

HereditCaRe: Dokumentations- und IT-Lösung eines spezialisierten Registers für erblichen Brust- und Eierstockkrebs

HereditCaRe: Documentation and IT Solution of a Specialized Registry for Hereditary Breast and Ovarian Cancer

Autoren

Christoph Engel^{1*}, Kerstin Wieland^{1*}, Silke Zachariae¹, Karolin Bucksch¹, Ute Enders¹, Ulrike Schoenwiese¹, Maryam Yahiaoui-Doktor¹, Katharina Keupp², Anke Waha², Eric Hahnen², Robert Remy², Corinna Ernst², Markus Loeffler¹, Rita K. Schmutzler²

Institute

- 1 Institut für Medizinische Informatik, Statistik und Epidemiologie, Universität Leipzig, Leipzig, Deutschland
- 2 Zentrum familiärer Brust- und Eierstockkrebs, Uniklinik Köln, Köln, Deutschland

Schlüsselwörter

erblicher Brust- und Eierstockkrebs, HereditCaRe, modellhafte Register in der Versorgungsforschung, Dokumentation, IT-Lösung

Key words

hereditary breast and ovarian cancer, HereditCaRe, model registers in health services research, documentation, IT solution

Bibliografie

Gesundheitswesen 2021; 83 (Suppl. 1): S12–S17

DOI 10.1055/a-1658-0313

ISSN 0949-7013

© 2021. Thieme. All rights reserved.

Georg Thieme Verlag, Rüdigerstraße 14, 70469 Stuttgart, Germany

Korrespondenzadresse

Privatdozent Dr. med. habil. Christoph Engel
Institut für Medizinische Informatik,
Statistik und Epidemiologie
Universität Leipzig
Härtelstraße 16-18
04107 Leipzig
Deutschland
christoph.engel@imise.uni-leipzig.de

ZUSAMMENFASSUNG

Das nationale Register „HereditCaRe“ für die Evaluation und Verbesserung der risiko-adjustierten Prävention bei erblichem Brust- und Eierstockkrebs ist eines von sechs vom BMBF geförderten „modell-

haften Registern in der Versorgungsforschung“. In diesem Beitrag beschreiben und diskutieren wir die zur standardisierten Datenerfassung gewählte Dokumentations- und IT-Lösung auf der Basis der zuvor definierten speziellen funktionalen Anforderungen. Die Dokumentation gliedert sich in verschiedene patientenindividuell einzusetzende Module, die auf einem zuvor festgelegten Merkmalskatalog beruhen. Aufgrund spezieller funktionaler Anforderungen wurde eine eigene Datenerfassungsanwendung auf der Basis von ORACLE und ORACLE Forms entwickelt und implementiert. Die speziellen Anforderungen umfassten u. a. die Einbindung grafischer Stammbaumdarstellungen, den strukturierten Upload von Stammbaumdaten und molekulargenetischen Informationen, die automatisierte Altdatenübernahme aus dem Vorgängersystem, sowie die freie Programmierbarkeit von beliebig komplexen Datenbankabfragen zur zentralen Datenqualitätsprüfung. In die Anwendung ist eine Datenbank zur patienten-unabhängigen Verwaltung genetischer Risikovarianten nahtlos integriert und mit den patientenbezogenen Daten verknüpft. Die Vor- und Nachteile der gewählten IT-Lösung werden kritisch diskutiert. Insgesamt kommen wir zu der Schlussfolgerung, dass es angesichts der komplexen Dokumentation und der speziellen Funktionsanforderungen alternativ keine fertigen Softwareprodukte zu der von uns gewählten Eigenentwicklung existieren.

ABSTRACT

The national registry „HereditCaRe“ for the evaluation and improvement of risk-adjusted prevention in hereditary breast and ovarian cancer is one of six „model registries in health services research“ funded by the BMBF. In this paper, we describe and discuss the documentation and IT solution chosen for standardized data collection based on the specific functional requirements previously defined. The documentation is divided into different modules to be used individually for each patient, which are based on a previously defined catalog of documentation items. Due to special functional requirements, a specific data entry application based on ORACLE and ORACLE Forms was developed and implemented. The specific requirements included the integration of graphical pedigree representations, the structured upload of pedigree data and molecular genetic information, the automated transfer of old data from the pre-

* gleichberechtigte Erstautoren.

vios system, as well as the free programmability of complex database queries for central data quality control. A database for patient-independent management of genetic risk variants was seamlessly integrated into the application and linked to the patient-related data. The advantages and disadvantages

of the chosen IT solution are critically discussed. Overall, we come to the conclusion that, in view of the complex documentation and the special functional requirements, there are no alternative ready-made software products to the in-house development we have chosen.

Einleitung

Brustkrebs ist die häufigste Krebserkrankung der Frau mit deutschlandweit derzeit jährlich ca. 69.000 Neuerkrankungen. Ungefähr jede achte Frau ist im Laufe ihres Lebens betroffen. Verbesserte Therapiemöglichkeiten haben die Überlebenswahrscheinlichkeit in den vergangenen Jahren deutlich verbessert. Die Früherkennung bleibt aber die wichtigste Voraussetzung für eine günstigere Prognose. Etwa 15 bis 20% der Brustkrebskrankungen sind auf ein familiäres Risiko zurückzuführen. Wenn nahe Verwandte an Brust- oder Eierstockkrebs erkrankt sind, dann steigt das eigene Erkrankungsrisiko. Darüber hinaus kann bei etwa 5 bis 10% der Brustkrebsfälle eine genetische Ursache nachgewiesen werden (hereditäre Krebsdisposition). Diese beruhen größtenteils auf hoch-penetranten pathogenen Keimbahnvarianten in den Genen *BRCA1* und *BRCA2*, die in den 1990er Jahren entdeckt wurden. Seitdem wurden weitere mittel- und niedrig-penetrante Risikogen-Varianten für Brustkrebs identifiziert. Einige dieser Gen-Varianten sind auch mit einem erhöhten Risiko für Eierstockkrebs verbunden.

Das Deutsche Konsortium Familiärer Brust- und Eierstockkrebs

Die optimale klinische Betreuung von Frauen mit Verdacht auf hereditären Brust- und Eierstockkrebs erfordert die strukturierte interdisziplinäre Zusammenarbeit von Gynäkologie, Humangenetik, Pathologie, Molekularbiologie und Radiologie. Zu diesem Zweck wurde 1996 mit Unterstützung der Deutschen Krebshilfe das „Deutsche Konsortium Familiärer Brust- und Eierstockkrebs“ (DK-FBREK) ins Leben gerufen. Das DK-FBREK ist ein deutschlandweiter Verbund von spezialisierten universitären Zentren. Ziel ist es, Frauen mit einer familiären Belastung für Brustkrebs und Eierstockkrebs bundesweit optimal zu betreuen. Hierzu bietet das Konsortium ein umfassendes Versorgungskonzept an, welches in Krankenkassenverträgen zur besonderen Versorgung nach §140a SGB V abgebildet ist. Dieses beinhaltet eine umfassende Beratung von Risikofamilien, eine individuelle Risikokalkulation, eine Genanalyse aller derzeit bekannten Risikogene, sowie die Durchführung präventiver Maßnahmen (intensiviertes Früherkennungs- und Nachsorgeprogramm, prophylaktische Operationen). Das Konsortium umfasst bundesweit derzeit 23 universitäre Zentren (► **Abb. 1**).

Das DK-FBREK hat frühzeitig eine zentrale Datenbank aufgebaut, um verbundweit einheitlich definierte Daten für Forschungszwecke qualitätskontrolliert zu dokumentieren. Standort dieser Datenbank ist seit 2002 das Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE) an der Medizinischen Fakultät der Universität Leipzig. Hierzu wurden am IMISE entsprechende IT-Werkzeuge zur dezentralen Datenerfassung entwickelt und bereitgestellt. Ferner stellt das IMISE für das Konsortium seine biometrische und klinisch-epidemiologische Methodenexpertise für wissenschaftliche Datenanalysen zur Verfügung. Im Konsortium arbeiten Ärzte und Wissenschaftler bei der Umsetzung

von Forschungsergebnissen in die klinische Versorgung eng zusammen und kooperieren auf nationaler und internationaler Ebene. Im Konsortium sind eine Reihe von relevanten Publikationen entstanden [1–12]. Daneben wurden konsortiale Daten in eine Vielzahl von internationalen Kooperationen eingebracht (<https://www.konsortium-familiaerer-brustkrebs.de/forschung-publikationen/publikationen/>).

Nationales Register für erblichen Brust- und Eierstockkrebs: „HerediCaRe“

Bis April 2019 wurde die wissenschaftliche Dokumentation von der Deutschen Krebshilfe gefördert. Im Jahr 2016 initiierte das Bundesministerium für Bildung und Forschung (BMBF) ein Förderprogramm zum Aufbau modellhafter Register in der Versorgungsforschung. Das DK-FBREK entwickelte daraufhin im Rahmen einer 9-monatigen Konzeptentwicklungsphase ein detailliertes Konzept zum Aufbau eines neuen patientenorientierten Registers für erblichen Brust- und Eierstockkrebs („HerediCaRe“, Hereditary Cancer Registry), welches seit



► **Abb. 1** Universitäre Standorte des Deutschen Konsortiums Familiärer Brust- und Eierstockkrebs (* zusätzlich Standort der zentralen Datenbank).

März 2019 in einer Realisierungsphase neben 5 anderen Registern für 5 Jahre bis Februar 2024 gefördert wird (<https://www.gesundheitsforschung-bmbf.de/de/modellhafte-register-realisierungsphase-9011.php>). Die Fördermaßnahme wird durch ein Projekt (REGISVF-AP) der Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V. (TMF) engmaschig begleitet (<https://www.gesundheitsforschung-bmbf.de/de/regisvf-ap-begleitprojekt-der-fordermassnahme-register-der-versorgungsforschung-10803.php>).

„HerediCaRe“ dient der Langzeitdokumentation von genetischen und klinischen Daten aus der Routineversorgung von Familien mit einer erblichen Belastung für Brust- und Eierstockkrebs, die in einem der konsortialen Zentren oder in einem kooperierenden zertifizierten Krebszentrum betreut werden. Dokumentiert werden auf der Grundlage einer informierten schriftlichen Einwilligung u. a. Angaben zur Konstellation von Krebserkrankungen innerhalb der Familie und Informationen über das Vorliegen von Risikofaktoren für Krebs. Zusätzlich werden Daten zur genetischen Diagnostik sowie zu Früherkennungs- bzw. Nachsorgeuntersuchungen erfasst und wissenschaftlich ausgewertet. Weiterhin wird DNA zu Forschungszwecken gewonnen und zentral gelagert (Standort: Uniklinik Köln).

Da viele Fragen zur Entstehung und zum Verlauf von erblich bedingten Krebserkrankungen noch nicht beantwortet sind, ist dieses spezialisierte Register von großer Bedeutung für die zukünftige Wissensgenerierung und Verbesserung der Versorgung. Es fehlen weiterhin Daten, um die Bedeutung neuer Risikogene und die Wirksamkeit von Präventionsmaßnahmen zu überprüfen. Therapeutische und vorbeugende Maßnahmen müssen stetig weiterentwickelt werden, um sie dem individuellen Risiko der Betroffenen anzupassen und damit Über- oder Unterversorgung zu vermeiden. Die wissenschaftliche Analyse der gesammelten Registerdaten soll daher u. a. Aufschluss geben über altersspezifische Erkrankungsraten und Krankheitsverläufe unterschiedlicher genetischer Tumorsubtypen, den langfristigen klinischen Nutzen und Schaden von präventiven und therapeutischen Maßnahmen, sowie über die Güte von Versorgungs- und Lebensqualität der betroffenen Patientinnen.

Die Etablierung des Registers erfolgt in Zusammenarbeit mit zertifizierten Krebszentren, den gesetzlichen Krankenkassen, der Deutschen Krebsgesellschaft (DKG) und der Selbsthilfegruppe BRCA-Netzwerk e.V. Es ist insbesondere eine Verknüpfung mit klinischen Krebsregistern vorgesehen, um Langzeitdaten zum Krankheitsverlauf zu erhalten.

Von zentraler Bedeutung für das Register ist das elektronische Datenerfassungs- und Datenbanksystem, welches am IMISE konzipiert, realisiert, und Ende September 2020 in den Produktivbetrieb gestellt wurde. Die Nutzung dieses Systems ist für alle Konsortialzentren verpflichtend. Die speziellen inhaltlichen Anforderungen an das Dokumentationssystem, die daraus resultierende technische Umsetzung und die damit verbundenen Vor- und Nachteile sollen nachfolgend dargestellt und diskutiert werden.

Realisierung

Dokumentationskonzept

Die Dokumentation erfolgt in „HerediCaRe“ in verschiedenen inhaltlich abgegrenzten Modulen, die flexibel eingesetzt werden können. ► **Tab. 1** gibt eine Übersicht über die Dokumentationsinhalte.

► **Abb. 2** illustriert die relationalen Beziehungen zwischen den einzelnen Dokumentationsmodulen. Für jeden registrierten Patienten ist obligat das Modul „PAT“ anzulegen. Die Familienzusammengehörigkeit von mehreren Patienten wird im Modul „FAM“ abgebildet. Der beim Patienten erhobene Familienstammbaum wird im Modul „STB“ erfasst. Dies kann pro Patient gegebenenfalls mehrfach erfolgen, um zeitliche Änderungen der familienanamnestischen Informationen dokumentieren zu können. Jeder Stammbaum wiederum beinhaltet individuelle Daten über jedes einzelne Familienmitglied („MIT“). Bei Patientinnen, die über zertifizierte Brustzentren (Kooperationspartner) an das Konsortium überwiesen werden, ist anstelle des „STB“-Moduls eine Checkliste („CLK“) mit aggregierten Daten zu aufgetretenen Krebserkrankungen in der Familie zu erfassen. Darüber hinaus können für jeden Patienten Module zu aufgetretenen Primärtumoren („TUM“), durchgeführten bildgebenden Untersuchungen im intensivierten Früherkennungsprogramm („IFP“), persönlicher Gesundheit („PHQ“) sowie Erkrankungssorgen („CWS“) dokumentiert werden. Ergebnisse von molekulargenetischen Untersuchungen werden im Modul „MGU“ erfasst. Diese Module können je nach individueller Erfordernis mehrfach pro Patient erfasst werden. Das Modul „VAR“ repräsentiert personenunabhängig die Liste der jemals im Konsortium identifizierten distinkten genetischen Varianten. Jede einzelne Variante ist durch einen eindeutigen Identifikator gekennzeichnet. Die Verknüpfung zu den patientenbezogenen molekulargenetischen Informationen erfolgt im Modul „MGU“, in welchem entsprechende Verweise auf die individuell identifizierten Genvarianten erfasst werden.

Die in ► **Tab. 1** dargestellten Dokumentationsinhalte werden in den klinischen Zentren durch behandelnde Ärzte, Labormitarbeiter und ggf. weitere Dokumentationskräfte erhoben und lokal vor Ort im elektronischen Online-Dokumentationssystem des Konsortiums erfasst und zentral am IMISE gespeichert.

Anforderungen an das elektronische Dokumentationssystem

Der allgemeine Anforderungskatalog an das elektronische Dokumentationssystem für „HerediCaRe“ umfasst eine übersichtliche und intuitive Bedienbarkeit, einen möglichst geringen Installationsaufwand in den klinischen Zentren, eine leichte Skalierbarkeit hinsichtlich der künftigen Einbindung neuer Zentren, und eine leichte Erweiterung der Eingabemasken und Datenbank bei Änderungen der Dokumentationsinhalte. Weitere Anforderungen betreffen die Integration einer Dashboard-Funktion zur Generierung einfacher Übersichten über den aktuellen Stand der Dokumentation und aggregierter Patientencharakteristika, sowie die Integration eines Query-Managements mit aussagekräftigen Übersichten und unterstützenden Navigationsmöglichkeiten. Daneben soll das System dem Dokumentationspersonal in den Zentren die Selbstadministration von Benutzerkonten ermöglichen. Darüber hinaus gibt es eine Reihe von speziellen Anforderungen, die für „HerediCaRe“ von besonderer Bedeutung sind:

- 1) Das System soll die Erfassung sämtlicher distinkter genetischer Varianten erlauben, die im Rahmen der genetischen Diagnostik im Konsortium jemals identifiziert wurden. Hierbei handelt es sich um nicht-personenbezogene Daten. Neben einer einheitlichen Annotation soll insbesondere auch die standardisierte

► **Tab. 1** Dokumentationsmodule und -inhalte.

Modul	Inhalt (vereinfacht)
FAM	Verwandtschaftliche Beziehungen der erfassten Registerpatienten (und ggf. verbindender Familienmitglieder)
PAT	Obligate Basisdokumentation für Registerpatienten: Zeitpunkt und Details der Einwilligung, Geburtszeitpunkt (monatsgenau), Geschlecht, Vitalstatus, ggf. Todeszeitpunkt, Art der Zuweisung, Art der zuständigen Krankenkasse, Einschlusskriterien, Beratungen, frühere Tumor-erkrankungen und Organentfernungen, Angaben zu Gendiagnostik/Risikoberechnung/Früherkennung
STB	Daten zum Stammbaum des Registerpatienten: Zeitpunkt der Erhebung, nicht-genetische Risikofaktoren und polygener Risikoscore gemäß BOADICEA-/CanRisk-Spezifikation ^a
MIT	erforderliche Daten zur Durchführung einer Risikoberechnung gemäß BOADICEA-/CanRisk-Spezifikation ^a : explizite Verwandtschaftsverhältnisse (strukturierter Stammbaum), Geburtsjahr, Krebserkrankungen (Art, Diagnosealter), Vitalstatus, Todesalter, ggf. Mutationsstatus
MGU	Ergebnis einer molekulargenetischen Untersuchung: Zeitpunkte für Probennahmen/Anforderung/Befund/Ergebnismitteilung, Art des Materials, Art der Untersuchung (vollständige Genanalyse oder gezielte Testung), Art des Genpanels, untersuchte Gene, Ergebnisse (Verknüpfung zu VAR)
CLK	Checkliste zur Erfassung einer familiären Belastung für Brust- und Eierstockkrebs (nur zu verwenden für Fälle, die über einen Kooperationspartner an ein DK-FBREC-Zentrum verwiesen wurden) ^b
TUM	Beschreibung von Primärtumor-Ereignissen: Diagnosezeitpunkt, Lokalisation (ICD-10), Seite, Rezidive und Fernmetastasen, Art der Entdeckung, Staging (TNM), Histologie (ICD-O), Grading, Rezeptorstatus, Therapie/Operationen
IFP	Durchführung und Ergebnis der Bildgebungsuntersuchungen in der intensivierten Früherkennung: Art der Bildgebungen, Gesamtbefund nach BI-RADS, mammographische Dichte, Empfehlung gemäß Gesamtbefund, histologische Abklärung inklusive Ergebnis
FUP	Follow-up: Änderungen des Vitalstatus, neue Tumorerkrankungen, neue Operationen
PHQ	Personal Health Questionnaire 4 (PHQ-4, 4-Item Selbstausfüller-Fragebogen)[17]
CWS	Cancer Worry Scale (5-Item Selbstausfüller-Fragebogen) [18]
IDT ^c	Identifizierende Merkmale (IDAT)
VAR	Liste aller distinkten genetischen Varianten in den relevanten Risikogenen: Gen, Transkript, genomische Position und Beschreibung der Sequenzänderung, HGVS-Annotation (DNA, Protein), Pathogenitätsbeurteilung (Class 1–5) [19] inklusive Beschreibung der Evidenz

^a <https://canrisk.org/>. ^b <https://www.konsortium-familiaerer-brustkrebs.de/informationen/informative-links/>. ^c Diese Merkmale werden nur erhoben und erfasst, wenn der Registerpatient explizit in das optionale Record-Linkage mit Krebsregister- und Krankenkassendaten eingewilligt hat. Die identifizierenden Merkmale werden strikt getrennt von den übrigen Registerdaten bei einer unabhängigen Treuhandstelle gespeichert.

Erfassung von Informationen zur Evidenz der Pathogenität einzelner Genvarianten möglich sein, die von einem Expertengremium des DK-FBREC kontinuierlich erstellt werden [13]. Es soll eine nahtlose Verknüpfung dieser Genvarianten-Datenbank mit den individuellen (personenbezogenen) Genbefunden möglich sein (siehe Erläuterung oben zur Beziehung zwischen dem Modul „VAR“ und „MGU“).

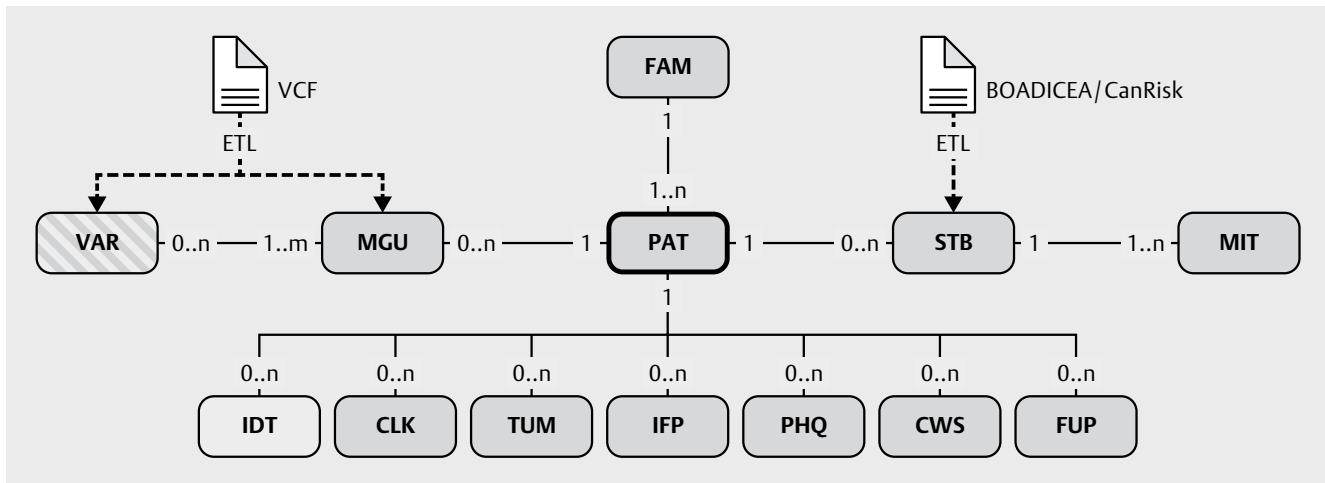
- 2) Mit der 2015 eingeführten Genpanel-Diagnostik wird zeitgleich eine Vielzahl von Risikogenen untersucht. Aufgrund der dadurch deutlich gestiegenen Zahl von festgestellten Genvarianten ist deren manuelle Erfassung sehr zeit- und arbeitsintensiv geworden. Aus diesem Grund, und auch um eine einheitliche Annotation der Genbefunde zu gewährleisten, sollte das Erfassungssystem das Hochladen und Einlesen von Dateien im „Variant Call Format“ (VCF) ermöglichen. Praktisch alle in den Konsortialzentren verwendeten Laborsysteme und deren bioinformatische Analysesoftware können dieses standardisierte textbasierte Dateiformat generieren.
- 3) Die Dokumentation von Stammbäumen ist von großer Bedeutung für „HerediCaRe“, da diese Daten eine wesentliche Grundlage für die Validierung des im klinischen Beratungskontext verwendeten Risikoberechnungsmodells „BOADICEA/CanRisk Web Tool“ (<https://canrisk.org/>) darstellen [14, 15]. Die Erfassung der per SOP geforderten strukturierten 3-Generationen-Stammbäume, d. h. aller einzelnen Familienmitglieder mit ihren expliziten verwandtschaftlichen Beziehungen, ihrem Vitalstatus, sowie aller aufgetretenen Krebserkrankungen (Art und Diagnosealter) ist eine sehr zeit- und arbeitsintensive Dokumentationsaufgabe. Das Dokumentationssystem soll daher eine Funktionalität

zum Hochladen und Einlesen von Stammbaum-Dateien im standardisierten „BOADICEA“-Format bereitstellen.

- 4) Um die Korrektheit der Stammbäume überprüfen zu können, wird eine grafische Anzeigemöglichkeit der dokumentierten Stammbäume gefordert.
- 5) Das System soll es erlauben, im Rahmen des Query-Managements beliebig komplexe und frei programmierbare Überprüfungen der zentralen Daten definieren zu können.
- 6) Das System soll die automatische Überführung von Altdaten aus der bisherigen Datenbank erlauben (inklusive der dazu erforderlichen Strukturtransformation)

IT-Lösung

Die Produktiv-Datenbank des Registers beruht auf ORACLE, aktuell Version 10. Eine Migration auf Version 19c ist in Vorbereitung. Die produktive Datenerfassungsanwendung beruht aktuell auf ORACLE Forms 12.2. Zur Nutzung der Datenerfassungsanwendung muss auf Seiten der klinischen Zentren eine Java 8 Laufzeitumgebung („Java Runtime Environment“, JRE) in aktueller Version installiert werden. Die dafür erforderliche Lizenzierung ist über die Dokumentationszentrale abgedeckt. Das Datenerfassungswerkzeug wird als Java-Web-Start-Anwendung über einen zentralen Applikationsserver bereitgestellt. Für den Austausch von VCF- und Stammbaumdateien zwischen Client und Server kommt Oracle Forms WebUtil zum Einsatz. Die grafische Darstellung von Stammbaumdateien beruht auf der freien Anwendung „CraneFoot“ (<http://www.finndiane.fi/software/cranefoot/>) [16]. Für den strukturierten Import von molekularanalytischen Informationen (VCF-



► **Abb. 2** Relationale Beziehungen zwischen den einzelnen Dokumentationsmodulen. Siehe ► **Tab. 1** für eine Beschreibung der jeweiligen Dokumentationsinhalte.

Format) und Stammbäumen (BOADICEA/CanRisk-Format) wurden spezielle Parser und ETL-Prozesse entwickelt und implementiert. Für Report-Funktionalitäten wird ORACLE Reports genutzt. Die IT-Lösung wird im IT-Verbund des IMISE/ZKS betrieben. Im Juli 2020 wurden umfassende Online-Schulungen für ca. 150 Mitarbeiter der klinischen Zentren zur Nutzung des Erfassungswerkzeugs durchgeführt.

Monitoring der Datenqualität

Im „HerediCaRe“-Register wird ausschließlich ein zentrales Datenmonitoring zur Überprüfung der Datenqualität durchgeführt. Ein Vor-Ort-Monitoring im Sinne einer Prüfung der Übereinstimmung von Quelldaten (Patientenakte) und Registerdaten ist derzeit nicht finanziert. Das zentrale Monitoring basiert wesentlich auf einer automatisierten Prüfung der Vollständigkeit, Plausibilität und Konsistenz aller zentral erfassten Informationen auf Basis eines in SQL programmierten Satzes von Prüfregele. Derzeit sind 278 verschiedene Prüfregele implementiert, anhand derer täglich der aktuelle Datenbestand überprüft wird. Neben einfachen Prüfungen auf Beantwortung mandatorischer Datenelemente spielen komplexe Überprüfungen eine besondere Rolle, z. B. Überprüfung der Stammbaumstruktur auf Fragmentierungen oder unplausible Konsanguinitäten (Zyklen), oder die regelkonforme Erfassung von TNM-Tumorformeln. Die Ergebnisse dieser Prüfungen werden den Dokumentaren in den klinischen Zentren tagesaktuell über das Datenerfassungssystem in Form von „Queries“ mit der Bitte um Lösung bereitgestellt. Eine Dokumentationsassistentin (50 % VK) steht am IMISE für Rückfragen der Dokumentare zur Verfügung.

Die Vergütung der Dokumentationsleistung in den klinischen Zentren erfolgt gemäß den Förderrichtlinien in Form von Fallpauschalen. Diese werden nur ausgezahlt, wenn durch die zentrale Dokumentation fallbezogen eine vorab bestimmte Datenqualität erreicht ist. Dies ist in der Regel dann der Fall, wenn keinerlei Queries für den jeweiligen Fall mehr vorliegen.

Diskussion

Aufgrund der speziellen Funktionsanforderungen wurde für „HerediCaRe“ die Eigenentwicklung einer geeigneten IT-Lösung favorisiert. Ausschlaggebend hierfür war, dass keines der uns bekannten am Markt befindlichen Softwareprodukte sämtliche dieser Anforderungen ohne Weiteres erfüllte. Zum anderen konnte das DK-FBREC vor der jetzigen BMBF-Förderung auf eine langjährige und äußerst positive Vorerfahrung einer selbst entwickelten IT-Lösung, basierend auf ORACLE und ORACLE Forms, zurückblicken. Das Datenbanksystem ORACLE bildet seit über 25 Jahren am IT-Verbund IMISE/ZKS die Grundlage fast aller dort betreuten klinischen und epidemiologischen Studien und hat sich im Einsatz als sehr zuverlässig und leistungsstark erwiesen, insbesondere bei komplexen epidemiologischen Studien und Registern. Auch für ORACLE Forms existieren am IMISE seit vielen Jahren gute Erfahrungen, auch wenn dieses Entwicklungswerkzeug aufgrund der Verfügbarkeit von modernen und funktionsstarken kommerziellen CDMS für klinische Studien mittlerweile kaum noch eingesetzt wird. Forms erlaubt die Realisierung datenbanknaher Eingabe-Anwendungen. Die zugrundeliegende Datenbanklogik wird automatisch generiert. Seit Version 12 sind Forms-Anwendungen ohne Browser lauffähig. Es ist lediglich die clientseitige Installation einer JRE erforderlich. Weitere Vorteile aus unserer Sicht sind das flexible und übersichtliche Maskendesign hinsichtlich Anordnung und Layout der Eingabefelder, und die Möglichkeit der Schaffung einer monolithischen Anwendung hinsichtlich der Einbindung heterogener Tools für spezielle Funktionalitäten. Dies hat nach unserer langjährigen Erfahrung zu einer guten Akzeptanz bei den Anwendern geführt. Eine systematische Nutzerbefragung zur Zufriedenheit ist geplant. Es gibt jedoch auch eine Reihe von Nachteilen, die mit der gewählten Lösung einhergehen. So erfordern die Anwendungsentwicklung und das Maskendesign IT-Fachwissen (Datenbank-Programmierer erforderlich). Aufgrund der Komplexität unserer Anwendung kann der vorhandene Maskengenerator (Layout Wizard) nicht genutzt werden. Zudem erlaubt die Forms-Entwicklungsumgebung nicht den Import von tabellarisch vorstrukturier-

ten Maskenspezifikationen. Datenausleitungen müssen selbst programmiert werden. Es sei auch bemerkt, dass erfahrenes Personal mit fundierten Kenntnissen der Forms-Entwicklungsumgebung zunehmend schwerer zu finden ist. Es wurde zu Beginn des Projektes die Möglichkeit einer Neuentwicklung der Benutzeroberfläche auf der Basis alternativer Entwicklungsumgebungen diskutiert. Aufgrund der sehr kurzen zur Verfügung stehenden Entwicklungszeit bis zur erforderlichen Einsatzfähigkeit des Dokumentationssystems konnte diese Option jedoch nicht verfolgt werden.

In langfristig angelegten Register-Studien sind die wissenschaftlichen Ziele oft zunächst sehr allgemein formuliert und konkretisieren sich erst zu einem späteren Zeitpunkt oder können sich auch langfristig ändern. Daher gilt es hinsichtlich der Dokumentationsinhalte eine gewisse Balance zu halten zwischen einerseits einer ressourcen- und datenschutzbedingten Erfordernis zur Sparsamkeit und andererseits einer möglichst breiten Dokumentation, die es erlaubt, auch künftig neue Fragestellungen zu beantworten. Änderungen oder Anpassungen der wissenschaftlichen Zielsetzungen können daher jederzeit Modifikationen des Merkmalkatalogs zur Folge haben. Das Datenerfassungssystem wurde Ende September 2020 in den Produktivbetrieb genommen und verfügt in der aktuellen Ausbaustufe über fast alle im Anforderungskatalog genannten Funktionalitäten. Die Funktion zum Einlesen der standardisierten Stammbaumdateien ist teilweise implementiert und wird derzeit vervollständigt. Mit Stand September 2021 wurden bereits über 13 000 Patienten neu erfasst oder aus dem Vorgängersystem migriert. Derzeit ist die Anbindung an eine unabhängige Treuhandstelle in Vorbereitung, welche die Kenndaten der Patienten verwaltet wird. Diese Kenndaten werden zur Durchführung eines geplanten Record-Linkage mit klinischen Krebsregistern benötigt. Zukünftig sollen außerdem Schnittstellen zu Laborinformationssystemen geschaffen werden, um Doppeleingaben zu vermeiden. Die Erfahrungen seit der Produktivstellung zeigen, dass trotz der zuvor langen und sorgfältigen Entwicklung des Merkmalskatalogs eine Reihe von nachträglichen Ergänzungen und Korrekturen erforderlich waren und diese künftig auch weiter zu erwarten sind. Da unsere Datenbanklösung die volle Kontrolle über die zugrundeliegende Datenstruktur erlaubt, können jedoch auch komplexere Änderungen am Datenmodell zeitnah umgesetzt werden.

Zusammenfassend lässt sich sagen, dass sich die Entwicklung einer eigenen IT-Lösung zur Erfassung von wissenschaftlichen Daten im komplexen Gebiet der hereditären Tumorerkrankungen aufgrund der sehr spezifischen Funktionsanforderungen bewährt hat und die genannten Vorteile die zu bedenkenden Nachteile klar überwiegen.

Interessenkonflikt

Die Autorinnen/Autoren geben an, dass kein Interessenkonflikt besteht.

Literatur

- [1] Engel C, Fischer C, Zachariae S et al. Breast cancer risk in BRCA1/2 mutation carriers and noncarriers under prospective intensified surveillance. *Int J Cancer* 2020; 146: 999–1009
- [2] Bick U, Engel C, Krug B et al. High-risk breast cancer surveillance with MRI: 10-year experience from the German consortium for hereditary breast and ovarian cancer. *Breast Cancer Res Treat* 2019; 175: 217–228
- [3] Engel C, Rhiem K, Hahnen E et al. Prevalence of pathogenic BRCA1/2 germline mutations among 802 women with unilateral triple-negative breast cancer without family cancer history. *BMC Cancer* 2018; 18: 265
- [4] Kast K, Rhiem K, Wappenschmidt B et al. Prevalence of BRCA1/2 germline mutations in 21 401 families with breast and ovarian cancer. *J Med Genet* 2016; 53: 465–471
- [5] Kast K, Schmutzler RK, Rhiem K et al. Validation of the Manchester scoring system for predicting BRCA1/2 mutations in 9,390 families suspected of having hereditary breast and ovarian cancer. *Int J Cancer* 2014; 135: 2352–2361
- [6] Rhiem K, Engel C, Graeser M et al. The risk of contralateral breast cancer in patients from BRCA1/2 negative high risk families as compared to patients from BRCA1 or BRCA2 positive families: a retrospective cohort study. *Breast Cancer Res* 2012; 14: R156
- [7] Graeser MK, Engel C, Rhiem K et al. Contralateral breast cancer risk in BRCA1 and BRCA2 mutation carriers. *J Clin Oncol* 2009; 27: 5887–5892
- [8] Meindl A, Hellebrand H, Wiek C et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet* 2010; 42: 410–414
- [9] Hauke J, Horvath J, Gross E et al. Gene panel testing of 5589 BRCA1/2-negative index patients with breast cancer in a routine diagnostic setting: results of the German Consortium for Hereditary Breast and Ovarian Cancer. *Cancer Med* 2018; 7: 1349–1358
- [10] Borde J, Ernst C, Wappenschmidt B et al. Performance of breast cancer polygenic risk scores in 760 female CHEK2 germline mutation carriers. *J Natl Cancer Inst* 2021; 113: 893–899
- [11] Weber-Lassalle N, Borde J, Weber-Lassalle K et al. Germline loss-of-function variants in the BARD1 gene are associated with early-onset familial breast cancer but not ovarian cancer. *Breast Cancer Res* 2019; 21: 55
- [12] Weber-Lassalle N, Hauke J, Ramser J et al. BRIP1 loss-of-function mutations confer high risk for familial ovarian cancer, but not familial breast cancer. *Breast Cancer Res* 2018; 20: 7
- [13] Wappenschmidt B, Hauke J, Faust U et al. Criteria of the German Consortium for Hereditary Breast and Ovarian Cancer for the Classification of Germline Sequence Variants in Risk Genes for Hereditary Breast and Ovarian Cancer. *Geburtshilfe Frauenheilkd* 2020; 80: 410–429
- [14] Lee A, Mavaddat N, Wilcox AN et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med* 2019; 21: 1708–1718
- [15] Carver T, Hartley S, Lee A et al. CanRisk Tool-A Web Interface for the Prediction of Breast and Ovarian Cancer Risk and the Likelihood of Carrying Genetic Pathogenic Variants. *Cancer Epidemiol Biomarkers Prev* 2021; 30: 469–473
- [16] Makinen VP, Parkkonen M, Wessman M et al. High-throughput pedigree drawing. *Eur J Hum Genet* 2005; 13: 987–989
- [17] Lowe B, Wahl I, Rose M et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord* 2010; 122: 86–95
- [18] Rippinger N, Fischer C, Haun MW et al. Cancer surveillance and distress among adult pathogenic TP53 germline variant carriers in Germany: A multicenter feasibility and acceptance survey. *Cancer* 2020; 126: 4032–4041
- [19] Plon SE, Eccles DM, Easton D et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 2008; 29: 1282–1291