

Aus dem Institut für Medizinische Biometrie und Informatik

Direktor: Prof. Dr. sc. hum. Meinhard Kieser

Sektion Medizinische Informatik

Leiter: Prof. Dr.-Ing. Hartmut Dickhaus

**Patientenübergreifende, multiple Verwendung
von Patientendaten für die klinische Forschung
unter Nutzung von Archetypen**

Inauguraldissertation

zur Erlangung des Doctor scientiarum humanarum

an der

Medizinischen Fakultät Heidelberg

der

Ruprecht-Karls-Universität

vorgelegt von

Christian Dominik Kohl

geboren in

Heilbronn-Neckargartach

2012

Dekan: Herr Prof. Dr. med. Claus R. Bartram
Doktormutter: Frau Prof. Dr. sc. hum. Petra Knaup-Gregori

für Tina

IN SCIENTIA SALUS

Heilung durch Wissenschaft

Vincenz Czerny (1842 - 1916)

Inhaltsverzeichnis

Abbildungsverzeichnis	X
Tabellenverzeichnis	XIII
Abkürzungsverzeichnis und Glossar	XV
1 Einleitung	1
1.1 Gegenstand und Motivation.....	1
1.2 Problemstellung	2
1.3 Zielsetzung.....	3
2 Grundlagen	5
2.1 Begriffsdefinitionen	5
2.2 Klinische Studien.....	6
2.3 Klassifikation von Daten	7
2.4 Semantische und syntaktische Interoperabilität.....	8
2.5 Zwei-Ebenen-Modell zur Repräsentation von Gesundheitsdaten.....	9
2.6 <i>openEHR</i>	11
2.7 Bestehende Ansätze zur multiplen Nutzung von Daten aus der Routineversorgung.....	14
3 Vorgehen und Methodik	19
3.1 Generisches Vorgehensmodell für die multiple Nutzung von Daten in Forschung und Versorgung.....	19
3.2 Fallstudie.....	21
3.3 Prototyp.....	23
3.4 Archetypen zur Dokumentation von und in klinischen Studien	24
3.5 Abbildung auf Standards	26
3.6 Zusammenfassung des Vorgehens.....	27
4 Ergebnisse	29
4.1 Fallstudie.....	29
4.2 Prototyp.....	35
4.3 Referenzarchitektur.....	39
4.3.1 <i>Anforderungen an die Referenzarchitektur</i>	39
4.3.2 <i>Spezifikation der Referenzarchitektur</i>	41
4.3.3 <i>Vorgehen gemäß Referenzarchitektur</i>	46
4.4 Archetypen zur Dokumentation von und in klinischen Studien	47
4.4.1 <i>Entwicklung von Archetypen für Studiendaten</i>	47
4.4.2 <i>Entwicklung von Archetypen für Studien-Metadaten</i>	54
4.5 Abbildung auf bestehende Standards.....	59
4.6 Zusammenfassung der Ergebnisse.....	62

5	Diskussion	63
5.1	Diskussion von Vorgehen und Methodik	63
5.2	Diskussion der Ergebnisse.....	68
5.2.1	<i>Fallstudie</i>	68
5.2.2	<i>Prototyp</i>	71
5.2.3	<i>Referenzarchitektur</i>	71
5.2.4	<i>Archetypen für Studiendaten</i>	75
5.2.5	<i>Archetypen für Studien-Metadaten</i>	77
5.2.6	<i>Abbildung auf Standards</i>	80
6	Fazit und Ausblick	83
7	Zusammenfassung	85
	Literaturverzeichnis.....	87
	Eigene Veröffentlichungen	95
	Anhang.....	97
1.	Abbildung CDASH auf <i>openEHR</i> -Archetypen.....	98
2.	Abbildung DRKS und WHO/ICTRP Merkmalsarten zu Beschreibung einer Studie auf <i>openEHR</i> -Archetypen	104
3.	Abbildung <i>openEHR</i> -Archetypen auf BRIDG-Modell	126
4.	Beispiel: Semantische Annotation von ODM-Elementen mittels Archetypen	149
5.	<i>openSDMS</i> : Prozessabläufe	163
	Lebenslauf.....	165
	Danksagungen	167

Abbildungsverzeichnis

Abb. 2-1:	Merkmalsarten, welche das Konzept ‚Blutdruck‘ beschreiben - dargestellt als Mindmap des openEHR-Archetyps ‚Blood Pressure‘.	10
Abb. 2-2:	Beziehungen zwischen <i>openEHR</i> -Templates, -Archetypen und -Referenzmodell sowie weiteren Elementen.	13
Abb. 2-3:	AQL-Abfrage zur Ermittlung der aktuellen Medikation eines Patienten.	13
Abb. 3-1:	Generisches Vorgehensmodell für die multiple Nutzung von Daten in Forschung und Versorgung.	19
Abb. 3-2:	Darstellung der für Präsentation, Geschäftslogik und Persistenz in Opereffa genutzten Technologien.	24
Abb. 4-1:	Ablauf der Datenerhebung in der analysierten Studie.	29
Abb. 4-2:	Verteilung der Ausprägungstypen der 1452 im KIS dokumentierbaren Merkmalsarten.	30
Abb. 4-3:	Verteilung der Ausprägungstypen der 200 Merkmalsarten der cTAG-Studie.	30
Abb. 4-4:	Ergebnis der Analyse, welcher Anteil der 200 in der cTAG-Studie benötigten Merkmalsarten im KIS dokumentiert werden kann.	33
Abb. 4-5:	Verteilung der notwendigen Transformationen, falls alle 70 der im KIS vorhandenen und für die cTAG Studie relevanten Merkmalsarten im KIS elektronisch vorlägen.	33
Abb. 4-6:	Verteilung der 70 Merkmalsarten, die für die cTAG-Studie multipel genutzt werden könnten, auf die elektronischen und konventionellen Subsysteme des KIS.	34
Abb. 4-7:	Ansicht ‚Patientenakte‘ von openSDMS.	36
Abb. 4-8:	Ansicht ‚Studie / CRF‘ von openSDMS.	37
Abb. 4-9:	Ansicht ‚Studie / Metadaten‘ von openSDMS.	37
Abb. 4-10:	Datenflüsse bei der Nutzung von openSDMS als Integrationssystem.	39
Abb. 4-11:	Auf Archetypen basierende Referenzarchitektur für die multiple Nutzung von KIS-Daten in klinischen Studien.	42
Abb. 4-12:	Beispiel für die Nutzung des ODM-Alias-Mechanismus zur semantischen Annotation mit Hilfe eines Archetypen.	43
Abb. 4-13:	Archetypen mit Relevanz bei der Dokumentation von Medikationsangaben.	48
Abb. 4-14:	AQL-Abfrage zur Ermittlung der ersten Gabe einer bestimmten Medikation.	50
Abb. 4-15:	Definition des CDASH CM-Templates im Ocean Template Designer.	52
Abb. 4-16:	HTML-Eingabemaske generiert aus dem openEHR CDASH CM-Template.	53

Abb. 4-17:	Generische Struktur einer klinischen Studie.....	54
Abb. 4-18:	Neu erstellte Archetypen für Studien-Metadaten.	55
Abb. 4-19:	Verzeichnisstruktur für Studiendaten und Studienmetadaten.....	56
Abb. 4-20:	Abbildung von Studien-Metadaten mittels COMPOSITION und ADMIN_ENTRY-Archetypen in die generische Verzeichnisstruktur.....	57
Abb. 4-21:	AQL-Abfrage zur Ermittlung, ob ein Patient die Einschlusskriterien einer Studie erfüllt.....	58
Abb. 4-22:	Beispiel für die multiple Nutzung des Alias-Mechanismus zur semantischen Annotation eines ODM-Elements.	59
Abb. 5-1:	Mögliche Varianten bei der Abbildung von Studien-Metadaten mittels Archetypen.....	79

Tabellenverzeichnis

Tab. 4-1:	Übersicht der in der cTAG-Studie verwendeten eCRF und der vorgesehenen Visiten.....	32
Tab. 4-2:	Mögliche Abbildungen von CDASH CM-Datenelementen auf openEHR-Archetypen.....	49
Tab. 4-3:	Gegenüberstellung und Zuordnung der Definition der Merkmalsart ‚primärer Endpunkt‘ einer Studie in den Spezifikationen von DRKS, WHO, openEHR und BRIDG-Modell.	61

Abkürzungsverzeichnis und Glossar

ADL	Archetype Definition Language Die ADL ermöglicht es, Archetypen in Textform zu definieren und mit Hilfe von Parsern maschinell zu verarbeiten (siehe Kapitel 2.6).
AJaX	Asynchronous JavaScript and XML AJaX ist eine Möglichkeit, um asynchrone Daten zwischen einem Web-Browser und –Server zu übertragen.
AOM	Archetype Object Model Das AOM ist ein Meta-Modell, um Archetypen zu beschreiben. Das AOM ist das objektorientierte Äquivalent zur ADL (siehe Kapitel 2.6).
AQL	Archetype Query Language (früher EQL – EHR Query Language) Abfragesprache für Daten, die in auf den <i>openEHR</i> -Spezifikationen basierenden Gesundheitsakten gespeichert sind (siehe Kapitel 2.6).
BMBF	Bundesministerium für Bildung und Forschung Das BMBF ist unter anderem zuständig für die Förderung der Forschung in allen wissenschaftlichen Bereichen (http://www.bmbf.de/).
BRIDG	Biomedical Research Integrated Domain Group Model - siehe Kapitel 2.7 -
caBIG	cancer Biomedical Informatics Grid - siehe Kapitel 2.7 -
CDASH	Clinical Data Acquisition Standards Harmonization - siehe Kapitel 2.7 -
CDISC	Clinical Data Interchange Standards Consortium Das CDISC entwickelt offene Standards für den Austausch von Daten im Bereich der klinischen Forschung (http://www.cdisc.org/).
CM	Prior and Concomitant Medications (dt.: Erfassung von früherer und Begleitmedikationen) Mit CM wird in dieser Arbeit die CDASH Domänentabelle ‚Prior and Concomitant Medications‘ referenziert.
CIMI	Clinical Information Modeling Initiative Im Rahmen der CIMI arbeiten verschiedene Organisationen (unter anderem CDISC, HL7 und die <i>openEHR</i> -Foundation) auf internationaler Ebene zusammen, um die Interoperabilität im Gesundheitswesen zu verbessern.
CRF	Case Report Form (dt.: Datenerhebungsbogen) Papierbasiertes Formular zur Dokumentation von Beobachtungen in klinischen Studien.
cTAG	conformable Thoracic Stent Grafts Studienname: Details siehe Kapitel 3.2.
DRKS	Deutsches Register Klinischer Studien Das DRKS ermöglicht die Registrierung und Recherche von in Deutschland durchgeführten klinischen Studien (http://www.germanctr.de/).

EAV	Entity-Value-Attribute Das EAV-Modell ist ein Datenmodell, bei welchem aus Datenbanksicht alle Merkmale in einer einzigen Tabelle abgelegt werden. Dabei werden in jeder Zeile dieser Tabelle die Bezeichnung der Merkmalsart und die zugehörige Merkmalsausprägung gespeichert.
eCRF	electronic Case Report Form Elektronischer CRF – siehe ‚CRF‘.
EGA	Elektronische Gesundheitsakte Eine EGA „soll verteilt bei Leistungserbringern und Patienten anfallende klinische und gesundheitsbezogene Daten eines Menschen zusammenfassen und diese omnipräsent, lebenslang, unabhängig von Ort und Zeit allen am Behandlungsprozess Beteiligten (inkl. der Patienten!) bedarfsgerecht präsentieren.“ ([WARDA 2006], S. 374).
EMA	European Medicines Agency (dt.: Europäische Arzneimittelagentur) Die EMA ist zuständig für die Zulassung und Überwachung von Arzneimitteln in der Europäischen Union (http://www.ema.europa.eu/).
EPA	Elektronische Patientenakte Eine EPA ist die computerbasierte Variante einer Patientenakte.
FDA	Food and Drug Administration Die FDA ist die für die Überwachung von Lebensmitteln und die Zulassung von Arzneimitteln zuständige Behörde in den Vereinigten Staaten von Amerika (http://www.fda.gov/).
GCP	Good Clinical Practice GCP bezeichnet international anerkannte, nach ethischen und wissenschaftlichen Gesichtspunkten aufgestellte Regeln für die Durchführung klinischer Studien.
HIMSS	Healthcare Information and Management Systems Society (dt.: Gesellschaft für Gesundheitsinformations- und Managementsysteme) Die HIMSS ist eine gemeinnützige Organisation, die darauf ausgerichtet ist, die Gesundheitsversorgung durch den sinnvollen Einsatz von Informationstechnologie- und Managementsystemen zu verbessern (http://www.himss.org).
HL7	Health Level 7 „Der speziell für das Gesundheitswesen entwickelte internationale Kommunikationsstandard HL7 ... ermöglicht die Kommunikation und Kooperation zwischen nahezu allen Institutionen und Bereichen des Gesundheitswesens“ ([NORGALL 2009]) (http://www.hl7.org/).
HL7 RCRIM TC	HL7 Regulated Clinical Research Information Management Technical Committee Ziel des Komitees ist die Entwicklung von Standards, um das Informationsmanagement in der klinischen Forschung zu verbessern.
HL7 RIM	HL7 Reference Information Model Das HL7 RIM ist ein Referenz-Informationsmodell, welches das Gesundheitswesen in seiner gesamten Breite abzudecken versucht und den Grundstein des HL7 Ansatzes in der Version 3 bildet.

-
- HTML** **Hypertext Markup Language**
(dt.: **Hypertext-Auszeichnungssprache**)
Die HTML ist eine Auszeichnungssprache zur Definition von Dokumenten, welche von Web-Browsern dargestellt werden können.
- i2b2** **Informatics for Integrating Biology and Bedside**
- siehe Kapitel 2.7 -
- ICD** **International Statistical Classification of Diseases and Related Health Problems (dt.: Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme)**
Die ICD ist ein von der Weltgesundheitsorganisation herausgegebenes Klassifikationssystem für Diagnosen. In Deutschland wird aktuell die Version ICD-10-GM Version 2010 eingesetzt.
- ICH** **International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use**
Die ICH strebt die Harmonisierung von Kriterien zur Beurteilung von Arzneimitteln im Rahmen von Arzneimittelzulassungen in Europa, den USA und Japan an.
- ICTRP** **International Clinical Trials Registry Platform**
Die ICTRP ermöglicht die Registrierung und Recherche von klinischen Studien auf internationaler Ebene (<http://www.who.int/ictrp/>).
- IHE** **Integrating the Healthcare Enterprise**
Im Rahmen von IHE arbeiten Anwender und Hersteller gemeinsam an der Harmonisierung des Datenaustauschs zwischen computergestützten Systemen im Gesundheitswesen (<http://www.ihe.net/>).
- ISO** **International Organization for Standardization**
(dt.: **Internationale Organisation für Normung**)
Die ISO erarbeitet als internationale Vereinigung von Normungsorganisationen Normen in zahlreichen Bereichen, mit dem Ziel dass diese von den einzelnen Mitgliedsorganisationen unverändert übernommen werden (<http://www.iso.org>).
- KIS** **Krankenhausinformationssystem**
„Ein Krankenhausinformationssystem ist das Informationssystem der Einrichtung ‚Krankenhaus‘“ ([AMMENWERTH und HAUX 2005], S. 8).
Ein Informationssystem kann sich sowohl aus rechnerunterstützten als auch aus konventionellen Subinformationssystemen zusammensetzen. Ein typisches rechnerunterstütztes Subsystem eines KIS ist die EPA.
- KKS** **Koordinierungszentrum für Klinische Studien**
Die KKS sind Einrichtungen, deren Aufbau vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wurden. Ziel war es, an Universitätsklinika die personellen und logistischen Ressourcen zur Verfügung zu stellen, um Planung, Durchführung und Auswertung von klinischen Studien nach international anerkannten Qualitätsstandards zu unterstützen.

NCI	National Cancer Institute Das NCI ist das Krebsforschungszentrum in den Vereinigten Staaten von Amerika, welches zu den staatlichen National Institutes of Health gehört (http://www.cancer.gov/).
ODM	Operational Data Model - siehe Kapitel 2.7 -
openEHR	open Electronic Health Record - siehe Kapitel 2.6 -
Opereffa	openEHR Reference Framework and Application - siehe Kapitel 3.3 -
openSDMS	open Study Data Management System Im Rahmen dieser Arbeit entwickeltes integriertes System, das die Nutzung von Einträgen in Patientenakten in klinischen Studien erlaubt (siehe Kapitel 4.2).
RDE	Remote Data Entry RDE bezeichnet die Erfassung von Daten mit elektronischen Formularen an verschiedenen Orten bei gleichzeitig zentraler Speicherung.
RFD	Retrieve Form for Data-capture - siehe Kapitel 2.7 -
SAE	Serious Adverse Event (dt.: Schwerwiegendes unerwünschtes Ereignis – SUE) Ein SAE ist ein schwerwiegendes Ereignis (wie Tod oder Hospitalisation), das im Rahmen einer klinischen Studie mit Arzneimitteln bei einem Probanden auftritt.
SDMS	Studiendatenmanagementsystem Ein SDMS dient dazu, medizinische und demographische Daten im Rahmen klinischer Studien zu erfassen und zu speichern. Aktuelle Systeme sind häufig als RDE-Systeme realisiert.
SDTM	Study Data Tabulation Model Das SDTM ist ein CDISC Standard, welcher die Struktur von Tabellen, zur Repräsentation von Studiendaten beschreibt. Das SDTM ist gedacht zur Auswertung und Einreichung von Studiendaten bei Zulassungsbehörden.
SOA	Serviceorientierte Architektur SOA ist eine Architektur zur Nutzung von Funktionalitäten in verteilten Systemen.
TMF	TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V. Als Dachorganisation medizinischer Forschungsverbände versucht die TMF durch gezielte Projekte und Dienstleistungen, die organisatorische, rechtliche und technologische Infrastruktur der medizinischen Forschung in Deutschland zu verbessern (http://www.tmf-ev.de/). <i>Hinweis:</i> Im Kontext klinischer Studien steht die Abkürzung auch für Trial Master File .

- UML** **Unified Modeling Language**
Die UML ist eine graphische Modellierungssprache zur Spezifikation und Dokumentation von (Anwendungs)systemen und Prozessen.
- VPN** **Virtuelles Privates Netzwerk**
VPN erlauben die (geschützte) Verknüpfung von mehreren privaten Netzen bzw. Arbeitsstationen über öffentliche Netze.
- WHO** **World Health Organization (dt.: Weltgesundheitsorganisation)**
Die WHO ist eine Sonderorganisation der Vereinten Nationen, deren verfassungsmäßiger Zweck es ist, „allen Völkern zur Erreichung des bestmöglichen Gesundheitszustandes zu verhelfen“
(<http://www.who.int/>).
- XML** **Extensible Markup Language**
XML ist eine ‚Auszeichnungssprache‘, welche es erlaubt, Textdaten hierarchisch strukturierter darzustellen.
- XSLT** **Extensible Stylesheet Language Transformation**
XSLT ist eine Sprache zur Beschreibung von Transformationsregeln für XML-Dokumente.

1 Einleitung

1.1 Gegenstand und Motivation

Hintergrund

„In scientia salus“ – „Heilung durch Wissenschaft“. Diesen Wahlspruch ließ der bedeutende Chirurg, Strahlentherapeut und Krebsforscher Vincenz Czerny vor über 100 Jahren am Samariterhaus in Heidelberg anbringen (nach [TUFFS und PIETZSCH 2008]). Bis heute hat dieser Spruch nichts an Aktualität verloren. Im Gegenteil: Mit neuen Methoden in den Naturwissenschaften eröffnen sich auch der medizinischen Forschung und Versorgung immer neue Möglichkeiten.

Klinische Studien sind dabei ein zentrales Element, um die Wirksamkeit oder Überlegenheit von Medikamenten und anderen therapeutischen Interventionen zu erforschen und zu evaluieren (vgl. [GAWLIK, ABHOLZ et al. 1998]). Hierzu werden Daten unter definierten Bedingungen erhoben und ausgewertet.

Ebenso wie in anderen Bereichen werden im Gesundheitswesen immer mehr Daten elektronisch übermittelt, verarbeitet und gespeichert (vgl. [DORDA, GALL et al. 2002]) – auch über Einrichtungsgrenzen hinweg (vgl. [HAAS 2006], S. 2 und [KRÜGER-BRAND 2007]). Dabei entsteht eine heterogene Menge von Datensammlungen an unterschiedlichen Orten. Diese verstärkte Verfügbarkeit von Informations- und Kommunikationstechnologien im Gesundheitswesen führt zu Schlagworten wie ‚E-Health‘ und ‚Gesundheitstelematik‘ (siehe Kapitel 2.1).

Problematik

In klinischen Studien ist es jedoch nach wie vor aufwändig und teuer, größere Mengen phänotypischer Daten in ausreichend guter Qualität zu erheben (vgl. [KOHANE und ALTMAN 2005] und [SAX 2008]). Eine hohe Qualität in der Forschung erfordert qualitativ hochwertige Daten und Prozesse zur Datenerhebung gemäß den Good Clinical Practices (GCP – siehe [ICH 2002]).

Aufwand

Obwohl in Routineversorgung wie auch in klinischen Studien vielfach Daten elektronisch verarbeitet werden, ist eine wechselseitige Nutzung bereits vorhandener Daten häufig nicht möglich. In beiden ‚Welten‘ haben sich unterschiedliche Standards etabliert. Meist stehen die HL7- und DICOM-Standards in der Versorgung den CDISC-Standards im Studienumfeld gegenüber (vgl. [OHMANN und KUCHINKE 2009]) und selbst die Interoperabilität zwischen verschiedenen Systemen eines Bereichs erweist sich häufig als schwierig.

Mangelnde Interoperabilität

So muss ein Prüfarzt mitunter eine Medikamentenverordnung, welche er bereits im Rahmen der Routineversorgung in der elektronischen Patientenakte (EPA) eines Studienteilnehmers dokumentiert hat, in einem elektronischen Case Report Forms (eCRF) für die jeweilige Studie erneut erfassen. Aber nicht nur zwischen Versorgung und Forschung, sondern auch innerhalb der Forschung sind Systeme oft noch nicht interoperabel. Tritt bei dem Studienteilnehmer etwa ein ‚schwerwiegendes unerwünschtes Ereignis‘ (SAE) auf, müssen die verordneten Medikamente gegebenenfalls im eingesetzten SAE-Managementsystem ein drittes Mal erfasst werden.

Verschiedene Bemühungen, um Routinedaten aus Krankenhausinformationssystemen (KIS) und Studiendaten aus Studiendatenmanagementsystemen (SDMS) zu verknüpfen, konzentrieren sich meist auf ein konkretes KIS/SDMS-Szenario (so zum Beispiel [KATZER, RÖHRIG et al. 2006]). Gerade für multizentrische Studien, welche eine Datenerhebung in verschiedenen Häusern mit oft unterschiedlichen KIS erfordern, wäre jedoch eine weiterreichende Interoperabilität wünschenswert.

Datenschutz

Um den einzelnen Bürger zu schützen, schränken in Deutschland Gesetze und Vorschriften die multiple Verwendung von Gesundheitsdaten zum Teil ein (vgl. etwa SGB V, § 291a Nr. 8). Wenn ein Datenaustausch zwischen Forschung und Versorgung ermöglicht werden soll, muss sorgfältig geprüft werden, ob der jeweils gewählte Ansatz mit dem Datenschutz vereinbar ist. Um bei der multiplen Nutzung von Patientendaten die Anforderungen des Datenschutzes zu erfüllen, müssen gegebenenfalls zusätzliche technische und organisatorische Maßnahmen getroffen werden.

Datenqualität

Darüber hinaus ist eine multiple Verwendung von Daten aus der Versorgung in der Forschung nur möglich, wenn sichergestellt ist, dass die Qualität der Daten für die Sekundärnutzung ausreichend hoch ist. Gerade klinische Studien stellen häufig hohe regulatorische Anforderungen an die Qualität und Nachvollziehbarkeit der erhobenen Daten.

Motivation

Die in Forschung und Versorgung vorhandenen Datensammlungen können auch für den jeweils anderen Bereich einen wichtigen ‚Informationsschatz‘ darstellen. Wären vorhandene Daten multipel und damit effizient nutzbar, könnten Prozesse in Forschung und Versorgung vereinfacht und Kosten eingespart werden. So kann beispielsweise bereits die Rekrutierung von Studienteilnehmern mit Hilfe der Daten aus Krankenhausinformationssystemen unterstützt werden (vgl. [DUGAS, LANGE et al. 2010]). Indem dieselben Daten nicht mehrfach erfasst werden müssen, können ferner Übertragungsfehler vermieden werden. In diesem Sinne muss angestrebt werden, Routinedaten aus der Versorgung und Studiendaten gemeinsam zu erheben bzw. – soweit möglich – wechselseitig zu nutzen (vgl. [WINTER, FUNKAT et al. 2007], [EMBI und PAYNE 2009], [PROKOSCH und GANSLANDT 2009], [OHMANN und KUCHINKE 2009] sowie [BREIL, SEMJONOW et al. 2011]). Dies ist besonders interessant vor dem Hintergrund einer Offenheit für persönliche, elektronische Gesundheitsakten in der Bevölkerung (vgl. [HOERBST, KOHL et al. 2009]) – auch wenn entsprechende Akten aktuell noch nicht verbreitet genutzt werden (siehe [ÄRZTE ZEITUNG 2011] und [INTERCOMPONENTWARE AG (ICW) 2011]).

Auf Archetypen basierende klinische Informationssysteme (siehe Kapitel 2.5) versprechen eine ‚nachhaltige Interoperabilität‘, weil diese Systeme große Flexibilität bieten und einfach an geänderte Rahmenbedingungen angepasst werden können. Da diese Informationssysteme gleichzeitig semantische Interoperabilität (siehe Kapitel 2.4) ermöglichen, könnte sich dieser Ansatz gut für die Verknüpfung von Systemen aus Forschung und Versorgung eignen.

1.2 Problemstellung**Problem 1**

Archetypen wurden bisher fast ausschließlich zur Dokumentation in der Versorgung eingesetzt (zum Beispiel [HAGGLUND, CHEN et al. 2010], [AUSTIN, LIM et al. 2011] oder [XIAO, COUSINS et al. 2011]). Auch wenn bereits in [GARDE, KNAUP et al. 2005] vorgeschlagen wurde, Archetypen zur Dokumentation von Forschungsdaten zu nutzen, wurde noch nicht systematisch und detailliert untersucht, ob sich Archetypen zur Repräsentation von Forschungsdaten eignen. Daher ist zurzeit noch unklar, ob auf Archetypen basierende Systeme geeignet sind, Daten für Forschung und Versorgung zu integrieren.

Problem 2

Entsprechend gibt es bisher auch keine Referenzarchitektur für auf Archetypen basierende Systeme, die eine Übernahme von Daten aus Krankenhausinformationssystemen in Studiendatenmanagementsysteme unterstützen.

1.3 Zielsetzung

Übergeordnetes Ziel dieser Arbeit ist es, basierend auf den *openEHR*-Spezifikationen und -Archetypen (siehe Kapitel 2.6) Ansätze zu erarbeiten und zu bewerten, die einen Übergang von der multiplen Erfassung zur multiplen Verwendung von Daten in Forschung und Versorgung ermöglichen.

Konkret soll eine auf Archetypen basierende Referenzarchitektur entworfen werden als generische Lösung zur Integration von Forschungs- und Versorgungsdaten. Im Gegensatz zu bestehenden Ansätzen, soll diese Architektur speziell die Integration verschiedener, heterogener KIS unterstützen. Dies ist besonders im Kontext von multizentrischen Studien von großer Bedeutung. In einem evolutionären Ansatz sollen Anforderungen an diese Referenzarchitektur (bottom-up) erhoben und eine entsprechende Architektur entworfen werden. In einem zweiten Schritt soll die Realisierbarkeit der relevanten Bausteine dieser Referenzarchitektur (top down) geprüft werden. Mit Hilfe der Referenzarchitektur soll aus informationsverarbeitender Sicht eine Brücke zwischen Versorgungs- und Studienwelt geschlagen werden.

Auf Archetypen basierende Referenzarchitektur zur Integration von Forschung und Versorgung: *Anforderungen* **Ziel 1**

Ermitteln von Anforderungen an eine auf Archetypen basierende Referenzarchitektur zur Integration von Krankenhausinformations- und Studiendatenmanagementsystemen im Rahmen eines evolutionären Analyseprozesses. Hierbei sollen vor allem folgende Fragen geklärt bzw. Aufgaben bearbeitet werden:

1. Ermitteln von generellen Anforderungen an eine Referenzarchitektur zur Integration von Forschung und Versorgung durch eine Fallstudie.
2. Ermitteln weiterer Anforderungen durch eine prototypische Implementierung eines auf Archetypen basierenden Systems zur Integration von Forschungs- und Versorgungsdaten.

Auf Archetypen basierende Referenzarchitektur zur Integration von Forschung und Versorgung: *Spezifikation* **Ziel 2**

Ausgehend von den zu Ziel 1 gesammelten Anforderungen und Erfahrungen soll eine auf Archetypen basierende Referenzarchitektur zur Integration von Forschung und Versorgung spezifiziert werden. Dazu werden vor allem folgende Fragen beantwortet:

3. Wie müsste eine auf Archetypen basierende Referenzarchitektur beschaffen sein, die es ermöglicht, Daten aus Krankenhausinformations- und Studiendatenmanagementsystemen multipel zu nutzen?
4. Wie müsste vorgegangen werden, um entsprechend dieser Architektur Daten aus Forschung und Versorgung multipel nutzen zu können? (Vorgehensmodell)

Auf Archetypen basierende Referenzarchitektur zur Integration von Forschung und Versorgung: *Realisierbarkeit* **Ziel 3**

Überprüfen der Realisierbarkeit der zentralen Komponenten der vorgeschlagenen, auf Archetypen basierenden Referenzarchitektur. Dabei soll insbesondere geprüft werden, ob sich Archetypen nicht nur zur Repräsentation von Daten aus der Versorgung sondern auch für Studiendaten eignen. Hierzu sollen für die Referenzarchitektur benötigte Archetypen spezifiziert bzw. falls notwendig neu entwickelt werden. Darüber hinaus soll geprüft werden, ob die entsprechenden Archetypen mit etablierten Standards und Modellen aus der

klinischen Forschung harmonisieren bzw. sich gemäß dieser entwickeln lassen. Es werden vor allem folgende Fragen beantwortet und Aufgaben gelöst:

5. Identifizieren und Entwickeln von benötigten Archetypen, um die vorgeschlagene Referenzarchitektur umsetzen zu können.
6. Prüfen, ob sich die auf Archetypen basierenden Repräsentationen der einzelnen Datenarten auf etablierte klinische Daten- und Informationsmodelle im Forschungsumfeld abbilden bzw. (falls entsprechende Archetypen noch nicht vorhanden sind) aus diesen ableiten lassen:
 - Studien-Metadaten: BRIDG-Modell (siehe Kapitel 2.7).
 - Studiendaten (medizinische Daten): Clinical Data Acquisition Standards Harmonization (siehe Kapitel 2.7).
 - Verknüpfung der auf Archetypen basierenden Repräsentation von Studiendaten und Studien-Metadaten mit dem Operational Data Model (siehe Kapitel 2.7).

2 Grundlagen

2.1 Begriffsdefinitionen

Im Folgenden werden für diese Arbeit wesentliche Begriffe und die hierfür verwendeten Bezeichnungen erläutert.

Gesundheitstelematik

„Gesundheitstelematik als *Kunstwort* aus Gesundheitswesen, Telekommunikation und Informatik umfasst ... alle einrichtungübergreifenden und ortsunabhängigen Anwendungen der Informations- und Kommunikationstechnologie im Gesundheitswesen zur Überbrückung von Raum und Zeit“ ([HAAS 2006], S. 6). Der Begriff ‚Telemedizin‘ wurde früher synonym verwendet, bezeichnet jedoch mittlerweile nur noch bestimmte Teilaspekte der Gesundheitstelematik – beispielsweise Teleradiologie, Telepathologie, Telechirurgie oder Telemonitoring (vgl. [HAAS 2006], S. 6).

E-Health

‚E-Health‘ ist ein modernes Schlagwort, welches schwer zu definieren ist, da es unterschiedlich verwendet wird (vgl. [EYSENBACH 2001]). Die HIMSS (Healthcare Information and Management Systems Society) beschreibt E-Health als „The application of Internet and other related technologies ... to improve the access, efficiency, effectiveness, and quality of clinical and business processes utilized by healthcare organizations, practitioners, patients, and consumers to improve the health status of patients“ ([HIMSS 2003], S 1). In dieser Arbeit wird E-Health als Synonym zum Begriff ‚Gesundheitstelematik‘ verstanden.

Konzept

‚Konzept‘ wird in dieser Arbeit als Synonym zu ‚Objekttyp‘, ‚Begriff‘, ‚Idee‘, und ‚Klasse‘ (im Sinne der objektorientierten Programmierung) verwendet. E. Ammenwerth definiert in ([AMMENWERTH und HAUX 2005], S. 10) Objekttyp wie folgt:

„Ein Objekt stellt einen Ausschnitt aus der wahrnehmbaren oder vorstellbaren Welt dar. Jedes einzelne Objekt weist eine Menge von Eigenschaften (Attribute) auf. Durch die Ermittlung gemeinsamer Eigenschaften lässt sich mittels Abstraktion eine Menge gleichartiger Objekte zu einer Denkeinheit zusammenfassen, die als Objekttyp (oder Klasse) bezeichnet wird. ...“

Definition

Zu einem (gedachten) Konzept bzw. Objekttyp können (reale) Instanzen bzw. Objekte existieren. Sowohl ein Konzept selbst als auch seine Instanzen verfügen dabei über dieselben kennzeichnenden Eigenschaften. Zusätzlich verfügt ein Konzept über spezifische Merkmalsarten, welche bei verschiedenen Instanzen unterschiedlich ausgeprägt sein können. So können verschiedene Instanzen desselben Konzepts unterschieden werden. Die folgenden Beispiele sollen dies veranschaulichen.

Konzept und Instanz

Jemand, der das Konzept ‚Auto‘ kennt, hat eine allgemeine Vorstellung davon, was ein Auto ist, bzw. welche Eigenschaften ein Auto kennzeichnen (hat vier Räder, fährt auf der Straße, dient der Beförderung von Menschen oder Gütern). Somit kann dieser Mensch auch ein Auto, welches er noch nie zuvor gesehen hat, als Auto erkennen und benennen. Dabei sind der Sportwagen, welcher vor dem Haus gegenüber steht und der Geländewagen, welcher gerade vorbeifährt, verschiedene reale Instanzen des gedachten Konzepts ‚Auto‘.

Merkmalsarten, über die alle Autos verfügen sind zum Beispiel ‚Farbe‘, ‚Höchstgeschwindigkeit‘ oder ‚Motorleistung‘. Je nach konkreter Instanz können diese unterschiedlich ausgeprägt sein – der Sportwagen gegenüber ist ‚blau‘, der Geländewagen ‚rot‘. Ein Konzept aus dem Bereich der Medizin ist der Blutdruck. Konkrete Instanzen wären der Blutdruck von Liz Miller oder Erika Gabler.

2.2 Klinische Studien

In klinischen Studien soll die Unbedenklichkeit, Wirksamkeit oder Überlegenheit von Medikamenten, medizinischen Geräten oder anderen Interventionen nachgewiesen werden. Klinische Studien sind eine wichtige Voraussetzung, um für einen Patienten die optimale Therapie auszuwählen und die Gabe unwirksamer oder schädlicher Substanzen zu vermeiden. So verbessern Studien nicht nur die Behandlung des einzelnen Patienten, sondern können auch Kosten im Gesundheitswesen senken.

Aufgrund ihrer großen Bedeutung – insbesondere bei der Zulassung neuer Medikamente – gibt es umfangreiche und strenge Regularien für die Durchführung solcher Studien, zum Beispiel die ‚Verordnung über die Anwendung der Guten Klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Arzneimitteln zur Anwendung am Menschen‘ (GCP-Verordnung). Diese Regularien stellen hohe Ansprüche an Datenqualität und Nachvollziehbarkeit der gewonnenen Ergebnisse. Durch entsprechendes Datenmanagement und Monitoring (siehe unten) wird versucht, diese Datenqualität zu gewährleisten.

Darüber hinaus muss eine klinische Studie systematisch geplant werden. Der Studienablauf muss detailliert im ‚Prüfplan‘ a priori festgelegt werden. Auch die Ein- und Ausschlusskriterien für potenzielle Probanden werden im Prüfplan definiert.

An der Erstellung und Umsetzung eines Prüfplans sind verschiedene Fachbereiche beteiligt, welche im Rahmen von klinischen Studien interdisziplinär zusammenarbeiten. Im Folgenden werden einzelne Bereiche kurz beschrieben.

Biometrie

Bereits bei der Planung ist die Biometrie von großer Bedeutung für die Qualität einer Studie. Typischerweise plant die Biometrie Studiendesign, Fallzahl, Randomisierungsverfahren, primäre und sekundäre Endpunkte und Auswertungsstrategie. Nachdem die benötigten Daten erhoben wurden, übernimmt die Biometrie die statistische Analyse und Auswertung (nach [LUNTZ, GORBAUCH et al. 2009]).

Datenmanagement

Die Integrität der Daten ist ein weiterer wichtiger Aspekt. Vor Beginn der Datenerhebung erstellt das Datenmanagement eine Studiendatenbank gemäß Datenmanagementplan. Sobald die Erhebung abgeschlossen ist, wird die Datenbank geschlossen und auf Inkonsistenzen sowie fehlende Werte geprüft. Werden Fehler erkannt, versucht das Datenmanagement diese durch Rückfragen (sogenannte ‚Queries‘) an die datenerhebenden Stellen zu klären und zu beseitigen (nach [LUNTZ, GORBAUCH et al. 2009]).

Monitoring

Das Monitoring ist ein wichtiges Instrument, um die (Daten)qualität in klinischen Studien zu gewährleisten. Das Monitoring überwacht, dass Prüfpläne, Gesetze und Standards wie etwa die Good Clinical Practice in den Prüfzentren eingehalten werden. Hierzu weisen Monitore die Mitarbeiter vor Ort in die Prüfpläne ein, zeigen strukturelle und organisatorische Fehler auf und vergleichen die für eine Studie erfassten Daten mit den sogenannten ‚Quelldaten‘ in den Patientenakten. Entdecken die Monitore beim Überprüfen der Studiendaten Lücken oder Inkonsistenzen stellen auch sie entsprechende Queries (nach [LUNTZ, GORBAUCH et al. 2009]).

Studienassistentz

Studienassistenten unterstützen die Prüffärzte vor Ort im Studienzentrum bei der Durchführung der Studien. Sie helfen, geeignete Patienten zu finden, unterstützen bei der Organisation und Durchführung der Studienvisiten (=Treffen zwischen Studienpersonal und Proband) oder dokumentieren Studiendaten. Oft sind sie nicht nur Ansprechpartner für Patienten und Ärzte, sondern insbesondere auch für die Monitore (nach [LUNTZ, GORBAUCH et al. 2009]).

Pharmakovigilanz

Die Pharmakovigilanz trägt dafür Sorge, dass die Anforderungen und Gesetze zur Arzneimittelsicherheit in klinischen Prüfungen beachtet werden. Die Pharmakovigilanz umfasst neben der kontinuierlichen Sicherheitsüberwachung einer klinischen Prüfung hauptsächlich Bewertungs- und Meldeprozesse für sogenannte ‚Serious Adverse Events‘ und jährliche Sicherheitsberichte (nach [LUNTZ, GORBAUCH et al. 2009]).

Projektmanagement, Qualitätsmanagement und IT

Neben den genannten Bereichen gibt es noch weitere Fachbereiche, die im Rahmen von klinischen Prüfungen zusammenarbeiten. So zum Beispiel das Projektmanagement (logistische Planung und Koordination), das Qualitätsmanagement (Review und Verwaltung der ‚Standard Operating Procedures‘ einer Einrichtung) oder die Informationstechnologie (Bereitstellen validierter Systeme zur Verarbeitung von Studiendaten) (nach [LUNTZ, GORBAUCH et al. 2009]).

2.3 Klassifikation von Daten

Wenn in dieser Arbeit von der ‚multiplen Verwendung von Daten in Forschung und Versorgung‘ gesprochen wird, werden unter der allgemeinen Bezeichnung ‚Daten‘ in der Regel verschiedene Arten von Daten zusammengefasst, die primär unterschiedlichen Zwecken dienen. Nachfolgend werden unterschiedliche Datenarten benannt und voneinander abgegrenzt.

Routinedaten

Unter ‚Routinedaten‘ werden in dieser Arbeit alle Daten verstanden, die im Rahmen der regulären ambulanten und stationären Patientenversorgung erhoben und typischerweise in einer Patientenakte dokumentiert werden. Diese Daten beschreiben den Gesundheitszustand eines Patienten sowie durchgeführte Interventionen. Hinzu kommen administrative Daten (=Patientenstammdaten) wie der Name, Geburtsdatum oder Versicherungsnummer des Patienten.

Versorgungsdaten

Unter ‚Versorgungsdaten‘ wird in der Regel die Teilmenge der Routinedaten verstanden, welche die Häufigkeit und Wirksamkeit bzw. Qualität der durchgeführten Maßnahmen beschreibt. So gehören beispielsweise die Liegedauer eines stationären Patienten oder die Dokumentation der durchgeführten Maßnahmen zu den Versorgungsdaten.

Studiendaten

Unter ‚Studiendaten‘ werden in dieser Arbeit alle Daten verstanden, die im Rahmen von klinischen Studien erhoben und dokumentiert werden, um den Gesundheitszustand eines Patienten, durchgeführte Interventionen bzw. die Wirkung einer Intervention (beispielsweise eines verabreichten Medikaments) zu beschreiben. An Studiendaten werden meist hohe Qualitätsanforderungen gestellt. Durch adäquates Datenmanagement und Monitoring (siehe Kapitel 2.2) wird versucht diesen Anforderungen gerecht zu werden.

Studien-Metadaten

„Unter Metadaten (‚Daten über Daten‘) versteht man strukturierte Daten, mit deren Hilfe eine Informationsressource beschrieben und dadurch besser auffindbar gemacht wird“ ([STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN 2001]). Studien-Metadaten beschreiben neben den zu erhebenden Merkmalen auch Rahmen, Design und involvierte Forscher einer klinischen Prüfung. Dies ist erforderlich, um die erhobenen Daten nachhaltig nutzen zu können (vgl. [ZPID 2010]). Studien-Metadaten sind beispielsweise notwendig, um studienübergreifende Auswertungen durchführen zu können. So kann etwa aus den Ein- und Ausschlusskriterien verschiedener Studien abgeleitet werden, ob die eingeschlossenen Patientenkollektive im Sinne einer bestimmten Fragestellung vergleichbar sind. Teilweise

wird unter ‚Studien-Metadaten‘ im engeren Sinne auch nur die Definition der in der Studie zu erhebenden Merkmalsarten und deren Formate verstanden.

Primärdaten

„Primärdaten sind Daten, die im Rahmen ihres originär vorgesehenen Verwendungszwecks aufbereitet und analysiert werden“ ([Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten der DGSMP und Arbeitsgruppe Epidemiologische Methoden der DGEpi der GMDS und der DGSMP 2008], S. 12).

Sekundärdaten

„Sekundärdaten sind Daten, die einer Auswertung über ihren originären, vorrangigen Verwendungszweck hinaus zugeführt werden. Maßgeblich für die Einstufung als Sekundärdaten sind Unterschiede zwischen dem primären Erhebungsanlass und der nachfolgenden Nutzung. Für die Einstufung ist es unerheblich, ob die weitergehende Nutzung durch den Dateneigner selbst oder durch Dritte erfolgt“ ([Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten der DGSMP und Arbeitsgruppe Epidemiologische Methoden der DGEpi der GMDS und der DGSMP 2008], S. 12).

2.4 Semantische und syntaktische Interoperabilität

Damit Daten multipel für Forschung und Versorgung genutzt werden können, müssen die jeweils eingesetzten Systeme interoperabel sein. Dabei können verschiedene Ebenen von Interoperabilität unterschieden werden.

Syntaktische Interoperabilität

Zwei oder mehr Computersysteme sind ‚syntaktisch interoperabel‘, wenn sie die Fähigkeit besitzen, Nachrichten auszutauschen: Alle beteiligten Systeme müssen über kompatible Schnittstellen und Protokolle bzw. Standards zum Datenaustausch verfügen. Syntaktische Interoperabilität impliziert jedoch nicht, dass Sender und Empfänger einer Nachricht die ausgetauschten Daten in jedem Fall gleich interpretieren. Syntaktische Interoperabilität ist dennoch die Grundlage für jede weitere Form von Interoperabilität. Das folgende Beispiel verdeutlicht die Grenzen von syntaktischer Interoperabilität.

Im Rahmen einer Studie sollte unter anderem der Blutzuckerwert der Probanden ermittelt werden. Die ermittelten Werte sollten dabei zunächst auf einem Server in Europa gespeichert werden. Zur Auswertung sollten die Werte in einer verschlüsselten Textdatei an einen Server in den USA geschickt werden. Beim Testen des Systems zeigte sich, dass die Werte auf dem Server in Europa in mmol/l gespeichert wurden, der Server in den USA jedoch die Werte in der Einheit mg/dl erwartete. Im Studienprotokoll war nicht angegeben worden, welche Einheit verwendet worden war (nach [IBERSON-HURST 2004]). Haben Sender und Empfänger einer Nachricht ein unterschiedliches Verständnis einer Größe bzw. allgemeiner eines Konzepts, birgt dies die Gefahr, dass beide alle Nachrichten, welche dieses Konzept nutzen, unterschiedlich interpretieren – insbesondere wenn es sich um einen automatisierten Austausch von Daten zwischen zwei Computersystemen handelt.

Nach [WALKER, PAN et al. 2005] sind ab einer bestimmten Stufe von Interoperabilität Schnittstellen erforderlich, welche in der Lage sind, empfangene Nachrichten aus dem Vokabular der sendenden ins Vokabular der empfangenden Organisation zu übersetzen. Im Allgemeinen führt dies häufig zu unvollständigen Übersetzungen, da die genutzten Vokabulare aufgrund unterschiedlicher Detaillierungsgrade oft nicht kompatibel sind.

Ein Beispiel aus dem menschlichen Sprachgebrauch soll verdeutlichen, dass selbst identische Bezeichnungen in verschiedenen Vokabularen zur Bezeichnung unterschiedlicher Begriffe benutzt werden können. Während im allgemeinen deutschen Sprachgebrauch der ‚Fuß‘ den untersten Teil des ‚Beins‘ eines Menschen bezeichnet, wird im süddeutschen Raum die Bezeichnung ‚Fuß‘ als Synonym und anstelle von ‚Bein‘ gebraucht. Wenn ein Patient aus Süddeutschland also seinem Orthopäden berichtet, er habe Schmerzen im Fuß, muss der Arzt diese Information in das übliche medizinische Vokabular übertragen (= ‚Fuß‘ als

‚Bein‘ interpretieren). Der Orthopäde muss sich im Klaren sein, dass die Schmerzen also von den Zehen bis zur Hüfte lokalisiert sein können.

‚Semantische Interoperabilität‘ bedeutet, dass zwischen verschiedenen Systemen Daten ausgetauscht *und* unmittelbar verarbeitet werden können – ohne diese Nachrichten übersetzen oder anpassen zu müssen. Es ist sichergestellt, dass Sender und Empfänger einer Nachricht die ausgetauschten Daten in gleicher Weise interpretieren. Dies kann beispielsweise dadurch erreicht werden, dass zwei oder mehr Systeme generell dieselbe Definition für verwendete Konzepte benutzen.

Semantische Interoperabilität

Von der ‚technischen semantischen Interoperabilität‘, die sich auf die Interoperabilität zwischen Computersystemen bezieht, kann die ‚menschliche semantische Interoperabilität‘ unterschieden werden. Ist letztere im Dialog zwischen verschiedenen Menschen nicht gegeben, reden die Betroffenen sprichwörtlich ‚aneinander vorbei‘. Dieser Arbeit beschäftigt sich hauptsächlich mit der technischen semantischen Interoperabilität.

2.5 Zwei-Ebenen-Modell zur Repräsentation von Gesundheitsdaten

In dieser Arbeit werden die Einsatzmöglichkeiten und möglichen Vorteile von Archetypen bei der multiplen Verwendung von Daten in Forschung und Versorgung untersucht. Nachfolgend wird daher das Konzept des Zwei-Ebenen-Modells ([BEALE, HEARD et al. 2006], [MICHELSEN, PEDERSEN et al. 2005]) zur Repräsentation von Gesundheitsdaten vorgestellt. Archetypen bilden eine Ebene dieses Modells. Dabei wird ein Schwerpunkt auf die *openEHR*-Spezifikationen gelegt, die ein solches Zwei-Ebenen-Modell sehr detailliert vorgeben.

Ebene 1: Archetypen

Die Bezeichnung ‚Archetyp‘ setzt sich zusammen aus den griechischen Wörtern ‚ἀρχή‘ (Beginn, Anfang) und ‚τύπος‘ (Vorbild, Skizze). Allgemein bedeutet dies so viel wie ‚Urtyp‘, ‚Prototyp‘ oder ‚Muster‘ (nach [WEBSTER 1971]). Dabei variiert die konkrete Bedeutung je nach Disziplin – wie beispielsweise Philosophie, Psychoanalyse oder Biologie. Im Kontext von Informationssystemen im Gesundheitswesen beschreiben Archetypen medizinische Konzepte. Konkret definiert ein Archetyp hier, welche Informationen zu einem bestimmten Konzept dokumentiert werden können. Abb. 2-1 zeigt, welche Merkmalsarten sinnvollerweise zu dem Konzept ‚Blutdruck‘ mit Hilfe des *openEHR*-Archetyps ‚Blood Pressure‘ dokumentiert werden können. Die Elemente unter ‚State‘ geben dabei als zusätzlichen Kontext zur Interpretation der Daten an, in welchem Zustand sich der Patient befand, als die unter ‚Data‘ erfassten Informationen erhoben wurden. ‚Protocol‘ gibt ferner darüber Auskunft *wie*, mit welchen Methoden und Instrumenten die Informationen erhoben wurden (vgl. [LESLIE UND HEARD 2006]).

Abstrakt gesehen, sind Archetypen wieder verwendbare, formale (und damit auch maschinenlesbare) Informationsmodelle der abgebildeten Konzepte: Ein Archetyp definiert, mit welchen Merkmalsarten konkrete Objekte des jeweiligen Konzepts beschrieben werden können. Darüber hinaus wird die Semantik der einzelnen Merkmalsarten sowie des modellierten Konzepts selbst beschrieben. Ziel ist es, so viele Merkmalsarten zu kombinieren, dass eine bestimmte Instanz eines Konzepts möglichst umfassend beschrieben werden kann. Aufgrund dieser Modellierungsvorgabe werden Archetypen im Englischen als ‚maximal data sets‘ bezeichnet (vgl. [LESLIE UND HEARD 2006] und Abb. 2-1). Auf diese Weise soll ein Archetyp für beliebige Anwendungsfälle geeignet sein. In einem konkreten Szenario werden jedoch nur die tatsächlich benötigten Merkmalsarten erhoben.

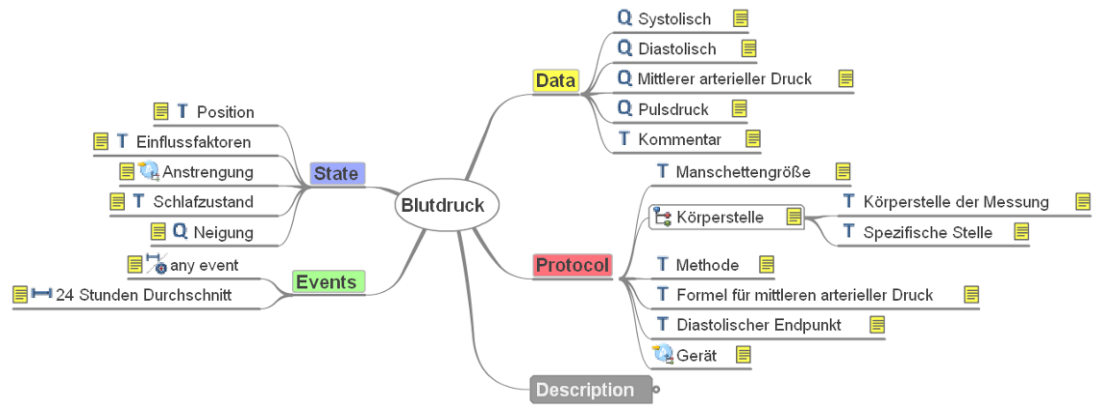


Abb. 2-1: Merkmalsarten, welche das Konzept ‚Blutdruck‘ beschreiben - dargestellt als Mindmap des openEHR-Archetyps ‚Blood Pressure‘ [Quelle: <http://www.openehr.org/knowledge/> – letzter Zugriff: 19.09.2011].

Ein Archetyp definiert also lediglich, mit welchen Merkmalsarten und gegebenenfalls erlaubten Ausprägungen ein bestimmtes Konzept beschrieben werden kann. Der Archetyp beschreibt nicht umfassend, die Eigenschaften, die bestimmtes Konzept kennzeichnen. Um das Beispiel aus Kapitel 2.1 erneut aufzugreifen: Ein Archetyp zur Beschreibung des Konzepts ‚Auto‘ würde keine erschöpfenden Informationen darüber enthalten, was allen Autos gemeinsam ist (hat vier Räder, fährt auf der Straße,...), sondern lediglich die Merkmalsarten aufzählen, mit welchen man ein konkretes Auto von einem anderen unterscheiden kann (also zum Beispiel ‚Farbe‘, ‚Höchstgeschwindigkeit‘ oder ‚Motorleistung‘).

Archetypen und Terminologien

Eine Terminologie beschreibt ein Konzept an sich – also die Eigenschaften, die auf alle Instanzen eines Konzepts zutreffen. Eine Terminologie könnte das Konzept ‚systolischer Blutdruck‘ beispielsweise beschreiben als ‚den Druck des Blutes in einer Arterie während der systolischen Herzphase‘. Ein Archetyp hingegen definiert die Strukturen bzw. Merkmale, welche zu dokumentieren sind, um konkrete Instanzen der jeweiligen Konzepte zu charakterisieren und voneinander zu unterscheiden. Um einen systolischen Blutdruck zu beschreiben, wäre gemäß dem Archetyp ‚Blutdruck‘ (vgl. Abb. 2-1) eine Größe zu dokumentieren, deren Maßzahl zwischen 0 und 1000 liegt und in der Maßeinheit mm[Hg] gemessen wird.

Archetypen sind ‚terminologieneutral‘. Dennoch können Terminologien an Archetypen oder einzelne Elemente eines Archetyps angebunden werden. Im Element ‚systolischer Blutdruck‘ des Archetyps ‚Blutdruck‘ könnte zur Beschreibung dieses Elements beispielsweise der vorgenannte Eintrag einer Terminologie referenziert werden. Darüber hinaus sind Archetypen unabhängig von menschlichen Sprachen. Ein ‚englischer Archetyp‘ könnte beispielsweise ins Deutsche übersetzt werden, ohne seine Bedeutung zu verlieren (nach [LESLIE UND HEARD 2006]).

Archetypen und Ontologien

Archetypen erheben nicht den Anspruch, die Realität exakt abzubilden, sondern beschreiben, was pragmatischer Weise zu einem klinischen Konzept zu dokumentieren ist. Eine Menge von Archetypen bildet daher auch keine Ontologie des Bereichs, aus welchem die beschriebenen Konzepte stammen (vgl. [openEHR 2007]).

Ebene 2: Referenzmodell

Archetypen zur Dokumentation medizinischer Konzepte basieren immer auf einem bestimmten Referenzmodell – beispielsweise dem openEHR-Referenzmodell oder dem Referenzmodell nach ISO 13606. Dieses Referenzmodell definiert eine kleine Menge elementarer Informationseinheiten wie zum Beispiel ‚Text‘, ‚Größe‘ bestehend aus Maßzahl

und Maßeinheit, ‚Uhrzeit‘ oder ‚Datum‘ sowie Möglichkeiten diese zu kombinieren, beispielsweise als Liste oder in einer Baumstruktur. So existieren theoretisch unendlich viele erlaubte Kombinationen dieser elementaren Informationseinheiten.

Ein Archetyp definiert die Merkmalsarten und gegebenenfalls erlaubte Ausprägungen, mit welchen konkrete Instanzen eines Konzepts beschrieben werden können, indem er Einschränkungen auf diesem Referenzmodell spezifiziert. Er legt fest, welche der möglichen Kombinationen der elementaren Informationseinheiten sinnvollerweise erlaubt sind, um ein bestimmtes Konzept zu beschreiben. Abb. 2-1 zeigt mit welchen Informationseinheiten beispielsweise das Konzept ‚Blutdruck‘ beschrieben werden kann. Im Kern werden hierzu die vier Größen ‚Systolisch‘, ‚Diastolisch‘, ‚mittlerer arterieller Druck‘ sowie ‚Pulsdruck‘ und ein Textfeld für Kommentare in einer Baumstruktur kombiniert.

Auf Archetypen basierende Informationssysteme

Informationssysteme, die Archetypen nutzen, müssen einerseits das zugrundeliegende Referenzmodell implementieren und andererseits in der Lage sein, Archetypen zu interpretieren. Während die technische Implementierung des Referenzmodells (Informationsebene) von System zu System beliebig variieren kann, sind die verwendeten Archetypen (Wissensebene) hiervon entkoppelt und systemübergreifend.

Durch die Trennung von logischer Wissensebene (Was soll gespeichert werden? ⇒ Archetypen) und physischer Informationsebene (Wie soll es gespeichert werden? ⇒ Referenzmodell) wird es leichter, ein System an sich änderndes Wissen anzupassen: Eine Änderung der Implementierung ist nicht erforderlich – das Referenzmodell ist statisch und ändert sich nicht; lediglich die entsprechenden Archetypen müssen adaptiert oder ergänzt werden.

Benutzen zwei oder mehr Informationssysteme dieselben Archetypen, ermöglicht dies prinzipiell nicht nur syntaktische, sondern auch semantische Interoperabilität zwischen diesen Systemen. In Archetypen ‚gekapselte‘ Daten haben an jedem Ort bzw. in jedem System die gleiche Bedeutung. Darüber hinaus können Archetypen genutzt werden, um erfasste Daten auf Gültigkeit zu prüfen und um Abfragen auf einem Datenbestand zu spezifizieren. Verschiedene auf Archetypen basierende Informationssysteme können nur dann vollständig semantisch interoperabel sein, wenn alle beteiligten Systeme dieselben Archetypen nutzen. Dies setzt voraus, dass die verwendeten Archetypen an einem zentralen Ort verwaltet werden. Dies impliziert wiederum, dass alle Beteiligten die Modellierung der verwendeten Konzepte so akzeptieren, wie dies durch die gemeinsam genutzten Archetypen vorgegeben wird (vgl. [KOHL, GARDE et al. 2008]).

Das Paradigma, Archetypen als Basis von Informationssystemen im Gesundheitswesen zu verwenden, hat sich auf europäischer und mittlerweile auch auf internationaler Ebene in der Norm ISO 13606 manifestiert. ISO 13606 regelt die standardisierte Kommunikation innerhalb und zwischen elektronischen Gesundheitsakten. Dazu spezifiziert die Norm, wie Auszüge aus elektronischen Gesundheitsakten (EGA) mittels eines Informationsmodells und zugehörigen Archetypen ausgetauscht werden sollen. Eine vollständige EGA wird – im Gegensatz zu den *openEHR*-Spezifikationen – nicht definiert (vgl. [SCHLOEFFEL, BEALE et al. 2006]). ISO 13606 basiert jedoch auf denselben Grundlagen wie die *openEHR*-Spezifikationen, weswegen sich die *openEHR*-Spezifikationen auf ISO 13606 abbilden lassen. Daher werden im Folgenden primär Archetypen und Informationssysteme gemäß den *openEHR*-Spezifikationen betrachtet.

2.6 openEHR

Hinter ‚*openEHR*‘ verbirgt sich eine Sammlung frei verfügbarer Spezifikationen zur Architektur von elektronischen Gesundheitsakten – *openEHR* ist also keine Anwendungssoftware. Dennoch wird anhand von Referenzimplementierungen geprüft, ob

die erarbeiteten Spezifikationen praktikabel sind und die Grundlage für weitere Implementierungen geschaffen. Ziel ist es, semantische Interoperabilität innerhalb und zwischen elektronischen Systemen im Gesundheitswesen zu schaffen, ohne ein proprietäres, herstellereigenes Datenformat zu verwenden. So soll eine nahtlose und qualitativ hochwertige Gesundheitsversorgung möglich werden. Hierzu werden alle klinischen Konzepte unabhängig von konkreten Softwaresystemen, in Form von *openEHR*-Archetypen modelliert und strukturiert gespeichert. Die verschiedenen Spezifikationen resultieren aus über 15 Jahren Forschung und internationalen Entwicklungen – einschließlich des ‚Good European Health Record‘. *OpenEHR*-Architektur und -Spezifikationen sind geistiges Eigentum der *openEHR*-Foundation. Die *openEHR*-Foundation ist eine ‚not-for-profit‘ Gesellschaft, welche vom University College London (CHIME) und der australischen Firma Ocean Informatics gegründet wurde. Eine beständig wachsende ‚Online-Community‘ mit über 1000 Mitgliedern aus 75 Ländern (Stand 2007) diskutiert und unterstützt die Weiterentwicklung von *openEHR* (nach [LESLIE 2007]).

***openEHR*-Archetypen**

Jeder *openEHR*-Archetyp bezieht sich auf eine Klasse innerhalb des *openEHR*-Referenzmodells. Entsprechend dieser Klassen lassen sich *openEHR*-Archetypen einteilen und benennen (etwa als COMPOSITIONSs, OBSERVATIONs oder INSTRUCTIONs).

COMPOSITIONS

In elektronischen Gesundheitsakten, die auf den *openEHR*-Spezifikationen basieren, repräsentieren COMPOSITIONs die elementaren Informationseinheiten, welche beim Bearbeiten einer Akte hinzugefügt, geändert oder gelöscht werden. Eine COMPOSITION wird ausgehend von der entsprechenden Klasse des Referenzmodells mit Hilfe eines COMPOSITION-Archetyps definiert. Hierzu wird in der Regel auf weitere Strukturarchetypen (beispielsweise SECTION-Archetypen) und Inhaltsarchetypen (so etwa OBSERVATION-Archetypen) zurückgegriffen. So nutzt die COMPOSITION *openEHR-EHR-COMPOSITION.medication_list.v1* beispielsweise die Archetypen *openEHR-EHR-ACTION.medication.v1* und *openEHR-EHR-INSTRUCTION.medication.v1*. Die jeweils vorangestellte Bezeichnung ‚*openEHR-EHR*[...]‘ zeigt dabei an, dass es sich um Archetypen handelt, die auf dem *openEHR* Referenzmodell basieren. Auf Ebene der COMPOSITIONs ist in *openEHR*-basierten Akten ferner ein Versionierungsmechanismus vorgesehen: Immer wenn ein Datenelement in einer Akte geändert wird, wird eine neue Version der COMPOSITION erzeugt, zu welcher das geänderte Element gehört.

Folders

Die *openEHR*-Spezifikationen legen fest, dass COMPOSITIONs in Verzeichnissen (engl. Folders) logisch organisiert werden können. Verzeichnisse könnten beispielsweise genutzt werden, um getrennt alle COMPOSITIONs eines Patienten zusammenzufassen, welche

- die demographischen Daten des Patienten repräsentieren, zum Beispiel das Alter des Patienten,
- die Informationen enthalten, die von längerfristiger Dauer sind, zum Beispiel Dauermedikationen und
- die Informationen enthalten, die nur eine kurze Zeit nach ihrer Erfassung von einiger Relevanz sind, zum Beispiel die Dokumentation eines vorübergehenden grippalen Infekts.

Dabei kann eine COMPOSITION auch in mehreren Verzeichnissen gleichzeitig referenziert werden (nach [BEALE, HEARD et al. 2008], S. 24f).

AOM / ADL

Das ‚Archetype Object Model‘ (AOM) ist das *openEHR*-Meta-Modell zur Beschreibung von Archetypen. Die ‚Archetype Definition Language‘ (ADL) ermöglicht es, einen Archetyp in Textform zu beschreiben. ADL und AOM sind zueinander äquivalent und lassen sich daher ineinander überführen. Mit einem Parser können eine ADL-Datei eingelesen und Objekte entsprechend dem AOM im Speicher erzeugt werden.

Templates

Um einen oder mehrere Archetypen lokalen Anforderungen, wie etwa in einem bestimmten Krankenhaus, anzupassen, werden sogenannte ‚Templates‘ benutzt. Ein Template kann

mehrere Archetypen zusammenfassen und die referenzierten Archetypen weiter einschränken. Zum Beispiel können dadurch optionale Elemente ausgeblendet oder Standardwerte vorgegeben werden. Ein Template entspricht häufig einem Bildschirmformular. So könnte ein Template ‚Vitalzeichen‘ die Archetypen ‚Blutdruck‘ und ‚Atmung‘ zusammenfassen. Abb. 2-2 zeigt wie Templates, Archetypen und Referenzmodell aufeinander aufbauen.

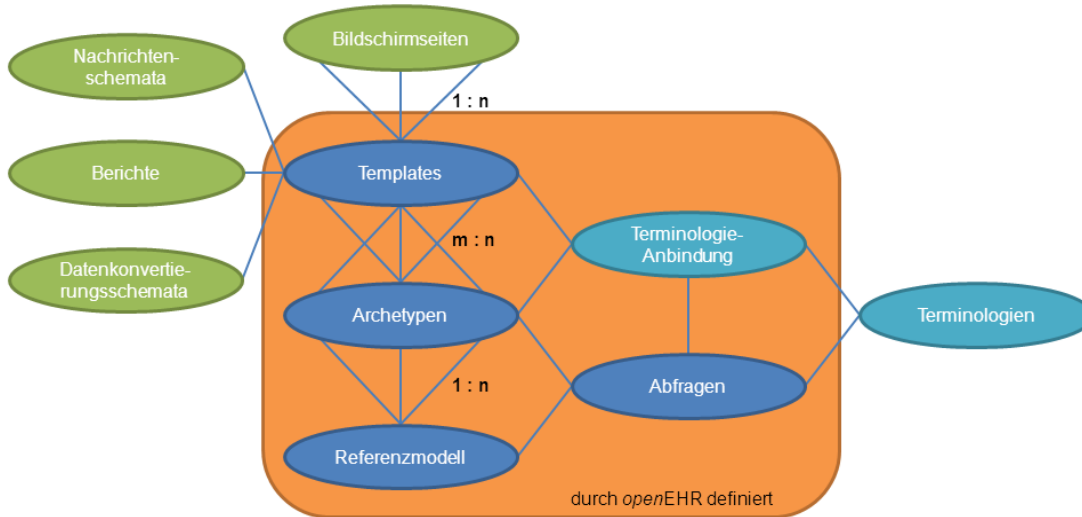


Abb. 2-2: Beziehungen zwischen *openEHR*-Templates, -Archetypen und -Referenzmodell sowie weiteren Elementen [nach <http://www.openehr.org/201-OE.html> - letzter Zugriff: 19.09.2011].

Zur Abfrage von in elektronischen Gesundheitsakten gespeicherten Daten wird die **AQL** (Archetype Query Language – früher EHR Query Language ‚EQL‘) entwickelt. Die AQL ist unabhängig von bestimmten EHR-Systemen oder Programmiersprachen und basiert lediglich auf der Semantik des *openEHR* AOM. Die AQL könnte prinzipiell auch für andere auf Archetypen basierende EHR-Systeme genutzt werden. Die Syntax von AQL ist eine Verschmelzung von SQL-Syntax und *openEHR* ADL Pfad-Syntax (nach [MA, FRANKEL et al. 2007; SACHDEVA und BHALLA 2009]; siehe auch <http://www.openehr.org/wiki/display/spec/Archetype+Query+Language+Description>).

Nach [MA, FRANKEL et al. 2007] könnte beispielsweise mit folgender Abfrage (Abb. 2-3) die aktuelle Medikation eines Patienten ermittelt werden:

```

SELECT c
FROM EHR e [ehr_id=$ehrId] CONTAINS
    COMPOSITION c [openEHR-EHR-COMPOSITION.medication_list.v1]
WHERE c/name/value = 'current medication list'
    
```

Abb. 2-3: AQL-Abfrage zur Ermittlung der aktuellen Medikation eines Patienten (nach [MA, FRANKEL et al. 2007]).

2.7 Bestehende Ansätze zur multiplen Nutzung von Daten aus der Routineversorgung

Klinische Studien verursachen durch hohe Anforderungen an die Datenqualität häufig nicht unerhebliche Kosten bei der Sammlung der benötigten Daten (vgl. [KOHANE und ALTMAN 2005] und [SAX 2008]). Daher besteht ein breites Interesse, bereits elektronisch vorhandene Daten für Studien zu nutzen (vgl. [WINTER, FUNKAT et al. 2007], [PROKOSCH und GANSLANDT 2009] und [OHMANN und KUCHINKE 2009]). Nachfolgend werden verschiedene lokale und überregionale Ansätze hierzu vorgestellt. Ferner werden in diesem Zusammenhang verfügbare Werkzeuge und Austauschformate betrachtet.

Single-Source-Ansatz

Die Idee des Single-Source-Ansatzes besteht darin, Daten für die klinische Forschung nicht in einem vom KIS unabhängigen System, sondern im KIS selbst zu dokumentieren. Hierzu werden im KIS notwendige Formulare ergänzt, welche die Dokumentation der für die jeweilige Studienfragestellung benötigten Merkmalsarten erlauben. Zur weiteren Verarbeitung und Auswertung werden die Daten nach dem Erfassen im KIS in der Regel in eine eigenständige Forschungsdatenbank oder ein Studiendatenmanagementsystem exportiert.

Vorteile

Hierdurch entfällt die manuelle Übernahme der Daten aus dem KIS in die Forschungsdatenbank bzw. das SDMS. Dies spart nicht nur Zeit, sondern vermeidet auch mögliche Übertragungsfehler. Darüber hinaus, können Ärzte und Pflegekräfte alle Daten in dem ihnen vertrauten System dokumentieren und müssen sich nicht in ein weiteres System einfinden. Wenn Studiendaten primär im KIS erfasst werden, können ferner keine Inkonsistenzen zwischen den Quelldaten im KIS und den Studiendaten im SDMS entstehen.

Nachteil

Dieser Ansatz ist nicht ohne weiteres auf eine multizentrische Studie übertragbar, da Forschungsdatenbasis und KIS oft proprietär beispielsweise auf Datenbankebene verknüpft wurden. Kämen durch andere Studienzentren verschiedene weitere KIS hinzu, müsste die Interoperabilität der Forschungsdatenbank mit jedem dieser Systeme einzeln geschaffen werden (vgl. etwa [KATZER, RÖHRIG et al. 2006]).

Beispiele

Single-Source-Ansätze wurden und werden aktuell in verschiedenen Häusern, vornehmlich Universitätsklinika, etabliert – so beispielsweise in Gießen, Erlangen, Münster oder Wien.

So wird am ‚Gießen Research Center in Infectious Diseases‘ des Universitätsklinikums Gießen versucht, aus Genomdaten und klinischen Daten prognostische Modelle für Entwicklung bzw. Verlauf einer Sepsis bei verschiedenen Krankheitsbildern zu entwickeln. Die hierzu benötigten klinischen Studiendaten werden automatisch aus dem dort etablierten elektronischen Intensivdokumentationssystem in eine Forschungsdatenbank übernommen (siehe [KATZER, RÖHRIG et al. 2006]).

Am Comprehensive Cancer Center des Universitätsklinikums Erlangen wird die Tumordokumentation in verschiedenen Bereichen vollständig im KIS durchgeführt. Von dort werden die Daten zur wissenschaftlichen Auswertung in ein i2b2-basiertes Data-Warehouse (siehe Kapitel 2.7) bzw. ein kommerzielles Studiendatenmanagementsystem überführt. Wenn i2b2 zur Auswertung genutzt wird, können in Erlangen erhobene Daten sogar mit Daten anderer Zentren zusammengeführt werden (siehe [PROKOSCH, RIES et al. 2011]).

Retrieve Form for Data-capture

Der ‚Retrieve Form for Data-capture‘-Leitfaden (RFD) wurde gemeinsam entwickelt von CDISC und Integrating the Healthcare Enterprise (IHE). RFD ermöglicht es, in Formulare einer Applikation Daten aus einer externen Quelle einzufügen, diese Daten zu ergänzen und wieder zurückzuspielen. Die Integration geschieht dabei nahtlos und für den Anwender unsichtbar (vgl. [INTEGRATING THE HEALTHCARE ENTERPRISE 2008] und [BAIN 2009]).

Der RFD-Leitfaden, in der IHE Terminologie als ‚Profil‘ bezeichnet, beschreibt die betrachtete Problematik in Form von Anwendungsfällen, Aktoren und Transaktionen und definiert einen auf bestehenden Standards und Techniken basierenden Lösungsansatz. So ist eine zentrale, im RFD-Profil genutzte Technologie der W3C Standard für elektronische Formulare ‚XForms‘ (<http://www.w3.org/MarkUp/Forms/>). Der RFD Leitfaden beschreibt zwar, wie Daten in Formulare aus einer externen Quelle eingefügt werden können, Voraussetzung hierfür ist die manuelle, teilweise aufwändige Zuordnung geeigneter Datenquellen zu den einzelnen Feldern eines Formulars.

Informatics for Integrating Biology and Bedside

Ein Ansatz zur Vernetzung von Versorgung und Forschung aus den USA ist das ‚Informatics for Integrating Biology and Bedside‘ Projekt (i2b2 – <http://www.i2b2.org>). i2b2 unterstützt die translationale Medizin, indem Werkzeuge bereitgestellt werden, um große, heterogene Bestände von phänotypischen Daten aus der Patientenversorgung sowie die zugehörigen Genomdaten für Forschungszwecke zu vernetzen. Der Fokus der vorliegenden Arbeit ist im Gegensatz dazu stärker ausgerichtet auf die multiple Nutzung phänotypischer Daten aus der Patientenversorgung, die in unterschiedlichen Systemen und in unterschiedlichen Einrichtungen entstehen. Diese Notwendigkeit entsteht insbesondere in multizentrischen Studien.

cancer Biomedical Informatics Grid

Speziell für die Krebsforschung wurde in den USA das ‚cancer Biomedical Informatics Grid‘ (caBIG – <http://cabig.nci.nih.gov>) etabliert. Das caBIG soll es ermöglichen, Daten aus verschiedenen Quellen zu integrieren und zu analysieren. Grundlage hierfür bilden eine Grid-Infrastruktur, entsprechende Applikationen sowie Leitlinien und Standards, um weitere Anwendungen zu entwickeln und integrieren (nach [OSTER, LANGELLA et al. 2008]). Das ‚Ökosystem‘ der verfügbaren caBIG Anwendungen und Dienste ist inzwischen relativ komplex. Der initialen Intention folgend, wurden die verfügbaren Werkzeuge primär für die Krebsforschung entwickelt. Der in dieser Arbeit untersuchte, auf Archetypen basierende Ansatz hat einen allgemeineren, breiteren Fokus.

Biomedical Research Integrated Domain Group Model

Das ‚Biomedical Research Integrated Domain Group‘ (BRIDG – <http://bridgmodel.org>) Model will Brücken schlagen zwischen Organisationen, technischen und fachlichen Experten sowie zwischen unterschiedlichen klinischen Informationsmodellen. Die folgenden Einrichtungen und Organisationen arbeiten gemeinsam an der Entwicklung des BRIDG-Modells (vgl. [BRIDG 2008], S. 4):

- Das Clinical Data Interchange Standards Consortium (CDISC),
- das HL7 Regulated Clinical Research Information Management Technical Committee (HL7 RCRIM TC),
- das amerikanische National Cancer Institute (NCI) und
- die amerikanische Food and Drug Administration (FDA).

Im Gegensatz zu den vorhergehend genannten Modellen handelt es sich beim BRIDG-Modell um ein ‚Domänen Analyse Modell‘. Übergeordnetes Ziel des BRIDG-Modells ist die semantische Interoperabilität insbesondere zwischen klinischer Forschung und Gesundheitsversorgung, aber auch mit anderen Bereichen wie etwa ‚Public Health‘. Die verschiedenen Standards in Forschung und Versorgung sowie innerhalb dieser Bereiche sollen harmonisiert und verschiedene Standardisierungsbemühungen integriert werden. Hierzu versucht das BRIDG-Modell eine gemeinsame Sicht auf

- Datenstrukturen (statische Semantik),
- Geschäftsprozesse (dynamische Semantik) sowie
- Beziehungen

zu schaffen. Der Fokus liegt dabei auf der ‚protokollgetriebenen Forschung und den zugehörigen regulatorischen Artefakten‘ ([ANGELES, EVANS et al. 2010], S. 8).

„Protokollgetriebene Forschung“ bezieht sich hier vornehmlich auf klinische Studien und bedeutet Forschung gemäß den exakten Vorgaben eines Studienprotokolls, wie es in klinischen Studien üblich ist. Elemente bzw. Konzepte aus dem Bereich der protokollgetriebenen Forschung sind beispielsweise Menschen, Tiere, Organisationen, Studiendesigns, Beobachtungen in klinischen Studien etc.

Um diese Konzepte zu modellieren, nutzt das BRIDG-Modell die Unified Modeling Language (UML). Konzepte, Attribute und Beziehungen werden mit Hilfe von UML-Klassen- und Instanz-Diagrammen dargestellt. Für Prozesse werden UML-Aktivitäts-, Status- und Sequenz-Diagramme eingesetzt. Die UML bietet den Vorteil, dass sie auch von nicht technischen Fachexperten relativ leicht erlernt werden kann. Die Überarbeitung und weitere Ausgestaltung des BRIDG-Modells wird von einem „Board of Directors“ und einem „Technical Harmonization Committee“ systematisch gesteuert und unterstützt (siehe [ANGELES, EVANS et al. 2010], S. 18ff).

Die Semantik des BRIDG-Modells kann prinzipiell abgebildet werden auf die CDISC Spezifikationen und das HL7 RIM. Obwohl es eine RIM-basierte Abbildung des BRIDG-Modells gibt, basiert das Modell selbst nicht auf dem HL7 RIM. Dies führt dazu, dass eine 1:1 Abbildung beispielsweise von Attributen zwischen BRIDG-Modell und HL7 RIM oft nicht möglich ist. Häufig wird ein BRIDG Attribut auf eine Kombination von RIM Attributen abgebildet. Die auf dem HL7 RIM basierende Abbildung des BRIDG-Modells ist darüber hinaus auch nicht als unmittelbare Grundlage für den auf HL7 basierenden Nachrichtenaustausch, sondern als Diskussionsgrundlage mit HL7 Nutzern gedacht (vgl. [ANGELES, EVANS et al. 2010], S. 42ff).

Praktisch kann das BRIDG-Modell genutzt werden bei der Entwicklung von Applikationen, um eine gemeinsame Semantik zwischen verschiedenen Anwendungen zu gewährleisten, etwa zwischen einem Studiendatenmanagementsystem und einem SAE-Reporting Werkzeug. So wird das BRIDG-Modell beispielsweise im caBIG „Clinical Trial Management System“ (<https://cabig.nci.nih.gov/workspaces/CTMS/>) genutzt.

Die BRIDG-Spezifikationen sind frei verfügbar (http://gforge.nci.nih.gov/frs/download.php/9082/BRIDG_Release_3.0.3_Package.zip). Für die Entwicklung von Software, welche das BRIDG-Modell nutzt, wird jedoch eine HL7-Lizenz benötigt, da die Attribute des statischen BRIDG-Modell auf HL7 Version 3 Datentypen basieren (vgl. [BRIDG 2008], S. 10f).

Clinical Data Acquisition Standards Harmonization

Der CDISC Clinical Data Acquisition Standards Harmonization Standard (CDASH – <http://www.cdisc.org/cdash/>) definiert Bezeichnungen, Beschreibungen und Ausfüllhinweise zu Datenelementen, die häufig in klinischen Studien erhoben werden. In sogenannten Domänentabellen werden logisch zusammengehörige Datenfelder als Basisdatensätze für die jeweiligen Anwendungen zusammengefasst, zum Beispiel für unerwünschte Ereignisse. Der CDASH Standard Version 1.1 beschreibt insgesamt 18 solcher Domänentabellen (vgl. [CDISC 2011]). Ziel von CDASH ist es, eine einheitliche und dadurch effiziente Datensammlung und -auswertung in klinischen Studien zu ermöglichen. Standardisierte Basisdatensätze vereinfachen jedoch nicht nur den Entwurf von papierbasierten und elektronischen CRF, sondern können darüber hinaus auch Metaanalysen unterstützen. Abgesehen von Empfehlungen zu Vorgehensweisen beim Erstellen von Datenerhebungswerkzeugen (siehe [CDISC 2010a]), nimmt CDASH jedoch keinen Einfluss auf das Erscheinungsbild oder Layout von CRF. Die definierten Basisdatensätze können auch genutzt werden, um die Schnittstellen zwischen Systemen in der Forschung und elektronischen Patientenakten zu verbessern.

Eine der 18 CDASH Domänen ist die Domäne „Prior and Concomitant Medications“ (CM), welche dazu dient, die Medikationen eines Studienteilnehmers zu erfassen. Die eigentliche Prüfmedikation einer Arzneimittelstudie wird hierbei jedoch nicht berücksichtigt und

gesondert behandelt. Es wird unterschieden zwischen Medikationen, die bereits vor Studienbeginn beendet wurden (prior), solchen, die erst im Laufe der Studie abgesetzt wurden (concomitant) und solchen, die über das Studienende hinaus bestanden (ongoing). Für die Domäne CM sind insgesamt 20 Felder definiert. Dabei kann unterschieden werden zwischen Feldern, welche eine Medikation

- qualitativ beschreiben (zum Beispiel die Angabe des Wirkstoffs),
- quantitativ beschreiben (zum Beispiel die verabreichte Dosis) sowie
- Datums- und Zeitangaben und
- ‚organisierenden‘ Feldern. Ein Beispiel dafür ist die Frage: ‚Ist überhaupt eine Medikationsangabe vorhanden?‘ Abhängig von der Beantwortung dieser Frage werden weitere Daten erhoben oder nicht.
- Darüber hinaus existieren Feldern zur Strukturierung des Datenerhebungsbogens bzw. Referenzfelder (zum Beispiel zur Referenz auf andere CRF – so etwa im Kontext von unerwünschten Ereignissen)

(nach [KRUEL 2010]. S17ff).

Ein enger Zusammenhang besteht zwischen CDASH und dem CDISC Study Data Tabulation Model (SDTM): CDASH bildet eine Erweiterung des SDTM, indem CDASH Anforderungen des SDTM und den Inhalt von Datenerhebungsbogen integriert. Zahlreiche CDASH-Datenelemente können daher auf das SDTM abgebildet werden. Ferner ist die CDASH-Terminologie eine Untermenge der SDTM-Terminologie (CDISC Controlled Terminology). Die CDASH-Terminologie stellt für eine Reihe von Datenfeldern Code-Listen mit häufig benutzten Fachterminen sowie deren Abkürzungen bereit. In CDASH Version 1.1 wurden die definierten Datenelemente außerdem um ein Attribut erweitert, welches den Bezug zum BRIDG-Modell herstellt, sofern dies sinnvoll und möglich war (vgl. [CDISC 2010a]).

Operational Data Model

Das Operational Data Model (ODM) ist ein vom Clinical Data Interchange Standards Consortium (CDISC) entwickelter, auf XML basierender, offener Standard für den Austausch, die Einreichung und Archivierung von Studiendaten. Das ODM bildet hierzu drei informationelle Aspekte klinischer Studien ab:

1. Die Studien-Metadaten (Definition von Merkmalsarten und Studienprotokoll),
2. administrative Daten (die beteiligten Nutzer und deren Zugriffsrechte),
3. die eigentlichen Studiendaten (vollständige Patientendatensätze samt Audit-Trail).

Bei ODM handelt es sich jedoch lediglich um einen ‚Transportstandard‘ – die Semantik der transportierten Daten ist primär nicht standardisiert (vgl. [IBERSON-HURST 2004]). Basierend auf dem ODM abgelegte Studiendaten lassen sich mit geeigneten Werkzeugen in andere Darstellungsformen wie etwa das CDISC Study Data Tabulation Model (SDTM) überführen (vgl. http://www.tmf-ev.de/Themen/Projekte/V034_01_SDTM_Wandler.aspx).

Werkzeuge der TMF e. V.

Im Rahmen von Projekten generiert die TMF als Dachorganisation medizinischer Forschungsverbände Methoden und Werkzeuge, um die vernetzte medizinische Forschung in Deutschland zu unterstützen.

So soll es das TMF Meta-Data-Repository beispielsweise ermöglichen, Strukturen, Merkmalsarten und Vokabulare zu definieren und einheitlich darzustellen. Hierdurch sollen, ähnlich wie auch durch CDASH, der Aufwand bei der Planung und Durchführung der Datenerhebung in der medizinischen Forschung reduziert und übergreifende Auswertungen erleichtert werden (nach [TMF 2009]). Hierzu implementiert das TMF Meta-Data-Repository ein Metamodell gemäß ISO/IEC 11179.

Meta-Data-Repository

**Datenschutz-
konzepte**

Um Verbundforschung gerade auch im Kontext der strengen, in Deutschland geltenden Datenschutzregularien zu ermöglichen, wurden von der TMF verschiedene, generische Datenschutzkonzepte entwickelt. Ergänzend wurden Werkzeuge implementiert, um die Konzepte einfach in der Praxis realisieren zu können (so beispielsweise Werkzeuge zur Pseudonymisierung und zum Identitätsmanagement wie den TMF PID-Generator).

Die beiden genannten TMF Projekte – Meta-Data-Repository und Datenschutzkonzept – können im Kontext der multiplen Verwendung von Gesundheitsdaten wertvolle Bausteine darstellen. Ferner engagiert sich die TMF direkt in der Projektgruppe ‚Nutzung von elektronischen Patientenakten für die klinische Forschung‘ der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS).

3 Vorgehen und Methodik

Im Folgenden werden das Vorgehen und die Untersuchungsmethoden beschrieben, die angewendet wurden, um die Ziele dieser Arbeit zu erreichen.

3.1 Generisches Vorgehensmodell für die multiple Nutzung von Daten in Forschung und Versorgung

An Daten, die in klinischen Studien ausgewertet werden, werden meist höhere Qualitätsanforderungen gestellt als an Daten, die im Rahmen der medizinischen Versorgung erhoben werden. Daten für klinische Studien müssen in der Regel in definierten Formaten und nach einem exakt definierten Vorgehen erhoben werden. Darüber hinaus kann es vorkommen, dass in klinischen Studien Daten benötigt werden, die nicht zum Spektrum der Routedokumentation eines Krankenhauses gehören. Daher muss vor einer multiplen Verwendung von KIS-Daten systematisch geprüft werden, welcher Anteil der in einer Studie geforderten Daten aus dem KIS übernommen werden könnte. Dabei müssen alle in der Studie zu erhebenden Merkmalsarten mit allen möglicherweise relevanten, im KIS dokumentierbaren Merkmalsarten verglichen werden. Die in einer Studie zu erhebenden Daten lassen sich dabei in der Regel aus Studienprotokoll, CRF und Studiendatenbank ableiten. KIS bestehen hingegen oft aus verschiedenen Subsystemen, zu welchen kaum semantische Beschreibungen der darin erfassbaren Merkmalsarten vorliegen. Dazu gehören explizit auch die konventionellen, also nicht rechnerunterstützten Subsysteme eines KIS.

Für den notwendigen systematischen Vergleich der Merkmalsarten aus KIS und Studie wurde ein generisches Vorgehensmodell entwickelt, das unabhängig ist von den genutzten Anwendungssystemen und für beliebige Studien genutzt werden kann. Die in Abb. 3-1 dargestellten Teilschritte dieses Vorgehensmodells werden nachfolgend im Detail beschrieben. Wesentlicher Bestandteil dieses Vorgehensmodell ist die Erhebung von Metadaten, welche die Eigenschaften der im KIS dokumentierbaren Merkmalsarten beschreiben. So kann geprüft werden, welche KIS-Daten die Anforderungen einer Studie erfüllen und sich für eine multiple Verwendung eignen.

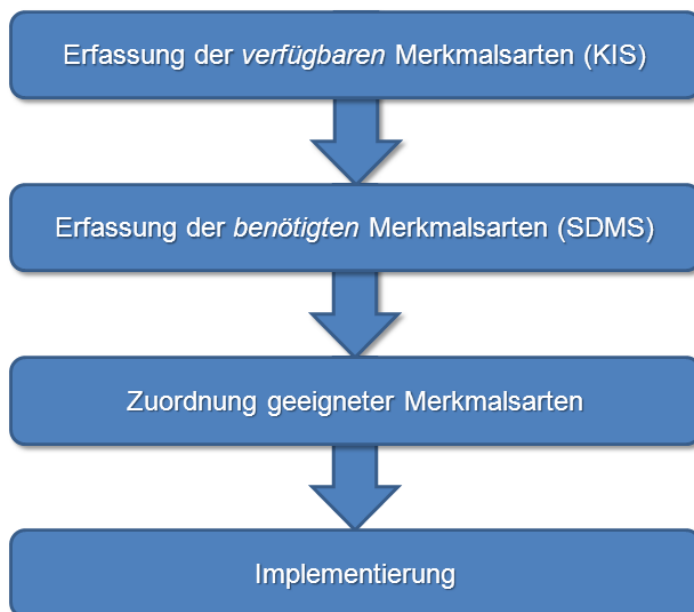


Abb. 3-1: Generisches Vorgehensmodell für die multiple Nutzung von Daten in Forschung und Versorgung.

Schritt 1:**Erfassung der verfügbaren Merkmalsarten**

Systematisches Erfassen aller Merkmalsarten, die im KIS typischerweise zu einem relevanten Patientenkollektiv dokumentiert werden. Ein solches Patientenkollektiv ergibt sich beispielsweise aus der Menge der Patienten, welche entsprechend den Ein- und Ausschlusskriterien in die jeweilige Studie eingeschlossen werden könnten.

Beginnend mit der im KIS vorhandenen Dokumentation zu einem Patienten werden alle Merkmalsarten, die im KIS zu diesem Patienten vorhanden sind, sowie zugehörige Meta-Merkmale erhoben. Diese Erhebung wird solange mit weiteren Patienten fortgeführt, bis die Menge der im KIS gefundenen Merkmalsarten ‚gesättigt‘ ist und nicht weiter wächst. Dabei sind für jede gefundene Merkmalsart folgende Meta-Merkmale zu erheben:

- **System bzw. Dokument**, welches genutzt wird, um die jeweilige Merkmalsart zu erfassen,
- die systeminterne **Bezeichnung** der Merkmalsart,
- der **Typ** der möglichen Merkmalsausprägungen
 - Ganzzahl
 - Gleitkommazahl
 - Datum
 - Uhrzeit
 - Code (zum Beispiel ICD 10-Codes)
 - Auswahlliste (zum Beispiel Liste von Stationen) und
 - Freitext
- sowie, falls zutreffend, die **Einheit** der Merkmalsausprägungen, zum Beispiel bei Laborwerten
- und falls vorgesehen auch alternative, erlaubte Einheiten.
- Ergänzende **Bemerkungen** werden bei Auffälligkeiten oder Besonderheiten dokumentiert. Dies können zum Beispiel Abhängigkeiten mehrerer Merkmalsarten untereinander oder Listendarstellungen sein. Ebenso wird hier die vollständige Bezeichnung einer Merkmalsart dokumentiert, wenn diese systemintern nur mit einer Abkürzung bezeichnet ist.
- Als weiteres Meta-Merkmal werden **semantische Integritätsbedingungen** dokumentiert, die beispielsweise regeln, wann eine Merkmalsart in der Studie zu erheben ist und wann nicht. Der Typ einer Fistel müsste etwa nur dann dokumentiert werden, wenn auch eine Fistel vorliegt.
- Ebenso werden **Plausibilitätsprüfungen** erfasst (zum Beispiel ‚Körpergröße muss zwischen 100 cm und 220 cm betragen‘).
- Als zusätzliches Meta-Merkmal wird zu den einzelnen Merkmalsarten im KIS erhoben, ob üblicherweise auch **Ausprägungen dokumentiert** werden.

Diese Meta-Merkmale dienen der Überprüfung der syntaktischen und semantischen Interoperabilität zwischen zwei Merkmalsarten. Ferner kann anhand dieser Meta-Merkmale geprüft werden, ob die in einer Studie geforderte Datenqualität eingehalten wird und ein entsprechendes Datenelement aus dem KIS in die Studie übernommen werden könnte.

Wurden alle Merkmalsarten, die in einem KIS dokumentiert werden können, vollständig analysiert, muss dieser Schritt nur einmal pro Studienzentrum durchgeführt werden. Der dabei entstandene Katalog von Merkmalsarten und zugehörigen Meta-Merkmalen kann bei der multiplen Nutzung der jeweiligen KIS-Daten für künftige Studien wiederverwendet werden.

Schritt 2:**Erfassung der benötigten Merkmalsarten**

Analog zu Schritt 1 werden in Schritt 2 alle Merkmalsarten sowie deren Meta-Merkmale systematisch erfasst, die in der Studie bzw. für die jeweilige Auswertungsfrage benötigt werden. Im Falle klinischer Studien ergeben sich diese Merkmalsarten in der Regel aus den CRF- bzw. Datenbankdefinitionen. Dieser Schritt muss einmal pro Studie durchgeführt werden.

Zuordnung geeigneter Merkmalsarten

Schritt 3:

Für jede in der Studie benötigte Merkmalsart wird geprüft, ob eine entsprechende Merkmalsart im KIS vorhanden ist, welche die jeweils benötigte Information liefern kann. Falls eine oder mehrere solcher Merkmalsarten im KIS erfasst werden, werden die entsprechenden Abbildungen dokumentiert. Bei diesen Abbildungen muss nicht nur die inhaltliche Konsistenz der Merkmalsarten geprüft werden, sondern auch ob Erhebungszeitpunkt, Typ und gegebenenfalls Format bzw. Einheit passend sind. Sind zwei Merkmalsarten in KIS und SDMS nicht identisch, kann es dennoch möglich sein, die für die Studie benötigten Informationen aus den KIS-Daten anzuleiten. Dazu müssen geeignete Transformationsvorschriften definiert werden. Ein Beispiel wäre die Berechnung des in einer Studie benötigten Body-Mass-Index aus den Merkmalsarten Körpergröße und Gewicht im KIS. Folgende Klassen von Transformationen können unterschieden werden:

- ‚1:1‘ – direkte Übernahme möglich / keine Transformation notwendig;
- ‚M:N‘ – eine Menge von m möglichen Merkmalsausprägungen im KIS muss eindeutig auf n mögliche Ausprägungen in der Studie abgebildet werden;
- ‚Umrechnung‘ – der für die Studie benötigte Wert muss aus einem anderen Wert im KIS errechnet werden (beispielsweise die Umrechnung eines Laborwerts in eine andere Einheit);
- ‚Logik‘ – der für die Studie benötigte Wert muss aus mehreren Werten im KIS errechnet werden (so etwa die Berechnung des Body-Mass-Index).

Darüber hinaus können die in einer Studie benötigten Informationen zwar im KIS erfasst sein, dort jedoch in wenig strukturierter Form, als Freitext vorliegen. Je nach Grad der Strukturierung kann es möglich sein, die benötigte Information mit einer speziellen Art der Transformation, dem ‚Textmining‘, aus dem unstrukturierten Text zu extrahieren.

Ist eine in der Studie benötigte Merkmalsart in mehreren Subsystemen des KIS verfügbar, wird primär die Merkmalsart vorgeschlagen, welches im engsten zeitlichen Kontext zur Erhebung der jeweiligen Studiendaten steht. Zusätzlich sollte berücksichtigt werden, welches Datenelement im KIS subjektiv die größte Validität aufweist oder am stärksten strukturiert vorliegt.

Implementierung

Schritt 4:

In Schritt 4 werden geeignete Schnittstellen zwischen KIS und SDMS basierend auf den vorhergehend identifizierten Abbildungen und Transformationen implementiert, um KIS-Daten in das SDMS übernehmen und beispielsweise CRF-Felder vorbelegen zu können. Alternativ, falls ein Single-Source-Ansatz angestrebt wird, werden Merkmalsarten im KIS ergänzt, welche in der Studie zusätzlich benötigt werden.

Nur wenn diese vier Schritte vollständig durchgeführt werden, kann gewährleistet werden, dass Daten, die aus dem KIS in eine Studie übernommen werden, die in der Studie geforderte Qualität aufweisen.

3.2 Fallstudie

Um zu ermitteln, welche Anforderungen an Daten in Forschung und Versorgung gestellt werden und in welchem Umfang eine multiple Verwendung realisierbar ist, wurde eine Fallstudie durchgeführt (vgl. [MAYRING 2007], S. 28ff und [AMMENWERTH, ILLER et al. 2003]). Die zentralen Auswertungsfragen dieser Fallstudie lauteten:

- Wie hoch wäre der Anteil der Daten, der heute schon aus dem KIS für eine konkrete Studie übernommen werden könnte?
- Wie hoch wäre dieser Anteil, wenn das KIS vollständig rechnerunterstützt wäre?
- In welchen Formaten werden die Merkmalsarten im KIS bzw. in der Studie dokumentiert?

Aufbauend auf den Antworten zu diesen Fragen wurden allgemeine Anforderungen an eine Systemarchitektur zur Integration von KIS und SDMS abgeleitet. In der Fallstudie wurde das

im vorhergehenden Kapitel eingeführte Vorgehensmodell angewendet. Als konkreter Fall wurden das Heidelberger KIS und die cTAG-Studie (conformable Thoracic Stent Grafts) ausgewählt. Das Heidelberger KIS, die cTAG-Studie sowie die daran durchgeführte Fallstudie werden im Folgenden im Detail dargestellt.

cTAG-Studie

Die cTAG-Studie ist eine gefäßchirurgische Studie, in welcher die Anwendung von verformbaren Stents in der Aorta untersucht wurde. Diese Stents können sich dem Gefäßverlauf anpassen und so bei der Überdeckung von Aneurysmen im Aortenbogen Vorteile bieten gegenüber herkömmlichen Stents. Die cTAG Studie ist eine einarmige, multizentrische, industriegesponserte Studie, die in Heidelberg und in vier weiteren europäischen Zentren durchgeführt wird. Dabei wird vom Sponsor angestrebt, insgesamt ca. 100 Patienten in die Studie einzuschließen. Die im Rahmen dieser Arbeit durchgeführte Fallstudie wurde an einer Stichprobe der in Heidelberg rekrutierten Patienten durchgeführt.

Prozessanalyse

In semistandardisierten Gesprächen mit dem Prüfarzt und der Study Nurse am Studienzentrum Heidelberg wurden der Ablauf der Datenerhebung für die cTAG-Studie und dabei auftretende, typische Probleme erfasst. Hierbei wurden auch die bei der Datenbestandsanalyse zu berücksichtigenden Dokumentationssysteme ermittelt.

Datenbestandsanalyse

Gemäß Vorgehensmodell (siehe Kapitel 3.1) wurden mit Hilfe von Datenbestandsanalysen zunächst alle Merkmalsarten, die im KIS konventionell wie auch elektronisch dokumentiert werden können, erfasst und mit dem Satz von Meta-Merkmalen charakterisiert. Anhand der eCRF der cTAG-Studie und der zugehörigen Studiendatenbank wurden dann die für die Studie zu erhebenden Merkmalsarten analysiert, dokumentiert und den im KIS erfassbaren Merkmalsarten systematisch gegenübergestellt. Als Werkzeug zur Dokumentation der einzelnen Merkmalsarten, der zugehörigen Merkmalsausprägungen, der identifizierten Abbildungen und der notwendigen Transformationen wurden Microsoft Excel-Tabellen entwickelt.

Waren zu einer für die Studie benötigte Merkmalsart in mehreren Subsystemen des KIS geeignete Datenelemente verfügbar, wurden neben der primär vorgeschlagenen Abbildung auch die weiteren möglichen Abbildungen dokumentiert. Die identifizierten Abbildungen wurden gemeinsam mit der für die cTAG-Studie zuständigen Study Nurse validiert.

Krankenhausinformationssystem

Am Universitätsklinikum Heidelberg wird eine weitgehend papierlose Dokumentation und Archivierung angestrebt. Während des stationären Aufenthalts eines Patienten werden alle vorhandenen und anfallenden papierbasierten Dokumente in der sogenannten Präsenzakte (klassische, papierbasierte Patientenakte) gesammelt. Sobald der Patient entlassen wird, werden alle Dokumente aus der Präsenzakte gescannt und anschließend vernichtet (vgl. [KNAUP, PILZ et al. 2006]). Auf die gescannten Dokumente kann dann, wie auch auf die übrigen, direkt elektronisch erzeugten Dokumente, über das klinische Arbeitsplatzsystem i.s.h.med bzw. ein darin eingebundenes digitales Archiv zugegriffen werden. Eine Texterkennung beim Scannen von ursprünglich papierbasierten Dokumenten findet zurzeit nicht statt. Konkret wurden am Studienzentrum Heidelberg folgende Dokumentationssysteme in die Untersuchung eingeschlossen:

Elektronische Subsysteme:

- SIEMENS i.s.h.med: Klinisches Arbeitsplatzsystem – inklusive der Daten aus den Laborinformationssystemen.
- PHILIPS CareVue: Patientendatenmanagementsystem – in Heidelberg für die Chirurgie adaptiert und auch nur dort eingesetzt.
- GE Centricity PACS: Bildarchivierungs- und Kommunikationssystem.

Konventionelle Subsysteme:

- Präsenzakte: Papierbasierte Patientenakte, die für die Dauer des stationären bzw. ambulanten Aufenthalts, alle konventionellen Dokumente aufnimmt (wird in Form von digitalen Kopien archiviert).
- Ambulanzakte: Papierbasierte Patientenakte (zur nachstationären Nutzung).

Bei der Untersuchung wurden nur Dokumente bzw. Merkmalsarten im KIS von Patienten berücksichtigt, die in die cTAG-Studie eingeschlossen worden waren. Um ein möglichst umfassendes Bild zu erhalten, wurden alle Merkmalsarten berücksichtigt, die im KIS vorhanden waren, nachdem der cTAG implantiert, der Patient aus der Klinik entlassen und mindestens zum ersten Follow-Up-Termin wieder vorstellig geworden war. Dabei wurde angenommen, dass einmal erfasste Daten nicht mehr aus dem KIS gelöscht werden und dementsprechend die Menge der verfügbaren Informationen zu einem Patienten monoton wächst.

Einschlusskriterien

Nicht berücksichtigt wurden Dokumente externen Ursprungs (beispielsweise Arztbriefe oder Laborbefunde der zuweisenden Ärzte), da diese von Patient zu Patient stark variieren. Auch am Universitätsklinikum Heidelberg selbst erzeugte, behandlungsspezifische Formulare bzw. Datenelemente im KIS wurden nicht berücksichtigt, wenn diese aus vorhergehenden oder nachfolgenden, von der cTAG-Implantation vollkommen unabhängigen Behandlungsfällen stammten. Ferner blieben Formulare ohne medizinischen Fallbezug unberücksichtigt (so etwa die Dokumentation des Sozialdienstes), da sie in diesem Fall keine studienrelevanten Merkmalsarten enthalten.

Ausschlusskriterien

In der geplanten Untersuchung sollte nicht nur eine Stichprobe der im klinischen Informationssystem dokumentierbaren Merkmalsarten eingeschlossen werden, sondern die Grundgesamtheit *aller* dort erfassbaren Datenelemente. Daher wurde die komplette elektronische und papierbasierte Dokumentation mehrerer Studienpatienten analysiert. Die Gruppe von Studienpatienten wurde dabei gemäß Vorgehensmodell solange erweitert, bis die Menge der papierbasierten und elektronischen Dokumente, die im Krankenhausinformationssystem zu den einzelnen Patienten vorgefunden wurde, saturiert war und nicht weiter wuchs. Die Patienten wurden dabei zufällig ausgewählt. Zufallsmoment war die Reihenfolge ihres Einschlusses in die cTAG-Studie.

Stichprobe

War dasselbe Dokument mehrfach vorhanden, wurden die darin enthaltenen Merkmalsarten nur einmal berücksichtigt und gezählt. So etwa im Fall von Arztbriefen, die sowohl konventionell in der Ambulanzakte als auch eingescannt als elektronisches Dokument vorlagen. Auch mehrfach vorhandene Dokumente desselben Typs wie mehrere Erhebungen derselben Laborparameter wurden nur einmal berücksichtigt

3.3 Prototyp

Zum Identifizieren weiterer (technischer) Anforderungen an eine auf Archetypen basierende Systemarchitektur zur Integration von Forschung und Versorgung wurde prototypisch das **open Study Data Management System (openSDMS)** entwickelt. Als Basis für openSDMS wurde das ‚*openEHR Reference Framework and Application*‘ (Opereffa) System genutzt. Darüber hinaus wurden die Abläufe bei der Benutzung von openSDMS beschrieben.

Opereffa (<http://opereffa.chime.ucl.ac.uk>) ist eine auf den *openEHR*-Spezifikationen basierende elektronische Gesundheitsakte, welche am University College London entwickelt wird. Opereffa selbst basiert auf der *openEHR* Java Referenz Implementierung (vgl. [CHEN und KLEIN 2007]) und präsentiert sich dem Nutzer als Web-Anwendung, die Java Server Faces in der Präsentationsschicht nutzt. Zusätzlich wird das Dojo-Toolkit (<http://dojotoolkit.org/>) verwendet, um ‚Asynchronous JavaScript and XML‘ (AJAX) sowie erweiterte Funktionalitäten für die grafische Benutzeroberfläche nutzen zu können ([ARIKAN und INGRAM 2009]). Zur physikalischen Datenspeicherung nutzt Opereffa eine PostgreSQL-

Datenbank (<http://www.postgresql.org/>), welche über Hibernate (<http://www.hibernate.org/>) angebunden ist ([ARIKAN und INGRAM 2009]). Dabei werden alle Datenelemente, die von den genutzten Archetypen spezifiziert werden, in einem ‚Entity-Attribute-Value-Ansatz‘ (EAV) in einer einzigen, generischen Tabelle gespeichert. Das bedeutet, dass Opereffa eine einzige Tabelle nutzt, um die medizinischen Daten aller im System vorhandenen Patienten zu speichern. Jede Zeile dieser Tabelle enthält genau ein Datenelement, welches für einen bestimmten Patienten zu einer bestimmten Zeit dokumentiert wurde und durch einen bestimmten Archetypen exakt beschrieben ist. Einfach ausgedrückt besteht ein Datensatz bzw. eine Zeile dieser Tabelle aus folgenden Elementen:

- Patienten-ID (identifiziert eindeutig einen bestimmten Patienten),
- Sitzungs-ID (identifiziert eindeutig einen bestimmten Dokumentationszusammenhang),
- Name des Archetyps, welcher das jeweilige Datenelement repräsentiert und
- Pfad des Datenelements innerhalb des Archetypen, durch Archetyp-Name und Archetyp-Pfad wird ein Datenelement eindeutig identifiziert,
- dem eigentlichen Wert des Datenelements (=Merkmalsausprägung).

Für die Implementierung von openSDMS wurde Opereffa um eine spezielle ‚Studiensicht‘ samt zugehöriger Programmlogik erweitert. Hierzu wurden alle Ebenen des Opereffa Systems (Datenbank, Anwendungslogik und Präsentationsschicht) adaptiert – siehe Abb. 3-2. Als Plattform für die Entwicklung wurde ‚Eclipse‘ der Eclipse Foundation genutzt (Version 3.5 / Galileo).

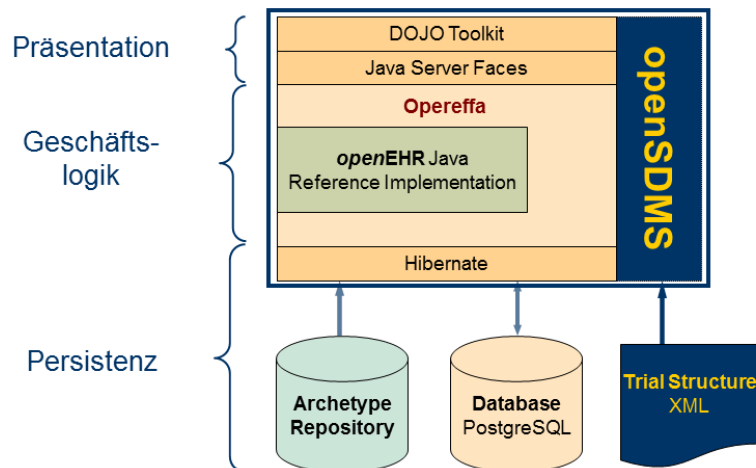


Abb. 3-2: Darstellung der für Präsentation, Geschäftslogik und Persistenz in Opereffa genutzten Technologien. Den Kern von Opereffa bildet die openEHR Java Reference Implementation. Für openSDMS wurden alle Ebenen des Opereffa Systems erweitert. Ferner wurde eine XML-basierte Studiendefinition integriert.

3.4 Archetypen zur Dokumentation von und in klinischen Studien

Eine auf Archetypen basierende, multiple Verwendung von Patientendaten für die klinische Forschung kann nur realisiert werden, wenn geeignete Archetypen zur Dokumentation von Studiendaten und Studien-Metadaten verfügbar sind. Analog der Unterscheidung von Studiendaten und Studien-Metadaten (vgl. Kapitel 2.3) können auch Archetypen für diese Klassen von Daten unterschieden werden. Unter Studien-Metadaten werden nachfolgend explizit die studienbeschreibenden Daten verstanden und nicht die Definition der in einer Studie zu erhebenden Merkmalsarten und deren Formate.

In einem ersten Schritt wurde daher geprüft, ob bereits spezielle Archetypen für Studiendaten oder Studien-Metadaten existieren. Hierzu wurde systematisch mit Hilfe des Ocean Clinical Knowledge Managers (www.openehr.org/knowledge/) nach geeigneten

Archetypen gesucht. Als Suchbegriffe wurden ‚clinical trial‘, ‚clinical study‘, ‚trial‘, und ‚study‘ verwendet.

Archetypen für Studiendaten

In der Regel variieren die meisten in klinischen Studien erhobenen Merkmalsarten stark von Studie zu Studie. Bestimmte Merkmalsarten werden jedoch, zumindest in ähnlicher Weise, regelmäßig in Studien erhoben. Die CDASH Spezifikation (siehe Kapitel 2.7) will genau solche Merkmalsarten standardisieren. Daher wurde systematisch untersucht, in wie weit sich die in CDASH (Version 1.1) definierte Datenelemente mit Hilfe von *openEHR*-Archetypen abbilden lassen.

Hierzu wurden exemplarisch die in den vier CDASH-Domänen

- Common Identifier Variables (CI)
- Common Timing Variables (CT)
- Adverse Events (AE) und
- Prior and Concomitant Medications (CM)

definierten Datenelemente (=Merkmalsarten) mit *openEHR*-Archetypen modelliert. Werden in einer elektronischen Patientenakte (EPA), welche auf *openEHR*-Archetypen basiert, Merkmalsarten mit denselben Archetypen repräsentiert, die zur Modellierung der CDASH-Domänen genutzt wurden, handelt es sich um die gleichen Merkmalsarten. Entsprechende EPA-Daten wären daher prinzipiell für die Übernahme in den jeweiligen, CDASH-konformen Datenerhebungsbogen einer Studie geeignet.

Hierzu wurden zunächst mit Hilfe des Ocean Clinical Knowledge Managers (www.openehr.org/knowledge/) systematisch alle Archetypen zusammengestellt, welche zum Zeitpunkt der Untersuchung vorhanden waren und einen Bezug zu den Konzepten der jeweiligen CDASH-Domäne erkennen ließen (im Falle der Domäne ‚Prior and Concomitant Medications‘ beispielsweise den Bezug zum Konzept ‚Medikation‘). Anschließend wurde für jedes CDASH-Datenelement geprüft, ob hierzu einer der identifizierten Archetypen ein korrespondierendes Datenelement enthält. Zusätzlich wurde für jede mögliche Abbildung geprüft, ob die von CDASH bzw. *openEHR* vorgegebenen Datentypen und Wertebereiche in Übereinstimmung gebracht werden können.

Alle korrespondierenden Felder wurden mittels entsprechender Abbildungen *CDASH-Element* ↔ *Archetyp-Name* und *-interne Pfadangabe / Feldname* dokumentiert. Diese Abbildungen sollen es ermöglichen, auf Archetypen basierende CDASH-konforme CRF für die betrachteten Domänen zu implementieren. Für jedes CDASH Feld, für welches sich eine Abbildung als schwierig erwies bzw. keine auf Archetypen basierende Abbildung gefunden wurde, wurden die Abbildungsprobleme analysiert und mögliche Lösungsansätze aufgezeigt. Wo sinnvoll und notwendig wurden bestehende Archetypen in Form von Spezialisierungen erweitert bzw. neue Archetypen entwickelt. Darüber hinaus wurden die Abbildungen so definiert, dass sie auch genutzt werden können, um Daten aus KIS, die auf den *openEHR*-Spezifikationen basieren, zu extrahieren und damit CDASH-basierte Felder in eCRF vorzubelegen. In Fällen, in welchen sich Abbildungen als schwierig erwiesen, etwa weil mehrere unterschiedliche Archetypen gleichzeitig berücksichtigt werden mussten, wurde die AQL (siehe Kapitel 2.6) genutzt. Das Feld CMYN der Domäne ‚Prior and Concomitant Medications‘ ist ein Beispiel für einen solchen Fall:

Feld **CMYN** (CM-Feldnummer 1)

Feldinformation: Wurden irgendwelche Medikamente eingenommen? („Ja“, „Nein“)

Das Feld zeigt an, ob eine Medikation eingenommen wurde. Ist dies nicht der Fall, wird lediglich diese Information dokumentiert – alle weiteren CM-Domänenfelder entfallen. Zur Dokumentation, dass von einem Patienten eine Medikation bzw. Substanz eingenommen wurde, könnten in einer auf *openEHR*-Archetypen basierenden Patientenakte unterschiedliche Archetypen genutzt werden. Um das CDASH-Feld CMYN in einem CRF

basierend auf einer Patientenakte vorbelegen zu können, müsste diese Akte nach unterschiedlichen medikationsbezogenen Archetypen durchsucht werden.

OpenEHR-Templates ermöglichen es, Bildschirmformulare ähnlich einem eCRF zu definieren, um Daten basierend auf Archetypen erfassen zu können. Für jede der betrachteten CDASH Domänen wurde ein solches Template erstellt, ausgehend von den identifizierten sowie den für CDASH neu entwickelten und angepassten Archetypen. Diese Templates können als Vorlage für eCRF in auf Archetypen basierenden SDMS genutzt werden. Zum Entwickeln der Templates wurde der Ocean Template Designers in der Version 2.6 eingesetzt.

Archetypen für Studien-Metadaten

Bei der Dokumentation klinischer Studien werden neben klinischen Daten auch Informationen erfasst, welche eine Studie und die dort erhobenen Daten an sich beschreiben wie etwa die Bezeichnung oder die Fragestellung einer Studie (Studien-Metadaten). Die papierbasierte Dokumentation einer Studie kann nur dann sinnvoll in einem elektronischen System abgebildet werden, wenn auch die entsprechenden Studien-Metadaten dort strukturiert dokumentierbar sind. Nur so können die erhobenen Daten nachhaltig, beispielsweise im Rahmen von Metastudien, genutzt werden. Ziel des weiteren Vorgehens war es daher, ergänzend zur Repräsentation von Studiendaten, auch die auf Archetypen basierende Repräsentation von Studien-Metadaten zu ermöglichen.

Grundlage des Entwurfs von Archetypen für Studien-Metadaten war die Definition einer generischen Struktur, mit welcher sich nahezu alle klinischen Studien abbilden lassen. Hierzu wurde zunächst analysiert, welche Objekttypen im Rahmen einer klinischen Prüfung relevant sind, wie diese zueinander in Beziehung stehen und durch welche Merkmalsarten (=Studien-Metadaten) diese Objekttypen charakterisiert werden können. Um relevante Objekttypen und kennzeichnende Merkmalsarten zu identifizieren, wurde ein systematischer Vergleich aller Merkmalsarten durchgeführt, welche für das Deutsche Register Klinischer Studien (DRKS – register.germanct.de/drks) und für die International Clinical Trials Registry Platform (ICTRP – apps.who.int/trialsearch) der Weltgesundheitsorganisation (WHO) erhoben werden. Zusätzlich wurden Merkmalsarten berücksichtigt, die im Koordinierungszentrum für Klinische Studien (KKS) Heidelberg (www.kks-hd.de) zur Dokumentation der dort betreuten Studien erhoben werden. Die Liste der im KKS erhobenen Merkmalsarten war in einem iterativen Prozess über einen längeren Zeitraum hinweg etabliert worden. Alle Merkmalsarten und die jeweils vorgesehenen Ausprägungen aus diesen Quellen wurden in einer Tabelle zusammengefasst und gegenübergestellt. Basierend auf dieser tabellarischen Zusammenfassung von Merkmalsarten, wurden Gruppen von logisch zusammengehörigen Merkmalsarten identifiziert und eine generische Studienstruktur definiert. Darauf aufbauend wurden *openEHR*-Archetypen modelliert, mit denen konkrete Studien beschrieben werden können. Die zuvor erstellte, tabellarische Gegenüberstellung der studienbeschreibenden Merkmalsarten von WHO, DRKS und KKS wurde erweitert um Referenzen auf die entsprechenden Datenpunkte in den neu entwickelten Archetypen. Zur Modellierung der Archetypen wurde der Archetype Editor der Firma Ocean Informatics in der Version 2.1 eingesetzt.

3.5 Abbildung auf Standards

Um zu prüfen, ob sich die gemäß Kapitel 3.4 entwickelten Archetypen tatsächlich zur Dokumentation von und in klinischen Studien eignen, wurde untersucht, inwieweit diese Archetypen mit in der klinischen Forschung etablierten Modellen und Standards harmonisieren und sich auf diese abbilden lassen.

Studiendaten

Als Ausgangspunkt für die Identifikation und Entwicklung von Archetypen zur Repräsentation von (medizinischen) Studiendaten waren die CDASH-Spezifikationen genutzt worden (vgl. Kapitel 3.4.1). Somit ist eine Abbildbarkeit auf CDASH gegeben und es waren keine weiteren Untersuchungen notwendig.

CDASH

Ergänzend wurde geprüft, wie die auf Archetypen basierende Repräsentation von Studiendaten und Studien-Metadaten im Operational Data Model (siehe Kapitel 2.7) abgebildet werden kann. Durch die Abbildung auf das ODM wird zumindest eine syntaktische Interoperabilität zwischen auf Archetypen basierenden SDMS und konventionellen Systemen im Umfeld klinischer Studien ermöglicht.

ODM

Studien-Metadaten

Ausgehend von den Archetypen für Studien-Metadaten wurde geprüft, an welchen Stellen im BRIDG-Modell (siehe Kapitel 2.7) sich die in den Archetypen definierten Merkmalsarten wiederfinden. Alle identifizierten Abbildungsmöglichkeiten wurden dabei in der tabellarischen Zusammenstellung der Studien-Metamerkmale ergänzt, welche in Vorbereitung der Entwicklung der Archetypen für Studien-Metadaten erstellt worden war (vgl. Kapitel 3.4). Zur Analyse und Darstellung der Archetypen sowie des BRIDG-Modells wurden die *openEHR ADL-Workbench* (Version 1.4) bzw. der *SPARX Enterprise Architect Light* (Version 8.0) genutzt. Eine Einführung in das BRIDG-Modell lieferte der zugehörige *User's Guide* ([ANGELES, EVANS et al. 2010]).

BRIDG

3.6 Zusammenfassung des Vorgehens

Um zu ermitteln, welche Anforderungen an Daten in Forschung und Versorgung gestellt werden und in wie weit eine wechselseitige, multiple Verwendung realisierbar ist, wurde eine Fallstudie durchgeführt. Hierbei wurden, einem zuvor definierten generischen Vorgehensmodell folgend, die im Heidelberger KIS dokumentierbaren und die in der cTAG-Studie benötigten Merkmalsarten vollständig erfasst und beschrieben. In Vorbereitung einer multiplen Nutzung der KIS-Daten wurde geprüft, welche in der cTAG-Studie benötigten Informationen aus dem KIS extrahiert werden können.

Das Opereffa-System, eine auf Archetypen basierende EPA, wurde dem Single-Source-Ansatz folgend zu einem rudimentären Studiendatenmanagementsystem (*openSDMS*) erweitert. Hierbei wurden die prinzipielle Realisierbarkeit eines auf Archetypen basierenden SDMS sowie die damit verbundenen technischen Anforderungen geprüft.

Aufbauend auf den Erkenntnissen aus Fallstudie und der Implementierung von *openSDMS* wurde in einem iterativen Prozess eine Referenzarchitektur zur Anbindung verschiedener KIS an ein SDMS entwickelt und beschrieben.

Um die entwickelte Referenzarchitektur realisieren zu können, werden geeignete Archetypen benötigt. Daher wurden Archetypen sowohl zur Dokumentation von Studiendaten als auch von Studien-Metadaten bereitgestellt. Hierzu wurden sowohl bestehende Archetypen genutzt und erweitert, als auch neue Archetypen speziell für das Studiumfeld entwickelt.

In einem weiteren Schritt, wurden alle von den bereitgestellten Archetypen definierten Merkmalsarten auf die im Bereich der klinischen Forschung etablierten Modelle BRIDG, CDASH und ODM abgebildet. So wurde sowohl die Eignung der bereitgestellten Archetypen für den Einsatz im Studiumfeld geprüft, als auch deren Kompatibilität zu etablierten Modellen hergestellt.

4 Ergebnisse

Im Folgenden werden die Ergebnisse beschrieben, welche sich aus den verschiedenen Untersuchungen und methodischen Ansätzen ergaben. Diese umfassen

- eine Fallstudie zur multiplen Verwendbarkeit von KIS-Daten,
- den Prototyp eines auf Archetypen basierenden SDMS,
- eine auf Archetypen basierende Referenzarchitektur zur Integration von Forschung und Versorgung sowie
- die Entwicklung von Archetypen für Studiendaten und Studien-Metadaten und
- die Verknüpfung dieser Archetypen mit etablierten Modellen im Bereich der klinischen Forschung.

4.1 Fallstudie

Ziel der Fallstudie anhand der cTAG-Studie und des Heidelberger KIS (siehe Kapitel 3.2) war es, den Anteil der für die Studie multipel nutzbaren KIS-Daten zu ermitteln. Weiteres Ziel war es, Anforderungen an eine Referenzarchitektur zur Integration von Forschung und Versorgung abzuleiten. Dazu wurden insbesondere, die Anforderungen bezüglich Datenqualität und -formaten in Studie und KIS analysiert und verglichen.

Die Datenerhebung für Forschung und Versorgung im Studienzentrum läuft wie folgt ab: Wenn ein Patient in die cTAG-Studie eingeschlossen wird, erhebt ein (Prüf-)Arzt die benötigten Daten im Rahmen der Aufnahme sowie direkt nach der Operation (OP). Weitere Daten werden während der Follow-Up-Visiten erhoben. Alle Daten werden in der Patientenakte bzw. – wenn das Erfassen der jeweiligen Daten in der Patientenakte nicht oder nicht in der benötigten Form vorgesehen ist – auf studienspezifischen Quelldatenbögen dokumentiert. Die Study Nurse überträgt die Daten dann in das eCRF-System und kontrolliert gleichzeitig die Vollständigkeit und Qualität der erfassten Daten. Gegebenenfalls recherchiert sie fehlende Daten in der elektronischen und papierbasierten Patientenakte oder erhebt diese, falls möglich, beim Patienten selbst. Anschließend kontrolliert der Prüfarzt die Eintragungen im eCRF und signiert diese (vgl. Abb. 4-1).

Datenerhebung cTAG-Studie

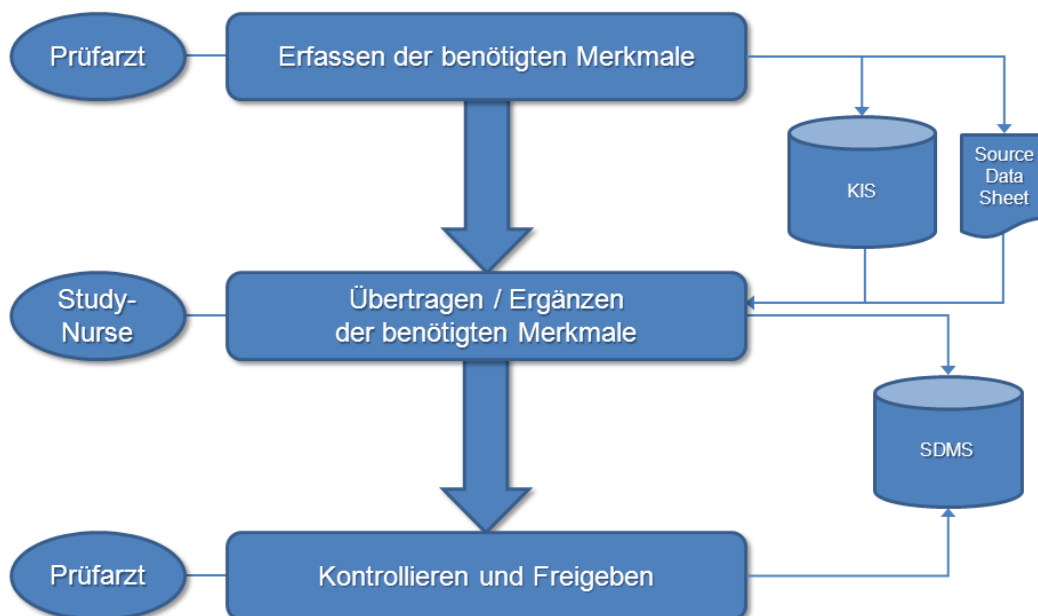


Abb. 4-1: Ablauf der Datenerhebung in der analysierten Studie (ohne eine multiple Nutzung von Datenelementen aus dem KIS).

Teilweise werden dieselben Merkmalsarten redundant in verschiedenen Subsystemen des KIS dokumentiert. So finden sich beispielsweise die Haupt- und Nebendiagnosen eines Patienten sowohl im OP-Bericht, wie auch im Bericht des Aufwachraums, im Bericht der ‚Intermediate-Care-Station‘ und im Entlassbrief. Dabei treten mitunter Inkonsistenzen zwischen den einzelnen Dokumenten auf. Dies erschwert die Arbeit der Study Nurse bei der Recherche nach den Studienmerkmalen.

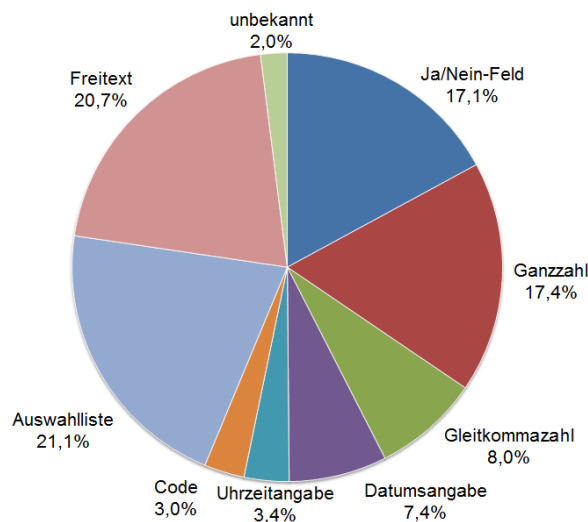


Abb. 4-2: Verteilung der Ausprägungstypen der 1452 im KIS dokumentierbaren Merkmalsarten (elektronisch und konventionell).

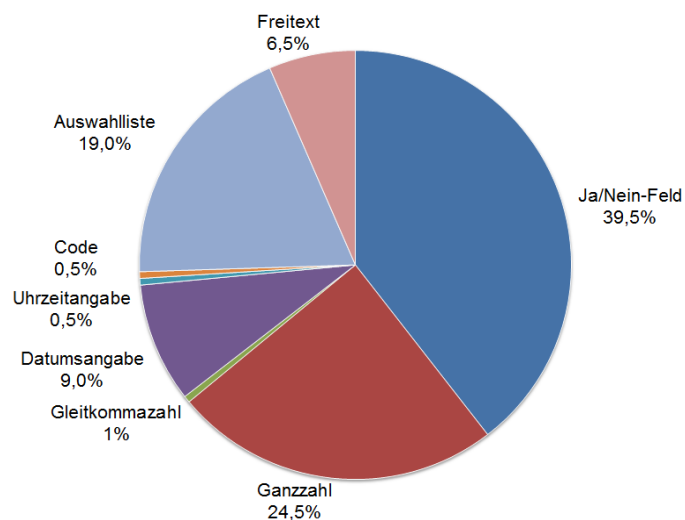


Abb. 4-3: Verteilung der Ausprägungstypen der 200 Merkmalsarten der cTAG-Studie.

Im Anschluss an die Analyse der Abläufe bei der Datenerhebung im Studienzentrum wurden gemäß des generischen Vorgehensmodells für die multiple Nutzung von Daten in Forschung und Versorgung (siehe Kapitel 3.1) die im KIS erfassbaren sowie die in der cTAG-Studie benötigten Merkmalsarten systematisch analysiert und einander gegenübergestellt. Zur Identifikation und Beschreibung der im KIS dokumentierbaren Merkmalsarten musste die Dokumentation von insgesamt vier in die cTAG-Studie eingeschlossenen Patienten analysiert werden. Bei der Analyse der Dokumentation des vierten Patienten zeigte sich, dass die Menge der identifizierten Merkmalsarten saturiert war. Die Gruppe der zufällig ausgewählten Patienten setzten sich zusammen aus einer Frau und einem Mann zwischen 40 und 50 Jahren sowie einer Frau und einem Mann zwischen 70 und 80 Jahren.

KIS Heidelberg

In welchen Formaten werden die Merkmalsarten im KIS dokumentiert?

Im Heidelberger KIS wurden 1452 Merkmalsarten identifiziert, zu welchen in 1138 Fällen (78,4 %) auch Ausprägungen dokumentiert worden waren. Mehr als die Hälfte der Merkmalsausprägungen (762 / 52,5 %) lagen elektronisch vor. Bei den KIS-Datenelementen werden die Ausprägungen bei 306 Merkmalsarten (21,1 %) durch die Selektion eines Wertes aus einer Auswahlliste erfasst; bei ähnlich vielen Merkmalsarten (300 / 20,7 %) werden die Ausprägungen in Form von Freitexten erfasst, in 253 Fällen (17,4 %) als Ganzzahl, in 248 Fällen (17,1 %) als ‚Ja/Nein‘-Felder. Die restlichen Merkmalsarten teilen sich wie folgt auf: Gleitkommazahlen 116 Merkmalsarten (8,0 %), Datumsangaben 107 Merkmalsarten (7,4 %), Angaben von Uhrzeiten 49 Merkmalsarten (3,4 %) sowie codierte Merkmalsausprägungen bei 44 Merkmalsarten (3,0 %). In 29 Fällen (2,0 %) konnte der Typ der Ausprägung mittels der Aktenanalyse nicht eindeutig bestimmt werden. Dabei handelte es sich ausschließlich um Merkmalsarten, zu welchen keine Ausprägungen in den Subsystemen des KIS dokumentiert worden waren – vgl. Abb. 4-2.

Anhand der eCRF der cTAG-Studie und der zugehörigen Studiendatenbank wurden die für die Studie zu erhebenden Merkmalsarten analysiert. Tab. 4-1 gibt eine Übersicht der in der cTAG-Studie vorgesehenen Visiten, der jeweils ausgefüllten eCRF sowie der Zahl der mit den einzelnen eCRF erhobenen Merkmalsarten. Das eCRF zur Dokumentation von unerwünschten Ereignissen (Adverse Event) wird nur bei Bedarf eingesetzt. Zu jedem unerwünschten Ereignis werden mit diesem eCRF 30 Merkmalsarten erhoben.

cTAG-Studie

Bei der cTAG-Studie werden in der Regel acht eCRF je Patient ausgefüllt. Dies geschieht im Rahmen von sechs Visiten. Bei der ‚registration visit‘ wird neben einem Registrierungsbogen und einem allgemeinen eCRF ein spezielles eCRF ausgefüllt, je nachdem, welche Indikation eine cTAG-Implantation notwendig machte (‚Aneurysm‘, ‚Traumatic Transaction‘, ‚Dissection‘ bzw. ‚Other Aortic Pathologies‘ – vgl. Tab. 4-1). Samt des eCRF für unerwünschte Ereignisse besteht die Studie somit aus insgesamt 10 verschiedenen Datenerhebungsbogen.

In welchen Formaten werden die Merkmalsarten in der Studie dokumentiert?

Bei der Datenbestandsanalyse wurden 200 Merkmalsarten identifiziert, die im Rahmen der cTAG-Studie erhoben werden. Von den identifizierten Merkmalsarten in der Studie werden in 79 Fällen (39,5 %) die Ausprägungen als ‚Ja/Nein‘-Felder erfasst, in 49 Fällen (24,5 %) als Ganzzahl-Werte und in 38 Fällen (19,0 %) durch die Selektion eines Wertes aus einer Auswahlliste. Die verbleibenden 34 Merkmalsarten (17,0 %) verteilen sich auf Datumsangaben (18 Merkmalsarten / 9,0 %), Freitextangaben (13 Merkmalsarten / 6,5 %) sowie Uhrzeitangaben, codierte Merkmalsausprägungen und Gleitkommazahlen (je 1 Merkmalsart / 0,5 %) – vgl. Abb. 4-3.

Tab. 4-1: Übersicht der in der cTAG-Studie verwendeten eCRF und der vorgesehenen Visiten. Wird ein eCRF bei einer Visite ausgefüllt, ist angegeben, wie viele Merkmalsarten bei dieser Visite mit dem jeweiligen eCRF erhoben werden.

Visite:	Registration visit	Intervention visit	Follow-Up 1	Follow-Up 2	Follow-Up 3	Close-out visit
Subject Registration	3					
Pre-Treatment	38					
Aneurysm	14					
Traumatic Transaction	14					
Dissection	17					
Other Aortic Pathologies	17					
Treatment		50				
Follow-Up			9	9	9	
End of Trial						12
Adverse Event						

Wie hoch wäre der Anteil der Daten, der heute schon aus dem KIS für eine konkrete Studie übernommen werden könnte?

Von den 200 Merkmalsarten, die für die cTAG-Studie erhoben werden, könnten die vier Merkmalsarten Alter, Geschlecht, Gewicht und ‚Operation als Notfall‘ (2,0 %) direkt aus den elektronischen Subsystemen des KIS übernommen werden. Indem geeignete Transformationen zwischengeschaltet werden, würde es möglich 12 weitere Datenelemente pro Proband (6,0 %) aus dem KIS zu übernehmen. Weitere 23 (11,5 %) der für die Studie benötigten Merkmalsarten liegen im KIS als wenig strukturierte Informationen in Freitextelementen vor. Diese könnten gegebenenfalls mit Hilfe von Textmining-Algorithmen extrahiert werden (vgl. Abb. 4-4).

Wie hoch wäre dieser Anteil, wenn das KIS vollständig rechnerunterstützt wäre?

Darüber hinaus stünden 29 Merkmalsarten (14,5 %) zur Verfügung, wenn diese nicht konventionell, sondern elektronisch im KIS dokumentiert würden. Von diesen 29 Merkmalsarten wären wiederum nur 4 (2,0 % der in der cTAG-Studie zu erhebenden Merkmalsarten) direkt nutzbar. Ausgehend von der aktuellen papierbasierten Dokumentation müssten 16 dieser 29 Merkmalsarten vor einer multiplen Verwendung ebenfalls transformiert werden. Bei 9 dieser 29 Merkmalsarten wäre ein Textmining notwendig. Zwei Datenelemente (1,0 %) – ein konventionelles und ein elektronisches – könnten nach einer entsprechenden Transformation zusätzlich genutzt werden, wenn zu diesen tatsächlich Ausprägungen im KIS dokumentiert worden wären.

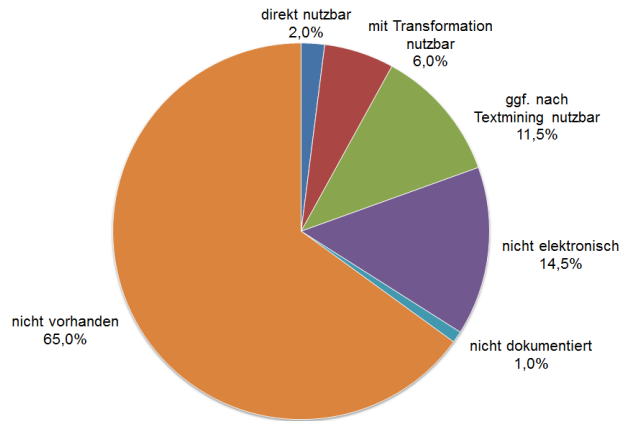


Abb. 4-4: Ergebnis der Analyse, welcher Anteil der 200 in der cTAG-Studie benötigten Merkmalsarten im KIS dokumentiert werden kann und unter welchen Voraussetzungen eine multiple Nutzung möglich bzw. nicht möglich ist.

Lägen alle Datenelemente im KIS elektronisch vor und würde zu allen vorgesehenen und für die Studie relevanten Merkmalsarten eine Ausprägung dokumentiert werden, könnten insgesamt 70 Datenelemente (35,0 %) für die cTAG-Studie multipel genutzt werden. Abb. 4-5 stellt dar, welche Anteile dieser 70 Merkmalsarten im KIS vor einer Übernahme in ein eCRF der cTAG-Studie wie transformiert werden müssten. So könnten 8 der 70 Datenelemente (11,4 %) direkt ‚1:1‘ übernommen werden. Bei 30 Merkmalsarten (42,9 %) müsste eine geeignete Transformation zwischengeschaltet werden. Bei weiteren 32 Merkmalsarten (45,7 %) wäre ein ‚Textmining‘ notwendig. In der cTAG-Studie werden 21 dieser 32 Merkmalsarten mit Hilfe von ‚Ja/Nein‘-Feldern erfasst, 3 Merkmalsarten als ‚Ganzzahl‘-Werte und 7 Merkmalsarten als ‚Auswahl: ‚1 aus n‘. Eine dieser Merkmalsarten wird zwar in der Studie ebenfalls als Freitext erhoben, dennoch wäre auch in diesem Fall eine Transformation in Form eines Textminings notwendig, um unnötige Textelemente zu entfernen.

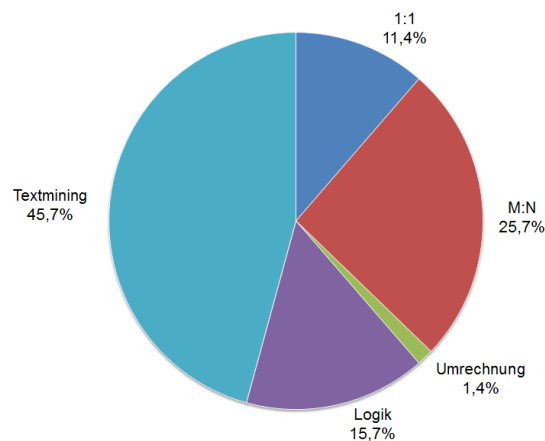


Abb. 4-5: Verteilung der notwendigen Transformationen, falls alle 70 der im KIS vorhandenen und für die cTAG Studie relevanten Merkmalsarten im KIS elektronisch vorlägen.

Über drei Viertel der 70 potenziell wiederverwendbaren Datenelemente im KIS könnten der papierbasierten Präsenzakte (53 Merkmalsarten / 75,7 %) entnommen werden. Dabei könnten mindestens 30 dieser 53 Datenelemente aus dem in der Präsenzakte abgelegten, papierbasierten OP-Bericht extrahiert werden. Darüber hinaus werden bestimmte Merkmalsarten wie zum Beispiel ‚Diagnosen‘ im OP-Bericht und in allen Arztbriefen aufgeführt. Aus dem i.s.h.med könnten 14 Datenelemente (20,0 %) und aus dem Radiologieinformationssystem 3 Datenelemente (4,3 %) multipel genutzt werden – vgl. Abb. 4-6. Einige Merkmalsarten sind mehrfach in verschiedenen Subsystemen verfügbar (beispielsweise Diagnosen oder Laborbefunde). Daher könnten in bestimmten Fällen auch CareVue oder die Ambulanzakte potenziell als Datenquellen genutzt werden.

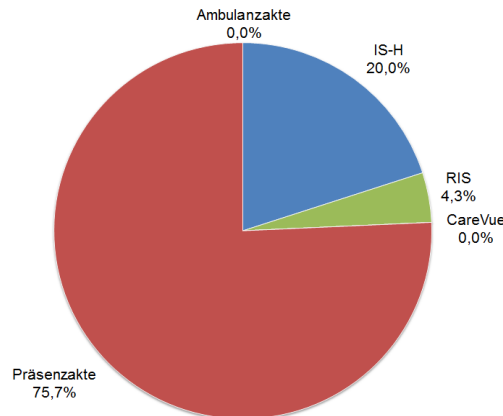


Abb. 4-6: Verteilung der 70 Merkmalsarten, die für die cTAG-Studie multipel genutzt werden könnten, auf die elektronischen und konventionellen Subsysteme des KIS.

Weitere 12 der in der Studie benötigten Merkmalsarten (6,0 %) könnten ebenfalls automatisch ermittelt werden. Einerseits könnten hierzu Systeminformationen genutzt werden (beispielsweise das aktuelle Datum oder der Name des angemeldeten Benutzers). Andererseits könnten die benötigten Informationen auch aus zuvor erfassten Studiendaten abgeleitet werden (so zum Beispiel die Gesamtzahl der eingesetzten Stents als Summe der Anzahlen der eingesetzten Stents unterschiedlichen Typs). Insgesamt könnten auf diese Weise 82 Datenelemente (41,0 %) multipel verwendet werden.

In der Fallstudie wurden, dem generischen Vorgehensmodell gemäß Kapitel 3.1 folgend, 1452 Merkmalsarten des Heidelberger KIS und 200 in der cTAG-Studie zu erhebenden Merkmalsarten erfasst und analysiert. Dabei zeigte sich, dass dieses Vorgehen sehr arbeitsintensiv werden kann. Daher sind rechnerbasierte Werkzeuge wünschenswert, die helfen, die einzelnen Arbeitsschritte zu vereinfachen und zu beschleunigen. Im Folgenden werden die Anforderungen an entsprechende Werkzeuge bezogen auf die einzelnen Arbeitsschritte (vgl. Kapitel 3.1) aufgelistet:

Schritt 1

Erfassung der verfügbaren Merkmalsarten (KIS)

Systematisches Erfassen aller Merkmalsarten, die im KIS typischerweise zu einem bestimmten Patientenkollektiv dokumentiert werden können: Ein unterstützendes System sollte dabei in der Lage sein, die einzelnen KIS-Merkmalsarten jeweils samt beschreibender Meta-Merkmale, zum Beispiel Datentyp und mögliche Ausprägungen, strukturiert in einer Datenbank zu erfassen. Neben einrichtungs- bzw. systeminterner Benennung sollte jeweils auch eine Bezeichnung gemäß einer standardisierten Terminologie erfassbar sein. Dies kann die Suche nach geeigneten Zuordnungen vereinfachen – Schritt 3 des Vorgehensmodells. Zusätzlich wäre es sinnvoll, Informationen darüber zu erfassen, mit welchem Verfahren ein

Datenelement im KIS erhoben wird. So kann abgeschätzt werden, wie verlässlich eine dokumentierte Ausprägung ist. Ferner sollte auch der zeitliche Kontext der Erfassung eines Datenelements innerhalb des Behandlungsprozesses geeignet dokumentiert werden (vgl. [DORDA, GALL et al. 2002]).

Erfassung der benötigten Merkmalsarten (Studie)

Schritt 2

Systematisches Erfassen aller Merkmalsarten sowie deren Meta-Merkmale, die in der Studie bzw. für die jeweilige Auswertungsfrage benötigt werden: Dies erfolgt analog zur Erfassung der im KIS dokumentierten Merkmalsarten im vorhergehenden Schritt.

Zuordnung geeigneter Merkmale

Schritt 3

Ermitteln geeigneter Zuordnungen von Merkmalsarten im KIS zu den für die Studie bzw. Auswertungsfrage benötigten Merkmalsarten: Das System soll hierbei die Suche nach geeigneten Abbildungen unterstützen. So könnten beispielsweise basierend auf den erfassten Meta-Merkmalen geeignete Zuordnungen automatisch vorgeschlagen werden. Ein angeschlossenes medizinisches Terminologiemanagementsystem könnte dabei unterstützen, korrespondierende Merkmalsarten auch dann anhand deren Bezeichnungen zu identifizierend, wenn synonyme Bezeichnungen verwendet wurden.

Über die Unterstützung bei der Identifikation der Merkmalsarten hinaus sollte das System in der Lage sein, die ermittelten Zuordnungen strukturiert abzuspeichern. Trotz Unterstützung beim Finden geeigneter Abbildungen, wird der Aufwand für diesen Schritt – je nach Umfang der jeweiligen Studie – nicht unerheblich sein. Indem Zuordnungen abgespeichert werden, können diese selbst multipel genutzt und so Aufwand eingespart werden – beispielsweise dann, wenn in Studien in einem medizinischen Fachgebiet immer wieder dieselben Merkmalsätze erhoben werden.

Direkte Zuordnungen der Merkmalsarten im KIS zu für die Studie bzw. Auswertungsfrage benötigten Merkmalsarten sind nicht immer möglich. Im untersuchten Fall waren direkte Abbildungen nur in 8 Fällen (11,4 % der 70 nutzbaren Merkmalsarten im KIS) der identifizierten Zuordnungen möglich. Sehr häufig müssen daher geeignete Transformationen zwischengeschaltet werden, um die im Rahmen einer Zuordnung vorgesehene multiple Nutzung erst zu ermöglichen. Wünschenswert wäre hier ein System, welches es erlaubt, entsprechende Transformationen zu definieren und ergänzend zu den erfassten Zuordnungen ebenfalls strukturiert abzuspeichern. Auf diese Weise könnten neben den Daten nicht nur die Zuordnungen multipel genutzt werden, sondern auch die hierzu benötigten Transformationen.

Implementierung

Schritt 4

Implementierung der identifizierten Zuordnungen und Transformationen: Falls ein Single-Source-Ansatz oder eine durch ‚pre-population‘ unterstützte Datensammlung in einem eigenständigen SDMS vorgesehen ist, müssen gegebenenfalls das KIS erweitert oder entsprechende Schnittstellen zwischen den beteiligten Systemen geschaffen werden. Das unterstützende Werkzeug sollte die Implementierung entsprechender Schnittstellen samt gegebenenfalls notwendigen Transformationen basierend auf den zuvor erfassten Zuordnungen und Transformationsvorschriften automatisieren oder zumindest vereinfachen.

4.2 Prototyp

Mit dem open Study Data Management System (openSDMS) wurde exemplarisch ein auf Archetypen basierendes System zur Integration von Forschung und Versorgung entwickelt. Diese Entwicklung bildete die Grundlage für die Ableitung weiterer (technischer) Anforderungen an eine auf Archetypen basierende Systemarchitektur zur Integration von Forschung und Versorgung.

OpenSDMS basiert auf dem *openEHR* Reference Framework and Application (Opereffa) System (siehe Kapitel 3.3). OpenSDMS wurde so implementiert, dass zusätzlich zu der von Opereffa bereitgestellten Funktionalität einer EPA (siehe Abb. 4-7) die Möglichkeit besteht, basierend auf Archetypen Studien abzubilden (Erfassen von Studien-Metadaten) und Studiendaten zu den vorhandenen Patienten zu erheben. Hierzu wurden zwei neue Sichten generiert: ‚Studie / CRF‘ (siehe Abb. 4-8) und ‚Studie / Metadaten‘ (siehe Abb. 4-9). In der Ansicht ‚Studie / CRF‘ können nicht nur Studiendaten zu einem Patienten erfasst werden, sondern auch geeignete Daten, die zuvor in der Patientenakte von openSDMS erfasst wurden, ‚per Klick‘ in das CRF übernommen werden. Somit müssen nur noch die Studiendaten manuell erfasst werden, die nicht aus der EPA übernommen werden können. Die einzelnen Abläufe bei der Nutzung von openSDMS werden in Anhang 5 detailliert beschrieben.

In der Ansicht ‚Studie / Metadaten‘ können alle Daten erfasst werden, welche Studie, Studienarme und Visiten beschreiben – also beispielsweise der Name oder das Akronym einer Studie. Da auch die Studien-Metadaten durch hierfür entwickelten Archetypen (vgl. Kapitel 4.4.2) repräsentiert werden, können systemintern zum Speichern der Studien-Metadaten dieselben Mechanismen genutzt werden wie zum Speichern der Studiendaten. Anstelle der Patienten-Identifikationsnummern wird für Studien-Metadaten in der EAV-Tabelle eine Studien-Identifikationsnummern gekennzeichnet. Patienten-Identifikationsnummern und Studien-Identifikationsnummern werden in getrennten Listen verwaltet, um diese unterscheiden zu können. Da in der EAV-Tabelle jedoch technisch nicht zwischen Daten von Patienten und Studien unterschieden wird, muss jedes Element der Vereinigungsmenge aus Patienten-Identifikationsnummern und Studien-Identifikationsnummern von den übrigen Elementen verschieden sein. Jedes Element muss eindeutig einen Patienten *oder* eine Studie identifizieren.

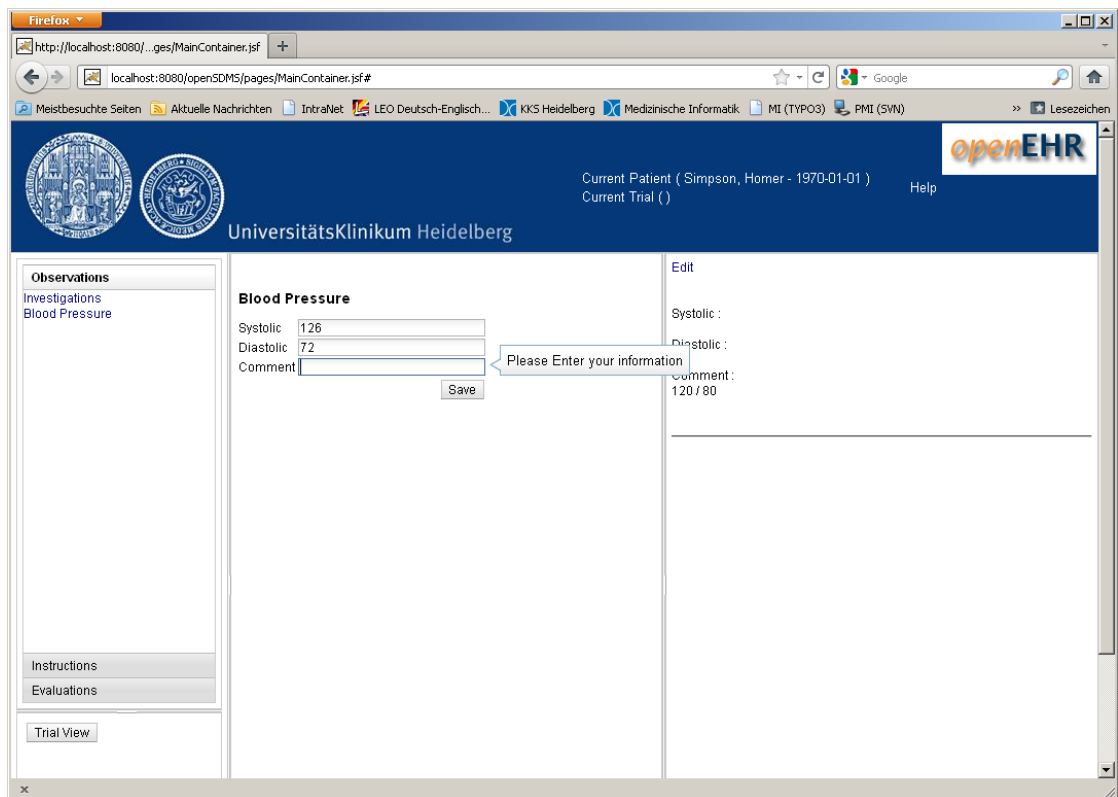


Abb. 4-7: In der Ansicht ‚Patientenakte‘ von openSDMS können Daten aus der Routineversorgung eines Patienten erfasst, eingesehen und falls notwendig verändert werden.

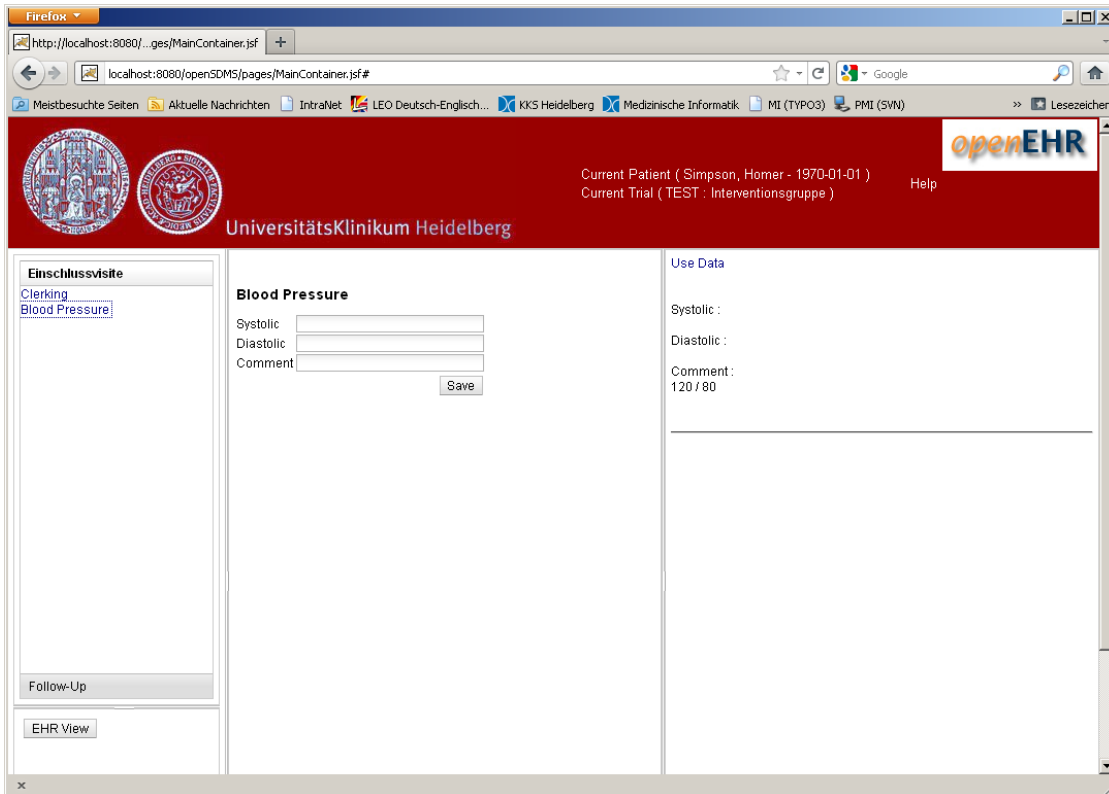


Abb. 4-8: In der Ansicht ‚Studie / CRF‘ von openSDMS können Studiendaten zu einem Patienten erfasst, eingesehen und korrigiert werden. Ferner ist es möglich, Daten aus der EPA des Patienten in eine Studie zu übernehmen.

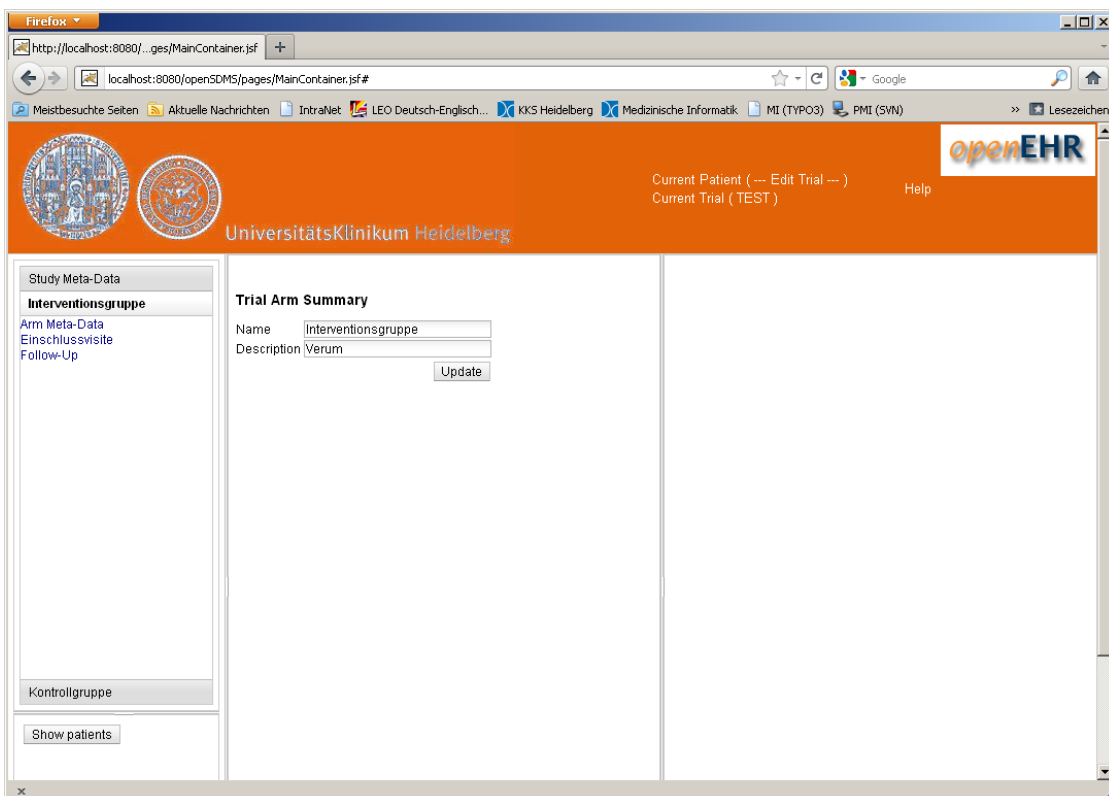


Abb. 4-9: In der Ansicht ‚Studie / Metadaten‘ von openSDMS können Studien-Metadaten erfasst, eingesehen und verändert werden.

Die systeminterne Datenhaltung ist als Single-Source-Lösung realisiert. Studiendaten werden physisch auf dieselbe Art und Weise gespeichert wie Einträge in der EPA. Dementsprechend werden studienbezogene Einträge zu einem Patienten automatisch auch in dessen EPA angezeigt. Daten aus der EPA werden durch eine logische Verknüpfung einer bestimmten Studie, einem Studienarm und einer darin vorgesehenen Visite zugeordnet und so als Studiendaten gekennzeichnet. Zur Implementierung dieser Funktionalität wurde die originale EAV-Tabelle von Opereffa um drei zusätzliche Spalten erweitert. Entsprechend der generischen Studienstruktur (vgl. Kapitel 4.4.2) können in openSDMS zu einem Element zusätzlich zu Patienten-ID, Archetypen-Bezeichner und –Pfad sowie eigentlichem Wert auch

- Eine Studien-ID,
- die ID eines Studienarms und
- die ID einer Visite

gespeichert werden. So kann ein Element eindeutig einem bestimmten Erhebungszeitpunkt innerhalb einer Studie zugeordnet werden. Bei Daten, die lediglich im Rahmen der Funktionalität des Systems als Patientenakte dokumentiert wurden, bleiben die zusätzlichen Spalten leer.

Da sowohl EPA- als auch Studiendaten von denselben Archetypen beschrieben werden, kann direkt ermittelt werden, welche der in der EPA vorhandenen Daten in einer Studie multipel nutzbar sind. In der EPA von openSDMS muss lediglich nach Daten gesucht werden, die auf demselben Archetyp basieren wie das jeweilige Feld im CRF.

Zur Strukturierung der Daten nutzt openSDMS das Konzept der *openEHR*-Folders (siehe Kapitel 2.6). Die Struktur einer Studie (Anzahl der Arme, Anzahl der Visiten je Arm, zu erfassende Daten innerhalb der Visite) werden innerhalb von openSDMS über eine spezielle XML-Datei (*structure.xml*) konfiguriert. Im Gegensatz zum statisch konfigurierten Opereffa ist es in openSDMS auch möglich, über diese XML-Datei die Struktur der EPA-Komponente zu konfigurieren (vgl. Abb. 3-2). Bei jedem Start von openSDMS wird die Datei *,structure.xml‘* gegen die korrespondierende Schema-Datei *,structure.xsd‘* validiert.

Um die Zuordnung von Patienten zu einer Studie im System speichern zu können (M:N-Relation), wurde für openSDMS in der Opereffa Datenbank eine zusätzliche Tabelle angelegt (*patient_trial*), welche diese M:N-Relation abbildet.

Anwendungsszenarien von openSDMS

OpenSDMS kombiniert die Funktionalität einer EPA und eines SDMS und ermöglicht so die multiple Verwendung von Daten aus der EPA in Studien. Neben der Nutzung von openSDMS als eigenständiges System (**Stand-Alone-System**) ist es auch denkbar, Daten aus anderen KIS, mit Hilfe geeigneter Mechanismen in die EPA von openSDMS zu importieren (**Integrationssystem**). Das Szenario ‚Integrationssystem‘ wäre in multizentrischen Studien hilfreich, um die in den KIS der beteiligten Zentren vorhandenen Daten für eine multiple Nutzung zusammenzuführen (siehe Abb. 4-10). Die EPA von openSDMS würde dabei nur noch als Zwischenspeicher für die KIS-Daten der angeschlossenen Studienzentren genutzt, um diese Daten für eine multiple Nutzung anbieten zu können. Die manuelle Datenerfassung in der EPA von openSDMS entfiel. Der Anwender würde nur noch die SDMS Funktionalität nutzen, um importierte Daten in die Studie zu übernehmen oder um Studiendaten manuell zu ergänzen. Ein echter Single-Source-Ansatz wäre dabei aber nicht realisierbar, da die Quelldaten in ihren Quellsystemen in den einzelnen Studienzentren und die Studiendaten hiervon getrennt in openSDMS erfasst würden.

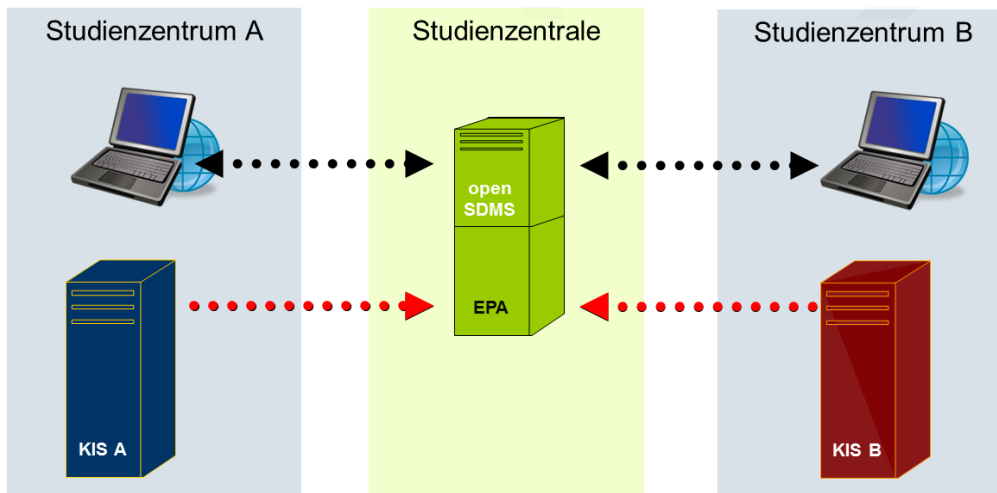


Abb. 4-10: Datenflüsse bei der Nutzung von openSDMS als Integrationssystem, welches die multiple Nutzung von Daten aus verschiedenen KIS für Studienzwecke erlaubt (hierzu müsste openSDMS um geeignete Import-Schnittstellen erweitert werden).

4.3 Referenzarchitektur

4.3.1 Anforderungen an die Referenzarchitektur

Basierend auf den Erkenntnissen aus der Fallstudie und der Entwicklung von openSDMS ergeben sich folgende Anforderungen an eine Referenzarchitektur zur Integration von KIS und SDMS. Dabei werden besonders auch multizentrische Studien berücksichtigt.

- Zu den in einem KIS dokumentierbaren Merkmalsarten müssen semantische Beschreibungen vorliegen. Ist dies nicht gegeben, sollten Werkzeuge verfügbar sein, die das Erstellen eines Katalogs von semantischen Beschreibungen der in einem KIS dokumentierbaren Merkmalsarten unterstützen.
- In verschiedenen KIS vorhandene Daten müssen strukturiert und über eine einheitliche Schnittstelle zugreifbar sein. Nur so ist es möglich, die für eine multiple Nutzung potenziell geeigneten KIS-Daten aufzulisten. Diese Schnittstelle muss auch sehr feingranulare und komplexe Abfragen zulassen, wie zum Beispiel ‚liefere alle systolischen Blutdruckwerte des aktuellen Patienten, welche innerhalb der letzten 3 Stunden durch invasive arterielle Messung ermittelt wurden‘. Je exakter Daten im KIS ausgewählt werden können, desto besser. Die Menge der gefundenen Daten wird so in der Regel kleiner. Die Wahrscheinlichkeit, dass ein im KIS gefundenes Datenelement tatsächlich in der jeweiligen Studie nutzbar ist, wächst. Der Aufwand des Prüfarztes oder der Study Nurse beim Prüfen und Auswählen der für die Studie geeigneten Datenelemente sinkt.
- Während in KIS Patienten meist über Ihren Name und Geburtsdatum identifiziert werden, werden die Probanden in SDMS üblicherweise über ein Pseudonym identifiziert. Wenn eine Schnittstelle genutzt werden soll, um Daten aus einem KIS in ein SDMS zu übernehmen, muss diese Schnittstelle in der Lage sein, das Pseudonym eines Patienten im SDMS in die korrekte Patienten-ID des KIS zu überführen und umgekehrt.
- Soweit möglich sollten zu den einzelnen KIS-Daten Kontextinformationen (beispielsweise zum Vorgehen oder den eingesetzten Geräten bei der Datenerhebung) abrufbar sein. Diese sind notwendig, um entscheiden zu können, ob ein Datenelement aus dem KIS tatsächlich für die multiple Nutzung in einer Studie geeignet ist.

- Im SDMS müssen Studiendaten hoch strukturiert dokumentiert werden, um für exakt definierte Auswertungsfragen bzw. Analyseverfahren nutzbar zu sein. Neben Studiendaten sollte das SDMS auch in der Lage sein, Studien-Metadaten wie Studientitel, Indikation oder Ein-/Ausschlusskriterien strukturiert zu erfassen. Ferner wäre es wünschenswert auch den Satz der Studien-Metadaten bei Bedarf ähnlich flexibel wie ein eCRF erweitern zu können.
- Die Definition der eCRF im SDMS (= Definition der zu erhebenden Studiendaten) muss so gestaltet sein, dass einzelne Datenelemente und ganze eCRF in verschiedenen Studien wiederverwendet werden können. Wenn in zwei Studien die gleiche Merkmalsart erhoben werden soll, muss gewährleistet werden, dass hierfür dieselbe, eindeutige Elementdefinition verwendet wird.
- Die Definition der zu erhebenden Studiendaten im SDMS und die Schnittstelle zum Zugriff auf im KIS gespeicherte Daten, müssen derselben Systematik folgen. So können teilweise aufwändige Transformationen bei der Übernahme von KIS-Daten in das SDMS vermieden werden.
- Werden KIS-Daten in ein SDMS übernommen, müssen neben den eigentlichen Daten auch Kontextinformationen wie Quellsystem und Erhebungsmethode in das SDMS übernommen werden können. Indem das Quellsystem (KIS) eines ins SDMS importierten Datenelements bis auf Subsystemebene eindeutig identifizierbar bleibt, werden regulatorische Anforderungen erfüllt (siehe folgenden Abschnitt). Ebenso wird ein späterer Quelldatenabgleich erleichtert.
- Wurden KIS-Daten in ein SDMS übernommen, können nachfolgende Änderungen der Daten in KIS oder SDMS zu Inkonsistenzen mit dem jeweils anderen System führen. Daher sind technische und organisatorische Maßnahmen notwendig, die dafür sorgen, dass solche Inkonsistenzen vermieden bzw. erkannt und beseitigt werden.
- Unabhängig von der technischen Realisierung eines SDMS sollte es möglich sein, Studiendaten von jedem Rechner mit Internetzugang mittels Webschnittstelle und ohne die Installation zusätzlicher Softwarekomponenten zu erfassen. Ferner sollte es auf diese Art auch möglich sein, Daten aus KIS, welche an das SDMS angeschlossen sind, in eine Studie zu übernehmen. Auf diese Weise lassen sich Kollisionen zwischen den Voraussetzungen zur Nutzung des SDMS und den Sicherheits- und Firewall-Richtlinien in den einzelnen Studienzentren vermeiden. Ist beispielsweise die Installation einer speziellen Zugangsoftware oder der Aufbau eines virtuellen privaten Netzwerks (VPN) notwendig, um das SDMS nutzen zu können, kann dies den Einsatz des Systems in manchen Studienzentren erschweren. Trotz der zu Recht meist sehr restriktiven Firewall-Konfigurationen und Sicherheitsrichtlinien in Gesundheitsversorgungseinrichtungen, ist ein Zugriff auf Web-Seiten in der Regel jedoch problemlos möglich.
- Eine Referenzarchitektur zur Verknüpfung von KIS und SDMS sollte etablierte Modelle wie CDASH, BRIDG und ODM soweit möglich integrieren. So wird eine Kompatibilität zu bestehenden Systemen, welche diese Modelle implementieren, ermöglicht.

Regulatorische Anforderungen

Ergänzend zur durchgeführten Fallstudie und der Entwicklung von openSDMS wurde geprüft, welche regulatorischen Anforderungen an Systeme zur Integration von Forschung und Versorgung gestellt werden. Hierbei wurde von den Anforderungen ausgegangen, die generell an Systeme gestellt werden, mit welchen studienbezogene Daten verarbeitet werden. Solche Anforderungen ergeben sich aus den Richtlinien 2001/20/EG und 2005/28/EG der Europäischen Union, welche wiederum die Good Clinical Practices (Note for Guidance on Good Clinical Practices CPMP/ICH/135/95 – [ICH 2002]) referenzieren und konkretisieren. In Deutschland setzen das Arzneimittelgesetz und die GCP-Verordnung diese Richtlinien in nationales Recht um. Die Einhaltung der GCP wird darüber hinaus immer öfter auch bei öffentlichen Ausschreibungen als Voraussetzung für eine Förderung gefordert. Eine gute Zusammenfassung der technischen Anforderung an studiendatenverarbeitende elektronische

Systeme, die sich aus den europäischen Richtlinien ergeben, liefert ([EMA 2010]). In den Vereinigten Staaten von Amerika hat die dort zuständige Food and Drug Administration ihre Erwartungen auf ähnliche Weise zusammengefasst ([FDA 2010]). Die von EMA und FDA genannten Anforderungen sind dabei im Wesentlichen deckungsgleich. Aus [ICH 2002] als Grundlage und [EMA 2010] ergeben sich für computerbasierte SDMS und KIS vor allem folgende Anforderungen:

- Die eingesetzten Systeme und ihre Schnittstellen müssen validiert sein.
- Es muss ein Audit-Trail geführt werden, der es erlaubt jede Änderung und somit die gesamte Historie der Datenerfassung in ihrer originalen Chronologie nachzuvollziehen. Jederzeit muss eindeutig und verlässlich nachvollziehbar sein, wer wann welche Daten eingetragen bzw. wie und warum geändert hat.
- Darüber hinaus müssen in allen KIS, in denen Daten von Patienten verarbeitet werden, die an entsprechenden klinischen Studien teilnehmen, die folgenden Eigenschaften nach [EMA 2010] in Bezug auf Quelldaten (englisch: source data) gewährleistet sein. Die Daten müssen
 - genau bzw. fehlerfrei sein, mindestens in gleichen Maß wie bei der papierbasierten Erhebung
 - unmittelbar lesbar und für einen Menschen zum Zeitpunkt der Ein- und Ausgabe verständlich sein
 - gleichzeitig mit der Beobachtung erfasst werden
 - original in dem Sinn sein, dass es sich um den ursprünglich erzeugten Datensatz handelt
 - zuordenbar sein. Das bedeutet, dass eine Aktion – etwa eine Dateneingabe – eindeutig der ausführenden Person, zum Beispiel einem Arzt oder einer Pflegekraft, zuordenbar sein muss. Gleichzeitig muss durch geeignete Maßnahmen wie beispielsweise der Eingabe von Benutzername und Passwort sichergestellt sein, dass diese Aktion tatsächlich von der angegebenen Person durchgeführt wurde.
 - vollständig sein
 - einheitlich und widerspruchsfrei sein
 - dauerhaft und sicher gespeichert sein sowie
 - bei Bedarf verfügbar sein.

4.3.2 Spezifikation der Referenzarchitektur

Ausgehend von den in Kapitel 4.3.1 genannten Anforderungen und basierend auf den Erkenntnissen aus der Fallstudie sowie der Planung und Entwicklung von openSDMS konnte eine auf Archetypen basierende Referenzarchitektur zur Integration von Forschung und Versorgung spezifiziert werden. Systeme, die gemäß dieser Architektur implementiert werden, ermöglichen es, Daten aus Krankenhausinformations- und Studiendatenmanagementsystemen multipel zu nutzen. Die einzelnen Komponenten der spezifizierten Referenzarchitektur sowie das Zusammenwirken dieser Komponenten werden im Folgenden detailliert beschrieben. Als Beispiel wird dazu ein Szenario gewählt, in welchem die KIS-Daten aus zwei Studienzentren (KIS A und KIS B) für eine multiple Nutzung in das SDMS der Studienzentrale importiert werden. Um die Unterschiede beim Import aus auf Archetypen basierenden und aus nicht auf Archetypen basierenden KIS darzustellen, wird angenommen, dass nur KIS B auf Archetypen basiert (vgl. Abb. 4-11).

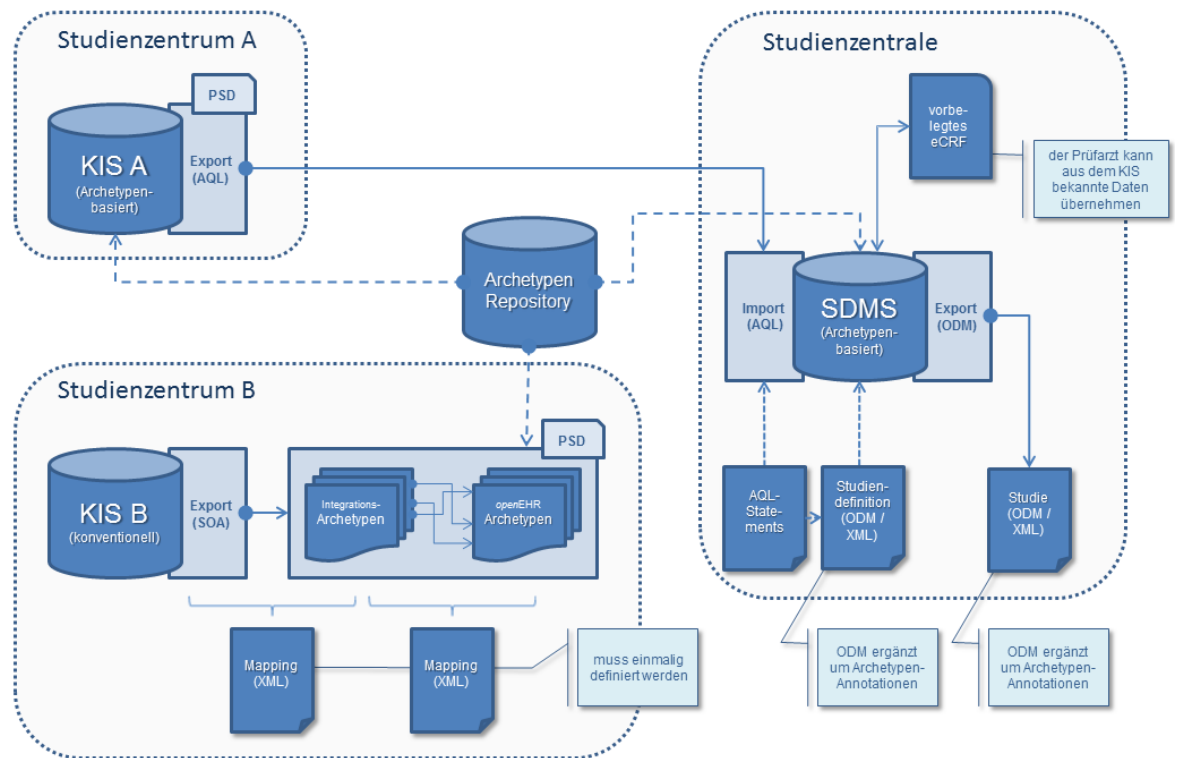


Abb. 4-11: Auf Archetypen basierende Referenzarchitektur für die multiple Nutzung von KIS-Daten in klinischen Studien.

Der spezifizierten Architektur folgend können prinzipiell beliebig viele KIS an ein SDMS angebunden werden. Ebenso wäre es möglich die Daten der angeschlossenen KIS für eine multiple Nutzung in verschiedene SDMS in unterschiedlichen Studienzentralen zu importieren. Die vorgeschlagene Referenzarchitektur ist so aufgebaut, dass je nach Situation und je nachdem ob bereits vorhandene Systeme auf Archetypen basieren oder nicht, die benötigten Komponenten ausgewählt werden können. Die zentralen Komponenten der Architektur bilden das SDMS und das Archypen-Repository. Nachfolgend wird zunächst der Aufbau des SDMS näher betrachtet.

ODM mit Referenzen auf Archetypen

Ein wesentlicher Bestandteil der vorgeschlagenen Referenzarchitektur ist die Erweiterung des Operational Data Model (vgl. Kapitel 2.7) um semantische Beschreibungen der einzelnen Datenelemente mit Hilfe von Referenzen auf Archetypen. Dazu wird der im ODM enthaltene ‚Alias-Mechanismus‘ genutzt, der es erlaubt, die Beschreibung von sämtlichen ODM-Elementen (vor ODM-Version 1.3.1 nur von ItemDefs und ItemGroupDefs) um zusätzliche Bezeichnungen zu ergänzen. Ein Alias setzt sich dabei zusammen aus einer Kontextangabe (Context), welche den Anwendungsbereich für den jeweiligen Alias spezifiziert und dem eigentlichen Aliasname (Name) – vgl. [CDISC 2010b] Kapitel 3.1.1.3.6.6. Der Referenzarchitektur folgend, wird eine mit Archetypen annotierte ODM-Datei genutzt, um die Struktur der Studie und die eCRF samt den zu erhebenden Datenelementen im SDMS zu definieren.

```

[...]
<ItemDef      OID="ID.ANYCONMED"
              Name="Any Concomitant Medications"
              DataType="text" Length="1">
  <Description>
    <TranslatedText xml:lang="en">
      Any Concomitant Medications
    </TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">
      Is the Subject taking any Concomitant
      Medications?
    </TranslatedText>
  </Question>
  <Alias Context="Archetype"
        Name="openEHR-EHR-COMPOSITION.cdash_cm.v1/
              context/other_context/items[at0010]/
              value/value"/>
  <CodeListRef CodeListOID="CL.YESNO"/>
</ItemDef>
[...]

```

Abb. 4-12: Beispiel für die Nutzung des ODM-Alias-Mechanismus zur semantischen Annotation mit Hilfe eines Archetypen.

Abb. 4-12 veranschaulicht die Nutzung des ODM-Alias-Mechanismus zur semantischen Annotation mit Hilfe eines Archetypen am Beispiel der in klinischen Studien typischen Frage nach Begleitmedikationen (eine umfangreichere ODM-Datei mit weiteren semantischen Annotationen findet sich in Anhang 4). Der Kontext des Alias wurde mit ‚Archetype‘ benannt und als Name der Verweis auf die entsprechende Merkmalsart im Archetyp `openEHR-EHR-COMPOSITION.cdash_cm.v1` angegeben. Dieser Verweis setzt sich zusammen aus Archetyp-Name und Pfadangabe innerhalb des Archetypen. Gemäß den *openEHR*-Spezifikationen kann aus dem Namen des Archetypen auf das zugrundeliegende Referenzmodell (hier ‚openEHR-EHR‘) geschlossen werden. Somit ist das dargestellte ODM-Element mit Hilfe des Archetyps `openEHR-EHR-COMPOSITION.cdash_cm.v1` durch den Alias eindeutig semantisch beschrieben. Der Archetyp definiert das ODM-Element als Indikator, welcher anzeigt, ob der jeweilige Studienpatient parallel zur Studie eine gemäß Studienprotokoll relevante Medikation bzw. Therapie erhalten hat.

Alternativ zur direkt ODM-basierten Definition der einzelnen eCRF-Seiten könnte ferner das *openEHR*-Template-Model genutzt werden. Unter Hinzunahme weiterer Strukturinformationen (Anzahl von Studienarmen sowie den jeweils vorgesehenen Visiten) könnte ein Anwendungsbaustein implementiert werden, der aus auf *openEHR*-Templates basierenden Bildschirmformularen automatisch mit Archetypen-Referenzen annotierte ODM-basierte Studiendefinition generiert.

Ein weiterer Baustein der Referenzarchitektur ist der Export der erfassten Studiendaten aus dem SDMS in das ODM-Format. Indem eine mit Archetypen annotierte ODM-Datei zur Definition von Studien- und eCRF-Struktur im SDMS genutzt wurde, bildet sich diese Semantik auch auf den ODM-basierten Export der erfassten Studiendaten ab.

Beispiel

openEHR- Templates

ODM-Export

Pre-population

Ein zentrales Element der Referenzarchitektur, ist die Implementierung eines Mechanismus im SDMS, welcher beim Ausfüllen der eCRF zu jedem Eingabefeld geeignete Daten aus den KIS vorschlägt, die dann direkt (per Mausklick) bestätigt und übernommen werden können. Voraussetzung ist, dass die zum jeweiligen Probanden benötigten Informationen bereits in einem angebotenen KIS erfasst wurden. Damit ein Prüfarzt entscheiden kann, ob ein vorhandener Wert tatsächlich in einer Studie nutzbar ist, sollten Kontextinformationen zu diesem Wert angezeigt werden. So beispielsweise Informationen zur Untersuchungsmethode bzw. dem Messgerät, mit welchem das Merkmal erhoben wurde oder zum Zustand des Patienten während der Beobachtung soweit diese im jeweiligen KIS dokumentiert wurden – zumindest der Erhebungszeitpunkt sollte angezeigt werden.

Alle in einer Studie benötigten Informationen, die nicht aus einem KIS abgeleitet werden können, müssen manuell im SDMS erfasst werden. Das SDMS muss alle Funktionen zum Übernehmen der in den KIS vorhandenen Daten sowie zum Erfassen der in einer Studie zusätzlich benötigten Informationen mit Hilfe einer Web-basierte Benutzerschnittstelle ermöglichen.

AQL

Um Daten aus KIS übernehmen zu können, benötigt das SDMS eine auf Archetypen basierende Importschnittstelle. Diese Schnittstelle muss es ermöglichen, mittels Archetype Query Language (vgl. Kapitel 2.6) für eine Studie und einen Patienten geeignete, exakt definierte Datensätze aus den angeschlossenen KIS abzurufen – die KIS müssen hierzu einheitliche Exportschnittstellen bereitstellen. Für jedes CRF-Feld, das eine Übernahme von KIS-Daten erlauben soll, muss eine dedizierte auf AQL basierende KIS-Abfrage hinterlegt werden. Auf diese Weise kann die Liste der angebotenen Werte, auf diejenigen beschränkt werden, die alle Anforderungen einer Studie erfüllen und somit für eine multiple Nutzung geeignet sind. Muss in einer Studie etwa der Blutdruck der Probanden invasiv in definierten Zeitabständen gemessen werden, kann mittels AQL die Menge der zu einem Patienten aus dem KIS ins SDMS importierten Blutdruckmessungen auf die invasiv durchgeführten Messungen und das entsprechende Zeitintervall eingeschränkt werden. Mit der AQL ist es ferner möglich, komplexe Anfragen zu formulieren, die es erlauben, Ausprägungen von Studienmerkmalen aus den Datenbeständen der KIS abzuleiten, die dort nur implizit vorliegen. So könnte bei entsprechender Datenlage aus den KIS-Daten beispielsweise abgeleitet werden, ob ein Patient eine bestimmte Kombination von Medikamenten einnimmt.

Pro Studie müssen zusätzlich zur Studiendefinition für alle Studienmerkmale, für welche multipel nutzbare KIS-Daten ermittelt werden sollen, geeignete AQL-Abfragen in einer eigenen Datei oder Datenbank hinterlegt werden. KIS-spezifische Anpassungen sind dabei nicht notwendig, da auf alle KIS über eine einheitliche Schnittstelle zugegriffen wird. Im einfachsten Fall können die AQL-Abfragen in einer strukturierten Textdatei oder in einer weiteren XML-Datei hinterlegt werden. Die Zuordnung einer AQL-Abfrage zu einem Datenelement in der Studie erfolgt mit Hilfe der OID, mit welcher das jeweilige Element in der ODM-basierten Studiendefinition eindeutig bezeichnet ist. Das Erstellen der benötigten AQL-Abfragen sollte unbedingt durch Generatorwerkzeuge unterstützt werden.

**Daten-
übertragung**

Für den physikalischen Datenaustausch zwischen KIS und SDMS erlaubt die Referenzarchitektur verschiedene Ansätze: Angefangen von einer direkten Netzwerkverbindung, über den in E-Mails gekapselten Austausch von AQL-Abfragen und Ergebnisdatensätzen bis hin zur manuellen Übertragung von Abfragen und Ergebnisdatensätzen mittels USB-Stick. Die gewählte Realisierung muss sich dabei an den jeweiligen Gegebenheiten wie Firewall- und Sicherheitsrichtlinien orientieren. Erfolgt eine Datenübertragung über das Internet muss eine geeignete Methode zur Verschlüsselung der ausgetauschten Daten (beispielsweise ein Virtual Private Network) implementiert werden.

**Auf Archetypen
basierende KIS**

Jedes auf Archetypen basierende KIS, welches an das SDMS angebunden werden soll, muss eine entsprechende AQL-basierte Abfrage- und Exportschnittstelle bereitstellen. Darüber hinaus sollte mit dieser Abfrageschnittstelle ein Pseudonymisierungsdienst verknüpft sein.

Während in KIS in der Regel die realen identifizierenden Daten eines Patienten angezeigt werden, werden in Studien Probanden über ein Pseudonym identifiziert. Die Zuordnungsliste von Patienten und Pseudonymen befindet sich dabei im Studienzentrum. Dementsprechend muss der Pseudonymisierungsdienst in der Lage sein, das in einer AQL-Anfrage enthaltene, studienbezogene Pseudonym in die Patienten-ID des jeweiligen KIS umzuwandeln und entsprechende Zuordnungslisten zu verwalten.

Bei der Umsetzung der *openEHR*-Spezifikationen muss hierbei berücksichtigt werden, dass die in den Systemen genutzten versionierten COMPOSITIONs über einen ‚Unique Identifier‘ verfügen. Wenn Daten aus einem auf Archetypen basierenden KIS in ein ebenfalls auf Archetypen basierendes SDMS exportiert würden, bestünde die Gefahr, dass zusätzlich zur Vergabe einer pseudonymisierten Kennung für die Studienteilnehmer, der Unique Identifier der übertragenen COMPOSITIONs als weiteres Pseudonym die Rückidentifizierung eines Patienten theoretisch ermöglicht. Um dies zu verhindern, müssen beim Datenexport alle Unique Identifier eindeutig durch ein kryptographisch geschütztes weiteres Pseudonym ersetzt werden.

Jedes *nicht* auf Archetypen basierende KIS muss prinzipiell über eine beliebige Export-Schnittstelle verfügen, um Daten extrahieren zu können. In Abb. 4-11 wird für das nicht auf Archetypen basierende KIS beispielsweise die Möglichkeit des Datenexports über eine SOA-Schnittstelle (serviceorientierte Architektur) angenommen und dargestellt; ebenso ist auch ein direkter SQL-basierter Zugriff möglich, falls die KIS-Daten in einer relationalen Datenbank gespeichert werden oder eine proprietäre Exportschnittstelle. Zusätzlich muss ein dediziertes ‚Integrationsmodul‘ bereitgestellt werden. Dieses Modul kapselt die KIS-spezifischen Export-Schnittstellen und erlaubt den AQL-basierten Zugriff wie auf ein Archetypen-basiertes KIS. Hierzu transformiert das Integrationsmodul die KIS-spezifische Repräsentation der dort gespeicherten Daten in eine auf Archetypen basierende Repräsentation. Für jedes nicht auf Archetypen basierende KIS muss ein eigenes Integrationsmodul bereitgestellt werden, welches auf die KIS-Daten zugreifen und diese entsprechend der AQL-Abfragen des SDMS exportieren kann. Hierzu müssen innerhalb des Integrationsmoduls Abbildungs- und Transformationsvorschriften definiert werden, welche den im KIS erfassten Datenelementen die entsprechenden, auf Archetypen basierenden Repräsentationen zuordnen.

Zur Implementierung dieser Transformationen können im Fall von *openEHR*-Archetypen sogenannte ‚Integration Archetypes‘ und das ‚*openEHR* Integration Information Model‘ ([BEALE 2006]) genutzt werden. Das *openEHR* Integration Information Model sieht vor, dass Daten aus Systemen, die nicht auf Archetypen basieren, zunächst mit Hilfe von Integrations-Archetypen repräsentiert werden. Darin soll die ursprüngliche Datenstruktur so gut wie möglich mit generischen Elementen abgebildet werden. Ausgehend von den Integration-Archetypen kann dann eine Abbildung der Daten auf die tatsächlich genutzten *openEHR*-Archetypen erfolgen. Dieses zweistufige Vorgehen hat den Vorteil, dass Abbildungen, die gegebenenfalls eine Datentransformation erfordern, nur zwischen *openEHR*-Archetypen definiert werden müssen – nämlich zwischen den Integration-Archetypen und den tatsächlich genutzten Archetypen (vgl. [BEALE 2006], S. 11f). Die Definition und konkrete Realisierung der Abbildungsvorschriften zwischen KIS und Integrations-Archetypen muss sich an den technischen Rahmenbedingungen des jeweiligen KIS orientieren. Dabei ist sowohl eine dynamische Transformation denkbar – definiert beispielsweise in einer XSLT-Datei (Extensible Stylesheet Language Transformation) – als auch eine statische Implementierung.

In diesem Fall ist, ebenso wie bei einem vollständig auf Archetypen basierenden KIS, ein Pseudonymisierungsdienst zu integrieren. Besteht ein KIS sowohl aus Subsystemen, die auf Archetypen basieren, als auch aus klassischen, nicht Archetyp-basierten Subsystemen, ist eine Kombination der beiden geschilderten Szenarien denkbar.

**Nicht auf
Archetypen
basierende KIS**

Archetypen-Repository

Ein zentrales Element der hier vorgestellten Systemarchitektur ist das Archetypen-Repository, welches von allen beteiligten, auf Archetypen basierenden Komponenten genutzt werden muss. Nur wenn alle involvierten Komponenten dieselben Archetypen nutzen, ist semantische Interoperabilität auf einfache Art möglich. Werden nicht dieselben Archetypen genutzt, ist auch hier eine Abbildung erforderlich. Dabei kann angenommen werden, dass die Abbildung der Datenelemente eines KIS ohne semantische Beschreibung auf eine auf Archetypen basierende Repräsentation schwieriger zu generieren ist, als die Abbildung des Archetypen-Sets eines Systems auf das Archetypen-Set eines anderen Systems. Speziell für die Abbildung von *openEHR*-Archetypen auf ISO 13606 Archetypen zeigten [MARTINEZ-COSTA, MENARGUEZ-TORTOSA et al. 2010], dass es möglich ist, aus *openEHR*-Archetypen automatisiert ISO 13606-Archetypen zu generieren und umgekehrt (siehe hierzu auch [MARTINEZ-COSTA, MENARGUEZ-TORTOSA et al. 2011]).

Die Referenzarchitektur wurde so umfassend und generisch gestaltet, dass es je nach Situation und Gegebenheiten möglich ist, die benötigten Komponenten zu kombinieren. Beispielsweise könnte es sinnvoll sein, über die ODM-Import-Schnittstelle eines bereits vorhandenen, jedoch nicht auf Archetypen basierenden SDMS im KIS gespeicherte Daten zu importieren. Die gemäß der Referenzarchitektur neu zu implementierende Komponente müsste in der Lage sein, entsprechende Daten per AQL aus dem KIS zu extrahieren (AQL-Import) und im ODM-Format für das vorhandene SDMS bereitzustellen (ODM-Export). Ein vollständiges, auf Archetypen basierendes SDMS müsste in diesem Fall nicht realisiert werden.

4.3.3 Vorgehen gemäß Referenzarchitektur

Um in einer Studie die im KIS vorhandenen Daten gemäß der hier vorgestellten Referenzarchitektur nutzen zu können, sind folgende Schritte notwendig:

Schritt 1

Anbindung der zu integrierenden KIS

Im Falle eines auf Archetypen basierenden KIS muss eine AQL-basierte Abfrageschnittstelle bereitgestellt werden. Im Falle eines nicht auf Archetypen basierenden KIS muss ein entsprechendes Integrationsmodul bereitgestellt werden. Hierbei ist es denkbar, das Integrationsmodul iterativ jeweils um die in der aktuellen Studie benötigten Archetypen zu erweitern (dabei kann gemäß dem generischen Vorgehensmodell – vgl. Kapitel 3.1 – verfahren werden). Dieser Schritt muss einmal pro KIS durchgeführt werden.

Schritt 2

Bereitstellen der Studiendefinition

Definition von Studienstruktur und zu erhebenden Daten mit Hilfe des ODM. Die einzelnen, zu erhebenden Datenelemente werden dabei durch Referenzen auf Archetypen mit dem ODM-Alias-Mechanismus eindeutig semantisch beschrieben. Der auf Archetypen basierende Entwurf von eCRF-Seiten kann in einem vollständig auf Archetypen basierenden SDMS mit Hilfe von *openEHR*-Templates erfolgen. Aus Templates und Studienstruktur kann dann die ODM-Datei automatisch generiert werden. Dieser Schritt muss einmal pro Studie durchgeführt werden. Dabei können einmal definierte eCRF/Merkmalgruppen (zum Beispiel Basisdatensätze) leicht in unterschiedlichen Studien wiederverwendet werden.

Schritt 3

Bereitstellen von entsprechenden AQL-Abfragen

Definition einer AQL-Abfrage zu jedem für die Studie definierten ODM-Datenelement, um die jeweils benötigte Information aus den einzelnen KIS extrahieren zu können. Die Zuordnung einer AQL-Abfrage zu einem ODM-Datenelement erfolgt über die jeweilige OID. Dieser Schritt muss einmal pro Studie durchgeführt werden. Wird ein ODM-Datenelement wiederverwendet, welches bereits für eine andere Studie definiert wurde, kann in der Regel auch die entsprechende AQL-Abfrage wiederverwendet werden. Allerdings muss geprüft werden, ob für eine Beobachtung in der aktuellen Studie dieselben Untersuchungsmethoden vorgesehen sind wie in der Studie, für welche die AQL-Abfrage ursprünglich erstellt wurde. Gegebenenfalls muss die Abfrage entsprechend angepasst werden.

Darüber hinaus müssen verschiedene technische und organisatorische Rahmenbedingungen definiert werden. Ausgehend von den Gegebenheiten in den einzelnen Studienzentralen muss beispielsweise geregelt werden, wann bzw. wie häufig Daten aus den KIS in das SDMS überführt werden: Nur manuell, automatisch in definierten Zeitintervallen oder automatisch und in Echtzeit, wenn ein CRF ausgefüllt wird. Dementsprechend muss weiter festgelegt werden, wann, wie und wie oft die im SDMS erfassten Daten mit den Quelldaten in den KIS abgeglichen werden und wie im Falle von Inkonsistenzen verfahren wird. Diese Rahmenbedingungen müssen in der Regel einmal pro Studienzentrum definiert werden. Sobald diese Schritte abgearbeitet sind, können mit dem SDMS Studiendaten erfasst und geeignete vorhandene Daten aus den angeschlossenen KIS übernommen werden.

4.4 Archetypen zur Dokumentation von und in klinischen Studien

Die zentralen Komponenten der im vorhergehenden Kapitel vorgeschlagenen Referenzarchitektur basieren auf Archetypen. Um diese Architektur implementieren und einsetzen zu können, müssen entsprechende Archetypen zur Dokumentation im Kontext klinischer Studien verfügbar sein. Dies setzt voraus, dass sich Archetypen nicht nur zur Repräsentation von Daten aus der Versorgung sondern auch für Studiendaten und Studien-Metadaten eignen.

Eine systematische Suche mit den Suchbegriffen ‚clinical trial‘, ‚clinical study‘, ‚trial‘, und ‚study‘ ergab, dass bisher keine speziellen Archetypen zur Dokumentation im Kontext klinischer Studien existieren.

4.4.1 Entwicklung von Archetypen für Studiendaten

Eine Aussage, ob sich Archetypen generell auch zur Dokumentation von Studiendaten eignen ist nicht möglich, da die in klinischen Studien erhobenen Merkmalsarten meist von Studie zu Studie variieren. Bestimmte Merkmalsarten werden, zumindest in ähnlicher Weise, regelmäßig in Studien erhoben. Die CDASH Spezifikation (siehe Kapitel 2.7) standardisiert genau solche Merkmalsarten.

Wenn Archetypen identifiziert werden, welche sich zur Repräsentation der einzelnen CDASH Felder eignen, lässt sich daraus schließen, an welchen Stellen in *openEHR*-basierten Patientenakten entsprechende Daten gefunden werden können. Diese Daten sind damit potenziell geeignet für eine multiple Verwendung in korrespondierenden, CDASH-basierten eCRF-Datenfeldern. Gemäß der vorgeschlagenen Referenzarchitektur können diese Daten dann durch AQL-Abfragen aus den KIS extrahiert werden.

Ausgehend von den CDASH Domänen

- Common Identifier Variables (CI)
- Common Timing Variables (CT)
- Adverse Events (AE) und
- Prior and Concomitant Medications (CM)

wurde daher untersucht, ob und gegebenenfalls welche verfügbaren *openEHR*-Archetypen sich eignen, um die dort definierten Merkmalsarten zu dokumentieren. Für CDASH-Datenelemente, zu welchen keine passenden Archetypen identifiziert werden konnten, wurden bestehende Archetypen erweitert bzw. neue Archetypen entworfen. Nachfolgend werden die hierbei erzielten Ergebnisse und gewonnenen Erkenntnisse beispielhaft anhand der CDASH Domäne ‚Prior and Concomitant Medications‘ dargestellt. Die Ergebnisse für die anderen Domänen finden sich in Anhang 1.

In einem ersten Schritt wurden alle verfügbaren *openEHR*-Archetypen ermittelt, die beim Dokumentieren von Medikationsangaben relevant sein können. Das Ergebnis dieser Recherche ist in Abb. 4-13 dargestellt.

- openEHR-EHR-**OBSERVATION**.
 - substance_use.v1 (SU)
- openEHR-EHR-**EVALUATION**.
 - check_list.v1
 - Spezialisierung von check_list.v1:*
 - check_list-medication.v1 (CL)
 - excluded-adverse.v1 (EA)
 - exclusion-medication.v1 (EM)
- openEHR-EHR-**INSTRUCTION**.
 - medication.v1 (IM)
- openEHR-EHR-**ACTION**.
 - medication.v1 (AM)
- openEHR-EHR-**COMPOSITION**.
 - medication_list.v1 (ML)
 - prescription.v1 (PR)
- openEHR-EHR-**ITEM_TREE**.
 - medication.v1 (TM)
 - medication-formulation.v1 (MF)
 - medication-vaccine.v1 (MV)

Abb. 4-13: Archetypen mit Relevanz bei der Dokumentation von Medikationsangaben (die in Klammern angegebenen Abkürzungen korrespondieren mit den Angaben in Tab. 4 2).

Tab. 4-2 stellt, basierend auf den vorhergehend aufgelisteten Archetypen, die identifizierten Abbildungsmöglichkeiten dar. Der für CDASH-Elemente teilweise angegebene Datentyp ‚Text (kodiert)‘ zeigt an, dass es sich um ein Textfeld handelt, für welches nur vordefinierte Codes gemäß der CDISC Controlled Terminology zugelassen sind. Weitere Details finden sich in Anhang 1.

Die CDASH-Domäne CM besteht aus 20 Datenelementen. Für 12 dieser Elemente konnten direkte Abbildungen (*CDASH CM-Element ↔ auf Archetypen basierendes Attribut/Merkmal*) identifiziert werden. Um auch die übrigen 8 CDASH CM-Elemente basierend auf Archetypen dokumentieren zu können, wurden vorhandene Archetypen spezialisiert und um fehlende Felder erweitert – so beispielsweise für das Element CMDOSTOT zur Dokumentation der täglichen Gesamtdosis eines Medikaments. Zusätzlich wurde für die Domäne CM ein eigener CDASH-bezogener COMPOSITION-Archetyp neu entworfen. In 3 der 8 Fälle ohne direkte Abbildung (CMSPID, CMAENO und CMMHNO) waren keine geeigneten Archetypen vorhanden, da diese Elemente für studienbezogene Referenzen und Strukturinformationen vorgesehen und daher in klassischen Patientenakten nicht vorhanden sind. Ebenso war für das CDASH-Feld CMINGRD nur im Archetyp openEHR-EHR-OBSERVATION.substance_use.v1 eine Entsprechung vorhanden, nicht jedoch im Archetyp openEHR-EHR-ITEM_TREE.medication.v1. Das Element CMDOSTOT (Tagesdosis einer Medikation) fand sich nicht als eigenständiges Feld in den vorhandenen Archetypen. Es ist jedoch möglich, diese Information unter Nutzung mehrerer Felder (‚Dose‘ und ‚Dose duration‘) des Archetyps openEHR-EHR-ITEM_TREE.medication.v1 abzubilden.

Auch die übrigen 3 Elemente CMOCCUR, CMPRIOR und CMONGO wurden von keinem bereits vorhandenen Archetypen beschrieben, da sich diese aus anderen Einträgen einer Patientenakte implizit ergeben. So kann etwa das Element CMPRIOR, welches anzeigt, ob eine Medikation bereits vor Studienbeginn begonnen wurde, abgeleitet werden aus dem Beginn einer Medikation und dem Datum des Einschlusses des jeweiligen Studienteilnehmers in die Studie.

Tab. 4-2: Mögliche Abbildungen von CDASH CM-Datenelementen auf openEHR-Archetypen (aufbauend auf [KRUEL 2010] S. 57). Die Bezeichnungen der Archetypen entsprechen den in Abb. 4-13 eingeführten Kürzeln. Typ der Abbildung: ‚direkt‘ – es gibt einen Archetypen, welcher eine dem CDASH-Element entsprechende Merkmalsart definiert; ‚n. a.‘ bedeutet nicht anwendbar – die Information, die für das jeweilige CDASH-Feld benötigt wird, findet sich in der Regel nicht in einer Patientenakte – es gibt keinen entsprechenden Archetypen; ‚AQL‘ – es gibt keinen Archetypen für das jeweilige CDASH-Feld, die benötigte Information könnte aber mit einer AQL-Abfrage aus der Patientenakte abgeleitet werden.

CDASH CM Datenelement			openEHR-Archetyp		
Nr.	Name	Datentyp	Archetyp	Attribute	Typ der Abbildung
1	CMYN	Text (kodiert)	EM	global statements	direkt
2	CMSPID		-	-	n. a.
3	CMTRT	Text	TM	name of medication	direkt
4	CMOCCUR	Text (kodiert)	AM	-	AQL
5	CMINGRD	Text (Liste)	MF	name of formulation	direkt
6	CMINDC	Text (Liste)	TM	indication	direkt
7	CMAENO		-	-	n. a.
8	CMMHNO		-	-	n. a.
9	CMDSTXT	Text	TM	dose	direkt
10	CMDOSTOT	Text	TM	-	AQL
11	CMDOSU	Text (kodiert)	TM	dose unit	direkt
12	CMDOSFRM	Text (kodiert)	TM	Form	direkt
13	CMDOSFRQ	Text (kodiert)	TM	maximum dose unit frequency	direkt
14	CMROUTE	Text (kodiert)	TM	route	direkt
15	CMSTDAT	Datum	TM	date (time) of first administration	direkt
16	CMPRIOR	Text (kodiert)	AM, TM	-	AQL
17	CMSTTIM	Uhrzeit	TM	date (time) of first administration	direkt
18	CMENDAT	Datum	TM	date (time) of last administration	direkt
19	CMONGO	Text (kodiert)	AM, TM	-	AQL
20	CMENTIM	Uhrzeit	TM	date (time) of last administration	direkt

AQL-Abfragen

Zum Entwurf von CRF-Seiten zur Dokumentation der CDASH CM-Datenelemente CMOCCUR, CMPRIOR und CMONGO in einem auf Archetypen basierenden SDMS können die eigens hierfür entworfenen bzw. erweiterten Archetypen genutzt werden. Diese Archetypen sind CDASH-spezifisch und werden üblicherweise nicht zur Dokumentation in Patientenakten verwendet. Daher können diese Archetypen nicht genutzt werden, um in

openEHR-basierten Patientenakten nach Einträgen zu suchen, die sich für eine multiple Verwendung im Studienkontext eignen. Da sich die genannten CDASH-Elemente jedoch implizit aus Archetypen ergeben, die in *openEHR*-basierten Patientenakten genutzt werden, können auch die Informationen für diese CDASH-Elemente mit AQL aus Patientenakten extrahiert werden (Vermerk ‚AQL‘ in Tab. 4-2). Da in der Regel mehrere Archetypen berücksichtigt werden müssen, sind diese AQL-Abfragen jedoch meist komplexer als direkte Abbildungen. Zu allen Datenelementen der untersuchten CDASH-Domänen, die sich nur implizit aus *openEHR*-basierten Patientenakten ableiten lassen, wurden nicht nur CDASH-spezifische Archetypen entwickelt, sondern auch entsprechende AQL-Abfragen definiert (siehe Anhang 1). Das folgende Beispiel (Abb. 4-14) zeigt die Abfrage für das Feld CMPRIOR:

Bei den mit ‚\$‘ beginnenden Bezeichnern handelt es sich um Variablen, die vor der Abfrage mit Werten belegt werden müssen. Im Fall von ‚\$ehrId_of_current_patient‘ ist dies die systeminterne Identifikation des aktuellen Patienten; ‚\$medication_of_interest‘ muss die Bezeichnung der aktuell betrachteten Medikation zugewiesen werden. Als Ergebnis liefert die Abfrage das Datum der ersten Gabe der betrachteten Medikation – dieses kann dann in Beziehung zum Datum des Studieneinschlusses des betrachteten Patienten gesetzt werden.

```

SELECT itemTreeMedication/data/items[at0018]/items[0019]/value AS
    Date_Time_of_First_Administration
FROM EHR e [ehr_id/value=$ehrId_of_current_patient]
    CONTAINS (COMPOSITION composition
        CONTAINS (ACTION
            actionMedication [openEHR-EHR-ACTION.medication.v1]
                CONTAINS (ITEM_TREE
                    itemTreeMedication [openEHR-EHR-ITEM_TREE.medication
                        (-[a-zA-Z0-9_+]*.v1))))
WHERE
    itemTreeMedication/items[at0001]/value=
        $medication_of_interest

```

Abb. 4-14: AQL-Abfrage zur Ermittlung der ersten Gabe einer bestimmten Medikation.

Die CDASH Domäne CM konnte unter Hinzunahme der vorgeschlagenen Spezialisierungen und Ergänzungen vollständig mit vorhandenen *openEHR*-Archetypen abgebildet werden. Bei der Abbildung von Daten aus Patientenakten, die auf Archetypen basieren, auf CDASH-konforme eCRF müssen in Einzelfällen jedoch spezifische Hürden gemeistert werden. Das folgende Beispiel soll dies verdeutlichen:

Beispiel CMDSTXT

Für das Feld **CMDSTXT** (Einzeldosis einer Medikation) ist die Abbildung auf das Feld ‚Dosis‘ im Archetypen *openEHR-EHR-ITEM_TREE.medication.v1* nicht eindeutig: Je nachdem, ob eine Dosisangabe dokumentiert wird als

- eine Masseneinheit samt Maßzahl oder Intervall, als
- eine Volumeneinheit samt Maßzahl oder Intervall oder als
- Anzahl, Bruchteil oder Intervall einer Dosiseneinheit (beispielsweise ‚6 Beutel‘)

wird die entsprechende Angabe in einem anderen Attribut des Archetyps abgelegt. Um in einer *openEHR*-basierten Akte einen Wert zu finden, welcher für das Feld **CMDSTXT** in

einen eCRF übernommen werden könnte, müssen immer alle relevanten Attribute berücksichtigt werden. Wurde mit dem Archetyp ein Dosisintervall dokumentiert, müsste darüber hinaus eine individuelle Abbildungsvorschrift (zum Beispiel Berechnung des Mittelwerts) festgelegt werden, um aus dem Intervall den Absolutwert einer Einzeldosis für das CDASH-Feld berechnen zu können. Umgekehrt müsste bei der Repräsentation von CDASH-Feldern mit Archetypen berücksichtigt werden, dass das CDASH-Feld CMDSTXT eine reine Zahlenangabe verlangt (für die Dosisseinheit ist ein eigenes Feld CMDOSU vorgesehen), der Archetyp hier jedoch nur ein kombiniertes Attribut aus Maßzahl und Maßeinheit vorsieht.

Zur Erleichterung der Abbildung von CDASH-Elementen mit *openEHR*-Archetypen wäre es teilweise sinnvoll, die bestehenden Archetypen für Patientenakten noch weiter zu ergänzen bzw. neue Archetypen zu entwickeln. Nachfolgend wird beispielhaft die Abbildung für das Element CMINGRD betrachtet.

In Bezug auf das Element **CMINGRD** (Aufzählung aller Wirkstoffe eines Arzneimittels) wäre es hilfreich, die beiden Archetypen

- *openEHR-EHR-ITEM_TREE.medication.v1* und
- *openEHR-EHR-ITEM_TREE.medication-vaccine.v1*

jeweils um eine direkt mit diesem CDASH-Element korrespondierende optionale Merkmalsart ‚Inhaltsstoffe‘ (analog zum Archetyp *openEHR-EHR-ITEM_TREE.medication-formulation.v1*) zu erweitern. So könnten auch diese Archetypen einfach zur Dokumentation des CDASH-Feldes genutzt werden. In einer Patientenakte, die nicht auf Archetypen basiert, ist eine solche Erweiterung nicht unbedingt notwendig, da hier die Wirkstoffe eines Fertigarzneimittels bei Bedarf in entsprechenden Datenbanken recherchiert werden können. In diesem Zusammenhang sollte außerdem beachtet werden, dass die Bezeichnungen ‚Wirkstoff‘ und ‚Inhaltsstoff‘ oft synonym gebraucht werden, obwohl nicht jeder Inhaltsstoff auch notwendiger Weise ein Wirkstoff sein muss. Sowohl der CDASH-Standard als auch die Definition dieses Archetyp-Feldes sollten daher semantisch präziser gefasst werden (nach [KRUEL 2010] S. 64).

Basierend auf den vorhergehend identifizierten Archetypen wurde ein *openEHR*-Template zur Dokumentation des CDASH CM-Datensatzes entworfen. *OpenEHR*-Templates können in einem auf *openEHR*-Archetypen basierenden SDMS zur Definition von CRF-Seiten genutzt werden. Abb. 4-15 zeigt einen Ausschnitt der Definition dieses Templates im Ocean Template Designer – Abb. 4-16 eine hieraus automatisch generierte HTML-basierte Eingabemaske.

**Beispiel
CMINGRD**

**CDASH CM
Template**

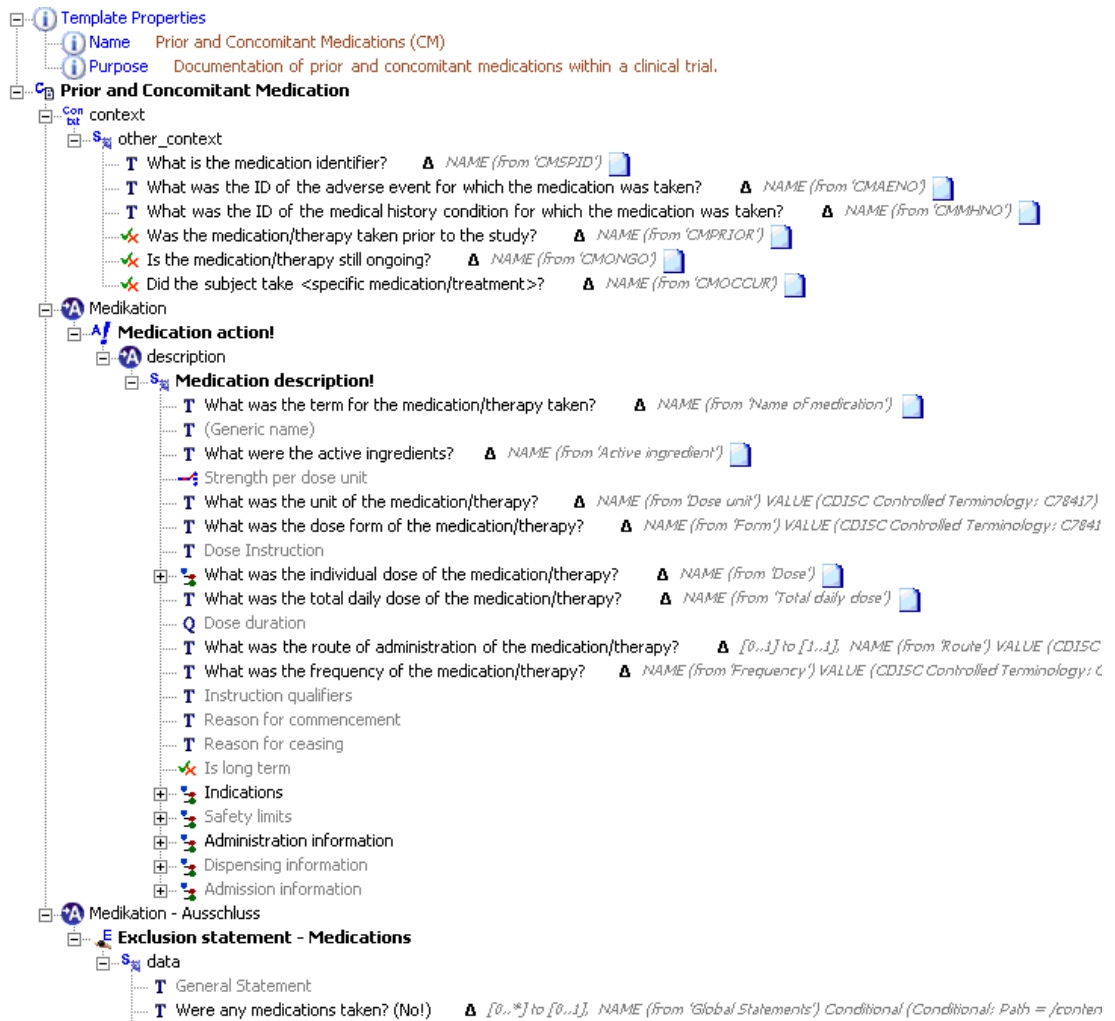


Abb. 4-15: Definition des CDASH CM-Templates im Ocean Template Designer.

Abb. 4-15 veranschaulicht, wie der *openEHR*-Template-Mechanismus genutzt wurde, um basierend auf verfügbaren und adaptierten bzw. neue generierten Archetypen CRF gemäß der CDASH-Spezifikation zu generieren:

- Von den Archetypen definierte, optionale Felder, die gemäß CDASH-Vorgabe nicht benötigt werden, wurden ausgeblendet (graue Schrift).
- Für die Darstellung am Bildschirm wurden die in den Archetypen hinterlegten Feldnamen durch die von CDASH vorgegebenen Fragen ersetzt (zum Beispiel „What was the individual dose of the medication/therapy? ▲ NAME (from ‚Dose‘) □“)
- Falls in einem verwendeten Archetypen ein Feld als optional, in CDASH jedoch als verpflichtend gekennzeichnet ist, wurde dieses auch im entsprechenden Template als verpflichtend deklariert („[0..1] to [1..1]“).
- Für Elemente, für welche CDASH bestimmte CDISC Code-Listen als mögliche Ausprägungen vorsieht, wurden diese Code-Listen mit Hilfe des Templates referenziert („VALUE (CDISC Controlled Terminology: C12345)“).
- Darüber hinaus (in Abb. 4-15 nur durch die „□“-Symbole ersichtlich) wurden alle im Template definierten Felder mit dem entsprechenden, durch CDASH vorgegebenen Feldcode annotiert (zum Beispiel ‚CMPRIOR‘).
- Wo sinnvoll wurden Bedingungen definiert, um Inkonsistenzen bei der Eingabe zu vermeiden: Die CDASH Felder CMENDAT und CMENTIM beispielsweise geben Datum bzw. Uhrzeit des Endes der jeweiligen Medikation an (für beide Felder sieht der

Archetyp `openEHR-EHR-ITEM_TREE.medications.v1` ein kombiniertes Datums- und Zeitfeld vor). Falls eine Medikation auch nach Studienende noch fortbesteht, darf kein Endzeitpunkt erfasst werden. Stattdessen muss im Feld `CMONGO` (ongoing medication) der Wert `„true“` gesetzt werden. Daher wurde im Template für das Eingabefeld zur Erfassung des Enddatums einer Medikation die Bedingung hinterlegt, dass dort nur dann ein Wert eingetragen werden darf, wenn `„CMONGO“` den Wert `„false“` enthält und umgekehrt.

PRIOR AND CONCOMITANT MEDICATIONS (CM)

Purpose
 ① Documentation of prior and concomitant medications within a clinical trial.

Archetype ID	openEHR-EHR-COMPOSITION.cdash_cm.v1
Template ID	2f42d583-4644-464a-a353-ffab16f5be69
MetaDataSet: Sample Set	Template metadata sample set

C **PRIOR AND CONCOMITANT MEDICATION**

Collapse All Show Annotations

[+/-] **other_context [1]**

T What is the medication identifier?

T What was the ID of the adverse event for which the medication was taken?

T What was the ID of the medical history condition for which the medication was taken?

Was the medication/therapy taken prior to the study?

Is the medication/therapy still ongoing?

Did the subject take <specific medication/treatment>?

[+/-] **Medication action!**

[+/-] **Medication description! [1]**

T What was the term for the medication/therapy taken? [1] The name of the intervention - which may be coded

T What were the active ingredients?

T What was the unit of the medication/therapy?

T What was the dose form of the medication/therapy?

[+/-] What was the individual dose of the medication/therapy?

Abb. 4-16: HTML-Eingabemaske generiert aus dem openEHR CDASH CM-Template.

Zur Definition des CDASH CM-Templates wurden einerseits die zu Anfang dieses Abschnitts aufgezählten, allgemein verfügbaren und für Patientenakten entworfenen Archetypen genutzt.

Darüber hinaus wurden der entwickelte ‚CDASH-Archetyp‘ openEHR-EHR-COMPOSITION.cdash_cm.v1 sowie die CDASH-Spezialisierungen der Archetypen

- openEHR-EHR-ACTION.medication.v1 und
- openEHR-EHR-ITEM_TREE.medication.v1

verwendet.

Vergleichbare Ergebnisse wurden auch für die anderen Domänen (Common Identifier Variables, Common Timing Variables und Adverse Events) erarbeitet. Insgesamt wurden somit vier Templates entworfen, für welche insgesamt zwei COMPOSITION-Archetypen neu entwickelt und drei bestehenden Archetypen (je ein ACTION-, ein EVALUATION- und ein ITEM_TREE-Archetyp) spezialisiert werden mussten. Im Rahmen der neu entwickelten und spezialisierten Archetypen mussten insgesamt 23 Merkmalsarten basierend auf Archetypen neu definiert werden.

4.4.2 Entwicklung von Archetypen für Studien-Metadaten

Zur Umsetzung der Referenzarchitektur werden ferner Archetypen benötigt, die Studien-Metadaten repräsentieren. Aus der Synopse von Merkmalsarten, welche von der International Clinical Trials Registry Platform (apps.who.int/trialsearch), vom Deutschen Register Klinischer Studien (register.germanctr.de/drks) sowie KKS-intern (www.kks-hd.de) erhoben werden, wurden 82 Merkmalsarten identifiziert, die Studien und zugehörige Objekttypen umfassend beschreiben. So ist eine Studie beispielsweise gekennzeichnet durch ein Studienakronym, ihren primären Endpunkt oder ihre Ein- und Ausschlusskriterien.

Als Grundlage für die Abgrenzung der einzelnen Archetypen wurde folgende generische Studienstruktur (Abb. 4-17) erarbeitet:

- Eine Studie hat mindestens einen Studienarm.
- Ein Studienarm besteht aus mindestens einer Visite.
- Im Rahmen einer Visite werden studienrelevante Daten erhoben und/oder Interventionen (Medikamentengaben, Operationen etc.) durchgeführt.

Während eine Beobachtungsstudie in der Regel nur einen Studienarm hat, verfügt eine randomisierte, kontrollierte Studie über mindestens zwei Studienarme.

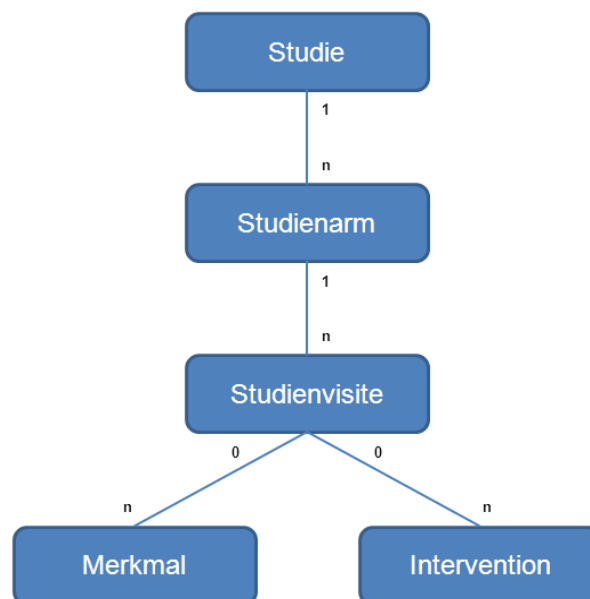


Abb. 4-17: Generische Struktur einer klinischen Studie.

Ausgehend von den identifizierenden Objekttypen und den zugehörigen Merkmalsarten wurden *openEHR*-Archetypen für Studien-Metadaten modelliert. Insgesamt wurden die 16 in Abb. 4-18 dargestellten *openEHR*-Archetypen entworfen und in englischer und deutscher Sprache beschrieben. Diese Archetypen ermöglichen es, Studien-Metadaten in *openEHR*-basierten Systemen analog zu Patientendaten zu speichern und zu verarbeiten. Die Zuordnung der einzelnen DRKS-, ICTRP- und KKS-Studien-Metamerkmale zu den entworfenen *openEHR*-Archetypen kann Anhang 2 entnommen werden.

- openEHR-EHR-**COMPOSITION**.
 - clinical_trial.v1 (Klinische Studie)
 - trial_arm.v1 (Studienarm)
 - trial_visit.v1 (Studiensite)
 - study_site.v1 (Studienzentrum)
- openEHR-EHR-**ADMIN_ENTRY**.
 - study_contact_data.v1 (Studienbezogene Kontaktdaten)
 - study_content.v1 (Studieninhalt)
 - study_design.v1 (Studiendesign)
 - study_ethics_committee_vote.v1 (Ethikvotum)
 - study_identification.v1 (Studienidentifikation)
 - study_length.v1 (Studienumfang)
 - study_site_description.v1 (Studienzentrum -Beschreibung-)
 - study_sponsor.v1 (Sponsor)
 - study_state.v1 (Studienstatus)
 - study_type.v1 (Studientyp)
 - trial_arm_description.v1 (Studienarm -Beschreibung-)
 - trial_visit_description.v1 (Studiensite -Beschreibung-)

Abb. 4-18: Neu erstellte Archetypen für Studien-Metadaten.

Die *openEHR*-Architektur bietet die Möglichkeit COMPOSITIONs in *openEHR*-Folders (siehe Kapitel 2.6) zu organisieren. Diese Verzeichnisse können beim Entwurf eines auf Archetypen basierenden SDMS genutzt werden, um alle zu einer bestimmten Studie gehörenden Metadaten und die Studiendaten zusammenzuführen und gemäß der dargestellten generischen Studienstruktur zu ordnen. Abb. 4-19 stellt dies schematisch dar.

Folders

Für jede Studie wird basierend auf der generischen Studienstruktur eine eigene Verzeichnishierarchie angelegt. Dabei enthält das Wurzelverzeichnis dieser Hierarchie genau ein COMPOSITION-Objekt ‚clinical_trial‘, welches die jeweilige Studie identifiziert und beschreibt. Weiter enthält das Wurzelverzeichnis jeweils ein COMPOSITION-Objekt ‚study_site‘ zur Beschreibung der Studienzentren sowie ein Unterverzeichnis pro Studienarm.

Studie

In den die Studienarme repräsentierenden Verzeichnissen befindet sich analog zum ‚Studienverzeichnis‘ ein COMPOSITION-Objekt ‚trial_arm‘, welches den jeweiligen Studienarm identifiziert und beschreibt. Weiter enthalten die Verzeichnisse der einzelnen Studienarme ein Unterverzeichnis für jede geplante Studiensite.

Arm

Auch in den die Studiensites repräsentierenden Verzeichnissen befindet sich analog zum Studienverzeichnis ein COMPOSITION-Objekt ‚trial_visit‘, welches die jeweilige Site identifiziert und beschreibt. Darüber hinaus enthalten die Visiten-Verzeichnisse COMPOSITION-Objekte, welche die medizinischen Daten aller Patienten repräsentieren, die bei der jeweiligen Site erhoben wurden. So führt ein Visitenverzeichnis, die Studiendaten aller Studienteilnehmer zusammen, die zu einer bestimmten Site (in einem bestimmten Studienarm in einer bestimmten Studie) erhoben wurden, inklusive der Metadaten, welche die Site selbst beschreiben.

Visite

Ergänzt um die Beschreibung der Studienstruktur und die einzelnen Studien-Metadaten ergeben die medizinischen Inhalte aller Visiten aller Arme der Studie den vollständigen Datensatz der jeweiligen Studie (vgl. Abb. 4-19).

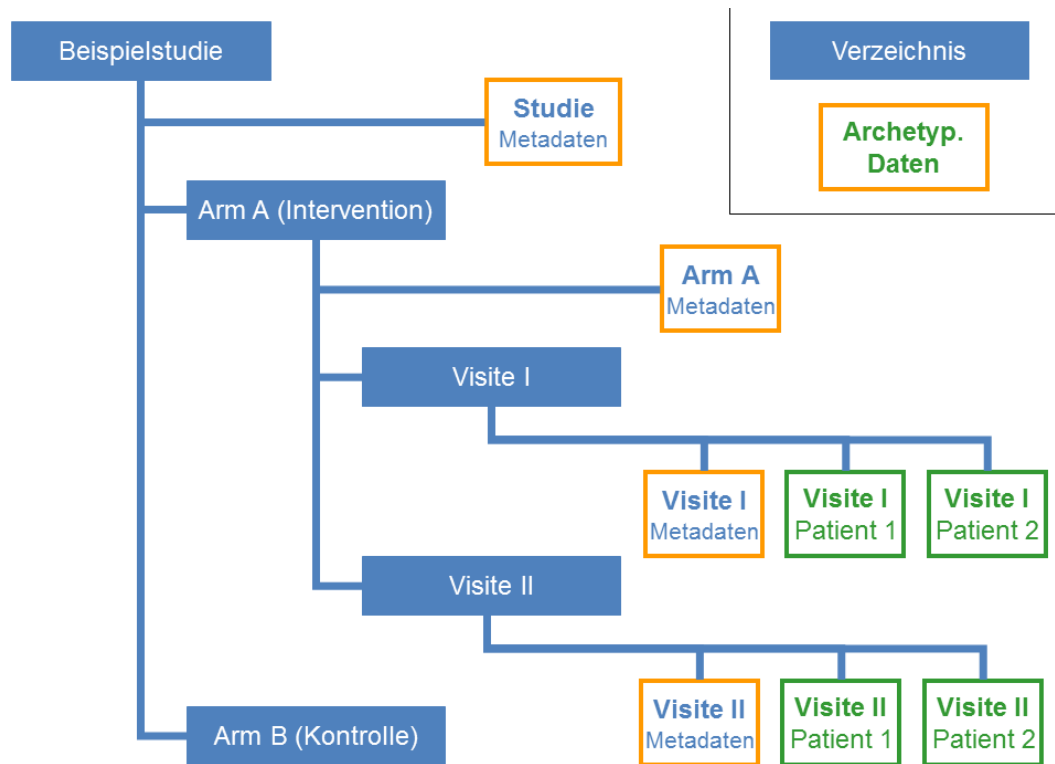


Abb. 4-19: Verzeichnisstruktur für Studiendaten und Studienmetadaten (blau: Verzeichnisstruktur; orange/blau: Studien-Metadaten; grün: eigentliche Studiendaten).

Für die Objekttypen ‚Studie‘, ‚Studienarm‘ und ‚Visite‘ wurde dabei jeweils ein COMPOSITION-Archetyp und ein oder mehrere zugehörige ADMIN_ENTRY-Archetypen definiert. Die COMPOSITION-Archetypen enthalten als CONTEXT eine Merkmalsart ‚ID‘, welche es erlaubt, das jeweilige Objekt eindeutig zu referenzieren. Darüber hinaus sind SECTIONS vorhanden, um die zugehörigen ADMIN_ENTRY-Archetypen einzubinden. Die eigentlichen Meta-Merkmale sind – von den jeweiligen IDs abgesehen – in den ADMIN_ENTRY-Archetypen enthalten. Abb. 4-20 veranschaulicht dies schematisch. Um die Übersichtlichkeit zu verbessern, werden nicht alle modellierten Merkmalsarten abgebildet. Der Archetyp openEHR-EHR-ADMIN_ENTRY.study_identification.v1 definiert Felder für Studienname, Studienakronym und Versionsnummer des Studienprotokolls.

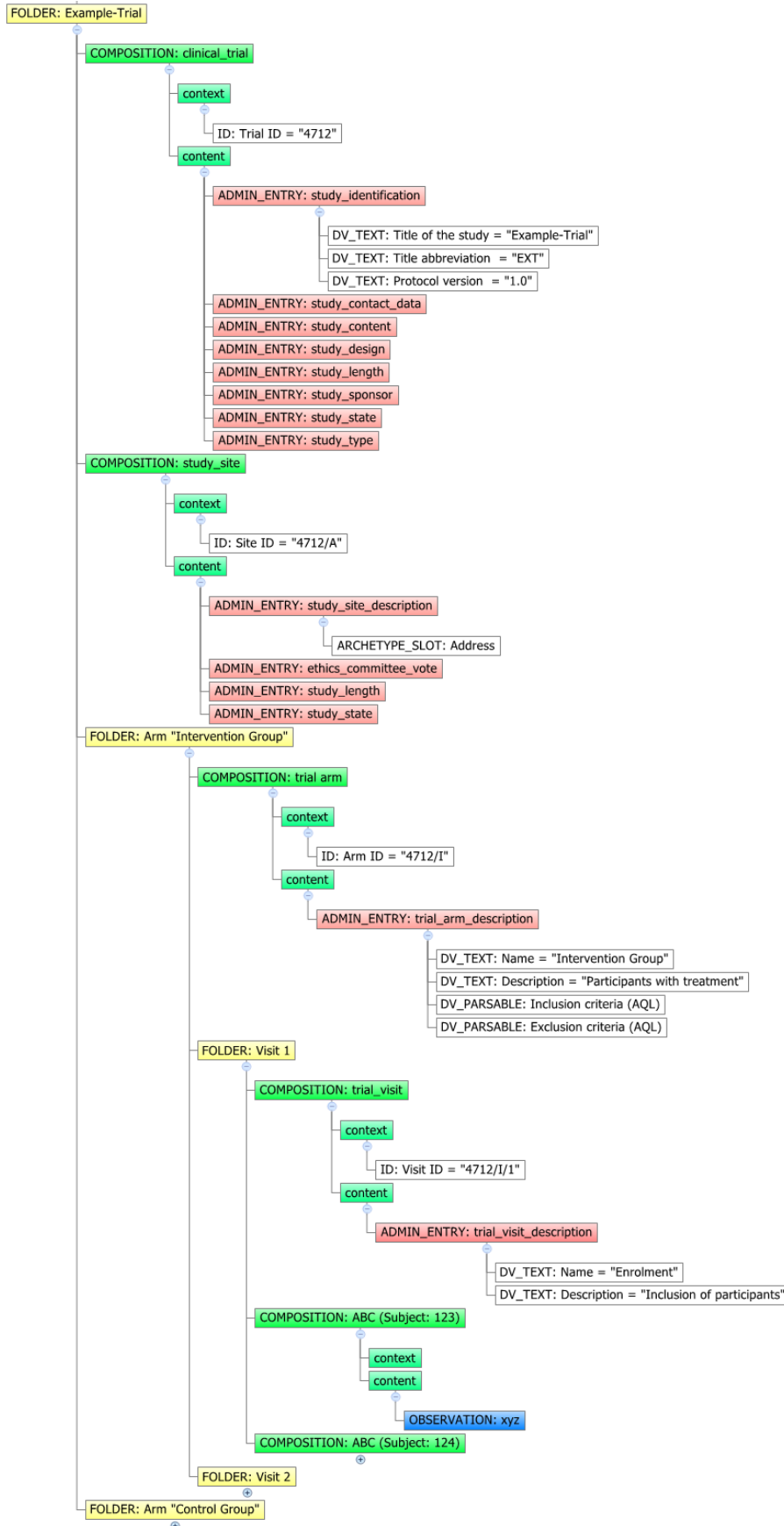


Abb. 4-20: Abbildung von Studien-Metadaten mittels COMPOSITION und ADMIN_ENTRY-Archetypen in die generische Verzeichnisstruktur.

Ein- und Ausschlusskriterien

Zur strukturierten Erfassung von Ein- und Ausschlusskriterien wurde die Archetype Query Language gewählt. Der Archetyp `openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1` enthält je ein Textfeld zur Definition von Ein- und die Schlusskriterien, welche in AQL formulierten Ausdrücke aufnehmen können. Bei Bedarf können so für jeden Studienarm individuelle Ein- und Ausschlusskriterien exakt und maschinenlesbar definiert werden. Die so definierten Kriterien können direkt genutzt werden, um in einem auf den *openEHR*-Spezifikationen basierenden KIS die Gruppen der Patienten zu ermitteln, welche alle Ein- bzw. mindestens ein Ausschlusskriterium erfüllen.

Folgende AQL-Abfrage (Abb. 4-21 - adaptiert nach einem Beispiel aus [MA, FRANKEL et al. 2007]) kann genutzt werden, um zu ermitteln, ob ein Patient die fiktiven Einschlusskriterien

- erkrankt an ‚Diabetes mellitus‘ und
- HbA1c größer als 7,0 in den vergangenen 12 Monaten

einer Studie erfüllt. Die im Beispiel dargestellte Abfrage liefert die Anzahl (COUNT...) der Patientenakten, welche die geforderten Kriterien erfüllen. Da die Abfrage immer auf die Akte eines bestimmten Patienten (`ehr_id/value=$ehrId_of_current_patient`) eingeschränkt wird, ist das Ergebnis entweder 0 (=Einschlusskriterien werden nicht erfüllt) oder 1 (=Einschlusskriterien werden erfüllt):

```

SELECT COUNT (e/ehr_id)
FROM EHR e [ehr_id/value=$ehrId_of_current_patient]
    CONTAINS (COMPOSITION
        cp1 [openEHR-EHR-COMPOSITION.problem_list.v1]
            CONTAINS EVALUATION
                epd [openEHR-EHR-EVALUATION.problemdiagnosis.v1])
    AND (COMPOSITION
        cr [openEHR-EHR-COMPOSITION.report.v1]
            CONTAINS OBSERVATION
                o1 [openEHR-EHR-OBSERVATIO.laboratoryhba1c.v1])
WHERE
    epd/data/items[at0002.1]/value/value='diabetes mellitus'
AND
    cr/context/other_context/items[at0006]/items[at0013]/value >
        current-date() - P1Y
AND
    o1/data/events[at0002]/data/items[at0013.1]/value/numerator > 7

```

Abb. 4-21: AQL-Abfrage zur Ermittlung, ob ein Patient die Einschlusskriterien einer Studie erfüllt (adaptiert nach einem Beispiel aus [MA, FRANKEL et al. 2007]).

Die AQL kann auch unabhängig von Systemen, die auf den *openEHR*-Spezifikationen basieren, genutzt werden, um Ein- und Ausschlusskriterien für klinische Studien einheitlich und systematisch zu definieren. Zum Speichern der entsprechenden Ausdrücke können einfache Textdateien genutzt werden. Zusätzlich müssten die referenzierten Archetypen vorgehalten werden, um die Semantik der formulierten AQL-Statements nachvollziehen zu können.

4.5 Abbildung auf bestehende Standards

Im Umfeld klinischer Studien haben sich die CDISC Standards etabliert. Exemplarisch und als wichtige Vertreter seien ODM (siehe Kapitel 2.7 und [KASPAR 2008], S. 19f), SDTM und CDASH sowie das BRIDG-Modell (siehe Kapitel 2.7) genannt. Während das ODM Studien-Metadaten, administrative Daten und Studiendaten in einer hierarchischen, XML-basierten Struktur abbildet, nutzt das SDTM einen Tabellen-basierten Ansatz. ODM standardisiert jedoch nicht die Semantik der repräsentierten Daten. Hier setzt CDASH an: Datenelemente, die in bestimmten Bereichen klinischer Studien typischer Weise benötigt werden, werden von CDASH inhaltlich beschrieben und samt Ausfüllhinweisen für CDAHS-konforme CRF definiert. Um ODM-Elemente mit studienübergreifender Semantik zu versehen, kann in ODM beispielsweise auf CDASH-Definition verwiesen werden. Das BRIDG-Modell ist auf einer weiteren, übergeordneten Ebene einzuordnen: Es beschreibt standardübergreifend typische Konzepte im Kontext klinischer Studien und deren Beziehungen zueinander.

Zu jedem Datenelement in einer ODM-Datei existiert eine Definition, die das jeweilige Element näher beschreibt. Diese Elementdefinitionen enthalten zunächst eher ‚technische‘ Informationen wie Datentyp oder mögliche Ausprägungen der jeweiligen Elemente und unterstützen damit die syntaktische Interoperabilität. Über den Alias-Mechanismus kann jede Elementdefinition in einer ODM-Datei zusätzlich mit einem CDASH-Datenelement aber auch mit von Archetypen definierten Merkmalsarten verknüpft werden. Durch diese Verknüpfungen wird eine semantische Definition an die jeweiligen Elemente gebunden. Wurde eine ODM-basierte Studiendefinition mit Hilfe von Archetypen semantisch annotiert, bleibt dieser Bezug auch bestehen, wenn die erhobenen Studiendaten in eine ODM-Datei exportieren werden. Diese Eigenschaft wird im Rahmen der in Kapitel 4.3 vorgestellten Referenzarchitektur genutzt.

ODM

[...]

```
<ItemDef      OID="ID.ANYCONMED"
              Name="Any Concomitant Medications"
              DataType="text" Length="1">
  [...]
  <Alias context="CDASH"
        Name="CMYN"/>
  <Alias context="Archetype"
        Name="openEHR-EHR-COMPOSITION.cdash_cm.v1/
              context/other_context/items[at0010]/
              value/value"/>
  <CodeListRef CodeListOID="CL.YESNO"/>
</ItemDef>
```

[...]

Abb. 4-22: Beispiel für die multiple Nutzung des Alias-Mechanismus zur semantischen Annotation eines ODM-Elements.

Ähnlich wie CDASH-Annotationen, ermöglichen Archetypen somit semantische Interoperabilität beim Austausch dieser Datenelemente. Im Gegensatz zu CDASH beschränken sich Archetypen jedoch nicht auf bestimmte Themenbereiche, sondern können prinzipiell beliebige medizinische Daten repräsentieren und sind bei Bedarf flexibel erweiter- und anpassbar. Wenn möglich sollten ODM-Elemente mit CDASH-Referenzen *und* mit den entsprechenden Referenzen auf Archetypen annotiert werden. So können diese ODM-Datenelemente sowohl von Systemen semantisch interpretiert werden, die kompatibel

zu CDASH sind, als auch von solchen, die auf Archetypen basieren. Die ODM Spezifikation lässt multiple Aliasse zu, wie Beispiel (Abb. 4-22) veranschaulicht.

Bestimmte Meta-Informationen zu einem ODM-Datenelement, wie beispielsweise Angaben zum Datentyp eines Elements, können sowohl in einem verknüpften Archetyp als auch mit Hilfe des ODM selbst definiert werden. Ebenso können Code-Listen für die möglichen Ausprägungen eines Elements oder die Kennzeichnung von Feldern als Pflichtfelder sowohl in Archetypen als auch im ODM definiert werden. Um in ODM-Dateien, die mit Archetypen annotiert wurden, Inkonsistenzen durch sich widersprechende Definitionen zu vermeiden, sollte zur Beschreibung eines ODM-Elements wenn möglich immer nur ein Mechanismus genutzt werden. Verglichen mit den ODM-Mechanismen bieten Archetypen dabei tendenziell mehr Möglichkeiten zur Definition und Beschreibung eines Datenelements.

Falls es die Kompatibilität zu unterschiedlichen Systemen dennoch erfordert, dass Elementdefinitionen redundant mittels ODM-Mechanismen und basierend auf Archetypen erfolgen, muss sichergestellt werden, dass sich diese nicht widersprechen. Dabei müssen die Definitionen, welche vom datenerhebenden (=Daten produzierenden) System interpretiert werden können, immer die am engsten gefassten sein. Dies entspricht der Anforderung innerhalb der *openEHR*-Spezifikationen, dass Templates oder Spezialisierungen von Archetypen Merkmaldefinitionen einschränken, aber nicht weiter fassen dürfen. Das folgende Beispiel soll dies verdeutlichen.

Angenommen ein System zur Datenerhebung kann lediglich ODM-Dateien interpretieren und in einer entsprechenden Datei wird ein Datenelement vom Typ ‚String‘ mit einer Code-Liste von Termen als möglichen Ausprägungen definiert. Ferner sei dieses Datenelement mit einer auf Archetypen basierenden Annotation versehen, wobei der referenzierte Archetyp lediglich ein Datenelement vom Typ ‚Text‘ ohne vorgegebene Liste von Ausprägungen vorsieht. In diesem Fall kann ein nachgeordnetes, ebenfalls auf Archetypen basierendes System alle gemäß der ODM-Definition dieses Elements erhobenen Daten korrekt verarbeiten. Wenn der Archetyp jedoch ebenfalls die zulässigen Ausprägungen auf eine eigene Liste beschränkt, welche sich von der in der ODM-Datei definierten Liste unterscheidet, ist keine unmittelbare semantische Interoperabilität mehr gegeben.

CDASH

Die Abbildung der zur Repräsentation von **Studiendaten** entwickelten und identifizierten Archetypen auf die jeweiligen CDASH Domänen ist gegeben, da diese Archetypen ausgehend von CDASH zusammengestellt bzw. speziell entwickelt wurden (siehe Kapitel 4.4.1). Die einzelnen Abbildungen *CDASH-Element ↔ auf Archetypen basierende Merkmalsart* können Anhang 1 entnommen werden. Bei der Abbildung der CDASH Datenelemente auf *openEHR*-Archetypen wurden zwei Punkte deutlich, die bei einer technischen Umsetzung der definierten Abbildungen berücksichtigt werden müssen: Während die *openEHR*-Spezifikation Uhrzeit und Datumsangaben gemäß ISO 8601 vorsehen (also eine Datumsangabe der Form JJJJ-MM-TT, zum Beispiel 2006-09-19) ist in CDASH das Datumsformat DD-MMM-YYYY (zum Beispiel 19-SEP-2006) vorgesehen. Dieser Unterschied lässt sich technisch durch entsprechende Transformationen einfach überwinden. Weitere Probleme können auftreten wenn CDASH und die entsprechenden *openEHR*-Archetypen unterschiedliche Code-Listen für ein korrespondierendes Datenelement vorsehen. Hierbei muss im Einzelfall und mit entsprechendem Fachwissen geprüft werden, ob und wie die einzelnen Elemente der Code-Listen aufeinander abgebildet werden können. Während CDASH beispielsweise für den Schweregrad eines unerwünschten Ereignisses (Datenelement ‚AESEV‘) die Ausprägungen ‚mild‘, ‚moderate‘ und ‚severe‘ vorsieht, definiert der korrespondierende Archetyp *openEHR-EHR-EVALUATION.adverse.v1* die Ausprägungen ‚mild‘, ‚disabling‘ und ‚life threatening‘.

BRIDG

Ausgehend von den in Kapitel 4.4.2 vorgestellten Archetypen für **Studien-Metadaten** wurde für jede der dort definierten Merkmalsarten geprüft, wo sich diese im BRIDG-Modell (siehe Kapitel 2.7) wiederfindet. Die einzelnen Abbildungen wurden in einer Tabelle erfasst

und um teilweise notwendige Bedingungen zur Steuerung dieser Abbildungen ergänzt - siehe Anhang 3. Bis auf wenige Ausnahmen, wie etwa spezielle Angaben zum gesetzlichen Rahmen einer Studie oder der ergänzenden Beschreibung des Studiendesigns, konnten alle in den Archetypen für Studien-Metadaten definierten Merkmalsarten auf das BRIDG-Modell abgebildet werden.

Tab. 4-3 stellt am Beispiel der Merkmalsart ‚primärer Endpunkt‘ einer Studie diese Abbildung dar. Zusätzlich zur Abbildung der *openEHR*-Archetypen auf das BRIDG-Modell werden auch die Abbildungen auf die Bezeichnungen und Definitionen von DRKS und WHO dargestellt. Die Vollständige Liste aller 115 identifizierten Zuordnungen bzw. Gegenüberstellungen findet sich in Anhang 2 (Abbildung der DRKS und WHO/ICTRP Merkmalsarten zur Beschreibung einer Studie auf *openEHR*-Archetypen) und Anhang 3 (Abbildung von *openEHR*-Archetypen auf das BRIDG-Modell). In Anhang 2 sind zusätzlich auch die im KKS Heidelberg definierten Merkmalsarten berücksichtigt. Die gemeinsame Gegenüberstellung aller in Tab. 4 3 genannter Attribute für alle 115 Abbildungen steht in einer eigenständigen Tabelle in der elektronischen Fassung dieser Arbeit zur Verfügung.

Tab. 4-3: Gegenüberstellung und Zuordnung der Definition der Merkmalsart ‚primärer Endpunkt‘ einer Studie in den Spezifikationen von DRKS, WHO, *openEHR* und BRIDG-Modell.

Attribut	Eintrag
DRKS: Parameter (Name)	Primärer Endpunkt
DRKS: Beschreibung (Datentyp)	[Freitext]
DRKS: Weitere Beschreibung	-
WHO: Parameter (Name)	Primary Outcome(s)
WHO: Beschreibung	Outcomes are events, variables, or experiences that are measured because it is believed that they may be influenced by the intervention. [...]
<i>openEHR</i>: Archetyp	<i>openEHR</i> -EHR-ADMIN_ENTRY.study_content.v1
<i>openEHR</i>: Pfad (Archetyp-intern)	/data/items[at0026]
<i>openEHR</i>: Name der Merkmalsart	Primary outcome
<i>openEHR</i>: Datentyp	DV_TEXT
BRIDG: Klassenname	StudyOutcomeMeasure
BRIDG: Attribut (der Klasse)	Name
BRIDG: Datentyp	ST (=String)
BRIDG: Pfad der Abbildung	StudyProtocolVersion > StudyObjective > StudyOutcomeMeasure.name WHERE (Bedingung) StudyOutcomeMeasure.primaryIndicator = "TRUE" (BL = Boolean)
BRIDG: Kommentar/Anmerkung	-

Die in den vorhergehenden Schritten identifizierten und entwickelten Archetypen eignen sich zur Repräsentation sowohl von Studiendaten als auch von Studien-Metadaten. Dies wurde gezeigt, indem diese Archetypen auf etablierte Modelle aus der klinischen Forschung abgebildet bzw. mit diesen verknüpft wurden.

4.6 Zusammenfassung der Ergebnisse

Die durchgeführte Fallstudie hat gezeigt, dass bei der cTAG-Studie eine multiple Nutzung von KIS-Daten prinzipiell möglich und sinnvoll ist. Die hierzu notwendige Analyse der Merkmalsarten im KIS ist jedoch sehr zeitaufwändig. Ferner wurde deutlich, dass in der Studie die Daten stärker strukturiert erhoben werden als im KIS.

Mit openSDMS steht der Prototyp eines auf Archetypen basierenden integrierten EPA- und Studiendatenmanagementsystems zur Verfügung. OpenSDMS implementiert hierzu einen Single-Source-Ansatz. Dabei zeigten sich die prinzipielle Realisierbarkeit des auf Archetypen basierenden Ansatzes sowie dessen Vorteile – so beispielsweise die einfache Identifikation von Daten, die sich für eine multiple Nutzung eignen.

Aufbauend auf den Erkenntnissen aus Fallstudie und der Implementierung von openSDMS wurde eine Referenzarchitektur zur Anbindung verschiedener KIS an ein SDMS spezifiziert. Dabei wird sowohl die Integration von auf Archetypen basierenden als auch von nicht auf Archetypen basierenden, klassischen KIS beschrieben. Die Bausteine dieser Architektur können je nach konkretem Szenario für verschiedene Studien adaptiert und eingesetzt werden. Die Referenzarchitektur wurde explizit auch auf multizentrische Studien ausgelegt.

Um die entwickelte Referenzarchitektur realisieren zu können, sind geeignete Archetypen notwendig. Daher wurden exemplarisch Archetypen zur Dokumentation aller Datenelemente der vier CDASH Domänen ‚Common Identifier Variables‘, ‚Common Timing Variables‘, ‚Adverse Events‘ sowie ‚Prior and Concomitant Medications‘ spezifiziert (Studiendaten). Hierzu wurden insgesamt 23 Merkmalsarten basierend auf Archetypen neu definiert, wozu wiederum drei bestehenden Archetypen spezialisiert und zwei COMPOSITION-Archetypen neu entwickelt wurden. Zur Definition von CDASH-konformen eCRF für die betrachteten Domänen wurden basierend auf den spezifizierten Archetypen vier *openEHR*-Templates entworfen.

Ergänzend wurden 71 Merkmalsarten in 16 Archetypen zur Dokumentation von Studien-Metadaten definiert. Die neu entworfenen Archetypen wurden jeweils sowohl in englischer als auch in deutscher Sprache beschrieben.

In einem weiteren Schritt wurden alle von den bereitgestellten Archetypen definierten Merkmalsarten auf die im Bereich der klinischen Forschung etablierten Modelle BRIDG, CDASH und ODM abgebildet. Für das ODM wurde eine entsprechende Verknüpfung im Rahmen der Referenzarchitektur beschrieben (siehe Kapitel 4.3.2). Für CDASH und BRIDG wurden Abbildungstabellen erstellt (siehe Anhang 1 und 3).

5 Diskussion

5.1 Diskussion von Vorgehen und Methodik

Fallstudie

In Anlehnung an ([MAYRING 2007], S. 28ff und [AMMENWERTH, ILLER et al. 2003]) wurde eine Fallstudie durchgeführt. Im Gegensatz P. Mayrings qualitativem Ansatz wurden bei den in dieser Arbeit durchgeführten Untersuchungen primär quantitative Auswertungen durchgeführt, wie beispielsweise zur Zahl der multipel nutzbaren Merkmalsarten. Als konkreter Fall wurde die cTAG-Studie aus dem Bereich der Gefäßchirurgie gewählt, da in diesen Bereich – im Gegensatz etwa zu onkologischen oder psychiatrischen Studien – meist gut überschaubare Mengen an Merkmalen erhoben werden. Durch die Fallstudie war es möglich, Abläufe und assoziierte Schwierigkeiten detailliert zu untersuchen, um daraus allgemeine Empfehlungen und Anforderungen an Systeme zur Integration von Forschung und Versorgung abzuleiten.

Obwohl es sich bei der cTAG-Studie um eine multizentrische Studie handelte, wurde die Fallstudie ausschließlich am Heidelberger Studienzentrum durchgeführt. Diese Einschränkung beruht einerseits auf der Detailtiefe der durchgeführten Untersuchung. Andererseits beruht sie auf der in Heidelberg vorhandenen, notwendigen Unterstützung durch die Studiendurchführenden und der Zugangsmöglichkeit zum KIS. Das Heidelberger KIS war dabei insofern interessant, da es sich um ein über lange Jahre gewachsenes und aus zahlreichen papierbasierten wie auch elektronischen Subsystemen bestehendes System handelt.

Für die Fallstudie wurden alle Schritte des ‚generischen Vorgehensmodell für die multiple Nutzung von Daten in Forschung und Versorgung‘ (Kapitel 3.1) durchgeführt. Die einzelnen Teilschritte des Modells erwiesen sich als geeignet. Die systematische Analyse der elektronischen und konventionellen Subsysteme des KIS sowie der eCRF und der Studiendatenbank ermöglichte die Identifikation multipel verwendbare Datenelemente. Indem die Stichprobe der Patienten, deren Dokumentation analysiert wurde, solange erweitert wurde, bis die Menge der jeweils im KIS vorgefundenen Dokumente saturiert war, wurden dieselben Dokumententypen teilweise mehrfach überprüft. So konnten Fehler beim Erheben der einzelnen Merkmalsarten erkannt und korrigiert werden. Im Rahmen der Fallstudie nicht explizit erfasst wurden Informationen zum zeitlichen Kontext und zur Art der Erhebung der einzelnen KIS-Merkmale. Diese Informationen waren aufgrund der zu jeder Merkmalsart dokumentierten Quellen (Arztbrief, OP-Bericht, RIS,...) und der Kenntnis der zugehörigen Prozesse implizit bekannt.

Die im Rahmen von Schritt 3 des Vorgehensmodells identifizierten Zuordnungen von KIS- und Studiendaten wurden gemeinsam mit der Heidelberger Study Nurse validiert. Dabei wurde deutlich, dass nur ein sehr geringer Anteil an Zuordnungen falsch definiert worden war. So zeigte sich etwa, dass der Prüfarzt bestimmte Abstände selbst in den CT-Aufnahmen der Patienten nachmisst, obwohl die benötigte Angabe bereits im entsprechenden Radiologie-Befund enthalten wäre. Grund für diese ‚Doppelbefundung‘ ist die in der Studie benötigte Genauigkeit der entsprechenden Angabe. Alle fehlerhaften Zuordnungen konnten korrigiert werden.

Mit Blick auf den letzten Schritt des Vorgehensmodells wurde mit openSDMS ein auf *openEHR*-Archetypen basierendes System implementiert, welches es erlaubt, Daten aus der Versorgung für Studienzwecke wiederzuverwenden (siehe Kapitel 4.2). Um dieses für eine konkrete Studie nutzen zu können, müsste ein geeignetes Importmodul für das betroffene KIS implementiert werden. Darüber hinaus muss geprüft werden, ob für alle zu erhebenden Studiendaten bereits geeignete Archetypen existieren oder ob diese noch zu entwickeln sind.

Die Suche nach potenziell geeigneten Zuordnungen zwischen Merkmalsarten im KIS und der Studie konnte mit der zum Erfassen der Daten eingesetzten Tabellenkalkulation (Microsoft Excel 2003) nicht automatisiert werden. Um etwa Abkürzungen oder Synonyme (wie ‚RR‘ und ‚Blutdruck‘) korrekt zuordnen zu können, war Fachwissen notwendig. Gleichwohl erwies sich die Tabellenkalkulation als sehr hilfreich bei der systematischen Erfassung und Auswertung der identifizierten Merkmalsarten.

Generell müssen alle im KIS bzw. seinen Subsystemen erfassten Merkmalsarten vor einer möglichen multiplen Verwendung einmal semantisch analysiert und katalogisiert werden. In monozentrischen Studien muss nur ein KIS analysiert werden. In multizentrischen Studien muss die Analyse in der Regel in jedem Studienzentrum wiederholt werden. Im Rahmen dieser Arbeit wurde die Analyse am Zentrum Heidelberg durchgeführt. Als Nebenergebnis der Fallstudie wurde dabei für die Gefäßchirurgie der Katalog der im Heidelberger KIS dokumentierbaren Merkmalsarten erstellt. Dieser Katalog kann dort bei jeder neuen Studie genutzt werden, um Zuordnungen zu generieren. Wenn bei den in einem Bereich durchgeführten Studien immer wieder gleiche Merkmalsarten erhoben werden, können auch die Zuordnungen von KIS-Daten zu Studiendaten multipel verwendet werden.

Kosten-Nutzen-Analyse

Das Analysieren der KIS bzw. der Studiendefinition und definieren der Zuordnungen stellt einen nicht unerheblichen Aufwand dar. Demgegenüber steht der eingesparte Aufwand bei der Studiendurchführung (keine Doppelerfassung und Übertragungsfehler), der im Einzelfall abgewogen werden muss (vgl. hierzu [MÜLLER-MIELITZ, OHMANN et al. 2010]).

Prototyp

OpenSDMS (siehe Kapitel 4.2) ist ein Prototyp, der sowohl Funktionalitäten einer EPA als auch eines SDMS bereitstellt. Zusätzlich ermöglicht openSDMS die Nutzung von Daten aus der EPA als Studiendaten. Weitere Funktionen, wie beispielsweise der Import von Daten aus anderen Systemen und der Export von Studiendaten, wurden bisher nicht implementiert. Ebenso fehlen zahlreiche für SDMS typische Funktionen, etwa Funktionen zur Unterstützung des Monitoring wie beispielsweise ‚Source Data Verification Marks‘.

Dennoch wurden bei der Implementierung von openSDMS die Stärken des auf Archetypen basierenden Ansatzes für eine multiple Datennutzung ersichtlich. So konnten etwa die Pre-population- und Single-Source-Funktionalität von openSDMS relativ leicht implementiert werden. Um für ein CRF-Element geeignete Werte zur Vorbelegung zu ermitteln, muss lediglich in der dem Entity-Attribute-Value-Ansatz folgenden Datenhaltung von openSDMS gefiltert werden nach

- Patienten-ID
- Archetypen-Bezeichner und
- Pfad des Datenelements innerhalb des jeweiligen Archetypen.

Aus der Ergebnismenge kann dann beispielsweise noch der aktuelle Eintrag herausgefiltert werden.

Bei der Entwicklung von openSDMS wurde auf das bestehende Opereffa System aufgebaut. Da sich Opereffa selbst noch in der Entwicklung befand, musste für die Implementierung von openSDMS eine frühe Version von Opereffa genutzt werden. Diese hatte den Nachteil, dass in Opereffa selbst nicht alle geplanten Funktionen implementiert waren. Insgesamt setzte Opereffa die *openEHR*-Spezifikationen noch nicht vollständig um. Daten wurden zwar basierend auf Archetypen gespeichert und abgerufen, dabei wurden die einzelnen Datenelemente eines Archetypen aber direkt gespeichert, ohne das von den *openEHR*-Spezifikationen vorgesehene Konstrukt der COMPOSITIONs zu nutzen. In Opereffa können alle Datenelemente, die einer Patientenakte zu einem bestimmten Zeitpunkt gleichzeitig hinzugefügt wurden, über ihren Zeitstempel identifiziert werden. Weitergehende Informationen wie sie durch COMPOSITIONs bereitgestellt werden, wie beispielsweise eine Versionierung, waren jedoch noch nicht verfügbar. Ferner war aufgrund des frühen Entwicklungsstadiums von Opereffa nur wenig Dokumentation zu dem Framework

verfügbar und der Quellcode teilweise noch recht unübersichtlich. Dies erschwerte die Einarbeitung in Opereffa und die Implementierung von openSDMS. Hinzu kam die sinnvolle, aber spezielle Kombination von Java Server Faces und Dojo, um AJAX-Funktionalität zu implementieren und die Möglichkeiten der grafischen Benutzerschnittstelle zu erweitern. Insgesamt bot Opereffa dennoch eine hilfreiche Ausgangsbasis für die Entwicklung von openSDMS. Alle von Opereffa zur Verfügung gestellten Funktionen (im Wesentlichen die Funktionen einer elektronischen Patientenakte) wurden in openSDMS übernommen.

OpenEHR-Folders ermöglichen es, COMPOSITIONs logisch zu gruppieren. Da Opereffa noch keine vollständige Unterstützung für COMPOSITIONs bot, konnte das Konstrukt der *openEHR*-Folders in openSDMS nicht wie spezifiziert implementiert werden. Stattdessen wurden die einzelnen Datenelemente, dem generischen Studienmodell aus Kapitel 4.4.2 folgend, in virtuellen Verzeichnissen der Form ‚Studie/Arm/Visite‘ gruppiert. Die Verzeichnisstruktur wurde durch das Einfügen entsprechender Spalten in der EAV-Tabelle statisch im System implementiert. Dieses Vorgehen erwies sich als geeignet.

Die in openSDMS definierten virtuellen Verzeichnisse fassen nicht nur verschiedene Daten eines Patienten zusammen, sondern auch Daten von verschiedenen Patienten. So werden beispielsweise die Blutdruckwerte aller Patienten einer Studie, die bei der Einschlussvisite erhoben wurden, im selben Verzeichnis abgelegt. Es bleibt unklar, ob eine solche Nutzung von Verzeichnissen durch den *openEHR*-Spezifikationen intendiert wurde. Die *openEHR*-Spezifikationen stellen die Möglichkeit COMPOSITIONs mit Hilfe von Verzeichnissen logisch zu gruppieren insgesamt relativ knapp dar. Aus pragmatischer Sicht hat sich der gewählte Ansatz jedoch bewährt.

Die einfache Erweiterung der EAV-Tabelle von Opereffa um Daten in openSDMS einer Studie zuzuordnen und in logischen Verzeichnissen gruppieren zu können, bringt jedoch auch einen Nachteil mit sich: Jedes im System zu einem Patienten gespeicherte Datenelement kann maximal einer Studie zugeordnet werden. Soll ein einmal erfasstes Datenelement in mehreren Studien mehrfach multipel verwendet werden können, müsste eine zusätzliche Tabelle angelegt werden, um die M:N-Relation zwischen Patient / Datenelement und Studie / Arm / Visite abbilden zu können. Wenn sichergestellt ist, dass ein Patient immer nur einem Arm einer Studie zugeordnet sein kann, könnte die explizite Angabe des Studienarms entfallen. Der Studienarm ließe sich dann aus Patienten-ID und Studien-ID eindeutig ableiten.

Das *openEHR*-Referenzmodell samt Archetypen ist primär auf die Dokumentation von Daten und die Abbildung von Szenarien aus der Patientenversorgung ausgelegt. Die in Kapitel 4.4.2 vorgestellten Archetypen für Studien-Metadaten, ermöglichen es in openSDMS, für Studien-Metadaten dieselben Mechanismen zum Speichern, Anzeigen, Bearbeiten und Austauschen zu verwenden, die bereits für die EPA-Daten implementiert worden waren. Dabei werden Studien systemintern analog zu Patienten behandelt und dargestellt. Dieses Verfahren hat sich in openSDMS als praktikabel erwiesen.

Referenzarchitektur

Ein zentrales Element der vorgeschlagenen Referenzarchitektur ist der Import von Daten aus KIS, um diese zur Übernahme in CRF anbieten zu können. J. Aerts von XML4Pharma schlug in [AERTS 2009] vor, *openEHR*-Exporte aus KIS mit einem Abbildungsmechanismus wie XSLT in eine ODM-Datei mit CDASH-Annotation zu überführen. Diese ODM-Dateien könnten dann genutzt werden, um entsprechende CRF-Felder vorzubelegen und anschließend die ergänzten Werte aufzunehmen.

Die in dieser Arbeit vorgeschlagene Architektur geht weit über den Ansatz von J. Aerts hinaus. Während in [AERTS 2009] zur Auswahl und zum Export von KIS-Daten, die sich für die Übernahme in ein bestimmtes Feld auf einem CRF eignen, XSLT vorgeschlagen wird, ist hierfür gemäß der vorgestellten Referenzarchitektur (Kapitel 4.3) die AQL vorgesehen. Die

Formulierung der Abfragen von KIS-Daten mittels AQL erscheint einfacher und intuitiver als mittels XSLT. Ferner ist AQL speziell auf *openEHR*-Systeme angepasst. Darüber hinaus sieht die Referenzarchitektur eine vollständig auf Archetypen basierende Repräsentation von Studiendaten vor. Diese bietet beispielsweise den Vorteil, dass auch Datenelemente, die nicht von CDASH beschrieben werden, semantisch annotierbar sind.

Gleichwohl schließen sich die hier vorgestellte Referenzarchitektur und die in [AERTS 2009] oder von IHE (siehe Kapitel 2.7) vorgeschlagenen Ansätze und Technologien nicht aus. Diese Ansätze und Technologien können vielmehr genutzt werden, um die einzelnen Bausteine der vorgestellten Referenzarchitektur zu realisieren. So könnte XSLT etwa bei der Implementierung von Integrationsmodulen für nicht auf Archetypen basierende KIS genutzt werden.

Archetypen zur Dokumentation von und in klinischen Studien

Studiendaten

An vier exemplarischen CDASH Domänen wurden, soweit möglich, Archetypen zur Dokumentation der für diese Domänen definierten Datenelemente identifiziert. Es handelt sich dabei um Datenelemente, die häufig in klinischen Studien erhoben werden. Die vier Domänen Common Identifier Variables, Common Timing Variables, Adverse Events und Prior and Concomitant Medications wurden ausgewählt, da es sich um Bereiche handelt, die in den meisten klinischen Studien relevant sind.

Der aktuelle CDASH-Standard (Version 1.1) definiert insgesamt 235 Felder. Zum Teil befanden sich diese Felder noch in der Entwicklung bzw. Diskussion. Ebenso befanden sich die meisten der identifizierten, korrespondierenden Archetypen noch im Entwurfsstatus. In dieser Arbeit wurden 53 Felder in vier Domänen untersucht. Um zu prüfen, ob ähnliche Abbildungen auch bei den übrigen CDASH Domänen möglich sind, sollte die Untersuchung für diese Domänen weitergeführt werden.

Studien-Metadaten

Ausgangspunkt für die Entwicklung der Archetypen zur Repräsentation von Studien-Metadaten war ein generisches Modell klinischer Studien (vgl. Kapitel 4.4.2). Theoretisch wäre es auch möglich gewesen, Archetypen ausgehend vom BRIDG-Modell (vgl. Kapitel 2.7) zu definieren. Aufgrund seiner Zielsetzung, ein Domänen Analyse Modell zu sein, besitzt das BRIDG-Modell jedoch eine relativ hohe Komplexität und eine andere primäre Zielsetzung als Archetypen. Dies hätte die Entwicklung von Archetypen erschwert. Die Entwicklung der Archetypen basierend auf einem einfachen, aus der praktischen Erfahrung heraus entwickelten Modell hat sich als sinnvoll erwiesen. Dies entspricht der Intention des Archetypen-Paradigmas (vgl. Kapitel 2.5): Ein Archetyp definiert, was zu einem bestimmten Konzept in der täglichen Praxis dokumentiert werden kann.

Ein- und Ausschlusskriterien

Ein besonderes Augenmerk erforderte die auf Archetypen basierende Modellierung von Ein- und Ausschlusskriterien in klinischen Studien. Anstelle der in der Referenzarchitektur vorgeschlagenen Definition mit Hilfe der AQL, waren im Archetyp EHR-ADMIN_ENTRY.clinical_trial.v1 hierfür zunächst Archetyp-Slots vorgesehen worden. Mit Hilfe von auf Archetypen basierenden Datenmodellen und zugehörigen Wertemengen sollten Ein- und Ausschlusskriterien im Sinne eines ‚Query-by-Example‘-Ansatzes beschrieben werden. Sollte beispielsweise ein Studienteilnehmer eine bestimmte Erkrankung nicht haben, wäre der Archetyp zur Dokumentation der entsprechenden Erkrankung dem Slot für die Ausschlusskriterien hinzugefügt worden. Die nicht erwünschten Werte wären dann mit den durch diesen Archetyp definierten Attributen dokumentiert worden. Auf diese Weise hätten jedoch keine komplexen Ein-/Ausschlusskriterien mit Beziehung und Abhängigkeiten von Kriterien definiert werden können. Beispiel: ‚Ein Patient, der Krankheit A hat darf nur eingeschlossen werden, wenn er auch Krankheit B, aber nicht Krankheit C hat‘.

Abbildung auf Standards

Die im Rahmen der Referenzarchitektur vorgeschlagene Verknüpfung der auf Archetypen basierenden Repräsentation von Studiendaten und Studien-Metadaten mit dem ODM erwies sich als unproblematisch. Dies lag darin begründet, dass das ODM solche Verknüpfung durch den Alias-Mechanismus unmittelbar unterstützt. **ODM**

Die Abbildung der Archetypen für Studiendaten auf CDASH ergab sich automatisch, da diese Archetypen basierend auf CDASH entwickelt und zusammengestellt wurden. Das Vorgehen hierbei entsprach dem Vorgehen beim Entwurf eines auf Archetypen basierenden Systems auf Grundlage eines vorhandenen Dokumentationssystems (vgl. [BUCK, GARDE et al. 2009]). Dieses Vorgehen erwies sich als zielführend. **CDASH**

Die Abbildung der zur Repräsentation von Studien-Metadaten entworfenen Archetypen auf das schon sehr ausgereifte BRIDG-Model, ermöglichte es, die Plausibilität dieser Archetypen zu prüfen. Dabei zeigten sich noch vereinzelte fehlende Merkmalsarten, die zur Erweiterung der erstellten Archetypen führten. Diese Punkte waren besonders wertvoll, da Archetypen den Anspruch erheben, ein Konzept, wenn auch pragmatisch, so doch umfassend zu beschreiben. **BRIDG**

5.2 Diskussion der Ergebnisse

5.2.1 Fallstudie

Am Beispiel des Heidelberger KIS und der cTAG-Studie wurde geprüft, welche Anforderungen an Daten in Forschung und Versorgung gestellt werden und in welchem Umfang eine multiple Verwendung realisierbar ist. Gleichzeitig wurde geprüft, welche Herausforderungen hierbei bewältigt werden müssen. Dabei ergaben sich folgende Erkenntnisse:

In welchen Formaten sind die Daten im KIS dokumentiert?

Im Heidelberger KIS werden die dort dokumentierbaren Merkmalsarten zu jeweils rund 20% als Freitexte, als Zahlenwerte (Ganz- und Gleitkommazahlen), in Form von Auswahllisten sowie als ‚Ja/Nein‘-Felder erfasst. Datums- und Zeitangaben schlagen zusammen mit gut 10 % zu Buche (vgl. Abb. 4-2).

In welchen Formaten werden Daten für klinische Studien dokumentiert?

Studiendaten müssen in der Regel hoch strukturiert dokumentiert werden, da mit ihnen exakt definierte Auswertungsfragen beantwortet werden bzw. Analyseverfahren durchgeführt werden sollen. Im Fall der cTAG-Studie werden 93,5 % der Studiendaten strukturiert erhoben (vgl. Abb. 4-3). Ein Viertel der Studiendaten wird quantitativ (als Ganz- oder Gleitkommazahl) erhoben. Weitere knapp 10 % als Datums- oder Uhrzeitangaben. Dabei fällt auf, dass es sich bei mehr als einem Drittel der in der Studie zu erhebenden Daten um ‚Ja/Nein‘-Felder handelt. Dies könnte jedoch ein Spezifikum der betrachteten Studie sein.

Die Auswertungsfragen, zu deren Beantwortung Daten im KIS erhoben werden, sind in der Regel allgemeiner, als dies in klinischen Studien der Fall ist. Der Strukturierungsgrad der Daten im KIS ist daher häufig geringer, als der Strukturierungsgrad vergleichbarer Studiendaten. Dies deckt sich mit Erkenntnissen von W. DORDA¹, der feststellt, dass Daten in KIS oft unregelmäßig, uneinheitlich und nicht zielgerichtet erfasst werden.

Sind die Daten im KIS ausreichend und vollständig, um diese für Studien nutzen zu können?

Generell problematisch bei der multiplen Nutzung von KIS-Daten ist eine unregelmäßige und nicht einheitliche Erfassung dieser Daten. Während viele KIS-Daten im Rahmen der Routineversorgung oft mit eher allgemeinerer Zielsetzung beispielsweise zur Verlaufskontrolle oder als Gedächtnisstütze erhoben werden, werden Daten in klinischen Studien anhand exakt definierter Protokolle und für genau vorgegebene Fragestellungen erfasst.

Über 40 % der für die cTAG-Studie benötigten Datenelemente (82 Merkmalsarten) könnten basierend auf im KIS vorhandenen Daten und Systemparametern erhoben werden. Hierbei handelt es sich um einen theoretischen Wert. Lediglich vier (Alter, Geschlecht, Gewicht und ‚Operation als Notfall‘) von potenziell 70 im Heidelberger KIS vorhandenen Merkmalsarten sind unmittelbar multipel für die cTAG-Studie nutzbar. Die Einführung einer vollständig elektronischen Dokumentation wäre eine wesentliche Voraussetzung, um 29 im KIS bisher konventionell dokumentierte Merkmalsarten überhaupt multipel verwenden zu können. Bei 62 Merkmalsarten müsste bei einer Übernahme aus dem KIS in die Studie ferner eine geeignete Transformation zwischengeschaltet werden. In etwas mehr als der Hälfte dieser Fälle (bei 32 Merkmalsarten) wäre dabei die Extraktion von strukturierten Informationen aus

¹ Berichtet im Vortrag DORDA W (2010) ‚Nutzung der Elektronischen Krankenakte des AKH Wien für die Klinische Forschung‘ sowie der anschließenden Diskussion im Rahmen der 3. Sitzung der GMDS Projektgruppe ‚Nutzung von Elektronischen Krankenakten für die klinische Forschung‘ am 20. Mai 2010 in Berlin.

Freitexten (Textmining) notwendig. Dies verwundert nicht, werden doch im KIS mehr als dreimal so viele der dort dokumentierbaren Merkmalsarten (20,7 %) als Freitext erhoben werden, als in der cTAG-Studie (6,5 %). Lediglich zu einer elektronischen und einer konventionellen Merkmalsart, die im KIS vorgesehen und für die Studie nutzbar gewesen wären, waren keine Ausprägungen dokumentiert worden. Dieses spezielle Problem könnte durch eine entsprechende Arbeitsanweisung gelöst werden.

Eine weitere Möglichkeit KIS-Daten multipel zu verwenden und einen Mehrwert aus dem hierfür notwendigen Aufwand zu generieren, besteht darin, das KIS zu nutzen, um die Rekrutierung von Studienpatienten zu unterstützen (vgl. [DUGAS, LANGE et al. 2010]). So wären beispielsweise typische von DUGAS und Kollegen genannte Merkmalsarten wie ICD 10 codierte Aufnahme-diagnose, Alter und Geschlecht der Patienten im Heidelberger KIS vorhanden.

Welche Anforderungen ergeben sich an Daten, die in der Routineversorgung erhoben werden, damit sie auch in Studien genutzt werden können?

- Medizinische Daten sollten möglichst in maschinenlesbarer Form dokumentiert werden.
- Medizinische Daten sollten weitestgehend strukturiert im KIS erfasst werden.

Bevor medizinische Daten, die bisher als Freitext erfasst wurden, in stärker strukturierter Form in KIS erhoben werden, sollte sorgfältig geprüft werden, in wieweit sich eine stärkere Strukturierung mit den Gewohnheiten und Wünschen der jeweiligen Anwender (vornehmlich Ärzte und Pflegekräfte) vereinbaren lässt.

Im konkreten Fall des Universitätsklinikums Heidelberg könnte eine stärkere Strukturierung der Freitextelemente von OP-Bericht und Arztbriefen ein sinnvoller erster Schritt sein. Alleine 30 der 200 für die cTAG-Studie benötigten Merkmalsarten (15,0 %), könnten aus dem OP-Bericht extrahiert werden. Darüber hinaus werden bestimmte Merkmalsarten wie etwa ‚Diagnosen‘ im OP-Bericht sowie in allen Arztbriefen aufgeführt. Indem diese Informationen nur einmal strukturiert erhoben werden, könnten sowohl Aufwand gespart als auch Inkonsistenzen vermieden werden,

Welche Anforderungen ergeben sich an Daten, die im Rahmen von klinischen Studien erhoben werden, damit diese auch in der Routineversorgung genutzt werden können?

In der Fallstudie ‚KIS Heidelberg/cTAG-Studie‘ zeigte sich, dass der überwiegende Anteil, der für die cTAG-Studie erfassten Merkmalsarten (187 Merkmalsarten / 93,5 %) strukturiert erhoben wird. Dies verwundert nicht, da an Studiendaten in der Regel hohe Qualitätsansprüche und konkrete Auswertungsfragen geknüpft sind. Aus diesen Gründen werden die in Studien zu erhebenden Merkmalsarten und das Vorgehen beim Erfassen dieser Daten in der Regel exakt definiert. Daher erscheint es problemlos möglich, Studiendaten multipel in der Routineversorgung zu nutzen. Daten, die beispielsweise im Rahmen von studienspezifischen Follow-Up-Untersuchungen erhoben wurden, können helfen, ergänzend zu den Routineuntersuchungen den Genesungsprozess eines Patienten zu beurteilen.

Um die teilweise aufwändigen Transformationen und das Textmining bei der Übernahme von Daten aus dem KIS in eine Studie zu vermeiden, könnte man die Datenerhebung im KIS stärker strukturieren und den Erfordernissen der Studie anpassen. Liegen Daten bereits strukturiert in einem KIS vor Beginn der Studie vor, könnte man alternativ überlegen, die Daten für eine Studie so zu erheben, wie diese im KIS angeboten werden. Diese Überlegung erscheint jedoch vor dem Hintergrund von multizentrischen Studien wenig realistisch. Typischerweise verfügen die einzelnen Studienzentren über heterogene KIS, die sich in Struktur und Semantik der dort gespeicherten Datenelemente unterscheiden. Auch die Anpassung der Datenerhebung im KIS an eine bestimmte Studie kann zu Schwierigkeiten führen, wenn ein Datenelement im KIS in mehreren unterschiedlichen Studien benötigt wird.

Alternative Überlegungen

Alternativ könnten die in einer Studie erhobenen Daten multipel verwendet werden, um die jeweiligen KIS mit entsprechenden Informationen zu füllen (vgl. [KUSH, ALSCHULER et al. 2007]). Die Übernahme von Studiendaten in ein KIS führt jedoch nur zu einer Aufwandsersparnis, wenn die Studiendaten erfasst wurden, bevor die gleichen Daten im Rahmen einer Routinemaßnahme ermittelt werden. Darüber hinaus müsste ein KIS für Merkmalsarten, die in einer Studie erhoben werden, im KIS bisher zur Dokumentation aber noch nicht vorgesehen sind, erweitert werden.

Welche Anforderungen ergeben sich daraus an ein System zur Integration?

Unter einem System zur Integration kann einerseits ein System verstanden werden, welches den planerischen Prozess der Integration von Daten aus klinischen Studien und Routineversorgung erleichtert und beschleunigt. Anforderungen an solche Werkzeuge wurden in Kapitel 4.1 detailliert beschrieben.

Unter einem System zur Integration kann andererseits auch ein System verstanden werden, welches die Integration von Daten aus Forschung und Versorgung technisch realisiert. In Kapitel 4.3.1 werden die Anforderungen, die an diese Art von Systemen bzw. eine Referenzarchitektur für solche Systeme gestellt werden, aufgelistet. Eine zentrale Anforderung besteht darin, dass alle Schritte, die notwendig sind, um KIS-Daten in ein SDMS übernehmen zu können, so gestaltet werden müssen, dass die Verknüpfung KIS/SDMS für möglichst viele Studien nutzbar ist. Die durchgeführte Fallstudie hat gezeigt, dass die Analyse eines KIS zur Vorbereitung der multiplen Verwendung der dort erfassten Datenelemente einen größeren zeitlichen Aufwand erfordert. So mussten im Rahmen der Fallstudie insgesamt 1452 Merkmalsarten, die im Heidelberger KIS dokumentiert werden können, katalogisiert und beschrieben werden. Zum Beschreiben jeder Merkmalsart wurden bis zu zehn Metra-Merkmale erfasst. Dabei wurden die katalogisierten Merkmalsarten mehrfach, anhand der Dokumentation von vier verschiedenen Patienten, überprüft.

Übertragbarkeit der Ergebnisse

Die vorhergehend zusammengefassten Ergebnisse wurden im Rahmen einer Fallstudie erhoben. Weitere, vergleichbare Untersuchungen müssen durchgeführt werden – insbesondere auch in anderen medizinischen Fachgebieten, um die Übertragbarkeit der Ergebnisse für andere Studienfragestellungen und -protokolle zu bestätigen.

Gleichwohl zeigte die Untersuchung, dass die multiple Nutzung von Daten aus der Versorgung generell möglich ist und somit helfen kann, Aufwand in der klinischen Forschung einzusparen. Es zeigte sich, dass aktuelle Krankenhausinformationssysteme, welche dem Heidelberger KIS vergleichbar sind, noch nicht vollständig auf eine multiple Datennutzung vorbereitet und hierfür geeignet sind. In der Untersuchung waren alle relevanten Subsysteme des Heidelberger KIS als potenzielle Datenquellen berücksichtigt worden, auch konventionelle. Es liegt die Vermutung nahe, dass die Beobachtungen bezüglich des KIS auch auf andere medizinische Fachbereiche zutreffen. So etwa, dass Informationen im KIS öfter als unstrukturierter Text erfasst werden und daher beispielsweise erst nach einem entsprechenden Textmining multipel verwendet werden können.

STARBRITE

Die in der vorgestellten Fallstudie gewonnenen Erkenntnisse korrelieren mit Erkenntnissen der STARBRITE Proof-of-Concept Studie zur Implementierung eines Single-Source-Ansatzes [KUSH, ALSCHULER et al. 2007]. Auch dort waren die medizinischen Notizen im KIS weniger stark strukturiert als die entsprechenden CRF-Einträge. Die Fallstudie ‚KIS Heidelberg/cTAG-Studie‘ hat analog zur STARBRITE Proof-of-Concept Studie gezeigt, dass weder die für die Studie zu erhebenden Daten noch die im KIS dokumentierten Daten eine Obermenge der jeweils anderen Daten sind. Im Unterschied zur STARBRITE Proof-of-Concept Studie [KUSH, ALSCHULER et al. 2007], die eine Überlappung von etwa 75 % feststellt, konnten in der Fallstudie ‚KIS Heidelberg/cTAG-Studie‘ nur rund 40 % an Datenelementen identifiziert werden, die multipel verwendet werden könnten. Dies könnte sowohl an den unterschiedlichen Bereichen liegen, aus welchen die betrachteten Studien stammen (Kardiologie bzw. Gefäßchirurgie), als auch den unterschiedlich gestalteten KIS.

An den Universitätsklinika Erlangen und Münster wurde untersucht, in welchem Umfang Daten in KIS genutzt werden können, um die Patientenrekrutierung in klinischen Studien zu unterstützen [PROKOSCH, DUGAS et al. 2009]. Die Untersuchungen zeigten, dass 71 % bzw. 73 % der betrachteten Ein- und Ausschlusskriterien potenziell in den KIS verfügbar wären. Dabei wurden über 700 Merkmalsarten aus 63 Studienprotokollen berücksichtigt. Auch dies ist sehr viel mehr, als die in der Fallstudie ‚KIS Heidelberg/cTAG-Studie‘ identifizierten rund 40 %. Das kann an der horizontalen Ausrichtung der Untersuchung in Erlangen und Münster liegen, in welche die Ein- und Ausschlusskriterien vieler verschiedener Studien einbezogen wurden. Im Gegensatz dazu war die Fallstudie ‚KIS Heidelberg/cTAG-Studie‘ vertikal, das heißt *alle in einer Studie* erhobenen Merkmalsarten wurden betrachtet.

**Erlangen /
Münster**

5.2.2 Prototyp

OpenSDMS – siehe Kapitel 4.2– wurde entwickelt, um (technische) Anforderungen an eine auf Archetypen basierende Systemarchitektur zur Integration von Forschung und Versorgung zu identifizieren. Es wurde deutlich, dass speziell zur Dokumentation von Studien-Metadaten geeignete Archetypen benötigt werden. Gleichzeitig wurden bei der Entwicklung von openSDMS als integriertem klinischem Informations- und Studiendaten-managementsystems Vorteile des auf Archetypen basierenden Ansatzes deutlich. Beispielsweise erleichtert die auf denselben Archetypen basierende Repräsentation von Studiendaten und Daten in der EPA die Suche nach EPA-Daten, die multipel für eine Studie nutzbar sind. Darüber hinaus wurde mit der Implementierung von openSDMS die prinzipielle Realisierbarkeit eines auf Archetypen basierenden Studiendatenmanagement-systems gezeigt.

5.2.3 Referenzarchitektur

Die in Kapitel 4.3 beschriebene Referenzarchitektur für die Integration von Forschung und Versorgung ermöglicht es, Daten aus verschiedenen KIS zusammenzuführen, um diese multipel für klinische Studien zu nutzen. Dies vermeidet die unnötige erneute Eingabe von Studiendaten, die bereits in einem KIS im Rahmen der Routineversorgung erfasst wurden. Dabei stand das Ziel im Vordergrund, mögliche Vorteile und Einsatzmöglichkeiten von Archetypen bei der multiplen Nutzung von KIS-Daten zusammenzuführen, gleichzeitig aber auch bestehende Standards zu berücksichtigen. Im Unterschied zu bestehenden Ansätzen ist das hier vorgestellte Konzept speziell auch auf multizentrische Studien und die Integration verschiedener KIS ausgerichtet, unabhängig davon, ob diese auf Archetypen basieren oder nicht. Je nach Situation und Gegebenheiten ist es möglich, ausgehend von der entwickelten Referenzarchitektur, die jeweils sinnvollen Bausteine zu kombinieren. Anstatt ein vollständig auf Archetypen basierendes SDMS zu implementieren, könnten die in einem SDMS erfassten Studiendaten beispielsweise lediglich um Referenzen auf Archetypen ergänzt werden. Dies kann relativ einfach durch entsprechend annotierte ODM-Dateien zur Studiendefinition erfolgen.

Sobald ein SDMS Daten aus den elektronischen Subsystemen eines KIS nutzt, muss ein Mechanismus vorgesehen werden, der bei der Übernahme nicht nur die benötigten Datenelemente, sondern auch Informationen zum jeweiligen Quellsystem erfasst. Indem das Quellsystem eines Datenelements bis auf Subsystemebene eindeutig identifizierbar bleibt, werden sowohl regulatorische Anforderungen erfüllt als auch ein Quelldatenabgleich erleichtert. Ferner ist es nur so möglich, bei der Übernahme der KIS-Daten in ein SDMS Implausibilitäten und Fehler, wie sie beispielsweise durch Queries identifiziert werden, auch in den jeweiligen Quelldaten zu korrigieren. Dies kann unterstützt werden, indem das SDMS automatisch Berichte zu den notwendigen Datenkorrekturen in den Quellsystemen generiert. Um eine Kontrolle und gegebenenfalls Korrektur zu erleichtern, müssen diese Berichte zu jedem potenziell fehlerhaften Quelldatenelement den betroffenen Patienten, das

Anforderungen

entsprechende Subsysteme des KIS sowie eine eindeutige Identifikation des betroffenen Datenelements inklusive zeitlicher Einordnung nennen. Ferner sollte das SDMS in der Lage sein, die Überprüfung und Korrektur der identifizierten Fehler nachzuverfolgen.

Da die vorgeschlagene Architektur auf multizentrische Studien ausgerichtet wurde und keinen Single-Source-Ansatz realisiert, besteht generell das Problem, dass sich Daten in den KIS ändern können, nachdem diese ins SDMS importiert wurden. Aus diesem Grund sollte bei importierten Daten im SDMS das Importdatum gespeichert werden und in den KIS ein Audit-Trail geführt werden. So wäre ersichtlich, ob ein Datenelement im KIS erst nach dem Export ins SDMS verändert wurde. Die Inkonsistenz zwischen KIS und SDMS wäre damit noch nicht gelöst. Alternativ könnte in KIS und SDMS die Änderung von multipel genutzten Daten gesperrt werden. Diese Lösung erscheint jedoch wenig praxistauglich. Die optimale Lösung besteht darin, alle KIS-Daten, die für eine Studie übernommen wurden, regelmäßig vom SDMS im Sinne eines automatischen Quelldatenabgleichs auf Konsistenz mit dem jeweiligen KIS prüfen zu lassen. Auch hierfür müssen die jeweiligen Quellsysteme bekannt sein. Bei Abweichung könnten automatisch Queries generiert werden, welche der zuständige Prüfarzt dann auflösen muss.

**SDMS:
Vorteile durch
*openEHR***

Im Gegensatz zu klassischen SDMS nutzt das hier vorgeschlagene System *openEHR*-Archetypen zur Definition und Speicherung von Studiendaten und Studien-Metadaten. Die *openEHR*-Spezifikationen bieten den Vorteil, dass nahezu alle der in den vorhergehenden Abschnitten benannten Anforderungen in diesen Spezifikationen bereits berücksichtigt werden. So ist es beispielsweise vorgesehen, beim Import von Daten aus nicht *openEHR*-Systemen ein spezielles ‚feeder audit‘ mit Informationen über das Quellsystem und die importierten Daten zu generieren. Beim Import von Daten aus *openEHR*-basierten Systemen werden die in *openEHR* regulär vorgesehenen Audit-Mechanismen genutzt. Daher sollte ein auf den *openEHR*-Spezifikationen basierendes SDMS per se in der Lage sein, zusätzlich zu importierten Daten auch Informationen über Importzeitpunkt und Quellsystem zu erfassen.

Darüber hinaus ermöglicht ein solches SDMS nicht nur semantische Interoperabilität mit anderen auf Archetypen basierenden Systemen (insbesondere KIS). Wenn eCRF und damit die Datenerhebung in Studien basierend auf Archetypen erstellt werden, kann immer wieder auf dieselben Element-Definitionen (dieselben Archetypen) in verschiedenen Studien zurückgegriffen werden. Dies führt einerseits zu einer Vergleichbarkeit der erhobenen Merkmalsarten über Studiengrenzen hinweg, wodurch Metastudien erleichtert werden. Andererseits entstehen semantisch klar definierte Sammlungen von Merkmalsarten, welche als Basisdatensätze für Studien desselben medizinischen Bereiches wiederverwendet werden können. Ein auf Archetypen basierendes SDMS verfügt damit über ähnliche Funktionen und Vorteile wie sie auch Meta-Daten-Repositorys für die klinische Forschung bieten (vgl. beispielsweise [TMF 2009]). Im Gegensatz zur bloßen semantischen Annotation von Datenelementen mit Hilfe eines Meta-Data-Repositorys bieten auf Archetypen und den *openEHR*-Spezifikationen basierende Informationssysteme jedoch weitergehende Funktionalitäten und Möglichkeiten. So definiert ein Archetyp beispielsweise, wie ein Datenelement in einem auf Archetypen basierenden System gespeichert werden muss. Ferner ist mit der AQL eine Abfragesprache vorgesehen, welche die feingranulare Abfrage gespeicherter Daten ermöglicht. Den größten Vorteil bei der multiplen Nutzung von KIS-Daten in klinischen Studien bieten Archetypen, wenn die KIS ebenfalls auf Archetypen basieren.

**KIS:
Vorteile durch
*openEHR***

Alle in einem auf Archetypen basierenden KIS dokumentierbaren Merkmalsarten sind durch die zugrunde liegenden Archetypen eindeutig beschrieben. Daher ist es nicht notwendig, wie in Schritt 1 des generischen Vorgehensmodells für die multiple Nutzung von KIS-Daten (siehe Kapitel 3.1) vorgegeben, einen Katalog dieser Merkmalsarten zu erstellen. Ein auf Archetypen basierendes KIS ist selbstdokumentierend. Wenn in diesem KIS bisher noch nicht erfassbare Merkmalsarten dokumentiert werden sollen, muss lediglich ein neuer

Archetyp eingebunden werden. Mit der Definition des neuen Archetyps liegt automatisch auch die Beschreibung des repräsentierten Konzepts und der zugehörigen Merkmalsarten vor.

Aufgrund der durch die Archetypen vorgegebenen einheitlichen Semantik von Daten in SDMS und KIS entfällt darüber hinaus die manuelle Zuordnung von KIS- und Studiendaten (Schritt 3 des generischen Vorgehensmodells für die multiple Nutzung von KIS-Daten). Dementsprechend ist auch kein spezielles Integrationsmodul erforderlich. Dies würde auch in multizentrischen Studien funktionieren, an welchen Studienzentren mit ganz unterschiedlichen KIS teilnehmen, solange alle KIS dieselben Archetypen nutzen bzw. auf dasselbe Archetypen-Repository zugreifen. Falls unterschiedliche Archetypen-Repositoryn genutzt werden, wäre zunächst eine Abbildung zwischen den unterschiedlichen Archetypen erforderlich. Durch die semantische Beschreibung der repräsentierten Konzepte in den Archetypen sowie deren Merkmalsarten, würde dieser Schritt jedoch wesentlich erleichtert.

Sollen Daten aus einem KIS für eine klinische Studie in ein SDMS übernommen werden, bieten auf Archetypen basierende KIS noch einen weiteren Vorteil gegenüber klassischen Systemen: Damit Daten aus KIS multipel in klinischen Studien genutzt werden können, müssen diese Daten in der Regel ein bestimmtes Qualitätsniveau aufweisen, gemäß einer definierten Methode erhoben worden sein sowie in einem bestimmten zeitlichen Kontext zu der jeweiligen Studie stehen. Während klassische Systeme oft nur eine zeitliche Einordnung der erhobenen Daten erlauben, sind Archetypen darauf ausgelegt, auch weitere Kontextinformationen zu den verschiedenen Merkmalsarten eines klinischen Konzepts zu erfassen. So verfügen beispielsweise OBSERVATION-Archetypen über die zusätzlichen Attribute ‚Protocol‘ und ‚State‘. Unter ‚Protocol‘ können die Methoden und Instrumente beschrieben werden, die zum Erheben der Ausprägungen der beschreibenden Merkmalsarten genutzt wurden. Unter ‚State‘ kann der Zustand des jeweiligen Patienten während der Beobachtung dargestellt werden, um eine korrekte Interpretation der erhobenen Information zu ermöglichen. Beispielsweise definiert der OBSERVATION-Archetyp ‚Blutdruck‘ neben Feldern für den systolischen und den diastolischen Blutdruck auch Felder zur Dokumentation der angewandten Messmethode (Attribut ‚Protocol‘ – mögliche Ausprägungen: Auskultation, Palpation, Maschine, Invasiv) und zur Dokumentation der Position des Patienten während der Messung (Attribut ‚State‘ – mögliche Ausprägungen: Stehend, sitzend, zurückgelehnt, liegend, nach links geneigt liegend) – vgl. Kapitel 2.5 und Abb. 2-1.

Dieser Vorteil kommt nur dann zum Tragen, wenn die entsprechenden Protocol- und State-Informationen auch erfasst werden, wodurch ein zusätzlicher Aufwand bei der Dokumentation im KIS entsteht. Dem kann entgegengewirkt werden, indem der Nutzer nur dann ergänzende Informationen explizit angeben muss, wenn diese vom üblichen Wert abweichen. Als Position eines Patienten während einer Blutdruckmessung wird im Archetyp ‚Blutdruck‘ beispielsweise die Ausprägung ‚sitzend‘ angenommen, sofern keine andere Position angegeben ist. Ferner können zu Messwerten, welche beispielsweise auf Intensivstationen von Überwachungsmonitoren aufgezeichnet werden, Informationen über die eingesetzten Messgeräte automatisch aufgezeichnet werden.

Die vorgestellte Referenzarchitektur sieht auch die Integration von *nicht* auf Archetypen basierenden KIS vor. Hierzu muss für jedes dieser KIS ein dediziertes Integrationsmodul bereitgestellt werden. Die Definition hierfür benötigter Abbildungs- und Transformationsvorschriften stellt einen gewissen Aufwand dar. Dieser Arbeitsschritt muss nur einmal durchgeführt werden und kann dann für beliebig viele Studien bzw. Abfragen genutzt werden. Das Integrationsmodul ist ein Adapter, der es ermöglicht, auf ein nicht auf Archetypen basierendes KIS wie auf ein den *openEHR*-Spezifikationen folgendes System mittels AQL zuzugreifen.

AQL	<p>Mit Hilfe der AQL kann der Export von Daten aus den elektronischen Subsystemen eines KIS für Studien feingranular gesteuert werden. Dabei müssen entsprechende AQL-Abfragen auch in multizentrischen Studien nur einmal pro Datenelement spezifiziert werden – alle angeschlossenen KIS können diese dann interpretieren. Darüber hinaus eignet sich die AQL gut, um im Rahmen der Patientenrekrutierung nach geeigneten Studienteilnehmern in den angebundenen KIS zu suchen. Umgekehrt kann die AQL auch genutzt werden, um Ein- und Ausschlusskriterien in Studien strukturiert und maschinenlesbar zu definieren.</p>
Pseudonymisierung	<p>Die vorgeschlagene Referenzarchitektur sieht Pseudonymisierungsdienste vor, die auf Seiten der KIS angesiedelt sind. Dies bietet zweierlei Vorteile: Einerseits verbleiben die Pseudonymisierungslisten so auch technisch in den Studienzentren. Andererseits werden die Daten aus den elektronischen Subsystemen des KIS direkt beim Export pseudonymisiert, so dass ausschließlich pseudonymisierte Daten zwischen KIS und SDMS übertragen werden müssen. Dies verringert auf einfache Art das Risiko von unbefugter Kenntnisnahme bei der Datenübertragung – gleichwohl sollte diese durch einen geeigneten Verschlüsselungsmechanismus gesichert werden. Ein anderes, in der Regel aufwändigeres Pseudonymisierungssystem wird notwendig, wenn es möglich sein soll, dass ein Patient seine behandelnde Einrichtung wechselt und dennoch alle zu ihm erfassten Daten dasselbe Pseudonym erhalten sollen. Lösungsansätze für dieses Szenario bieten die generischen Datenschutzkonzepte der TMF (siehe [RENG, DEBOLD et al. 2006] und [POMMERENING, DREPPER et al. 2009]).</p>
Datenaustausch	<p>Für den physikalischen Datenaustausch zwischen KIS und SDMS sind verschiedene Realisierungen möglich: Automatisch via Netzwerkverbindung bzw. E-Mail oder auch manuell mittels USB-Stick. Hierzu ist anzumerken, dass durch einen manuellen Datentransfer, die KIS-Daten im zentralen SDMS erst mit entsprechendem, zeitlichen Versatz verfügbar sind. Dies kann eine multiple Nutzung in Studien erschweren. Ein automatisierter Datentransfer ist daher die bevorzugte Lösung.</p>
Archetypen Repository	<p>Um eine unmittelbare semantische Interoperabilität zu ermöglichen, sollten alle beteiligten Systeme dieselben Archetypen nutzen (vgl. Kapitel 2.5): Die einzelnen auf Archetypen basierenden KIS, die einzelnen Module zur Integration von nicht auf Archetypen basierenden Systeme sowie das eigentliche SDMS. Hierzu muss ein zentrales Archetypen-Repository bereitgestellt werden. Dieses kann auf lokaler, nationaler oder internationaler Ebene angesiedelt sein. Je größer der Einzugsbereich eines solchen Archetypen-Repositories ist, desto größer ist in der Regel auch der Kreis der Systeme, zwischen welchen semantische Interoperabilität gegeben ist. Gleichzeitig wird es jedoch auf organisatorischer Ebene umso schwieriger, ein Repository zu etablieren und zu verwalten (siehe hierzu auch [KOHL, GARDE et al. 2008]).</p> <p>Speziell wenn die auf Archetypen basierenden semantischen Annotationen von Studiendaten im Sinne eines Meta-Data-Repository genutzt werden, um studienübergreifende Auswertungen zu unterstützen, kann auch ein lokal gepflegtes Archetypen-Repository sinnvoll sein. Primär sollte aber immer ein möglichst breit anerkanntes und genutztes Archetypen-Repository angestrebt werden.</p> <p>Wie müsste vorgegangen werden, um entsprechend dieser Architektur Daten aus Forschung und Versorgung multipel nutzen zu können?</p> <p>Der in Kapitel 4.3.3 beschriebene Dreischritt</p> <ol style="list-style-type: none"> 1. Anbindung der zu integrierenden KIS 2. Bereitstellen der Studiendefinition 3. Bereitstellen der benötigten AQL-Abfragen <p>ist notwendig, um basierend auf einer konkreten Implementierung der vorgeschlagenen Referenzarchitektur eine Studie durchführen und dabei im KIS vorhandene Daten multipel nutzen zu können.</p>

Wurden ein oder mehrere KIS und SDMS gemäß der vorgestellten Referenzarchitektur integriert und eine Studie samt entsprechender AQL-Abfragen definiert, können in dieser Studie KIS-Daten multipel genutzt werden. Der Prüfarzt oder Study-Nurse kann beim Ausfüllen eines Feldes auf einem eCRF aus einer Liste von potenziell passenden, zum aktuellen Patienten in den KIS vorhandenen Datenelementen auswählen und in die Studie übernehmen. Dabei werden – soweit verfügbar – zu jedem Datenelement Kontextinformationen wie beispielsweise Quellsystem, Erhebungszeitpunkt und Erhebungsart angezeigt.

Darüber hinaus ist es denkbar, die Verbindung von KIS und SDMS zu nutzen, um auch umgekehrt ausschließlich für eine Studie erhobene Daten aus dem SDMS in ein angeschlossenes KIS zu überspielen. So stünden studienbezogene Untersuchungsergebnisse auch in der Routineversorgung zur Verfügung. Ausgehend von der vorgestellten Referenzarchitektur muss das SDMS hierfür um eine Exportschnittstelle erweitert werden, welche den selektiven Export der im SDMS manuell erfassten Daten basierend auf Archetypen erlaubt. Jedes KIS muss um eine entsprechende Importschnittstelle erweitert werden. Für alle nicht auf Archetypen basierende KIS müsste hierbei, analog zum Export von KIS-Daten, ein auf das jeweilige System angepasstes Import-Modul erstellt werden.

Multiple Nutzung von Studiendaten in KIS

5.2.4 Archetypen für Studiendaten

Bei einer ersten Suche nach speziellen *openEHR*-Archetypen für den Bereich klinischer Studien mit Hilfe des Clinical Knowledge Managers wurden keine Archetypen gefunden. Dies verwundert nicht: Archetypen sind darauf ausgerichtet, ein klinisches Konzept nicht nur umfassend und strukturiert, sondern auch im Kontext beliebiger Anwendungsfälle zu beschreiben (vgl. [BEALE und HEARD 2007], S. 10f). Daher liegt es nahe, dass alle Archetypen, die der Dokumentation von Daten in der medizinischen Versorgung dienen, prinzipiell auch zur Dokumentation von Studiendaten genutzt werden können. Diese Vermutung hat sich bei der Abbildung von CDASH Domänen auf *openEHR*-Archetypen (Kapitel 4.4.1) wie auch bei der Entwicklung von *openSDMS* (Kapitel 4.2) größtenteils bestätigt. Es hat sich aber auch gezeigt, dass es Sonderfälle gibt, in welchen zur Repräsentation von Studiendaten spezielle Studienarchetypen erforderlich sind.

Für die vier CDASH Domänen

- Common Identifier Variables (CI)
- Common Timing Variables (CT)
- Adverse Events (AE) und
- Prior and Concomitant Medications (CM)

wurde untersucht, welche verfügbaren *openEHR*-Archetypen sich eignen, um die dort definierten Datenelemente zu dokumentieren, welche regelmäßig in klinischen Studien erhoben werden (vgl. Kapitel 4.4.1). Waren semantisch naheliegende Archetypen vorhanden, welche die CDASH Vorgaben jedoch nicht vollständig abdecken konnten, wurden diese Archetypen gemäß der CDASH Spezifikation spezialisiert. Waren keine geeigneten Archetypen vorhanden, wurden neue entwickelt. Studienbezogenen Referenzen, die in Patientenakten normalerweise nicht relevant sind, wurden etwa in speziellen für CDASH entwickelten COMPOSITIONs abgebildet – so beispielsweise das CDASH-Element ‚CMAENO‘, welches die Verknüpfung eines unerwünschten Ereignisses in einer Studie mit einer Begleitmedikation erlaubt. Alle für CDASH spezialisierten und neu entwickelten Archetypen sind gedacht für die Nutzung in einem auf Archetypen basierenden SDMS. Diese Archetypen wurden daher auch für das in Anhang 4 dargestellte Beispiel einer mit Archetypen annotierten ODM-Datei genutzt. Aufbauend auf diesen Sets von Archetypen zur Repräsentation der CDASH Domänen wurde analog einem eCRF für jede betrachtete Domäne ein *openEHR*-Template erstellt (siehe Abb. 4-15 und Abb. 4-16).

Sollen in einem auf Archetypen basierenden SDMS die Datenfelder der CDASH-konformen Templates bzw. eCRF mit im KIS vorhandenen Daten belegt werden, können die beschriebenen Archetypen-Sets genutzt werden, um die benötigten Informationen zu extrahieren. In den KIS muss lediglich nach Daten gesucht werden, welche mit denselben Archetypen repräsentiert werden, die auch zur Dokumentation der CDASH-Elemente vorgesehen sind. Für alle CDASH-Datenelemente, für welche Archetypen studienspezifisch spezialisiert oder neu entwickelt werden mussten, funktioniert dies jedoch nicht, da diese Archetypen für den Einsatz in SDMS, nicht aber in Patientenakten bzw. KIS gedacht sind. Dies bedeutet jedoch nicht, dass für diese Felder keine geeigneten Informationen aus dem KIS extrahiert werden können. In der Regel können über komplexere AQL-Abfragen auch für diese Felder Informationen aus den KIS-Daten abgeleitet werden. Für die betroffenen Datenelemente der betrachteten CDASH-Domänen wurden soweit möglich die AQL-Abfragen definiert (siehe Anhang 1).

Wenn ein Krankenhaus ein auf Archetypen basierendes KIS nutzt, wird es in der Regel möglich sein, die verwendeten Archetypen auch in einem auf Archetypen basierenden SDMS zur Dokumentation in klinischen Studien zu nutzen. Dies vereinfacht gleichzeitig eine unmittelbare multiple Verwendung der im KIS erfassten Daten: Indem zur Beschreibung der Studiendaten im SDMS dieselben Archetypen wie zur Routinedokumentation im KIS genutzt werden, ist direkt ersichtlich, an welchen Stellen im KIS die Informationen erfasst sind, die in einer Studie benötigt werden. Zwar hat sich gezeigt, dass die Dokumentation in Studien oft stärker strukturierte Daten erfordert als in der Routineversorgung (vgl. Kapitel 4.1), doch kann davon ausgegangen werden, dass in auf Archetypen basierenden KIS die Informationen bereits relativ stark strukturiert dokumentiert werden. Dies liegt begründet in der Konstruktion der Systeme basierend auf einem Referenzinformationsmodell einerseits und den darauf aufbauenden Archetypen andererseits (vgl. Kapitel 2.5).

Spezialisierungen und Templates

Dennoch ist es möglich, dass ein Archetyp im KIS ein Freitextfeld an einer Stelle vorsieht, an welcher in einer Studie ein Datenelement mit bestimmten Ausprägungen erhoben werden muss. Beispielsweise ist im eCRF eine Auswahlliste mit den erlaubten Ausprägungen ‚nicht vorhanden‘, ‚mild‘ und ‚schwer‘ für eine bestimmte Merkmalsart vorgesehen. Soll nun in einem SDMS oder einem erweiterten single-source-fähigen KIS dieser vorhandene Archetyp dennoch zur Dokumentation in der entsprechenden Studie genutzt werden, gibt es zwei Lösungsansätze:

- 1.) Für die Anwendung im SDMS kann der bestehende **Archetyp spezialisiert** werden. Die erfassten Daten können nach wie vor von dem System, welches den spezialisierten Archetypen nutzt, an andere Systeme übermittelt werden, welche den ursprünglichen Archetypen nutzen. Spezialisierte Archetypen sind immer ‚aufwärtskompatibel‘ zur allgemeineren Definition. Bestehende Vorgaben dürfen bei der Spezialisierung daher auch nur konkreter und enger gefasst werden, nie aber gelockert oder gestrichen werden. Neue Felder dürfen ergänzt, vorhandene aber nicht entfernt werden. Optionale Felder dürfen in Pflichtfelder umgewandelt werden jedoch nicht umgekehrt.
- 2.) Alternativ könnten die bei der Datenerfassung in einer Studie unerwünschten Spielräume durch **Templates** minimiert werden. Auch Templates ermöglichen es, die Vorgaben, welche ein Archetyp für ein Datenelement definiert, enger zu fassen. Darüber hinaus können Sie genutzt werden, um nicht benötigte optionale Felder auszublenden (vgl. [BEALE und HEARD 2007]. S. 13f und Kapitel 2.6). Mit Hilfe eines Templates können einem Archetyp jedoch ebenfalls keine neuen Felder hinzugefügt werden.

Im vorhergehend genannten Beispiel könnte für die studienbezogene Datenerfassung ein Template genutzt werden, welches die erlaubten Ausprägungen des betreffenden Freitextfelds auf die in der Studie erlaubten Terme einschränkt.

Der Einsatz von spezialisierten Archetypen oder Templates bringt auch einen Nachteil mit sich: Wurden Daten nicht unter Einsatz des spezialisierten Archetyps bzw. des entsprechenden Templates erfasst ist eine direkte multiple Nutzung dieser Daten nicht möglich. Dies kann der Fall sein, wenn die Daten erfasst wurden, bevor der Archetyp spezialisiert wurde oder die Daten in einem System erfasst wurden, welches das Template nicht nutzt. Vor einer Übernahme in die Studie müsste zunächst geprüft werden, ob die bereits erhobenen Daten den Anforderungen des für die Studie spezialisierten Archetypen bzw. des Templates entsprechen.

Ein weiterer Bedarf nach speziellen Archetypen entsteht, wenn dokumentiert werden muss, dass etwas nicht untersucht wurde. Betrachtet man etwa einen Schwangerschaftstest, so können sich in einer Patientenakte hierzu folgende Einträge finden:

- Es wurde ein Test durchgeführt – das Ergebnis war positiv.
- Es wurde ein Test durchgeführt – das Ergebnis war negativ.
- Es wurde kein Test durchgeführt, aber eine Schwangerschaft kann aus anderen Gründen ausgeschlossen werden.
- Es findet sich hierzu kein Eintrag, was impliziert, dass kein Test durchgeführt wurde.

Archetypen sind zunächst darauf ausgelegt, Beobachtungen oder Handlungen zu dokumentieren. Wurde ein Test oder eine Beobachtung nicht durchgeführt, kann dies häufig nur implizit daraus geschlossen werden, dass kein Wert dokumentiert ist. Gerade in klinischen Studien ist es jedoch häufig erforderlich, dass der letzte Fall (Test bzw. Beobachtung wurde nicht durchgeführt) explizit dokumentiert wird. Hierfür können spezielle ‚Ausschluss-Archetypen‘ genutzt werden, so zum Beispiel `openEHR-EHR-EVALUATION.exclusion-medication.v1` zur Dokumentation dass keine Medikation oder nur eine bestimmte Medikation nicht eingenommen wurde.

In klinischen Studien werden Informationen häufig in Form von ‚Ja/Nein‘-Fragen erfasst, zum Beispiel ‚Nimmt der Patient Medikament XY?‘. Dies wurde auch in der durchgeführten Fallstudie deutlich. Beim Entwurf von eCRF in auf Archetypen basierenden SDMS muss berücksichtigt werden, dass es bei solchen Fragen erforderlich sein kann, den positiven bzw. negativen Fall mit unterschiedlichen Archetypen zu dokumentieren. Ein ‚Ja/Nein‘-Feld auf einem eCRF muss so konstruiert werden, dass je nach Antwort der richtige Archetyp zur Repräsentation der erfassten Information gewählt wird.

5.2.5 Archetypen für Studien-Metadaten

In Kapitel 4.4.2 wurden Archetypen beschrieben, die es ermöglichen, Studien-Metadaten zu dokumentieren und gleichzeitig (in Kombination mit *openEHR*-Folders) die erhobenen Studiendaten zu strukturieren. Hierzu stellt sich die Frage, ob eine auf Archetypen basierende Repräsentation von Studien-Metadaten überhaupt notwendig ist. Ein auf Archetypen basierendes SDMS könnte beispielsweise Archetypen ausschließlich nutzen, um Studiendaten zu erfassen. Zum Speichern der Studien-Metadaten könnte auch eine proprietäre, relationale Tabellenstruktur oder soweit möglich das XML-basierte ODM genutzt werden.

Dennoch bietet die auf Archetypen basierende Repräsentation von Studien-Metadaten verschiedene Vorteile. In einem auf Archetypen basierenden SDMS, können alle Funktionen, die zur Dokumentation und zum Austausch von Daten mit anderen Systemen sowie zum Bearbeiten der Archetypen selbst bereitgestellt wurden, unmittelbar auch für Studien-Metadaten genutzt werden. Gleiches gilt für single-source-fähige KIS, die auf Archetypen basieren.

Ferner ist eine strukturierte Beschreibung von Studienfragestellung, Ein- und Ausschlusskriterien oder der Anzahl der eingeschlossenen Patienten notwendig, um zu entscheiden, ob Studiendaten etwa in einer Metaanalyse berücksichtigt werden können.

Spezielle Archetypen

Indem Studien-Metadaten strukturiert erfasst werden, wird eine nachhaltige Nutzung der Studiendaten erleichtert. Strukturiert vorliegende Studien-Metadaten bilden darüber hinaus die Grundlage von Auswertungen, die in studierendurchführenden Einrichtungen häufig benötigt werden, wie etwa zum Rekrutierungsverlauf in den einzelnen Zentren einer Studie. Die auf Archetypen basierende Definition von Studien-Metadaten ermöglicht – entsprechend der Natur der Archetypen – die strukturierte Erfassung und Abfrage dieser Daten, ist dabei aber gleichzeitig relativ leicht erweiterbar.

Wie die Bemühungen im Rahmen des BRIDG-Modells (siehe Kapitel 2.7) zeigen, ist im Bereich der protokollgetriebenen klinischen Forschung nicht nur technische, sondern auch menschliche semantische Interoperabilität notwendig. Die vorgestellten Archetypen für Studien-Metadaten leisten hierzu einen Beitrag ebenso wie das BRIDG-Modell. Sie ermöglichen es, klinische Studien und assoziierte Objekte exakt zu beschreiben und diese Beschreibungen auszutauschen. Dabei wird die Bedeutung der einzelnen Merkmalsarten selbst detailliert beschrieben. Im Fall der Archetypen für Studien-Metadaten stellen die Archetyp-internen, für Menschen lesbaren Beschreibungen der einzelnen Merkmalsarten die Definitionen der Studien-Metadaten dar. Somit handelt es sich gewissermaßen um ‚Meta-Metadaten‘.

Die betrachteten Studien-Metadaten und die hierfür entworfenen Archetypen fokussieren auf die Beschreibung von Studien, zum Beispiel auf Studienfragestellung oder Ein- und Ausschlusskriterien. Unter ‚Studien-Metadaten‘ wird häufig in einem engeren Sinne aber auch die formale Beschreibung der in einer Studie erhobenen Merkmalsarten verstanden, zum Beispiel die Datentypen und –formate der Studiendaten. Die Erfassung dieser Art von Studien-Metadaten wird in einem auf Archetypen basierenden Studiendatenmanagementsystem per se in Form der genutzten Archetypen definiert und gespeichert.

Gegenstand der *openEHR*-Spezifikationen ist eine elektronische Gesundheitsakte. Die Spezifikationen sind daher auf Dokumentationssysteme ausgerichtet, in deren Mittelpunkt der Patient steht. Dementsprechend gibt es ausgehend vom *openEHR*-Referenzmodell auch keine klar definierte Vorgehensweise zur Dokumentation von nicht patientenbezogenen Daten wie etwa Studien-Metadaten. Für die Studien-Metadaten lässt sich, im Gegensatz zu den typischen Daten in einer Gesundheitsakte, daher nicht eindeutig bestimmen, welche Klasse des Referenzmodells die richtige ist, um die verschiedenen Konzepte (wie beispielsweise eine Studienvisite oder ein Prüfzentrum) zu repräsentieren.

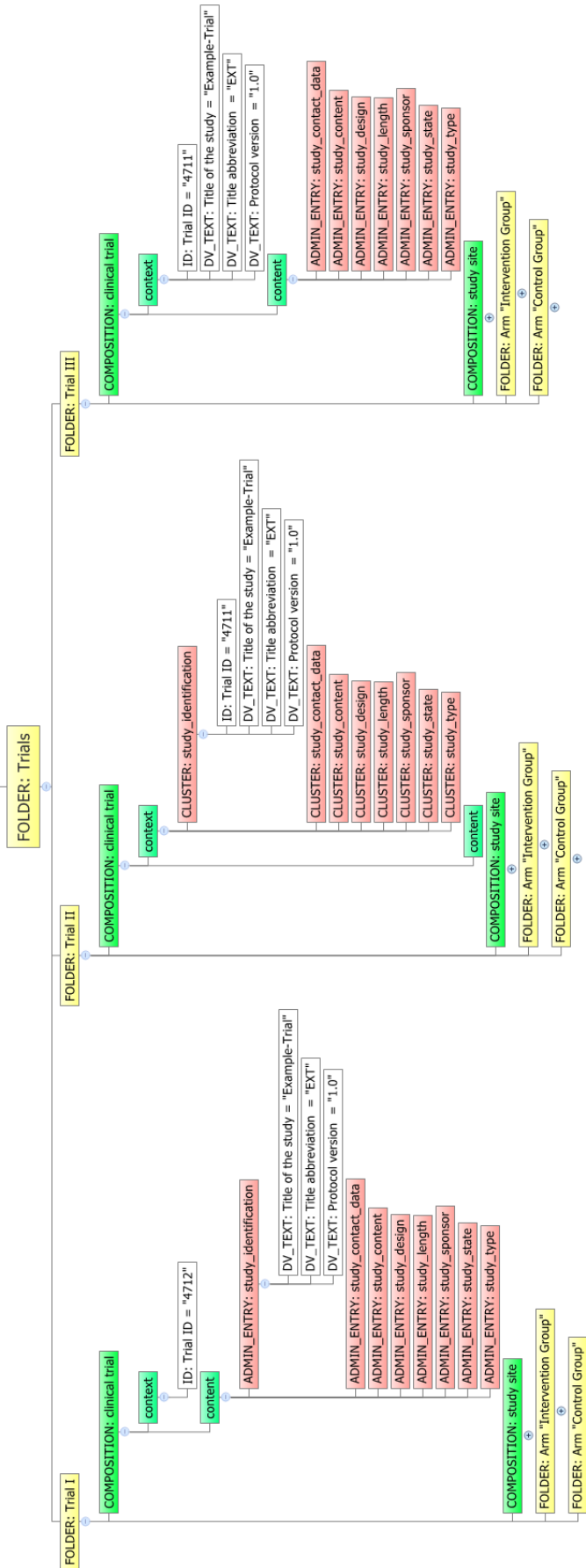


Abb. 5-1: Mögliche Varianten bei der Abbildung von Studien-Metadaten mittels Archetypen.

Alternative Ansätze

Bei den für Studien-Metadaten entwickelten Archetypen handelt es sich um ADMIN_ENTRY-Archetypen (siehe Abb. 5-1 – Trial I). Es wäre jedoch auch möglich gewesen, CLUSTER-Archetypen zu nutzen (siehe Abb. 5-1 – Trial II). Auch bei der Einbindung in die entsprechenden COMPOSITON-Archetypen standen verschiedene Varianten zur Auswahl: Anstatt alle Studien-Metadaten mit Hilfe von ADMIN_ENTRY- oder CLUSTER-Archetypen abzubilden, die dann in die einzelnen COMPOSITION-Archetypen eingebunden werden, könnte zumindest ein Teil der Studien-Metadaten direkt als ‚content‘ der COMPOSITIONS-Archetypen modelliert werden (siehe Abb. 5-1 – Trial III).

Der hier gewählte Ansatz der ‚ADMIN_ENTRY-Archetypen‘, die alle Studien-Metadaten aufnehmen (Abb. 5-1 – Trial I), spiegelt die Idee wider, Studien in Form virtueller Patienten in einem auf Archetypen basierenden System abzubilden. Dabei werden Studien-Metadaten zu diesen virtuellen Patienten (=Studien) auf die gleiche Art gespeichert, wie medizinische Daten zu realen Patienten. Auch openSDMS (siehe Kapitel 4.2) basiert auf diesem Ansatz. Darüber hinaus sorgt der Ansatz, alle Studien-Metadaten mit Hilfe von ADMIN_ENTRY-Archetypen zu modellieren, welche wiederum von entsprechenden COMPOSITON-Archetypen zusammengefasst werden, für eine gewisse Flexibilität. Die ADMIN_ENTRY-Archetypen (zum Beispiel ADMIN_ENTRY.study_center.v1) können bei Bedarf in dem vorgestellten Modell von verschiedenen COMPOSITIONS referenziert oder auch in anderen Modellen genutzt werden.

Eventuell wird das openEHR-Referenzmodell bei einer zukünftigen Revision um den Aspekt nicht patientenbezogener Daten erweitert. Wenn für diese Daten eindeutig geregelt ist, wie sie basierend auf dem openEHR-Referenzmodell repräsentiert werden können, müssten die hier vorgestellten Archetypen für Studien-Metadaten gegebenenfalls entsprechend angepasst werden. Wünschenswert wäre eine solche Erweiterung in jedem Fall. Die in openEHR beschriebene Kombination von Referenzmodell und den darauf basierenden Archetypen bietet neben semantischer Interoperabilität eine große Flexibilität. Diese kann auch im Fall von nicht patientenbezogenen Daten von großem Nutzen sein, wie die vorgelegte Untersuchung für Studien-Metadaten gezeigt hat.

5.2.6 Abbildung auf Standards

In den vorhergehenden Abschnitten wurde deutlich, dass die auf Archetypen basierende Repräsentation von Studiendaten und Studien-Metadaten die multiple Nutzung von Daten in Forschung und Versorgung unterstützen kann. Der vorgeschlagene, auf Archetypen basierende Ansatz soll jedoch nicht dazu führen, dass ein neues, zu etablierten Modellen für die Repräsentation von Studiendaten und Studien-Metadaten inkompatibles Modell entsteht. Durch die Verknüpfung mit bestehenden Modellen soll erreicht werden, dass Systeme, die gemäß der vorgeschlagenen Referenzarchitektur aufgebaut wurden, auch mit Systemen, die nicht auf Archetypen sondern auf bestehenden Standards im Forschungsumfeld setzen, Daten austauschen können. Ferner konnte durch die Abbildung auf etablierte Modelle gezeigt werden, dass sich die entwickelten Archetypen prinzipiell zur Repräsentation von Studiendaten und Studien-Metadaten eignen.

BRIDG

In Kapitel 4.5 wurde gezeigt, dass alle Merkmalsarten, die durch die in Kapitel 4.4.2 vorgestellten Archetypen für Studien-Metadaten definiert werden, auf das BRIDG-Modell abbildbar sind. Hierzu wurde untersucht, ob und wo im BRIDG-Modell sich die von den Archetypen definierten Merkmalsarten wiederfinden. Die identifizierten Abbildungen wurden mit Hilfe einer Abbildungstabelle (siehe Anhang 3) dokumentiert. Systeme, welche die Archetypen für Studien-Metadaten nutzen, bieten so den Vorteil der Abbildbarkeit auf das BRIDG-Modell und damit auf alle Modelle, die ebenfalls auf BRIDG abgebildet werden können. Basierend auf diesen Abbildungen können Austauschmechanismen zwischen Systemen, die diese Modelle nutzen, implementiert werden. Dies entspricht exakt der Intention des BRIDG-Modells (vgl. [ANGELES, EVANS et al. 2010] Kapitel 2.8). Gleichzeitig

kann die Flexibilität der Archetypen genutzt werden, um relativ einfach, etwa lokal benötigte, in BRIDG jedoch noch nicht abgebildete Studien-Metamerkmale mit Hilfe spezialisierter Archetypen zu ergänzen.

Das BRIDG-Modell stellt Konzepte, Attribute und Beziehungen der Domäne der protokollgetriebenen Forschung mit Hilfe von UML-Klassendiagrammen dar. So bildet es eine Ontologie des modellierten Bereichs. Archetypen hingegen beschreiben einzelne Konzepte und nicht deren Beziehungen zueinander. Durch die Abbildung auf das BRIDG-Modell werden die von diesem definierten Beziehungen und Zusammenhänge zwischen seinen Elementen auf die Archetypen und deren Elemente übertragen.

6 Fazit und Ausblick

In der vorliegenden Arbeit wurde gezeigt, dass es möglich ist, Studiendaten und Studien-Metadaten basierend auf Archetypen zu erfassen. Während zur Repräsentation von Studien-Metadaten Archetypen neu definiert werden mussten, konnten für Studiendaten vielfach vorhandene Archetypen, die für die Nutzung in Patientenakten entworfen worden waren, genutzt werden. Für die übrigen Studiendaten konnten neue Archetypen entworfen oder vorhandene Archetypen spezialisiert werden.

Archetypen für Studien-(Meta)daten

Ergänzend zu klinischen Daten gewinnen biomedizinische Daten wie Genomdaten oder beschreibende Daten zu Proben in Biobanken immer mehr Bedeutung. Daher sollte künftig auch verstärkt untersucht werden, ob und wie biomedizinische Daten mit Hilfe von Archetypen beschrieben werden können. So wird in [SPÄTH und GRIMSON 2010] berichtet, dass der Ansatz Archetypen zur Modellierung eines Biobank-Informationssystems zu nutzen, zwar vielversprechend, gleichzeitig aber auch mit verschiedenen Schwierigkeiten verbunden sei.

Archetypen für biomedizinische Daten

In der vorliegenden Arbeit wurde darüber hinaus eine auf Archetypen basierende Referenzarchitektur zur multiplen Verwendung von Daten aus Krankenhausinformationssystemen in Studiendatenmanagementsystemen vorgestellt. Diese Architektur ist ausdrücklich auch für die Integrationen von KIS und SDMS in multizentrischen Studien vorgesehen. Kernidee der Architektur ist die auf Archetypen basierende Modellierung von KIS-Inhalten und Case Report Forms. Der Einsatz der Archetype Query Language erlaubt es, Informationen, die zu einzelnen Patienten im KIS vorhanden sind, exakt und feingranular zu extrahieren. Da die in Studien erhobenen Daten durch die Archetypen semantisch beschrieben werden, werden darüber hinaus studienübergreifende Meta-Analysen erleichtert. Zusätzlich können die vorgeschlagenen Technologien genutzt werden, um abzuschätzen, ob vorgegebene Rekrutierungsziele realistisch sind und um die Rekrutierung von Studienteilnehmern zu unterstützen.

Referenzarchitektur

Ein Großteil des Aufwands für eine multiple Nutzung von Daten aus der Routineversorgung resultiert aus der systematischen Untersuchung, welche Merkmalsarten in einem KIS vorhanden und für eine Studie nutzbar sind. Darüber hinaus müssen für Informationen, die zwar im KIS vorliegen, aber nicht unmittelbar nutzbar sind, weil sie zum Beispiel aus verschiedenen Merkmalsarten abgeleitet werden müssen, geeignete Transformationen definiert und implementiert werden. Da diese Aufgaben in einer multizentrischen Studie für jedes Studienzentrum durchgeführt werden müssen, erhöht sich der Aufwand entsprechend. Dabei besteht die Gefahr, dass der Aufwand unangemessen groß wird und die Vorteile einer multiplen Verwendung klinischer Daten dadurch nicht zum Tragen kommen. Die vorgeschlagene Architektur ermöglicht es daher, im KIS dokumentierbare Merkmalsarten, Abbildungen und Transformationen strukturiert zu erfassen und in verschiedenen Studien wiederzuverwenden.

Sollen der vorgeschlagenen Architektur folgend, ein klassisches, nicht auf Archetypen basierendes KIS und ein SDMS integriert werden, müssen die im KIS dokumentierbaren Merkmalsarten zunächst analysiert werden. In einem zweiten Schritt muss dann ein KIS spezifisches Integrationsmodul implementiert werden. Im Falle eines auf Archetypen und den *openEHR*-Spezifikationen basierenden KIS können sowohl dessen Analyse des KIS wie auch die Implementierung eines Integrationsmoduls entfallen. Daher ist es wünschenswert, dass auf Archetypen basierende KIS künftig eine weitere Verbreitung finden. Dieser Wunsch wird auch von [LOVIS, BALL et al. 2010] gestützt. Die Entscheidung der ‚Clinical Information Modeling Initiative‘ zur Definition klinischer Informationsmodelle die Archetype Definition Language bzw. das Archetype Object Model zu nutzen, ist ebenfalls ein Schritt in diese Richtung. Im Rahmen der Clinical Information Modeling Initiative arbeiten verschiedene Organisationen (unter anderem CDISC, HL7 und die openEHR-

Auf Archetypen basierende KIS

Foundation) auf internationaler Ebene zusammen, um die Interoperabilität im Gesundheitswesen zu verbessern.

Darüber hinaus ist es wünschenswert, dass in KIS Daten möglichst strukturiert gespeichert werden. Nur dann können diese Daten multipel, zum Beispiel in einer Studie, genutzt werden. Zwar ist es mitunter möglich, aus elektronisch gespeicherten Freitexten mit Hilfe entsprechender Textmining-Algorithmen Informationen zu extrahieren, dies ist aber meist mit einem großen Aufwand verbunden. Auch in diesem Zusammenhang sind auf Archetypen aufgebaute Informationssysteme von Vorteil. Bedingt durch die auf Archetypen basierende Erfassung, werden Daten in diesen Systemen häufig stärker strukturiert vorliegen als in klassischen Systemen. Darüber hinaus ermöglichen Archetypen bei Bedarf eine einfache und flexible Erweiterung dieser Systeme.

Auf Archetypen basierende SDMS

Im Gegensatz zu auf Archetypen basierenden KIS ist bisher noch kein ausgereiftes, auf Archetypen basierendes SDMS verfügbar. Als Grundlage für eine konkrete Implementierung der vorgeschlagenen Referenzarchitektur sollte daher ein vollständiges, auf Archetypen basierendes SDMS entwickelt werden. SDMS sind spezialisierte Systeme, die, verglichen mit einer EPA, in der Regel wesentlich mehr Funktionen bereitstellen, so beispielsweise zur Unterstützung des Monitoring oder zur flexiblen Erstellung von Eingabemasken (eCRF) durch den Anwender. Aus diesem Grund und basierend auf den Erfahrungen bei der Implementierung von openSDMS, erscheint es am einfachsten, die auf Archetypen basierende Datenrepräsentation in ein bestehendes SDMS zu integrieren.

Weitere Nutzung der Referenzarchitektur

Die Referenzarchitektur zur Integration von Forschung und Versorgung wurde primär ausgelegt auf die Integration von KIS und SDMS. Anstelle von KIS könnten aber auch andere Systeme basierend auf Archetypen integriert werden, so beispielsweise Register oder persönliche Gesundheitsakten. Bedingt durch den sogenannten demographischen Wandel muss künftig die Versorgung einer größer werdenden Gruppe älterer von einer kleiner werdenden Gruppe jüngerer Menschen geschultert werden. Setzt sich die aktuelle Entwicklung fort, wird im Jahr 2050 die Zahl der über 65-jährigen doppelt so groß sein, wie die der unter 20-jährigen (vgl. [EISENMENGER, PÖTZSCH et al. 2006], S. 36). Um diesem Problem begegnen zu können, sollten gerade auch neue Datenquellen wie persönliche elektronische Gesundheitsakten für die Forschung zugänglich gemacht werden.

Umgekehrt könnten mit Hilfe der Referenzarchitektur anstelle eines SDMS auch andere Systeme mit verschiedenen KIS verknüpft werden, so beispielsweise Systeme zur Gesundheitsüberwachung und zur Früherkennung von Epidemien. Hierzu kann im Einzelfall geprüft werden, inwieweit die Architektur geeignet ist und welche Modifikationen sich als hilfreich erweisen.

7 Zusammenfassung

Sowohl in der Routineversorgung als auch in klinischen Studien werden immer mehr Daten elektronisch verarbeitet. Trotzdem ist ein Austausch von Daten zwischen beiden Bereichen häufig noch nicht etabliert. Dies führt dazu, dass Daten mehrfach erfasst werden müssen. Die redundante Datenerfassung ist zeitaufwändig und kann zu Inkonsistenzen zwischen Krankenhausinformationssystem (KIS) und Studiendatenmanagementsystem (SDMS) führen. Obwohl ein Datenaustausch zwischen Forschung und Versorgung oft technisch möglich wäre, scheitert er meist noch an mangelnder semantischer Interoperabilität.

Archetypen sind ein innovatives Konzept zur Gestaltung von flexiblen und leicht erweiterbaren elektronischen Gesundheitsakten. Sie ermöglichen semantische Interoperabilität zwischen Systemen, welche dieselben Archetypen nutzen. Das Archetypen-Konzept hat mittlerweile auch Eingang in internationale Standards gefunden (ISO 13606). Die *openEHR*-Spezifikationen definieren ein mit ISO 13606 kompatibles jedoch weitergehendes Modell für elektronische Gesundheitsakten. Bisher wurden Archetypen hauptsächlich für Informationssysteme in der Routineversorgung und weniger für die klinische Forschung entwickelt und genutzt. Ziel dieser Arbeit war es daher, basierend auf den *openEHR*-Spezifikationen und Archetypen generische Ansätze zu erarbeiten, die eine multiple Verwendung von Daten aus der Versorgung in der Forschung ermöglichen und deren Umsetzbarkeit zu prüfen.

In einer Voruntersuchung wurde ermittelt, dass 35 % der in der betrachteten Studie zu erhebenden Merkmalsarten aus dem untersuchten KIS übernommen werden könnten, wenn die Daten dort elektronisch und ausreichend strukturiert vorlägen. In einem zweiten Schritt wurde mit *openSDMS* der Prototyp eines auf Archetypen basierenden integrierten elektronischen Gesundheitsakten- und Studiendatenmanagementsystems zur Verfügung gestellt. Aus der Voruntersuchung und der Implementierung von *openSDMS* wurden Anforderungen abgeleitet und eine auf *openEHR*-Archetypen basierende Referenzarchitektur entwickelt, welche die Nutzung von Daten aus KIS in klinischen Studien unterstützt. Dabei wird sowohl die Integration von KIS beschrieben, die auf Archetypen basieren, als auch von klassischen KIS. Kernkomponenten dieser Architektur sind auf Archetypen basierende semantische Annotationen von Studiendaten sowie Import- und Exportmodule, welche die Archetype Query Language nutzen. Die vorgestellte Referenzarchitektur ermöglicht den Übergang von der multiplen Erfassung hin zur multiplen Verwendung von Daten in Forschung und Versorgung.

Um die entwickelte Referenzarchitektur realisieren zu können, werden geeignete Archetypen auch für Forschungsdaten benötigt. Daher wurden Archetypen zur Dokumentation aller Datenelemente der vier CDASH Domänen ‚Common Identifier Variables‘, ‚Common Timing Variables‘, ‚Adverse Events‘ sowie ‚Prior and Concomitant Medications‘ spezifiziert (Studiendaten). Hierzu wurden insgesamt 23 Merkmalsarten basierend auf Archetypen neu definiert, wozu drei bestehende Archetypen spezialisiert und zwei neu entwickelt wurden. Zur Definition von CDASH-konformen elektronischen Datenerhebungsbogen für die betrachteten Domänen wurden, basierend auf den spezifizierten Archetypen, vier *openEHR*-Templates entworfen. Ferner wurden 71 Merkmalsarten in 16 Archetypen zur Dokumentation von Studien-Metadaten definiert. Alle neu entworfenen Archetypen wurden jeweils in englischer und deutscher Sprache beschrieben und können nun als Referenzinformationsmodell für Forschungsdaten genutzt werden. Ergänzend wurden alle von den bereitgestellten Archetypen definierten Merkmalsarten auf die im Bereich der klinischen Forschung etablierten Modelle BRIDG, CDASH und ODM abgebildet.

Summary

More and more data is collected electronically in patient care and in clinical trials. However, the exchange of data between these areas is not well-established. This often causes a multiple collection of the same data. Redundant data collection is time consuming and may lead to inconsistencies between hospital information system (HIS) and clinical trial data management system. Although the exchange of data between research and patient care is often possible on a technical level, it is hard to achieve semantic interoperability.

Archetypes are an innovative approach to build flexible and easily adaptable electronic health records. Archetypes facilitate semantic interoperability among all systems which use the same archetypes. The archetype approach strongly influenced the international ISO 13606 standard on electronic health record exchange. The *openEHR*-specifications define a model for electronic health records which is compatible with ISO 13606 but more comprehensive. Up to now, archetypes have been developed and used mainly for regular clinical care but not for clinical research. The aim of this work was to develop generic approaches based on the *openEHR*-specifications and archetypes which facilitate a multiple use of data collected in patient care for research and to verify the feasibility of this approach.

In a preparatory study it was found that 35 % of the data items required in the analyzed trial could be reused from the HIS if all data were highly structured and stored electronically. In a second step *openSDMS* as a prototype of a combined electronic health record and clinical trial data management system has been implemented. Based on the preparatory study and the implementation of *openSDMS* requirements have been analyzed and a reference architecture based on *openEHR*-archetypes has been developed. This architecture includes the integration of both, archetype based and conventional HIS. Core elements of this architecture are import and export modules which make use of the archetype query language and the semantic annotation of study data based on archetypes. This reference architecture facilitates the transition from a multiple collection to a multiple use of data in clinical research and care.

Archetypes for research data are required to be able to implement the proposed architecture. Therefore, archetypes have been developed to represent all data elements of the CDASH domains 'Common Identifier Variables', 'Common Timing Variables', 'Adverse Events' and 'Prior and Concomitant Medications'. For 23 attributes for which no archetype based representation existed, two archetypes had to be newly defined and three had to be specialized. An *openEHR*-template has been generated for each of the four investigated CDASH domains to be able to define CDASH-compliant electronic case report forms based on archetypes. To represent study meta data 71 attributes in 16 archetypes have been defined. All newly built archetypes have been provided in English and German and may be used as reference information model for clinical research data. In addition, all attributes defined by these archetypes have been mapped to BRIDG, CDASH and ODM as established models in clinical research.

Literaturverzeichnis

- Aerts J (2009).
Newsletter - February 2009. Bimonthly newsletter of XML4Pharma.
XML4Pharma, Singen
- Ammenwerth E und Haux R (2005)
IT-Projektmanagement in Krankenhaus und Gesundheitswesen. Einführendes Lehrbuch
und Projektleitfaden für das taktische Management von Informationssystemen.
Schattauer, Stuttgart
- Ammenwerth E, Iller C und Mansmann U (2003)
Can evaluation studies benefit from triangulation? A case study.
International Journal of Medical Informatics 70: 237-248
- Angeles B, Evans J, Hastak S, McKenzie L, Mead C, Sandberg S, Ver Hoef W und
Woodcock M (2010).
BRIDG Model - Release 3.0.2 User's Guide.
BRIDG Semantic Coordination Committee (SCC) (Hrsg.)
Biomedical Research Integrated Domain Group
- Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten der DGSMP und Arbeitsgruppe
Epidemiologische Methoden der DGEpi der GMDS und der DGSMP (2008).
GPS - Gute Praxis Sekundärdatenanalyse: Revision nach grundlegender Überarbeitung.
- Arikan S und Ingram D (2009)
Web Technologies used in Opereffa.
[Online im Internet] URL: <http://opereffa.chime.ucl.ac.uk/technologies.jsf> [Stand:
23.09.2009]
- Ärzte Zeitung (2011)
Aus für digitale Krankenakte Google Health.
[Online im Internet] URL: <http://www.aerztezeitung.de/panorama/article/660558/digitale-krankenakte-google-health.html> [Stand: 23.11.2011]
- Austin T, Lim YS, Nguyen D und Kalra D (2011)
Design of an Electronic Healthcare Record Server Based on Part 1 of ISO EN 13606.
Journal of Healthcare Engineering 2: 143-160
- Bain L (2009).
RFD. CDISC Newsletter.
Clinical Data Interchange Standards Consortium, Austin
- Beale T (2006).
The *openEHR* Reference Model. Integration Information Model.
openEHR Foundation
- Beale T und Heard S (2007).
openEHR. Archetype Definitions and Principles.
openEHR Foundation
- Beale T, Heard S, Kalra D und Lloyd D (2006).
openEHR Architecture. Architecture Overview.
openEHR Foundation
- Beale T, Heard S, Kalra D und Lloyd D (2008).
The *openEHR* Reference Model. EHR Information Model.
openEHR Foundation
- Breil B, Semjonow A, Muller-Tidow C, Fritz F und Dugas M (2011)
HIS-based Kaplan-Meier plots - A single source approach for documenting and reusing
routine survival information.
BMC Med Inform Decis Mak 11.

- BRIDG (2008).
BRIDG Frequently Asked Questions.
Biomedical Research Integrated Domain Group
- Buck J, Garde S, Kohl CD und Knaup-Gregori P (2009)
Towards a comprehensive electronic patient record to support an innovative individual care concept for premature infants using the openEHR approach.
International Journal of Medical Informatics 78: 521-531
- CDISC (2010a).
Clinical Data Acquisition Standards Harmonization (CDASH).
Clinical Data Interchange Standards Consortium
- CDISC (2010b).
Specification for the Operational Data Model (ODM) - Version 1.3.1 Production.
Clinical Data Interchange Standards Consortium
- CDISC (2011)
CDASH - The Clinical Data Acquisition Standards Harmonization Standard [Online im Internet] URL: <http://www.cdisc.org/cdash> [Stand: 06.02.2011]
- Chen R und Klein G (2007)
The *openEHR* Java reference implementation project. 58-62.
In Kuhn KA, Warren JR und Leong T-Y (Hrsg.):
MEDINFO 2007 - Proceedings of the 12th World Congress on Health (Medical) Informatics – Building Sustainable Health Systems.
Studies in Health Technology Informatics 129.
IOS Press, Amsterdam
- Dorda W, Gall W und Duftschmid G (2002)
Clinical data retrieval: 25 years of temporal query management at the University of Vienna Medical School.
Methods of Information in Medicine 41: 89-97
- Dugas M, Lange M, Muller-Tidow C, Kirchhof P und Prokosch HU (2010)
Routine data from hospital information systems can support patient recruitment for clinical studies.
Clinical Trials 7: 183-189
- EMA (2010).
Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials.
GCP Inspectors Working Group (GCP IWG) (Hrsg.)
European Medicines Agency, London
- Embi PJ und Payne PR (2009)
Clinical research informatics: challenges, opportunities and definition for an emerging domain.
Journal of the American Medical Informatics Association 16: 316-327
- Eysenbach G (2001)
What is e-health?
Journal of Medical Internet Research 3.
- FDA (2010).
Guidance for Industry - Electronic Source Documentation in Clinical Investigations [DRAFT GUIDANCE].
Food and Drug Administration, Rockville

- Garde S, Knaup P, Schuler T und Hovenga E (2005)
Can openEHR Archetypes Empower Multi-Centre Clinical Research? 971-976.
In Engelbrecht R, Geissbuhler A, Lovis C und Mihalas G (Hrsg.):
Proceedings of MIE 2005 - Connecting Medical Informatics and Bio-Informatics.
Studies in Health Technology Informatics 116.
IOS Press, Amsterdam
- Gawlik CS, Abholz H-H, Burkhard BB, Gaus W, Köbberling J und Sehrt-Ricken UM (1998)
Beurteilung klinischer Therapiestudien: Mindeststandards für den Arbeitsalltag.
Deutsches Ärzteblatt 95: 1155-1160
- Haas P (2006)
Gesundheitstelematik. Grundlagen, Anwendungen, Potenziale.
Springer, Berlin Heidelberg New York
- Hagglund M, Chen R und Koch S (2010)
Modeling shared care plans using CONTsys and openEHR to support shared homecare of
the elderly.
Journal of the American Medical Informatics Association 18: 66-69
- HIMSS (2003).
HIMSS E-Health SIG White Paper.
HIMSS E-Health Special Interest Group (Hrsg.)
Healthcare Information and Management Systems Society
- Hoerbst A, Kohl CD, Knaup P und Ammenwerth E (2009)
Attitudes and behaviors related to the introduction of electronic health records among
Austrian and German citizens.
International Journal of Medical Informatics 79: 81-89
- Iberson-Hurst D (2004)
The CDISC Operational Data Model: Ready to Roll?
Applied Clinical Trials 13: 48-53
- ICH (2002).
ICH Topic E 6 (R1) - Guideline for Good Clinical Practice International Conference on
Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human
Use (Hrsg.)
European Medicines Agency, London
- Integrating the Healthcare Enterprise (2008).
IHE IT Infrastructure (ITI) Technical Framework Supplement: Retrieve Form for Data
Capture (RFD).
- InterComponentWare AG (ICW) (2011)
LifeSensor.
[Online im Internet] URL: <http://www.icw-global.com/de/intercomponentware-ag/lifesensor-gesundheitsakte.html> [Stand: 23.11.2011]
- Kaspar M (2008).
Integration genomischer Daten in Systeme der klinischen Versorgung und der klinischen
Forschung.
Masterarbeit, Georg-August-Universität Göttingen
- Katzer C, Röhrig R, Weismüller K, Little S, Langefeld T, Meier M, Menges T, Chakraborty
T und Hempelmann G (2006).
Erfassung, Verarbeitung und Qualitätssicherung von Studiendaten mit einem klinischen
Informationssystem – Ein Erfahrungsbericht. 344-345.
In: Löffler M und Winter A (Hrsg.):
51. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und
Epidemiologie (gmds), Leipzig
German Medical Science, Düsseldorf Köln

- Knaup P, Pilz J und Thalheimer M (2006)
Can temporary paper-based patient records sensibly complete an electronic patient record? 207-212.
In: Hasman A, Haux R, van der Lei J, De Clercq E, Roger-France F (Hrsg.):
Proceedings of MIE2006 - Ubiquity: Technologies for Better Health in Aging Societies.
Studies in Health Technology Informatics 124.
IOS Press, Amsterdam
- Kohane IS und Altman RB (2005)
Health-information altruists - a potentially critical resource.
New England Journal of Medicine 353: 2074-2077
- Kohl CD, Garde S und Knaup P (2008)
Facilitating the openEHR approach - organizational structures for defining high-quality archetypes. 437-442.
In Andersen S, Klein G, Schulz S, Aarts J und Mazzoleni M (Hrsg.):
MIE 2008 - Proceedings of the 21st International Congress of the European Federation
for Medical Informatics: eHealth Beyond the Horizon – Get IT There.
Studies in Health Technology Informatics 136.
IOS Press, Amsterdam
- Kruel R (2010).
Abbildung von CDASH-Feldern auf Archetypen zur Nutzung von Daten aus
Patientenakten für die medizinische Forschung.
Masterarbeit, Ruprecht-Karls-Universität Heidelberg
- Krüger-Brand HE (2007)
MEDICA 2007: Gesundheitstelematik. Stark im Trend: Patientenakten & Co.
Deutsches Ärzteblatt 104
- Kush R, Alschuler L, Ruggeri R, Cassells S, Gupta N, Bain L, Claise K, Shah M und Nahm M (2007)
Implementing Single Source: the STARBRITE proof-of-concept study.
Journal of the American Medical Informatics Association 14: 662-673
- Leslie H (2007)
OpenEHR: The World's Record.
Pulse IT: 50-55
- Leslie H und Heard S (2006)
Archetypes 101. 18-23.
In Westbrook J und Callen J (Hrsg.):
HIC 2006 - Bridging the Digital Divide: Clinician, consumer and computer.
Health Informatics Society of Australia Ltd (HISA), Brunswick East
- Lovis C, Ball M, Boyer C, Elkin PL, Ishikawa K, Jaffe C, March A, Marin H, Mykkanen J, Rienhoff O, Silva J, Sittig DF und Talmon J (2010)
Hospital and Health Information Systems - Current Perspectives. Contribution of the
IMIA Health Information Systems Working Group.
Yearbook of Medical Informatics 6: 73-82
- Luntz SP, Gorbauch T, Schröder B und Seibert-Grafe M (2009)
Aufgaben der Koordinierungszentren für Klinische Studien.
Bundesgesundheitsblatt 52: 444-450

- Ma C, Frankel H, Beale T und Heard S (2007)
EHR query language (EQL) - a query language for archetype-based health records. 397-401.
In Kuhn KA, Warren JR und Leong T-Y (Hrsg.):
MEDINFO 2007 - Proceedings of the 12th World Congress on Health (Medical) Informatics - Building Sustainable Health Systems.
Studies in Health Technology Informatics 129.
IOS Press, Amsterdam
- Martinez-Costa C, Menarguez-Tortosa M und Fernandez-Breis JT (2010)
An approach for the semantic interoperability of ISO EN 13606 and OpenEHR archetypes.
Journal of Biomedical Informatics 43: 736-746
- Martinez-Costa C, Menarguez-Tortosa M und Fernandez-Breis JT (2011)
Clinical data interoperability based on archetype transformation.
Journal of Biomedical Informatics 44: 869-880
- Mayring P (2007)
Qualitative Inhaltsanalyse: Grundlagen und Techniken.
9. Aufl. Beltz, Weinheim und Basel
- Michelsen L, Pedersen SS, Tilma HB und Andersen SK (2005)
Comparing different approaches to two-level modelling of electronic health records. 113-118.
In Engelbrecht R, Geissbuhler A, Lovis C und Mihalas G (Hrsg.):
Proceedings of MIE 2005 – The XIXth International Congress of the European Federation for Medical Informatics.
Studies in Health Technology Informatics 116.
IOS Press, Amsterdam
- Müller-Mielitz S, Ohmann C und Goldschmidt AJW (2010).
Kosten-Nutzen-Analyse eines eSource Projekts: Ein Beispiel aus dem Kompetenznetz Angeborene Herzfehler. 99-106.
In: Schreier G, Hayn D und Ammenwerth E (Hrsg.):
eHealth 2010 - Wien
Österreichische Computergesellschaft (OCG)
- Norgall T (2009)
HL7 Benutzergruppe in Deutschland e.V.
[Online im Internet] URL: <http://www.hl7.de/index.php> [Stand: 09.02.2009]
- Ohmann C und Kuchinke W (2009)
Future developments of medical informatics from the viewpoint of networked clinical research. Interoperability and integration.
Methods of Information in Medicine 48: 45-54
- openEHR (2007)
What is the difference between archetypes and ontologies?
[Online im Internet] URL: <http://www.openehr.org/119-OE.html> [Stand: 17.02.2009]
- Oster S, Langella S, Hastings S, Ervin D, Madduri R, Phillips J, Kurc T, Siebenlist F, Covitz P, Shanbhag K, Foster I und Saltz J (2008)
caGrid 1.0: an enterprise Grid infrastructure for biomedical research.
Journal of the American Medical Informatics Association 15: 138-149
- Pommerening K, Drepper J, Ganslandt T, Helbing K, Müller T, Sax U, Semler S und Speer R (2009).
Das TMF-Datenschutzkonzept für medizinische Datensammlungen und Biobanken
In: Fischer S, Maehle E und Reischuk R (Hrsg.):
INFORMATIK 2009 - Im Focus des Lebens, Lübeck
Gesellschaft für Informatik

- Prokosch HU, Ries M, Beyer A, Schwenk M, Seggewies C, Kopcke F, Mate S, Martin M, Barthlein B, Beckmann MW, Sturzl M, Croner R, Wullich B, Ganslandt T und Burkle T (2011)
IT infrastructure components to support clinical care and translational research projects in a comprehensive cancer center. 892-896.
In Moen A, Andersen SK, Aarts J und Hurlen P (Hrsg.):
Proceedings of MIE 2011 - User Centred Networked Health Care.
Studies in Health Technology Informatics 169.
IOS Press, Amsterdam
- Prokosch HU, Dugas M, Breil B, Müller M und Beyer A (2009).
Analyse der Ein-/Ausschlusskriterien einer größeren zufälligen Stichprobe von Studien - Abschlussbericht TMF-Projekt V054-01. TMF-Projekt ‚IT-Strategie‘.
TMF - Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., Berlin
- Prokosch HU und Ganslandt T (2009)
Perspectives for medical informatics. Reusing the electronic medical record for clinical research.
Methods of Information in Medicine 48: 38-44
- Reng C-M, Debold P, Specker C und Pommerening K (2006)
Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin.
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin
- Sachdeva S und Bhalla S (2009).
Implementing High-Level Query Language Interfaces for Archetype-Based Electronic Health Records Database
In: 15th International Conference on Management of Data COMAD 2009, Mysore, India
Computer Society of India
- Sax U (2008).
Anforderungen der modernen Verbundforschung an die Biomedizinische Informatik.
Vortrag (Heidelberger Kolloquium Medizinische Biometrie, Informatik und Epidemiologie)
- Schloeffel P, Beale T, Hayworth G, Heard S und Leslie H (2006)
The relationship between CEN 13606, HL7, and openEHR. 24-28.
In Westbrook J und Callen J (Hrsg.):
HIC 2006 - Bridging the Digital Divide: Clinician, consumer and computer.
Health Informatics Society of Australia Ltd (HISA), Brunswick East
- Späth MB und Grimson J (2010)
Applying the archetype approach to the database of a biobank information management system.
International Journal of Medical Informatics 80: 205-226
- Staats- und Universitätsbibliothek Göttingen (2001)
Metadata Server - Einführung in Metadaten.
[Online im Internet] URL: <http://www2.sub.uni-goettingen.de/intrometa.html> [Stand: 25.08.2010]
- TMF (2009)
Metadata Repository.
[Online im Internet] URL: http://www.tmf-ev.de/Themen/Projekte/V063_01_Metadata_Repository.aspx [Stand: 10.01.2012]
- Tuffs A und Pietzsch J (2008).
In scientia salus: Heilung durch Wissenschaft.
Universitätsklinikum Heidelberg, Heidelberg

-
- Walker J, Pan E, Johnston D, Adler-Milstein J, Bates DW und Middleton B (2005)
The value of health care information exchange and interoperability.
Health Affairs Suppl Web Exclusives: W5-10-W15-18
- Warda F (2006)
Patienten-Empowerment durch den Einsatz elektronischer Gesundheitsakten. 374-378.
In Jäckel A (Hrsg.):
Telemedizinführer Deutschland 2006.
7. Aufl. Minerva, Darmstadt
- Webster N (1971)
Webster's third new international dictionary of the English language: unabridged.
G. & C. Merriam Company, Springfield
- Winter A, Funkat G, Haeber A, Mauz-Koerholz C, Pommerening K, Smers S und Stausberg J (2007)
Integrated information systems for translational medicine.
Methods of Information in Medicine 46: 601-607
- Xiao L, Cousins G, Courtney B, Hederman L, Fahey T und Dimitrov BD (2011)
Developing an electronic health record (EHR) for methadone treatment recording and decision support.
BMC Med Inform Decis Mak 11
- ZPID (2010)
Metadaten - Den Forschungskontext dokumentieren.
[Online im Internet] URL: <http://psychdata.zpid.de/index.php?main=give&sub=metadaten> [Stand: 25.08.2010]

Eigene Veröffentlichungen

Medline-referenzierte Originalarbeiten

Kohl CD, Garde S, Knaup P (2010):

Facilitating secondary use of medical data by using openEHR archetypes. 1117-1121.
 In: Safran C, Reti S, Marin HF (Hrsg.):
 Medinfo 2010 - Proceedings of the 13th International Congress on Medical and Health Informatics: Partnerships for effective e-Health solutions - Innovative collaborations promote solutions to health challenges.
 Studies in Health Technology and Informatics 160.
 IOS Press, Amsterdam

Hoerbst A, Kohl CD, Knaup P und Ammenwerth E (2009)

Attitudes and behaviors related to the introduction of electronic health records among Austrian and German citizens.
 International Journal of Medical Informatics 79: 81-89

Buck J, Garde S, Kohl CD und Knaup-Gregori P (2009)

Towards a comprehensive electronic patient record to support an innovative individual care concept for premature infants using the openEHR approach.
 International Journal of Medical Informatics 78: 521-531

Kohl CD, Garde S, Knaup P (2008):

Facilitating the openEHR approach - organizational structures for defining high-quality archetypes. 437-442.
 In: Andersen SK, Klein GO, Schulz S, Aarts J, Mazzoleni MC (Hrsg.):
 MIE 2008 - Proceedings of the 21st International Congress of the European Federation for Medical Informatics: eHealth Beyond the Horizon – Get IT There.
 Studies in Health Technology and Informatics 136.
 IOS Press, Amsterdam

Kohl CD, Schott C, Verveur D, Pöschl J, Knaup P (2007):

Introducing an Innovative, Multidisciplinary Concept of Care - Communication Problems and Possible Solutions. 241-246.
 In: Coiera E, Westbrook JI, Callen JL, Aarts J (Hrsg.):
 ITHC 2007 - Proceedings of the 3rd International Conference on Information Technology in Health Care: Socio-technical Approaches.
 Studies in Health Technology and Informatics 130.
 IOS Press, Amsterdam

Kohl CD, Schott C, Verveur D, Linderkamp O, Knaup-Gregori P (2007):

Can an EPR Support the Concept of Family-Centred, Individualized Developmental Care of Premature Infants and Newborns? 73-77.
 In: Kuhn KA, Warren JR, Leong T-Y (Hrsg.):
 Medinfo 2007 - Proceedings of the 12th World Congress on Health (Medical) Informatics: Building Sustainable Health Systems.
 Studies in Health Technology and Informatics 129.
 IOS Press, Amsterdam

Knaup P, Bott O, Kohl C, Lovis C, Garde S (2007):

Electronic patient records: Moving from islands and bridges towards electronic health records for continuity of care. 34-46.
 In: Geissbuhler A, Haux R, Kulikowski C (Hrsg.):
 IMIA Yearbook of Medical Informatics 2007.
 Methods of Information in Medicine 46: Suppl 1
 Schattauer, Stuttgart

Weitere Originalarbeiten in Zeitschriften

Tönshoff B, Plotnicki L, Feneberg R, Kohl C, Höcker B (2011):

Etablierung eines webbasierten Registers für die pädiatrische Nierentransplantation:
Das CERTAIN Registry.

Nieren- und Hochdruckkrankheiten 40: 320-326.

Knaup-Gregori P, Heid J, Kohl C, Zimmermann M, Pritsch M (2007):

Elektronische Prüfungen - mehr als nur ankreuzen!

MDI - Forum der Medizin_Dokumentation und Medizin_Informatik 9: 105-109.

Anhang

1. Abbildung CDASH auf openEHR-Archetypen

Die nachfolgenden Tabellen bilden die CDASH-Datenelemente der Domänen ‚Common Identifier Variables‘, ‚Common Timing Variables‘, ‚Adverse Events‘ und ‚Prior and Concomitant Medications‘ auf bestehende und neu entwickelte Archetypen ab (im Rahmen dieser Arbeit neu entwickelte bzw. spezialisierte Archetypen sind durch ein ‚x‘ in der Spalte ‚neu‘ gekennzeichnet). Alle für CDASH neu entwickelten oder spezialisierten Archetypen können in einem SDMS verwendet werden für die Generierung von auf Archetypen basierenden CRF zur Erhebung der jeweiligen CDASH-Felder. In einer Patientenakte werden diese Archetypen jedoch nicht genutzt werden. Daher kann in Patientenakten auch nicht ausgehend von diesen Archetypen nach Daten gesucht werden, welche sich für eine multiple Verwendung in Studien eignen. Um auch diese CDASH-Felder mit den Daten, die in einer auf openEHR-Archetypen basierenden Patientenakten gegebenenfalls vorhanden sind, vorbelegen zu können, werden in diesen Fällen, soweit möglich, zusätzlich AQL-Abfragen angegeben. Ferner wird, ergänzend zum Datentyp einer Merkmalsart, das im jeweiligen Archetypen festgelegte Vorkommen dieser Merkmalsart angegeben (zum Beispiel ‚1..1‘ – zu dieser Merkmalsart muss eine Ausprägung angegeben werden).

Common Identifier Variables

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)
1	SPONSOR	Text	O	openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	Primary sponsor	AT_SLOT	1..1	Person oder Organisation möglich.	Keine Entsprechung in einer Patientenakte.
2	STUDYID	Text	HR	openEHR-EHR-ADMIN_ENTRY.study_identification.v1	Unique trial ID	DV_Text	1..1		Keine Entsprechung in einer Patientenakte.
3	SITEID	Text	HR	openEHR-EHR-ADMIN_ENTRY.study_site_description.v1	Name of site	DV_Text	1..1		Keine Entsprechung in einer Patientenakte.
4	SUBJID	Text	HR	< record ID / system internal >					Keine Entsprechung in einer Patientenakte.
5	USUBJID	Text	O	< system external >					Keine Entsprechung in einer Patientenakte.
6	INVID	Text	O	< user ID / system internal >					Keine Entsprechung in einer Patientenakte.

Common Timing Variables

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)
1	VISITNUM	Text	O	openEHR-EHR-COMPOSITION.trial_visit.v1	Visit ID	DV_Identifier	0..1		Keine Entsprechung in einer Patientenakte.
	VISIT	Text	O	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	Name	DV_Text	1..1		Keine Entsprechung in einer Patientenakte.
2	VISDAT	Date	HR	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	Start of visit	DV_Date_Time	0..1	Falls nicht dokumentiert: < Date of recording >	Keine Entsprechung in einer Patientenakte.

3	VIENDAT	Date	O	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	End of visit	DV_Date_Time	0..1		<i>Keine Entsprechung in einer Patientenakte.</i>
4	VISTIM	Time	O	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	Start of visit	DV_Date_Time	0..1		<i>Keine Entsprechung in einer Patientenakte.</i>

Adverse Event

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)
1	AEYN	{YN}	O	openEHR-EHR-EVALUATION.exclusion-adverse.v1	Global Statement	DV_Coded_Text	0..*	Hier nur "nein" / Bedingung.	
2	AESPID	Text (Num)	O	openEHR-EHR-COMPOSITION.cdash_ae.v1	context/AESPID	DV_Text	0..1	x	<i>Keine Entsprechung in einer Patientenakte.</i>
3	AETERM	Text	HR	openEHR-EHR-EVALUATION.adverse.v1	Reaction description	DV_Text	1..1		
8 [sic.]	AEOCCUR	{YN}	O	openEHR-EHR-EVALUATION.exclusion-adverse.v1	General statement	DV_Text	0..*	Hier nur "nein".	
4	AESTDAT	{DATE}	HR	openEHR-EHR-EVALUATION.adverse.v1	Date of onset of reaction	DV_DATE_TIME	0..1	openEHR: Kombiniert Datum und Zeit	
5	AESTTIM	{TIME}	R/C	openEHR-EHR-EVALUATION.adverse.v1	Date of onset of reaction	DV_DATE_TIME	0..1	openEHR: Kombiniert Datum und Zeit	
6	AEENDAT	{DATE}	HR	openEHR-EHR-EVALUATION.adverse.v1	Duration of the reaction	Time unit	0..1	Endzeitpunkt muss aus Beginn und Dauer errechnet werden /	
7	AEENTIM	{TIME}	R/C	openEHR-EHR-EVALUATION.adverse.v1	Duration of the reaction	Time unit	0..1	openEHR: Kombiniert Datum und Zeit	
8	AEONGO	{YN}	O	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Ongoing	DV_BOOLEAN	0..1	x	SELECT evaluationAdverse/data/items[at0019]/items[0028]/value AS Duration FROM HER e [ehr_id/value= \$ehrId_of_current_patient] CONTAINS (COMPOSITION composition CONTAINS EVALUATION evaluationAdverse [openEHR-EHR-EVALUATION.adverse.v1]) WHERE composition/uid/value= \$composition_of_interest

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)	
9	AESEV	{AESEV}	HR	openEHR-EHR-EVALUATION.adverse.v1	Reaction severity	DV_Coded_Text	0..1	CDISC Codelist: C66769		
9a	AETOXGR	{TOXGR}	HR	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Toxicity grade	Terminologie	1..1	x	CDISC Codelist: C66784	<i>Keine Entsprechung in einer Patientenakte.</i>
10	AESER	{YN}	HR	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Seriousness	DV_BOOLEAN	1..1	x		<i>Keine Entsprechung in einer Patientenakte / ggf. Vorschlag durch komplexe Abfrage</i>
11a	AESCONG	{YN}	R/C	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Congenital anomaly or birth defect	DV_BOOLEAN	0..1	x		<i>Keine Entsprechung in einer Patientenakte.</i>
11b	AESDISAB	{YN}	R/C	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Significant disability	DV_BOOLEAN	0..1	x		<i>Keine Entsprechung in einer Patientenakte.</i>
11c	AESDTH	{YN}	R/C	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Death	DV_BOOLEAN	0..1	x		SELECT COUNT (d/ehr_id) AS Death_Data_Available FROM DEMOGRAPHIC d [ehr_id/value=\$ehrId_of_current_patient] CONTAINS (COMPOSITION composition CONTAINS CLUSTER clusterPersonDeathData [openEHR-DEMOGRAPHIC-CLUSTER.person_death_data_iso.v1])
11d	AESHOSP	{YN}	R/C	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Hospitalization	DV_BOOLEAN	0..1	x	Ggf. abgeleitet aus Leistungserbringender Organisation.	<i>Keine Entsprechung in einer Patientenakte.</i>
11e	AESLIFE	{YN}	R/C	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Life threatening	DV_BOOLEAN	0..1	x		<i>Keine Entsprechung in einer Patientenakte.</i>
11f	AESMIE	{YN}	R/C	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Other medically important event	DV_BOOLEAN	0..1	x		<i>Keine Entsprechung in einer Patientenakte.</i>
12	AEREL	Text (Spon)	HR	openEHR-EHR-EVALUATION.adverse.v1	Probability of causation	DV_Coded_Text	0...1			
13	AEACN	{ACN}	HR	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Action taken with study treatment	Terminologie	1..1	x	CDISC Codelist: C66767	<i>Keine Entsprechung in einer Patientenakte.</i>
14	AEACNOTH	Text	O	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Other action taken	DV_Text	0..1	x		<i>Keine Entsprechung in einer Patientenakte.</i>

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)	
15	AEOUT	{OUT}	HR	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Outcome	Terminologie	1..1	x	CDISC Codelist: C66768	Keine Entsprechung in einer Patientenakte.
16	AEDIS	{YN}	O	openEHR-EHR-EVALUATION.adverse-cdash_ae.v1	Caused study discontinuation	DV_BOOLEAN	0..1	x		Keine Entsprechung in einer Patientenakte.

Prior and Concomitant Medication

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)	
1	CMYN	{YN}	O	openEHR-EHR-EVALUATION.exclusion-medication.v1	Global Statements	DV_Coded_Text	0..*		Codierter Text lässt sich auf {YN} abbilden: N = "Never taken any medication".	
2	CMSPID	Text	O	openEHR-EHR-COMPOSITION.cdash_cm.v1	context/CMSPID	DV_Text	0..1	x		Keine Entsprechung in einer Patientenakte.
3	CMTRT	Text	HR	openEHR-EHR-ITEM_TREE.medication.v1	Name of medication	DV_Text	1..1			
4	CMOCCUR	{YN}	O	openEHR-EHR-COMPOSITION.cdash_cm.v1	context/CMOCCUR	DV_Boolean	0..1	x		SELECT COUNT (e/ehr_id) AS Medications FROM EHR e [ehr_id/value= \$ehrId_of_current_patient] CONTAINS (COMPOSITION composition CONTAINS ACTION actionMedication [openEHR- EHR-ACTION.medication.v1])
5	CMINGRD	Text	O	openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1	Active ingredient	DV_Text	0..*	x	Ggf. über externes Arzneimittel-informationssystem; Alternative: openEHR-EHR-ITEM_TREE.medication-formulation.v1.	Keine Entsprechung in einer Patientenakte.
6	CMINDC	Text	R/C	openEHR-EHR-ITEM_TREE.medication.v1	Indication	Cluster / DV_Text v DV_URI	0..*			
7	CMAENO	Text	O	openEHR-EHR-COMPOSITION.cdash_cm.v1	context/CMAENO	DV_Text	0..1	x		Keine Entsprechung in einer Patientenakte.
8	CMMHNO	Text	O	openEHR-EHR-COMPOSITION.cdash_cm.v1	context/CMMHNO	DV_Text	0..1	x		Keine Entsprechung in einer Patientenakte.

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)	
9	CMDSTXT	Text	O	openEHR-EHR-ITEM_TREE.medication.v1	Dose	Cluster	0..1	In openEHR wird zwischen ‚absolut quantity‘ und ‚dose units‘ unterschieden.		
10	CMDOSTOT	Text	O	openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1	Total daily dose	DV_Text	0..1	x	For general medication, it is not recommended to use "Total Daily Dose". Instead, this can be calculated or derived from other fields such as Units, Dose, and Frequency.	<i>Kann gegebenenfalls berechnet werden bairend auf den Feldern ‚Dose‘ und ‚Dose duration‘ des Archetyps openEHR-EHR-ITEM_TREE.medication.v1</i>
11	CMDOSU	{UNIT}	O	openEHR-EHR-ITEM_TREE.medication.v1	Dose unit	DV_Text	0..1		openEHR: Terminologie (erlaubt sind alle Dosisangaben)	
12	CMDOSFRM	{FRM}	O	openEHR-EHR-ITEM_TREE.medication.v1	Form	DV_Text	0..1			
13	CMDOSFRQ	{FREQ}	O	openEHR-EHR-ITEM_TREE.medication.v1	Maximum dose unit frequency	DV_Quantität	0..1		<i>Achtung: Maximal erlaubten Frequenz in openEHR.</i>	
14	CMROUTE	{ROUTE}	R/C	openEHR-EHR-ITEM_TREE.medication.v1	Route	DV_Text	0..1			
15	CMSTDAT	{Date}	HR	openEHR-EHR-ITEM_TREE.medication.v1	Date (time) of first administration	DV_Date_Time	0..1		Kombiniert mit Zeit (CMSTTIM)	
16	CMPRIOR	{YN}	R/C	openEHR-EHR-COMPOSITION.cdash_cm.v1	context/ CMPRIOR	DV_Boolean	0..1	x	SELECT itemTreeMedication/data /items[at0018]/items[0019] /value AS Date_Time_of_First Administration FROM EHR e [ehr_id/value= \$ehrId_of_current_patient] CONTAINS (COMPOSITION composition CONTAINS (ACTION actionMedication [openEHR- EHR-ACTION.medication.v1] CONTAINS (ITEM_TREE itemTreeMedication [openEHR-EHR- ITEM_TREE.medication(-[a-zA- Z0-9_]+)*.v1]))))	

Nr.:	CDASH	Typ	Core	Archetype	Field-Name	Datatype	neu	Bemerkung	Akte (AQL)
									WHERE itemTreeMedication /items[at0001]/value= \$medication_of_interest
17	CMSTTIM	{Time}	R/C	openEHR-EHR- ITEM_TREE.medication.v1	Date (time) of first administration	DV_ Date_ Time	0..1	Kombiniert mit Datum (CMSTDAT)	
18	CMENDAT	{Date}	HR	openEHR-EHR- ITEM_TREE.medication.v1	Date (time) of last administration	DV_ Date_ Time	0..1	Kombiniert mit Zeit (CMENTIM)	
19	CMONGO	{YN}	O	openEHR-EHR- COMPOSITION.cdash_cm.v1	context/ CMSONGO	DV_ Boolean	0..1	x	SELECT itemTreeMedication/data/items[at 0018]/items[0032]/value AS Date_Time_of_Last_ Administration FROM EHR e [ehr_id/value= \$ehrId_of_current_patient] CONTAINS (COMPOSITION composition CONTAINS (ACTION actionMedication [openEHR- EHR-ACTION.medication.v1] CONTAINS (ITEM_TREE itemTreeMedication [openEHR-EHR- ITEM_TREE.medication(-[a-zA- Z0-9_]+)*.v1])))) WHERE itemTreeMedication /items[at0001]/value= \$medication_of_interest
20	CMENTIM	{Time}	R/C	openEHR-EHR- ITEM_TREE.medication.v1	Date (time) of last administration	DV_Date_ Time	0..1	Kombiniert mit Datum (CMENDAT)	

2. Abbildung DRKS und WHO/ICTRP Merkmalsarten zu Beschreibung einer Studie auf openEHR-Archetypen

Die nachfolgende Tabelle bildet Studien-Metamerkmale aufeinander ab, die in folgenden Systemen erhoben werden können:

- Orange: Deutsches Register Klinischer Studien (DRKS).
- Blau: International Clinical Trials Registry Platform (ICTRRP) der Weltgesundheitsorganisation (WHO) – internationales Pendant zum DRKS.
- Grün: KKS Heidelberg - intern genutzten Liste zur Verwaltung der dort betreuten Studienprojekte (es werden nur die Merkmalsarten aufgeführt, die nicht bereits im Rahmen von DRKS und/oder ICTRP genannt wurden).
- Rot: openEHR-Archetypen, die im Rahmen dieser Arbeit entwickelt wurden.

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
DRKS-ID der Studie	Wird automatisch vergeben				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
Registrierungsdatum im DRKS	[Datum] Wird automatisch vergeben		Date of Registration in Primary Registry	Date when trial was officially registered in the Primary Registry. If relevant, also include the date of registration in the Partner Registry.	openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0008]	Date of registration	DV_DATE
Registrierungsdatum im Partner-Register	[Datum]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0008]	Date of registration	DV_DATE
Wissenschaftsinitiierte Studie (IST/IIT)	[Ja Nein]				openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0015]	Type of financing	CODE_PHRASE
Antragsdatum bei der (federführenden) Ethikkommission	[Datum]				openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0011]	Leading Ethics Committee	DV_BOOLEAN
					openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0003]	Request	DV_DATE
Datum des positiven Votums /der zustimmenden Bewertung der (federführenden) Ethikkommission	[Datum]				openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0004]	Vote	DV_DATE

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
					openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0005]	Vote	DV_CODED_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0011]	Leading Ethics Committee	DV_BOOLEAN
Votum der Ethikkommission	[Positives Votum/Zustimmende Bewertung Abschlägige Bewertung Votum steht noch aus Studie vor Votumserteilung zurückgezogen]				openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0005]	Vote	DV_CODED_TEXT
Name und Vorlage-Nr. der (federführenden) Ethikkommission	[Freitext] Vorlagennummer	[Liste deutscher Ethikkommissionen] zum Auswählen; [Freitext] Nur eintragen, wenn zuvor 'andere' (federführende) Ethikkommission ausgewählt wurde			openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0013] Name kann aus Adresse ausgelesen werden.	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_+]*).v1)
Gesetzlicher Rahmen	[AMG MGP Non-AMG/non-MPG]				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0007]	Drug	DV_BOOLEAN
					openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0008]	Narcotic Substance	DV_BOOLEAN
					openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0009]	Medical Device	DV_BOOLEAN
Weitere gesetzliche Rahmen	[RöV StrlSchV GenTG Transfusionsgesetz Andere]				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0010]	Ionizing radiation	DV_BOOLEAN
					openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0013]	Transfusion	DV_BOOLEAN

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
					openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0014]	Genetic engineering	DV_BOOLEAN
Wenn andere gesetzliche Rahmen, bitte hier eintragen	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0029]	Legal context	DV_TEXT
UTN (Universal trial number)	[Nummer]		Secondary Identifying Numbers	Other identifiers besides the Trial Identifying Number allocated by the Primary Registry, if any. These should include: The Universal Trial Number (UTN), Identifiers assigned by the sponsor (record sponsor name and sponsor-issued trial number (e.g., protocol number)), Other trial registration numbers), Identifiers issued by funding bodies, collaborative research groups, regulatory authorities, ethics committees / institutional review boards, etc. All secondary identifiers will have 2 elements: an identifier for the issuing authority (eg NCT, ISRCTN, ACTRN) plus a number. There is no limit to the number of secondary identifiers that can be provided.	openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
EudraCT-Nr.	[Nummer]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
Primär-Register-ID	[Nummer]		Trial Identifying Number	Unique ID number assigned by the Primary Registry to this trial.	openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
Primär-Register-ID/Name	[Freitext]		Primary Registry	Name of Primary Registry.	openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
Partner-Register-ID	[Nummer]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Partner-Register-ID/Name	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
BfArM-Nummer	[Nummer]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
PEI-Nummer	[Nummer]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
sonstige sekundäre IDs	[Primär-Register-ID Partner-Register-ID Sponsor-ID sonstige sekundäre IDs]	[Freitext] Name des Sponsors/Name des Registers/Bezeichnung zur ID			openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT
					openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT
Primärer Sponsor (AMG/MPG)/Studienleiter (Non-AMG/non-MPG)	[Adresse]	Anrede, Akademischer Titel, Vorname, Nachname, Einrichtung, Straße und Hausnummer/Postfach, PLZ, Stadt, Land, Telefon mit Vorwahl, FAX mit Vorwahl, E-Mail, URL der Einrichtung	Primary Sponsor	The individual, organization, group or other legal entity which takes responsibility for initiating, managing and/or financing a study. The Primary Sponsor is responsible for ensuring that the trial is properly registered. The Primary Sponsor may or may not be the main funder.	openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0002]	Primary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_+]*\v1)

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
					openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0002]	Primary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_+]*).v1)
Sekundärer Sponsor	[Adresse]	(Falls zutreffend)	Secondary Sponsor(s)	Additional individuals, organizations or other legal persons, if any, that have agreed with the primary sponsor to take on responsibilities of sponsorship. A secondary sponsor may have agreed: to take on all the responsibilities of sponsorship jointly with the primary sponsor; or to form a group with the primary sponsor in which the responsibilities of sponsorship are allocated among the members of the group; or to act as the sponsor's legal representative in relation to some or all of the trial sites; or to take responsibility for the accuracy of trial registration information submitted.	openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0003]	Secondary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_+]*).v1)
					openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0003]	Secondary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_+]*).v1)

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
			Contact for Public Queries	Email address, telephone number, or postal address of the contact who will respond to general queries, including information about current recruitment status.	openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0002]	Contact for public queries	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_]+)*\v1 ODER openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_]+)*\v1)
Kontakt für wissenschaftliche Anfragen	[Adresse]		Contact for Scientific Queries	Email address, telephone number, or postal address, and affiliation of the person to contact for scientific queries about the trial (e.g., principal investigator, medical director employed by the sponsor). For a multi-center study, enter the contact information for the lead Principal Investigator or overall scientific director.	openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0003]	Contact for scientific queries	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_]+)*\v1 ODER openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_]+)*\v1)

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Kontakt für Studienteilnehmer	[Adresse]				openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0004]	Contact for trial participants	ARCHETYP E_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_]+)*\v1 ODER openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_]+)*\v1)
Leiter Klinische Prüfung (multizentrisch)/Hauptprüfer (monozentrisch)	[Adresse]				openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0005]	Principal investigator	ARCHETYP E_SLOT(openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_]+)*\v1)

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Ansprechpartner für Register	[Adresse]				openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0002] (falls es sich um ein Register handelt)	Contact for public queries	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_+]*).v1 ODER openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_+]*).v1)
(federführende) Ethikkommission	[Adresse]				openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0013]	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_+]*).v1)
Sonstige Adressen	[Typ] zum Auswählen und [Adresse]				openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0007]	Further addresses	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_+]*).v1)
Finanzierungsquelle	[Kommerziell (Pharma, med.-techn. Industrie u. a.) Öffentliche Förderinstitutionen, aus Steuermitteln getragene Institutionen (wie	[Adresse]	Source(s) of Monetary or Material Support	Major source(s) of monetary or material support for the trial (e.g., funding agency, foundation, company).	openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0004]	Primary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_+]*).v1)

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
	DFG, BMBF u. a.) Private Gelder (Stiftungen, Studiengesellschaften etc.) Haushaltsmittel, keine fremden Finanzmittel (Budget des Studienleiters/Antragstellers)]							
					openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0004]	Primary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_+]*).v1)
weitere Finanzierungsquellen	[Kommerziell (Pharma, med.-techn. Industrie u. a.) Öffentliche Förderinstitutionen, aus Steuermitteln getragene Institutionen (wie	[Adresse]			openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0005]	Secondary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.organisation(-[a-zA-Z0-9_+]*).v1)
	DFG, BMBF u. a.) Private Gelder (Stiftungen, Studiengesellschaften etc.) Haushaltsmittel, keine fremden							

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
	Finanzmittel (Budget des Studienleiters/Antragstellers)				openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0004]	Primary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.individual_professional(-[a-zA-Z0-9_+]*).v1)
Titel der Studie	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0002]	Title of the study	DV_TEXT
Studienakronym	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0003]	Title abbreviation	DV_TEXT
Internetseite der Studie	[Freitext]	URL?			openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0009]	URL	DV_URI
Allgemeinverständliche Kurzbeschreibung	[Freitext]		Public Title	Title intended for the lay public in easily understood language.	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0006]	Generally understandable abstract	DV_TEXT
Wissenschaftliche Kurzbeschreibung	[Freitext]		Scientific Title	Scientific title of the study as it appears in the protocol submitted for funding and ethical review. This title should contain information on population, intervention, comparator and outcome(s). Include trial acronym if available.	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0005]	Scientific abstract	DV_TEXT
Selbst vergebene Schlagwörter	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0030]	Keyword	DV_TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Untersuchte Krankheit/ Gesundheitsproblem	[ICD 10 MedDRA Freitext Anderes Codesystem] und [Freitext] für Freitext oder Code, falls nicht ICD 10		Health Condition(s) or Problem(s) Studied	Primary health condition(s) or problem(s) studied (e.g., depression, breast cancer, medication error). If the study is conducted in healthy human volunteers belonging to the target population of the intervention (e.g. preventive or screening interventions), enter the particular health condition(s) or problem(s) being prevented.	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0003]	Health condition	DV_TEXT
Weitere Untersuchte Krankheit/ Gesundheitsproblem	[ICD 10 MedDRA Freitext Anderes Codesystem] und [Freitext] für Freitext oder Code, falls nicht ICD 10				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0003]	Health condition	DV_TEXT
Intervention Arm 1	[Freitext]		Intervention(s)	For each arm of the trial record a brief intervention name plus an intervention description.	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0002]	Name	DV_TEXT
				Intervention Name: For drugs use generic name; for other types of interventions provide a brief descriptive name.	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0003]	Description	DV_TEXT
Intervention Arm 2	[Freitext]			* For investigational new drugs that do not yet have a generic name, a chemical name, company code or serial number may be used on a temporary basis. As soon as the generic name has been established, update the associated registered records accordingly.	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0002]	Name	DV_TEXT
				* For non-drug intervention types,	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0003]	Description	DV_TEXT
weitere Interventionsarme	[Freitext]				openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0002]	Name	DV_TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
				<p>provide an intervention name with sufficient detail so that it can be distinguished from other similar interventions. Intervention Description: Must be sufficiently detailed for it to be possible to distinguish between the arms of a study (e.g., comparison of different dosages of drug) and/or among similar interventions (e.g., comparison of multiple implantable cardiac defibrillators). For example, interventions involving drugs may include dosage form, dosage, frequency and duration.</p> <p>* If the intervention is one or more drugs then use the International Non-Proprietary Name for each drug if possible (not brand/trade names). For an unregistered drug, the generic name, chemical name, or company serial number is acceptable.</p> <p>* If the intervention consists of several separate treatments, list them all in one line separated by commas (e.g., 'low-fat diet, exercise').</p> <p>* For controlled trials, the identity of the control arm should be clear. The control intervention(s) is/are the interventions against which the study intervention is evaluated (e.g., placebo, no treatment, active control). If an active control is used, be sure to enter in the name(s) of that intervention, or enter 'placebo' or 'no treatment' as applicable. For each intervention, describe other intervention details as applicable (dose, duration, mode of administration, etc).</p>	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0003]	Description	DV_TEXT
Studientyp	[Interventionell Nicht-interventionell]		Study Type	<p>Study type consists of:</p> <p>* Type of study (interventional or observational)</p>	openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0002]	Type of study	DV_CODED_TEXT

DRKS		WHO / ICTRP		openEHR-Archetypes				
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Studientyp nicht-interventionell	[Beobachtungsstudie Epidemiologische Studie Sonstige]			* Study design including: <ul style="list-style-type: none"> o Method of allocation (randomized/non-randomized) o Masking (is masking used and, if so, who is masked) o Assignment (single arm, parallel, crossover or factorial) o Purpose * Phase (if applicable)	openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0002]	Type of study	DV_CODED_TEXT
Studiendesign Zuteilung	[Einarmige Studie Kontrollierte, nicht randomisierte Studie Kontrollierte, randomisierte Studie Sonstige]				openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0003]	Randomization	DV_BOOLEAN
Verblindung	[Offen Einfach verblindet Doppelt oder mehrfach verblindet]				openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0002]	Blinding	DV_CODED_TEXT
Kontrolle	[Unkontrolliert/einarmig Placebo Aktive Kontrolle Historisch Keine Therapie Andere]				openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0005] und Logik: Information über Studienarme	Comment	DV_TEXT
Studienzweck	[Therapie Prävention Diagnose Prognose Supportive care Screening Gesundheitssystemforschung Grundlagenforschung/physiologische Studie Pharmakogenetik Pharmakogenomik Gesundheitsökonomie Anderer]				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0004]	Purpose	DV_CODED_TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Gruppenzuteilung	[Einarmig Parallelverteilung Crossover Design Faktoriell andere]				openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0004]	Randomization algorithm	DV_PARSABLE
Studienphase	[Nicht zutreffend 0 I-II II Iia Iib II-III III IIIa IIIb IV]				openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0005]	Phase	DV_CODED_TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Primärer Endpunkt	[Freitext]		Primary Outcome(s)	<p>Outcomes are events, variables, or experiences that are measured because it is believed that they may be influenced by the intervention. The Primary Outcome should be the outcome used in sample size calculations, or the main outcome(s) used to determine the effects of the intervention(s). Most trials should have only one primary outcome.</p> <p>For each primary outcome provide:</p> <ul style="list-style-type: none"> * The name of the outcome (do not use abbreviations) * The metric or method of measurement used (be as specific as possible) * The time point(s) of primary interest <p>Example: Outcome Name: Depression Metric/method of measurement: Beck Depression Score Timepoint: 18 weeks following end of treatment</p>	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0026]	Primary outcome	DV_TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Sekundärer Endpunkt	[Freitext]		Key Secondary Outcomes	<p>Secondary outcomes are outcomes which are of secondary interest or that are measured at timepoints of secondary interest. A secondary outcome may involve the same event, variable, or experience as the primary outcome, but measured at timepoints other than those of primary interest.</p> <p>As for primary outcomes, for each secondary outcome provide:</p> <ul style="list-style-type: none"> * The name of the outcome (do not use abbreviations) * The metric or method of measurement used (be as specific as possible) * The time point(s) of interest 	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0027]	Secondary outcome	DV_TEXT
Länder in denen Studienteilnehmer rekrutiert werden	[Auswahlliste]		Countries of Recruitment	The countries from which participants will be, are intended to be, or have been recruited at the time of registration.	openEHR-EHR-ADMIN_ENTRY.study_center_description.v1	/data/items[at0002] Kann aus der lokalisation der teilnehmenden Studienzentren abgeleitet werden.	Address	ARCHETYP E_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_]+)*\).v1)
Publikation der Studie	[Veröffentlichung Abstract zur Studie]	die Publikation wird als [Freitext] erfasst			openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0028]	Publication	DV_TEXT
(Geplantes/tatsächliches Datum) Einschluss des ersten Studienteilnehmers	[Datum]		Date of First Enrollment	Anticipated or actual date of enrollment of the first participant.	openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0003]	First-Patient-In (scheduled)	DV_DATE
Geplant/Tatsächlich	[geplant tatsächlich]				openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0004]	First-Patient-In (actual)	DV_DATE

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Geplante Studienteilnehmeranzahl gesamt	[Zahl]		Target Sample Size	Number of participants that this trial plans to enrol in total.	openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0008]	Target sample size	DV_COUNT (INTEGER)
Tatsächliches Datum des Studienabschlusses (LPLV)	[Datum]				openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0006]	Last-Patient-Out (acutal)	DV_DATE
					openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0006]	Last-Patient-Out (acutal)	DV_DATE
Monozentrisch/Multizentrisch	[monozentrisch multizentrisch]				openEHR-EHR-COMPOSITION.study_center.v1	Logik: Anzahl der Erfassten Zentren		
National/International	[national international]				openEHR-EHR-ADMIN_ENTRY.study_center_description.v1	/data/items[at0002]	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_]+)*\..v1)
Einschlusskriterium Geschlecht	[Männlich Weiblich Beide, männlich und weiblich]				openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)
Einschlusskriterium Mindestalter	[Zahl] und [Zeiteinheit] zum Auswählen	[Jahre Monate Wochen Tage Schwangerschaftswoche Kein Mindestalter]			openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)
Einschlusskriterium Höchstalter	[Zahl] und [Zeiteinheit] zum Auswählen	[Jahre Monate Wochen Tage Schwangerschaftswoche Kein Mindestalter]			openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)
Einschlusskriterien	[Freitext]		Key Inclusion and Exclusion Criteria	Inclusion and exclusion criteria for participant selection, including age and sex. Other	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Ausschlusskriterien	[Freitext]			selection criteria may relate to clinical diagnosis and co-morbid conditions; exclusion criteria are often used to ensure patient safety. If the study is conducted in healthy human volunteers not belonging to the target population (e.g., a preliminary safety study), enter 'healthy human volunteer'.	openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0005]	Exclusion criteria	DV_PARSABLE (AQL statement)
Status der Rekrutierung	[Rekrutierung geplant Rekrutierung läuft Rekrutierung temporär ausgesetzt Rekrutierung abgeschlossen, follow up läuft noch Rekrutierung abgeschlossen, follow up abgeschlossen Rekrutierung permanent eingestellt (nach Rekrutierungsstart) Rekrutierung zurückgezogen (vor Rekrutierungsstart)]		Recruitment Status	[Pending: participants are not yet being recruited or enrolled at any site Recruiting: participants are currently being recruited and enrolled Suspended: there is a temporary halt in recruitment and enrollment Complete: participants are no longer being recruited or enrolled Other]	openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0002]	Current state	DV_CODED _TEXT

DRKS			WHO / ICTRP		openEHR-Archetypes			
Attribute Name	Description	Further description	Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Falls 'Status der Rekrutierung' 'permanent eingestellt' oder 'zurückgezogen' hier bitte den Grund angeben.	[fehlende Finanzierung Auflage einer Behörde Projektleiter hat gewechselt Ungenügende Rekrutierung Kein positive Votum/zustimmende Bewertung Anderer]				openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0003]	Comment	DV_TEXT
Falls Grund des Rekrutierungsstopps 'anderer' bitte hier den Grund angeben.	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0003]	Comment	DV_TEXT
Tatsächliche Gesamtzahl Studienteilnehmer in Deutschland nach abgeschlossener Rekrutierung	[Zahl]				openEHR-EHR-COMPOSITION.study_center.v1.adl	/data/items[at0002] Logik: Zählen aller Studienteilnehmer, die einem Studienzentrum zugeordnet sind, welches in Deutschland lokalisiert ist.	Address	ARCHETYP E_SLOT (openEHR-EHR-CLUSTER. address(-[a-zA-Z0-9_+]*).v1)
Tatsächliche Gesamtzahl Studienteilnehmer aller Zentren nach abgeschlossener Rekrutierung	[Zahl]				openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0011] (Precondition: /data/items[at0002]/value = 'completed')	Enrolled participants	DV_COUNT (INTEGER)
KKS Heidelberg (zusätzlich)					openEHR-Archetypes			
Attribute Name	Description				Archetype	Path	Attribute Name	Data Type
Studiendesign (ergänzende Beschreibung)	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0005]	Comment	DV_TEXT
Studienfragestellung	[Freitext]				openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0002]		

KKS Heidelberg (zusätzlich)		openEHR-Archetypes			
Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Chirurgische Intervention	[Ja Nein]	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0011]	Surgical intervention	DV_BOOLEAN
Pädiatrische Studie	[Ja Nein]	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0012]	Pediatric trial	DV_BOOLEAN
Psychiatrische Studie	[Ja Nein]	openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0015]	Psychiatric trial	DV_BOOLEAN
Protokoll-Version	[Freitext]	openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0004]	Protocol version	DV_TEXT
Studienbeginn (Beginn der Planungen)	[Datum]	openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0002]	Begin of trial	DV_DATE
		openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0002]	Begin of trial	DV_DATE
Geplantes Studienende	[Datum]	openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0005]	Last-Patient-out (scheduled)	DV_DATE
		openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0005]	Last-Patient-out (scheduled)	DV_DATE
Studienende (inklusive aller Tätigkeiten)	[Datum]	openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0007]	End of trial	DV_DATE
		openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0007]	End of trial	DV_DATE
Visite (Name)	[Freitext]	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0002]	Name	DV_TEXT
Visite (Beschreibung)	[Freitext]	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0003]	Description	DV_TEXT
Visite (minimale Anzahl Tage nach Einschluss)	[Zahl]	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0004]	Minimal number of days after enrolment	DV_COUNT (INTEGER)

KKS Heidelberg (zusätzlich)		openEHR-Archetypes			
Attribute Name	Description	Archetype	Path	Attribute Name	Data Type
Visite (maximale Anzahl Tage nach Einschluss)	[Zahl]	openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0005]	Maximal number of days after enrolment	DV_COUNT (INTEGER)

3. Abbildung openEHR-Archetypen auf BRIDG-Modell

Die nachfolgende Tabelle bildet Studien-Metamerkmale, die mit den im Rahmen dieser Arbeit entwickelten *openEHR*-Archetypen dokumentiert werden können (rot), auf im BRIDG-Modell Version 3.0.3 (blau) beschriebene Klassen und Attribute ab.

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier > StudyRegistry.name WHERE StudyRegistry.acronym = 'DRKS' (ST)	Alternatively: PerformedStudyAdministrativeActivity > Documentidentifier > StudyRegistry.name (ST)
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier.identifier WHERE Documentidentifier.typeCode = 'registry identifier' (CD)	Alternatively: PerformedStudyAdministrativeActivity > Documentidentifier.identifier WHERE Documentidentifier (is assigned by) StudyRegistry
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0008]	Date of registration	DV_DATE	PerformedStudyAdministrativeActivity	statusDate	TS.DATETIME	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier > PerformedStudyAdministrativeActivity.statusDate WHERE PerformedStudyAdministrativeActivity.statusCode = 'completed' (CD) AND PerformedStudyAdministrativeActivity > Documentidentifier > StudyRegistry.acronym = 'DRKS' (ST)	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0008]	Date of registration	DV_DATE	PerformedStudyAdministrativeActivity	statusDate	TS. DATETIME	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier > PerformedStudyAdministrativeActivity.statusDate WHERE DocumentIdentifier.typeCode = 'registry identifier' (CD) AND DocumentIdentifier.primaryIndicator = 'FALSE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0015]	Type of financing	CODE_PHRASE				DERIVED (if exist): StudyProtocolVersion > StudyLegalSponsor > HealthcareProvider > BiologicEntity.Person WHERE StudyLegalSponsor.primaryIndicator = 'TRUE' (BL)	Cave: Uncertain criteria!
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0011]	Leading Ethics Committee	DV_BOOLEAN	StudySite	leadIndicator	BL	OversightAuthority.OversightCommittee > StudySite.leadIndicator WHERE StudySite.leadIndicator = 'TRUE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0003]	Request	DV_DATE	Submission	receiptDate	TS. DATETIME	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > SubmissionUnit > Submission.receiptDate	
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0004]	Vote	DV_DATE	Regulatory Assessment	completionDate	TS. DATETIME	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > SubmissionUnit > Submission > RegulatoryAssessment.completionDate	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0005]	Vote	DV_CODED_TEXT	Regulatory Assessment	resultCode	CD	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > SubmissionUnit > Submission > RegulatoryAssessment.resultCode WHERE RegulatoryAssessment.resultCode = 'approved' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0011]	Leading Ethics Committee	DV_BOOLEAN	StudySite	leadIndicator	BL	OversightAuthority.OversightCommittee > StudySite.leadIndicator WHERE StudySite.leadIndicator = 'TRUE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0005]	Vote	DV_CODED_TEXT	Regulatory Assessment	resultCode	CD	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > SubmissionUnit > Submission > RegulatoryAssessment.assessmentCode	
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0013] Name kann aus Adresse ausgelesen werden.	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\address(-[a-zA-Z0-9_]*).v1)	Organization	name	DSET <EN.ON>	StudyProtocolVersion > StudyProtocol > StudyExecution > StudySite > OversightAuthority.OversightCommittee > Organization.name	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0007]	Drug	DV_BOOLEAN	DefinedSubstance Administration	-	-	DERIVED: EXISTS StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedProcedure.DefinedSubstanceAdministration	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0008]	Narcotic Substance	DV_BOOLEAN	Drug	code		DERIVED: EXISTS StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedProcedure.DefinedSubstanceAdministration > Product.Drug WHERE Drug.code = the code of a narcotic substance (CD)	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0009]	Medical Device	DV_BOOLEAN	Device	-	-	DERIVED: EXISTS StudyProtocolVersion > StudyAgent > Product.Device	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0010]	Ionizing radiation	DV_BOOLEAN	DefinedProcedure	nameCode	CD	DERIVED: EXISTS StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedProcedure WHERE DefinedProcedure.nameCode = 'Irradiation' (CD) OR DefinedProcedure.nameCode = 'X-ray' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0013]	Transfusion	DV_BOOLEAN	DefinedProcedure	nameCode	CD	DERIVED: EXISTS StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedProcedure WHERE DefinedProcedure.nameCode = 'Transfusion' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0014]	Genetic engineering	DV_BOOLEAN	Product	classCode	DSET<CD>	DERIVED: EXISTS StudyProtocolVersion > StudyAgent > Product.classCode WHERE Product.classCode = 'Genetically Modified Organism' (DSET<CD>)	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0029]	Legal context	DV_TEXT					No mapping?

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier > StudyRegistry.name WHERE StudyRegistry.name = name of primary registry (ST)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier.identifier WHERE DocumentIdentifier.typeCode = 'registry identifier' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier > StudyRegistry.name WHERE StudyRegistry.acronym = 'EMA' (ST)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier.identifier WHERE DocumentIdentifier.typeCode = 'registry identifier' (CD)	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier.identifier WHERE DocumentIdentifier.typeCode = 'registry identifier' (CD) AND DocumentIdentifier.primaryIndicator = 'TRUE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier > StudyRegistry.name	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier.identifier WHERE DocumentIdentifier.typeCode = 'registry identifier' (CD) AND DocumentIdentifier.primaryIndicator = 'FALSE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005]/items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier > StudyRegistry.name	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier > StudyRegistry.name WHERE StudyRegistry.acronym = 'BfArM' (ST)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier.identifier WHERE Documentidentifier.typeCode = 'registry identifier' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier > StudyRegistry.name WHERE StudyRegistry.acronym = 'PEI'	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier.identifier WHERE Documentidentifier.typeCode = 'registry identifier' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0006]	Register	DV_TEXT	StudyRegistry	name	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > Documentidentifier > StudyRegistry.name	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0005] /items[at0007]	Unique trial ID	DV_TEXT	DocumentIdentifier	identifier	II	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion > Document > DocumentIdentifier.identifier WHERE DocumentIdentifier.typeCode = 'registry identifier' (CD) AND DocumentIdentifier.primaryIndicator = 'FALSE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0002]	Primary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)*.v1)	Organization	name	DSET <EN.ON>	StudyProtocolVersion > StudyLegalSponsor > Organization WHERE StudyLegalSponsor.primaryIndicator = 'TRUE' (BL)	Concrete mapping: openEHR-EHR-CLUSTER.organisation.v1/data/items[at0001] (DV_TEXT) <=> Organization.name (ST)
				Organization	postalAddress	AD		openEHR-EHR-CLUSTER.organisation.v1/data/items[at0008] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Organization.postalAddress (AD)
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0002]	Primary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)*.v1)	Person	name	DSET <EN.ON>	StudyProtocolVersion > StudyLegalSponsor > HealthcareProvider > BiologicEntity.Person WHERE StudyLegalSponsor.primaryIndicator = 'TRUE' (BL)	Concrete mapping: openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0015] -> openEHR-EHR-CLUSTER.person_name.v1/data/items[at0013]/items[at0001] (DV_TEXT) <=> Person.name (ST)
				Person	postalAddress	AD		openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Person.postalAddress (AD)

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0003]	Secondary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)\.v1)	Organization	name	DSET <EN.ON>	StudyProtocolVersion > StudyLegalSponsor > Organization WHERE StudyLegalSponsor.primaryIndicator = 'FALSE' (BL)	Concrete mapping: openEHR-EHR-CLUSTER.organisation.v1/data/items[at0001] (DV_TEXT) <=> Organization.name (ST)
				Organization	postalAddress	AD		openEHR-EHR-CLUSTER.organisation.v1/data/items[at0008] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Organization.postalAddress (AD)
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0003]	Secondary sponsor	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)\.v1)	Person	name	DSET <EN.ON>	StudyProtocolVersion > StudyLegalSponsor > HealthcareProvider > BiologicEntity.Person WHERE StudyLegalSponsor.primaryIndicator = 'FALSE' (BL)	Concrete mapping: openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0015] -> openEHR-EHR-CLUSTER.person_name.v1/data/items[at0013]/items[at0001] (DV_TEXT) <=> Person.name (ST)
				Person	postalAddress	AD		openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[at0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Person.postalAddress (AD)

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0002]	Contact for public queries	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)*\v1 ODER openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)*\v1)	StudyPersonnel	postalAddress	AD	StudyExecution > StudyPersonnel.postalAddresses WHERE StudyPersonnel.roleCode = 'Public Queries' (CD)	openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> StudyPersonnel.postalAddress (AD) - What about organizations?
openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0003]	Contact for scientific queries	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)*\v1 ODER openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)*\v1)	StudyPersonnel	postalAddress	AD	StudyExecution > StudyPersonnel.postalAddresses WHERE StudyPersonnel.roleCode = 'Scientific Queries' (CD)	openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> StudyPersonnel.postalAddress (AD) - What about organizations?

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0004]	Contact for trial participants	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)\.v1 OR openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)\.v1)	StudyPersonnel	postalAddress	AD	StudyExecution > StudyPersonnel.postalAddresses WHERE StudyPersonnel.roleCode = 'Contact for participants' (CD)	openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> StudyPersonnel.postalAddress (AD) - What about organizations?
openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0005]	Principal investigator	ARCHETYPE_SLOT(openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)\.v1)	Person	-	BRIDG Class	StudyExecution > StudyPersonnel.StudyInvestigator > HealthcareProvider > Person WHERE StudyPersonnel.StudyInvestigator.primaryIndicator = 'TRUE' (BL)	Concrete mapping (examples): openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0015] -> openEHR-EHR-CLUSTER.person_name.v1/data/items[at0013]/items[at0001] (DV_TEXT) <=> Person.name (DSET<EN.ON>); openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Person.postalAddress (AD)

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0002] (falls es sich um ein Register handelt)	Contact for public queries	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)*\v1 ODER openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)*\v1)	StudyPersonnel	postalAddress	AD	StudyExecution > StudyPersonnel.postalAddresses WHERE StudyPersonnel.roleCode = 'Scientific Queries' (CD)	openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> StudyPersonnel.postalAddress (AD) - What about organizations?
openEHR-EHR-ADMIN_ENTRY.study_ethics_committee_vote.v1	/data/items[at0013]	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\address(-[a-zA-Z0-9_]+)*\v1)	Organization	postalAddress	AD	StudyExecution > StudySite > OversightAuthority.OversightCommittee > Organization.postalAddress WHERE OversightCommittee.typeCode = 'IEC' (CD) AND StudySite.leadindicator = 'TRUE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_contact_data.v1	/data/items[at0007]	Further addresses	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\address(-[a-zA-Z0-9_]+)*\v1)	Organization	postalAddress	AD	Organization.postalAddress	
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0004]	Primary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)*\v1)	Organization	name	DSET <EN.ON>	StudyExecution > StudyResource > Resource.Funding > ResourceProvider > Organization WHERE Resource.typeCode = 'Financial Sponsor' (CD)	Concrete mapping: openEHR-EHR-CLUSTER.organisation.v1/data/items[at0001] (DV_TEXT) <=> Organization.name (ST)

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
				Organization	postalAddress	AD		openEHR-EHR-CLUSTER.organisation.v1/data/items[at0008] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Organization.postalAddress (AD)
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0004]	Primary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\individual_professional(-[a-zA-Z0-9_]+)*\v1)	Person	name	DSET <EN.ON>	StudyExecution > StudyResource > Resource.Funding > ResourceProvider > Person WHERE Resource.typeCode = 'Financial Sponsor' (CD)	Concrete mapping: openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0015] -> openEHR-EHR-CLUSTER.person_name.v1/data/items[at0013]/items[at0001] (DV_TEXT) <=> Person.name (ST)
				Person	postalAddress	AD		openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[at0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Person.postalAddress (AD)
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0005]	Secondary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\organisation(-[a-zA-Z0-9_]+)*\v1)	Organization	name	DSET <EN.ON>	StudyExecution > StudyResource > Resource.Funding > ResourceProvider > Organization WHERE Resource.typeCode = 'Financial Sponsor' (CD)	Concrete mapping: openEHR-EHR-CLUSTER.organisation.v1/data/items[at0001] (DV_TEXT) <=> Organization.name (ST)
				Organization	postalAddress	AD	No differentiating between primary and secondary source!	openEHR-EHR-CLUSTER.organisation.v1/data/items[at0008] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Organization.postalAddress (AD)

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_sponsor.v1	/data/items[at0004]	Primary funding source	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\individual_professional(-[azAZ0-9_+]*).v1)	Person	name	DSET <EN.ON>	StudyExecution > StudyResource > Resource.Funding > ResourceProvider > Person WHERE Resource.typeCode = 'Financial Sponsor' (CD)	Concrete mapping: openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0015] -> openEHR-EHR-CLUSTER.person_name.v1/data/items[at0013]/items[at0001] (DV_TEXT) <=> Person.name (ST)
				Person	postalAddress	AD	No differentiating between primary and secondary source!	openEHR-EHR-CLUSTER.individual_professional.v1/data/items[at0003]/items[at0019] -> openEHR-EHR-CLUSTER.address.v1 (ARCHETYPE) <=> Person.postalAddress (AD)
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0002]	Title of the study	DV_TEXT	StudyProtocolDocumentVersion	officialTitle	ST	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion.officialTitle	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0003]	Title abbreviation	DV_TEXT	StudyProtocolVersion	acronym	ST	StudyProtocolVersion.acronym	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0009]	URL	DV_URI	StudyProtocolDocumentVersion	uniform RessourceLocator	URL	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion.uniform RessourceLocator	Cave: StudyProtocolDocumentVersion.uniformRessourceLocator is the location of a certain document and not a general information source. However, a general URL for a trial might be derivable.
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0006]	Generally understandable abstract	DV_TEXT	StudyProtocolDocumentVersion	publicTitle	ST	StudyProtocolVersion > StudyProtocolDocumentVersion.publicDescription	WHO attribute fits better to StudyProtocolDocumentVersion.publicTitle
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0005]	Scientific abstract	DV_TEXT	StudyProtocolDocumentVersion	scientific Description	ST	StudyProtocolVersion > StudyProtocolDocumentVersion.scientificDescription	WHO attribute fits better to StudyProtocolDocumentVersion.officialTitle
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0030]	Keyword	DV_TEXT	StudyProtocolDocumentVersion	keywordText	DSET<ST>	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion.keywordText	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0003]	Health condition	DV_TEXT	StudyCondition	code	CD	StudyProtocolVersion > StudyCondition.code	See also: StudyProtocolVersion.populationDescription AND DefinedEligibilityCriterion.requireResponse
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0003]	Health condition	DV_TEXT	StudyCondition	code	CD	StudyProtocolVersion > StudyCondition.code	See also: StudyProtocolVersion.populationDescription AND DefinedEligibilityCriterion.requireResponse
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0002]	Name	DV_TEXT	Arm	name	ST	StudyProtocolDocumentVersion > Arm.name	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0003]	Description	DV_TEXT	Arm	description	ST	StudyProtocolDocumentVersion > Arm.description	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0002]	Name	DV_TEXT	Arm	name	ST	StudyProtocolDocumentVersion > Arm.name	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0003]	Description	DV_TEXT	Arm	description	ST	StudyProtocolDocumentVersion > Arm.description	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0002]	Name	DV_TEXT	Arm	name	ST	StudyProtocolDocumentVersion > Arm.name	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0003]	Description	DV_TEXT	Arm	description	ST	StudyProtocolDocumentVersion > Arm.description	
openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0002]	Type of study	DV_CODED_TEXT	StudyProtocolVersion	Type of Subclass		DERIVED: StudyProtocolVersion IS A (InterventionalStudyProtocolVersion OR ObservationalStudyProtocolVersion OR ExpandedAccessProtocolVersion)	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0002]	Type of study	DV_CODED_TEXT	StudyProtocolVersion	Type of Subclass		DERIVED: StudyProtocolVersion IS A (InterventionalStudyProtocolVersion OR ObservationalStudyProtocolVersion OR ExpandedAccessProtocolVersion)	
openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0003]	Randomization	DV_BOOLEAN	InterventionalStudyProtocolVersion	allocationCode	CD	StudyProtocolVersion.InterventionalStudyProtocolVersion.allocationCode	
openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0002]	Blinding	DV_CODED_TEXT	InterventionalStudyProtocolVersion	blindingSchemaCode	CD	StudyProtocolVersion.InterventionalStudyProtocolVersion.blindingSchemaCode	
openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0005]	Comment und Logik: Information über Studienarme	DV_TEXT	InterventionalStudyProtocolVersion	controlTypeCode	CD	StudyProtocolVersion.InterventionalStudyProtocolVersion.controlTypeCode	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0004]	Purpose	DV_CODED_TEXT	StudyProtocolVersion	primaryPurposeCode	CD	StudyProtocolVersion.primaryPurposeCode	
openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0004]	Randomization algorithm	DV_PARSABLE	StudyProtocolVersion	designConfigurationCode	CD	StudyProtocolVersion.designConfigurationCode	
openEHR-EHR-ADMIN_ENTRY.study_type.v1	/data/items[at0005]	Phase	DV_CODED_TEXT	StudyProtocolVersion	phaseCode	CD	StudyProtocolVersion.phaseCode	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0026]	Primary outcome	DV_TEXT	StudyOutcomeMeasure	name	ST	StudyProtocolVersion > StudyObjective > StudyOutcomeMeasure.name WHERE StudyOutcomeMeasure.primaryIndicator = 'TRUE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0027]	Secondary outcome	DV_TEXT	StudyOutcomeMeasure	name	ST	StudyProtocolVersion > StudyObjective > StudyOutcomeMeasure.name WHERE StudyOutcomeMeasure.primaryIndicator = 'FALSE' (BL)	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_center_description.v1	/data/items[at0002] Kann aus der lokalisation der teilnehmenden Studienzentren abgeleitet werden.	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_]+)*).v1)	StudyProtocolVersion	participatingCountryCode	DSET<CD>	StudyProtocolVersion. participatingCountryCode	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0028]	Publication	DV_TEXT	ReferenceToStudyResults	citationDescription	ST	StudyExecution > ReferenceToStudyResults. citationDescription	
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0003]	First-Patient-In (scheduled)	DV_DATE	ScheduledActivity	dateRange	IVL<TS.DATETIME>	StudyProtocolVersion > StudyActivity > PlannedActivity > ScheduledActivity.dateRange. low WHERE Activity.ScheduledActivity.comment = 'Enroll first subject' (ST)	In the archetype-based representation this item is related to one study site => all sites have to be checked to find the lowest date.
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0004]	First-Patient-In (actual)	DV_DATE	PerformedAdministrativeActivity	dateRange	IVL<TS.DATETIME>	StudyProtocolVersion > PerformedActivity.PerformedAdministrativeActivity.dateRange.low WHERE Activity.PerformedActivity.PerformedAdministrativeActivity.comment = 'Enroll first subject' (ST)	In the archetype-based representation this item is related to one study site => all sites have to be checked to find the lowest date.
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0008]	Target sample size	DV_COUNT (INTEGER)	StudyProtocolVersion	targetAccrualNumberRange	URG<INT.NONNEG>	StudyProtocolVersion. targetAccrualNumberRange. low	In the archetype-based representation the target sample size is intended to be the minimum sample size. CAVE: In the archetype-based representation this item may be related to a single study site.

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0006]	Last-Patient-Out (acutal)	DV_DATE	StudySite	statusDate	TS. DATETIME	StudyExecution > StudySite.statusDate WHERE StudySite.statusCode = 'Closed to Accrual and Intervention' (CD)	Actual LPLV has to be derived from LPLV of the single sites. NOTE from BRIDGE documentation: The status of a study site may overlap with the study site accrual status and overall status of the study. This overlap needs to be clearly differentiated. (See tracker issue 29398).
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0006]	Last-Patient-Out (acutal)	DV_DATE	StudyOverallStatus	Date	TS. DATETIME	StudyExecution > StudyOverallStatus.date WHERE StudyOverallStatus.code = 'Closed to Accrual and Intervention' (CD)	LPLV using StudyOverallState.
openEHR-EHR-COMPOSITION.study_center.v1	Logik: Anzahl der Erfassten Zentren			StudyExecution		Association with StudySite	DERIVED: Number of study sites.	
openEHR-EHR-ADMIN_ENTRY.study_center_description.v1	/data/items[at0002]	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER\address(-[a-zA-Z0-9_]+)*\v1)	HealthcareFacility	postalAddress	AD	DERIVED (check location of all sites): StudyExecution > StudySite > HealthCareFacility.postalAddress	
openEHR-EHR-ADMIN_ENTRY.trial_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)	DefinedInclusionCriterion	Required Response	ANY	StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedObservation.DefinedEligibilityCriterion.DefinedInclusionCriterion.requiredResponse WHERE DefinedInclusionCriterion.nameCode = 'Inclusion criteria gender' (CD)	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)	DefinedInclusionCriterion	Required Response	ANY	StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedObservation.DefinedEligibilityCriterion.DefinedInclusionCriterion.requiredResponse WHERE DefinedInclusionCriterion.nameCode = 'Inclusion criteria minimum age' (CD)	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)	DefinedInclusionCriterion	Required Response	ANY	StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedObservation.DefinedEligibilityCriterion.DefinedInclusionCriterion.requiredResponse WHERE DefinedInclusionCriterion.nameCode = 'Inclusion criteria limiting age' (CD)	
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0004]	Inclusion criteria	DV_PARSABLE (AQL statement)	StudyProtocolVersion	Population Description	ST	DERIVED (text): StudyProtocolVersion.populationDescription	Alternatively: StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedObservation.DefinedEligibilityCriterion.DefinedInclusionCriterion.requiredResponse
openEHR-EHR-ADMIN_ENTRY.terminal_arm_description.v1	/data/items[at0005]	Exclusion criteria	DV_PARSABLE (AQL statement)	StudyProtocolVersion	Population Description	ST	DERIVED (text): StudyProtocolVersion.populationDescription	Alternatively: StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedObservation.DefinedEligibilityCriterion.DefinedExclusionCriterion.requiredResponse
openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0002]	Current state	DV_CODED_TEXT	StudyRecruitmentStatus	Code	CD	StudyExecution > StudyRecruitmentStatus.code	NOTE from BRIDGE documentation: The recruitment status of a study may overlap with the study site status and accrual status. This overlap needs to be clearly differentiated. (See tracker issue 29398).
openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0003]	Comment	DV_TEXT	StudyOverallStatus	studyStoppedReasonCode	CD	StudyExecution > StudyOverallStatus.studyStoppedReasonCode	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0003]	Comment	DV_TEXT	StudyOverallStatus	studyStoppedReasonCode	CD	StudyExecution > StudyOverallStatus.studyStoppedReasonCode	
openEHR-EHR-COMPOSITION.study_center.v1.adl	/data/items[at0002] Logik: Zählen aller Studienteilnehmer, die einem Studienzentrum zugeordnet sind, welches in Deutschland lokalisiert ist.	Address	ARCHETYPE_SLOT (openEHR-EHR-CLUSTER.address(-[a-zA-Z0-9_]+)*).v1)	StudySubject	-	BRIDG Class	DERIVED (number of instances): StudyProtocolVersion > StudySiteProtocolVersionRelationship > StudySubjectProtocolVersionRelationship > StudySubject WHERE StudySiteProtocolVersionRelationship > StudySite > Organization.postalAddress.country = "Germany"	Precondition: StudyExecution > StudyRecruitmentStatus.code = 'completed'
openEHR-EHR-ADMIN_ENTRY.study_state.v1	/data/items[at0011] (Precondition: /data/items[at0002] /value = 'completed')	Enrolled participants	DV_COUNT (INTEGER)	StudySubject	-	BRIDG Class	DERIVED (number of instances): StudyProtocolVersion > StudySiteProtocolVersionRelationship > StudySubjectProtocolVersionRelationship > StudySubject	Precondition: StudyExecution > StudyRecruitmentStatus.code = 'completed'
openEHR-EHR-ADMIN_ENTRY.study_design.v1	/data/items[at0005]	Comment	DV_TEXT					No direct mapping?
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0002]			StudyObjective	description	ST	StudyProtocolVersion > StudyObjective.description	
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0011]	Surgical intervention	DV_BOOLEAN	DefinedProcedure	nameCode	CD	DERIVED (if exists): StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedProcedure.nameCode WHERE DefinedProcedure.nameCode (CD) = code of a surgical intervention	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0012]	Pediatric trial	DV_BOOLEAN	DefinedInclusionCriterion	requiredResponse	ANY	DERIVED (if exists): StudyProtocolVersion > StudyActivity > DefinedActivity.DefinedObservation.DefinedEligibilityCriterion.DefinedInclusionCriterion.requiredResponse WHERE DefinedInclusionCriterion.nameCode = 'Inclusion criteria limiting age' (CD) AND DefinedInclusionCriterion.requiredResponse = '18'	Alternatively: DERIVED (text): StudyProtocolVersion.populationDescription
openEHR-EHR-ADMIN_ENTRY.study_content.v1	/data/items[at0015]	Psychiatric trial	DV_BOOLEAN	StudyCondition	code	CD	DERIVED (if exists): StudyProtocolVersion > StudyCondition.code WHERE StudyCondition.code = code of a psychiatric disease (CD)	
openEHR-EHR-ADMIN_ENTRY.study_identification.v1	/data/items[at0004]	Protocol version	DV_TEXT	StudyProtocolDocumentVersion	numberText	ST_SIMPLE	StudyProtocolVersion > DocumentVersion.StudyProtocolDocumentVersion.numberText	
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0002]	Begin of trial	DV_DATE	StudySite	statusDate	TS_DATETIME	DERIVED (from all sites): StudyExecution > StudySite.statusDate WHERE StudySite.statusCode = 'pending' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0002]	Begin of trial	DV_DATE	StudyOverallStatus	date	TS_DATETIME	StudyExecution > StudyOverallStatus.date WHERE StudyOverallStatus.code = 'Started' AND StudyOverallStatus.anticipateIndicator = 'FALSE' (BL)	
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0005]	Last-Patient-out (scheduled)	DV_DATE	StudySite	plannedDuration	PQ.TIME	DERIVED (from all sites): StudyExecution > StudySite.plannedDuration. end	

openEHR-Archetypes				BRIDG 3.0.3				
Archetype	Path	Attribute Name	Data Type	Class Name	Attribute Name	Data Type	Mapping Path	Comments
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0005]	Last-Patient-out (scheduled)	DV_DATE	StudyOverallStatus	date	TS.DATETIME	StudyExecution > StudyOverallStatus.date WHERE StudyOverallStatus.code = 'Closed to Accrual and Intervention' AND StudyOverallStatus.anticipateIndicator = 'TRUE' (BL)	NOTE from BRIDGE documentation: The actual overall status of a study may be derived if it is possible to roll-up the site-specific status.
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0007]	End of trial	DV_DATE	StudySite	statusDate	TS.DATETIME	DERIVED (from all sites): StudyExecution > StudySite.statusDate WHERE StudySite.statusCode = 'Administratively complete' (CD)	
openEHR-EHR-ADMIN_ENTRY.study_length.v1	/data/items[at0007]	End of trial	DV_DATE	StudyOverallStatus	code	TS.DATETIME	StudyExecution > StudyOverallStatus.date WHERE StudyOverallStatus.code = 'Administratively complete' AND StudyOverallStatus.anticipateIndicator = 'FALSE' (BL)	NOTE from BRIDGE documentation: The actual overall status of a study may be derived if it is possible to roll-up the site-specific status.
openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0002]	Name	DV_TEXT	PlannedSubject ActivityGroup	name	ST	StudyProtocolVersion > StudyActivity > PlannedActivity.PlannedSubjectActivityGroup.name	
openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0003]	Description	DV_TEXT	PlannedSubject ActivityGroup	description	ST	StudyProtocolVersion > StudyActivity > PlannedActivity.PlannedSubjectActivityGroup.description	
openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0004]	Minimal number of days after enrolment	DV_COUNT (INTEGER)	PlannedSubject ActivityGroup	studyDayRange	IVL<INT>	StudyProtocolVersion > StudyActivity > PlannedActivity.PlannedSubjectActivityGroup.studyDayRange	
openEHR-EHR-ADMIN_ENTRY.trial_visit_description.v1	/data/items[at0005]	Maximal number of days after enrolment	DV_COUNT (INTEGER)	PlannedSubject ActivityGroup	studyDayRange	IVL<INT>	StudyProtocolVersion > StudyActivity > PlannedActivity.PlannedSubjectActivityGroup.studyDayRange	

4. Beispiel: Semantische Annotation von ODM-Elementen mittels Archetypen

Das folgende Beispiel veranschaulicht wie ODM-Elemente durch Archetypen semantisch annotiert werden können. Die ursprüngliche ODM-Datei stammt von der Internetseite www.eclinicalopinion.blogspot.com (letzter Zugriff 22.12.2011). Die für dieses Beispiel eingefügten Referenzen auf Archetypen sind durch Fettdruck und blaue Einfärbung hervorgehoben. Neben Referenzen auf bestehende sowie im Rahmen dieser Arbeit entwickelte Archetypen, wird mit dem Alias-Mechanismus auch auf Systemvariablen des jeweiligen Aktensystems verwiesen (Context ‚SYSTEM‘ / Name = *Systemparameter*). Zum Beispiel:

```
<Alias Context="SYSTEM" Name="Trial_Subject_ID"/>
```

An wenigen Stellen waren Referenzen notwendig auf (noch) nicht vorhandene Archetypen bzw. deren Attribute. Diese Referenzen wurden fiktiv angegeben. Um zu kennzeichnen, dass es sich dabei um Referenzen auf zu diesem Zeitpunkt nicht vorhandene Archetypen handelt, wurde für diese jeweils die Versionsnummer 99 angegeben. Zum Beispiel:

```
<Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-DEMOGRAPHIC-
CLUSTER.person_age.v99]/items[0002]/value"/>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Demonstration of openEHR annotations.
The following file has been used as basis for this demonstration:
-----
SAMP-101 (Samplestra) Example ODM File.
Copyright eClinicalOpinion
www.eclinicalopinion.blogspot.com
v1.2
- Updated to change encoding to UTF-8 and to remove schema location
  references per suggestion of Jozef Aerts XML4Pharma.
v1.1
- Updated after comments from Jozef Aerts regarding required ordering
  of elements in the ODM file (caused by definition of schema)
v1.0
- First version.
-->
<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3"
  xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  ODMVersion="1.3" FileOID="000-00-0000" FileType="Snapshot"
  CreationDateTime="2009-02-20T06:56:00">
  <Study OID="ST.001">
    <GlobalVariables>
      <StudyName>SAMP-101</StudyName>
      <StudyDescription>A sample study to determine the usefulness of the CDISC ODM
1.3 Snapshot Metadata Standard</StudyDescription>
      <ProtocolName>SAMPLESTRA (SAMP-101)</ProtocolName>
    </GlobalVariables>
    <BasicDefinitions>
      <MeasurementUnit Name="Centimeters" OID="MU.CM">
        <Symbol>
          <TranslatedText xml:lang="en">
            cm
          </TranslatedText>
        </Symbol>
      </MeasurementUnit>
      <MeasurementUnit Name="Kilograms" OID="MU.KG">
        <Symbol>
          <TranslatedText xml:lang="en">
            kg
          </TranslatedText>
```

```

    </Symbol>
  </MeasurementUnit>
  <MeasurementUnit Name="Beats/Minute" OID="MU.BPM">
    <Symbol>
      <TranslatedText xml:lang="en">
        bpm
      </TranslatedText>
    </Symbol>
  </MeasurementUnit>
  <MeasurementUnit Name="Degrees C" OID="MU.DEG_C">
    <Symbol>
      <TranslatedText xml:lang="en">
        °C
      </TranslatedText>
    </Symbol>
  </MeasurementUnit>
  <MeasurementUnit Name="Millimeters of Mercury" OID="MU.MMHG">
    <Symbol>
      <TranslatedText xml:lang="en">
        mmHG
      </TranslatedText>
    </Symbol>
  </MeasurementUnit>
</BasicDefinitions>
<MetaDataVersion OID="MD.001" Name="Study Version 1">
  <Protocol>
    <StudyEventRef OrderNumber="1" StudyEventOID="SE.INIT" Mandatory="Yes"/>
    <StudyEventRef OrderNumber="2" StudyEventOID="SE.VISIT2" Mandatory="Yes"
CollectionExceptionConditionOID="CD.NOT_CONT_LAST_VISIT"/>
    <StudyEventRef OrderNumber="3" StudyEventOID="SE.VISIT3" Mandatory="Yes"
CollectionExceptionConditionOID="CD.NOT_CONT_LAST_VISIT"/>
    <StudyEventRef OrderNumber="4" StudyEventOID="SE.CM" Mandatory="Yes"/>
    <StudyEventRef OrderNumber="5" StudyEventOID="SE.AE" Mandatory="Yes"/>
  </Protocol>
  <StudyEventDef OID="SE.INIT" Name="Initiation" Repeating="No" Type="Scheduled"
Category="Screening">
    <FormRef OrderNumber="1" FormOID="FD.VISITINFO" Mandatory="Yes"/>
    <FormRef OrderNumber="2" FormOID="FD.DEMOG" Mandatory="Yes"/>
    <FormRef OrderNumber="3" FormOID="FD.PREG" Mandatory="Yes"
CollectionExceptionConditionOID="CD.IS_NOT_FERTILE_FEMALE"/>
    <FormRef OrderNumber="4" FormOID="FD.VITALS" Mandatory="Yes"/>
  </StudyEventDef>
  <StudyEventDef OID="SE.VISIT2" Name="Visit 2" Repeating="No" Type="Scheduled"
Category="Treatment">
    <FormRef OrderNumber="1" FormOID="FD.VISITINFO" Mandatory="Yes"/>
    <FormRef OrderNumber="2" FormOID="FD.VITALS" Mandatory="Yes"/>
  </StudyEventDef>
  <StudyEventDef OID="SE.VISIT3" Name="Visit 3" Repeating="No" Type="Scheduled"
Category="Treatment">
    <FormRef OrderNumber="1" FormOID="FD.VISITINFO" Mandatory="Yes"/>
    <FormRef OrderNumber="2" FormOID="FD.VITALS" Mandatory="Yes"/>
  </StudyEventDef>
  <StudyEventDef OID="SE.CM" Name="Concomitant Medications" Repeating="No"
Type="Common" Category="Treatment">
    <Description>
      <TranslatedText xml:lang="en">All Concomitant Medications are collected in
this event.</TranslatedText>
    </Description>
    <FormRef OrderNumber="1" FormOID="FD.CM" Mandatory="Yes"/>
  </StudyEventDef>
  <StudyEventDef OID="SE.AE" Name="Adverse Events" Repeating="No" Type="Common"
Category="Treatment">
    <Description>
      <TranslatedText xml:lang="en">All Adverse Reactions are collected in this
event.</TranslatedText>
    </Description>
    <FormRef OrderNumber="1" FormOID="FD.AE" Mandatory="Yes"/>
  </StudyEventDef>
<!--

  Visit Info Form

-->
<FormDef OID="FD.VISITINFO" Name="Visit Information" Repeating="No">
  <Description>
    <TranslatedText xml:lang="en">The Visit Info form.</TranslatedText>
  </Description>
  <ItemGroupRef ItemGroupOID="IG.VISITINFO" Mandatory="Yes"/>
</FormDef>
<FormDef OID="FD.CM" Name="Concomitant Medications" Repeating="No">
  <Description>
    <TranslatedText xml:lang="en">
      Concomitant Medications Form
    </TranslatedText>
  </Description>
  <ItemGroupRef OrderNumber="1" ItemGroupOID="IG.ANYCONMED" Mandatory="Yes"/>

```

```

    <ItemGroupRef OrderNumber="2" ItemGroupOID="IG.CONMED" Mandatory="Yes"
CollectionExceptionConditionOID="CD.NO_CONMEDS"/>
  </FormDef>
  <FormDef OID="FD.DEMOG" Name="Demography" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Subject Demography Form</TranslatedText>
    </Description>
    <ItemGroupRef ItemGroupOID="IG.DEMOG" Mandatory="Yes"/>
  </FormDef>
  <FormDef OID="FD.PREG" Name="Pregnancy Test" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Pregnancy Test Form</TranslatedText>
    </Description>
    <ItemGroupRef ItemGroupOID="IG.PREG" Mandatory="Yes"/>
  </FormDef>
  <FormDef OID="FD.VITALS" Name="Vital Signs" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Vital Signs Form</TranslatedText>
    </Description>
    <ItemGroupRef OrderNumber="1" ItemGroupOID="IG.HEIGHTWEIGHT"
Mandatory="Yes"/>
    <ItemGroupRef OrderNumber="2" ItemGroupOID="IG.BLOODVITALS" Mandatory="Yes"/>
    <ItemGroupRef OrderNumber="3" ItemGroupOID="IG.BODYTEMP" Mandatory="Yes"/>
  </FormDef>
  <FormDef OID="FD.AE" Name="Adverse Events" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">
        Adverse Events Form
      </TranslatedText>
    </Description>
    <ItemGroupRef OrderNumber="1" ItemGroupOID="IG.ANYAE" Mandatory="Yes"/>
    <ItemGroupRef OrderNumber="2" ItemGroupOID="IG.AE" Mandatory="Yes"
CollectionExceptionConditionOID="CD.NO_AES"/>
  </FormDef>
  <ItemGroupDef OID="IG.VISITINFO" Name="Visit Information" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Visit Information</TranslatedText>
    </Description>
    <ItemRef OrderNumber="1" ItemOID="ID.CNTR" Mandatory="Yes"/>
    <ItemRef OrderNumber="2" ItemOID="ID.SUBID" Mandatory="Yes"/>
    <ItemRef OrderNumber="3" ItemOID="ID.VDAT" Mandatory="Yes"/>
    <ItemRef OrderNumber="4" ItemOID="ID.FEEL" Mandatory="Yes"/>
    <ItemRef OrderNumber="5" ItemOID="ID.CONT" Mandatory="Yes"/>
  </ItemGroupDef>
  <ItemGroupDef OID="IG.DEMOG" Name="Demography" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Demography</TranslatedText>
    </Description>
    <ItemRef OrderNumber="1" ItemOID="ID.DOB" Mandatory="Yes"/>
    <!-- Age is calculated by Method MD.AGE -->
    <ItemRef OrderNumber="2" ItemOID="ID.AGE" Mandatory="Yes"
MethodOID="MD.AGE"/>
    <ItemRef OrderNumber="3" ItemOID="ID.SEX" Mandatory="Yes"/>
    <ItemRef OrderNumber="4" ItemOID="ID.FERTILE" Mandatory="Yes"/>
CollectionExceptionConditionOID="CD.IS_NOT_MALE"/>
    <ItemRef OrderNumber="6" ItemOID="ID.RACE" Mandatory="Yes"/>
    <ItemRef OrderNumber="7" ItemOID="ID.RACEOTH" Mandatory="Yes"
CollectionExceptionConditionOID="CD.RACE_NOT_OTH"/>
  </ItemGroupDef>
  <ItemGroupDef OID="IG.PREG" Name="Pregnancy Test" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Pregnancy Test</TranslatedText>
    </Description>
    <ItemRef OrderNumber="1" ItemOID="ID.TESTPERFORMED" Mandatory="Yes"/>
    <ItemRef OrderNumber="2" ItemOID="ID.TESTRESULT" Mandatory="Yes"
CollectionExceptionConditionOID="CD.PREG_NOT_TESTED"/>
  </ItemGroupDef>
  <ItemGroupDef OID="IG.BLOODVITALS" Name="vitals: Blood" Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Blood-related vital signs</TranslatedText>
    </Description>
    <ItemRef OrderNumber="1" ItemOID="ID.PULSE" Mandatory="Yes"/>
    <ItemRef OrderNumber="2" ItemOID="ID.SYS" Mandatory="Yes"/>
    <ItemRef OrderNumber="3" ItemOID="ID.DIA" Mandatory="Yes"/>
  </ItemGroupDef>
  <ItemGroupDef OID="IG.HEIGHTWEIGHT" Name="Vitals: Height & weight"
Repeating="No">
    <Description>
      <TranslatedText xml:lang="en">Vital Signs: Height &
weight</TranslatedText>
    </Description>
    <ItemRef OrderNumber="1" ItemOID="ID.HEIGHT" Mandatory="Yes"/>
    <ItemRef OrderNumber="2" ItemOID="ID.WEIGHT" Mandatory="Yes"/>
    <ItemRef OrderNumber="3" ItemOID="ID.BMI" Mandatory="Yes"
MethodOID="MD.BMI"/>
  </ItemGroupDef>
  <ItemGroupDef OID="IG.BODYTEMP" Name="Vitals: Body Temperature" Repeating="No">

```

```

    <Description>
      <TranslatedText xml:lang="en">vital signs: Body
temperature</TranslatedText>
    </Description>
  </ItemGroupDef>
</ItemGroupDef>
<ItemGroupDef OID="IG.ANYCONMED" Name="Any Conmeds" Repeating="No">
  <Description>
    <TranslatedText xml:lang="en">
      Concomitant Medication Question
    </TranslatedText>
  </Description>
  <ItemRef OrderNumber="1" ItemOID="ID.ANYCONMED" Mandatory="Yes"/>
</ItemGroupDef>
<ItemGroupDef OID="IG.CONMED" Name="Concomitant Medications" Repeating="Yes">
  <Description>
    <TranslatedText xml:lang="en">
      Concomitant Medications detail
    </TranslatedText>
  </Description>
  <ItemRef OrderNumber="1" ItemOID="ID.DRUGNUM" Mandatory="Yes"
MethodOID="MD.ASSIGN_DRUGNUM" KeySequence="1"/>
  <ItemRef OrderNumber="2" ItemOID="ID.DRUGNAME" Mandatory="Yes"/>
  <ItemRef OrderNumber="3" ItemOID="ID.DOSE" Mandatory="Yes"/>
  <ItemRef OrderNumber="4" ItemOID="ID.DOSEUNITS" Mandatory="Yes"/>
  <ItemRef OrderNumber="5" ItemOID="ID.FREQ" Mandatory="Yes"/>
  <ItemRef OrderNumber="6" ItemOID="ID.STARTDAT" Mandatory="Yes"/>
  <ItemRef OrderNumber="7" ItemOID="ID.ENDDAT" Mandatory="Yes"/>
  <ItemRef OrderNumber="8" ItemOID="ID.DRUGCONT" Mandatory="Yes"/>
  <ItemRef OrderNumber="9" ItemOID="ID.DIAGNOSIS" Mandatory="Yes"/>
</ItemGroupDef>
<ItemGroupDef OID="IG.AE" Name="Adverse Events" Repeating="Yes">
  <Description>
    <TranslatedText xml:lang="en">
      Adverse Events detail
    </TranslatedText>
  </Description>
  <ItemRef OrderNumber="1" ItemOID="ID.AENUM" Mandatory="Yes"
MethodOID="MD.ASSIGN_AENUM" KeySequence="1"/>
  <ItemRef OrderNumber="2" ItemOID="ID.AEDESC" Mandatory="Yes"/>
  <ItemRef OrderNumber="3" ItemOID="ID.AESTARTDAT" Mandatory="Yes"/>
  <ItemRef OrderNumber="4" ItemOID="ID.AEENDDAT" Mandatory="Yes"/>
  <ItemRef OrderNumber="5" ItemOID="ID.AECONT" Mandatory="Yes"/>
  <ItemRef OrderNumber="6" ItemOID="ID.AESEVERITY" Mandatory="Yes"/>
  <ItemRef OrderNumber="7" ItemOID="ID.AESERIOUS" Mandatory="Yes"/>
  <ItemRef OrderNumber="8" ItemOID="ID.AERELATION" Mandatory="Yes"/>
  <ItemRef OrderNumber="9" ItemOID="ID.AEACTION" Mandatory="Yes"/>
</ItemGroupDef>
<ItemGroupDef OID="IG.ANYAE" Name="Any Adverse Events" Repeating="No">
  <Description>
    <TranslatedText xml:lang="en">
      Any Adverse Events Question
    </TranslatedText>
  </Description>
  <ItemRef OrderNumber="1" ItemOID="ID.ANYAE" Mandatory="Yes"/>
</ItemGroupDef>
<ItemDef OID="ID.CNTR" Name="Center Number" DataType="text" Length="4">
  <Description>
    <TranslatedText xml:lang="en">
      Center Number
    </TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">
      CENTER
    </TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_site.v1/context/other/items[0004]/value"/>
  <!-- Comparator is not used for FormalExpressions,
bug in ODM 1.3 means it is required
-->
  <RangeCheck SoftHard="Hard" Comparator="EQ">
    <FormalExpression Context="ISVL"><![CDATA[
//Center number must be all digits.
IsNumeric(:this)
]]></FormalExpression>
    <ErrorMessage>
      <TranslatedText xml:lang="en">
        Center number must be four digits.
      </TranslatedText>
    </ErrorMessage>
  </RangeCheck>
</ItemDef>
<ItemDef OID="ID.SUBID" Name="Subject Number" DataType="text" Length="5">
  <Description>
    <TranslatedText xml:lang="en">

```

```

        Subject Number
      </TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">
        Subject Number
      </TranslatedText>
    </Question>
    <Alias Context="SYSTEM" value="Trial_Subject_ID"/>
  </ItemDef>
  <ItemDef OID="ID.VDAT" Name="Visit Date" DataType="date">
    <Description>
      <TranslatedText xml:lang="en">Date of visit</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Visit Date</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.trial_visit.v1/context/other_context/items[0004]/value"/>
    <!-- Comparator is not used for FormalExpressions,
        bug in ODM 1.3 means it is required
    -->
    <RangeCheck SoftHard="Hard" Comparator="EQ">
      <FormalExpression Context="ISVL"><![CDATA[
//Cannot be a future date
not(:this > :localtime)
]]></FormalExpression>
      <ErrorMessage>
        <TranslatedText xml:lang="en">
          Visit date cannot be a future date!
        </TranslatedText>
      </ErrorMessage>
    </RangeCheck>
  </ItemDef>
  <ItemDef OID="ID.FEEL" Name="Subject Feeling Report" DataType="text"
Length="1">
    <Description>
      <TranslatedText xml:lang="en">Subject Reported Feeling</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Subject Reported Feeling</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.well_being.v1]
/data/events[0002]/data/items[0003]/value"/>
    <CodeListRef CodeListOID="CL.SUB_FEELING"/>
  </ItemDef>
  <ItemDef OID="ID.CONT" Name="Continue to next visit" DataType="text"
Length="1">
    <Description>
      <TranslatedText xml:lang="en">Continue to next visit?</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Will the subject continue to the next
visit?</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-ADMIN_ENTRY.trial_visit_information.v1]
/data/items[0002]/value"/>
    <CodeListRef CodeListOID="CL.YESNO"/>
  </ItemDef>
  <!--
    Demography Form
  -->
  <ItemDef OID="ID.DOB" Name="Date of Birth" DataType="date">
    <Description>
      <TranslatedText xml:lang="en">Date of Birth</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Date of Birth</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-DEMOGRAPHIC-ITEM_TREE.person_details.v1]
/items[0001]/items[0010]/value"/>
    <!-- Comparator is not used for FormalExpressions,
        bug in ODM 1.3 means it is required
    -->
    <RangeCheck SoftHard="Hard" Comparator="EQ">
      <FormalExpression Context="ISVL"><![CDATA[
//Cannot be a future date
not(:this > :localtime)
]]></FormalExpression>
      <ErrorMessage>
        <TranslatedText xml:lang="en">
          Date of Birth cannot be a future date!
        </TranslatedText>
      </ErrorMessage>
    </RangeCheck>
  </ItemDef>

```

```

        </TranslatedText>
    </ErrorMessage>
</RangeCheck>
</ItemDef>
<ItemDef OID="ID.AGE" Name="Subject age at visit date" DataType="integer"
Length="3">
    <Description>
        <TranslatedText xml:lang="en">Age at visit date</TranslatedText>
    </Description>
    <Question>
        <TranslatedText xml:lang="en">Age</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-DEMOGRAPHIC-CLUSTER.person_age.v99]
/items[0002]/value"/>
    <RangeCheck SoftHard="Soft" Comparator="GE">
        <CheckValue>18</CheckValue>
        <ErrorMessage>
            <TranslatedText xml:lang="en">
                Subject should be at least 18 years
                of age at this visit
            </TranslatedText>
        </ErrorMessage>
    </RangeCheck>
    <RangeCheck SoftHard="Soft" Comparator="LE">
        <CheckValue>120</CheckValue>
        <ErrorMessage>
            <TranslatedText xml:lang="en">
                Subject should be less than 120 years
                of age at this visit
            </TranslatedText>
        </ErrorMessage>
    </RangeCheck>
</ItemDef>
<ItemDef OID="ID.SEX" Name="Sex" DataType="text" Length="1">
    <Description>
        <TranslatedText xml:lang="en">Sex</TranslatedText>
    </Description>
    <Question>
        <TranslatedText xml:lang="en">Sex</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-DEMOGRAPHIC-ITEM_TREE.person_details.v1]/
/items[0001]/items[0017]/value"/>
    <CodeListRef CodeListOID="CL.SEX"/>
</ItemDef>
<ItemDef OID="ID.FERTILE" Name="Fertile Female" DataType="text" Length="1">
    <Description>
        <TranslatedText xml:lang="en">Female of childbearing
potential</TranslatedText>
    </Description>
    <Question>
        <TranslatedText xml:lang="en">If Female, of childbearing
potential</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EVALUATION.childbearing_potential.v99]
/data/items[0002]/value"/>
    <CodeListRef CodeListOID="CL.YESNO"/>
</ItemDef>
<ItemDef OID="ID.RACE" Name="Race" DataType="text" Length="1">
    <Description>
        <TranslatedText xml:lang="en">Race</TranslatedText>
    </Description>
    <Question>
        <TranslatedText xml:lang="en">Race</TranslatedText>
    </Question>
    <CodeListRef CodeListOID="CL.RACE"/>
</ItemDef>
<ItemDef OID="ID.RACEOTH" Name="Race Other Specify" DataType="text"
Length="40">
    <Description>
        <TranslatedText xml:lang="en">Race Other Specify</TranslatedText>
    </Description>
    <Question>
        <TranslatedText xml:lang="en">If Race Other, Specify:</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-DEMOGRAPHIC-ITEM_TREE
.person_details_trial.v99]/items[0036]/value"/>
</ItemDef>
<!--

    Pregnancy Test Form

-->

```

```

    <ItemDef OID="ID.TESTRESULT" Name="Pregnancy Test Result" DataType="text"
Length="1">
    <Description>
    <TranslatedText xml:lang="en">Test result</TranslatedText>
    </Description>
    <Question>
    <TranslatedText xml:lang="en">Test result</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.pregnancy_test.v1]
/data/events[at0002]/data/item[at0004]/value/defining_code"/>
    <CodeListRef CodeListOID="CL.POSNEG"/>
    </ItemDef>
    <ItemDef OID="ID.TESTPERFORMED" Name="Pregnancy Test Performed" DataType="text"
Length="1">
    <Description>
    <TranslatedText xml:lang="en">Pregnancy Test Performed</TranslatedText>
    </Description>
    <Question>
    <TranslatedText xml:lang="en">was a pregnancy test
performed?</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION
.pregnancy_test-trial.v99]/data/events[at0002]/data/item[at0013]/value"/>
    <CodeListRef CodeListOID="CL.YESNO"/>
    </ItemDef>
    <!--

    Vitals Form

    -->
    <ItemDef OID="ID.HEIGHT" Name="Height" DataType="float" Length="5"
SignificantDigits="2">
    <Question>
    <TranslatedText xml:lang="en">Height</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.height.v1]
/data/events[at0002]/data/item[at0004]/value"/>
    <MeasurementUnitRef MeasurementUnitOID="MU.CM"/>
    </ItemDef>
    <ItemDef OID="ID.WEIGHT" Name="Weight" DataType="float" Length="4"
SignificantDigits="1">
    <Question>
    <TranslatedText xml:lang="en">Weight</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.body_weight.v1]
/data/events[at0003]/data/item[at0004]/value"/>
    <MeasurementUnitRef MeasurementUnitOID="MU.KG"/>
    </ItemDef>
    <ItemDef OID="ID.BMI" Name="Height" DataType="float" Length="3"
SignificantDigits="1">
    <Question>
    <TranslatedText xml:lang="en">BMI</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.body_mass_index.v1]
/data/events[at0002]/data/item[at0004]/value"/>
    </ItemDef>
    <ItemDef OID="ID.PULSE" Name="Pulse" DataType="integer" Length="3">
    <Question>
    <TranslatedText xml:lang="en">Pulse</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.heart_rate-pulse.v1]
/data/events[at0003]/data/items[at0004]/value"/>
    <MeasurementUnitRef MeasurementUnitOID="MU.BPM"/>
    </ItemDef>
    <ItemDef OID="ID.SYS" Name="Systolic Blood Pressure" DataType="integer"
Length="3">
    <Question>
    <TranslatedText xml:lang="en">Sys</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.blood_pressure.v1]
/data/events[at0006]/data/items[at0004]/value"/>
    <MeasurementUnitRef MeasurementUnitOID="MU.MMHG"/>
    </ItemDef>
    <ItemDef OID="ID.DIA" Name="Diastolic Blood Pressure" DataType="integer"
Length="3">
    <Question>
    <TranslatedText xml:lang="en">Dia</TranslatedText>
    </Question>

```

```

    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.blood_pressure.v1]
/data/events[at0006]/data/items[at0005]/value"/>
    <MeasurementUnitRef MeasurementUnitOID="MU.MMHG"/>
  </ItemDef>
  <ItemDef OID="ID.TEMP" Name="Body Temperature" DataType="float" Length="3"
SignificantDigits="1">
    <Question>
      <TranslatedText xml:lang="en">Body Temperature</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.study_data.v1/content[openEHR-EHR-OBSERVATION.body_temperature.v1]
/data/events[at0003]/data/items[at0004]/value"/>
    <MeasurementUnitRef MeasurementUnitOID="MU.DEG_C"/>
  </ItemDef>
  <!--
    Concomitant Medications Form
  -->
  <ItemDef OID="ID.ANYCONMED" Name="Any Concomitant Medications" DataType="text"
Length="1">
    <Description>
      <TranslatedText xml:lang="en">Any Concomitant Medications</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">
        Is the Subject taking any Concomitant Medications?
      </TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/context/other_context/items[at0010]/value/value"/>
    <CodeListRef CodeListOID="CL.YESNO"/>
  </ItemDef>
  <ItemDef OID="ID.DRUGNUM" Name="Drug ID" DataType="text" Length="4">
    <Description>
      <TranslatedText xml:lang="en">Drug repeat number</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">ID</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/context/other_context/items[at0003]/value"/>
  </ItemDef>
  <!-- ItemDefs for all CM elements-->
  <ItemDef OID="ID.DRUGNAME" Name="Drug name" DataType="text" Length="50">
    <Description>
      <TranslatedText xml:lang="en">Drug name</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Drug Name</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]/items[at0001]/value"/>
  </ItemDef>
  <ItemDef OID="ID.DOSE" Name="Drug Dose" DataType="text" Length="4">
    <Description>
      <TranslatedText xml:lang="en">Drug dose</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Dose</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]
/items[at0033]/items[at0034]/items[at0005]/value"/>
  </ItemDef>
  <ItemDef OID="ID.DOSEUNITS" Name="Dose Units" DataType="text" Length="4">
    <Description>
      <TranslatedText xml:lang="en">Dose units</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Units</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]/items[at0006]/value"/>
  </ItemDef>
  <ItemDef OID="ID.FREQ" Name="Frequency" DataType="text" Length="1">
    <Description>
      <TranslatedText xml:lang="en">Dosing Frequency</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Frequency</TranslatedText>
    </Question>

```



```

    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]/items[at0.77]/value"/>
    <CodeListRef CodeListOID="CL.DOSEFREQ"/>
  </ItemDef>
  <ItemDef OID="ID.STARTDAT" Name="Medication Start Date" DataType="date">
    <Description>
      <TranslatedText xml:lang="en">Date medication started</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Start Date</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]
/items[at0018]/items[at0019]/value"/>
    <!--
      Comparator not required for FormalExpression type range check
      but bug in ODM 1.3 requires it.
    -->
    <RangeCheck SoftHard="Hard" Comparator="EQ">
      <FormalExpression Context="ISVL"><![CDATA[
        //Cannot be a future date
        not(:this > :localtime)
      ]]></FormalExpression>
      <ErrorMessage>
        <TranslatedText xml:lang="en">
          Medication Start Date cannot be a future date!
        </TranslatedText>
      </ErrorMessage>
    </RangeCheck>
  </ItemDef>
  <ItemDef OID="ID.ENDDAT" Name="Medication End Date" DataType="date">
    <Description>
      <TranslatedText xml:lang="en">Date medication ended</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">End Date</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]
/items[at0018]/items[at0032]/value"/>
    <!--
      Comparator required because of bug in ODM 1.3 specification
      Should not be required for a FormalExpression check
    -->
    <RangeCheck SoftHard="Hard" Comparator="EQ">
      <FormalExpression Context="ISVL"><![CDATA[
        //Cannot be a future date
        not(:this > :localtime)
      ]]></FormalExpression>
      <ErrorMessage>
        <TranslatedText xml:lang="en">
          Medication End Date cannot be a future date!
        </TranslatedText>
      </ErrorMessage>
    </RangeCheck>
    <!--
      Comparator required because of bug in ODM 1.3 specification
      Should not be required for a FormalExpression check
    -->
    <RangeCheck SoftHard="Soft" Comparator="EQ">
      <FormalExpression Context="ISVL"><![CDATA[
        //Cannot be prior to start date
        not(:this < :this.itemgroup.field(name->"ID.STARTDAT"))
      ]]></FormalExpression>
      <ErrorMessage>
        <TranslatedText xml:lang="en">
          Medication End Date cannot be prior to Start Date
        </TranslatedText>
      </ErrorMessage>
    </RangeCheck>
  </ItemDef>
  <ItemDef OID="ID.DRUGCONT" Name="Drug Continuing" DataType="text" Length="1">
    <Description>
      <TranslatedText xml:lang="en">Drug continuing</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Continuing?</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/context/other_context/items[at0009]/value/value"/>
    <!--
      Comparator required because of bug in ODM 1.3 specification
      Should not be required for a FormalExpression check
    -->

```

```

<RangeCheck SoftHard="Soft" Comparator="EQ">
  <FormalExpression Context="ISVL"><![CDATA[
    //Must not be continuing if End Date Supplied
    (Decode(:this) == "Yes")
    AND
    Not Empty(:this.itemgroup.field(name->"ID.ENDDAT"))
  ]]></FormalExpression>
  <ErrorMessage>
    <TranslatedText xml:lang="en">
      Cannot be continuing if End Date is set.
    </TranslatedText>
  </ErrorMessage>
</RangeCheck>
<CodeListRef CodeListOID="CL.YESNO"/>
</ItemDef>
<ItemDef OID="ID.DIAGNOSIS" Name="Diagnosis" DataType="text" Length="100">
  <Description>
    <TranslatedText xml:lang="en">Diagnosis</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">Diagnosis</TranslatedText>
  </Question>
</ItemDef>
<Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_cm.v1/content[openEHR-EHR-ACTION.medication-cdash_cm.v1]
/description[openEHR-EHR-ITEM_TREE.medication-cdash_cm.v1]
/items[at0010]/items[at0011]/value"/>
<!--

      Adverse Events Form

-->
<ItemDef OID="ID.ANYAE" Name="Any Adverse Events" DataType="text" Length="1">
  <Description>
    <TranslatedText xml:lang="en">Any Adverse Events</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">
      Have any Adverse Events been reported?
    </TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/context/other_context/items[at0010]/value/value"/>
  <CodeListRef CodeListOID="CL.YESNO"/>
</ItemDef>
<ItemDef OID="ID.AENUM" Name="AE #" DataType="integer" Length="4">
  <Description>
    <TranslatedText xml:lang="en">AE repeat number</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">#</TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/context/other_context/items[at0003]/value"/>
</ItemDef>
<ItemDef OID="ID.AEDESC" Name="Adverse Event Description" DataType="text"
Length="50">
  <Description>
    <TranslatedText xml:lang="en">Adverse Event Description</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">AE Description</TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse.v1]
/data/items[at0019]/items[at0022]/value"/>
</ItemDef>
<ItemDef OID="ID.AESTARTDAT" Name="AE Start Date" DataType="date">
  <Description>
    <TranslatedText xml:lang="en">Date AE started</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">Start Date</TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse.v1]
/data/items[at0019]/items[at0027]/value"/>
<!--
      Comparator not required for FormalExpression type range check
      but bug in ODM 1.3 requires it.
-->
<RangeCheck SoftHard="Hard" Comparator="EQ">
  <FormalExpression Context="ISVL"><![CDATA[
    //Cannot be a future date
    not(:this > :localtime)
  ]]></FormalExpression>

```

```

    <ErrorMessage>
      <TranslatedText xml:lang="en">
        Adverse Event Start Date cannot be a future date!
      </TranslatedText>
    </ErrorMessage>
  </RangeCheck>
</ItemDef>
<ItemDef OID="ID.AEENDDAT" Name="AE End Date" DataType="date">
  <Description>
    <TranslatedText xml:lang="en">Date AE ended</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">End Date</TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse.v1]
/data/items[at0019]/items[at0021]/value"/>
  <!--
    Comparator required because of bug in ODM 1.3 specification
    Should not be required for a FormalExpression check
  -->
  <RangeCheck SoftHard="Hard" Comparator="EQ">
    <FormalExpression Context="ISVL"><![CDATA[
      //Cannot be a future date
      not(:this > :localtime)
    ]]></FormalExpression>
    <ErrorMessage>
      <TranslatedText xml:lang="en">
        Adverse Event End Date cannot be a future date!
      </TranslatedText>
    </ErrorMessage>
  </RangeCheck>
  <!--
    Comparator required because of bug in ODM 1.3 specification
    Should not be required for a FormalExpression check
  -->
  <RangeCheck SoftHard="Soft" Comparator="EQ">
    <FormalExpression Context="ISVL"><![CDATA[
      //Cannot be prior to start date
      not(:this < :this.itemgroup.field(name->"ID.AESTARTDAT"))
    ]]></FormalExpression>
    <ErrorMessage>
      <TranslatedText xml:lang="en">
        Adverse Event Date cannot be prior to Start Date
      </TranslatedText>
    </ErrorMessage>
  </RangeCheck>
</ItemDef>
<ItemDef OID="ID.AECONT" Name="AE Continuing" DataType="text" Length="1">
  <Description>
    <TranslatedText xml:lang="en">Adverse Event continuing</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">Continuing?</TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/context/other_context/items[at0009]/value/value"/>
  <!--
    Comparator required because of bug in ODM 1.3 specification
    Should not be required for a FormalExpression check
  -->
  <RangeCheck SoftHard="Soft" Comparator="EQ">
    <FormalExpression Context="ISVL"><![CDATA[
      //Must not be continuing if End Date Supplied

      (Decode(:this) == "Yes")
      AND
      Not Empty(:this.itemgroup.field(name->"ID.AEENDDAT"))
    ]]></FormalExpression>
    <ErrorMessage>
      <TranslatedText xml:lang="en">
        Cannot be continuing if End Date is set.
      </TranslatedText>
    </ErrorMessage>
  </RangeCheck>
  <CodeListRef CodeListOID="CL.YESNO"/>
</ItemDef>
<ItemDef OID="ID.AESEVERITY" Name="AE Severity" DataType="text" Length="1">
  <Description>
    <TranslatedText xml:lang="en">Adverse Event Severity</TranslatedText>
  </Description>
  <Question>
    <TranslatedText xml:lang="en">Severity</TranslatedText>
  </Question>
  <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse.v1]
/data/items[at0019]/items[at0023]/value/defining_code"/>

```

```

    <CodeListRef CodeListOID="CL.AESEVERITY"/>
  </ItemDef>
  <ItemDef OID="ID.AESERIOUS" Name="Is AE Serious" DataType="text" Length="1">
    <Description>
      <TranslatedText xml:lang="en">Is Adverse Event classified
Serious?</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Serious</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse-trial.v1]/
data/items[at0019]/items[at0035]/value/value"/>
    <CodeListRef CodeListOID="CL.YESNO"/>
  </ItemDef>
  <ItemDef OID="ID.AERELATION" Name="AE Relationship=" DataType="text"
Length="1">
    <Description>
      <TranslatedText xml:lang="en">Adverse Event Relationship to
Study Drug</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Relationship</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse-trial.v1]
/data/items[at0019]/items[at0036]/value/value"/>
    <CodeListRef CodeListOID="CL.AERELATION"/>
  </ItemDef>
  <ItemDef OID="ID.AEACTION" Name="AE Action Taken" DataType="text" Length="1">
    <Description>
      <TranslatedText xml:lang="en">Action taken on Adverse
Event</TranslatedText>
    </Description>
    <Question>
      <TranslatedText xml:lang="en">Action Taken</TranslatedText>
    </Question>
    <Alias Context="Archetype" Name="openEHR-EHR-
COMPOSITION.cdash_ae.v1/content[openEHR-EHR-EVALUATION.adverse-trial.v1]
/data/items[at0019]/items[at0037]/value/value"/>
    <CodeListRef CodeListOID="CL.AEACTION"/>
  </ItemDef>
<!--

Code-Lists:

-->
<CodeList OID="CL.YESNO" Name="Yes No CodeList" DataType="text">
  <CodeListItem CodedValue="Y" Rank="1">
    <Decode>
      <TranslatedText xml:lang="en">Yes</TranslatedText>
    </Decode>
  </CodeListItem>
  <CodeListItem CodedValue="N" Rank="2">
    <Decode>
      <TranslatedText xml:lang="en">No</TranslatedText>
    </Decode>
  </CodeListItem>
</CodeList>

[...]
```

```

  <!--

Condition-Definitions:

-->
<ConditionDef OID="CD.IS_NOT_MALE" Name="Is not male">
  <Description>
    <TranslatedText xml:lang="en">
Returns True if the subject is not male (i.e. they are female)
    </TranslatedText>
  </Description>
  <FormalExpression Context="ISVL"><![CDATA[
Decode(event->"SE.INIT",
form->"FD.DEMOG",
item->"ID.SEX") != "Male"
]]></FormalExpression>
</ConditionDef>

[...]
```

```

  <!--

Method-Definitions:

-->
```

```
<MethodDef OID="MD.ASSIGN_AENUM" Name="Assign AE repeat number"
Type="Computation">
  <Description>
    <TranslatedText xml:lang="en">Assign an AE repeat number
    unique within the repeats of the ItemGroup.</TranslatedText>
  </Description>
  <FormalExpression Context="ISVL"><![CDATA[
    //Get the next available number

    itemnums = select(event->"SE.AE",
                      form->"FD.AE",
                      item->"ID.AENUM")

    aenum = 1
    while true {
      if not In(daenum,itemnums) {
        break
      }
      aenum += 1
    }

    //return the new number
    aenum
  ]]></FormalExpression>
</MethodDef>

[...]
```

```
</MetaDataVersion>
</Study>
</ODM>
```


5. openSDMS: Prozessabläufe

Im Folgenden werden die vorgesehenen Anwenderinteraktionen innerhalb der verschiedenen Einsatzszenarien von openSDMS (siehe Kapitel 4.2) im Detail dargestellt.

a) Stand-Alone-System

Nutzung der EPA-Komponente:

EPA-Struktur definieren

- Die EPA-Struktur von openSDMS wird definiert, indem der Anwender die einzelnen, auf Archetypen basierenden Eingabeformulare, in der zum System gehörenden Datei ‚structure.xml‘ auflistet.

Neuen Patienten anlegen

- Ein neuer Patient wird im System angelegt, indem ein neuer Datensatz (Patienten-ID, Vorname, Nachname, Geburtsdatum des Patienten) in der Tabelle ‚patient_data‘ angelegt wird.

Medizinische Daten eines Patienten einsehen, ändern oder ergänzen

- Der Anwender selektiert in der Patientenliste, den Patienten, dessen Daten eingesehen, geändert oder ergänzt werden sollen. Falls bereits ein Patient ausgewählt ist, kann die Patientenliste durch wählen des Links ‚Current Patient (<Name des aktuell ausgewählten Patienten>‘ aufgerufen werden.
- Der Anwender wählt links im Navigationsmenü das gewünschte (auf Archetypen basierende) Eingabeformular aus: Ein leeres Erfassungsformular wird geöffnet. Parallel werden alle bereits gespeicherten Datensätze angezeigt, welche auf diesem Formular basieren.
- Der Anwender kann nun einen neuen Datensatz in das noch leere Formular eintragen und abspeichern oder einen der zuvor erfassten Datensätze auswählen, um diesen zu ändern.

Nutzung der SDMS-Komponente:

Studie anlegen

- Eine neue Studie wird angelegt, indem eine entsprechende Studiendefinition in die Datei ‚structure.xml‘ eingefügt wird. Dieser Datensatz definiert die Struktur der Studie gemäß Abb. 4-19. Der Verweis auf Studie, Studienarme und –visiten erfolgt mit Hilfe der zugehörigen, eindeutigen ID.

Studien-Metadaten erfassen oder ändern

- Der Anwender wählt zunächst den Link ‚Current Trial ()‘ – die Liste aller im System angelegten Studien wird angezeigt.
- Dann selektiert der Anwender die gewünschte Studie (das System wechselt in die Sicht ‚Studie / Metadaten‘).
- Nun kann der Anwender links im Navigationsmenü das gewünschte Element bearbeiten (beispielsweise einen bestimmten Studienarm und dann dessen Beschreibung): Falls zu dem jeweiligen Formular noch kein Datensatz vorhanden ist, wird ein leeres Erfassungsformular geöffnet. Falls bereits ein Datensatz erfasst wurde, wird dieser angezeigt. (Im Gegensatz zu medizinischen Daten, die mehrfach erfasst werden können, werden Studien-Metamerkmale wie etwa der Studienname genau einmal pro Studie erfasst.)

Der Nutzer kann nun entweder einen neuen Datensatz in das noch leere Formular eintragen und abspeichern oder – falls bereits ein Datensatz erfasst wurde – diesen ändern.

Patient einer Studie zuordnen

- Ein Patient wird einer Studie zugeordnet, indem ein neuer Datensatz in die Tabelle ‚patient_trial‘ eingefügt wird. Dieser Datensatz besteht aus Patienten-ID und Studien-ID sowie der ID des Studienarms, welchem der jeweilige Nutzer zugeordnet ist. Hinzu kommt außerdem die ID des Nutzers, welcher den Datensatz angelegt hat.

Studiendaten eines Patienten einsehen, ändern oder ergänzen

- Der Anwender selektiert in der Patientenliste, den Patienten, dessen Daten er einsehen, ändern oder ergänzen möchte. Falls bereits ein Patient ausgewählt ist, kann die Patientenliste durch wählen des Links ‚Current Patient (<Name des aktuell ausgewählten Patienten>)‘ aufgerufen werden. Alternativ kann zu jeder Studie eine Liste aller Patienten abgerufen werden, welche der jeweiligen Studie zugeordnet sind – auch in dieser Liste kann bei Bedarf ein Patient ausgewählt werden.
- Das System wechselt in die Sicht ‚Studie / CRF‘.
- Der Nutzer wählt die gewünschte Studienvisite aus. (In der Regel ist ein Patient nur einem Arm einer Studie zugeordnet. Im Studienmodell von openSDMS wird ferner angenommen, dass ein Patient an maximal einer Studie gleichzeitig teilnimmt. Daher ist durch die Auswahl einer bestimmten Visite die entsprechende Studie samt Studienarm eindeutig bestimmt.)
- Der Nutzer wählt das gewünschte auf Archetypen basierende Eingabeformular aus: Falls zu dem jeweiligen Formular noch kein Datensatz vorhanden ist, wird ein leeres Erfassungsfomular geöffnet. Falls bereits ein Datensatz erfasst wurde, wird dieser angezeigt. Im Gegensatz zu medizinischen Daten in einer EPA, die mehrfach erfasst werden können, werden Merkmalsarten in Studien – wie auch die Studien-Metamerkmale – genau einmal pro Visite pro Patient bzw. wie in der Studiendefinition vorgegeben erfasst. Wird eine Merkmalsart innerhalb einer Visite mehrfach benötigt – etwa ein Blutdruckwert jeweils vor und nach Belastung – so muss dies bereits beim Entwurf des entsprechenden Erhebungsformulars berücksichtigt werden.
- Der Nutzer kann nun einen neuen Datensatz in das noch leere Formular eintragen und abspeichern oder einen der zuvor erfassten Datensätze auswählen, um diesen zu ändern.
- Ist noch kein Datensatz vorhanden, sucht openSDMS in der EPA des ausgewählten Patienten nach geeigneten, bereits erfassten Daten. Sind passende Einträge vorhanden, schlägt das System diese zur Übernahme in die Studie vor (pre-population). Anstelle des manuellen Erfassens von Werten kann der Anwender in diesem Fall einen vorgeschlagenen Wert einfach durch Selektieren übernehmen. (Daten die bereits in einer Studie verwendet wurden, werden nicht mehr vorgeschlagen. Werden mehrere potenziell geeignete Datenelemente gefunden, wird die Liste mit den Vorschlägen anhand des Erfassungszeitpunkts der einzelnen Daten chronologisch sortiert.)
- Alle im Rahmen einer Studie zu einem Patienten erfassten Werte erscheinen automatisch auch in der EPA dieses Patienten (Single Source Konzept).

b) Integrationssystem

Wird openSDMS im Rahmen einer multizentrischen Studie als Studiendatenmanagementsystem eingesetzt, so funktioniert der Datenübernahme-Mechanismus nur, wenn die Daten der Studienteilnehmer aus den elektronischen Patientenakten der beteiligten Zentren in die EPA-Komponente von openSDMS importiert wurden. Da zu einem Patienten, insbesondere bei stationären Behandlungen, in der Regel kontinuierlich Daten in seiner EPA erfasst werden, muss der Export nach openSDMS in regelmäßigen Zyklen erfolgen. Hierfür muss zunächst ein geeigneter Mechanismus implementiert werden.

Die übrigen Prozesse sind analog zu a) ‚Stand-Alone-System‘.

Lebenslauf**Christian Dominik Kohl**

* 28. August 1980 in Heilbronn-Neckargartach

Anschrift

Gartenstr. 18
74172 Neckarsulm

Schulischer Werdegang

**08/1987 –
07/1991** Grundschule – Neubergschule Neckarsulm

**09/1991 –
06/2000** Gymnasium – Albert-Schweitzer-Gymnasium Neckarsulm

06/2000 Abitur

Zivildienst

**07/2000 –
05/2001** Caritas Altenzentrum „Parkwohnheim St. Vinzenz“ – Neckarsulm
im Bereich „Technik“

Universitärer Werdegang

**09/2001 –
08/2006** Studium der Medizinischen Informatik an der
Universität Heidelberg / Hochschule Heilbronn

08/2006 Diplom in Medizinischer Informatik

Beruflicher Werdegang

**seit
10/2006** Universitätsklinikum Heidelberg
Institut für Medizinische Biometrie und Informatik
Sektion Medizinische Informatik

**seit
11/2006** Universitätsklinikum Heidelberg
Koordinierungszentrum für Klinische Studien (KKS) Heidelberg
Abteilung IT

**seit
11/2010** Leiter der IT Abteilung des KKS Heidelberg

Auslandspraktikum

09/2007 Praktikum zum Thema ‚open Electronic Health Record‘
an der Central Queensland University
Faculty of Business and Informatics / Informatics Research Group
am Austin Centre for Applied Clinical Informatics
in Melbourne, Australien

Danksagungen

Ohne die Unterstützung und den Rat vieler verschiedener Menschen wäre diese Arbeit nicht möglich gewesen – ihnen allen gilt mein herzlichster Dank.

Besonders danke ich Frau Professor Dr. Petra Knaup, die diese Arbeit mit vielen guten fachlichen und praktischen Ideen sowie viel Engagement begleitet hat.

Großer Dank gilt auch Frau Dr. Heather Leslie, Herrn Dr. Ian McNicoll und Herrn Dr. Sebastian Garde von Ocean Informatics, die immer ein offenes Ohr für meine Fragen hatten und mir kompetent und engagiert Antworten auf meine Fragen gaben. Ferner danke ich Herrn Dr. Jozef Aerts von XML4Pharma und dem Schreiber des Blogs ‚www.eclinicalopinion.blogger.com‘.

Ebenso danke ich Herrn Professor Dr. Dittmar Böckler herzlich für die Möglichkeit eine Studie aus seiner Abteilung als Fallstudie für diese Arbeit zu nutzen. In diesem Zusammenhang danke ich ebenfalls Frau Petra Hoppek, die mir geduldig alle meine Fragen beantwortet hat.

Ferner danke ich den Kollegen aus dem KKS Heidelberg, die mir sowohl den Rücken für diese Arbeit freigehalten haben, als auch für inhaltliche Fragen immer zur Verfügung standen. Besonders nennen möchte ich hierbei meine Kolleginnen und Kollegen aus der IT-Abteilung Frau Dr. Katja Krockenberger, Frau Heidrun Binder, Frau Sonja Früh, Herrn Andreas Klinger und Herrn Andreas Dörrzapf sowie die Kollegen Erika Müller und Gregor Ottawa. Mein besonderer Dank gilt dem Leiter des KKS Heidelberg, Herrn Dr. Steffen Luntz, der diese Arbeit auf vielfältige Weise unterstützt hat.

Ebenso danke ich meinen Kolleginnen und Kollegen aus dem Institut für Medizinische Biometrie und Informatik herzlich für ihre Unterstützung und den guten fachlichen Austausch. Mein besonderer Dank gilt dem Leiter der Sektion Medizinische Informatik Herrn Professor Dr. Hartmut Dickhaus, dessen Rat mir immer sehr wertvoll war.

Sehr herzlich danke ich meinen Eltern, die mir mein Studium und damit auch diese Arbeit erst ermöglicht und die mich auch während dieser Arbeit vielfach unterstützt und ermutigt haben.

Ein ganz besonderer Dank gilt auch meiner Freundin, Frau Christina Weber, die mir ebenfalls den Rücken freigehalten und vor allen Dingen diese Arbeit mitgetragen hat.