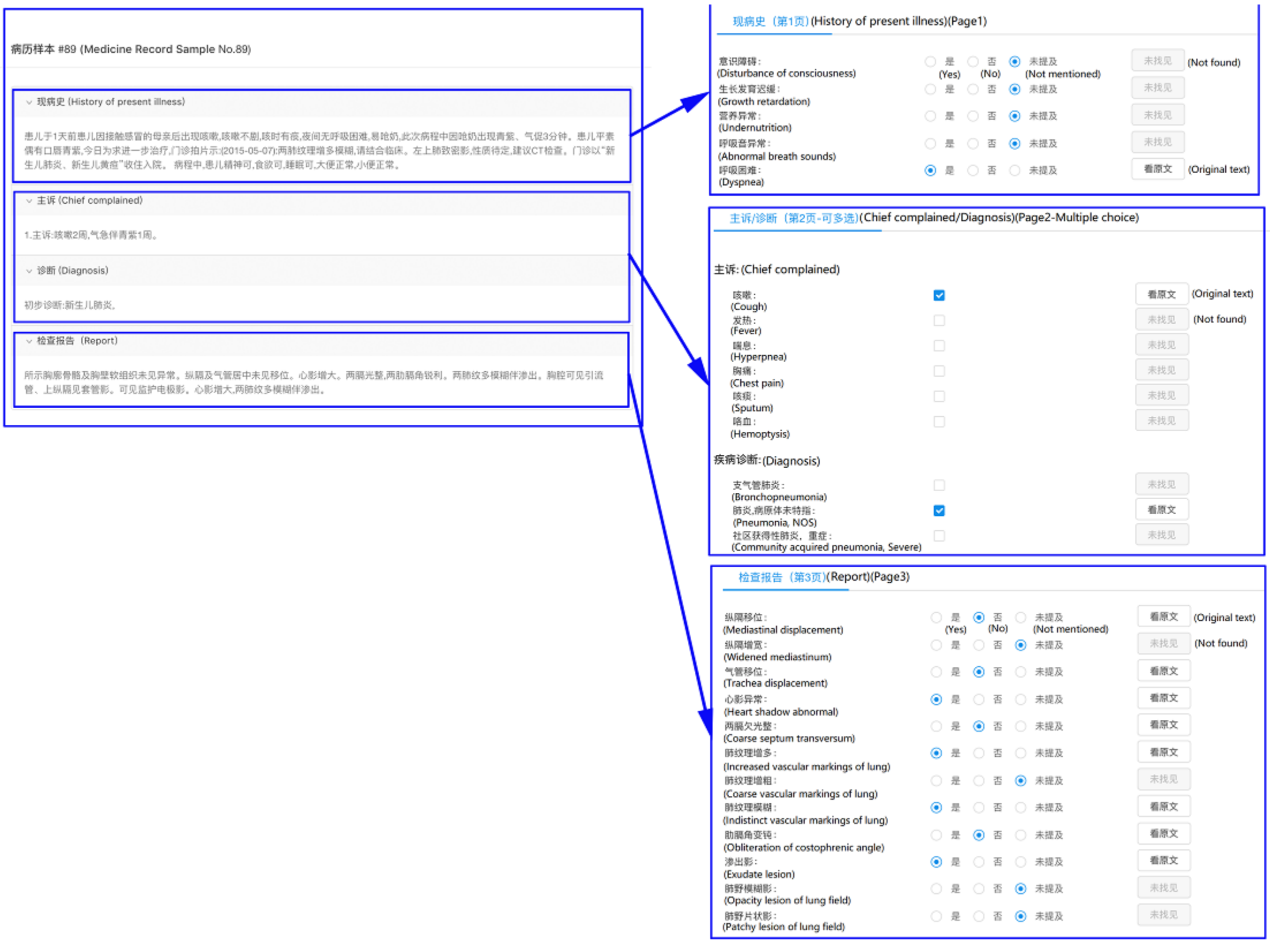
检查报告（第3页）(Report)（Page3）

右室扩大： (Right ventricular enlargement)	● 是 (Yes)	○ 否 (No)	○ 未提及 (Not mentioned)	看原文	(Original text)
右房扩大： (Right atrial enlargement)	○ 是	○ 否	● 未提及	未找见	(Not found)
左室扩大： (Left ventricular enlargement)	○ 是	○ 否	● 未提及	未找见	
左房扩大： (Left atrial enlargement)	● 是	○ 否	○ 未提及	看原文	
左室壁运动异常： (Left ventricular wall motion abnormality)	○ 是	● 否	○ 未提及	看原文	
主动脉增宽： (Dilated aorta)	○ 是	● 否	○ 未提及	看原文	
肺动脉增宽： (Dilated pulmonary artery)	○ 是	○ 否	● 未提及	未找见	
二尖瓣增厚： (Mitral valve thickening)	○ 是	● 否	○ 未提及	看原文	
二尖瓣反流： (Mitral regurgitation)	○ 是	○ 否	● 未提及	未找见	
室间隔缺损： (Ventricular septal defect)	● 是	○ 否	○ 未提及	看原文	
房间隔缺损： (Atrial septal defect)	○ 是	● 否	○ 未提及	看原文	
卵圆孔未闭： (Patent foramen ovale)	○ 是	○ 否	● 未提及	未找见	
主动脉瓣反流： (Aortic regurgitation)	○ 是	○ 否	● 未提及	未找见	
三尖瓣反流： (Tricuspid reguritation)	● 是	○ 否	○ 未提及	看原文	
动脉导管未闭： (Patent ductus arteriosus)	○ 是	● 否	○ 未提及	看原文	

检查报告-数值（第4页）(Report-Numeric value)(Page4)

二尖瓣反流束宽： (Beam width of mitral regurgitation)		cm	未找见	(Not found)
房间隔缺损大小： (Lesion size of atrial septal defect)		cm	未找见	
室间隔缺损大小： Lesion size of ventricular septal defect)	0.79	cm	看原文	(Original text)
卵圆孔未闭大小： Lesion size of patent foramen ovale)		cm	未找见	
动脉导管内径： Diameter of ductus arteriosus)		cm	未找见	

Figure 3. Electronic case report form design for pneumonia.



病历样本 #89 (Medicine Record Sample No.89)

⌄ 现病史 (History of present illness)

患儿于1天前患儿因接触感冒的母亲后出现咳嗽,咳嗽不剧,咳时有痰,夜间无呼吸困难,易呛奶,此次病程中因呛奶出现青紫、气促3分钟。患儿平素偶有口唇青紫,今日为求进一步治疗,门诊拍片示:(2015-05-07):两肺纹理增多模糊,请结合临床。左上肺致密影,性质待定,建议CT检查。门诊以“新生儿肺炎、新生儿黄疸”收住入院。 病程中,患儿精神可,食欲可,睡眠可,大便正常,小便正常。

⌄ 主诉 (Chief complained)

1.主诉:咳嗽2周,气急伴青紫1周。

⌄ 诊断 (Diagnosis)

初步诊断:新生儿肺炎。

⌄ 检查报告 (Report)

所示胸廓骨骼及胸壁软组织未见异常。纵隔及气管居中未见移位。心影增大。两膈光整,两肋膈角锐利。两肺纹理多模糊伴渗出。胸腔可见引流管、上纵隔见套管影。可见监护电极影。心影增大,两肺纹多模糊伴渗出。

现病史（第1页）(History of present illness)(Page1)

意识障碍：(Disturbance of consciousness) ○ 是 (Yes) ○ 否 (No) ● 未提及 (Not mentioned) 未找见 (Not found)
生长发育迟缓：(Growth retardation) ○ 是 ○ 否 ● 未提及 未找见
营养异常：(Undernutrition) ○ 是 ○ 否 ● 未提及 未找见
呼吸音异常：(Abnormal breath sounds) ○ 是 ○ 否 ● 未提及 未找见
呼吸困难：(Dyspnea) ● 是 ○ 否 ○ 未提及 看原文 (Original text)

主诉/诊断（第2页-可多选）(Chief complained/Diagnosis)(Page2-Multiple choice)

主诉: (Chief complained)

咳嗽：(Cough) ☑ 看原文 (Original text)
发热：(Fever) ☐ 未找见 (Not found)
喘息：(Hyperpnea) ☐ 未找见
胸痛：(Chest pain) ☐ 未找见
咳痰：(Sputum) ☐ 未找见
咯血：(Hemoptysis) ☐ 未找见

疾病诊断: (Diagnosis)

支气管肺炎：(Bronchopneumonia) ☐ 未找见
肺炎,病原体未特指：(Pneumonia, NOS) ☑ 看原文
社区获得性肺炎，重症：(Community acquired pneumonia, Severe) ☐ 未找见

检查报告（第3页）(Report)(Page3)

纵隔移位：(Mediastinal displacement) ○ 是 (Yes) ● 否 (No) ○ 未提及 (Not mentioned) 看原文 (Original text)
纵隔增宽：(Widened mediastinum) ○ 是 ○ 否 ● 未提及 未找见 (Not found)
气管移位：(Trachea displacement) ○ 是 ● 否 ○ 未提及 看原文
心影异常：(Heart shadow abnormal) ● 是 ○ 否 ○ 未提及 看原文
两膈欠光整：(Coarse septum transversum) ○ 是 ● 否 ○ 未提及 看原文
肺纹理增多：(Increased vascular markings of lung) ● 是 ○ 否 ○ 未提及 看原文
肺纹理增粗：(Coarse vascular markings of lung) ○ 是 ○ 否 ● 未提及 未找见
肺纹理模糊：(Indistinct vascular markings of lung) ● 是 ○ 否 ○ 未提及 看原文
肋膈角变钝：(Obliteration of costophrenic angle) ○ 是 ● 否 ○ 未提及 看原文
渗出影：(Exudate lesion) ● 是 ○ 否 ○ 未提及 看原文
肺野模糊影：(Opacity lesion of lung field) ○ 是 ○ 否 ● 未提及 未找见
肺野片状影：(Patchy lesion of lung field) ○ 是 ○ 否 ● 未提及 未找见

We further divided the data element true-false into two parts based on where the elements should be retrieved from: admission records (true-false I) or imaging reports (true-false II). All data elements were bound with SNOMED-CT or RadLex concepts and relations, such as disturbance of consciousness (concept, medical problem, SNOMED-CT ID: 3006004), cardiac murmur (concept, medical problem, SNOMED-CT ID: 42842009), lesion size (concept, observable, SNOMED-CT ID: 246116008) of (relation, medical problems and observables) ventricular septal defect (concept, medical problem, RadLex ID: RID3277).

Medical Text From the Electronic Medical Record System

For the congenital heart disease condition, we included admission records and ultrasonic cardiogram reports from pediatric patients aged 2 hours to 14 years with congenital heart disease (including atrial septal defect, ventricular septal defect, patent ductus arteriosus, patent foramen ovale, etc) attending the department of cardiothoracic surgery of Shanghai Children's Medical Center from July 1, 2016, to July 1, 2017.

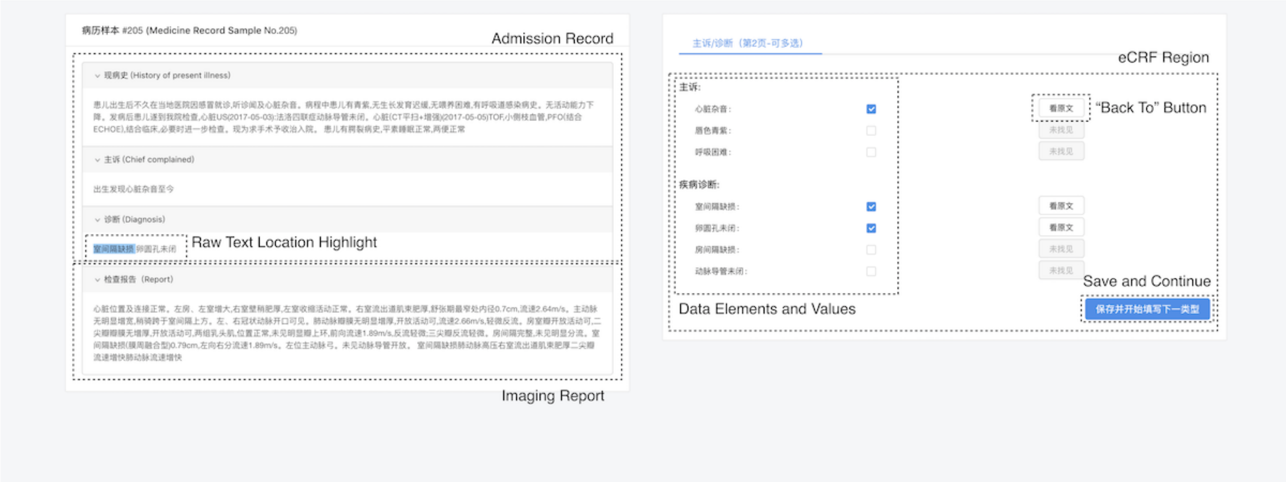
For the pneumonia condition, we included admission records and chest x-ray reports from pediatric patients aged 6 months to 14 years with pneumonia (including bronchopneumonia, viral pneumonia, bacterial pneumonia, mycoplasma pneumonia, lobar pneumonia, lobular pneumonia, etc) attending the department of respiratory medicine of Shanghai Children's Medical Center from July 1, 2016, to July 1, 2017.

All medical texts were from the EMR system of Shanghai Children's Medical Center and were de-identified. We randomly selected 60 patient cases for each condition. A total of 120 cases and 240 medical texts were included.

System Functions and Human-Computer Interaction

We developed a graphical user interface for easy browsing of imported patient medical texts as shown in Figure 4. User can see imported admission records, imaging reports, and eCRFs on the screen. When NLP-MIES was enabled, our system automatically scanned the texts, found medical concepts mentioned in raw texts, identified assertion or value information, and prepopulated the data elements accordingly. Our system recorded the raw text location where each medical concept was extracted. When necessary, user could directly click the "back to" button to highlight the location for further data verification. Each eCRF was divided into three or four parts according to the types of data elements (Figures 2 and 3). During the experiment, the elapsed time for finishing each part was automatically recorded by system.

Figure 4. Graphic user interface for electronic case report form (eCRF) data entry.



病历样本 #205 (Medicine Record Sample No.205)

现病史 (History of present illness)

患儿出生后不久在当地医院因感冒就诊,听诊闻及心脏杂音。病程中患儿有青紫,无生长发育迟缓,无喂养困难,有呼吸道感染病史。无活动能力下降。发病后患儿遂到我院检查,心脏US(2017-05-03):法洛四联症动脉导管未闭。心脏(CT平扫+增强)(2017-05-05)TOF,小侧枝血管,PFO(结合ECHOE),结合临床,必要时进一步检查。现为求手术予收治入院。 患儿有腭裂病史,平素睡眠正常,两便正常

主诉 (Chief complained)

出生发现心脏杂音至今

诊断 (Diagnosis)

室间隔缺损 卵圆孔未闭

检查报告（Report）

心脏位置及连接正常。左房、左室增大,右室壁稍肥厚,左室收缩活动正常。右室流出道肌束肥厚,舒张期最窄处内径0.7cm,流速2.64m/s。主动脉无明显增宽,稍骑跨于室间隔上方。左、右冠状动脉开口可见。肺动脉瓣膜无明显增厚,开放活动可,流速2.66m/s,轻微反流。房室瓣开放活动可,二尖瓣瓣膜无增厚,开放活动可,两组乳头肌,位置正常,未见明显瓣上环,前向流速1.89m/s,反流轻微;三尖瓣反流轻微。房间隔完整,未见明显分流。室间隔缺损(膜周融合型)0.79cm,左向右分流速1.89m/s。左位主动脉弓。未见动脉导管开放。 室间隔缺损肺动脉高压右室流出道肌束肥厚二尖瓣流速增快肺动脉流速增快

主诉/诊断（第2页-可多选）

主诉:
心脏杂音: ☑ 看原文
唇色青紫: ☐ 未找见
呼吸困难: ☐ 未找见

疾病诊断:
室间隔缺损: ☑ 看原文
卵圆孔未闭: ☑ 看原文
房间隔缺损: ☐ 未找见
动脉导管未闭: ☐ 未找见

保存并开始填写下一类型

Gold Standard

The ground truth results of eCRFs for all 120 cases were provided by three clinical researchers involved in the eCRF design. We used a two-step strategy to create our gold standard. First, two invited researchers independently extracted data from medical texts and populated eCRFs using an eCRF application but without the support of NLP-MIES. Our system automatically recorded the populated values and elapsed time for each data entry. Second, for pairs in which the two researchers did not have complete agreement, a third researcher resolved inconsistent data extraction between the two researchers.

Study Design

We conducted a randomized and controlled field experiment at Shanghai Children's Medical Center to evaluate whether the NLP-MIES group was more effective and efficient than the manual group in the data entry process of eCRF. Participants holding medical degrees, having clinical research experience, or working as clinicians were eligible for inclusion and recruited in this study. The study was approved by the Human Research Ethics Committees of Shanghai Children's Medical Center. Written informed consent was obtained from all participants prior to randomization.

We randomly allocated the volunteers to two groups by using a completely randomized digital table:

- Manual group: participants should check the data elements in the eCRF, find related information in the medical text, and click or enter values accordingly.
- NLP-MIES–supported group: NLP-MIES prepopulated the data elements in the eCRF. Participants should check the data elements, find related information in the medical text, and verify or correct values accordingly.

Before the experiment, all participants were authorized and trained to use the system and eCRF-based data entry. We chose a relatively quiet place for the experiment to reduce the potential effect of other environmental factors. Each participant was provided with a laptop and asked to complete all cases from 2:00 pm to 5:00 pm. Participants failing to complete the eCRFs in that time frame were excluded from the data analysis. The order of the 120 cases was randomly shuffled for each participant.

Outcomes and Statistical Analysis

We calculated average accuracy and elapsed time for each participant to finish all assigned eCRFs and compared the differences between the manual and NLP-MIES–supported group. To further analyze data entry errors made by participants under the support of NLP-MIES, we performed a post hoc error analysis for the results provided by NLP-MIES–supported group. We calculated the percentages of two types of data entry errors: error with modification and error without modification. We defined an error with modification as a data entry error made when a participant incorrectly modified a prepopulated result and an error without modification as a data entry error made when a participant kept an incorrect prepopulated result.

Educational and psychological studies have indicated that the distributions of the measurements of how many points participants could get in a certain test and how much time it would take a participant to respond to a certain stimulus (reaction time) were right-skewed [25-27]. Thus, we expected the data for each participant's average accuracy and elapsed time for finishing eCRFs would not be normally distributed and described them using their median and interquartile range. To evaluate the differences between groups, we made a logarithmic transformation of the data and performed independent group *t* tests with SAS 9.2 (SAS Institute) software. *P* value, logarithmic mean difference (MD), ratio of change in geometric mean (exponential of logarithmic mean difference), and corresponding 95% confidence interval were calculated [28]. We considered two-sided *P* values <.05 as statistically significant.

Results

Participant Characteristics

We recruited a total of 24 eligible participants, 12 for the manual group and 12 for the NLP-MIES–supported group. All the participants successfully completed the eCRFs within the required time. The mean age of participants was 24.66 (SD 2.30) years (manual group 24.70 [SD 2.47] years, NLP-MIES group 24.48 [SD 2.36] years; *P*=.73); 33% (8/24) of participants were men and 67% (16/24) were women. There were no significant differences between the characteristics of the participants in the two groups.

The overall interoperator consistency rate was 96.85% (1627/1680) for the congenital heart disease condition and 94.82% (1081/1440) for the pneumonia condition (Multimedia Appendix 1).

Accuracy

The overall average accuracy for the congenital heart disease and pneumonia eCRFs was significantly higher in the NLP-MIES–supported group than the manual group (congenital heart disease, *P*=.03; pneumonia, *P*=.008; Table 1). For the congenital heart disease eCRFs, the logarithmic MD of average accuracy between groups was 0.14 (95% CI 0.03-0.25), corresponding to an increase of 15% (95% CI 4%-120%) in geometric mean. Similarly, for the pneumonia eCRFs, the logarithmic MD was 0.17 (95% CI 0.06-0.28), corresponding to an increase of 18% (95% CI 6%-32%) in geometric mean. Comparing by types of data elements, the average accuracy was significantly higher in the NLP-MIES–supported group for all types except true-false II and fill-in-the-blank on the congenital heart disease eCRFs. The average accuracy of NLP-MIES prepopulation was slightly higher than median average accuracy of the manual group but lower than that of the NLP-MIES–supported group for most data element types.

Elapsed Time

The overall average time elapsed for congenital heart disease and pneumonia eCRFs was significantly lower in the NLP-MIES–supported group than the manual group (congenital heart disease, *P*<.001; pneumonia, *P*<.001; Table 2). For the congenital heart disease eCRFs, the logarithmic MD of average time elapsed was –0.40 (95% CI –0.55 to –0.25), corresponding to a reduction of 33% (95% CI 22% to 42%) in geometric mean. For the pneumonia eCRFs, the logarithmic MD was –0.37 (95% CI –0.53 to –0.21), corresponding to a reduction of 31% (95% CI 19% to 41%) in geometric mean. Comparing by types of data elements, the average elapsed time was significantly lower in the NLP-MIES–supported group for all types.

Error Analysis

Post hoc error analysis showed that errors without modification held the majority of error cases in all types of data elements (Table 3), and the overall percentage of errors without modification was almost 2.5 time higher than the percentage of errors with modification.

Table 1. Average accuracy for electronic case report form data entry.

Type of disease and data element	NLP[a] only	Manual group (median, IQR[b])	NLP-MIES[c] group (median, IQR)	Logarithmic mean difference (95% CI)	Ratio of change in geometric mean (95% CI)	*P* value
Congenital heart disease						
True-false I[d]	97.50	79.17 (66.74, 84.17)	96.81 (95.69, 97.29)	0.41 (0.04 to 0.79)	1.51 (1.03 to 2.20)	.04
True-false II[e]	92.00	95.39 (92.67, 95.89)	97.78 (97.19, 98.44)	0.21 (–0.01 to 0.10)	1.10 (0.99 to 1.24)	.10
Multiple choice	89.33	82.80 (73.13, 85.83)	95.00 (94.58, 97.42)	0.29 (0.10 to 0.49)	1.34 (1.10 to 1.63)	.009
Fill-in-the-blank	94.17	96.33 (95.25, 97.00)	97.00 (95.83, 97.42)	0.01 (–0.01 to 0.02)	1.01 (0.99 to 1.02)	.22
Overall	92.77	90.42 (87.75, 92.68)	97.17 (96.83, 97.44)	0.14 (0.03 to 0.25)	1.15 (1.04 to 2.20)	.03
Pneumonia						
True-false I	88.00	70.83 (65.25, 77.75)	88.17 (87.25, 89.00)	0.30 (0.11 to 0.50)	1.35 (1.11 to 1.65)	.009[f]
True-false II	94.44	91.25 (88.26, 93.78)	95.83 (95.21, 96.81)	0.11 (0.01 to 0.21)	1.12 (1.01 to 1.23)	.04
Multiple choice	80.83	67.50 (50.21, 72.50)	81.25 (77.92, 85.00)	0.33 (0.14 to 0.52)	1.39 (1.15 to 1.68)	.003[f]
Overall	84.15	84.21 (80.53, 86.23)	92.19 (91.49, 93.20)	0.17 (0.06 to 0.28)	1.18 (1.06 to 1.32)	.008

[a] NLP: natural language processing.

[b] IQR: interquartile range.

[c] NLP-MIES: NLP-driven medical information extraction system.

[d] True-false I: data elements retrieved from admissions records.

[e] True-false II: data elements retrieved from imaging reports (ultrasonic cardiogram or chest x-ray).

[f] Independent group *t* test.

Table 2. Average elapsed time for electronic case report form data entry.

Type of disease and data element	Manual group seconds (median, IQR[a])	NLP-MIES[b] group seconds (median, IQR)	Logarithmic mean difference (95% CI)	Ratio of change in geometric mean (95% CI)	*P* value
Congenital heart disease					
True-false I[c]	26.43 (21.43, 30.24)	13.84 (11.83, 16.06)	–0.71 (–1.02 to –0.39)	0.49 (0.36 to 0.68)	<.001
True-false II[d]	49.48 (43.08, 51.44)	35.47 (31.34, 38.63)	–0.29 (–0.46 to –0.11)	0.75 (0.63 to 0.89)	.003
Multiple choice	9.70 (10.61, 12.29)	7.34 (7.47, 8.55)	–0.36 (–0.53 to –0.19)	0.70 (0.59 to 0.82)	<.001
Fill-in-the-blank	18.41 (17.35, 19.60)	12.38 (11.38, 14.70)	–0.34 (–0.50 to –0.17)	0.71 (0.60 to 0.84)	<.001
Overall	103.79 (94.59, 109.39)	69.73 (60.91, 79.66)	–0.40 (–0.55 to –0.25)	0.67 (0.58 to 0.78)	<.001
Pneumonia					
True-false I	28.71 (25.61, 32.61)	15.82 (14.36, 16.88)	–0.64 (–0.97 to –0.30)	0.53 (0.38 to 0.74)	.001
True-false II	31.59 (28.29, 32.49)	25.22 (22.07, 28.80)	–0.19 (–0.35 to –0.03)	0.83 (0.71 to 0.97)	.02
Multiple choice	11.02 (10.65, 12.05)	8.61 (8.05, 9.25)	–0.33 (–0.51 to –0.15)	0.72 (0.60 to 0.86)	.001
Overall	73.28 (65.80, 74.47)	49.42 (44.33, 53.88)	–0.37 (–0.53 to –0.21)	0.69 (0.59 to 0.81)	<.001

[a] IQR: interquartile range.

[b] NLP-MIES: NLP-driven medical information extraction system.

[c] True-false I: data elements retrieved from admissions records.

[d] True-false II: data elements retrieved from imaging reports (ultrasonic cardiogram or chest x-ray).

Table 3. Error analysis for natural language processing–driven medical information extraction system–supported data entry.

Types	Errors, n (%)			
	True-false (n=1167)	Multiple choice (n=439)	Fill-in-the-blank (n=121)	Total (N=1727)
Errors with modification	325 (27.85)	158 (36.00)	16 (13.22)	499 (28.89)
Errors without modification	842 (72.15)	281 (64.01)	105 (86.78)	1228 (71.11)

Discussion

Principal Findings

In this field experiment, we created a mock-up eCRF application with NLP-supported data entry and simulated a real-world eCRF completion process. Results showed a consistent trend across all eCRF topics and data element types indicating NLP-MIES could significantly improve the accuracy and efficiency of data entry. In quantitative evaluation, data entry under the support of NLP-MIES could increase accuracy by approximately (relative change in geometric mean is similar to the change in arithmetic mean) [29] 15% to 18% and reduce elapsed time by one-third.

Many potential factors could contribute to the increased accuracy and efficiency of NLP-MIES–aided data entry. First, we considered NLP-MIES–aided data entry as in essence a process of double-checking—an NLP-MIES check followed by a manual check. In clinical practice, double-checking is a widely used and trusted approach that could significantly reduce medical errors [30,31]. Second, we tried several ways to establish participant trust in NLP-MIES: ensuring NLP-MIES entry accuracy (not worse or even better than manual entry), providing better interpretability (one-click back to raw text), and simplifying system interaction [28]. Third, the overall time elapsed for the manual group was about 50% more than the NLP-MIES–supported group. In our study, higher accuracy was achieved for pneumonia cases than congenital heart disease cases; it may be that extracted information on congenital heart disease cases was more complicated than that of pneumonia cases.

In our post hoc error analysis, we considered errors with modification as cognitive errors. Participants made cognitive errors because they failed to find correct answers (due to limitation of knowledge or lack of training) even though they noticed prepopulated answers were wrong. We considered most errors without modification as commission errors. Participants made commission errors because they followed the prepopulated answers that were incorrect. The result of error analysis indicated that commission errors dominated the data entry quality under the support of NLP-MIES. Overreliance could be a key factor for commission errors and as a side effect of participant trust in NLP-MIES [29]. One possible solution to this problem could be to use NLP-MIES as an independent investigator. In real-world clinical research data management, at least two investigators independently enter data for each case to reduce commission errors and then submit the entries to the clinical research associates (CRAs). The CRAs review and verify the entries to ensure data completeness and quality [30]. In our scenario, the NLP system could act as an independent investigator and provide data entry directly to CRAs rather than prepopulate data for other investigators, and CRAs could make final decisions based on both NLP-MIES–supported and manual entries.

Strengths and Limitations

As far as we know, this is the first study, especially in Chinese language settings, that quantitatively evaluated how NLP technologies could improve the efficiency and efficacy of data collection of clinical research. We believe NLP technologies would be a vital link in the great chain of data exchange between EHRs and EDC. It can potentially extract and transform data from medical text in real time and pose fewer restrictions on clinician freedom of expressions and workflows. In addition, our mock-up NLP-driven eCRF application provided graphical user interface for easy browsing and validation of source text data and data entries to ensure data quality. We believe that the results of our study can provide guidance of future research and development of NLP-driven EDC systems as well as the integration of EDC and EMR systems.

Although the results of our field experiment demonstrated beneficial outcomes for NLP-MIES–supported data entry, there were limitations. First, we did not evaluate the efficacy of NLP-MIES under different prepopulation. Early research has indicated that improving accuracy of the automation system itself may not necessarily improve the performance of human-computer collaboration [31]. Moreover, some studies suggest that automation systems with low accuracy can affect human-computer collaboration and trust [32]. Second, there might be significant differences between our eCRFs and real-world CRFs in contents and types of data elements. Thus, it is inappropriate to extrapolate our quantitative results to real-world settings. Third, since NLP-MIES was designed for Chinese medical records and tested in Chinese eCRFs only, the efficiency of this methodology based on the i2b2 reference standard needs further evaluation in other languages.

Conclusions

In this study, we developed an NLP-driven medical information extraction system based on i2b2 reference standards to facilitate the data entry process of eCRFs for clinical research. We conducted a randomized and controlled field experiment to simulate a real-world data entry process and evaluated the efficacy of our system. The results of our study showed NLP-MIES could significantly improve the accuracy and efficiency of data entry.

Acknowledgments

We would like to thank Gen Gu (software development engineer—natural language processing, Synyi Research, Shanghai, China), Junjie Cai (software development engineer—machine learning, Synyi Research), and Xiaopeng Jia (software development engineer—backend, Synyi Research) for their help and advice during the development of NLP-MIES and the eCRF application. This work was supported by the Shanghai Collaborative Innovation Center for Translational Medicine (TM201720), National Science Foundation of China (81872637, 81728017, 81602868), Shanghai Municipal Commission of Health and Family Planning (201840324, 20164Y0095), National Science and Technology Commission for the Association of Diabetes and Nutrition in Adolescents (2016YFC1305203), Shanghai Children’s Health Service Capacity Construction (GDEK201708), National Human Genetic Resources Sharing Service Platform (2005DKA21300), Science and Technology Development Program of Pudong Shanghai New District (PKJ2017-Y01), Medical and Engineering Cooperation Project of Shanghai Jiao Tong University (YG2017ZD15), Shanghai Professional and Technical Services Platform (18DZ2294100), and the 2019 Science and Technology Innovation–Biomedical Supporting Program of the Shanghai Science and Technology Committee (19441904400).

Authors' Contributions

JH and KC drafted the manuscript and contributed equally to this work. SL, KC, and SZ designed the study. JH and LF collected the data. SL and LZ obtained the funding. LF and KC were involved in data cleaning and verification, and KC analyzed the data. SL, KC, LZ, and FW contributed to the interpretation of the results and critical revision of the manuscript for important intellectual content. SL had the primary responsibility for the final content. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interoperator agreement and elapsed time for each electronic case report form topic.

[PDF File (Adobe PDF File), 36KB - medinform_v7i3e13331_app1.docx]

References

1. ClinicalTrials.gov. Trends, charts, and maps URL: https://clinicaltrials.gov/ct2/resources/trends [accessed 2019-06-12] [WebCite Cache ID 75C1n01tO]
2. Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. PLoS One 2011;6(9):e25348 [FREE Full text] [doi: 10.1371/journal.pone.0025348] [Medline: 21966505]

3. Bellary S, Krishnankutty B, Latha MS. Basics of case report form designing in clinical research. Perspect Clin Res 2014 Oct;5(4):159-166 [FREE Full text] [doi: 10.4103/2229-3485.140555] [Medline: 25276625]
4. Fleischmann R, Decker A, Kraft A, Mai K, Schmidt S. Mobile electronic versus paper case report forms in clinical trials: a randomized controlled trial. BMC Med Res Methodol 2017 Dec 01;17(1):153 [FREE Full text] [doi: 10.1186/s12874-017-0429-y] [Medline: 29191176]
5. Dillon DG, Pirie F, Rice S, Pomilla C, Sandhu MS, Motala AA, African Partnership for Chronic Disease Research (APCDR). Open-source electronic data capture system offered increased accuracy and cost-effectiveness compared with paper methods in Africa. J Clin Epidemiol 2014 Dec;67(12):1358-1363 [FREE Full text] [doi: 10.1016/j.jclinepi.2014.06.012] [Medline: 25135245]
6. Ene-Iordache B, Carminati S, Antiga L, Rubis N, Ruggenenti P, Remuzzi G, et al. Developing regulatory-compliant electronic case report forms for clinical trials: experience with the demand trial. J Am Med Inform Assoc 2009;16(3):404-408 [FREE Full text] [doi: 10.1197/jamia.M2787] [Medline: 19261946]
7. Le Jeannic A, Quelen C, Alberti C, Durand-Zaleski I, CompaRec Investigators. Comparison of two data collection processes in clinical studies: electronic and paper case report forms. BMC Med Res Methodol 2014 Jan 17;14:7 [FREE Full text] [doi: 10.1186/1471-2288-14-7] [Medline: 24438227]
8. Thriemer K, Ley B, Ame SM, Puri MK, Hashim R, Chang NY, et al. Replacing paper data collection forms with electronic data entry in the field: findings from a study of community-acquired bloodstream infections in Pemba, Zanzibar. BMC Res Notes 2012;5:113 [FREE Full text] [doi: 10.1186/1756-0500-5-113] [Medline: 22353420]
9. Food and Drug Administration. Use of electronic health record data in clinical investigations: guidance for industry URL: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM501068.pdf [accessed 2019-06-12] [WebCite Cache ID 75C2iUaBX]
10. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting information from textual documents in the electronic health record: a review of recent research. Yearb Med Inform 2008:128-144. [Medline: 18660887]
11. Matsumura Y, Hattori A, Manabe S, Takahashi D, Yamamoto Y, Murata T, et al. Case report form reporter: a key component for the integration of electronic medical records and the electronic data capture system. Stud Health Technol Inform 2017;245:516-520. [Medline: 29295148]
12. El Fadly A, Rance B, Lucas N, Mead C, Chatellier G, Lastic P, et al. Integrating clinical research with the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform. J Biomed Inform 2011 Dec;44 Suppl 1:S94-S102 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.007] [Medline: 21888989]
13. Patrick J, Li M. High accuracy information extraction of medication information from clinical notes: 2009 i2b2 medication extraction challenge. J Am Med Inform Assoc 2010 Oct;17(5):524-527 [FREE Full text] [doi: 10.1136/jamia.2010.003939] [Medline: 20819856]
14. Uzuner O, South BR, Shen S, DuVall SL. 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. J Am Med Inform Assoc 2011;18(5):552-556 [FREE Full text] [doi: 10.1136/amiajnl-2011-000203] [Medline: 21685143]
15. Xu Y, Liu J, Wu J, Wang Y, Tu Z, Sun J, et al. A classification approach to coreference in discharge summaries: 2011 i2b2 challenge. J Am Med Inform Assoc 2012;19(5):897-905 [FREE Full text] [doi: 10.1136/amiajnl-2011-000734] [Medline: 22505762]
16. Sun W, Rumshisky A, Uzuner O. Evaluating temporal relations in clinical text: 2012 i2b2 Challenge. J Am Med Inform Assoc 2013;20(5):806-813 [FREE Full text] [doi: 10.1136/amiajnl-2013-001628] [Medline: 23564629]
17. Jagannathan V, Mullett CJ, Arbogast JG, Halbritter KA, Yellapragada D, Regulapati S, et al. Assessment of commercial NLP engines for medication information extraction from dictated clinical notes. Int J Med Inform 2009 Apr;78(4):284-291. [doi: 10.1016/j.ijmedinf.2008.08.006] [Medline: 18838293]
18. Xu H, Stenner SP, Doan S, Johnson KB, Waitman LR, Denny JC. MedEx: a medication information extraction system for clinical narratives. J Am Med Inform Assoc 2010;17(1):19-24 [FREE Full text] [doi: 10.1197/jamia.M3378] [Medline: 20064797]
19. Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system. BMC Med Inform Decis Mak 2006;6:30 [FREE Full text] [doi: 10.1186/1472-6947-6-30] [Medline: 16872495]
20. Lei J, Tang B, Lu X, Gao K, Jiang M, Xu H. A comprehensive study of named entity recognition in Chinese clinical text. J Am Med Inform Assoc 2014;21(5):808-814 [FREE Full text] [doi: 10.1136/amiajnl-2013-002381] [Medline: 24347408]
21. Jiang M, Chen Y, Liu M, Rosenbloom ST, Mani S, Denny JC, et al. A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries. J Am Med Inform Assoc 2011;18(5):601-606 [FREE Full text] [doi: 10.1136/amiajnl-2011-000163] [Medline: 21508414]
22. Rink B, Harabagiu S, Roberts K. Automatic extraction of relations between medical concepts in clinical texts. J Am Med Inform Assoc 2011;18(5):594-600 [FREE Full text] [doi: 10.1136/amiajnl-2011-000153] [Medline: 21846787]
23. Ashish N, Dahm L, Boicey C. University of California, Irvine-Pathology Extraction Pipeline: the pathology extraction pipeline for information extraction from pathology reports. Health Informatics J 2014 Dec;20(4):288-305. [doi: 10.1177/1460458213494032] [Medline: 25155030]

24. Chen Y, Lu H, Li L. Automatic ICD-10 coding algorithm using an improved longest common subsequence based on semantic similarity. PLoS One 2017;12(3):e0173410 [FREE Full text] [doi: 10.1371/journal.pone.0173410] [Medline: 28306739]
25. Bedard K, Ferrall C. Wage and test score dispersion: some international evidence. Economics of Education Review 2003 Feb;22(1):31-43. [doi: 10.1016/s0272-7757(01)00060-7]
26. Ratcliff R. Methods for dealing with reaction time outliers. Psychol Bull 1993 Nov;114(3):510-532. [Medline: 8272468]
27. Lo S, Andrews S. To transform or not to transform: using generalized linear mixed models to analyse reaction time data. Front Psychol 2015;6:1171 [FREE Full text] [doi: 10.3389/fpsyg.2015.01171] [Medline: 26300841]
28. Keene ON. The log transformation is special. Stat Med 1995 Apr 30;14(8):811-819. [Medline: 7644861]
29. Friedrich JO, Adhikari NKJ, Beyene J. Ratio of geometric means to analyze continuous outcomes in meta-analysis: comparison to mean differences and ratio of arithmetic means using empiric data and simulation. Stat Med 2012 Jul 30;31(17):1857-1886. [doi: 10.1002/sim.4501] [Medline: 22438170]
30. Schwappach DLB, Taxis K, Pfeiffer Y. Oncology nurses' beliefs and attitudes towards the double-check of chemotherapy medications: a cross-sectional survey study. BMC Health Serv Res 2018 Dec 17;18(1):123 [FREE Full text] [doi: 10.1186/s12913-018-2937-9] [Medline: 29454347]
31. Ross LM, Wallace J, Paton JY. Medication errors in a paediatric teaching hospital in the UK: five years operational experience. Arch Dis Child 2000 Dec;83(6):492-497 [FREE Full text] [Medline: 11087283]
32. Montague EN, Kleiner BM, Winchester WW. Empirically understanding trust in medical technology. Int J Industr Ergonomics 2009 Jul;39(4):628-634. [doi: 10.1016/j.ergon.2009.01.004]
33. Parasuraman R, Riley V. Humans and automation: use, misuse, disuse, abuse. Hum Factors 2016 Nov 23;39(2):230-253. [doi: 10.1518/001872097778543886]
34. Krishnankutty B, Bellary S, Kumar NBR, Moodahadu LS. Data management in clinical research: an overview. Indian J Pharmacol 2012 Mar;44(2):168-172 [FREE Full text] [doi: 10.4103/0253-7613.93842] [Medline: 22529469]
35. Sorkin RD, Woods DD. Systems with human monitors: a signal detection analysis. Hum–Comput Interact 2009 Nov 11;1(1):49-75. [doi: 10.1207/s15327051hci0101_2]
36. Dzindolet MT, Peterson SA, Pomranky RA, Pierce LG, Beck HP. The role of trust in automation reliance. Int J Hum-Comput Stud 2003 Jun;58(6):697-718. [doi: 10.1016/S1071-5819(03)00038-7]

Abbreviations

CRA: clinical research associate
CRF: case report form
eCRF: electronic case report form
EDC: electronic data capture
EHR: electronic health record
EMR: electronic medical record
i2b2: informatics for integrating biology and the bedside
IQR: interquartile range
MD: mean difference
NLP: natural language processing
NLP-MIES: natural language processing–driven medical information extraction system
SNOMED-CT: Systematized Nomenclature of Medicine–Clinical Terms
VA: Veterans Affairs

Edited by G Eysenbach; submitted 18.01.19; peer-reviewed by L Cui, J Zheng; comments to author 28.03.19; revised version received 13.05.19; accepted 29.05.19; published 16.07.19.

Please cite as:
Han J, Chen K, Fang L, Zhang S, Wang F, Ma H, Zhao L, Liu S
Improving the Efficacy of the Data Entry Process for Clinical Research With a Natural Language Processing–Driven Medical Information Extraction System: Quantitative Field Research
JMIR Med Inform 2019;7(3):e13331
URL: http://medinform.jmir.org/2019/3/e13331/
doi:10.2196/13331
PMID:

©Jiang Han, Ken Chen, Lei Fang, Shaodian Zhang, Fei Wang, Handong Ma, Liebin Zhao, Shijian Liu. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 16.07.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.