

Creating together
the digital health environment
of tomorrow.
Now.



MEDIZININFORMATIK - INITIATIVE

Schlussbericht HiGHmed 01ZZ1802A

Berichtszeitraum: 01.01.2018 - 31.12.2023

Konsortialleitung:

Prof. Dr. Roland Eils

Medizinische Fakultät der Universität Heidelberg /

Universitätsklinikum Heidelberg

Im Neuenheimer Feld 267

D-69120 Heidelberg

Telefon: + 49 6221 54 51290

Fax: + 49 6221 54 51488

roland.eils@bioquant.uni-heidelberg.de

Inhalt

1. Zusammenfassende Darstellung des gesamten Projektverlaufs	3
2. Eingehende Darstellung	6
2.1. Verwendung der Zuwendung und des erzielten Ergebnisses im Einzelnen, mit Gegenüberstellung der vorgegebenen Ziele	6
2.2 Das Medizinische Datenintegrationszentrum des Universitätsklinikums Heidelberg	7
2.2.1 Organisation.....	7
2.2.2 Rollen und Aufgaben.....	7
2.2.3 Architektur, Schnittstellen und Datenschutz	9
2.2.4 OpenEHR Clinical Data Repository.....	12
2.2.5 FHIR Clinical Data Repository.....	12
2.2.6 ETL-Prozesse und Semantik.....	12
2.2.7 Mitentwicklung, Anbindung und Nutzung des Data Sharing Frameworks .	13
2.2.8 Record Linkage und Pseudonymisierung.....	14
2.2.9 Qualitätsmanagement	14
2.3 Weitere Strukturen und Komponenten	15
2.3.1 Use & Access Committee	15
2.3.2 Datennutzungsanträge.....	16
2.3.3 Consent Management.....	17
2.3.4 Verteiltes Rechnen.....	18
2.3.5 Patientenportal	18
2.3.6 Nutzung des Datenintegrationszentrums im Rahmen der MII und HiGHmed	18
2.4 Wichtigste Positionen des zahlenmäßigen Nachweises.....	23
2.5 Notwendigkeit und Angemessenheit der geleisteten Arbeit	23
2.6 Verwertbarkeit der Ergebnisse im Sinne des fortgeschriebenen Verwertungsplans.....	23

2.7 Während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen	24
2.8 Erfolgte Veröffentlichungen der Ergebnisse	24
3. Das Datenintegrationszentrum für „Genomics“- und „Radiomics“-Daten (OmicsDIC) des Universitätsklinikums Heidelberg	26
3.1. Kurzzusammenfassung	26
3.1.1 Aufgabenstellung	26
3.1.2 Planung und Ablauf	27
3.1.3 Stand der Wissenschaft und Technik	29
3.1.4 Projektpartner	30
3.2. Eingehende Darstellung OmicsDIC	30
3.2.1 Ziele und Ergebnisse	31
3.2.2 Notwendigkeit und Angemessenheit der Projektarbeiten	37
3.2.3 Verwendung der Zuwendung	37
3.2.4 Fortschreibung des Verwertungsplans	37
3.2.5 Veröffentlichung der Ergebnisse	39
ABBILDUNG 1: MEDIC SYSTEMLANDSCHAFT BASIEREND AUF AUFEINANDER AUFBAUENDEN SCHICHTEN.....	11
ABBILDUNG 2: MEILENSTEINPLAN UKL-HD	22
ABBILDUNG 3: MEILENSTEINPLAN OMICSDIC	29

1. Zusammenfassende Darstellung des gesamten Projektverlaufs

Als Ergebnis der neunmonatigen Konzeptphase (01.08.2016 – 30.04.2017) der Medizininformatik-Initiative hat das Konsortium HiGHmed nach den Vorgaben der BMBF-Ausschreibung die Grundlage für die Aufstellung eines Medizininformatikverbundes zwischen den beteiligten Universitätskliniken und weiteren Partner erarbeitet. Dies war Grundlage für die Erarbeitung des HiGHmed-Konzeptantrages für die Ausbau- und Vernetzungsphase (01.01.2018 – 31.12.2022). Dieser beschreibt im Detail den Aufbau von Medizinischen Datenintegrationszentren (MeDICs), die unter anderem die Grundlage für drei medizinische Use Cases aus den Bereichen Onkologie, Kardiologie und Infektiologie bilden. Die Use Cases konnten über die geplante Laufzeit verlängert werden, wodurch ein Teil der Arbeiten aus diesem Bericht zum Universitätsklinikum Heidelberg in das Jahr 2023 reichen.

Ziel des HiGHmed Konsortiums in der Aufbau- und Vernetzungsphase war der Aufbau von MeDICs, um Daten aus der Versorgung, Forschung und externen Quellen zu integrieren und somit die Entwicklung neuer technischer Lösungen für medizinische Datenanalysen zu vereinfachen. Um dieses Vorhaben umsetzen zu können, wurden verschiedene Arbeitspakete (APs) definiert, auf deren Aufgaben, Ziele und Verlauf während der vergangenen Förderperiode (2018 – 2023) im Folgenden näher eingegangen werden soll.

Das Universitätsklinikum Heidelberg arbeitet in dem Verbund mit weiteren Stellen zusammen. Die lokale Bearbeitung und Umsetzung der Arbeitspakete (AP) findet als Teil der übergeordneten APs aus dem Verbund statt. Der Konsortialverbund ist im Verlauf der Förderphase angewachsen und darüber lief die Zusammenarbeit mit nachfolgenden weiteren Stellen.

Zusammenarbeit im Verbund:

FKZ	Projektpartner	Projektleiter
01ZZ1802A	Universitätsklinikum Heidelberg (UKL-HD)	Prof. Dr. Roland Eils
01ZZ1802B	Universitätsmedizin Göttingen (UMG)	Prof. Dr. Ramin Yahyapour
01ZZ1802C	Medizinische Hochschule Hannover (MHH)	Prof. Dr. Dr. Michael Marschollek
01ZZ1802D	Deutsches Krebsforschungszentrum (DKFZ)	Prof. Dr. Martin Lablans

01ZZ1802E	Hochschule Heilbronn (HHN)	Prof. Dr. Christian Fegeler / Prof. Dr. Rolf Bendl
01ZZ1802F	Universität Heidelberg (Uni-HD)	Prof. Dr. Christoph Dieterich
01ZZ1802G	TU Darmstadt (TU-D)	Dr. Tatiana von Landesberger / Prof. Dr. Kay Hamacher
01ZZ1802H	Hasso-Plattner-Institut für Digital Engineering (HPI)	Dr.-Ing. Matthieu-P. Schapranow
01ZZ1802I	Robert Koch-Institut (RKI)	Dr. Tim Eckmanns
01ZZ1802J	InterComponentWare (ICW)	Dr. Ralf Brandner
01ZZ1802K	Siemens Healthcare GmbH*	Dr. Michael Meyer
01ZZ1802M	NEC Laboratories Europe GmbH*	Dr. Jürgen Quittek
01ZZ1802N	Helmholtz-Zentrum für Infektionsforschung (HZI)	Prof. Alice McHardy
01ZZ1802O	HAWK Hochschule für angewandte Wissenschaft und Kunst (HAWK)	Prof. Dr. Christoph Rußmann
01ZZ1802P	Hochschule Hannover (HsH)	Prof. Dr. Oliver J. Bott
01ZZ1802Q	TU Braunschweig (TU-BS)	Prof. Dr. Thomas M. Deserno
01ZZ1802S	Ada Health GmbH*	Dr. Martin Christian Hirsch
01ZZ1802T	Universitätsklinikum Schleswig-Holstein (UKSH) (Campus Kiel)	Prof. Dr. med. Björn Bergh
01ZZ1802U	Universitätsklinikum Köln (UKK)	Prof. Dr. Andreas Beyer
01ZZ1802V	Westfälische Wilhelms-Universität Münster (WWU)	Prof. Dr. med. Julian Varghese
01ZZ1802W	Universitätsklinikum Würzburg (UKW) und Medizinische Fakultät der Universität Würzburg (JMU)	Prof. Dr. Georg Ertl / Prof. Dr. Heuschmann
01ZZ1802X	Charité-Universitätsmedizin Berlin	Prof. Dr. Axel Radlach Pries / Fabian Prasser
01ZZ1802Y	Dell GmbH	Dr. Marten Neubauer
01ZZ1802Z	Universitätsklinikum Schleswig-Holstein (UKSH) (Campus Lübeck)	Prof. Dr. Josef Ingenerf
01ZZ180AB	Carl-Thiem-Klinikum Cottbus (CTK)	Prof. Dr. Dr. Kurt J.G. Schmailzl

Die technischen APs befassten sich mit der Schaffung der semantischen und syntaktischen/technischen Interoperabilität als Rahmen für den kompatiblen Datenaustausch und Analysen innerhalb des UKHD sowie zwischen den HiGHmed MeDICs. Dies sollte durch die Entwicklung gemeinsamer Informationsmodelle und einer gemeinsamen technischen Infrastruktur, basierend auf IHE Profilen, gewährleistet werden. Zudem war die institutionsübergreifende Entwicklung und Nutzung von gemeinsamen Analysetools, User Interfaces (UIs) und Software (für zum Beispiel National Language Processing (NLP)) vorgesehen. Die APs wurden in enger Zusammenarbeit mit weiteren Schwerpunktthemen, wie den APs bezüglich Ethik und Datenschutz, konzipiert. Die Erstellung von Datensicherheitskonzepten und Datenschutzkonzepten in Einklang mit konsortienübergreifenden Strukturen wie der TMF und dem NSG (Nationales Steuerungsgremium) waren ebenso Inhalt wie die Entwicklung von Prozessen und Empfehlungen für den Umgang mit personalisierten Daten (Nutzungs- und Zugangsprozesse). Die in HiGHmed entwickelten technischen Lösungen sollten in drei Use Cases (UCs) erprobt werden. Hierbei handelte es sich um den Use Case Oncology (UCO), der sich mit der Einrichtung einer Infrastruktur für einen standortübergreifenden Austausch bezüglich Therapiemöglichkeiten von HPB Krebspatienten/Krebspatientinnen (hepato-pankreato-biliär) durch ein gemeinsames Tumorboard befasste. In diesem Rahmen sollte unter anderem ein sogenannter Data Mart mit Zugriff auf relevante onkologische Daten eingerichtet werden. Der Use Case Cardiology (UCC) hatte die Aufgabe, innovative Datenkonzepte und standortübergreifende Analysemöglichkeiten zu identifizieren und zu erproben sowie einen Mehrwert für Patienten/Patientinnen aufgrund verbesserter Follow-ups (u.a. durch Sensordaten und Apps) zu demonstrieren. Im Use Case Infection Control (UCIC) stand die Entwicklung einer technischen Lösung im Fokus, mithilfe derer Häufungen von Pathogenen (insbesondere multiresistenten bakteriellen Erregern) und mögliche Übertragungsketten in Krankenhäusern mittels einer Clusterfrüherkennung detektiert werden können. Durch interaktive Visualisierung von Patientenbewegungsdaten und automatisierter Erkennung von Erregerclustern sollte Hygienefachpersonal bei der Aufgabe gestärkt werden, Ausbrüche einzudämmen oder sogar zu verhindern. Begleitet wurde die gesamte Entwicklung im HiGHmed Projekt durch das zentrale Projektmanagement (PM) und Arbeitspaketen mit Fokus auf die

Stärkung der Medizininformatik sowie die Etablierung von Roll-Out Strukturen. Ziel des PMs war die standortübergreifende Koordination des HiGHmed Konsortiums mittels einer gemeinsamen PM-Plattform, der Etablierung eines Technical Coordination Boards (TCB), der Organisation diverser Meetings auf Arbeitsebene und Governanceebene, sowie interner und externer Kommunikation. Das AP Lehre hatte die Stärkung der Medizininformatik durch die Gestaltung eines innovativen, institutionsübergreifenden Lehrangebots (in Kooperation mit Industriepartnern) zum Ziel. Um von Beginn an ein Roll-Out auf weitere Kliniken und Partner gewährleisten zu können, befasste sich das AP Roll-Out mit der Erarbeitung von Roll-Out Konzepten zur Integration weiterer Partner in das HiGHmed Konsortium, sowie zur Integration weiterer Datenressourcen.

2. Eingehende Darstellung

2.1. Verwendung der Zuwendung und des erzielten Ergebnisses im Einzelnen, mit Gegenüberstellung der vorgegebenen Ziele

Ziel des UKHDs war es, am Ende der Entwicklungs- und Vernetzungsphase Behandlern und Forschern den Zugriff auf große und verschiedenartige lokale Datenbestände durch die Erstellung eines medizinischen Datenintegrationszentrums (MeDIC) zu ermöglichen. Die fragmentierten Datenbestände des Klinikums sollten hierfür über das MeDIC in einer zentralen Infrastruktur im UKHD integriert werden. Die datenschutzrechtlichen Bestimmungen, IT-Sicherheit und Patienteneinwilligungen sollten hierfür vollumfänglich beachtet und umgesetzt werden. Hierfür sollte seitens der Hochschule Heilbronn eine datenschutzkonforme Pseudonymisierung entwickelt werden und zum Einsatz kommen. Die MeDIC Plattform sollte des Weiteren mithilfe von Industriepartnern unter Verwendung internationaler, offener Standards etabliert werden. Die Zusammenführung der Datenbestände des Klinikums sollte zu einer integrierten Versorgung und stark erweiterten, interdisziplinären Forschungsmöglichkeiten führen. Der Wert der momentan fragmentierten Datenbestände sollte sich in der Integration und interdisziplinären Verwendung und Zusammenführung der verfügbaren Daten zeigen.

Neben den benötigten Funktionalitäten für die drei Use Cases Kardiologie, Onkologie und Infektionskontrolle sollten konsortiumsweit Analyse- und Abfragewerkzeuge für

Forscher und Behandler zur Verfügung stehen. Weitere, lokale Infrastrukturen in einzelnen Abteilungen des Klinikums sollten nicht benötigt werden.

Für Patienten sollte ein Patientenportal zur Verfügung stehen, welches ihnen Einblicke in ihre Datenbestände gibt und von Behandlern und Forschern getätigte Zugriffe auf ihre Daten anzeigen kann. Des Weiteren sollten Einwilligungserklärungen von Patienten in diesem Patientenportal eingeholt werden können, welche nach Erstellung, Änderung oder Widerruf ohne Zeitverzögerung in einem zentralen Einwilligungsmanagement konsortiumsweit effektiv werden sollten. Hierfür sollte eigens ein Einwilligungsmanagement entwickelt und implementiert werden.

Neben der MeDIC-Plattform sollte auch ein OmicsDIC am UKHD entstehen.

2.2 Das Medizinische Datenintegrationszentrum des Universitätsklinikums Heidelberg

2.2.1 Organisation

Mit Beginn der Aufbau- und Vernetzungsphase wurde mit dem praktischen Aufbau des Medizinische Datenintegrationszentrum (MeDIC) begonnen. Im Februar 2019 wurde das erste Datennutzungsprojekt bearbeitet.

Organisatorisch war das MeDIC zunächst Bestandteil der Abteilung Medizinische Informationssysteme. Mit der Gründung des neuen Instituts für Medizinische Informatik wurde das MeDIC offiziell als dazugehörige wissenschaftliche Einheit im Mai 2021 gegründet. Im Herbst 2023 wurde die Zuständigkeit des MeDICs auch auf die Tochterunternehmen des UKHDs erweitert.

2.2.2 Rollen und Aufgaben

Wesentliche Aufgaben des MeDICs bestehen darin, im Einklang mit den Vorgaben und Zielen der nationalen Medizininformatikinitiative, die folgenden Aufgaben zu erfüllen:

1. An Standorten des UKHDs und dessen Tochterunternehmen anfallende klinische Daten aus den Informationssystemen der primären Routedokumentation zusammenzuführen, zu harmonisieren, semantisch zu annotieren, wenn möglich strukturiert zu speichern. Zu diesen Primärsystemquellen zählen das patientenführende System IS-H/ish.med, das digitale Archiv, das Radiologieinformationssystem, das PACS, das Laborinformationssystem, die PEPA sowie weitere Spezialsysteme.

2. Diese Daten für Anwendungen der Biomedizinischen Forschung oder Versorgung bereit zu stellen. Die momentan fragmentierten Datenbestände des UKHDs und dessen Tochterunternehmen werden hierfür in einer zentralen Plattform des MeDICs zusammengeführt. Die Zusammenführung erfolgt qualitätsgesichert in dedizierten Datenrouten mithilfe von internationalen, offenen Standards wie HL7v2, HL7 FHIR, DICOM, W3C und openEHR. Dabei werden so weit wie möglich Profile der internationalen Initiative Integrating the Healthcare Enterprise (IHE) genutzt. Zudem werden hieraus der Kerndatensatz der MII (HL7 FHIR) sowie der GECCO-Datensatz (HL7 FHIR) der NUM-Codex-Plattform exportiert.
3. Sicherstellung der rechtskonformen Nutzung durch Etablierung folgender Strukturen:
 - Eine Treuhandstelle, die zuständig ist für
 - einen Broad Consent als Rechtsgrundlage für die Erhebung und Bereitstellung des Kerndatensatz der MII sowie des GECCO-Datensatzes und dessen Verwaltung inkl. Widerrufsmöglichkeit (Consent Management)
 - die Deidentifizierung der Daten für Forschungsfragestellungen (Pseudonymisierung bzw. Anonymisierung)
 - Ein Use & Access Committee zur Entscheidung über Datennutzungsanträge
 - Eine Transferstelle zur Koordination und Bearbeitung von Datennutzungsanträgen sowie Datenbereitstellung
4. Das MeDIC erbringt seine Leistungen für die klinischen Partner des UKHDs und dessen Tochterunternehmen sowie externe Partner im Rahmen von Verbundforschungsprojekten.

Dazu wurden innerhalb des MeDICs folgende Rollen definiert:

- MeDIC Leitung
- Datenschutzverantwortlicher
- Qualitätsmanagementverantwortlicher
- IT-Sicherheitsverantwortlicher
- ETL – Team
- Datentransferstelle

2.2.3 Architektur, Schnittstellen und Datenschutz

Die Architektur des MeDICs ist in fünf aufeinander aufbauende Schichten aufgeteilt: Service Layer, Access Layer, Data Layer, Integration Layer und Source Layer.

Service Layer

Der Service Layer stellt Dienste zur Unterstützung von Versorgung, Qualitätssicherung und Forschungsunterstützung zur Verfügung. Dazu gehören unter anderem die Services der Datentransferstelle für lokale wie auch für MII übergreifende Datennutzungsanfragen einschließlich der Beantwortung von Machbarkeitsanfragen. Dazu ist das MeDIC über das Data Sharing Framework an das Forschungsdatenportal für Gesundheit angeschlossen.

Access Layer

Der Access Layer dient dem Zugriffsschutz auf die im darunterliegenden Data Layer gespeicherten Daten. Er lässt sich in die beiden Blöcke Datenzugriffsmanagement und Interne Treuhandstelle aufteilen. Die Systeme der internen Treuhandstelle stellen PatientenIDs-, Pseudonyme, demographische Daten und die erhobenen Consente zur Verfügung. Das Datenzugriffsmanagement sorgt dafür, dass niemand unberechtigter Weise auf die Daten zugreifen darf.

Data Layer

Der Data Layer dient zur Speicherung der Daten. Der Datenspeicher beinhaltet verschiedene Speicher (einen Simple Storage Service (S3-Speicher), einen ElasticSearch Service/Speicher, eine rel. DB, eine BiosignalDB, ein FHIR Clinical Data Repository (CDR), sowie ein openEHR (Better Platform) Repository, welche medizinische Daten in verschiedenen Bearbeitungsschritten und somit verschiedener Qualität enthalten. Der S3-Speicher und der ElasticSearch Service/Speicher stellen dabei die unterste Stufe (unstrukturierte bzw. semi-strukturierte Daten) dieser Datenhaltung dar. Im Gegensatz dazu beinhalten das openEHR (Better Platform) Repository und das FHIR CDR die Daten mit der höchsten Qualität (strukturierte Daten). Diese Datenspeicher bilden zusammen den Medical Data Lake des MeDIC. Zusätzlich befinden sich im Data Layer das Consent Repository, welches die vom Patienten gegebenen Einwilligungen speichert, sowie ein Document Repository, das

die dokumentenorientierte Ablage medizinischer Inhalte ermöglicht. Zusammengefasst werden diese beiden Repositories über eine sogenannte Document Registry, die die Metadaten zu den in den Repositories gespeicherten medizinischen Dokumenten speichert. Document Registry und Consent Repository stellen mit dem Policy Decision Point (PDP) und dem Policy Administration Point (PAP) zwei weitere Komponenten bereit, die im Rahmen des Autorisierungs- und Zugriffsmanagements zum Einsatz kommen. Der PAP verwaltet Zugriffsberechtigungs-Policies, die der PDP zur Bewertung von Zugriffsanfragen heranzieht, um darauf basierend eine Zugriffsentscheidung treffen zu können. Die getroffene Entscheidung wird dann zum PEP kommuniziert, der sie entsprechend durchsetzen soll.

Integration Layer

Der Integration Layer stellt die Schnittstelle zwischen Source- und Data Layer dar. Über einen dreistufigen Prozess werden die relevanten Daten aus den Primärsystemen extrahiert, durch Transformationen an neue Datenschemata angepasst und schließlich, basierend auf der Qualität der Daten, in eines der Ziel-Repositories S3-Speicher, ElasticSearch, rel. DB oder openEHR (Better Platform) geladen. Die Softwaresysteme, welche hier zum Einsatz kommen, sind Talend und Apache Kafka in Kombination mit Apache Flink.

Source Layer

Der Source Layer stellt die Datengrundlage für das MeDIC dar. Er beinhaltet verschiedene Primärsysteme, welche für die im Versorgungskontext angelegten Dienste die benötigten Daten liefern. Dazu gehören unter anderem das i.s.h.med als klinisches Informationssystem, das Laborinformationssystem Swisslab, das PDMS-System Copra, das Tumordokumentationssystem Onkostar, das Archivsystem HydMedia, das Picture Archiving System (PACS), das Video- und Fotodokumentationssystem IMS, die „Digitale Kurve“ zur Dokumentation auf Station (Medikation, Fieberkurve), AiD und aisS zur Dokumentation von Medikation, die PEPA, sowie die dezentralen Kardiologiesysteme MRT-Datenbank, Herzkatheter-Datenbank, und Echokardiogramm-Datenbank.

Identitäts- und Audit Management

Alle fünf Schichten des MeDICs werden zusätzlich von Systemen für das Identitäts- und Audit Management flankiert, welche zum einen die Anforderungen an Benutzerauthentifizierung und -autorisierung abdecken und zum anderen eine Protokollierung und Überprüfung der auf den Daten innerhalb des MeDIC durchgeführten Operationen ermöglichen.

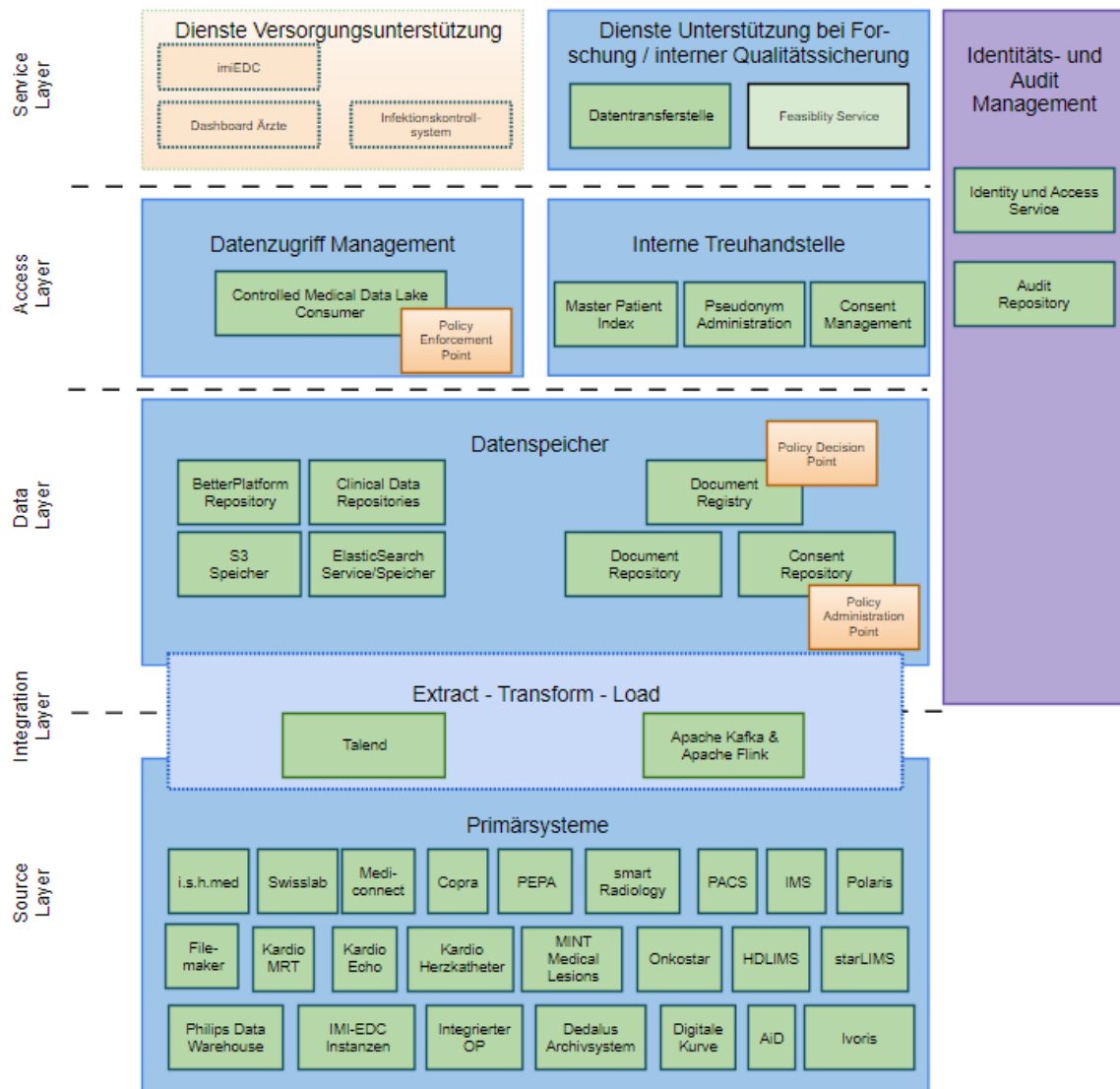


Abbildung 1: MeDIC Systemlandschaft basierend auf aufeinander aufbauenden Schichten

Das MeDIC hat über die verschiedensten Konnektoren Zugriff auf die wichtigsten Quellsysteme des Universitätsklinikums Heidelberg und verarbeitet diese je nach Quelle einmal am Tag bzw. kontinuierlich. Das Universitätsklinikum Heidelberg umfasst allerdings mehr als 250 Quellsysteme, weshalb der Ausbau der Anbindung aller Quellsysteme noch nicht abgeschlossen ist. Nach außen hin betreibt das MeDIC

vor allem die Schnittstelle zum Forschungsdatenportal für Gesundheit (FDPG), an das NUM Dashboard sowie für weitere Datennutzungsprojekte.

Für das MeDIC existiert ein von der Datenschutzbeauftragten abgesegnetes Datenschutzkonzept zur Integration von Daten aus dem Universitätsklinikum, das auch verschiedene technische – und organisatorische Maßnahmen zur Sicherung des Datenschutzes und der IT- Sicherheit beinhaltet. Alle Server sind in die regelmäßigen Backup-Prozesse des ZDIs integriert.

2.2.4 OpenEHR Clinical Data Repository

Gemäß den Konzepten des HiGHmed Konsortiums dient das OpenEHR (in unserem Fall der Firma Better) als zentraler Ort zur einheitlichen Speicherung, dem Abruf und dem Austausch der strukturierten und nach openEHR transformierten Daten. Voraussetzung dafür war die Modellierung und Standardisierung verschiedener Archetypen und Templates in den HiGHmed Modelling Groups, an denen sich das Universitätsklinikum Heidelberg inklusive des MeDIC intensiv beteiligte. Für die Modellierung wurde zunächst eine Software von Ocean Informatics später vor allem von der Firma Better genutzt. Die Archetypen und darauf aufbauenden Templates wurden in einer von den HiGHmed Partnern finanzierten Instanz des Clinical Knowledge Managers der Firma Ocean Informatics veröffentlicht. Das openEHR Repository ist für das Universitätsklinikum Heidelberg die zentrale Basis für die Transformation der Daten in FHIR.

2.2.5 FHIR Clinical Data Repository

Die Medizininformatikinitiative hat beschlossen, dass der konsortienübergreifende Austausch von Daten nicht auf der Basis von openEHR sondern von FHIR stattfinden soll. Daher wurden neben der openEHR Plattform noch weitere FHIR Clinical Data Repositories aufgebaut. Für die klinikumsinterne Nutzung der Daten wurde ein HAPI FHIR Server mit Daten von allen Patienten des Klinikums befüllt. Für die Nutzung in der MII wurde darüber hinaus ein Blaze FHIR Server etabliert, da es dieser erlaubt, auch Daten anhand der Abfragesprache CQL abzufragen. Zunächst wurde dieser Server mit Daten von Patienten befüllt, die dem Broad Consent zugestimmt haben.

2.2.6 ETL-Prozesse und Semantik

Für alle in HiGHmed Use Cases benötigten OpenEHR-Templates und auf MII-Ebene benötigten FHIR-Profile wurden die Primärsysteme mit den notwendigen Daten identifiziert und großteils als Datenquelle verfügbar gemacht. Für die Transformation der Daten wurden ETL Prozesse mit Hilfe von Talend geschrieben. Diese Software ermöglicht es, Daten direkt aus den Tabellen von SAP abzurufen.

Ein wesentlicher Bestandteil zur Sicherstellung der semantischen Interoperabilität ist neben der Nutzung gemeinsamer Archetypen oder FHIR Profile, die Nutzung einheitlicher Terminologien wie SNOMED, LOINC, ICD – 10- GM, ICD-O usw.. Gerade SNOMED und LOINC wurden vor dem Beginn der Medizininformatikinitiative kaum am Klinikum eingesetzt. Daher waren umfangreiche Mappings der hauseigenen Terminologien auf die von HiGHmed bzw. der MII vorgeschriebenen Terminologien notwendig. Das Mapping der TOP 300 Laboranalyte auf LOINC wurde extern in Auftrag gegeben und nach Review durch die zuständigen Laborärzte im Laborinformationssystem Swisslab hinterlegt. Weitere Laboranalyte werden nach Bedarf durch die Laborärzte gemappt. Um diese Terminologien und das Mapping auf die standardisierten Terminologien im Klinikum zentral verfügbar zu machen, wurde eine lokale Instanz des Terminologieservers der Firma Ontoserver angeschafft und aufgebaut. Dieser ist an den vom MeDIC des Universitätsklinikums Heidelberg mitfinanzierten zentralen HiGHmed Terminologieserver angebunden. Durch die Anbindung an den zentralen HiGHmed Terminologieserver können die von HiGHmed und der Medizininformatik genutzten Terminologien (v.a. ICD – 10 – GM, OPS, LOINC und SNOMED), die zum Teil mehrmals im Jahr aktualisiert werden, auch aktuell gehalten werden.

2.2.7 Mitentwicklung, Anbindung und Nutzung des Data Sharing Frameworks

Das Data Sharing Framework (DSF) ist ein wesentlicher Bestandteil zum Austausch von Daten über verschiedene Standorte hinweg. Es wird inzwischen nicht nur für HiGHmed, sondern von allen Konsortien genutzt. Mitarbeiter des Heidelberger MeDICs waren an dessen Konzeption und Entwicklung erheblich beteiligt.

Das MeDIC des UKHD betreibt einen lokalen DSF Knoten, der Zugriff auf den lokalen MII – FHIR Server und demographische Daten hat. Damit keine unberechtigten Systeme Daten aus dem MeDIC abrufen können, sind Komponenten zur Wahrung der IT-Sicherheit wesentlicher Bestandteil. Dazu gehören die gegenseitige Authentifizierung über Mutual TLS an Hand von Zertifikaten (X.509 Zertifikate des DFN

Typ Webserver und Typ 802.1X Client). Zusätzlich kommt eine administrierte und an die Organisationen synchronisierte Allow-List zum Einsatz, in welcher die am DSF Netzwerk teilnehmenden Organisationen hinterlegt sind. Aus Gründen der IT-Sicherheit und zur Erweiterung der Funktionalität wird das DSF ständig weiterentwickelt und die lokale Installation angepasst.

2.2.8 Record Linkage und Pseudonymisierung

Im Rahmen von HiGHmed wurde am UKHD vom ZDI eigens ein sogenannter Master Patient Index eingeführt. Dieser ist in der Lage, verschiedene Identitäten des Patienten auf einander abzubilden und eine zentrale ID zur Verfügung stellen. Darüber hinaus bietet er die Möglichkeit, die demographischen Daten getrennt von den übrigen Daten zu halten. Die Verwaltung des Klinikums ist für die Zusammenführung von Patienten-Dubletten zuständig.

Vor der Nutzung von Patientendaten zu Forschungszwecken ist eine Pseudonymisierung der Daten notwendig. Zur Generierung und Verwaltung der Pseudonyme wird am UKHD das Greifswalder Tool gPAS eingesetzt. Zur eigentlichen Pseudonymisierung wurden Skripte entwickelt.

Für das standortübergreifende Record Linkage und die standortübergreifende Pseudonymisierung haben die Universität Greifswald und die Hochschule Heilbronn im Rahmen von NUM-RDP ein gemeinsames Konzept entwickelt und dort erstmalig eingesetzt. Die Universität Greifswald spielt hierbei die Rolle einer föderierten Trusted Third Party, die das Record Linkage standortübergreifend vornimmt. Die datenschutzkonforme standortübergreifende Pseudonymisierung wird hierbei durch den NUM Transfer Hub aus Heilbronn durchgeführt. Dieser Prozess soll auch für die MII genutzt werden.

2.2.9 Qualitätsmanagement

Im Rahmen des Aufbaus und der Weiterentwicklung der MeDICs wurden etwa dreißig SOPs erstellt, um die Qualität der vom MeDIC erbrachten Leistungen (Erstellung und Betrieb von ETL – Strecken, Bearbeitung von Datennutzungsanfragen) und zur Erhebung, Erfassung und Verarbeitung (inkl. Widerruf) des Broad Consents sicherzustellen. Derzeit wird eine ISO 9001 Zertifizierung des gesamten Instituts angestrebt.

Das Monitoring der ETL Strecken erfolgt über ein eigenes Dashboard, das wichtige Kennzahlen der jeweiligen ETL – Strecken (wie ältestes Datum, neuestes Datum, Anzahl der vorhandenen Werte) kontinuierlich präsentiert. Darüber hinaus wurde ein Feedbackmechanismus etabliert, der die dokumentierende Organisationseinheit über fehlerhaft dokumentierte Daten informiert, damit sie dort berichtigt werden. Darüber hinaus gibt es SOPs zur Einarbeitung und Qualifizierung von Mitarbeitern.

Das MeDIC wurde 2021 im Rahmen der Medizininformatikinitiative erfolgreich von PwC vor Ort auditiert. Hierbei wurden 54 von 61 geprüften Auditkriterien als erfüllt bewertet (89%). Daher wurde dem MeDIC ein „weit fortgeschrittener Stand des Aufbaus des DIZ“ bescheinigt.

Vom Forschungsdatenportal für Gesundheit (FDPG) wird die Anbindung des MeDICs an das FDPG gemonitort. Das MeDIC Heidelberg war der erste angebundene Standort. Alle Pflichtressourcen sind im angeschlossenen FHIR Server verfügbar.

2.3 Weitere Strukturen und Komponenten

2.3.1 Use & Access Committee

Am Universitätsklinikum wurde im Rahmen von HiGHmed ein Use and Access Committee (UAC) etabliert. Seit der Gründung wurde die Geschäftsordnung zweimal angepasst, um die Zuständigkeit des Use and Access Committees zu erweitern. Es ist eine Einrichtung des UKHD und wird organisatorisch direkt dem Klinikumsvorstand zugeordnet. Das UAC entscheidet am UKHD in Vertretung des Klinikumsvorstandes über Herausgabe und Nutzung von (Patienten-) Daten, welche im MeDIC des UKHD und seiner Tochterunternehmen zusammengetragen und gespeichert werden. In der Umsetzung seiner Ziele arbeitet es eng mit den bestehenden Kliniken, den Medizinischen Versorgungszentren, Praxen, Therapiezentren und deren Abteilungen des UKHDs und der Tochterunternehmen zusammen.

Das UAC ist zuständig für die Begutachtung, qualifizierte Prüfung nach organisatorischen, (datenschutz-) rechtlichen, ethischen und wissenschaftlichen Aspekten, Genehmigung oder Ablehnung von Datennutzungsanträgen, die an das MeDIC des UKHD und seiner Tochterunternehmen gestellt werden. Anträge auf Nutzung und/oder Herausgabe von Biomaterialien, die an das UAC oder das MeDIC gestellt werden, leitet das UAC zuständigkeitshalber an die BioMaterialBank Heidelberg (BMBH) oder ggf. andere zuständige Biomaterialbanken weiter.

Personell besteht das UAC aus folgenden stimmberechtigten Vertretern:

- (1) Der/die Leiter/in des MeDIC (gleichzeitig LeiterIn) des UACs
- (2) Die CMO Vertretung des Klinikums (LÄD oder dessen Vertreter)
- (3) ein/e Vertreter/in des Dekanats der Medizinischen Fakultät Heidelberg oder ein/e von diesem entsandte/r Beauftragte/r und
- (4) Ein/e Vertreter/in aus Biometrie / Epidemiologie
- (5) der/die behördliche Datenschutzbeauftragte (DSB) oder ein/e Stellvertreter/in aus dem Datenschutzteam des UKHDs und seiner Tochterunternehmen
- (6) Ein/e Vertreter/in der Ethik (nicht Mitglied der Ethikkommission)
- (7) Patientenvertreter/in
- (8) Der/ Die lokale(r) MII Standortsprecher/in

Auf einer eigens eingerichteten Homepage wird die Zusammensetzung des UACs transparent veröffentlicht.

Das UAC wird nach außen vom Vorstand des UKHD vertreten.

Zu den Aufgaben des MeDIC Leiters gehören:

1. die Planung der Sitzungstermine für i. d. R. ein Jahr im Voraus
2. die Organisation und Durchführung der Sitzungen
3. die Verantwortung für das Sitzungsprotokoll
4. die jährliche Berichterstattung an den Klinikumsvorstand jeweils bis zum 30.06. des Folgejahres

Darüber hinaus ist das MeDIC für die Koordination der Datennutzungsverträgen mit den Datennutzern (teilweise über das FDPG) und der Rechtsabteilung verantwortlich. Zwischen Dezember 2021 und dem 31.12.2023 hat sich das UAC 13 Mal getroffen, um über Datennutzungsanträge zu entscheiden.

2.3.2 Datennutzungsanträge

Am MeDIC des Universitätsklinikums Heidelberg wurde ein Prozess zur Beantragung und Bearbeitung von Datennutzungsprojekten etabliert.

Datennutzungsanträge können sowohl über das Forschungsdatenportal für Gesundheit als auch über lokale Prozesse beantragt werden.

Bis Ende 2023 sind beim MeDIC des UKHD 81 Datennutzungsanträge eingegangen, davon 18 von extern. Die Tendenz der Datennutzungsanträge ist stark ansteigend. Auf dem Transparenzportal der MI-I ist das Universitätsklinikum bei zwölf Projekten als Datenlieferant angegeben. Von den 66 genehmigten Datennutzungsanträge wurden

an 53 Datennutzungsanträge Daten geliefert. Bei neun Anfragen sind die Bedingungen zum Datentransfer noch nicht erfüllt.

2.3.3 Consent Management

Da es am Universitätsklinikum keine zentrale Patientenaufnahme gibt, ist der Aufklärungsprozess zum Broad Consent sehr aufwändig.

Am Universitätsklinikum Heidelberg wurde der Broad Consent der MII zunächst in der Version 1.6f mit dem retrospektiven Modul aus 1.7 aber ohne Biobanken- und Krankenkassenmodul in der Ambulanz der Radioonkologie pilotiert. Seit dem 10.03.2021 werden in der Radioonkologie durch eine vom UKHD finanzierte Study Nurse Patienten über den Broad Consent aufgeklärt und ihre Einwilligung papierbasiert erhoben. Für die elektronische Erfassung des Broad Consents wurde von der Firma ICW ein Tool zur Erfassung des Broad Consents entwickelt. Dieses Tool ist nicht nur in der Lage den Broad Consent zu erfassen, sondern aus den Eingabemasken heraus auch basierend auf dem XACML Standard maschinell verarbeitbare Regeln zur Zugriffssteuerung abzuleiten. Dazu wurden vom MeDIC Policies für die Zugriffsregeln erstellt. Bis Ende 2023 wurden mit diesem Tool mehr als 10 000 Willensäußerungen der Patienten zum Broad Consent erfasst. Nach der elektronischen Erfassung der Einwilligung werden die Einwilligungsdokumente gescannt und archiviert. Um den Aufklärungsprozess zu unterstützen, wurde im IS-H vom Team des Zentrums für Digitalisierung eine Ampel eingebaut, die angibt, ob der Patient bereits seinen Willen zum Broad Consent geäußert hat. Diese Ampel umfasst drei Farben: Grün bedeutet, dass eine Willensäußerung des Patienten vorliegt, die jünger als fünf Jahre alt ist. Gelb bedeutet, dass eine Willensäußerung vorliegt, die aber schon älter als fünf Jahre ist. Damit neue Datenelemente im MeDIC erfasst werden können, muss der Patient neu aufgeklärt werden. Eine rote Ampel bedeutet, dass der Patient noch nie aufgeklärt wurde.

Daneben wurde am Klinikum eine Treuhandstelle aufgebaut, die die Pseudonyme der Broad Consent Patienten verwaltet und Widerrufe verarbeitet.

Während der Pilotierung in der radiologischen Ambulanz wurde eine Reihe von SOPs geschrieben, die die Consenterhebung, - erfassung aber auch Widerruf umfassen. Diese wurden kontinuierlich verbessert und dienen bei der Ausrollung auf andere Stationen als Vorgaben für eine einheitliche Prozessgestaltung. Bis Ende Dezember

2023 wurde die Broad Consenterfassung in weitere Kliniken ausgerollt. Außerdem wurde im Rahmen der Ausrollung auch Erfahrung mit der Erfassung des Consents auf Stationen statt Ambulanzen gesammelt.

Die Möglichkeit, den Consent über ein Patientenportal zu erheben, wird in 2024 in der Radioonkologie erprobt.

Darüber hinaus wurde in Kooperation mit ABIDE basierend auf dem MII Broad Consent 1.7.2 eine zweite Version des Heidelberger Broad Consents entworfen, in der das Biobankenmodul und das Krankenkassenmodul hinzugefügt wurden.

Damit die Consente auch vom FDPG Portal abgerufen werden können, wurde eine ETL – Strecke für die Transformation in FHIR erstellt.

2.3.4 Verteiltes Rechnen

Über das Forschungsdatenportal für Gesundheit können neben patientenbezogenen Daten auf Grundlage des Broad Consents auch Analysen der lokalen Daten beantragt werden. Man spricht hier im Rahmen der MII von verteiltem Rechnen. Auch das MeDIC Heidelberg hat sich an diesen Analysen beteiligt. Sie wurden vor allem im Rahmen der Use Cases CORD und POLAR, aber auch bereits initial im Jahr 2019 in der sogenannten Demonstratorstudie durchgeführt.

2.3.5 Patientenportal

Über die PEPA des Universitätsklinikums Heidelberg können Patienten eigene Dokumente, Antworten auf Fragebögen und Sensordaten für die Routineversorgung und das MeDIC bereitstellen. Die Möglichkeit, Patienten über dieses Portal Auskunft über ihre Datenbestände zu geben, konnte allerdings nicht realisiert werden. Stattdessen hat der Patient das Recht über den Datenschutz der UKHD oder über eine direkte Anfrage an das MeDIC, Auskunft über die von ihm gespeicherten Daten anzufordern. Er kann darüber hinaus anfragen, für welche Datennutzungsprojekte seine Daten genutzt wurden.

2.3.6 Nutzung des Datenintegrationszentrums im Rahmen der MII und HiGHmed

Hackathons, Projectathons und Nutzungsprojekte

Das MeDIC des UKHDs hat sich innerhalb des HiGHmed-Konsortiums an zahlreichen Hackathons und Projectathons beteiligt. Auf diese Weise konnten gemeinsam mit anderen HiGHmed- bzw. MII-Standorten neu etablierte Infrastrukturkomponenten hinsichtlich ihrer Funktionsfähigkeit getestet werden.

Über die Anbindung an das FDPG und die erfolgreiche Beantwortung von zahlreichen internen und externen Datennutzungsanfragen, hat das MeDIC des UKHDs bewiesen, dass Datennutzungsprojekte umgesetzt werden können und es vollautomatische Machbarkeitsanfragen beantworten kann.

Nutzung für HiGHmed Use Cases

Das Universitätsklinikum Heidelberg war an allen drei HiGHmed Use Cases beteiligt. Aufgabe des MeDICs war hier vor allem die Datenintegration auf Grundlage der openEHR Templates, aber auch die Anbindung der Spezialquellsysteme. Gemeinsam mit der Kardiologie und dem ZDI wurden auch Möglichkeiten zur verbesserten strukturierten Datenerhebung geschaffen.

Use Case Infektionskontrolle

Für den Use Case Infektionskontrolle sollten die Daten vor allem so aufbereitet werden, dass Bewegungs- und Infektionsdaten im Smart Infection Control System (SmICS) verfügbar waren, um automatisiert größere Ausbrüche von Infektionen im UKHD erkennen zu können. Für den Use Case wurden ETL – Strecken für Bewegungsdaten und Mikrobiologische Daten geschrieben. Mikrobiologiedaten werden nun vom Swisslab in Form von XML Befunden über den Kommserver an den eigens vom MeDIC gebauten Restconsumer in nahezu Echtzeit als XML gesendet. Darüber hinaus hat das MeDIC direkten Zugriff auf die Swisslab Statistik Datenbank erhalten, um auch ältere Mikrobiologiebefunde extrahieren zu können. Der direkte Zugriff auf ISH/i.s.h.med Tabellen hat sich in Bezug auf die Verarbeitung von Bewegungsdaten als einfacher herausgestellt als die Verarbeitung der HL7-ADT Ströme über den Kommunikationsserver. Alle empfangenen oder abgerufenen Daten werden im Staging Bereich des MeDICs zwischengespeichert und dann an Hand von weiteren ETL -Strecken in das openEHR Format transformiert und in die Betterplattform nach der Validierung der PatientenID hochgeladen. Leider stand keine geeignete Dokumentation für die Quellsysteme (vor allem Mikrobiologiebefunde) zur Verfügung. Daher war die Analyse der Quellformate sehr zeitintensiv.

Für die Entwicklung der ETL-Strecken waren neben den Mappings der Datenstrukturen auch Terminologiemappings notwendig. Diese wurden dem MeDIC vom Use Case Infection Control bereitgestellt. Das Mapping der Erregernamen von UKHD spezifischen Codes auf SNOMED Codes umfasst 1579 Einträge. Die meisten Mikrobiologiedaten und Bewegungsdaten konnten hiermit erfolgreich verarbeitet werden.

Darüber hinaus hat das MeDIC das SMICS installiert und an die lokale openEHR Plattform angebunden. Im SmICS werden die Daten per AQL (Archetype Query Language) aus dem openEHR-Repository des MeDICs abgefragt. Das SMICS wendet darauf Cluster-Analysealgorithmen an und gibt die aufbereiteten Daten an die Visualisierungsapplikation weiter.

Im Gegensatz zu den anderen Kliniken wurde in Heidelberg das SMICS an Hand des Zugriffs auf den Komplettdatenbestand des Klinikums durchgeführt. Dies führte allerdings u.a. wegen der hohen Anfragelast dazu, dass das System viele Fehlermeldungen generierte. Diese konnten bis zum Ende der Förderung trotz der Verbesserungen des SMICS durch das SMICS Team und Verbesserungen der lokalen ETL – Strecken durch das MeDIC nicht vollkommen behoben werden.

Use Case Kardiologie

Wie oben bereits erwähnt, wurden gemeinsam mit der Kardiologie und dem ZDI Möglichkeiten zur verbesserten strukturierten Datenerhebung in Form eines neuen strukturierten dynamischen Verlaufsdokuments (dPDF) geschaffen. Darüber hinaus wurden umfangreiche Änderungen an den Quellsystemen der kardiologischen Spezialsysteme (MRT, Echo, Herzkatheter) vorgenommen und ETL – Strecken sowohl in das MeDIC als auch intern im MeDIC etabliert. Die Daten aus Echo, dem Studiencockpit und dem dPDF laufen zuverlässig in das MeDIC und werden dort automatisch in openEHR (Studienteilnahmen von Patienten, Anamnese, Medikation, Follow up von kardiologischen Patienten) transformiert.

Im Rahmen der Sensorikstudie wurden zwei verschiedene Apps zur Erfassung von Patient Reported Outcomes (Lebensqualität, täglicher Fragebogen, Sensorikdaten) am Klinikum pilotiert und intensiv getestet. Die Datenübertragung der ersten App erfolgte direkt von den Smartphones an die PEPA. Darüber hinaus wurden ETL – Strecken, die diese Daten aus der PEPA in das MeDIC transferieren, etabliert. Dort wurde auch die Transformation in openEHR erfolgreich getestet. Zuletzt lief die zweite

(medizinproduktkonforme) App fehlerfrei, konnte aber aus Gründen des Datenschutzes nicht produktiv genommen werden.

Use Case Oncology

Am Universitätsklinikum Heidelberg liegen nur wenige onkologische Daten strukturiert vor. Daher wurde beschlossen, qualitativ höherwertige Daten vor allem aus dem Krebsregister zu übernehmen. Dazu wurde vom Krebsregister eine Restschnittstelle für Onkostar beauftragt und die Restschnittstelle des MeDICs so angepasst, dass sie Dokumente aus Onkostar empfangen kann. In Kooperation mit den anderen Standorten konnte das Konzept für die Patientenmanagement-Software erstellt werden.

Nutzung für weitere Projekte

Das MeDIC wurde für weitere lokale Projekte und Use Cases der MII genutzt.

In das MeDIC konnten über weitere Projekte weitere Datenquellen integriert werden:

- Hochfrequente Daten aus dem Intensivmonitoring (KoMED)
- OP-Videos und Informationen über OP-Geräte (PART; Surgomics)
- Covid Spezialdokumentation (NUM – CODEX, NUM - RDP)
- OrphaCodes für seltene Erkrankungen (CORD)

2.4 Wichtigste Positionen des zahlenmäßigen Nachweises

Der größte Anteil der Fördergelder auf das Förderkennzeichen insgesamt bezogen, wurde für die Finanzierung des Personals verwendet.

Die Unterstützung durch Eurice im Aufbau der Projektstrukturen und des Projektmanagements zu Beginn des Projektes stellt die größte Einzelposition dar.

Für den Aufbau und die Weiterentwicklung der technischen Infrastruktur wurden diverse Softwarelizenzen und Server über das Projektbudget finanziert, wie z.B. die openEHR Software von Better, die Clinical Knowledge Management Software und der Ontologie-Server von CSIRO

Das für den in HiGHmed genutzten Standard openEHR benötigte openEHR Clinical Data Repository stellt eine bedeutende Ausgabe dar.

Die Zertifizierung der App aus dem Use Case Kardiologie als Medizinprodukt stellt die drittgrößte Position dar.

2.5 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die geleisteten Projektarbeiten waren sinnvoll und notwendig, um Daten aus unterschiedlichen Quellsystemen zu identifizieren, standortübergreifend und in vereinheitlichter Form sowohl für die Versorgung als auch für Forschungsprojekte zur Verfügung zu stellen. Der Erfolg der Datenharmonisierung und des konsortiumweiten Austauschs wurde durch die Verwendung der Daten in diversen Use Cases demonstriert und bereits durch Publikationen belegt. Die geschaffene technische Infrastruktur wird zur Vernetzung mit Uniklinika auf konsortialer und MII-Ebene und zur Beteiligung an Datennutzungsanfragen zu Forschungszwecken benutzt.

Darüber hinaus wurden am UKHD dafür notwendige organisatorische Instanzen wie Treuhandstelle und Use and Access Committee neu etabliert.

2.6 Verwertbarkeit der Ergebnisse im Sinne des fortgeschriebenen Verwertungsplans

Das Datenintegrationszentrum am UKHD wurde mit flexibler institutionsübergreifender Rahmenarchitektur aufgebaut und basiert auf internationalen und offenen Standards. Zudem fördern offene Standards die Auswahl zwischen technischen Lösungen und somit den Wettbewerb am IT-Markt. Das Patientenportal konnte nur teilweise

umgesetzt werden. Dieses wird aber in der nächsten Förderphase weiter vorangetrieben.

Alle openEHR Templates sind der Öffentlichkeit zugänglich. Die unter der Leitung von Heidelberg entwickelte technische Spezifikation wurde den anderen Standorten und der AG Rollout zur Verfügung gestellt. Damit konnten diese von der Erfahrung Heidelbergs profitieren. Heidelberg hat allen UCC Standorten die Möglichkeit gegeben, die Sensorikapp zu testen. Das Data Sharing Framework ist frei verfügbar.

Die Verwendung der auf offenen Standards basierenden HiGHmed Architekturvorlage zum Aufbau des Datenintegrationszentrums in Heidelberg dient dem nachhaltigen Betrieb des MeDICs. Durch offene Standards können die verschiedenen medizinischen und biologischen Daten mit anderen Standorten und Partnern geteilt werden und Ergebnisse können in die klinische Forschung und Krankenversorgung zurückfließen. Durch die verbesserte Datenerhebung und -integration und die Bereitstellung von Daten für die Forschungsprojekte können die Patienten und das Gesundheitswesen stark profitieren.

Die erzielten Ergebnisse werden in die Lehre aufgenommen, um den wissenschaftlichen Nachwuchs zu fördern. Ergebnisse wurden auf nationalen und internationalen Konferenzen und Kongressen präsentiert und diskutiert und in Fachjournalen veröffentlicht.

Die neuen technischen Entwicklungen und Erkenntnisse werden dem UKHD bei der Planung von Qualifizierungs- und Personalentwicklungsmaßnahmen helfen.

2.7 Während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen

Derzeit existieren zahlreiche Projekte mit dem Ziel, Patientendaten elektronisch, langfristig und austauschbar zu speichern, diese sowohl zur Versorgung als auch für die Forschung zu nutzen und die Vernetzung beteiligter Kliniken und Forschungseinrichtungen voran zu treiben.

Eine Nutzung des Datenintegrationszentrums für andere Projekte wie beispielsweise die bwHealthCloud oder die Anbindung an AKTIN wurden daher aktiv vorangetrieben.

2.8 Erfolgte Veröffentlichungen der Ergebnisse

- Aguduri L S, Merzweiler A, Yüsekogul N, Meyer N, Brandner A, Heinze O. „Modeling Clinical Data Transformation for a Medical Data Integration Center: An openEHR Approach” doi: 10.3205/19gmids161
- Yüsekogul N, Meyer N, Aguduri L S, Merzweiler A, Heinze O. “Implementing ETL-processes for a Medical Data Integration Center - First experiences from the Heidelberg University Hospital” doi: 10.3205/19gmids169
- Bild R, Bialke M, Buckow K, Ganslandt T, Ihrig K, Jahns R, Merzweiler A, Roschka S, Schreiweis B, Stäubert S, Zenker S, Prasser F, “Towards a comprehensive and interoperable representation of consent-based data usage permissions in the German medical informatics initiative”. doi: 10.1186/s12911-020-01138-6
- Wettstein R, Merzweiler A, Klass M, Heinze O, „Using openEHR in XDS.b Environments – Opportunities and Challenges” (). doi: 10.3233/SHTI200554
- Wettstein R, Hund H, Fegeler C, Heinze O. “Data Sharing in Distributed Architectures - Concept and Implementation in HiGHmed”, Stud Health Technol Inform. 2021 Sep 21;283:111-118. doi: 10.3233/SHTI210548.
- Wettstein R, Hund H, Kobylinski I, Fegeler C, Heinze O. "Feasibility Queries in Distributed Architectures - Concept and Implementation in HiGHmed", Stud Health Technol Inform. 2021 May 24;278:134-141. doi: 10.3233/SHTI210061.
- Hund H, Wettstein R, Heidt CM, Fegeler C. "Executing Distributed Healthcare and Research Processes - The HiGHmed Data Sharing Framework", Stud Health Technol Inform. 2021 May 24;278:126-133. doi: 10.3233/SHTI210060.
- Wulff A, Baier C, Ballout S, Tute E, Sommer KK, Kaase M, Sargeant A, Drenkhahn C, Infection Control Study Group, Schlüter D, Marschollek M, Scheithauer S. Transformation of microbiology data into a standardised data representation using OpenEHR. Sci Rep. 2021 May 18;11(1):10556. doi: 10.1038/s41598-021-89796-y PubMed PMID: 34006956
- Wettstein R, Kussel T, Hund H, Fegeler C, Dugas M, Hamacher K. Secure Multi-Party Computation Based Distributed Feasibility Queries - A HiGHmed Use Case. Stud Health Technol Inform. 2022 Aug 17;296:41-49. doi: 10.3233/SHTI220802. PMID: 36073487.
- Sommer KK, Amr A, Bavendiek U, Beierle F, Brunecker P, Dathe H, Eils J, Ertl M, Fette G, Gietzelt M, Heidecker B, Hellenkamp K, Heuschmann P, Hoos JDE, Kesztyüs T, Kerwagen F, Kindermann A, Krefting D, Landmesser U, Marschollek M, Meder B, Merzweiler A, Prasser F, Pryss R, Richter J, Schneider P, Störk S, Dieterich C. Structured, Harmonized, and Interoperable Integration of Clinical Routine Data to Compute Heart Failure Risk Scores. Life (Basel). 2022 May 18;12(5):749. doi: 10.3390/life12050749. PMID: 35629415; PMCID: PMC9147139.

3. Das Datenintegrationszentrum für „Genomics“- und „Radiomics“-Daten (OmicsDIC) des Universitätsklinikums Heidelberg

3.1. Kurzzusammenfassung

Das Datenintegrationszentrum für „Genomics“- und „Radiomics“-Daten (OmicsDIC) des Universitätsklinikums Heidelberg unterstützte mit seinem Vorhaben das HiGHmed Konsortium bei der Harmonisierung und Standardisierung von Daten und Prozessen in den Bereichen Genomics und Radiomics.

3.1.1 Aufgabenstellung

Das Universitätsklinikum Heidelberg hatte in Zusammenarbeit mit der Ruprecht-Karls-Universität den Aufbau eines Datenintegrationszentrums für „Genomics“- und „Radiomics“-Daten (OmicsDIC) zur Aufgabe. Das OmicsDIC dient dabei als Serviceeinheit für die Verarbeitung, Auswertung, Integration, Speicherung, Analyse, Visualisierung und den Datenaustausch von „omics“-Daten innerhalb des Konsortiums. Im Einzelnen verfolgte der Aufbau des OmicsDIC die folgenden Ziele:

- Bereitstellen der Infrastruktur, Services und Speicherkapazitäten für die Prozessierung von „omics“-Daten
- Bereitstellen der Infrastruktur und Services für die Datenintegration, Datenanalyse und Datenföderation
- Sicherstellen der Prozess- und Datenqualität durch „Data Stewardship“
- Etablieren der organisatorischen und technischen Prozesse für den Datenaustausch und Datenzugriff
- Bereitstellen von Expertise, Dienstleistungen und Software für interne und externe Partner durch das Roll-out Konzept
- Bereitstellen von genomischen und radiomischen Daten für die Zentrums-übergreifende Diskussion von Patienten im Use Case Onkologie
- Beteiligung an der Etablierung von virtuellen Onkologie-Zentren durch Bereitstellen von Omics-Daten und Radiomik-assoziierten Informationen sowie Anbindung an externe Datenquellen

3.1.2 Planung und Ablauf

Zum Aufbau des Omics-Datenintegrationszentrums sah der Arbeitsplan die Beteiligung des Universitätsklinikum Heidelberg an folgenden fünf Arbeitspaketen vor. Dabei wurde das Universitätsklinikum Heidelberg von der Ruprecht-Karls-Universität Heidelberg sowie zu Beginn der Förderphase von SAP, nach deren Ausscheiden von DELL als weiterem Projektpartner unterstützt.

WPB01 Projektmanagement OmicsDIC

Dieses Arbeitspaket umfasste alle erforderlichen Projektmanagementaufgaben, um das OmicsDIC aufzubauen. Diese Aufgaben beinhalteten u.a. das Zeit-, Kosten-, Ressourcen- und Risikomanagement sowie die Koordinierung der internen und externen Kommunikation und Aktivitäten.

WPB03 Infrastruktur für Datenmodellierung

Dieses Arbeitspaket diente dem Aufbau der erforderlichen Datenmodellierungsinfrastruktur einschließlich Software sowie der Definition der Modellierungsprozesse für „Genomics“- und „Radiomics“-Daten.

WPB04 Datenintegration

Dieses Arbeitspaket umfasste alle Aktivitäten und Maßnahmen, die erforderlich waren, um die definierten Daten aus den Quellsystemen zu integrieren und die Grundlage für Datenanalysen zu generieren. Das Konzept „Data Stewardship“ wurde eingeführt, um ein qualitätsgesichertes Datenmanagement zu gewährleisten.

Abweichend von der initialen Planung einer zentralen Datenhaltung von Omics-Daten wurde das Konzept während der Projektlaufzeit auf den föderierten Ansatz der Medizininformatik Initiative angepasst. Nach dem Ausscheiden von SAP wurde das Universitätsklinikum Heidelberg durch den Projektpartner DELL bei der Anpassung der Lösungen unterstützt, die eine Verarbeitung und Speicherung der „omics“-Daten am Standort erlauben.

WPB05 Datenanalytik, Visualisierung und User Interface

Ein Hauptziel der Datenintegration von „omics“-Daten ist die Wiederverwendung der Daten für weitere Analysen im Rahmen der Forschung oder Patientenversorgung. Dieses Arbeitspaket beschäftigte sich mit der Auswahl und Anpassung von Tools und der Entwicklung von Benutzeroberflächen, um weitergehende Analysen und Visualisierung von „omics“-Daten zu ermöglichen.

WPB06 Anpassungen Use Cases

Dieses Arbeitspaket diente der konkreten Unterstützung der Use Cases im HiGHmed Konsortium. Der Fokus lag hierbei auf dem Use Case Oncology, insbesondere der Unterstützung des Virtual Oncology Centers (VOC). Die weiteren Use Cases wurden hinsichtlich der Verwendung von „omics“ Technologien beraten und unterstützt.

Hauptschwerpunkt des Arbeitspaketes war die weitergehende Automatisierung und funktionale Weiterentwicklung der „omics“-Pipelines mit dem Ziel, eine optimale Patientenversorgung zu gewährleisten und aktuelle Entwicklungen im Bereich Next Generation Sequencing und Radiomics zu berücksichtigen. Zusätzlich beinhaltete das Arbeitspaket die Datenmodellierung der „omics“-Daten in enger Zusammenarbeit mit den Use Case Experten sowie die weitergehende Standardisierung und Harmonisierung basierend auf den Standards openEHR und FHIR.

Die folgende Abbildung zeigt die initiale vorhabenbezogene Ressourcen-, Meilenstein- und Zeitplanung sowie die tatsächliche Realisierung.

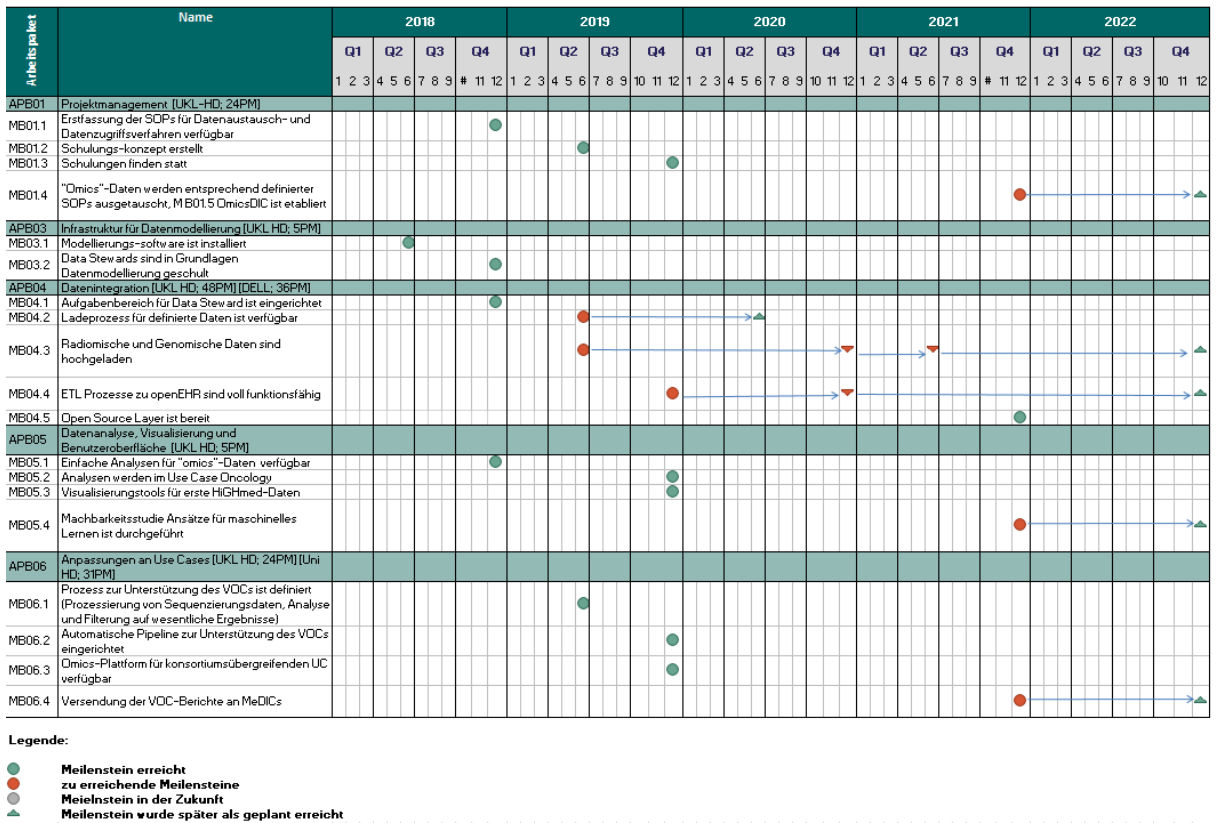


Abbildung 3: Meilensteinplan OmicsDIC

Beim Aufbau des OmicsDIC kam es bedingt durch die Corona-Pandemie zu Verzögerungen. Der verschobene Projektstart des neuen Partners DELL, Abhängigkeiten zu verzögerten Arbeitspaketen im UC Onkologie sowie der vorübergehende Einsatz personeller MII-Ressourcen im NUM CODEX-Projekt hatten direkte Auswirkungen auf die initiale Zeitplanung und Meilensteine mussten dementsprechend verschoben werden. Mit Abschluss der Förderphase konnten alle Meilensteine erreicht werden.

3.1.3 Stand der Wissenschaft und Technik

Zu Projektbeginn hatte die molekulare Diagnostik auf Basis von Genomsequenzierungen bereits durch den zunehmenden technischen Fortschritt weiter an Bedeutung gewonnen. In der Onkologie ist die Präzisionsmedizin ein wichtiges Beispiel dafür, wie Molecular-Profiling-Technologien zu einer verbesserten Diagnostik und damit zu einer besseren Patientenversorgung führen können. Ziel dieser Translation ist es, die in wissenschaftlichen Forschungsprojekten erworbenen Erkenntnisse schneller in die klinische Anwendung zu übertragen. Das Universitätsklinikum Heidelberg

hatte in Zusammenarbeit mit der Ruprecht-Karls-Universität und dem DKFZ zum Zeitpunkt des Projektstarts bereits eine erste Version einer Automatisierungsplattform für die Verarbeitung von Sequenzierungsdaten entwickelt, auf deren Ergebnissen weitere Analysen und Entscheidungen im Molekularen Tumorboard basieren. Diese Plattform wurde im Rahmen des Forschungsvorhabens erweitert.

Analog zur molekularen Diagnostik hatte die radiologische Diagnostik zu Projektbeginn mit Radiomics-Verfahren vielversprechende Ansätze erschlossen, um neue quantitative bildbasierte Biomarker zu entwickeln und damit bildgestützte Diagnostik und Therapie zu verbessern. Trotz breiter Forschung auf dem Gebiet, stand zu Projektbeginn noch ein großer Mangel an Standardisierung und Automatisierung welches die Translation in die klinische Anwendung kommt. Diese Probleme wurden im Rahmen des Forschungsvorhabens durch die Projektarbeiten adressiert.

Der Stand von Wissenschaft und Technik auf den von diesem Vorhaben berührten Arbeitsgebieten wurde durch aktuelle Informationsrecherchen von uns vorab ermittelt. Nach unserem Kenntnisstand war das Vorhaben nicht bereits Gegenstand anderweitiger Forschungen / Entwicklungen / Untersuchungen / Patente und es standen und stehen der Projektdurchführung und einer späteren Ergebnisverwertung keinerlei Schutzrechte und Schutzrechtsanmeldungen entgegen.

3.1.4 Projektpartner

Zur Umsetzung des Vorhabens fand eine intensive Zusammenarbeit mit der Ruprecht-Karls-Universität Heidelberg und dem Projektpartner DELL statt, insbesondere in den gemeinsamen Arbeitspaket B06 bzw. B04.

3.2. Eingehende Darstellung OmicsDIC

Der Schwerpunkt des Vorhabens des OmicsDIC des Universitätsklinikums Heidelberg lag in Zusammenarbeit mit den Projektpartnern der Ruprecht-Karls-Universität Heidelberg und DELL in der Entwicklung und technischen Umsetzung der Konzepte zum Aufbau eines Datenintegrationszentrums für Omics-Daten. Im Projektverlauf wurde von dem initialen Konzept des Aufbaus eines zentralen Datenintegrationszentrums für Omics-Daten aus

datenschutzrechtlichen und organisatorischen Gründen Abstand genommen und zugunsten eines dezentralen Konzeptes entsprechend des förderierten Ansatzes der Medizin-Informatik Initiative weiterentwickelt.

3.2.1 Ziele und Ergebnisse

Konkret war das Universitätsklinikum Heidelberg an der Umsetzung der folgenden, in der Vorhabenbeschreibung formulierten Ziele beteiligt. Dabei wurden folgende Ergebnisse erzielt.

Ziel: Bereitstellen der Infrastruktur, Services und Speicherkapazitäten für die Prozessierung von „omics“-Daten

Ergebnis: Durch die Umstellung des Architekturkonzeptes von zentraler nach dezentraler Datenhaltung haben die Standardisierung und Automatisierung der omics-Pipelines zur Verarbeitung genomischer und radiomischer Daten, die einfach an anderen Standorten installiert und betrieben werden können (Rollout), stark an Bedeutung gewonnen und standen deshalb im Fokus der Arbeiten des OmicsDIC. Nach dem förderierten Ansatz werden die omics-Daten am jeweiligen Standort prozessiert und in das medizinische Datenintegrationszentrum (MeDIC) übertragen. Entsprechend müssen die erforderliche Infrastruktur und Speicherkapazitäten am Standort bereitgestellt werden.

Alle erforderlichen technischen Erweiterungen und Anpassungen der omics-Pipeline „OneTouchPipeline „(OTP) zur automatisierten Verarbeitung von Whole-Genome-Sequenzierungsdaten sind erfolgt, um WGS-Sequenzierungsdaten nach aktuellen Erkenntnissen standardisiert prozessieren und analysieren zu können. Der Funktionsumfang wurde erweitert, sodass OTP jetzt zusätzlich auch die Verarbeitung von Single Cell Daten unterstützt. Der Hauptschwerpunkt der Entwicklungen während der Projektlaufzeit lag auf einer weitergehenden Standardisierung und Modularisierung der Prozesse. In Zusammenarbeit mit dem Projektpartner Dell wurde die Umstellung der internen Workflowsysteme auf Workflows nach GA4GH Workflow Execution Standard (<https://github.com/ga4gh/workflow-execution-service-schemas>) abgeschlossen. Dadurch lassen sich jetzt gängige Workflow Management Systeme wie Snakemake (<https://snakemake.github.io/>) und Nextflow (<https://nextflow.io/>) einfach in die automatische

Datenverarbeitung mit OTP integrieren. Durch Refaktorisierung und Modernisierung von Legacy-Komponenten wurde eine größere Flexibilisierung von OTP erreicht. Dies erlaubt z.B. das einfache Hinzufügen neuer Sequenzierungsverfahren. Für das Software Development Life-Cycle Management von OTP wurde eine Continuous Integration Pipeline aufgebaut sowie der Deployment-Prozess überarbeitet. Hier hat uns der Projektpartner Dell stark unterstützt, insbesondere was das Workflow Management in OTP betrifft.

Die OmicsDIC Pipeline-Software OTP ist open-source und unter folgendem Link zugänglich: <https://gitlab.com/one-touch-pipeline/otp>. OTP wird auch nach Projektende kontinuierlich in einem agilen Entwicklungsprozess an neue Anforderungen angepasst und optimiert.

Die Radiomics-Pipeline mit ihren Analyse-Workflows wurde in einem agilen Vorgehen kontinuierlich an die sich ändernden Anforderungen während des Projektverlaufs angepasst und weiterentwickelt. U.a. wurde ein Delphi-Prozess zusammen mit dem SPP Radiomics (SPP 2177) und der Experten-Community in Deutschland durchgeführt, um das Anforderungsprofil für den klinischen/translationalen Einsatz von Radiomics zu konkretisieren und einen gemeinsamen Standard für die Definition von Radiomics Workflows zu etablieren, um damit klinische Translation zu erleichtern. Dieser Consensus, sowie die gefunden Definitionskonflikte in der bestehenden Literatur und die Charakterisierung der wichtigsten Herausforderungen für die klinische Translation von Radiomics wurde in einem Peer-Reviewed.Journal veröffentlicht (Floca et al.2024; <https://doi.org/10.1186/s13244-024-01704-w>) und als Webresource (<https://github.com/MIC-DKFZ/radiomics-workflow-definition>) zur Verfügung gestellt.

Des Weiteren wurde u.a. basierend auf den Anforderungsanalysen die Radiomics-Software weiter optimiert. Diese Entwicklungen stehen mit der Software an MITK als Open Source Code unter folgendem Link zur Verfügung: <https://github.com/mitk/mitk>. MITK wird ebenfalls auch nach Projektende kontinuierlich aktualisiert und angepasst. Durch die Umstellung auf eine föderierte Datenhaltung, bei der die Daten am Standort verbleiben und dort

analysiert werden, ergaben sich zusätzliche erforderliche Softwareanpassungen zur Erleichterung des Deployments der Software an einem anderen Standort. Deshalb wurde die der Radiomics Workflow in die Medical-Imaging-Deployment-Plattform Kaapana integriert und steht unter <https://github.com/kaapana/kaapana> als Open Source Code zur Verfügung.

Ziel: Bereitstellen der Infrastruktur und Services für die Datenintegration, Datenanalyse und Datenföderation

Ergebnis: Durch die Umstellung auf das MII konforme Konzept der föderierten Datenhaltung mussten die Standorte in die Lage versetzt werden, die zentral entwickelten „omics“-Pipelines zur Harmonisierung und Standardisierung der Verarbeitung von „omics“-Daten möglichst einfach installieren und in Betrieb nehmen zu können. Entsprechend standen die Entwicklung und die Umsetzung eines entsprechenden Rollout-Konzeptes im Vordergrund (siehe nachfolgendes Ziel). Nach dem föderierten Ansatz werden die omics-Daten am jeweiligen Standort prozessiert, analysiert und in das medizinische Datenintegrationszentrum (MeDIC) integriert.

Ziel: Bereitstellen von Expertise, Dienstleistungen und Software für interne und externe Partner durch das Roll-out Konzept

Ergebnis: In Zusammenarbeit mit dem Projektpartner DELL und der Ruprecht-Karls-Universität Heidelberg wurde ein Konzept entwickelt, das aufbauend auf einem neuen Prozessmodell die Integration der Software One Touch Pipeline (OTP) zur automatisierten Verarbeitung von Sequenzierungsdaten in eine Appliance-Struktur vorsah. Durch die funktionelle Einheit von Hard- und Software ist eine einfachere Installation an den Standorten möglich.

Das dezentrale Konzept, bei dem die Daten am Standort verbleiben, erwies sich auch für die Radiomics-Pipeline als Vorteil, da die Übertragung großer Bilddateien als auch die Herausforderung einer datenschutzkonformen Überlassung von Bildmaterial nicht berücksichtigt werden mussten. Das erstellte Konzept für den Roll-out der Radiomics-Pipeline konzentrierte sich deshalb auf das einfache Deployment der Workflows an einem neuen Standort. Für die Umsetzung des Rollout-Konzeptes für die OTP-Pipeline standen insbesondere die softwareseitigen Anpassungen zur Umsetzung der Appliance

im Vordergrund. In Zusammenarbeit mit dem Projektpartner DELL und der Ruprecht-Karls-Universität wurden Softwareentwicklungsarbeiten durchgeführt, die zu einer Standardisierung, Modularisierung und Flexibilisierung der Software geführt haben. Die einzelnen Arbeiten sind im Abschnitt „Ziel: Bereitstellen der Infrastruktur, Services und Speicherkapazitäten für die Prozessierung von „omics“-Daten“ detaillierter beschrieben.

Durch die Standardisierungs- und Modularisierungsarbeiten wurden softwareseitig die Grundlagen für die Appliance geschaffen. Die Pipeline-Software OTP ist eigenständig als Software-Paket in Form eines Docker-Container für Dritte verfügbar (<https://gitlab.com/one-touch-pipeline/otp>). Die Containerisierung der Software war die Voraussetzung, um OTP auch als frei installierbare Software in der Cloud zur Verfügung zu stellen und damit für eine breitere Community zugänglich zu machen. Die erforderlichen Anpassungen zur Cloudfähigkeit erfolgten mit Unterstützung von DELL im Rahmen des European Open Science Cloud (EOSC) Projektes.

Die hardwareseitige Umsetzung der Appliance durch den Projektpartner DELL war für das Jahr 2022 vorgesehen, wurde aber aufgrund der Dynamik in der Entwicklung von IT- und Digitalisierungstechnologien zugunsten der Cloudlösung verworfen. Dies hatte keinen Einfluss auf das Erreichen der Ziele und Meilensteine, da die OTP-Software als frei installierbarer Docker-Container als auch in der Cloud für Dritte zur Verfügung steht.

Die OTP-Pipeline konnte erfolgreich beim Projektpartner Charité ausgerollt werden und ist dort im produktiven Betrieb.

Zur Erleichterung des Roll-Out-Prozesses der Radiomics-Pipeline an anderen Standorten wurde die Software ebenfalls containerisiert. Um verschiedene Deployment-Szenarien zu ermöglichen, wurden die Radiomics-Workflows in die Medical-Imaging-Deployment-Plattform Kaapana (<https://github.com/kaapana/kaapana>) integriert und Kaapana hinsichtlich der Anforderungen an eine Installation in MeDICs optimiert. Das Deployment der Radiomics-Pipeline wurde in verschiedenen Szenarien (Deployment VMWare, Openstack und „bare metal“) erfolgreich getestet und steht für den Roll-Out zur Verfügung.

Ziel: Sicherstellen der Prozess- und Datenqualität durch “Data Stewardship”

Ergebnis: Die Datenmodellierung und Datenaufbereitung der genomischen und radiologischen Primärdaten erfolgte durch Data Stewards. Durch die Beteiligung von Domain Experts, die über die inhaltliche Expertise aus den Ergebnissbereichen der Daten verfügten, wurde die inhaltliche Korrektheit der Datenmodelle sichergestellt. Die technische Umsetzung der Datenmodelle erfolgte durch fachlich geschulte Data Stewards auf Basis des Interoperabilitätsstandards openEHR unter Berücksichtigung der Kompatibilität zu entsprechenden FHIR-Ressourcen. Die Qualität der interoperablen Datenmodelle wurde in mehrfachen Review-Prozessen im Clinical Knowledge Manager (CKM) unter Beteiligung von Data Stewards der Community sichergestellt. Alle entwickelten Datenmodelle sind im HiGHmed CKM (<https://ckm.highmed.org/ckm/>) oder im internationalen CKM (<https://ckm.openehr.org/ckm/>) versioniert gespeichert und öffentlich zugreifbar.

Ziel: Bereitstellen von genomischen und radiomischen Daten für die zentrumsübergreifende Diskussion von Patienten im Use Case Onkologie

Ziel: Beteiligung an der Etablierung von virtuellen Onkologie-Zentren durch Bereitstellen von Omics-Daten und Radiomik-assoziierten Informationen sowie Anbindung an externe Datenquellen

Ergebnis: Durch die Umstellung auf die föderierte Datenhaltung verschob sich der Aufgabenschwerpunkt von einer zentralen Bereitstellung der genomischen und radiomischen Daten hin zu einer Bereitstellung von interoperablen Datenmodellen für diese Datenarten als Grundlage für die einheitliche Verarbeitung und Speicherung der Daten am Standort bzw. in den Onkologie-Zentren. Die im Rahmen des Förderprojekts entwickelten standardisierten Datenmodelle werden über öffentlich zugängliche Repositorien allen Beteiligten bereitgestellt.

Für den Use Case Oncology wurden in HiGHmed in Zusammenarbeit mit internationalen Partnern Datenmodelle, sogenannte Archetypen, für genomische Daten im openEHR Standard entwickelt, im internationalen Clinical Knowledge Manager (<https://ckm.openehr.org/ckm/>) öffentlich publiziert und kontinuierlich an aktuelle Entwicklungen angepasst. Die standardisierten openEHR-Datenmodelle für genomische Daten stehen global zur Verfügung.

Diese Archetypen dienen als Grundlage, um das Virtual Oncology Center im UC Oncology bei der openEHR konformen Modellierung der Empfehlung des Molekularen Tumorboards (<https://ckm.highmed.org/ckm/templates/1246.169.2198>) und der Weiterentwicklung des molekulargenetischen Befundes (<https://ckm.highmed.org/ckm/templates/1246.169.2198>) zu unterstützen. Die Ergebnisse dieser Modellierungsarbeiten und die dabei gewonnene Expertise sind in das FHIR basierte Erweiterungsmodul Molekulargenetischer Befundbericht der TF Kerndatensatz eingeflossen. Damit wurde sichergestellt, dass die HiGHmed openEHR-Datenmodelle mit den auf FHIR basierenden Kerndatensätzen auch in Bezug auf die molekulargenetischen Daten kompatibel sind und die Interoperabilität zukünftig gewährleistet ist.

Zur Harmonisierung und Standardisierung radiomischer Daten wurden in HiGHmed entsprechende openEHR basierte Datenmodelle (Archetypen) definiert. Auf Grundlage dieser Archetypen wurden die Berichte (Templates) „Radiomics Imaging Report“ (<https://ckm.highmed.org/ckm/templates/1246.169.2315>), „Radiomics Measurement Report“ (<https://ckm.highmed.org/ckm/templates/1246.169.2317>) und „Radiomics Segmentation Report“ (<https://ckm.highmed.org/ckm/templates/1246.169.2316>) modelliert, von der Community kommentiert und anschließend finalisiert.

Alle in HiGHmed entwickelten Archetypen und Templates sind im HiGHmed Clinical Knowledge Manager (<https://ckm.highmed.org/ckm/>) veröffentlicht.

Zusätzlich wurde der UC Onkologie im Bereich Radiomics in der „Bild-gestützten Vermessung/Kategorisierung von HCC-Läsionen für Monitoring und zur Beurteilung des Therapieansprechens“ unterstützt. Mit Hilfe eines ersten Satzes an Trainingsdaten wurde die Problemstellung genauer analysiert und verschiedene State-Of-The-Art Ansätze zur DL basierten Semantischen- und Instanz-Segmentierung erprobt.

Ziel: Etablieren der organisatorischen und technischen Prozesse für den Datenaustausch und Datenzugriff

Ergebnis: Das Konzept des omicsDIC unterstützt den föderierten Ansatz der Medizininformatik-Initiative, die Daten dezentral an den Standorten in den Medizinischen Datenintegrationszentren zu halten. Unter Einbeziehung der lokalen Use & Access Committees (UAC) behalten die Standorte die Hoheit darüber selbst zu entscheiden, an welchen Forschungsvorhaben sie sich beteiligen. Über das etablierte Verfahren des Deutschen Forschungsdatenportal für Gesundheit (<https://forschen-fuer-gesundheit.de/>) können Machbarkeitsanfragen und Datennutzungsanfragen auf Basis des Broad Consent, zukünftig voraussichtlich auch auf Basis des Gesundheitsdatennutzungsgesetzes gestellt werden. Über Erweiterungsmodule des Kerndatensatzes können genomische und radiomische Informationen abgefragt und ausgetauscht werden.

3.2.2 Notwendigkeit und Angemessenheit der Projektarbeiten

Die geleisteten Projektarbeiten für das Omics-DIC waren sinnvoll und notwendig, um die erforderliche Harmonisierung und Standardisierung von Daten und Prozessen in den Bereichen Genomics und Radiomics zu erreichen. Nach Projektabschluss stehen interoperable Datenmodelle, standardisierte Softwarepipelines für die einheitliche Verarbeitung genomischer und radiomische Daten sowie entsprechende Roll-out-Konzepte für eine einfache Installation und Inbetriebnahme der Softwarelösungen an weiteren Standorten zur Verfügung. Durch die Standardisierung wurden die Grundlagen für die Wiederverwendbarkeit und den Datenaustausch der genomischen und radiomischen Daten geschaffen und damit die Voraussetzungen für eine optimale Nutzung in Versorgung und Forschung.

3.2.3 Verwendung der Zuwendung

Die Zuwendung wurde insbesondere für Personal verwendet, das für die Software-Entwicklung, Konzepterstellung, Datenmodellierung sowie für begleitendes Projektmanagement in den oben genannten Arbeitspaketen eingesetzt wurde.

3.2.4 Fortschreibung des Verwertungsplans

Es sind keine Patente und Erfindungen angemeldet bzw. erteilt worden, es wurden auch keine in Anspruch genommen. Zum Verwertungsplan der Vorhabenbeschreibung haben sich keine Veränderungen ergeben.

Die während der Projektlaufzeit entwickelten open-source Softwarelösungen basieren auf standardisierten Datenmodellen mit standardisierten Schnittstellen. Dies ermöglicht die leichte Übertragbarkeit der Lösungen auf andere Einrichtungen bzw. die Verwendung in weiteren Projekten und sichert damit die Anschlussfähigkeit des Vorhabens. Konkret haben sich hier im wissenschaftlichen Umfeld bereits Anschlussfinanzierungen durch die drei EU Projekte EUCANCan, ECPDC und InstandNGS4P ergeben. In EUCANCan und ECPDC (Cancer related EU Horizon 2020 Projekte) fanden die in HiGHmed entwickelten openEHR Datenmodelle Berücksichtigung und wurden gemeinsam weiterentwickelt.

Im InstandNGS4P Call haben die HiGHmed Projektpartner DKFZ und Charité in dem Pre Commercial Procurement Call (PCP) auch die Förderung für die finale dritte Phase erhalten, die Ausschreibung für die beiden ersten Phasen bereits gewonnen und erfolgreich abgeschlossen werden konnte. In InstandNGS4P wird die durch HiGHmed unter Beteiligung des Universitätsklinikums Heidelberg mit entwickelte und mit finanzierte One Touch Pipeline (OTP) als ein integraler Bestandteil des omicsDIC für den Bereich der Whole Genome Sequenzierung als maßgebliche Infrastruktur eingesetzt.

Im Bereich Radiomics werden die Medical-Imaging-Deployment-Plattform Kaapana (kaapana.ai) sowie weitere Radiomics-Tools und Services des OmicsDIC in nationalen und internationalen Projekten verwendet und weiterentwickelt. Zu den Projekten gehören u.a. EUCAIM (cancerimaging.eu; pan-European federated infrastructure for AI innovations in cancer images), in dem die Arbeiten in den Arbeitspaketen „Data Federation and Interoperability Framework“ und „Federated data processing and analysis“ durch die Vorarbeiten unterstützt werden; das Projekt NUM RACOON (<https://racon.network/>), in dem u.a. der Radiomics-Workflow zum Einsatz kommt; sowie der MII Modul 3 Antrag PrivateAIM, in welchem wir in der nächsten MII Förderperiode die Integration von medizinischen Bilddaten in die MII-weite Lern- und Analyse-Infrastruktur sicherstellen. Die HiGHmed

Vorarbeiten haben einen wichtigen Betrag zu dem Erfolg und der Nachnutzung in den genannten Projekten geleistet.

Durch die enge Zusammenarbeit mit dem privatwirtschaftlichen Projektpartner Dell haben sich wirtschaftliche Anschlussmöglichkeiten ergeben. Die Grundlage für die wirtschaftliche Verwertung konnte in Form eines optimierten Roll-outs der entwickelten Softwarekomponenten zu anderen Partnern und Institutionen gelegt werden.

Es sind keine ähnlichen oder vergleichbaren Ergebnisse, welche dieses Vorhaben mittelbar oder unmittelbar betreffen, an anderer Stelle bekannt.

3.2.5 Veröffentlichung der Ergebnisse

Publikationen

Borchert F, Mock A, Tomczak A, Hügel J, Alkarkoukly S, Knurr A, Volckmar AL, Stenzinger A, Schirmacher P, Debus J, Jäger D, Longerich T, Fröhling S, Eils R, Bougatf N, Sax U, Schapranow MP. Knowledge bases and software support for variant interpretation in precision oncology. *Brief Bioinform.* 2021 Nov 5;22(6):bbab246. doi: 10.1093/bib/bbab246. Erratum for: *Brief Bioinform.* 2021 Nov 5;22(6): PMID: 34125166; PMCID: PMC8575010.

Cecilia Mascia, Francesca Frexia, Paolo Uva, Gianluigi Zanetti, Luca Pireddu, Gideon Giacomelli, Christina Jaeger-Schmidt, Aurelie Tomczak, Simon Schumacher, Florian Kraecher, Roland Eils, Silje Ljosland Bakke, Heather Leslie (2020). The openEHR Genomics Project. *Studies in Health Technology and Informatics, Volume 270: Digital Personalized Health and Medicine*, p. 443 – 447. DOI 10.3233/SHTI200199

Ralf Floca, Jonas Bohn, Christian Haux et al.. Radiomics workflow definition & challenges - German priority program 2177 consensus statement on clinically applied radiomics. *Insights Imaging* 15, 124 (2024). <https://doi.org/10.1186/s13244-024-01704-w>

Software

Open Source OmicsDIC Pipeline-Software OTP (<https://gitlab.com/one-touch-pipeline/otp>)

Open Source Radiomics Analyse Tool (MITK Phenotyping;
<https://github.com/MITK/MITK>)

Open Source Deployment-Plattform für Bildanalysen (Radiomics und DNN;
<https://github.com/kaapana/kaapana>)

Datenmodelle

openEHR Datenmodell (Archetypen) für genomische Daten
(<https://ckm.openehr.org/ckm/> Project „Genomics“)

openEHR Datenmodelle für molekulargenetische Daten (Archetypen
& Templates) (<https://ckm.highmed.org/ckm/> Project „Use Case Onkologie“)

openEHR Datenmodelle (Archetypen und Templates) für DICOM Bilddaten und
Radiomics Reports (<https://ckm.highmed.org/ckm/> „Project Use Case
Onkologie“)