ENHANCING MOBILE SPONTANEOUS ADVERSE DRUG EVENT REPORTING THROUGH ELECTRONIC HEALTH RECORDS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET KUBILAY KAHVECI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING

SEPTEMBER 2015

Approval of the thesis:

ENHANCING MOBILE SPONTANEOUS ADVERSE DRUG EVENT REPORTING THROUGH ELECTRONIC HEALTH RECORDS

submitted by **MEHMET KUBILAY KAHVECI** in partial fulfillment of the requirements for the degree of **Master of Science in Computer Engineering Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Adnan Yazıcı _____
Head of Department, **Computer Engineering**

Assist. Prof. Dr. İsmail Sengör Altıngövde _____
Supervisor, **Computer Engineering Dept., METU**

Prof. Dr. Asuman Doğaç _____
Co-supervisor, **SRDC Ltd.**

Examining Committee Members:

Assist. Prof. Dr. Engin Demir
Computer Engineering Department, UTAA _____

Assist. Prof. Dr. İsmail Sengör Altıngövde
Computer Engineering Department, METU _____

Prof. Dr. İsmail Hakkı Toroslu
Computer Engineering Department, METU _____

Prof. Dr. Ahmet Coşar
Computer Engineering Department, METU _____

Assoc. Prof. Dr. Pınar Karagöz
Computer Engineering Department, METU _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: MEHMET KUBILAY KAHVECI

Signature :

ABSTRACT

ENHANCING MOBILE SPONTANEOUS ADVERSE DRUG EVENT REPORTING THROUGH ELECTRONIC HEALTH RECORDS

Kahveci, Mehmet Kubilay
M.S., Department of Computer Engineering
Supervisor : Assist. Prof. Dr. İsmail Sengör Altıngövde
Co-Supervisor : Prof. Dr. Asuman Doğaç

September 2015, 91 pages

Post marketing surveillance for pharmaceutical drugs has been largely dependent on spontaneous reporting systems (SRSs) for quite some time. Although paper based reporting forms are broadly replaced by digital (online) counterparts, accessibility and usability of those systems still pose problems. Considering the fact that adverse drug event (ADE) reporting is mostly a voluntary action and it takes a lot of effort to complete an ADE report on current systems, outputs are usually in low quality and quantity.

On the other hand, individual case safety reports (ICSRs), generated by SRSs, contain contextual information such as patient's active medications, past medical history or past drug therapies, most of which is already available in patient's electronic health records (EHRs). Therefore, seamlessly accessing EHR sources to pre-fill ICSR forms would be a major improvement for spontaneous reporting process.

There have already been studies aiming to utilize EHR data for post market surveillance. However, rather than focusing on facilitating the reporting process, they target automated detection of adverse events and to the best of our knowledge, none of them aims mobile platforms or has usability concerns for the end user.

In this thesis, we address the issue of under-reporting and interoperability of those systems and demonstrate that EHR systems can be exploited in mobile SRSs to generate high quality reports at high rates and provide a better experience to the reporter as well. We have developed a scalable platform, integrable to existing reporting systems and EHR sources of different content models, which can semi-automatically pre-fill ADE reports using medical summary of patient available in EHR systems. The quality of reports produced using our tool and the amount of time spent reporting is a significant improvement compared to existing mediums.

Keywords: Spontaneous Reporting Systems, Post Market Surveillance, Pharmacovigilance, Electronic Health Records, Interoperability

ÖZ

MOBİL ADVERS ETKİ BİLDİRİM SİSTEMLERİNİN ELEKTRONİK SAĞLIK KAYITLARI YARDIMIYLA İYİLEŞTİRİLMESİ

Kahveci, Mehmet Kubilay
Yüksek Lisans, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi : Yrd. Doç. Dr. İsmail Sengör Altıngövde
Ortak Tez Yöneticisi : Prof. Dr. Asuman Doğaç

Eylül 2015 , 91 sayfa

İlaçların pazar sonrası denetimi çoğunlukla advers etki raporlama sistemlerine bağımlıdır. Bu sistemlerin dijital ve çevrim içi olarak sunulan yeni sürümleri, kağıt esaslı formlar üzerinden işleyen raporlama faaliyetlerinin yerini almış olsa da sistemlerin ulaşılabilirliği ve kullanılabilirliği hala sorun olabilmektedir. Advers etki raporlamanın çoğunlukla gönüllü olarak yapılan bir iş olduğu düşünüldüğünde ve bir yan etkiyi eksiksiz raporlayabilmek için harcanan efor ve zaman göz önüne alındığında, oluşturulan raporların nicelik ve nitelik olarak neden düşük seviyede kaldığı anlaşılabilir.

Halbuki, advers etki raporları, hastanın aktif olarak kullandığı ilaçlar, hastalık geçmişi ya da tamamlanmış ilaç tedavileri gibi halihazırda elektronik sağlık kayıtlarında (ESK) bulunabilecek birçok bilgiyi içermektedir. Bu ESK kaynaklarına sorunsuzca ulaşmak ve saptanan hasta verileriyle bireysel olgu güvenlilik raporlarını (BOGR) otomatik bir şekilde doldurmak, raporlama sürecine büyük

bir iyileştirme sağlayabilir.

Pazar sonrası denetim çalışmaları için ESK verilerinden faydalanmayı hedefleyen çalışmalar olmuştur. Fakat bu çalışmalar, genelde, raporlama sürecini kolaylaştırmak yerine ESK verilerini işleyerek advers etkileri otomatik tespit etmeye yöneliktir. Ayrıca, araştırmamız çerçevesinde gördüğümüz kadarıyla mobil platformları hedefleyen ve son kullanıcı için kullanılabilirlik kaygısı taşıyan bir çalışma olmamıştır.

Bu tez çalışmasında, advers etkilerin eksik raporlanması ve sistemlerin birlikte işlerlik sorunlarını ele alıyoruz. ESK'lardan faydalanarak mobil advers etki raporlama sistemlerinden yüksek nicelikte ve içerik olarak kaliteli raporlar üretirken kullanıcıya da iyi bir deneyim sağlanabileceğini gösteriyoruz. Bu amaçla, mevcut raporlama sitemlerine ve farklı içerik modelleri kullanan ESK kaynaklarına entegre edilebilir, esnek ve genişletilebilir bir sistem geliştirdik. Sistem, ESK kaynaklarında bulunan hasta geçmişini kullanarak advers etki bildirim raporlarını yarı-otomatik şekilde doldurabilmektedir. Üretilen raporların içeriklerinin zenginliği ve rapor hazırlama sürecinin çok daha basit ve zaman almayan bir şekilde tamamlanması var olan sistemlere kıyasla kayda değer bir iyileştirmedir.

Anahtar Kelimeler: Advers Etki Bildirim, Pazar Sonrası Denetim, Farmakovijilans, Elektronik Sağlık Kayıtları, Birlikte İşlerlik

To my family and

to the loving memory of my aunt,

Fikriye Kılıç...

ACKNOWLEDGMENTS

First and foremost, enormous gratitude is due to Prof. Dr. Asuman Doğaç for her encouragement and support throughout this study. I would also like to express my sincere appreciation to my supervisor Assist. Prof. Dr. İsmail Sengör Altıngövde for his support and guidance. To my committee, I am extremely grateful for your assistance and insightful comments.

I appreciate Mustafa Yüksel, Gökçe Banu Laleci Ertürkmen and Ali Anıl Sınacı for their invaluable advice and support. Without their guidance, this thesis would not have come into being. All colleagues at SRDC, I thank you for giving me the privilege and joy of working with you, and for all you have taught me.

Thank you to Gülfem Demir. There are no words that can express my gratitude and appreciation for all you have done and been for me. Without your encouragement, I would have never had the strength to complete this work.

Many thanks to my friends Arın Öztürk, Burak Soğukçam, Cem Pekdoğru, Anıl Paçacı, Bengi Demir, Oğuz Kartal, Işınsu Katırcıoğlu, Mert Gençtürk, Yasemin Kapucu, Çağıl Kirezci, Ceren Emre, Umut Ağıl, Berk Gürçay, İnan Özdemir, and Orkun Çolakoğlu. Their cheerful presence definitely made a big difference to my life.

Last but surely not least, I am deeply grateful to my family for their love, care, and faithful support over all these years.

The research leading to these results has received funding from the Innovative Medicines Initiative Joint Undertaking (IMI JU) under grant agreement n° 115632, resources of which are composed of financial contribution from the European Commission's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ADE	Adverse Drug Event
ADR	Adverse Drug Reaction
ANS	Adverse Drug Event Notification System
ASTER	Adverse Drug Event Spontaneous Triggered Event Reporting
ASTM	American Society for Testing and Materials
ATC	Anatomical Therapeutic Chemical Classification System
C-CDA	Consolidated Clinical Document Architecture
CCD	Continuity of Care Document
CCR	Continuity of Care Record
CDA	Clinical Document Architecture
CIP	Competitiveness and Innovation Programme
CPT	Current Procedural Terminology
E2B	Message Standard for Electronic Transmission of Individual Case Safety Reports
EC	European Commission
EDQM	European Directorate for the Quality of Medicines and Health-Care
EHR	Electronic Health Record
epSOS	Smart Open Services for European Patients
EU	European Union
FDA	Food and Drug Administration
GP	General Practitioner
HALMED	Agency for Medicinal Products and Medical Devices
HIPAA	Health Insurance Portability and Accountability Act
HL7	Health Level Seven
ICD	International Statistical Classification of Diseases and Related Health Problems
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use

ICSR	Individual Case Safety Report
ICT	Information and Communications Technology
IHE	Integrating the Healthcare Enterprise
iOS	iPhone Operating System
ISCO	International Standard Classification of Occupations
ISO	International Organization for Standardization
IT	Information Technologies
LOINC	Logical Observation Identifiers Names and Codes
MedDRA	Medical Dictionary for Regulatory Activities
MHRA	Medicines and Healthcare Products Regulatory Agency
NHS	National Health Service
NPfIT	National Programme for Information Technologies
OID	Object Identifier
PCC	Patient Care Coordination
PMS	Post Marketing Surveillance
PS	Patient Summary
RIM	Reference Information Model
SALUS	Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies
SCR	Summary Care Record
SNOMED	Systematized Nomenclature of Medicine
SNOMED CT	Systematized Nomenclature of Medicine - Clinical Terms
SRDC	Software Research and Development and Consultancy Ltd.
SRS	Spontaneous Reporting System
TGA	Therapeutic Goods Administration
UCUM	The Unified Code for Units of Measure
UK	United Kingdom
UMC	Uppsala Monitoring Centre
URL	Uniform Resource Locator
US	United States
WEB-RADR	Recognising Adverse Drug Reactions
WHO	World Health Organization
XML	eXtensible Markup Language
XSD	eXtensible Markup Language Schema Definition

CHAPTER 1

INTRODUCTION

1.1 Motivation

Adverse drug reactions (ADRs) are defined as unwanted, uncomfortable or dangerous effects that a drug may have. Adverse drug event (ADE), on the other hand, is an undesirable condition occurred after exposure to a drug, which is not necessarily a direct cause of the drug. ADEs are accounted for 5% of all hospital admissions, occur 10-20% of hospital inpatients and most importantly cause deaths in 0.1% of all medical inpatients [65]. Putting all those aside, ADEs may negatively affect patient's quality of life. As they mimic diseases, they cause unnecessary investigations and considerable increase in patient care cost as a result [19, 44, 65].

By the time a drug hit the market (i.e. licensed), it is exposed to 1500 patients. Whereas, in order to diagnose at least one patient with adverse drug reaction, it is needed to investigate 30,000 patients who are on this particular drug [34]. This dramatic gap puts the emphasis on post marketing surveillance (PMS) which is monitoring of pharmaceutical drugs for efficacy and safety after they have been released on the market.

Spontaneous reporting systems (SRSs) were developed as a cost-effective solution for monitoring drug safety and had success in early detection of signals of new and rare adverse drug events [75]. Thus, spontaneous reporting has become one of the most fundamental methods for post market surveillance of pharmaceutical drugs. Catching up with technology and considering the paperwork

burden, SRSs have started operating online in most of the developed and many developing countries; even have become compulsory in Sweden, France and Italy [56, 72, 76].

As social networks and mobile applications have been established as irreplaceable mediums for sharing information, ideas and experiences in real-time; many fields are progressing towards a mobile-driven, more interactive era, while eHealth studies are having a hard time to keep up with. Considering the sensitivity of healthcare data and the requirement for healthcare systems to be reliable and unerring in any circumstances, transformation pace is reasonable. There are already some initiatives to exploit social networks and mobile technology for pharmacovigilance purposes, such as WEB-RADR project. It aims at i) monitoring social networks for probable safety signals about pharmaceutical drug; ii) enabling healthcare professionals and patients to submit ADE reports via mobile devices [69]. Mining unstructured data that is publicly available on social networks and identifying information of interest for medicinal product surveillance is a great step towards earlier detection of potential ADEs. When existing attempts are proven and become robust solutions, they have the potential to be the game changer that redefines post market drug safety surveillance. Until then, spontaneous reporting systems would still be the go-to tool.

However, although online spontaneous reporting systems are established in many sites, they are not effectively utilized due to several problems; under-reporting of adverse drug events and qualitatively poor content of reports being the most prominent ones [17, 26, 34, 35]. Reported ADEs are estimated to be 5% of all events occurred [19].

There are multiple underlying causes that together create these problems, and eventually, unfruitful environments and systems. Firstly, filling adverse drug event reports require considerable amount of time, and healthcare professionals tend to deprioritize reporting due to the usual workload and lack of time [74]. This results in numerous side effects of drugs in the market reported at a low rate.

Another issue is that since spontaneous reporting systems do not benefit from

patient context information already available in various electronic health record (EHR) systems, not only the process of reporting is troublesome but also the quality of data collected through spontaneous reporting is very low.

Last but certainly not least, medical personnel and patients are not encouraged by regulatory authorities to report ADEs; it is quite often a voluntary action [74].

As a result, there is a clear need for facilitator tools to enhance drug safety surveillance and to boost pharmacovigilance activities. In these terms, increasing accessibility to spontaneous reporting tools with the help of mobile devices is promising and could definitely be exploited for two majors problems of spontaneous reporting: in-comprehensive report content and low reporting rates.

Individual case safety reports (ICSRs) define the data structure used in ADE reporting. They contain some contextual information such as patient's demographics, active medications, past medical history or past drug therapies in addition to basic information about the suspected drug and adverse event [40]. However, it takes significant effort to accurately complete all these information manually, both by health professionals and patients. On the other hand, most of this information about patient has been already collected as a part of routine care and available in EHR systems. Accessing EHR sources seamlessly to populate ICSRs using health records would be a major improvement for spontaneous reporting process, specifically addressing two problems aforementioned.

In this thesis, we propose an information and communications technology (ICT) toolkit which is adaptable to existing SRSs and EHR systems, to improve ADE reporting both quantitatively and qualitatively.

1.2 Challenges and Research Goals

Accessibility and usability of SRSs has long been an open issue attracting attention of many studies. There have also been efforts to facilitate secondary use of EHR in ADE detection [21, 30, 55]. Common approach is to detect possible

events or signals automatically using EHRs with the help of data mining and learning algorithms applied on top of relevant health records.

Although there exist well performing methods, they do not focus on accelareting the reporting process after detection, and to the best of our knowledge, none of them aims mobile platforms or has usability concerns for the end user. Hence, our work makes a difference in two points:

i. enhancing ADE reporting on mobile devices,

ii. improving usability of reporting systems rather than replacing them with computer based automated tools.

Our objective in this thesis to provide seamless access to EHR systems and reuse available patient data for making ADE report filling process easier thorough a mobile application. By automatically extracting relevant patient data from EHRs to pre-fill ADE reports, we aim to:

i. speed up the spontaneous reporting process,

ii. provide a better user experience to the form filler (can either be a healthcare professional, or patient herself),

iii. enrich the generated report content.

However, there are barriers, in both clinical research (e.g. post market surveillance) and clinical care (e.g. healthcare) domains, standing in the way of such an improvement.

First and foremost, EHR data is heterogenous. EHR sources differ to a great extent in terms of data format and data exchange transactions. Regional and national authorities define their own models for medical data to be exchanged within their borders and most of the time, try to align these definitions with international content and exchange standards. For example, Meaningful Use program introduced by the United States (US) government designated Consolidated CDA (C-CDA) as the standard for patient information exchange among

certified EHR systems [24]. As another national effort, the National Health Service (NHS) in the United Kingdom (UK) has developed Summary Care Record (SCR) which contains key information from an individual's primary care records [33]. There are some international patient data exchange efforts as well, such as epSOS Patient Summary (PS) which enables electronically exchanging health records across borders in Europe [10].

Similarly, data models in the pharmacovigilance domain do not conform to a common standard. ICH E2B is the international standard as the ICSR format in clinical research, whereas EHR systems have completely different representation of data. Hence, even though clinical care data could have been unified into a single standard, it will still not be interoperable with clinical research domain.

Lastly, medical terminology systems used in these content models are also quite disparate. For instance, ICH E2B uses MedDRA terminology to represent structured medical data, while EHR systems most widely use SNOMED CT, LOINC and ICD code systems; or ICH E2B codes units of measures with its own value set while UCUM terminology is adopted in EHR sources [45].

1.3 Contributions

In this thesis, we have developed an ICT toolkit that utilizes secondary use of EHRs for the purpose of providing high quality adverse drug event reports in a less time consuming, more user friendly manner. Our tools are built to be integrated with existing SRSs and EHR sources with minimum effort. To provide such an extendable and scalable architecture, we address barriers identified above by:

- designating a unified data model that completely covers required elements for individual case safety reports while also considering structure of EHR sources,
- employing a terminology server (benefiting from existing studies) for automatic conversion of terminology systems,

- building a tool that semi-automatically pre-fills ADE reports by locating relevant data in EHRs and convert terminologies accordingly,
- providing simple interfaces for integrating new source/input (EHR) and target/output (ICSR) data models.

We have also implemented an adapter on iOS platform, as a proof of concept, that is able communicate with our platform and simulated EHR sources via Web services and displays pre-filled ADE reports to demonstrate how a user can effortlessly report adverse drug events on mobile devices.

1.4 Publications

The work carried out for the completion of this thesis has resulted in a publication in eChallenges 2015 Conference.

1. Kubilay Kahveci, Mustafa Yuksel, Gokce Banu Laleci Erturkmen, "Enhancing Mobile Spontaneous Adverse Drug Event Reporting through Electronic Health Records", *eChallenges e-2015 Conference*, IEEE, November 2015. Accepted for publication.

1.5 Structure of the Thesis

The thesis is organized as follows:

- Chapter 1 presents the motivation behind this research, challenges, research goals and publications related to the thesis.
- Chapter 2 introduces data standards and profiles and enabling technologies of our work by giving a brief background information.
- Chapter 3 presents content model developed in the scope of the thesis. In particular, we explain how disparate models are mapped and become interoperable.

- Chapter 4 presents overall architecture of the system along with design and implementation details. Communication and transactions between internal and external components are also discussed.
- Chapter 5 demonstrates the workflow with a simple scenario.
- Chapter 6 presents conducted test results with brief explanations.
- Chapter 7 outlines the related studies on the topics of interest for this thesis.
- Chapter 8 concludes the thesis by giving final remarks, discussing issues about EHRs and SRSs, and suggesting possible directions for future researches in this field.
- Appendix A contains the list of data elements defined as the content model of this system.
- Appendix B contains an individual case safety report instance generated using our toolkit.
- Appendix C contains semantic definition of coded values for representing dose units in E2B (R2) profile.

CHAPTER 2

BACKGROUND ON ENABLING TECHNOLOGIES

Many technologies, standards and previous researches helped shaping our work and making it possible. In this section, background information about these technologies is presented. For the motivation behind depending on them and application specific details, please see Chapter 3 and 4.

2.1 ICH E2B (R2)

The International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) has started as a project with the aim of reducing the waste of resources in the development of new medicines and maintaining safeguards on efficacy, quality and safety [5]. ICH brings regulatory authorities and pharmaceutical industry together to develop guidelines about drug registration, which have been adopted as law in several countries [28].

ICH E2B (R2) is a standard that defines a data model for ICSRs and a protocol for electronic transmission of reports. Content model is composed of two main sections, see Figure 2.1 for a complete relational view of E2B data elements. Section A includes administrative information, primary sources of report, responsible parties in the transmission of the report, and case identifier. All details of reported case together with relevant patient context, which is most probably available in EHR sources, lie in Section B [45]. To be more specific, subsections that may already be available in patient's electronic health records are as follow:

- **Personal demographics:** Patient demographic details such as name, address, date of birth and medical record number.
- **Relevant medical history and concurrent conditions:** History of allergies, procedures, encounters and present conditions.
- **Relevant past drug history:** Past and active medications and history of immunizations, which are relevant to the reported event.
- **Measurements, vitals, tests:** Body weight, height and results of tests and procedures relevant to the investigation of the patient.
- **Reactions/events:** Details of observed reaction or event, such as description, start date, duration, and outcome.
- **Drug information:** Identification of the suspected drug(s), dosage information, pharmaceutical form, route of administration, etc.
- **Parental data:** Medical history of parents similar to the patient's in terms of structure.

ICH E2B (R2) is adopted by the Uppsala Monitoring Centre (UMC) as the Word Health Organization (WHO) Collaborating Centre for International Drug Monitoring and many regulatory authorities at regional and national level, e.g. Medicines & Healthcare Products Regulatory Agency (MHRA) in the UK [59], Agency for Medicinal Products and Medical Devices (HALMED) in Croatia [15], Therapeutic Goods Administration (TGA) in Australia [71], and Food and Drug Administration (FDA) in the US [31].

2.2 EHR Content Models

This section provides information about content models of EHR sources used in the thesis. To provide a complete understanding of standardization in electronic health records, base standards and extendable template architectures are introduced beforehand. Three main EHR content models used in our work, epSOS Patient Summary, HL7 Consolidated Clinical Document Architecture and

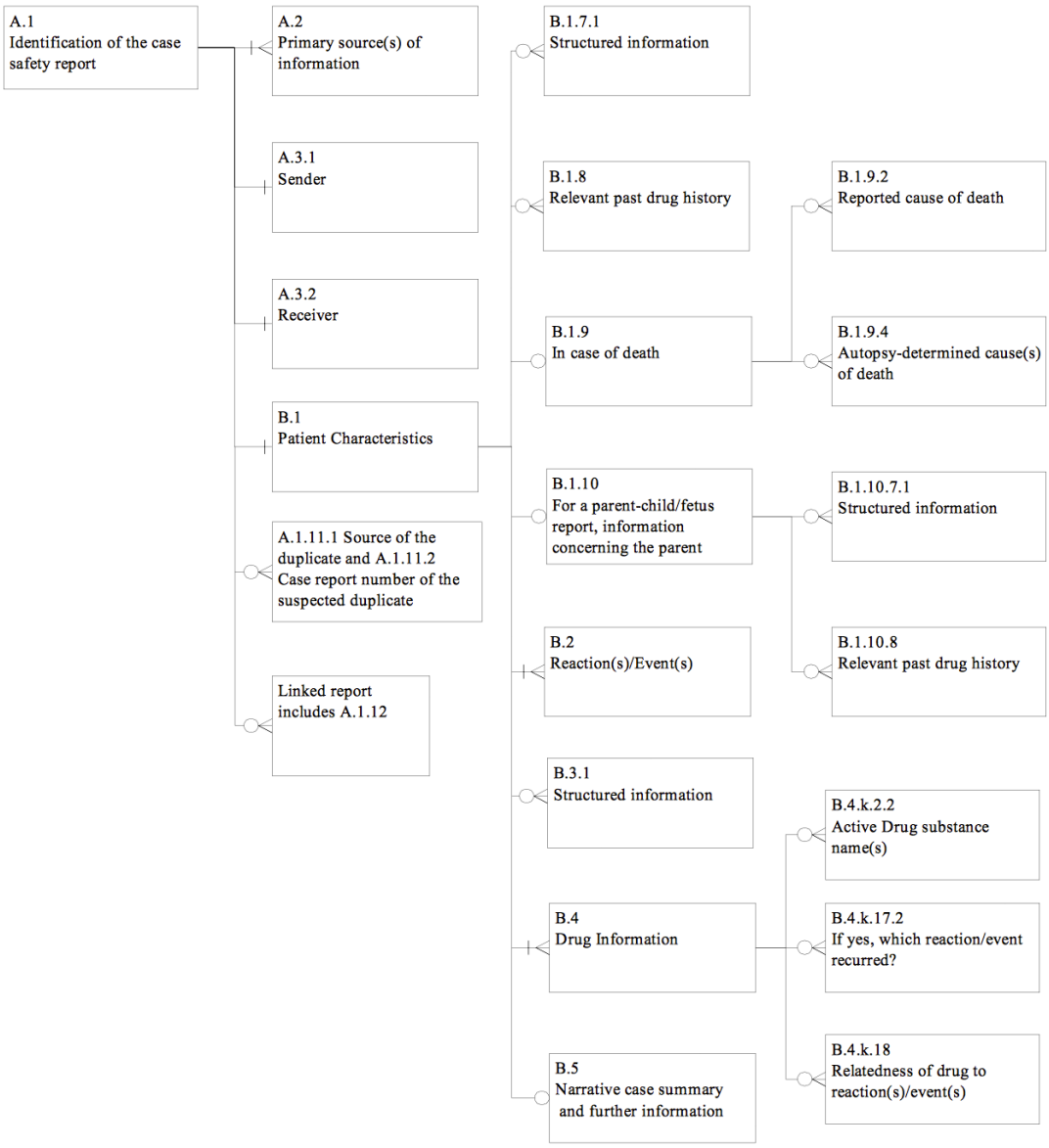

Figure 2.1: Relational view of E2B (R2) data elements

Summary Care Record are presented in Sections 2.2.4, 2.2.5 and 2.2.6, respectively. Former sections explain standards and infrastructure that enables the development of such content models for EHRs.

2.2.1 HL7 Clinical Document Architecture

Health Level Seven (HL7) is a not-for-profit organization with a goal of providing standards on electronic health information for exchange between health services

[3]. Clinical Document Architecture (CDA) is a document markup standard, developed by HL7, which defines structure and semantics of a clinical document which can contain various clinical content such as admission, pathology report, physical, and more [37]. CDA enables the exchange of clinical documents between parties involved in the patient care (i.e. healthcare providers, patients, vendors, EHR systems), promotes longevity of the information, and allows cost effective implementation across systems.

```
<ClinicalDocument>
    <!-- CDA Header -->
    ...
    <structuredBody>
    <section>
        <!-- Narrative Block -->
        <text>...</text>
        <!-- CDA Entries -->
        <observation>...</observation>
        <encounter>...</encounter>
        <procedure>...</procedure>
    </section>
    <section>...</section>
    </structuredBody>
</ClinicalDocument>
```

Figure 2.2: An example CDA excerpt

CDA documents are encoded in Extensible Markup Language (XML) and has machine processable meaning thanks to HL7 Reference Information Model (RIM) [41]. A CDA document consists of two major parts, namely header and body, wrapped by a root level `<ClinicalDocument>` element. The header contains identifier and classifier of the document, involved providers, and authentication information. The body, on the other hand, provides information about the clinical report. This can either be unstructured text, or structured block in a `<structuredBody>` element conforming to CDA XML Schema Definition (XSD). Structured body is divided into nested sections, which can contain a single narrative block and any number of CDA entries. In a section, narrative block contains human readable content, whereas CDA entries represent structured content to be processed by computer. Figure 2.2 shows the outline of a multi-section CDA

Table 2.1: Levels in CDA

Level	Description
CDA Level One	Document body may be human readable content (i.e. narrative) only. At this level, machine processable entries are not included.
CDA Level Two	Document body includes machine processable semantics on section level.
CDA Level Three	Document body has complete machine processable content on entry level.

document.

CDA has a notion of *levels* which is the basis of a hierarchical architecture. Table 2.1 presents three levels of document definition.

CDA, itself, draws up a generic and flexible specification which can be restricted using document-level, section-level and entry-level templates. Though, there is no requirement that CDA must be restricted.

2.2.2 HL7/ASTM Continuity of Care Document

Continuity of Care Document (CCD) is a collaborative effort between HL7 and American Society for Testing and Materials (ASTM), which describes constraints on the HL7 CDA to meet requirements of ASTM Continuity of Care Record (CCR) [43].

CCR is basically developed as a core data set of relevant demographic, administrative and clinical information about a patient's healthcare [16]. It allows healthcare professionals or systems to aggregate patient data and forward it to another professional or system. This ensures the continuity of care.

Similar to what CDA does, CCR also defines a content model as an XML schema for exchange purpose. Considering the fact that CDA is a generic structure that

can be constrained using templates, CCR can also be implemented in CDA, which is exactly where CCD comes into play. It establishes a set of templates on CDA so as to meet CCR requirements in CDA structure. Since CDA and its templates are universally accepted, CCD is widely preferred over CCR's own schema.

```
<section>
    <!-- Procedures section template -->
    <templateId root="2.16.840.1.113883.10.20.1.12"/>
    <code code="47519-4" codeSystem="2.16.840.1.113883.6.1"/>
    <title>Procedures</title>
    <!-- Narrative Block -->
    <text>...</text>
    <entry>
        <procedure classCode="PROC">
            <!-- Procedure activity template -->
            <templateId root="2.16.840.1.113883.10.20.1.29"/>
            <id root="e401f340-7be2-11db-9fe1-0800200c9a66"/>
            <code code="52734007" codeSystem="2.16.840.1.113883.6.96"
             ↪  displayName="Total hip replacement"/>
            <statusCode code="completed"/>
            <effectiveTime value="1998"/>
        </procedure>
    </entry>
</section>
```

Figure 2.3: A section from CCD templates

Regarding levels of CDA, CCD defines a document template (Level One), multiple section templates (Level Two), and entry level (Level Three) templates as well. According to those templates, CDA documents are required to have 16 sections containing core patient specific data based on common clinical conventions: Alerts, Encounters, Family History, Immunizations, Medical Equipment, Medications, Players, Plan of Care, Problem, Procedures, Purpose, Results, Social History, Vital Signs. Figure 2.3 shows an excerpt from Procedures section in a CDA document based on CCD template.

2.2.3 IHE Patient Care Coordination Templates

Integrating the Healthcare Enterprise (IHE) is a non-profit initiative by healthcare professionals and healthcare industry to improve interoperability of healthcare systems and to enable care providers to use and share information more effectively [4]. IHE promotes coordinated use of established standards such as HL7 CDA for specific clinical needs. With this motivation, they develop content templates under Patient Care Coordination (PCC) domain, similar to Continuity of Care Document (see Section 2.2.2).

Unlike CCD, PCC provides multiple templates on different levels rather than providing a single document template. As of August 2015, PCC has 34 document templates, 10 header templates, 144 section templates and 67 entry templates [46]. Table 2.2 presents some example templates at each level.

In Figure 2.4, a sample from History of Encounters section is presented:

- Since PCC templates aren't built as an alternative to existing solutions but to complement them instead, demonstrated section has two template

Table 2.2: PCC templates

Level	**Examples**
Document	Nursing Note, Maternal Discharge Summary, Emergency Department Referral, Transport Document, Immunization Detail, Care Plan, Labor and Delivery Summary, etc.
Header	Authorization, Language Communication, Consent Service Events, etc.
Section	Active Problems, Discharge Diagnosis, Social History, Medications, General Appearance, etc.
Entry	Pain Score Observation, Supply Entry, Blood Type Observation, Patient Transfer, etc.

```
<section>
    <!-- History of encounters section template -->
    <templateId root="2.16.840.1.113883.10.20.1.3"/>
    <templateId root="1.3.6.1.4.1.19376.1.5.3.1.1.5.3.3"/>
    <id root="0437bf50-7328-4e5c-9de4-3cb9d26b5a76"/>
    <code code="46240-8" displayName="History of encounters"
     ↪  codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC"/>
    <!-- Narrative Block -->
    <text>...</text>
    <entry>
        <encounter classCode="ENC" moodCode="EVN">
            <templateId root="1.3.6.1.4.1.19376.1.5.3.1.4.14"/>
            <templateId root="2.16.840.1.113883.10.20.1.21"/>
            ...
            <entryRelationship typeCode="RSON">
                <observation classCode="OBS" moodCode="EVN">
                    <templateId root="1.3.6.1.4.1.19376.1.5.3.1.4.5"/>
                    <templateId root="2.16.840.1.113883.10.20.1.28"/>
                    ...
                    <code code="282291009" displayName="Diagnosis"
                     ↪  codeSystem="2.16.840.1.113883.6.96"
                     ↪  codeSystemName="SNOMED CT"/>
                    <effectiveTime>
                        <low value="20090331"/>
                    </effectiveTime>
                    <value xsi:type="CD" code="578.9" displayName="Hemorrhage
                     ↪  of gastrointestinal tract, unspecified"
                     ↪  codeSystem="2.16.840.1.113883.6.2"
                     ↪  codeSystemName="ICD-9-CM"/>
                </observation>
            </entryRelationship>
        </encounter>
    </entry>
</section>
```

Figure 2.4: A section from PCC templates

identifiers (i.e. `<templateId>`): `2.16.840.1.113883.10.20.1.3` represents Encounters section in CCD template whereas `1.3.6.1.4.1.19376.1.5.3.1.1.5.3.3` is to identify Encounter Histories section in PCC templates. Thanks to this approach, PCC templates can be applied on HL7 CDA documents without breaking conformance to CCD templates.

- According to related PCC templates, section meanings are represented using codes in various code systems. Encounter Histories section have a predefined value of `46240-8` in LOINC terminology system, which represents *History of encounters*. Although human readable name for the terminology system is also provided, it is, in-fact, identified by universal object identifiers (OID), which is `2.16.840.1.113883.6.1` for LOINC terminology system.

- At the entry-level, encounter templates are used in this section. Similar to section templating, this entry conforms to both PCC specifications (via `1.3.6.1.4.1.19376.1.5.3.1.4.14`) and CCD specifications (via `2.16.840.1.113883.10.20.1.25`).

- As the element represents an encounter in patient's history, `classCode` shall be `ENC`. The `moodCode` on the other hand may have different values:

  `PRMS`: a scheduled appointment

  `ARQ`: request for an appointment that hasn't been yet scheduled

  `EVN`: event already occured

- Attached entry relationship with a `typeCode` of `RSON` indicates the reason of this encounter and includes a problem entry based on observation templates of PCC and CCD.

- The problem is classified as diagnosis using SNOMED CT terminology system. Other possible problem codes in SNOMED CT is listed in Table 2.3.

- The crucial information about this observation lies in `<value>` element, which is coded in a medical terminology system in this case. The element is not required to have coded values, though. It may include free text descriptions in a similar manner to narrative blocks at section-level. However, coded values are required for satisfying interoperability of healthcare systems.

- Entry-level templates of PCC, do not put a restriction on the terminology system to be used for representing the encounter. Depending on the im-

plementation, different terminology system may be used, which actually is the case in epSOS Patient Summary in Section 2.2.4. For details about available terminology systems suited for such use, please refer to Section 2.3.1.

Table 2.3: Problem codes in SNOMED CT

Code	Description
64572001	Condition
418799008	Symptom
404684003	Finding
409586006	Complaint
248536006	Functional limitation
55607006	Problem
282291009	Diagnosis

2.2.4 epSOS Patient Summary

Smart Open Services for European Patients (epSOS) project, co-funded by the European Commission Competitiveness and Innovation Programme (CIP), was an attempt to design and build a service infrastructure for seamless exchange of health records in Europe. With partners from 25 different European countries, the epSOS project team enabled cross-border interoperability between EHR systems across the continent. It has achieved this through the development of strategies, priorities, recommendations and guidelines designed to deliver eHealth in Europe in a coordinated way [10].

The epSOS project defined a Patient Summary (PS) template as EHR content model for exchange of clinical information. Thanks to pan-European, large scale PS implementation at pilot sites, for the first time, European citizens had the opportunity to use cross-border healthcare services. Several European Union (EU) countries have committed to continue epSOS services beyond the project duration. In November 2013, European Commission (EC) recognized efforts of

epSOS project and published European Patient Summary Guidelines to define patient summary dataset for cross-border electronic exchange of medical records, which is based totally on the epSOS PS content [27].

PS content model is based on HL7 CDA and content templates by HL7/ASTM CCD and IHE PCC, which enables EHR systems to be standards-based and interoperable at national level. Regarding the coding of values (e.g. problems, allergies, medications, country and languages, etc.), which PCC templates do not restrict, PS determines multiple terminology systems presented in Table 2.4.

As mentioned in Section 2.2.1, HL7 CDA provides standardization at the highest level of shared information, and supports extensibility via local definitions when corresponding representation is not available in CDA specification [38]. IHE states that the accepted practice for such implementations is extensions [47], to which epSOS PS conforms. It includes extensions in `epsos` namespace. In order to be compliant with the PCC template, low level template specifications are also met.

For a better understanding of CDA extensions, representation of medication strength in epSOS PS would be a good fit as an example. Ratio of the active ingredient(s) to a unit of medication represents the strength of the medication.

Table 2.4: Summary of main terminologies used in epSOS PS

Area of use	**Terminology system**
Field labels	LOINC
Problem list	ICD-10 (3 digit codes)
Medication list	ATC, EDQM, UCUM
Allergies	SNOMED
Surgical procedures	SNOMED
Medical devices	SNOMED
Country and languages	ISO
Professional role	ISCO

```
<epsos:ingredient classCode="ACTI">
    <!-- Strength 1% w/v -->
    <epsos:quantity>
        <epsos:numerator xsi:type="epsos:PQ" value="10" unit="mg"/>
        <epsos:denominator xsi:type="epsos:PQ" value="1" unit="ml"/>
    </epsos:quantity>
    <epsos:ingredient classCode="MMAT" determinerCode="KIND">
        <epsos:code code="S01AA01"
        ↪  codeSystem="1.3.6.1.4.1.12559.11.10.1.3.1.44.1"
        ↪  displayName="chloramphenicol"/>
        <epsos:name>Chloramphenicol</epsos:name>
    </epsos:ingredient>
</epsos:ingredient>
```

Figure 2.5: Ingredient extension in epSOS PS

However, in the case of medication coming in a variety of strength, CDA architecture does not allow expressing it separately. An example is eye-drops, where the medication is in a solution of particular strength and dose quantity is some number of drops. epSOS PS provides this information via an extension, presented in Figure 2.5.

- `epsos:ingredient` represents the active ingredient of the medication.
- `epsos:quantity` element represents the strength of the active ingredient as the ratio of the active ingredient to a unit of medication.
- `epsos:quantity` element contains the numerator and denominator of the strength ratio.

Our motivation for selecting epSOS PS as one of the EHR content models of our system is elaborated further in Chapter 3.

2.2.5 HL7 Consolidated Clinical Document Architecture

Implementing CDA documents conforming to various template specifications such as HL7/ASTM CCD and IHE PCC is a somewhat tedious and complicated

task. Arising from this need, HL7 introduced Consolidated Clinical Document Architecture (C-CDA) to organize all documentation in one place as the single source of truth. It harmonizes the original CCD by HL7, part of the IHE PCC and some other templates (clinical forms) to structure the document [39]. Compared to CDA, C-CDA introduces only slight changes in terms of content. However, it stands as the single guide to make it easier to analyze and implement.

The templates in C-CDA standard are, also, defined at three levels: document, section and entry. CCD, which is of our interest since it contains fundamental information as medical summary of patient, is an example template for defining the type of CDA document. Meaningful Use program in the US –which defines minimum standards for using and exchanging EHRs between healthcare providers, insurers and patients– has designated a patient summary dataset based on C-CDA CCD template [24]. Government-certified EHR technology in the US must conform to C-CDA CCD specifications to achieve meaningful use [25].

2.2.6 Summary Care Record

Summary Care Record (SCR) is another national effort for standardizing patient summary information, which is introduced by the National Health Service (NHS) in the UK. In the course of National Programme for IT (NPfIT), a central database adopting SCR as the content model is constructed to make patient data available anywhere. SCR is limited to hold only the essential information needed in an emergency, e.g. medications, allergies, adverse reactions, past procedures [68].

SCR documents are based on CDA and has document-level GP Summary template to be used patient's medical summary. At section-level and entry-level, unlike epSOS PS and C-CDA CCD, it does not conform to CCD or PCC templates, rather introduces its own [63].

As presented in Figure 2.6, `<section>` elements in the general CDA structure are replaced with nested `<pertinentInformation2>` and `<pertinentCREType>` elements.

```
<pertinentInformation2 typeCode="PERT" contextConductionInd="true"
 ↪  inversionInd="false" negationInd="false">
    <templateId root="2.16.840.1.113883.2.1.3.2.4.18.2"
     ↪  extension="CSAB_RM-NPfITUK10.pertinentInformation1"/>
    <pertinentCREType classCode="CATEGORY" moodCode="EVN">
        <code codeSystem="2.16.840.1.113883.2.1.3.2.4.15"
         ↪  code="163001000000103" displayName="Diagnoses"/>
        <component typeCode="COMP" contextConductionInd="true">
            <templateId root="2.16.840.1.113883.2.1.3.2.4.18.2"
             ↪  extension="CSAB_RM-NPfITUK10.component"/>
            <UKCT_MT144042UK01.Diagnosis classCode="OBS" moodCode="EVN">
                <id root="A1389221-83DF-47C7-8DE9-1B51A35ECFAE"/>
                <code code="66071002" displayName="Type B viral hepatitis"/>
                <statusCode/>
                <effectiveTime value="20050214"/>
            </UKCT_MT144042UK01.Diagnosis>
        </component>
    </pertinentCREType>
</pertinentInformation2>
```

Figure 2.6: Diagnoses section from GP Summary template

In a similar manner, template identifiers are assigned to each section in the document. Each section and component is extended from a base definition, which is akin to class inheritance in object oriented programming. Base template with the OID value of `2.16.840.1.113883.2.1.3.2.4.18.2` corresponds to approved NPfIT templates by HL7 [42]. Further refinements to this base template is described by extensions.

Sections are identified by coded values. `2.16.840.1.113883.2.1.3.2.4.15` is the identifier for UK Edition of SNOMED CT [42]. Code value of `163001000000103` in SNOMED CT terminology system indicates Diagnoses section.

2.3 Medical Terminologies

In technical domains, formal representation of information is usually restricted to terminology systems and dictionaries in order to form a common ground for communication of services and enable computers to process this informa-

tion. Similarly, in healthcare, concepts and medical terms are uniquely defined by terminologies. This constraint resolves ambiguity of information and creates semantics for applications to accurately exchange data. For instance, in SNOMED CT terminology system [11], *depression* is represented with concept code `D9-52000` with a fully specified name as `Depressive disorder (disorder)`. However, when a healthcare professional provides information in free text as "depressed", "depression", "melancholia" or "depressive episode", she means the same exact concept. That's why those concepts are also linked to concept code `D9-52000` as synonyms [36] (see Table 2.5) and instead of providing diagnosis in free text, healthcare professional selects it from a pre-defined value set which, in this case, is SNOMED CT. In this way, since the information becomes machine processable, applications that use SNOMED CT for the semantics of clinical terminology can interoperate seamlessly.

Table 2.5: D9-52000 concept in SNOMED CT

ID	Term	Label
767133013	Depressive disorder (disorder)	Fully specified name
59212011	Depressive disorder	Preferred term
486187010	Depressed	Synonym
416184015	Depression	Synonym
486185019	Depressive episode	Synonym
486186018	Depressive illness	Synonym
59218010	Melancholia	Synonym

2.3.1 Terminology Systems

ICH E2B (R2), epSOS Patient Summary, HL7 C-CDA CCD and SCR depend on a wide range of terminology systems [2, 39, 45]. In this section, very brief descriptions about some of those terminologies and purposes they serve are provided.

Since our work is not directly concerned with mapping of terminology systems,

we have made use of existing efforts in this field, details of which are explained in Section 2.3.2. That's why we find it hardly necessary to further discuss existing systems.

- **SNOMED CT**, Systemized Nomenclature of Medicine Clinical Terms [11]: It is one of the most comprehensive collection of clinical healthcare terms. It provides core concepts for EHRs in multiple languages –one of the main reasons why it is so widely used.

- **MedDRA**, Medical Dictionary for Regulatory Activities [8]: It is a rich medical terminology for regulatory information of medical products. MedDRA terms cover all information about medical products (e.g. pharmaceutical drugs, vaccines, devices) such as adverse drug reactions, symptoms, indications, relevant patient history. That's why it is adopted in ICH E2B (R2) individual case safety report standard and universally accepted by pharmaceutical companies and regulatory agencies as well.

- **LOINC**, Logical Observation Identifiers Names and Codes [7]: It is the go-to code system for identifying clinical observations, tests and measurements, which is used in many HL7 and IHE standards.

- **ICD**, International Statistical Classification of Diseases [6]: Maintained by World Health Organization, it is the standard for epidemiological and clinical information. Included concepts and terms classify diseases, signs, health problems and general health situation.

- **ATC**, The Anatomical Therapeutic Chemical Classification System [1]: It is the standard for identification of active ingredients of drugs, and their therapeutic and chemical characteristics. Since the classification is based on organs or systems on which the active ingredient act on, as a drug may have multiple codes, different brands may share the same code if their drug have the same active substance. By this way, the information that is meaningful in terms of patient care is extracted and standardized.

- **UCUM**, The Unified Code for Units of Measure [12]: It is a code system developed with the aim of gathering all units of measures being used in

science, engineering and business under a standard for electronic exchange.

2.3.2 Terminology Server

Terminology systems mention in Section 2.3.1 are mostly developed collaterally in different subdomains. Due to interdisciplinary areas in health, terminology systems have grown into something with large overlaps in-between. For example, ICH E2B (R2) content model uses MedDRA codes for the representation of adverse reactions, whilst EHR sources usually code this information using SNOMED CT and ICD-10. Therefore, two applications using those content models require a semantic mapping of different terminology systems to communicate and interoperate.

In the scope of Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies (SALUS) project [13], a terminology reasoning service, which provides mapping between different terminology systems, has been developed [78]. On top of this, a terminology server that stores terms, term mappings and serves Web interfaces for querying has been also implemented.

As far as we are concerned, current version of SALUS Terminology Server maintains term mappings between SNOMED CT, MedDRA and ICD-10 code systems and provides a RESTful Web service for developers.

2.4 Metadata Management

For organizations that stores data, transmits data in different structures and needs consistent definitions of data across time; tools for creating, defining an managing metadata is essential. Our case of data interoperability is no different. We need tools for creating content models, maintaining mappings of different standards to this model and meaning of information retained in this model. Metadata registries serve for this purpose. They provide authorization for individuals, and storage for data element semantics and representation.

ISO/IEC 11179, an international standard, is developed with the aim of providing metadata-driven data exchange in heterogeneous environments. Combining principles of semantic theory and data modeling, the standard defines the representation of metadata in a metadata registry [50].

2.4.1 Semantic Metadata Registry

An ISO/IEC 11179 compliant Semantic Metadata Registry (MDR) is developed in the context of SALUS project [13]. It allows users to query and browse the data model repository, and create their data models based on existing data models and hence increase the reusability as well as reducing the data model redundancy [67].

Specifically for our work, metadata of EHR sources and E2B data elements (e.g. age, birth data, allergy, etc.) can be stored in Semantic MDR together with data extraction specifications to locate data in the content model (e.g. XPath, SQL).

Semantic MDR has been released as open source[1].

2.4.2 IHE Data Element Exchange

In any field levering from the secondary use of electronic health records, a method is necessary to map EHR data to corresponding domain meanings. Semantic MDR (see Section 2.4.1) may serve as the storage for semantics of the data elements used in EHRs and post market surveillance domain. IHE Data Element Exchange (DEX) profile, on the other hand, makes sharing of machine processable metadata possible through standardized transactions [48]. Authored by SRDC Ltd., it defines a standard interaction with metadata registry to:

i. retrieve data element list of a selected domain (e.g. E2B data elements),

ii. retrieve metadata (e.g. extraction specifications) for a data element.

[1] `https://github.com/srdc/semanticMDR`

```
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
               xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                ↪  xmlns:wsa="http://www.w3.org/2005/08/addressing"
               xmlns:xsd="http://www.w3.org/2001/XMLSchema">
<soap:Header>
    <wsa:MessageID>urn:uuid:f43f7bda-a5f9-42b1-b8dc-e78be</wsa:MessageID>
    <wsa:Action>urn:ihe:qrph:dex:2013:RetrieveMetadata</wsa:Action>
</soap:Header>
<soap:Body>
    <dex:RetrieveMetadataRequest xmlns:dex="urn:ihe:qrph:dex:2013">
        <dex:id>19958a47-b1cb-4014-af0f-a9c5c86c2576</dex:id>
        <dex:registrationAuthority>SRDC</dex:registrationAuthority>
        <dex:version>0.1</dex:version>
    </dex:RetrieveMetadataRequest>
</soap:Body>
</soap:Envelope>
```

Figure 2.7: IHE DEX Retrieve Metadata request

In other words, given a profile, DEX is able to list the data elements included in it; given the data element list, it can also provide mappings of elements to a specified target profile. This enables dynamic pre-population of forms from the information provided in EHR sources.

The protocol for DEX transactions is based on Simple Object Access Protocol (SOAP) 1.2. In Figure 2.7 an example SOAP request of `RetrieveMetadata` transaction is provided.

2.5 IHE Retrieve Form for Data Capture

IHE Retrieve Form for Data Capture (RFD) is an interoperability profile of IHE defining where to retrieve an electronic form, how to pre-populate data in the form, how to fill it, and where to send the filled-in form in a standardized manner [49].

To illustrate with an oft-used example which overlaps with our case in this thesis, consider that a third party application utilizes EHRs of the patient, thus

requires access to EHR source. By implementing RFD profile, EHR provider can retrieve a data capture form from the third party application, fills the form, and returns it back to the third party application without leaving its (i.e. EHR source) local context.

RFD profile, itself, does not specify the content but mechanisms to populate it. RFD identifies four actors and four transactions between them, see Figure 2.8.

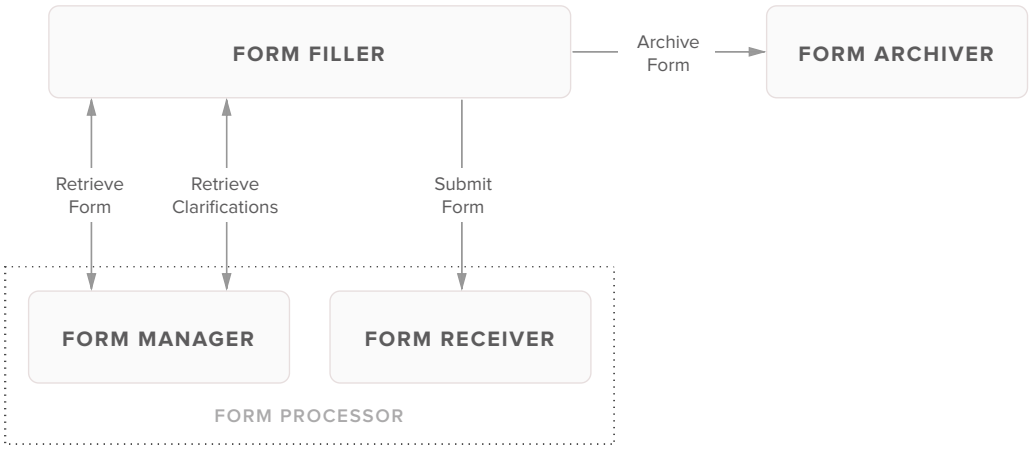

Figure 2.8: Actors and transactions in RFD

- **Form Manager:** Form supplier. It responds to `Retrieve Form` request by the form filler and provides the desired forms. It may maintain a storage for forms, or create these on the fly when requested.

- **Form Filler:** It acts as a transporter between actors; requests the form from form manager, processes it, then either submits it to form receiver or sends an `Archive Form` request to form archiver. When sending `Retrieve Form` request, it may also provide contextual information that is relevant for pre-population of form on the form manager.

- **Form Archiver:** It receives completed forms and records these in a persistent store.

- **Form Receiver:** It receives completed forms from the form filler and process these as desired. From receiver's responsibility is not specified in

RFD profile. Regulatory agencies collecting adverse drug event reports would act as the form receiver in an RFD based environment.

Actors and roles assigned to components in our architecture based on RFD will be further explained in Chapter 4.

CHAPTER 3

CONTENT MODEL AND MAPPING

Starting from this section, work carried out in the course of this thesis is demonstrated, justifications for selected enabling technologies are given in related sections. In this chapter, we discuss the process of developing a content model for our system and identifying extraction specifications for different data models.

3.1 EHR Sources

We started our research by defining the scope of EHR sources that will be able to operate with our system. As this study is conducted in as a part of WEB-RADR project [14], initial goal was to select EHR content models based on their applicability on the pilot sites of the project. Project consortium has two major regulatory agencies as partners, Medicines & Healthcare products Regulatory Agency (MHRA) in the United Kingdom and Agency for Medicinal Products and Medical Devices (HALMED) in Croatia, which are authorized authorities for submitting ADE reports at national level in their own countries. For possible exploitation opportunities in those countries, initial EHR content models are determined to be established standards in the UK and Croatia.

In the UK, Summary Care Record, introduced in Section 2.2.6, is the designated EHR standard at national level [33]. It includes general practitioner (GP) summary of patient, which we find fit as the data source to pre-populate ADE reports.

On the other hand, in Croatia, healthcare systems are lacking in a standard

data model for the exchange of patient records at national level [52]. However, as a partner in the epSOS project, Croatia has piloted electronic prescription and willing to implement epSOS Patient Summary in healthcare services [53]. Besides, EC has already published guidelines defining epSOS Patient Summary based datasets for electronic exchange of EHRs in Europe, which makes it a candidate for pan-European solution [27]. Therefore, epSOS PS (see Section 2.2.4) is selected to be integrated to our system as one of the initial content models.

Although project scope is limited to Europe, our tool is not dependent on predefined content models, can be extended easily. For the purpose of demonstrating that transatlantic exploitation is also possible, we have decided to integrate Consolidated CDA (explained in Section 2.2.5) based EHR sources, which is the standard in the United States, too [24]. There is an existing effort with the aim of establishing interoperability of electronic health records and patient summaries among the EU and US, in the scope of Trillium Bridge project [9, 22]. Funded by EC and having prestigious partners like HL7 and IHE, it establishes the foundations of transatlantic exchange of healthcare information based on epSOS PS and C-CDA. We have made use of mappings between epSOS PS and C-CDA CCD, which is delivered by the Trillium Bridge [73], in our system.

As there is no actual EHR provider beneficiary neither contributing to our study nor in the WEB-RADR project, we recognized that it would be impossible to fulfill the technical, organizational and ethical requirements of access to real patient data. Therefore, in order not to raise security and privacy concerns, we simulated the EHR sources with hand-crafted synthetic patient data (see Figure 3.1). Instead, we focused on building something technically ready for wide-scale integration; and complying with the national standards for patient summary representation and exchange mechanisms in our proof of concept prototype.

```
<!-- Patient weight representation in SCR -->
<component typeCode="COMP" contextConductionInd="true">
    <templateId root="2.16.840.1.113883.2.1.3.2.4.18.2"
     ↪  extension="CSAB_RM-NPfITUK10.component"/>
    <UKCT_MT144043UK02.Weight classCode="OBS" moodCode="EVN">
        <value unit="kg" value="95"/>
        <id root="7C257140-B85F-40DF-BF14-6C1271E314F8"/>
        <code code="27113001" displayName="body weight"/>
        <effectiveTime value="20070210"/>
    </UKCT_MT144043UK02.Weight>
</component>

<!-- Patient weight representation in PS -->
<component>
    <observation classCode="OBS" moodCode="EVN">
        <templateId root="1.3.6.1.4.1.19376.1.5.3.1.4.13"/>
        <templateId root="2.16.840.1.113883.10.20.1.31"/>
        <templateId root="1.3.6.1.4.1.19376.1.5.3.1.4.13.2"/>
        <id root="662ad0f1-c309-4a97-9e94-c97b84ebab8e"/>
        <code code="3141-9" displayName="Body weight (Measured)"
         ↪  codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC"/>
        <statusCode code="completed"/>
        <effectiveTime value="20100407"/>
        <value xsi:type="PQ" value="88" unit="kg"/>
    </observation>
</component>

<!-- Patient weight representation in C-CDA -->
<component>
    <observation classCode="OBS" moodCode="EVN">
        <templateId root="2.16.840.1.113883.10.20.22.4.27"
         ↪  extension="2014-06-09" />
        <id root="f4e729e2-a97f-4a7e-8e23-c92f9b6b55cf" />
        <code code="3141-9" codeSystem="2.16.840.1.113883.6.1"
         ↪  codeSystemName="LOINC" displayName="Patient Body Weight -
         ↪  Measured" />
        <statusCode code="completed" />
        <effectiveTime value="20120910" />
        <value xsi:type="PQ" value="86" unit="kg" />
        <interpretationCode code="N" codeSystem="2.16.840.1.113883.5.83" />
    </observation>
</component>
```

Figure 3.1: XML snippets showing synthetic patient data in SCR, PS and C-CDA

3.2 Unified Data Model for Reporting

Mainly depending on national specifications, regulatory agencies accept individual case safety reports in ICH E2B format [15, 31, 59, 71]. VigiBase, which is the WHO's global ICSR database, also expects ADE reports to be submitted as E2B files [58]. However, occasionally spontaneous reporting systems may collect supplementary information for deeper statistical analysis and research. That's why, while developing a content model for our system, although we have based it on E2B data elements, we have examined online SRSs to identify additional information asked, as well.

We have extracted required fields from forms in:

- Yellow Card Scheme reporting site, available in the UK [60].
- Medwatcher mobile reporting application, available in the US [29].
- Primary reporting tool of UMC, available in Croatia, Turkey and Venezuela [77].

Since our aim is to be as extendable and flexible as possible, we created a data model as a union of data elements extracted from these tools and required by E2B specification.

Designated unified data model for reporting contains 112 data elements, all of which is presented in Appendix A. Information contained in a complete E2B form can be represented using the new data model. Additionally, Yellow Card Scheme, Medwatcher and UMC's reporting tool can also utilize additional information fields.

Content model is based on data modeling methods defined by ISO/IEC 11179. It, in the simplest terms, combines a high-level *concept* with an *object class* to form a *data element concept* [50]. For example, "age" is a very wide concept that can have multiple meanings based on the context. When it (as a *concept*) is combined with an *object class*, such as "patient", formed *data element concept* (age of the patient) represents a more specific information. An example object

Table 3.1: Episode object class and its data elements

ID	Property	Data Type	Enumerated
0017	MedDRAVersion	characterstring	-
0018	name	MedDRACode	-
0019	startDateFormat	integer	102:CCYYMMDD, 203:CCYYMMDDHHMM, 602:CCYY, 610:CCYYMM
0020	startDate	characterstring	-
0021	continuing	integer	1:Yes, 2:No, 3:Unknown
0022	endDateFormat	integer	102:CCYYMMDD, 203:CCYYMMDDHHMM, 602:CCYY, 610:CCYYMM
0023	endDate	characterstring	-
0024	comments	characterstring	-

class of "Episode" is presented in Table 3.1 with its properties and data types. Episode class holds the information about past illnesses, surgeries, allergies or other adverse reactions of the patient. Each row in the table refers to a *data element* in ISO/IEC 11179 terms.

3.3 Mapping via Extraction Specifications

Common data element (CDE) is defined as the smallest meaningful data container in a given context based by Semantic MDR based on ISO/IEC 11179 specifications. CDE has one or more extraction specification(s) which is a machine processable expression that points to the exact place of the CDE in a content model [67]. In other words, information represented in a CDE can be located in a model (e.g. epSOS Patient Summary) using extraction specifications. Extraction specifications can be XPath expressions, SQL queries, SPARQL queries or any other expression that can be processed by a computer to locate information

in a data instance. As information may lay in multiple locations in the model, CDEs are not limited to have a single extraction specification. Consider the CDE for start date of a medical episode (in Table 3.1); a medical episode may be a procedure, an allergic reaction or an illness. These information exist in different sections with different entry structures in an EHR profile. Therefore, CDE for representing "start date of a medical episode" has to have multiple extraction specifications to locate the information properly in EHRs. This, concurrently, enables us to handle mapping of unified model to EHR sources and reporting formats by employing Semantic MDR.

Unified model, by itself does not specify the business logic to populate or consume data. As EHR sources and ICSRs are of different structures, they must be mapped to a common model, which would make their interoperability possible. Extraction specifications serve for this purpose by enriching a *dummy* model definition with some logic specific to source/target model for retrieving information.

Since HL7 CDA and ICH E2B is XML based document structures, we have used XPath expressions as extraction specifications. For each of 112 data elements, we have defined paths to locate this particular piece of information in epSOS PS, C-CDA CCD, SCR and E2B. Following the example above, Table 3.2 shows extraction specifications defined for each of these four content models. Note that this information is available in 4 and 3 different locations in an epSOS PS document and a C-CDA CCD record, respectively. Two XPath expressions from each are presented due to space constraints.

By such an approach, different content models do not depend each other to but to a common information model instead, which allows additional sources to be integrated to the system with ease. In order to make the system compatible with a new EHR source, extraction specifications of each CDE for that source shall be imported into Semantic MDR (see Section 3.4 for importing process). Same applies to adding new report models to the system. Once CDEs are mapped to the new source/target, the system becomes capable of handling that profile just like any other. In other words, system does not require any update on business

logic to add a new content model integration.

3.4 Semantic MDR Importer

Semantic MDR provides user interfaces to manage information models, create, update or delete data elements. However, it lacks an option for bulk operations such as importing a new context or updating all CDEs in the information model

Table 3.2: Extraction specifications for start date of a medical episode

Model	XPATH
E2B (R2)	/ichicsr/safetyreport/patient/medicalhistoryepisode/patientmedicalstartdate
epSOS PS	//hl7:section[hl7:templateId[@root='1.3.6.1.4.1.19376.1.5.3.1.3.8']]//hl7:observation[hl7:templateId[@root='1.3.6.1.4.1.19376.1.5.3.1.4.5']]/hl7:effectiveTime/hl7:low/@value
epSOS PS	//hl7:section[hl7:templateId[@root='1.3.6.1.4.1.19376.1.5.3.1.3.12']]//hl7:procedure[hl7:templateId[@root='1.3.6.1.4.1.19376.1.5.3.1.4.19']]/hl7:effectiveTime/hl7:low/@value
C-CDA CCD	//hl7:section[hl7:templateId[@root='2.16.840.1.113883.10.20.22.2.5.1']]//hl7:observation[hl7:templateId[@root='2.16.840.1.113883.10. 20.22.4.4']]/hl7:effectiveTime/hl7:low/@value
C-CDA CCD	//hl7:section[hl7:templateId[@root='2.16.840.1.113883.10.20.22.2.7.1']]//hl7:procedure[hl7:templateId[@root='2.16.840.1.113883.10.20.22.4.14']]/hl7:effectiveTime/@value
SCR	//hl7:pertinentInformation2/hl7:pertinentCREType[hl7:code[@code="163001000000103"]]/hl7:component/hl7:UKCT_MT144042UK01.Diagnosis/hl7:effectiveTime/hl7:low/@value

```
Context Name = WEBRADR
Concept Model = E2B (R2)
OID = 1.11.111.1.111111.1.7

...
0007   XPATH   /ichicsr/safetyreport/primarysource/reportertitle
0008   XPATH   /ichicsr/safetyreport/primarysource/reportergivename
0009   XPATH   /ichicsr/safetyreport/primarysource/reporterfamilyname
...
0043   LITERAL CCYYMMDD
0044   XPATH   /ichicsr/safetyreport/patient/test/testdate
0045   XPATH   /ichicsr/safetyreport/patient/test/testname
...
```

Figure 3.2: Extraction specification input for Semantic MDR importer

with new extraction specifications. When integrating a new data source, navigating through all data elements using the Web interface and adding XPath expressions to each one take considerable effort.

To overcome this limitation, we have implemented a wrapper around Semantic MDR API, and an importer on top. The software is capable of importing a new information model, integrating a data source to an existing model, or dropping already imported models and sources. Figure 3.2 shows an example input file for importing extraction specifications of E2B (R2) for WEBRADR information model. Header lines are to specify intended context and provide a name and OID for the new data source. The rest is tab delimited triplets in each line: data element identifier, extraction specification type and expression.

Current system supports constant values as extraction specifications. Type `LITERAL` indicates a constant value, value of which is not evaluated. XPath expressions, on the other hand, are evaluated on a given data instance.

Semantic MDR API wrapper and importer in Python is open-sourced and publicly available on GitHub[1].

[1] `https://github.com/mkubilayk/mdr-importer`

CHAPTER 4

SYSTEM ARCHITECTURE

This section presents overall architecture of the tools developed in the course of this work, communication between actors, and typical workflow.

The component roles and transactions between components are totally based on IHE Retrieve Form for Data Capture standard that is explained in Section 2.5. Form Manager and Form Filler are the two main actors in RFD transactions, supplying and processing forms, respectively. Form Receiver and Form Archiver actors can only receive completed forms and are not capable of making requests according to specifications.

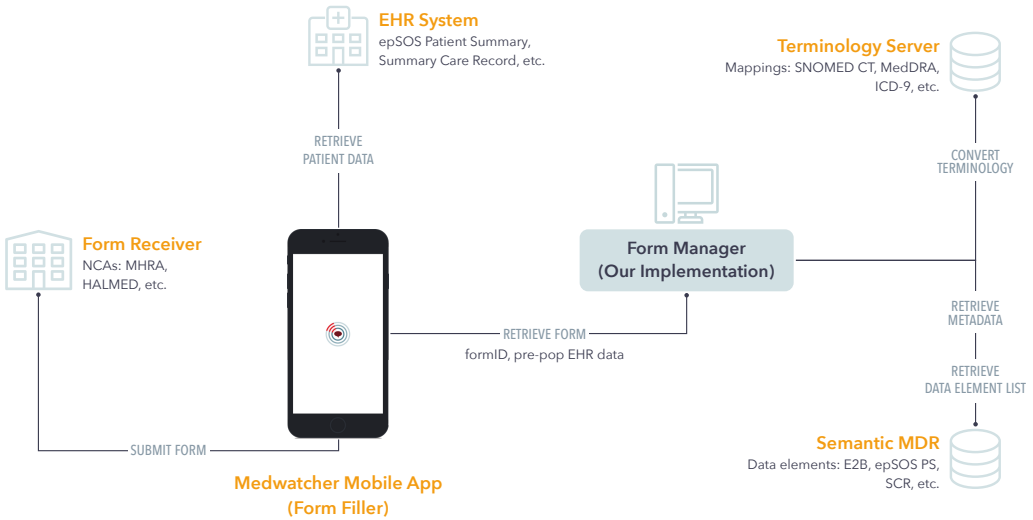

Figure 4.1: Overall architecture

Figure 4.1 shows overall system architecture. Terminology Server and Semantic MDR are third-party applications employed by our tools. EHR system would be a provider in real world however, as discussed, it is simulated in prototype implementation. For other actors of the system (Form Manager, Form Filler and Form Receiver), please refer to relevant section below.

4.1 Form Manager

In the general architecture, our platform, which does the heavy work of consuming heterogeneous inputs and converting them into a single model, plays the role of Form Manager. Form Manager implementation in this system, does not utilize persistent stores; it creates forms on the fly instead.[1] Therefore, security and privacy requirements of storing healthcare data does not concern the system. They are spontaneous reporting systems and EHR sources who are responsible for providing required environment.

```
<RetrieveFormRequest>
    <prepopData>
        <ClinicalDocument>
        ...
        </ClinicalDocument>
    </prepopData>
    <workflowData>
        <formID>00021</formID>
        <encodedResponse>false</encodedResponse>
        <context><text>C-CDA CCD</text></context>
    </workflowData>
</RetrieveFormRequest>
```

Figure 4.2: Request for Retrieve Form transaction in RFD

While creating forms as a response to Retrieve Form requests, Form Manager does not access any source out of its context for additional information. Prepopulation is done using the data provided in Retrieve Form request.

Figure 4.2 shows an example SOAP body of Retrieve Form request in an RFD

[1] RFD profile does not put a restriction on the type of storage to be used by Form Manager [49].

compliant system:

- `<prepopData>`: The information supplied by the Form Filler as a well-formed XML document (i.e. a valid EHR) for pre-populating fields.
- `<workflowData>`: Well-formed XML document containing workflow specific values.
- `<formID>`: Identifier of the requested form.
- `<encodedResponse>`: Tells the Form Manager to return a URL pointing to the form, or the form encoded in response.
- `<context>`: It may contain any XML document. In our implementation it is used to specify the content model of pre-population data. Provided information is in C-CDA CCD format in Figure 4.2.

Pre-population data shall contain EHRs of the patient that is being reported having an adverse drug event. Using EHR data and employing a meta-data registry and terminology server, Form Manager creates a pre-filled ICSR. The process of creating a new form is as follows:

1. Upon receiving a Retrieve Form request, Form Manager queries Semantic MDR for a list of data elements. Based on IHE DEX profile, it makes a Retrieve Data Element List transaction.
2. Having all data elements required for building an E2B document, Form Manager makes additional two requests (Retrieve Metadata, defined in IHE DEX) per data element to retrieve extraction specifications for i) E2B; ii) EHR source model, which is C-CDA CCD in the example but can be anything among integrated EHR content models (epSOS PS, C-CDA CCD, SCR). When all requests are completed, Form Manager is aware of where to find relevant information in the given EHR data, and where to put that information in a new E2B document.
3. In a top-down manner, starting from the root data element, Form Manager first executes extraction specifications on EHR data to locate relevant patient

```
<RetrieveFormResponse>
    <form>
        <URL>http://rdf.webradr.local/srs/forms/00021</URL>
        <instanceID>001425324</instanceID>
    </form>
</RetrieveFormResponse>
```

Figure 4.3: Response for Retrieve Form transaction in RFD

context; then, if the information is available in patient's EHR, it fills appropriate fields in E2B document. The process is straightforward for free text fields. However, coded values require post-processing for conversion between terminology systems, which is explained in Section 4.1.1.

4. When all data elements are processed, and coded values are converted, Form Manager sends the response to Form Filler containing an E2B report or a link to it, illustrated in Figure 4.3.

4.1.1 Converting Terminologies

Due to the fact that semantic coding is not consistent among content models used in this system, a terminology server, which contains relevant mappings, is employed to provide the conversion.

Terminology Server (see Section 2.3.2), serves semantic definitions of and the relations (i.e. semantic mappings) between terminologies [78]. Although Terminology Server has been loaded with relations between large code systems such as SNOMED CT, MedDRA and ICD-10 in the course of SALUS project, it does not contain semantic definitions of E2B codes, which are only used in ICSRs. For instance, instead of using a well-established terminology for coding dose units (e.g. UCUM), E2B (R2) specification defines its own enumerated value set. Similarly, route of administration of medications are coded in a defined value set of E2B (R2). Since those value sets include relatively low number of codes (see Table 4.1), it was possible for us to provide their semantic definitions in Terminology Server. Terminology Server accepts Resource Description

Framework (RDF) documents to load new terminologies to the system [54]. Using Simple Knowledge Organization System (SKOS) vocabulary, which provides a model for expressing controlled vocabularies as concept schemes in RDF [61], we defined E2B (R2) local terminology systems as concept schemes. SKOS also provides properties for linking semantically related concepts (i.e. codes) from different concept schemes, which exactly corresponds to mapping between terminology systems. Many of coded values in Table 4.1 do not require medical interpretation, indeed, can be mapped one-to-one to other terminology systems. Hence, `skos:exactMatch` property was enough to link two concepts so that concepts can be used interchangeably. An example concept scheme definition of E2B(R2) Dose Unit in RDF is available in Appendix C.

Table 4.1: Enumerated values in E2B (R2) defined as terminology systems

Concept Scheme	**OID**	**# of Terms**
AdministrationRoute	2.16.840.1.113883.3.989.2.1.2	67
DateFormat	2.16.840.1.113883.3.989.2.1.7	4
DoseUnit	2.16.840.1.113883.3.989.2.1.1	32
DurationUnit	2.16.840.1.113883.3.989.2.1.6	7
Gender	2.16.840.1.113883.3.989.2.1.5	3
Qualification	2.16.840.1.113883.3.989.2.1.3	5

Form Manager is the consumer of Terminology Server API endpoints serving semantic definitions and relations between concepts. To be used afterwards, XPath values defined in data elements extract coded system OIDs along with actual codes. When Form Manager generates an intermediate E2B form (not completely valid), it has coded values in the terminology system defined in its EHR source and identifier of the terminology system. Form Manager queries Terminology Server with triplets containing target concept scheme identifier, source concept scheme identifier and concept notation or label. Response is the corresponding coded value in the target concept system, which is –in our case– specified by E2B (R2). In this way, medical terminologies (SNOMED CT, ICD-10, etc.) are converted to MedDRA codes. Other coded values are also replaced

by their equivalent enumerated values according to E2B (R2) specifications.

Upon completion of this process, Form Manager has a structurally valid and semantically correct individual case safety report as an E2B (R2) document, which will be sent to Form Filler. For an example ADE report, please see Appendix B.

4.2 Form Filler

Spontaneous reporting systems correspond to Form Filler actors in RFD profile. SRSs may have their own implementation of Form Filler. It initiates communication with EHR sources to fetch patient records. By providing EHRs as pre-population data for the Form Manager, it retrieves an E2B document pre-filled with relevant patient context residing in EHRs.

As proof of concept, we have also implemented a mobile adaptor to be integrated with a mobile ADE reporting application on iOS devices. Form Filler communicates with our Web services serving hand-crafted synthetic patient data.[2] The prototype implementation displays the report retrieved from Form Manager in way that it is easy to exclude some of the pre-populated information and include new data. For a detailed scenario and screenshots of the prototype, please refer to Chapter 5.

4.3 Form Receiver

In a real life setting, regional and national pharmacovigilance authorities would be Form Receiver actors. In the scope of WEB-RADR project, MHRA (UK) and HALMED (Croatia) will be the agencies collecting adverse drug event reports at national level. In other words, mobile spontaneous reporting application (Medwatcher) will be localized for each pilot site and send ICSRs to national authorities. Since our research scope in this thesis is limited to pre-population of ICSR forms, we have not provided interfaces for submitting ADE reports to

[2] The reason behind this choice is discussed in Section 3.1.

From Receivers.

To be fully compliant with RFD profile, stakeholders in the role of Form Receiver shall be capable of responding to Submit Form transactions of Form Filler (i.e. mobile application).

CHAPTER 5

AN EXAMPLE SCENARIO

In this section, an example scenario for end users (i.e. doctors and patients) of the mobile ADE reporting application is presented.

- The user is a patient who has suffered an adverse reaction. She wants to report suspected incident.
- Or the user is a healthcare professional who diagnosed one of her patients with an adverse reaction and is going to report the case to pharmacovigilance agencies.

Either healthcare professional or patient, the user shall have the credentials to access reported patient's electronic health records. If credentials are not available, generated report would not contain any patient context. For the sake of simplicity, authentication mechanism in the prototype is not a secure implementation. Since we do not access real EHR data, it only asks for a patient identifier. In a complete integration, Form Filler on mobile application has to implement proper authentication and authorization methods while accessing EHR sources out of application context.

1. The user opens mobile ADE reporting application on her mobile device.
2. She provides an identifier to enable access to patient's health records, see the first screen in Figure 5.1.
3. Adaptor implementation (i.e. Form Filler) embedded into mobile application

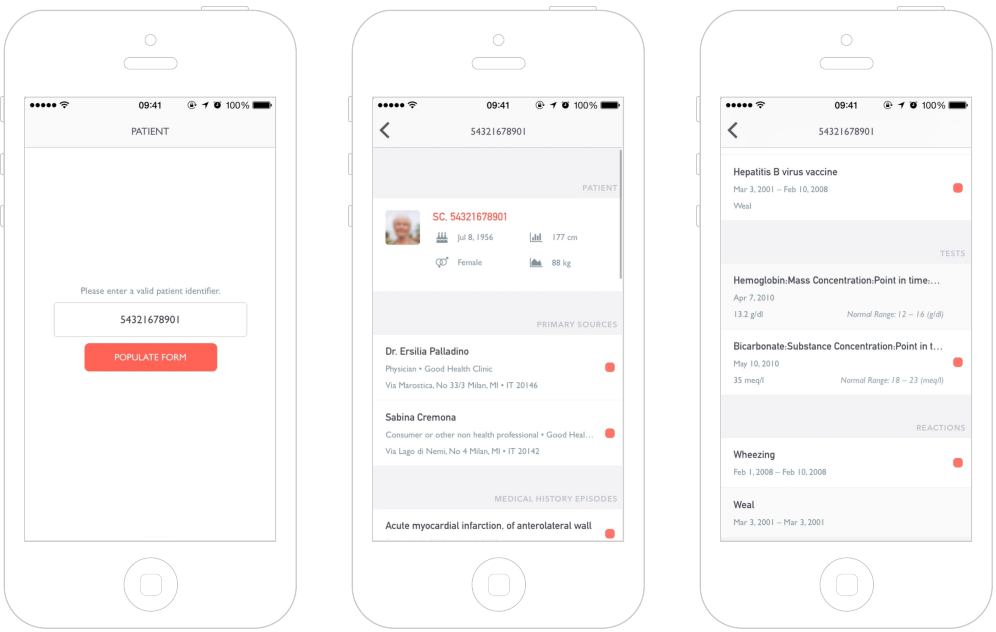

Figure 5.1: Screens of mobile prototype

retrieves a summary of patient healthcare information from the integrated EHR source.

4. Adaptor in the mobile application sends the request, which contains retrieved EHRs, for a pre-filled ADE report from Form Manager.

5. Form Manager, using our methods, filters the health records; locates reusable information such as past/active medications, diagnoses, vital signs; extracts and places these into the relevant fields of an E2B form.

6. Form Manager sends the E2B document as a response to Form Filler.

7. Mobile application displays the form to user for checking pre-filled information, see second and third screens in Figure 5.1. She is free to exclude any pre-populated data or to complete missing fields.

8. When ADE report is complete, mobile application sends it to national pharmacovigilance agency according to regulations.

CHAPTER 6

TESTS AND RESULTS

Our implementation is tested as a prototype and currently being integrated with Medwatcher, which is a USA based mobile application being adapted to European context within the scope of WEB-RADR project. Medwatcher will be tested in two pilot sites: the UK and Croatia. When pilot studies are conducted, we will also be able to report on results from the end-user validation activities.

Although, we have currently run some test revealing different aspects than end-user opinions, an example set of survey questions is also provided.

6.1 User Evaluation Survey

- To healthcare professionals:
  - How often do you report adverse drug events? (#/week)
  - How many hours do you spend on reporting adverse drug events every week?
  - Using Medwatcher app, how many adverse drug events have you reported per week on average?
  - Using Medwatcher app, how many hours have you spent reporting every week?
- To patients:
  - Have you ever reported an adverse drug event to regulatory bodies?

  - Using Medwatcher app, have you ever reported any adverse drug event? If yes, how many?

- To regulatory agencies:

  - How many adverse drug events do you receive daily?
  - What is the volume of reports submitted through Medwatcher app?
  - Compared to conventional reporting systems, how is the quality of reports submitted through Medwatcher app?

  Based on the answers of healthcare professionals, we expect to see an increase in the number of reports submitted every week and a decrease in the time spent reporting adverse events. Secondly, we also expect a raise in the number of patients getting involved in the reporting process voluntarily. Lastly, number of events reported to regulatory agencies and quality of those reports are expected to increase.

6.2 Coverage

For evaluating coverage of generated reports, we have created a dataset of 100.000 EHR documents, having different levels of comprehensiveness. Based on the number of sections available in the document, we have classified data quality in populated records. For example, a C-CDA document consists of 15 sections. Records having 0-2 sections are classified as "very poor", 3-6 sections as "poor", 7-9 sections as "moderate", 10-12 sections as "rich" and 13-15 sections as "almost complete". Each of these categories have 20.000 records which have been populated randomly of different source templates.

Responses of 100K requests containing EHR data have been stored. For each pre-filled E2B document, we have calculated a coverage rate, which is the rate of pre-populated data elements through EHR to all data elements of E2B template. One important point to note is that when calculating coverage, repeated blocks of the E2B form (e.g. medical history episodes, tests, drug therapies, etc.) were counted as one; i.e. even when we extract and map multiple instances of medical

Table 6.1: Coverage rates (%)

Data	Min	Mean	+/- SD	Median	Max
Very poor	31.92	36.71	7.89	32.07	61.32
Poor	32.07	50.89	11.25	52.83	70.75
Moderate	38.68	61.02	9.84	63.21	73.88
Rich	40.57	68.43	6.24	70.75	74.02
Almost complete	50.00	73.32	2.42	74.53	74.53
Total	**31.92**	**58.15**	**15.38**	**63.20**	**74.53**

history episodes, procedures, tests, reactions or drugs from patient EHR data, they are treated as binary values such that whether relevant block is covered or not. Considering that it is very common to have multiple diagnoses, procedures or drug therapies in EHRs, and these pieces of information is very meaningful in ADE reports for pharmacovigilance studies, coverage rates may be considered as much higher.

Table 6.1 presents minimum and maximum coverage rates achieved in each data quality level; and mean, median rates and standard deviation among all reports. Since even the least comprehensive data includes some information related to patient or administration, we were able to pre-fill ∼32% of an E2B report at the worst case. Using nearly complete patient medical history, we were able to pre-fill ∼73% of the ADR report. Even if, an electronic health record contains all the information it is designated to store, generated E2B reports may still have empty fields (∼25%) that contains administrative information or some more medical details which is not available in EHRs.

6.3 Load Tests

Since the proposed platform will serve for mobile applications, we have run load tests for benchmarking our server. The tests include 100 to 50K requests at four different concurrency levels. Note that these tests are conducted on a single

instance of server application running on a personal computer with 2.3 GHz Intel Core i7 processor and 16 Gb of memory. EHR data used in these tests are completely filled instances, which requires relatively more computational power.

Concurrency levels presented below correspond to the number of requests arrive to the server at the same time, concurrently. Table 6.2 presents the minimum and maximum connection times with mean, median and standard deviation of all requests. Since minimum and maximum values of connection times are calculated as average within concurrent request batches, as the number of concurrent requests increase the values approach to the mean.

Table 6.2: Connection times (ms)

Test	**Min**	**Mean**	**+/- SD**	**Median**	**Max**
N=100, C=1	405	467	53.60	456	760
N=5000, C=100	414	487	40.35	489	546
N=10000, C=500	394	494	43.27	491	510
N=50000, C=1000	423	481	41.60	483	490

1. Total requests: 100. Concurrency level: 1.
    - Requests per second: 2.14
    - Time per request: 467.023 ms
    - Transfer rate (received): 32.94 kb/s
    - Transfer rate (sent): 360.81 kb/s
2. Total requests: 5000. Concurrency level: 100.
    - Requests per second: 2.05
    - Time per request: 487.215 ms
    - Transfer rate (received): 31.57 kb/s
    - Transfer rate (sent): 345.86 kb/s
3. Total requests: 10000. Concurrency level: 500.

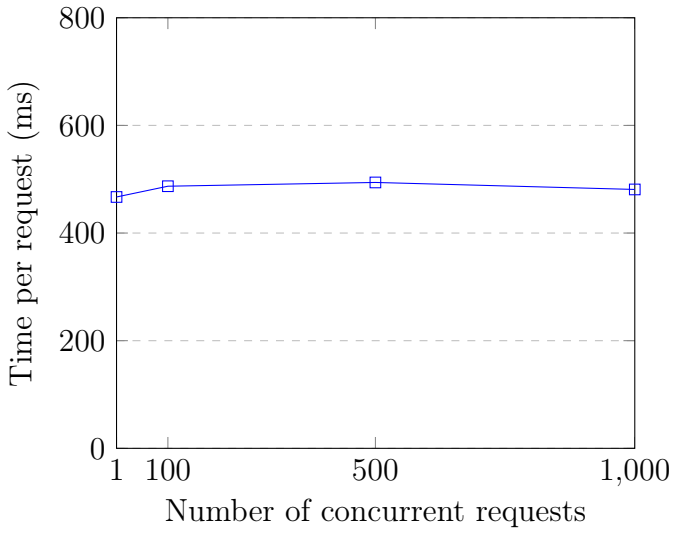

Figure 6.1: Time per requests vs. concurrency level

- Requests per second: 2.02
- Time per request: 494.924 ms
- Transfer rate (received): 30.16 kb/s
- Transfer rate (sent): 343.92 kb/s

4. Total requests: 50000. Concurrency level: 1000.

   - Requests per second: 2.08
   - Time per request: 481.421 ms
   - Transfer rate (received): 31.95 kb/s
   - Transfer rate (sent): 350.02 kb/s

Tests show that the service is not hampered by the increase of requests. Considering that a single node is able to respond to 1000 concurrent requests within a reasonable time, the platform is proved to be ready for production. By deploying multiple nodes and a load balancer, it can also be scaled to tens of thousands of concurrent users.

CHAPTER 7

RELATED WORK

As a part of the study, a research of state of the art in the topic of enhancing adverse drug event reporting has been conducted. In particular, we have focused on studies aiming at secondary use of electronic health records for better ADE detection and reporting. Some of the inspected works propose methods for automatic detection of adverse events over EHR data, some address the problems of spontaneous reporting process and facilitates it with EHR data. Each one has its own advantages and flaws. In this section, we present a review of the related work on utilization of EHR data for ADE detection and reporting.

In [21], voluntary reporting and medical chart reviews has been addressed as major problems of ADE reporting and a data mining approach for detecting ADE cases has been proposed. Building a data model containing diagnoses, drug therapies, test results and free text comments, they have utilized data in electronic health records. Since dataset has become too large to handle, they have developed aggregation methods to get a simpler representation. After extracting conditions and outcomes, they have run decision tree and association rules based learning algorithms in order to induct ADE detection rules, which have been validated by medical experts afterwards. However, some rules regarding the conditions that rarely occur or occur but not lead to an outcome has not been discovered by data mining. Although the method is innovative and provides a good alternative as a semi-automatic way to detect ADEs, being not capable of discovering new knowledge is its deficiency. Because, this leads to zero-day reactions being missed.

Another study, [57], was driven by the motivation arising from low ADE reporting rates to Food And Drug Administration (FDA) in the United States. They have proposed a system called ADE Spontaneous Triggered Event Reporting (ASTER), which automatically extracts data from EHR to directly submit ADE reports to regulatory agency. ASTER generates real-time reports when healthcare professionals discontinued a medication due an adverse event. However, before submitting reports to FDA, it requires healthcare professionals' review to provide some more information on the case. One major shortcoming of this system is that it has not been developed to be interoperable with multiple content models. As EHR sources are of disparate data models, a system specific to a particular source would have serious scalability issues. In addition to that, due its nature, the system generates high volume of reports on well-known non-serious ADE cases.

On the other hand, there exists studies with scalability and large-scale uptake concerns. In [55], an EHR-based ADE Notification System (ANS), which is not restricted to a particular clinical information system, has been proposed. ANS works on a comprehensive interoperability platform that enables analyzing heterogeneous data on patient's health records. Therefore, it has made a difference on applicability to various sources when compared to other relevant researches. ANS provides notification interfaces for clinicians. When an adverse event is detected based on EHRs, clinicians in the hospital are notified to review the case. As input is pre-populated from existing records, reporting the case is significantly less time consuming.

Up to this point, discussed researches have been engaged with secondary use of electronic health records for automatically detecting adverse drug events. Although they are promising in terms of increasing reporting rates, false positive notifications are a crucial problem of computer based detection systems. Therefore, spontaneous reporting systems still have their own advantages.

A recent study has addressed issues of spontaneous reporting systems and proposed facilitating existing systems with EHR data [23]. This research and [55] have been conducted in the scope of SALUS project [13]. Akin to approach

developed in [55], system has also included an extensive interoperability layer. Different data models in EHR sources and target data models for ADE reporting have been mapped via a common information model. This has been possible through pre-defined conversion rules on ontologies. Conversion between coded values and target value sets has been implemented using a terminology server, which coincides with our work in the thesis. According to their results, proposed methodology has been good enough to pre-populate large parts of an E2B report. Even though our approach is pretty similar, there are distinctive points. First of all, the system developed in [23] is intended for doctors, hence it works in the context of EHR systems, which eliminates the requirement for communication between SRS and EHR source. Our system, on the other hand, are open to patients as well. Secondly, ontological mapping, compared to XPath queries, brings more complicated process for non-technical personnel when integrating new EHR sources to the system. Lastly, as the system is not developed to be used by mobile reporting applications, processing time is not a critical issue, which was fundamental for us to be responsive to the end-user.

CHAPTER 8

CONCLUSIONS, DISCUSSIONS AND FUTURE WORK

For ensuring safety of pharmaceutical drugs and medical devices, post marketing surveillance has been occupying an important position for a long time. Compared to experiments and clinical trials conducted before a drug hit the market, it provides safety in a more cost-effective, evidence-based way [51]. Pre-market studies in controlled environments come short of simulating real world scene of high variability and inconsistency.

Although analysis based on secondary use of electronic health records has been becoming the direction pharmacovigilance researchers moving in, post market surveillance still largely depends on spontaneous reporting of adverse drug events. Due to many reasons such as reporting being a voluntary action and filling these reports being overwhelming, spontaneous reporting systems suffer from low reporting rates (i.e. under-reporting) and low report quality, which hamper their effectiveness severely [17, 20, 26, 32, 64].

The idea of secondary use of EHRs for post market safety studies is driven by the fact that EHR data contains (almost) complete patient history and are already available widely. The very same motivation is applicable to spontaneous reporting system, though. The information relevant to a reported ADE case, is usually available in patient's health records. So, instead of filling all fields of a long form, utilization of EHRs would help reporter ease her work by providing patient history (e.g. active/past medications, diagnoses, procedures, etc.). Facilitating spontaneous reporting systems with EHR system integration poses yet other challenges. As EHR data resides in sources with quite different content

models, capturing and processing it is an interoperability issue for computer based solutions.

Within the scope of this thesis, we have addressed the problem of under-reporting and report quality, and developed tools that is capable of extracting relevant patient context from electronic health records in the process of reporting adverse drug events through mobile mediums. Our tools are integrable to existing spontaneous reporting systems and extendable to work with multiple EHR sources. By building a standards-based and flexible architecture we have presented a solution to the main problem of EHR sources, which is heterogeneity in terms of content model.

As an initial step, We have provided integration with three EHR content models, each targeting different regions:

- Summary Care Record profile defines the content model for EHR systems in United Kingdom.
- Documents based on Consolidated Clinical Document Architecture and Continuity of Care Document templates are designated as the standard for government certified EHR systems in the United States.
- epSOS Patient Summary is developed in the scope of epSOS project, which is a large scale pilot for exchanging health records across Europe. It is integrated to our system as a potential pan-European solution.

Thanks to meta-data registry employed in our architecture, and unified data model developed, the number of EHR profiles that our tools are compatible with can be increased with minimum effort.

ADE Reports generated through our tools are also based on widely-established standards. Since spontaneous reporting systems are mainly based on ICH E2B profile, we have only provided E2B compatibility on this side. However, this can also be easily extended to include more profiles defining the content of an individual case safety report.

After all, our contribution in the course of this thesis can be summarized as follows:

- Building a pivot content model as a common ground for ICSR and EHR profiles to interoperate.
- Developing IHE Retrieve Form for Data Capture profile compliant tools that semi-automatically pre-populate ICSRs by extracting relevant information from patient's electronic health records.
- Providing mapping between E2B (R2) coded value sets and related terminology systems.
- Delivering easy to use intermediaries for extending the tools with new content models (e.g. Semantic MDR importer).
- Facilitating existing mobile spontaneous reporting tools so that it is significantly less time consuming to complete ADE reports, and submitted reports have richer content.

In this manner, we aim to increase the usability of mobile reporting tools and enhance them to generate high quality reports in high quantities.

8.1 Discussion

We are aware of the fact that EHR based solutions have some barriers for large scale uptake in real life settings; privacy concerns and universal applicability being the most obvious ones [18, 66].

Our primary obstacle in this work is the absence of an actual EHR provider. It is practically unlikely to get in contact with national authorities as EHR providers in our potential pilot sites (UK and Croatia) and convince them for such a study. Therefore, in order to address this barrier, we have demonstrated the feasibility of our implementation in a simulated environment, based on international and pan-European standards. For a large scale uptake, EHR providers (who have the data) should not be reluctant to take part in pharmacovigilance studies.

Regarding privacy concerns, there must be, for sure, solid and steady mechanisms implemented without leaving any ambiguous aspects about data exchange with EHR systems. State-of-the-art technology is capable of providing required robust authentication and authorization methods for patients to access their own health records [62, 70], which is realized in several countries through Patient Health Record (PHR) systems that are accessible as web applications. Providing access to health records for the purpose of semi-automatically filling ADE reports has no further privacy needs than those already addressed in the PHR systems.

8.2 Future Work

In the thesis, we have focused on one-way data exchange between spontaneous reporting systems and EHR sources. We have proposed tools to increase effectiveness of reporting systems with the help of EHR data. However, reverse communication for data exchange is also an promising topic. Spontaneous ADE reports can be pushed back to EHR systems so that patient's health records keep a complete medical summary. Establishing a two-way communication between SRSs and EHR systems might be a good path to follow for future studies in this field.

REFERENCES

[1] Anatomical Therapeutic Chemical Classification System (ACT). Available: http://www.whocc.no/atc/structure_and_principles/. Last visited on September 2015.

[2] epSOS Patient Summary. Available: https://decor.nictiz.nl/epsos/epsos-html-20131203T170006/tmp-1.3.6.1.4.1.12559.11.10.1.3.1.1.3-2012-05-03T000000.html. Last visited on September 2015.

[3] Health Level Seven International (HL7). Available: http://www.hl7.org. Last visited on September 2015.

[4] Integrating the Healthcare Enterprise (IHE). Available: http://www.ihe.net. Last visited on September 2015.

[5] International Conference on Harmonization of Technical Requirements. Available: http://www.ich.org. Last visited on September 2015.

[6] International Statistical Classification of Diseases (ICD). Available: http://www.who.int/classifications/icd/en/. Last visited on September 2015.

[7] Logical Observation Identifiers Names and Codes (LOINC). Available: https://loinc.org. Last visited on September 2015.

[8] MedDRA, Medical Dictionary for Regulatory Activities (MedDRA). Available: http://www.meddra.org. Last visited on September 2015.

[9] Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies (SALUS). Available: http://www.trilliumbridge.eu. Last visited on September 2015.

[10] Smart Open Services for European Patients (epSOS). Available: http://www.epsos.eu/. Last visited on September 2015.

[11] Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). Available: http://www.ihtsdo.org/snomed-ct. Last visited on September 2015.

[12] The Unified Code for Units of Measure (UCUM). Available: http://unitsofmeasure.org. Last visited on September 2015.

[13] Trillium Bridge, Bridging Patient Summaries across the Atlantic. Available: http://www.salusproject.eu. Last visited on September 2015.

[14] WEB-RADR, Recognising Adverse Drug Reactions. Available: http://web-radr.eu. Last visited on September 2015.

[15] Agency for Medicinal Products and Medical Devices. Electronic ADR reporting in E2B standard. Available: http://www.halmed.hr/?ln=en&w=farmakovigilancija&d=elektronicka_prijava_nuspojava. Last visited on September 2015.

[16] ASTM. Continuity of Care Record (CCR). Available: http://www.astm.org/Standards/E2369.htm. Last visited on September 2015.

[17] P. Bahri, P. Tsintis, and P. C. Waller. Regulatory pharmacovigilance in the eu. *Pharmacovigilance, Second Edition*, pages 185–198, 2007.

[18] A. Bakker. Access to EHR and access control at a moment in the past: A discussion of the need and an exploration of the consequences, 2004.

[19] D. W. Bates, R. S. Evans, H. Murff, P. D. Stetson, L. Pizziferri, and G. Hripcsak. Detecting adverse events using information technology. *J Am Med Inform Assoc*, 10(2):115–128, 2003.

[20] T. Brewer and G. A. Colditz. Postmarketing Surveillance and Adverse Drug Reactions: Current Perspectives and Future Needs. *JAMA*, 281(9):824–829, 1999.

[21] E. Chazard, G. Ficheur, S. Bernonville, M. Luyckx, and R. Beuscart. Data mining to generate adverse drug events detection rules. *IEEE Transactions on Information Technology in Biomedicine*, 15(6):823–830, 2011.

[22] C. Chronaki, A. Estelrich, G. Cangioli, M. Melgara, D. Kalra, Z. Gonzaga, L. Garber, E. Blechman, and J. Ferguson. Interoperability standards enabling cross-border patient summary exchange. *EHealth-For Continuity of Care: Proceedings of MIE2014*, 205:256, 2014.

[23] G. Declerck, S. Hussain, C. Daniel, M. Yuksel, G. Laleci, M. Twagirumukiza, M.-C. Jaulent, et al. Bridging data models and terminologies to support adverse drug event reporting using ehr data. *Methods Inf Med*, 54(1):24–31, 2015.

[24] Department of Health and Human Services. Medicare and medicaid programs, electronic health record incentive program, final rule. Available: http://www.gpo.gov/fdsys/pkg/FR-2010-07-28/pdf/2010-17207.pdf. Last visited on September 2015.

[25] Department of Health and Human Services. Medicare and medicaid programs; electronic health record incentive program—stage 2. Available: http://www.gpo.gov/fdsys/pkg/FR-2012-09-04/pdf/2012-21050.pdf, 2012. Last visited on September 2015.

[26] I. R. Edwards and C. Biriell. Who programme–global monitoring. *Pharmacovigilance, Second Edition*, pages 149–166, 2007.

[27] eHealth Network. Guidelines on minimum/non- exhaustive patient summary dataset for electronic exchange in accordance with the cross-border directive 2011/24/eu. Available: http://ec.europa.eu/health/ehealth/docs/guidelines_patient_summary_en.pdf, 2013. Last visited on September 2015.

[28] R. J. A. Elizabeth A. Bankert. *Institutional Review Board: Management and Function*. Jones and Bartlett, Sudbury, MA, 2 edition, 2006.

[29] Epidemico. Medwatcher. Available: https://medwatcher.org. Last visited on September 2015.

[30] R. S. Evans, S. L. Pestotnik, D. C. Classen, S. B. Bass, R. L. Menlove, R. M. Gardner, and J. P. Burke. Development of a computerized adverse drug event monitor. *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pages 23–27, 1991.

[31] Food and Drug Administration. Specifications for preparing and submitting electronic ICSRs and ICSR attachments. Available: http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM153588.pdf, 2013. Last visited on September 2015.

[32] C. D. Furberg, A. A. Levin, P. A. Gross, R. S. Shapiro, and B. L. Strom. The FDA and drug safety: a proposal for sweeping changes. *Archives of internal medicine*, 166(18):1938–1942, 2006.

[33] T. Greenhalgh, K. Stramer, T. Bratan, E. Byrne, J. Russell, H. W. Potts, et al. Adoption and non-adoption of a shared electronic summary record in england: a mixed-method case study. *Bmj*, 340, 2010.

[34] L. Hazell and S. A. W. Shakir. Under-reporting of adverse drug reactions: A systematic review, 2006.

[35] E. Heeley, J. Riley, D. Layton, L. V. Wilton, and S. A. Shakir. Prescription-event monitoring and reporting of adverse drug reactions. *Lancet*, 358(9296):1872–1873, 2001.

[36] A. Henriksson, M. Conway, M. Duneld, and W. W. Chapman. Identifying synonymy between SNOMED clinical terms of varying length using distributional analysis of electronic health records. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, 2013:600–9, 2013.

[37] HL7. Clinical Document Architecture (CDA). Available: `http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7`. Last visited on September 2015.

[38] HL7. Clinical Document Architecture (CDA) Extensibility. Available: `http://wiki.hl7.org/images/6/6a/1-4_CDA_Extensibility.docx`. Last visited on September 2015.

[39] HL7. Consolidated Clinical Document Architecture (C-CDA). Available: `http://www.hl7.org/implement/standards/product_brief.cfm?product_id=379`. Last visited on September 2015.

[40] HL7. Individual Case Safety Report (ICSR). Available: `http://www.hl7.org/implement/standards/product_brief.cfm?product_id=50`. Last visited on September 2015.

[41] HL7. Reference Information Model (RIM). Available: `http://www.hl7.org/implement/standards/rim.cfm`. Last visited on September 2015.

[42] HL7. UK Issued OIDs. Available: `http://www.hl7.org.uk/version3group/downloads/OidRootHl7UkOnly.html`. Last visited on September 2015.

[43] HL7/ASTM. Continuity of Care Document (CCD). Available: `http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6`. Last visited on September 2015.

[44] B. L. Hug, C. Keohane, D. L. Seger, C. Yoon, and D. W. Bates. The costs of adverse drug events in community hospitals. *Joint Commission journal on quality and patient safety / Joint Commission Resources*, 38(3):120–6, 2012.

[45] ICH. Maintenance of the ich guideline on clinical safety data management : Data elements for transmission of individual case safety reports e2b(r2). Available: `http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2B/Step4/E2B_R2__Guideline.pdf`, 2001. Last visited on September 2015.

[46] IHE. Patient Care Coordination (PCC) Technical Framework, Transactions and Content Modules. Available: `http://www.ihe.net/uploadedFiles/Documents/PCC/IHE_PCC_TF_Vol2.pdf`. Last visited on September 2015.

[47] IHE. Pharmacy Technical Framework Supplement. Available: ftp://ftp.ihe.net/Pharmacy/TF_Maintenance-2012/TF/IHE_Pharmacy_Suppl_PRE_Rev1.3_TI_2012-08-xx_mj.doc. Last visited on September 2015.

[48] IHE. Quality, Research and Public Health Technical Framework Supplement Data Element Exchange (DEX). Available: http://www.ihe.net/Technical_Framework/upload/IHE_QRPH_Suppl_DEX_Rev1-0_PC_2013-06-03.pdf. Last visited on September 2015.

[49] IHE. IT Infrastructure Technical Framework Supplement Retrieve Form for Data Capture (RFD). Available: http://www.ihe.net/Technical_Framework/upload/IHE_ITI_Suppl_RFD_Rev2-2_TI_2011-08-19.pdf, 2011. Last visited on September 2015.

[50] International Organization For Standardization. Information technology — Metadata registries (MDR) — Part 1: Framework, 2013.

[51] C. Ishiguro, M. Hall, G. A. Neyarapally, and G. Dal Pan. Post-market drug safety evidence sources: an analysis of fda drug safety communications. *Pharmacoepidemiology and drug safety*, 21(10):1134–1136, 2012.

[52] M. Katić, D. Soldo, Z. Ozvacić, S. Blazeković-Milaković, M. Vrcić-Keglević, B. Bergman-Marković, H. Tiljak, D. Lazić, V. C. Nekić, and G. Petricek. Information systems and the electronic health record in primary health care. *Informatics in primary care*, 15(3):187–192, 2007.

[53] P. Kierkegaard. E-Prescription across Europe. *Health and Technology*, 3(3):205–219, 2013.

[54] G. Klyne and J. J. Carroll. Resource description framework (RDF): Concepts and abstract syntax. 2006.

[55] T. Krahn, M. Eichelberg, S. Gudenkauf, G. B. Laleci Erturkmen, H. Appelrath, et al. Adverse drug event notification system: Reusing clinical patient data for semi-automatic ade detection. In *Computer-Based Medical Systems (CBMS), 2014 IEEE 27th International Symposium on*, pages 251–256. IEEE, 2014.

[56] M. Lapeyre-Mestre, S. Grolleau, and J.-L. Montastruc. Adverse drug reactions associated with the use of nsaids: a case/noncase analysis of spontaneous reports from the french pharmacovigilance database 2002–2006. *Fundamental & clinical pharmacology*, 27(2):223–230, 2013.

[57] J. A. Linder, J. S. Haas, A. Iyer, M. A. Labuzetta, M. Ibara, M. Celeste, G. Getty, and D. W. Bates. Secondary use of electronic health record data: Spontaneous triggered adverse drug event reporting. *Pharmacoepidemiology and Drug Safety*, 19(12):1211–1215, 2010.

[58] M. Lindquist. VigiBase, the WHO Global ICSR Database System: Basic Facts. *Drug Information Journal*, 42(5):409–419, 2008.

[59] Medicines and Healthcare products Regulatory Agency. Medicines, medical devices and blood regulation and safety – guidance. Available: `https://www.gov.uk/guidance/send-and-receive-information-on-adverse-drug-reactions-adrs`, 2014. Last visited on September 2015.

[60] MHRA. Reporting site for the Yellow Card Scheme. Available: `https://yellowcard.mhra.gov.uk`. Last visited on September 2015.

[61] A. Miles and S. Bechhofer. Skos simple knowledge organization system reference. *W3C recommendation*, 18:W3C, 2009.

[62] S. Narayan, M. Gagné, and R. Safavi-Naini. Privacy preserving EHR system using attribute-based infrastructure. *Proceedings of the 2010 ACM workshop on Cloud computing security workshop - CCSW '10*, page 47, 2010.

[63] NHS. Messaging Implementation Manual. Available: `https://isd.hscic.gov.uk/trud3/user/authenticated/group/0/pack/34/subpack/28/releases`. Last visited on September 2015.

[64] T. D. Piazza-Hepp and D. L. Kennedy. Reporting of adverse events to MedWatch. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists*, 52(13):1436–1439, 1995.

[65] M. Pirmohamed, A. M. Breckenridge, N. R. Kitteringham, and B. K. Park. Adverse drug reactions. *BMJ: British Medical Journal (International Edition)*, 316(7140):1295–1298, 1998.

[66] P. Ray and J. Wimalasiri. The need for technical solutions for maintaining the privacy of EHR. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pages 4686–4689, 2006.

[67] A. A. Sinaci and G. B. Laleci Erturkmen. A federated semantic metadata registry framework for enabling interoperability across clinical research and care domains. *Journal of Biomedical Informatics*, 46(5):784–794, 2013.

[68] P. K. Sinha, G. Sunder, P. Bendale, M. Mantri, and A. Dande. *Electronic health record: standards, coding systems, frameworks, and infrastructures*. John Wiley & Sons, 2012.

[69] R. Sloane, O. Osanlou, D. Lewis, D. Bollegala, S. Maskell, and M. Pirmohamed. Social media and pharmacovigilance: a review of the opportunities and challenges. *British journal of clinical pharmacology*, 2015.

[70] J. Sun, X. Zhu, C. Zhang, and Y. Fang. HCPP: Cryptography based secure EHR system for patient privacy and emergency healthcare. In *Proceedings - International Conference on Distributed Computing Systems*, pages 373–382, 2011.

[71] Therapeutic Goods Administration. Australian pharmacovigilance requirements and recommendations for medicine sponsors. Available: `http://bit.ly/tga-pharmacovigilance`, 2013. Last visited on September 2015.

[72] F. Thiessard, E. Roux, G. Miremont-Salamé, A. Fourrier-Réglat, F. Haramburu, P. Tubert-Bitter, and B. Bégaud. Trends in spontaneous adverse drug reaction reports to the French pharmacovigilance system (1986-2001). *Drug Safety*, 28(8):731–740, 2005.

[73] Trillium Bridge. Clinical model and terminology mappings: methodological approach. Available: `http://www.hl7italia.it/trillium/repository/Deliverables/FP7-SA610756-D3%201-20150210_v1.4_final.pdf`, 2014. Last visited on September 2015.

[74] A. Vallano, G. Cereza, C. Pedròs, A. Agustí, I. Danés, C. Aguilera, and J. M. Arnau. Obstacles and solutions for spontaneous reporting of adverse drug reactions in the hospital. *British Journal of Clinical Pharmacology*, 60(6):653–658, 2005.

[75] P. G. M. Van Der Heijden, E. P. Van Puijenbroek, S. Van Buuren, and J. W. Van Der Hofstede. On the assessment of adverse drug reactions from spontaneous reporting systems: The influence of under-reporting on odds ratios. *Statistics in Medicine*, 21(14):2027–2044, 2002.

[76] K. Wester, A. Jönsson, O. Spigset, and S. Hägg. Spontaneously reported fatal suspected adverse drug reactions: A 10-year survey from Sweden. *Pharmacoepidemiology and Drug Safety*, 16(2):173–180, 2007.

[77] WHO UMC. Primary reporting. Available: `https://primaryreporting.who-umc.org`. Last visited on September 2015.

[78] M. Yüksel. *A Semantic Interoperability Framework for Reinforcing Post Market Safety Studies*. PhD thesis, MIDDLE EAST TECHNICAL UNIVERSITY, 2013.

APPENDIX A

DATA ELEMENTS OF THE CONTENT MODEL

Table A.1: Data elements of the unified content model - 1

ID	Object Class	Property	Data Type
0001	CodedValue	codeSystem	string
0002	CodedValue	code	string
0003	MedDRACode	cv	CodedValue
0004	MedDRACode	oid	string
0005	VitalSign	date	string
0006	VitalSign	value	real
0007	Reporter	title	string
0008	Reporter	givenName	string
0009	Reporter	familyName	string
0010	Reporter	organization	string
0011	Reporter	street	string
0012	Reporter	city	string
0013	Reporter	state	string
0014	Reporter	postcode	string
0015	Reporter	country	string
0016	Reporter	qualification	integer
0017	Episode	MedDRAVersion	string
0018	Episode	name	MedDRACode
0019	Episode	startDateFormat	integer

Table A.2: Data elements of the unified content model - 2

ID	Object Class	Property	Data Type
0020	Episode	startDate	string
0021	Episode	continuing	integer
0022	Episode	endDateFormat	integer
0023	Episode	endDate	string
0024	Episode	comments	string
0025	DrugTherapy	name	string
0026	DrugTherapy	startDateFormat	integer
0027	DrugTherapy	startDate	string
0028	DrugTherapy	endDateFormat	integer
0029	DrugTherapy	endDate	string
0030	DrugTherapy	indicationMedDRAVersion	string
0031	DrugTherapy	indication	MedDRACode
0032	DrugTherapy	reactionMedDRAVersion	string
0033	DrugTherapy	reaction	MedDRACode
0034	Reaction	reaction	string
0035	Reaction	termHighlighted	integer
0036	Reaction	startDateFormat	integer
0037	Reaction	startDate	string
0038	Reaction	endDateFormat	integer
0039	Reaction	endDate	string
0040	Reaction	duration	integer
0041	Reaction	durationUnit	integer
0042	Reaction	outcome	integer
0043	Test	dateFormat	integer
0044	Test	date	string
0045	Test	name	string
0046	Test	result	string

Table A.3: Data elements of the unified content model - 3

ID	Object Class	Property	Data Type
0047	Test	unit	string
0048	Test	normalLowRange	string
0049	Test	normalHighRange	string
0050	Test	moreInformation	integer
0051	Drug	charactarization	integer
0052	Drug	name	string
0053	Drug	substance	string
0054	Drug	batchNumber	string
0055	Drug	dose	integer
0056	Drug	doseUnit	integer
0057	Drug	separateDose	integer
0058	Drug	doseInterval	integer
0059	Drug	doseIntervalUnit	integer
0060	Drug	cumulativeDose	integer
0061	Drug	cumlativeDoseUnit	integer
0062	Drug	doseText	string
0063	Drug	route	integer
0064	Drug	indicationMedDRAVersion	string
0065	Drug	indication	MedDRACode
0066	Drug	startDateFormat	integer
0067	Drug	startDate	string
0068	Drug	endDateFormat	integer
0069	Drug	endDate	string
0070	Drug	duration	integer
0071	Drug	durationUnit	integer
0072	Drug	actions	integer
0073	Drug	recur	integer

Table A.4: Data elements of the unified content model - 4

ID	Object Class	Property	Data Type
0074	Drug	additionalInformation	string
0075	Patient	id	string
0076	Patient	gpMedicalRecordNumber	string
0077	Patient	specialistRecordNumber	string
0078	Patient	hospitalRecordNumber	string
0079	Patient	investigationNumber	string
0080	Patient	dateOfBirthFormat	integer
0081	Patient	dateOfBirth	string
0082	Patient	age	integer
0083	Patient	ageUnit	integer
0084	Patient	weight	VitalSign
0085	Patient	height	VitalSign
0086	Patient	gender	integer
0087	Patient	medicalHistoryText	string
0088	Patient	medicalHistoryEpisode	Episode
0089	Patient	pastDrugTherapy	DrugTherapy
0090	Patient	reactions	Reaction
0091	Patient	tests	Test
0092	Patient	drugs	Drug
0093	Patient	reporterComments	string
0094	Receiver	receivertype	integer
0095	Sender	sendertype	integer
0096	Sender	senderorganization	string
0097	SafetyReport	country	string
0098	SafetyReport	serious	integer
0099	SafetyReport	death	integer
0100	SafetyReport	lifeThreatening	integer

Table A.5: Data elements of the unified content model - 5

ID	Object Class	Property	Data Type
0101	SafetyReport	hospitalization	integer
0102	SafetyReport	disabling	integer
0103	SafetyReport	congenitalAnomali	integer
0104	SafetyReport	other	integer
0105	SafetyReport	receiptdateformat	integer
0106	SafetyReport	receiptdate	string
0107	SafetyReport	primarySource	Reporter
0108	SafetyReport	sender	Sender
0109	SafetyReport	receiver	Receiver
0110	SafetyReport	patient	Patient
0111	ICHICSR	reports	SafetyReport
0112	ICHICSR	lang	string

APPENDIX B

AN EXAMPLE PRE-POPULATED E2B DOCUMENT

```
<ichicsr lang="en-US">
  <ichicsrmessageheader>
    <messagetype>ichicsr</messagetype>
    <messageformatversion>2.1</messageformatversion>
    <messageformatrelease>1.0</messageformatrelease>
    <messagenumb>0000</messagenumb>
    <messagesenderidentifier>WEB-RADR</messagesenderidentifier>
    <messagereceiveridentifier>MEDWATCHER</messagereceiveridentifier>
    <messagedateformat>204</messagedateformat>
    <messagedate>20150305132136</messagedate>
  </ichicsrmessageheader>
  <safetyreport>
    <safetyreportid>
      IT-SRDC-59891470955123915655239175118973013287
    </safetyreportid>
    <receiptdateformat>102</receiptdateformat>
    <receiptdate>20121105</receiptdate>
    <primarysource>
      <reportertitle>Dr.</reportertitle>
      <reportergivename>Ersilia</reportergivename>
      <reporterfamilyname>Palladino</reporterfamilyname>
      <reporterorganization>Good Health Clinic</reporterorganization>
      <reporterstreet>Via Marostica, No 33/3</reporterstreet>
      <reportercity>Milan</reportercity>
      <reporterstate>MI</reporterstate>
      <reporterpostcode>20146</reporterpostcode>
      <reportercountry>IT</reportercountry>
      <qualification>1</qualification>
    </primarysource>
    <primarysource>
      <reportergivename>Sabina</reportergivename>
      <reporterfamilyname>Cremona</reporterfamilyname>
      <reporterorganization>Good Health Clinic</reporterorganization>
```

```
  <reporterstreet>Via Lago di Nemi, No 4</reporterstreet>
  <reportercity>Milan</reportercity>
  <reporterstate>MI</reporterstate>
  <reporterpostcode>20142</reporterpostcode>
  <reportercountry>IT</reportercountry>
  <qualification>5</qualification>
</primarysource>
<sender>
  <sendertype>6</sendertype>
  <senderorganization>WEB-RADR</senderorganization>
</sender>
<receiver>
  <receivertype>4</receivertype>
</receiver>
<patient>
  <patientinitial>SC</patientinitial>
  <patientgpmedicalrecordnumb>54321678901</patientgpmedicalrecordnumb>
  <patientbirthdateformat>102</patientbirthdateformat>
  <patientbirthdate>19560708</patientbirthdate>
  <patientweight>88</patientweight>
  <patientheight>177</patientheight>
  <patientsex>1</patientsex>
  <patientmedicalhistorytext/>
  <medicalhistoryepisode>
    <patientepisodenamemeddraversion>
      13.0
    </patientepisodenamemeddraversion>
    <patientepisodename>10000891</patientepisodename>
    <patientmedicalstartdateformat>102</patientmedicalstartdateformat>
    <patientmedicalstartdate>20090401</patientmedicalstartdate>
    <patientmedicalenddateformat>102</patientmedicalenddateformat>
    <patientmedicalenddate>20090401</patientmedicalenddate>
    <patientmedicalcomment>Acute myocardial infarction, of anterolateral
    ↪  wall</patientmedicalcomment>
  </medicalhistoryepisode>
  <medicalhistoryepisode>
    <patientepisodenamemeddraversion>
      13.0
    </patientepisodenamemeddraversion>
    <patientepisodename>10003553</patientepisodename>
    <patientmedicalstartdateformat>102</patientmedicalstartdateformat>
    <patientmedicalstartdate>20030801</patientmedicalstartdate>
    <patientmedicalcomment>Asthma</patientmedicalcomment>
  </medicalhistoryepisode>
  <medicalhistoryepisode>
    <patientepisodenamemeddraversion>
```

```
    13.0
  </patientepisodenamemeddraversion>
  <patientepisodename>10051060</patientepisodename>
  <patientmedicalstartdateformat>102</patientmedicalstartdateformat>
  <patientmedicalstartdate>20080701</patientmedicalstartdate>
  <patientmedicalcomment>Total replacement of
   ↪  hip</patientmedicalcomment>
</medicalhistoryepisode>
<medicalhistoryepisode>
  <patientepisodenamemeddraversion>
    13.0
  </patientepisodenamemeddraversion>
  <patientepisodename>10055048</patientepisodename>
  <patientmedicalstartdateformat>102</patientmedicalstartdateformat>
  <patientmedicalstartdate>20080624</patientmedicalstartdate>
  <patientmedicalcomment>Allergy to substance, penicillins with
   ↪  extended spectrum</patientmedicalcomment>
</medicalhistoryepisode>
<medicalhistoryepisode>
  <patientepisodenamemeddraversion>
    13.0
  </patientepisodenamemeddraversion>
  <patientepisodename>10061958</patientepisodename>
  <patientmedicalstartdateformat>102</patientmedicalstartdateformat>
  <patientmedicalstartdate>20090624</patientmedicalstartdate>
  <patientmedicalenddateformat>102</patientmedicalenddateformat>
  <patientmedicalenddate>20100301</patientmedicalenddate>
  <patientmedicalcomment>Food intolerance, egg
   ↪  protein</patientmedicalcomment>
</medicalhistoryepisode>
<patientpastdrugtherapy>
  <patientdrugname>Albuterol 1 MG/ML Inhalant Solution
   ↪  [Ventolin]</patientdrugname>
  <patientdrugstartdateformat>102</patientdrugstartdateformat>
  <patientdrugstartdate>20080201</patientdrugstartdate>
  <patientdrugenddateformat>102</patientdrugenddateformat>
  <patientdrugenddate>20080210</patientdrugenddate>
  <patientindicationmeddraversion>
    13.0
  </patientindicationmeddraversion>
  <patientdrugindication>10003553</patientdrugindication>
  <patientdrgreactionmeddraversion>
    13.0
  </patientdrgreactionmeddraversion>
  <patientdrugreaction>10047924</patientdrugreaction>
</patientpastdrugtherapy>
```

```
<patientpastdrugtherapy>
  <patientdrugname>Hepatitis B virus vaccine</patientdrugname>
  <patientdrugstartdateformat>102</patientdrugstartdateformat>
  <patientdrugstartdate>20010303</patientdrugstartdate>
  <patientdrugenddateformat>102</patientdrugenddateformat>
  <patientdrugenddate>20080210</patientdrugenddate>
  <patientindicationmeddraversion>
    13.0
  </patientindicationmeddraversion>
  <patientdrugindication>10003553</patientdrugindication>
  <patientdrgreactionmeddraversion>
    13.0
  </patientdrgreactionmeddraversion>
  <patientdrugreaction>10037844</patientdrugreaction>
</patientpastdrugtherapy>
<reaction>
  <primarysourcereaction>Wheezing</primarysourcereaction>
  <reactionstartdateformat>102</reactionstartdateformat>
  <reactionstartdate>20080201</reactionstartdate>
  <reactionenddateformat>102</reactionenddateformat>
  <reactionenddate>20080210</reactionenddate>
</reaction>
<reaction>
  <primarysourcereaction>Weal</primarysourcereaction>
  <reactionstartdateformat>102</reactionstartdateformat>
  <reactionstartdate>20010303</reactionstartdate>
  <reactionenddateformat>102</reactionenddateformat>
  <reactionenddate>20010303</reactionenddate>
</reaction>
<test>
  <testdateformat>102</testdateformat>
  <testdate>20100407</testdate>
  <testname>Hemoglobin:Mass Concentration:Point in time:Blood
   ↪  arterial:Quantitative</testname>
  <testresult>13.2</testresult>
  <testunit>g/dl</testunit>
  <lowtestrange>12</lowtestrange>
  <hightestrange>16</hightestrange>
</test>
<test>
  <testdateformat>102</testdateformat>
  <testdate>20100510</testdate>
  <testname>Bicarbonate:Substance Concentration:Point in
   ↪  time:Serum:Quantitative</testname>
  <testresult>35</testresult>
  <testunit>meq/l</testunit>
```

```
        <lowtestrange>18</lowtestrange>
        <hightestrange>23</hightestrange>
      </test>
      <drug>
        <medicinalproduct>Albuterol 1 MG/ML Inhalant Solution
        ↪  [Ventolin]</medicinalproduct>
        <drugstructuredosagenumb>5</drugstructuredosagenumb>
        <drugstructuredosageunit>012</drugstructuredosageunit>
        <drugintervaldosageunitnumb>12</drugintervaldosageunitnumb>
        <drugintervaldosagedefinition>802</drugintervaldosagedefinition>
        <drugadministrationroute/>
        <drugindicationmeddraversion>13.0</drugindicationmeddraversion>
        <drugindication>10003553</drugindication>
        <drugstartdateformat>102</drugstartdateformat>
        <drugstartdate>20080201</drugstartdate>
        <drugenddateformat>102</drugenddateformat>
        <drugenddate>20080210</drugenddate>
      </drug>
      <drug>
        <medicinalproduct>Hepatitis B virus vaccine</medicinalproduct>
        <drugstructuredosagenumb>1</drugstructuredosagenumb>
        <drugstructuredosageunit>026</drugstructuredosageunit>
        <drugintervaldosageunitnumb>12</drugintervaldosageunitnumb>
        <drugintervaldosagedefinition>802</drugintervaldosagedefinition>
        <drugadministrationroute/>
        <drugindicationmeddraversion>13.0</drugindicationmeddraversion>
        <drugindication>10003553</drugindication>
        <drugstartdateformat>102</drugstartdateformat>
        <drugstartdate>20010303</drugstartdate>
        <drugenddateformat>102</drugenddateformat>
        <drugenddate>20080210</drugenddate>
      </drug>
    </patient>
  </safetyreport>
</ichicsr>
```

APPENDIX C

SEMANTIC DEFINITION OF E2B (R2) DOSE UNIT

```
# baseURI: http://ich.org/products/electronic-standards/E2BR2/DoseUnit

@prefix :
 ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit/> .
@prefix foaf:    <http://xmlns.com/foaf/0.1/> .
@prefix iso:     <uri:iso.org:9834#> .
@prefix owl:     <http://www.w3.org/2002/07/owl#> .
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos:    <http://www.w3.org/2004/02/skos/core#> .
@prefix E2B:
 ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit/> .
@prefix xsd:     <http://www.w3.org/2001/XMLSchema#> .

<http://ich.org/products/electronic-standards/E2BR2/DoseUnit>
    rdf:type skos:ConceptScheme ;
    rdfs:label "DoseUnit" ;
    foaf:name "E2B(R2) Dose Unit" ;
    iso:oid "2.16.840.1.113883.3.989.2.1.1" .

E2B:001
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "001" ;
    skos:prefLabel "kg" ;
    skos:label "kilogram(s)" ;
    rdfs:label "kilogram(s)" ;
    skos:label "kg" ;
    rdfs:label "kg" .

E2B:002
    rdf:type skos:Concept ;
```

```
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "002" ;
    skos:prefLabel "G" ;
    skos:label "gram(s)" ;
    rdfs:label "gram(s)" ;
    skos:label "G" ;
    rdfs:label "G" .

E2B:003
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "003" ;
    skos:prefLabel "Mg" ;
    skos:label "milligram(s)" ;
    rdfs:label "milligram(s)" ;
    skos:label "Mg" ;
    rdfs:label "Mg" .

E2B:004
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "004" ;
    skos:prefLabel "g" ;
    skos:label "microgram(s)" ;
    rdfs:label "microgram(s)" ;
    skos:label "g" ;
    rdfs:label "g" .

E2B:005
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "005" ;
    skos:prefLabel "ng" ;
    skos:label "nanogram(s)" ;
    rdfs:label "nanogram(s)" ;
    skos:label "ng" ;
    rdfs:label "ng" .

E2B:006
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
```

```
    skos:notation "006" ;
    skos:prefLabel "pg" ;
    skos:label "picogram(s)" ;
    rdfs:label "picogram(s)" ;
    skos:label "pg" ;
    rdfs:label "pg" .

E2B:007
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "007" ;
    skos:prefLabel "mg/kg" ;
    skos:label "milligram(s)/kilogram" ;
    rdfs:label "milligram(s)/kilogram" ;
    skos:label "mg/kg" ;
    rdfs:label "mg/kg" .

E2B:008
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "008" ;
    skos:prefLabel "g/kg" ;
    skos:label "microgram(s)/kilogram" ;
    rdfs:label "microgram(s)/kilogram" ;
    skos:label "g/kg" ;
    rdfs:label "g/kg" .

E2B:009
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "009" ;
    skos:prefLabel "mg/m2" ;
    skos:label "milligram(s)/sq. meter" ;
    rdfs:label "milligram(s)/sq. meter" ;
    skos:label "mg/m2" ;
    rdfs:label "mg/m2" .

E2B:010
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "010" ;
    skos:prefLabel "g/m2" ;
```

```
    skos:label "microgram(s)/sq. meter" ;
    rdfs:label "microgram(s)/sq. meter" ;
    skos:label "g/m2" ;
    rdfs:label "g/m2" .

E2B:011
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "011" ;
    skos:prefLabel "l" ;
    skos:label "litre(s)" ;
    rdfs:label "litre(s)" ;
    skos:label "l" ;
    rdfs:label "l" .

E2B:012
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "012" ;
    skos:prefLabel "ml" ;
    skos:label "millilitre(s)" ;
    rdfs:label "millilitre(s)" ;
    skos:label "ml" ;
    rdfs:label "ml" .

E2B:013
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "013" ;
    skos:prefLabel "l" ;
    skos:label "microlitre(s)" ;
    rdfs:label "microlitre(s)" ;
    skos:label "l" ;
    rdfs:label "l" .

E2B:014
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "014" ;
    skos:prefLabel "Bq" ;
    skos:label "becquerel(s)" ;
    rdfs:label "becquerel(s)" ;
```

```
    skos:label "Bq" ;
    rdfs:label "Bq" .

E2B:015
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "015" ;
    skos:prefLabel "GBq" ;
    skos:label "gigabecquerel(s)" ;
    rdfs:label "gigabecquerel(s)" ;
    skos:label "GBq" ;
    rdfs:label "GBq" .

E2B:016
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "016" ;
    skos:prefLabel "MBq" ;
    skos:label "megabecquerel(s)" ;
    rdfs:label "megabecquerel(s)" ;
    skos:label "MBq" ;
    rdfs:label "MBq" .

E2B:017
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "017" ;
    skos:prefLabel "Kbq" ;
    skos:label "kilobecquerel(s)" ;
    rdfs:label "kilobecquerel(s)" ;
    skos:label "Kbq" ;
    rdfs:label "Kbq" .

E2B:018
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "018" ;
    skos:prefLabel "Ci" ;
    skos:label "curie(s)" ;
    rdfs:label "curie(s)" ;
    skos:label "Ci" ;
    rdfs:label "Ci" .
```

```
E2B:019
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "019" ;
    skos:prefLabel "MCi" ;
    skos:label "millicurie(s)" ;
    rdfs:label "millicurie(s)" ;
    skos:label "MCi" ;
    rdfs:label "MCi" .

E2B:020
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "020" ;
    skos:prefLabel "Ci" ;
    skos:label "microcurie(s)" ;
    rdfs:label "microcurie(s)" ;
    skos:label "Ci" ;
    rdfs:label "Ci" .

E2B:021
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "021" ;
    skos:prefLabel "NCi" ;
    skos:label "nanocurie(s)" ;
    rdfs:label "nanocurie(s)" ;
    skos:label "NCi" ;
    rdfs:label "NCi" .

E2B:022
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "022" ;
    skos:prefLabel "Mol" ;
    skos:label "mole(s)" ;
    rdfs:label "mole(s)" ;
    skos:label "Mol" ;
    rdfs:label "Mol" .

E2B:023
```

```
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "023" ;
    skos:prefLabel "Mmol" ;
    skos:label "millimole(s)" ;
    rdfs:label "millimole(s)" ;
    skos:label "Mmol" ;
    rdfs:label "Mmol" .

E2B:024
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "024" ;
    skos:prefLabel "mol" ;
    skos:label "micromole(s)" ;
    rdfs:label "micromole(s)" ;
    skos:label "mol" ;
    rdfs:label "mol" .

E2B:025
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "025" ;
    skos:prefLabel "Iu" ;
    skos:label "international unit(s)" ;
    rdfs:label "international unit(s)" ;
    skos:label "Iu" ;
    rdfs:label "Iu" .

E2B:026
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "026" ;
    skos:prefLabel "Kiu" ;
    skos:label "iu(1000s)" ;
    rdfs:label "iu(1000s)" ;
    skos:label "Kiu" ;
    rdfs:label "Kiu" .

E2B:027
    rdf:type skos:Concept ;
```

```
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "027" ;
    skos:prefLabel "Miu" ;
    skos:label "iu(1,000,000s)" ;
    rdfs:label "iu(1,000,000s)" ;
    skos:label "Miu" ;
    rdfs:label "Miu" .

E2B:028
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "028" ;
    skos:prefLabel "iu/kg" ;
    skos:label "iu/kilogram" ;
    rdfs:label "iu/kilogram" ;
    skos:label "iu/kg" ;
    rdfs:label "iu/kg" .

E2B:029
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "029" ;
    skos:prefLabel "Meq" ;
    skos:label "milliequivalent(s)" ;
    rdfs:label "milliequivalent(s)" ;
    skos:label "Meq" ;
    rdfs:label "Meq" .

E2B:030
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "030";
    skos:prefLabel "%" ;
    skos:label "percent" ;
    rdfs:label "percent" ;
    skos:label "%" ;
    rdfs:label "%" .

E2B:031
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
```

```
    skos:notation "031" ;
    skos:prefLabel "Gtt" ;
    skos:label "drop(s)" ;
    rdfs:label "drop(s)" ;
    skos:label "Gtt" ;
    rdfs:label "Gtt" .

E2B:032
    rdf:type skos:Concept ;
    skos:inScheme
     ↪  <http://ich.org/products/electronic-standards/E2BR2/DoseUnit> ;
    skos:notation "032" ;
    skos:prefLabel "DF" ;
    skos:label "dosage form" ;
    rdfs:label "dosage form" ;
    skos:label "DF" ;
    rdfs:label "DF" .
```