Aspekte des datenschutzgerechten Managements klinischer Forschungsdaten

Thomas M. Deserno[1], Verena Deserno[2], Volker Lowitsch[3], Walter Franck[4], Joachim Willems[5], Helmut Löbner[6]

[1]Institut für Medizinische Informatik, Medizinische Fakultät, RWTH Aachen
[2]Clinical Trial Center Aachen (CTC-A), Universitätsklinikum der RWTH Aachen
[3]Geschäftsbereich IT-Direktion, Universitätsklinikum der RWTH Aachen
[4]Rechen- und Kommunikationszentrum, RWTH Aachen
[5]Datenschutzbeauftragter des Universitätsklinikums Aachen
[6]Datenschutzbeauftragter der RWTH Aachen

deserno@ieee.org

Abstract: Medizinische Forschungsdaten, die an universitären Fakultäten in klinischen Studien und Forschungsregistern erhoben werden, sind im Hinblick auf den Datenschutz besonders sensibel. Dennoch bietet die zentrale Krankenhaus-IT hier i.d.R. keine Unterstützung, und forschende Ärzte behelfen sich mit den verfügbaren Tools der Office-Palette. In diesem Beitrag präsentieren wir ein Datenschutz- und IT-Sicherheitskonzept für IT-Services, die am Clinical Trial Center Aachen (CTC-A) für forschende Wissenschaftler angeboten werden sollen. Wir führen eine Strukturanalyse, Schutzbedarfsanalyse und Bedrohungs- und Risikoanalyse durch, um basierend auf dem BSI Baustein 1.9 konkrete Maßnahmen einzuleiten, die verschiedenen IT-Verfahren abzusichern. Die vorgestellten Lösungswege orientieren sich an den allgemeinen Zielen des Datenschutzes (Datensparsamkeit, Vertraulichkeit, Integrität, Verfügbarkeit, Authentizität, Revisionsfähigkeit und Transparenz) und berücksichtigen das generische Datenschutzkonzept der TMF mit strikter Trennung identifizierender und medizinischer Daten, beinhalten den Pseudonymisierer des TMF. Die IT-Services des CTC-A werden demgemäß in einem 5-stufigen Modulkonzept realisiert.

1. Einleitung

Das Clinical Trial Center Aachen (CTC-A) wurde 2010 als eigenständige Einheit der Medizinischen Fakultät der RWTH Aachen mit dem Ziel neu konstituiert, die Strukturen zur Planung, Durchführung und Auswertung klinischer Studien an der Medizinischen Fakultät der RWTH Aachen zu verbessern. Satzungsgemäße Aufgabe ist die „Einführung zeitgemäßer elektronischer Techniken / IT-Lösungen für Daten-, Studien-, Projekt-, Kooperations- und Budgetmanagement“ [R11]. Diese Aufgabe wurde in ein modulares Stufenkonzept umgesetzt, welches fünf Stufen umfasst:

1. *Erfassung von administrativen Studiendaten*: Ziel dieser Stufe ist eine Bestandsaufnahme laufender und geplanter Studien am Universitätsklinikum Aachen. Struktu-

riert soll erfasst werden, welche Personen aus welchen Institutionen / Kliniken an welchen Studien in welchen Funktionen mit welchem Aufwand tätig sind.

2. *Erfassung von Patientenzahlen*: Ziel dieser Stufe ist es, nicht-rekrutierende laufende Studien zu identifizieren, um frühzeitig steuernd eingreifen zu können. Hierzu werden Einschlusszahlen und -zeitpunkte erfasst.

3. *Erfassung von Adressdaten der Patienten*: Ziel dieser Stufe ist es, die Arbeit der Study-Nurses (speziell zur Durchführung von Studien geschulte Krankenschwester/Pflegekraft) zu unterstützen. Für Kontaktaufnahme, Untersuchungsplanung und -terminierung werden die Adressdaten der Studienpatienten erhoben.

4. *Erfassung von medizinischen Studiendaten*: Das CTC-A bietet elektronische Case Report Forms (eCRF) für klinische Studien an, die insbesondere den sog. Investigator Initiated Trials (IIT), die typischerweise nur ein geringes Budget haben, zugute kommen.

5. *Erfassung von medizinischen Studiendaten*: Mit dieser Stufe ist ein vollständiger Studiensupport erreicht, denn Verblindung, Randominierung und Pseudonymisierung sind in die CTC-A IT-Infrastruktur integriert und über einen Single-Sign-On Mechanismus verfügbar.

Um diese Ziele zu erreichen, wurde das Institut für Medizinische Informatik, Medizinische Fakultät der RWTH Aachen mit der Entwicklung bzw. Bereitstellung geeigneter Software beauftragt. Die Eignung von Software bestimmt sich neben Attributen wie Funktionsumfang, Benutzbarkeit (Usability) oder dezentrale Verfügbarkeit vor allen aus den datenschutzrechlichen Aspekten.

Für den Datenschutz in Nordrhein-Westfalen sind das Landesdatenschutzgesetz (DSG NRW 2000) und das Gesundheitsdatenschutzgesetz (GDSG NRW 1999) maßgeblich [DG00, GG99]. Hierin werden folgende Ziele des Datenschutzes definiert:

- *Datensparsamkeit*: Es dürfen nur so wenig personenbezogene Daten erhoben und verarbeitet werden, wie zur Zweckerfüllung erforderlich ist. Fällt der Zweck weg (z.B. durch Abschluss der Studie) oder sind Aufbewahrungsfristen abgelaufen, so sind die personenbezogenen Daten zu löschen.

- *Vertraulichkeit*: Nur befugte Personen können personenbezogene Daten zur Kenntnis nehmen.

- *Integrität*: Personenbezogene Daten bleiben während der Verarbeitung unversehrt, vollständig und aktuell.

- *Verfügbarkeit*: Personenbezogene Daten stehen zeitgerecht zur Verfügung und können ordnungsgemäß verarbeitet werden.

- *Authentizität*: Personenbezogene Daten können jederzeit ihrem Ursprung zugeordnet werden. Herkunft und Verarbeitung werden dokumentiert.

– *Revisionsfähigkeit*: Es kann festgestellt werden, wer wann welche personenbezogenen Daten wie verarbeitet hat. Die Datenverarbeitung muss hinreichend protokolliert werden.

– *Transparenz*: Die Verfahrensweisen bei der Verarbeitung personenbezogener Daten sind vollständig, aktuell und so gut dokumentiert, dass sie in zumutbarer Zeit nachvollzogen werden können. Dies wird mit einem Verfahrensverzeichnis erreicht.

Um diese Ziele zu erreichen und dauerhaft zu gewährleisten, wurde in enger Abstimmung mit den Datenschutzbeauftragten des Universitätsklinikums Aachen sowie der RWTH Aachen ein Sicherheitskonzept für die elektronische Datenverarbeitung am CTC-A erarbeitet, das im folgenden vorgestellt werden soll.

2. Modellierung nach dem BSI Schichtenmodell

Datenschutz bezeichnet den Schutz des Einzelnen vor dem Missbrauch personenbezogener Daten. Werden in einem System personenbezogene Daten erfasst, so muss sichergestellt werden, dass diese nur demjenigen zugänglich sind, der hierauf auch zugreifen darf. Die Datensicherheit umfasst alle technischen und organisatorischen Maßnahmen zum Schutz von Daten gegen Beschädigung oder Verlust. Sie zielt dabei besonders auf die Sicherstellung der Verfügbarkeit, der Integrität und der Vertraulichkeit der Daten ab. Mit Systemverfügbarkeit werden Ausfallzeiten eines Systems, sowie die Dauer nach einer Störung bis zur Wiederinbetriebnahme beschrieben. Die Systemverfügbarkeit ist damit auch ein Teil der übergreifenden Datensicherheit.

Die Verfahrensmodellierung nach dem BSI Schichtenmodell[1] umfasst die Strukturanalyse, die Schutzbedarfsfeststellung, die Bedrohungs- und Risikoanalyse sowie die Erstellung eines datenschutz- und datensicherheitsgerechten Prüf- bzw. Entwicklungsplanes.

2.1. Strukturanalyse des Informationsverbundes

Das CTC-A ist die für den Betrieb seiner elektronischen Datenverarbeitung verantwortliche Stelle. Die Zusammenarbeit mit dem Institut für Medizinische Informatik (IMI) ist über einen Kooperationsvertrag geregelt [SS10]. Das IMI ist mit seiner IT-Infrastruktur wiederum an das Rechen- und Kommunikationszentrum (RKZ) der RWTH Aachen angebunden. Es wird vom Geschäftsbereich IT-Direktion des Universitätsklinikums Aachen (UKA) beraten (Abb. 1). Diese Einheiten können als Teil der RWTH bzw. des UKA autonom agieren.

2.2. Schutzbedarfsanalyse mit Schadensszenarien

Für die am CTC-A geplanten Verfahren sind Finanzdaten des CTC-A sowie personenbezogene Daten einzelner Mitarbeiter des CTC-A, die an den klinischen Studien betei-

[1] http://www.bsi.bund.de/

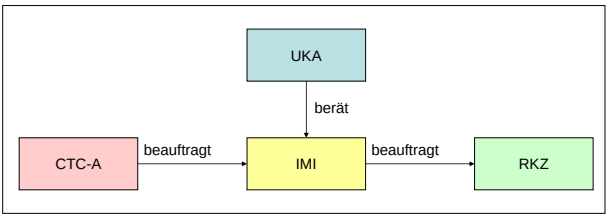

Abbildung 1: Die am Informationsverbund beteiligten Institutionen und deren Abhängigkeiten.

ligten Mitarbeiter des Universitätsklinikums Aachen sowie vor allem der in die Studien eingeschlossenen Patienten und Probanden relevant. Zur Schutzbedarfsfeststellung wurden nach Wirth einzelne Schadensszenarien analysiert [W08]:

- *Verstoß gegen Gesetze, Vorschriften und Verträge*: Neben den bereits genannten Landesdatenschutzgesetzen sind vor allem Grundgesetz, das Sozialgesetzbuch und das Strafgesetzbuch relevant. Als maßgebliche Verordnung gilt die EU-Verordnung zur Good Clinical Practice [GV04]. Wichtig ist hier, ob die Verstöße auf Vorsatz, oder (grobe) Fahrlässigkeit zurückgeführt werden können.

- *Beeinträchtigung des Rechts auf informationelle Selbstbestimmung*: Die vorgenannte Trennung in einzelne Realisierungs- bzw. Ausbaustufen vereinfacht die Bewertung potentieller Schäden, die durch unerlaubte Erfassung, fehlerhafte Verarbeitung oder zweckfremde Nutzung der Daten entstehen könnten. In jedem Fall werden Worst-Case-Szenarien betrachtet.

- *Beeinträchtigung der persönlichen Unversehrtheit*: In dieser Kategorie wurde analysiert, ob Fehler in der Datenverarbeitung zu psychischen oder physischen Beeinträchtigungen des Betroffenen – im Extremfall zum Tode führen könnten. Dies gilt für alle Personen, die mit den Systemen arbeiten, nicht nur die, deren Daten verarbeitet werden.

- *Beeinträchtigung der Aufgabenerfüllung*: Die Schwere des Schadens in dieser Kategorie richtet sich nach der Dauer des Systemausfalls. Insbesondere wurde die Frage geprüft, ob das CTC-A arbeitsunfähig werden kann, wenn die geplanten Softwaresysteme nicht mehr funktionieren.

- *Negative Außenwirkung*: Als Beispiele für dieses Schadensszenario können der Ansehensverlust der am Informationsverbund beteiligten Institutionen (CTC-A, IMI, RKZ) oder der übergeordneten Instanzen (Universitätsklinikum Aachen, RWTH Aachen) oder die Einbuße der Konkurrenzfähigkeit des CTC-A genannt werden. Die Höhe des Schadens wurde an der Schwere des Vertrauensverlustes und des Verbreitungsgrades der negativen Außenwirkung bemessen.

- *Finanzielle Auswirkungen*: Um dieses Schadensszenario zu bewerten wurde analysiert, ob ein möglicher finanzieller Schaden für die beteiligten Institutionen tolerabel, erheblich oder existenzbedrohend ist.

Ausbaustufe	Informationsart	Schutzbedarf	Skalenwert
1	Studienbezeichnung	öffentlich	0
	Referenznummern	normal	1
	Personen und Rollen	normal	1
	Institutionen und Rollen	normal	1
	Relevante Gesetze	öffentlich	0
	Laufzeit	öffentlich	0
	Finanzierung	hoch	2
	Relevante Dokumente	hoch	2
2	Anonyme Patientenkennung	normal	1
	Einschlussdatum	normal	1
	Status	normal	1
3	Patientenname	sehr hoch	3
	Patientengeburtsdaten	sehr hoch	3
	Patientenkontaktdaten	sehr hoch	3
	Klinische Studien des Patienten	sehr hoch	3
4	Medizinische Daten faktisch anonym	hoch	2
5	Medizinische Daten personenbezogen	sehr hoch	3

Tabelle 1: Schutzbedarf der jeweiligen Informationsarten einzelner IT-Ausbaustufen am CTC-A.

Die hieraus resultierende Einteilung in drei Schutzklassen „normal“, „hoch“ und „sehr hoch“ ist in Tabelle 1 dargestellt. „Öffentlich“ bedeutet, dass es sich hierbei um der Allgemeinheit bekannte Information handelt, ein Schutzbedarf also nicht vorhanden ist.

2.3. Bedrohungs- und Risikoanalyse

Nach der Modellierung des IT-Grundschutzes nach BSI wurden zunächst die fünf wesentlichen Bedrohungsszenarien für den Baustein B 1.9 „Hard- und Softwaremanagement“ betrachtet [BS11]:

1. *Höhere Gewalt*: Eine solche liegt nach deutscher Rechtsprechung genau dann vor, wenn das den Schaden verursachende Ereignis von außen einwirkt, also seinen Grund nicht in der Natur der gefährdeten Sache hat (objektive Voraussetzung) und das Ereignis auch durch die äußerst zumutbare Sorgfalt weder abgewendet noch unschädlich gemacht werden kann (subjektive Voraussetzung). Gefährdungen der höheren Gewalt ergeben sich z.B. aus krankheitsbedingtem Personalausfall oder Geräteschäden durch Feuer oder Wasser und beeinflussen vornehmlich die Verfügbarkeit der Daten.

2. *Organisatorische Mängel*: Der Begriff Organisation steht für den Prozess, durch den fortlaufende unabhängige Handlungen zu vernünftigen Folgen zusammengefügt werden, so dass sinnhafte Ergebnisse erzielt werden. Hierzu gehören also Regelungen und Ordnungen, die alle Anwender zu befolgen haben. Das Bedrohungspotential mangelnder Organisation erstreckt sich auf fast alle Schutzziele, insbesondere die Vertraulichkeit der Daten, deren Verfügbarkeit und Integrität.

3. *Menschliche Fehlhandlungen*: Durch Fehlverhalten von Personen aller Art kann vor allem der Vertraulichkeits- bzw. Integritätsverlust von Informationen und Daten herbeigeführt bzw. ermöglicht werden. Die Gefährdungen ergeben sich aus mensch-

lichen Handlungen, die versehentlich, d.h. also nicht vorsätzlich durchgeführt oder unterlassen werden.

4. *Technisches Versagen*: Die in den Verfahren zum Einsatz kommende Technik basiert auf Hard- und Softwarekomponenten sowie dem Zusammenspiel dieser. Alle drei Komponenten haben somit ein erhebliches Gefährdungspotential für Vertraulichkeit, Integrität, Verfügbarkeit und Authentizität.

5. *Vorsätzliche Handlungen*: Manipulation, Verfälschung und bewusste Zerstörung sind gefährdende vorsätzliche Handlungen, die von den Mitarbeitern des Informationsverbundes, den Benutzern der Software, oder auch von Dritten ausgehen können. Vorsätzliche Handlungen bedrohen vor allem die Verfügbarkeit (Hardware) und die Vertraulichkeit, Authentizität oder Integrität (Software) der personenbezogenen Daten.

Diese Risikoanalyse zeigte, dass alle Gefährdungsursachen für das CTC-A relevant sind. Das Eintrittsrisiko ist bei menschlichen und technischen Fehlern am größten, aber auch die gesamte Organisation birgt durch die translationale Struktur der klinischen Forschung erhebliche Risiken. Gezielte vorsätzliche Angriffe auf das CTC-A erscheinen weniger relevant, denn diese können nur von dem sehr kleinen Personenkreis der Programmierer am IMI vorgenommen werden und wären somit leicht mit dem Täter in Verbindung zu bringen. Weiterhin ist das CTC-A als einzelne Einheit im Vergleich zu anderen Bereichen der RWTH oder des UKA vergleichsweise unbedeutend. Risiken der höheren Gewalt werden im Informationsverbund, der das RKZ mit einschließt, auf anderen Ebenen behandelt.

3. Maßnahmen zur IT-Sicherheit und zum Datenschutz

Nach § 10 (3) DSG NRW ist die Wirksamkeit der Maßnahmen unter Berücksichtigung sich verändernder Rahmenbedingungen und Entwicklungen der Technik kontinuierlich zu überprüfen. Die sich daraus ergebenden notwendigen Anpassungen sind zeitnah umzusetzen. Das BSI bietet mit dem Schichtenkonzept hierzu einen geeigneten Leitfaden, wobei zunächst das generelle Realisierungskonzept der rechnergestützten IT-Anwendungen CTC-A als Ausgangspunkt dargestellt wird, bevor die Eigenentwicklungen des IMI für das CTC-A und weitere Maßnahmen nach BSI beschrieben werden.

3.1. Generelles Realisierungskonzept

Die Schutzbedarfsfeststellung mit Gefährdungs- und Risikoanalyse hat ergeben, dass die vom CTC-A eingeführten IT-Anwendungen auf einfache und klar verständliche Benutzung hin optimiert werden und möglichst ohne eigene lokale Hard- und Software betrieben werden sollten. Die neuartigen Web2.0-Technologien im Internet bieten hierfür entsprechende Lösungsmöglichkeiten. Heutzutage können im Internet komplexe Anwendungen realisiert werden, die in jedem Web-Browser ohne Softwareinstallation

sofort benutzt werden können und weitgehend unabhängig von der Art des benutzten Browsers (Microsoft Explorer, Mozilla Firefox, Google Chrome, etc.) sind.

Weiterhin gilt das generelle Konzept der Datensparsamkeit. Medizinische Forschungsdaten sollten grundsätzlich anonymisiert werden. Ist das nicht möglich ist, sollten sie pseudonym verarbeitet werden. Daher basiert das Realisierungskonzept der rechnergestützten Anwendungen des CTC-A auf dem generischen Datenschutzkonzept der Technologie- und Methodenplattform für die vernetzte medizinische Forschung (TMF) e.V., die vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wird [RD06].

3.1.1. Generisches Datenschutzkonzept für medizinische Forschungsnetze der TMF

Um in aktueller medizinischer Forschung international konkurrenzfähig zu bleiben, wurde vom BMBF die Bündelung medizinischer Kompetenzen in Netzwerken unterstützt. Die hierbei relevanten medizinischen Daten sind höchst-vertrauliche Patientendaten, die übergreifend zusammengeführt werden müssen. Um dies unter Berücksichtigung der ärztlichen Schweigepflicht und der in Deutschland gewachsenen Datenschutzgesetzgebung in Einklang zu bringen, hat das TMF generische Datenschutzmodelle entwickelt. Hierbei wird zwischen klinisch-wissenschaftlicher (Modell A) und wissenschaftlich-klinischer (Modell B) Orientierung des Forschungsverbundes unterschieden.

Im Modell A steht die klinische Forschung im Fordergrund, was den Aufgaben des CTC-A und den klinischen Studien entspricht. Das generische TMF Konzept verfolgt das Schutzziel der Datensparsamkeit: in der eigentlichen medizinischen Forschungsdatenbank werden keine identifizierenden Daten vorgehalten. Diese werden in einer separaten Patientenliste verwaltet. Damit ergeben sich die folgenden Paradigmen [RD06]:

1. Unterscheidung des klinischen Datenbestandes in identifizierende Patientenliste (IDAT) und medizinischen Behandlungsdatenbank (MDAT).
2. Physikalische, logische und organisatorische Trennung von IDAT und MDAT.
3. Referenzierung zwischen Patientenliste und Behandlungsdatenbank durch einen eineindeutigen Patientenidentifikator (PID).

Die Verknüpfung der Datensätze in Patientenliste und Behandlungsdatenbank erfolgt über die PID, die jedoch nur zwischen beiden Datenbanken kommuniziert wird. Nach außen wird lediglich eine temporäre ID (TempID) gegeben, um so dem Benutzer (z.B. behandelnder Arzt) bei autorisierter Abfrage die vorübergehende Zusammenführung des klinischen Datenbestandes des jeweiligen Patienten zu erlauben.

3.1.2. PID Generator der TMF

Neben der generischen Implementierungsstruktur unterstützt das TMF die klinische Forschung durch Bereitstellung spezieller Softwarekomponenten. Der PID-Generator ist ein wesentlicher Teil des Pseudonymisierungsdienstes der TMF, der als Open Source von der TMF verfügbar gemacht wird. Er erzeugt für jeden in eine Studie aufzunehmenden Patienten eine PID. Diese kann als einstufiges Pseudonym verwendet werden und

bildet zusätzlich einen Eingangsparameter für eine zweite Pseudonymisierung für die langfristige und weniger zweckgebundene Datenspeicherung in Forschungsdatenbanken entsprechend dem generischen Datenschutzkonzept B der TMF. Auf diese Weise werden Patienten, auch wenn sie verschiedene Ärzte kontaktieren, eindeutig unter demselben Pseudonym zusammengeführt. Ebenso können Daten aus verschiedenen Studien über das Pseudonym zusammengeführt werden, ohne dass die Identität des Patienten aufgedeckt wird. Dies ist für die IT-Anwendungen des CTC-A von besonderer Bedeutung, da die Patienten des UKA in mehreren Studien mitwirken können.

In der für die CTC-A Anwendungen eingesetzten Version 1.2 des TMF PID-Generators wurden die Prüf- und Match-Algorithmen von der TMF noch einmal verbessert und somit die Datenqualität in zentralen Datenpools optimiert [WG08]. Der PID Generator verbindet mehrstufige direkte, ungefähre und phonetische Vergleiche auf identifizierenden Feldern wie Name, Geburtsname und Geburtsdatum zu einem eindeutigen Pseudonym, das robust gegenüber Schreibfehlern, Buchstabendrehern und Datenänderungen ist.

3.1.3 Systemarchitektur der CTC-A Anwendungen

Wesentlich für Datenschutz und Datensicherheit der CTC-A Verfahren in den jeweiligen Realisierungsstufen ist die Verknüpfung der Komponenten Studienmanagement und Studiendaten mit dem Pseudonymisierungsdienst des TMF PID-Generators. Abbildung 3 zeigt die resultierende Systemarchitektur rechnergestützter Verfahren am CTC-A. Die farbigen Hinterlegungen entsprechen den Farben aus Abb. 1 und kennzeichnen die beteiligten Institutionen.

Seitens des CTC-A wird hierbei keinerlei spezielle Hard- oder Software benötigt. Das IMI betreibt drei unabhängige Web-Server in einer virtuellen Serverfarm, die mit eigenen IP-Adressen und festen Portkonfigurationen nur die für die CTC-A Anwendungen benötigten Dienste bereitstellen. Diese Server haben jeweils eigene und unabhängige Datenbanken, die mit MySQL, einem Open Source Datenbankmanagementsystem (DBMS) betrieben werden und über die vom RKZ angebotenen Dienste mit TSM dort unabhängig gesichert werden. Seit der Tivoli Client Version 5.3 wird eine AES 128-Bit Verschlüsselung der Backupdaten eingesetzt, so dass sie keinem Mitarbeiter des RKZ zugänglich sind.

3.1.4. Rollen- und Rechtekonzept

Kern der Internet-basierten CTC-A Verfahren ist ein differenziertes Rollen- und Rechtekonzept, das mit der Anmeldung über den Web-Browser aktiviert wird. In allen CTC-A Anwendungen können sich nur eingetragene Benutzer mit Username und Passwort anmelden. Der Datenzugriff ist über das Rollenmodell hierarchisch gestaltet. Die folgenden Zugriffs-Level sind derzeit definiert:

– *None*: Eine Person wird zwar im CTC-A Study-Management-Tool mit seinen Kontaktdaten (Adressdaten) geführt, hat aber keinen Zugriff auf die CTC-A Anwendungen. Dies ist für die überwiegende Mehrzahl der mittlerweile knapp 300 erfassten

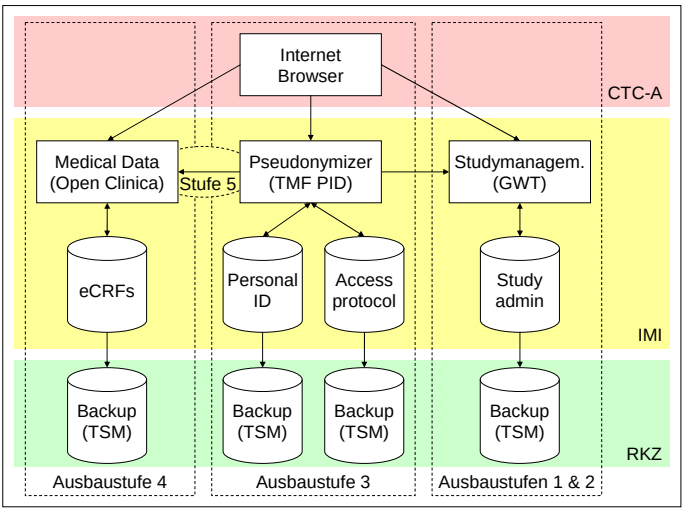

Abbildung 2: 4-Tier-Architecture der einzelnen Ausbaustufen.

Personen der Fall, die z.B. als externer Monitor oder Mitarbeiter einer CRO an klinischen Studien des UKA beteiligt sind.

- *Own*: Eine Person ist an klinischen Studien im UKA in einer gewissen Funktion direkt beteiligt und kann sich mit seiner E-Mail-Adresse als Benutzername und einem selbst gewählten Passwort anmelden. Derzeit sind insgesamt 27 Funktionen definiert, z.B. Coordinating Investigator, Principal Investigator (LKP), Study-Nurse Main, Project Manager, Monitor, Pharmacist, IT Manager. Das Level „Own“ erlaubt nur den Zugriff auf die Studien und die damit verknüpften Daten, an denen die Person selbst beteiligt ist. Sämtliche Ausgaben des Systems, sei es am Bildschirm im Browser-Interface oder über die gedruckten PDF-Reports, basieren auf dieser eingeschränkten Datensicht.

- *Department*: Jede Person ist genau einer Organisation zugeordnet. Mittlerweile sind 260 Organisationen mit ihren Adressdaten im CTC-A Study-Management-Tool erfasst. Das Zugriffs-Level „Department“ zeigt alle Studien und die damit verknüpften Daten, in denen mindestens eine Person desselben Departments eine der oben genannten standardisierten Funktionen ausübt. Diese Person muss – im Unterschied zum Level „Own“ – aber nicht mit dem User identisch sein.

- *All*: In diesem Level werden alle erfassten Daten angezeigt, unabhängig davon, welche Person in welcher Funktion mit einer Studie verknüpft ist.

- *Master*: Zusätzlich zum Level „All“ kann ein Master-User Nutzerrechte für Personen setzen bzw. modifizieren. Bekommt ein Benutzer Zugriffsrechte oder werden diese modifiziert, wird automatisch eine standardisierte E-Mail verschickt, aus der die Login-Informationen sowie Hinweise zum Setzen eines persönlichen Passwortes zu entnehmen sind. Weiterhin kann ein Master-User temporär in der Sicht eines an-

deren Users auf die Daten zugreifen. Dies ermöglicht eine stichprobenartige Prüfung der Vertraulichkeit der in den CTC-A Verfahren verarbeiteten Daten.

Der Zugriff ist generell nur lesend möglich. Darüber hinaus können die folgenden Rechte von einem Master-User für jeden anderen User individuell gesetzt werden:

- *Write*: In der je nach Zugriffs-Level unter Umständen eingeschränkten Sicht auf die Daten kann der Benutzer Daten ändern und hinzufügen.

- *Finance*: Ist dieses Recht gesetzt, werden dem Benutzer alle für die Abrechnung der Studie relevanten Daten angezeigt. Ein Nutzer mit dem Recht „Finance“ kann die Finanzdaten lesen und schreiben.

- *Patient*: Dieses Recht erlaubt ab der 3. Ausbaustufe die Entblindung der pseudonym verwalteten Patienten-Identifikationsdaten, also die Kommunikation mit der IDAT Applikation und dem TMF PID Generator. Weiterhin wird es mit diesem Recht in der 5. Ausbaustufe möglich sein, das Pseudonym der medizinischen Datenbank (MDAT), die in Stufe 4 mit Open Clinica etabliert wird, direkt aus der IDAT zu erfragen, um in der MDAT die medizinischen Daten dem richtigen Pseudonym zuzuordnen.

Dieses einfache und übersichtliche Konzept ist so gestaltet, dass Vertraulichkeitsverluste aufgrund von fälschlich gewährten Rechten minimiert werden. Dennoch ist es aufgrund seiner Modularität hinreichend flexibel und erweiterbar, um auch künftigen Anforderungen datenschutzkonform gerecht werden zu können.

3.1.5. Technische Regelungen zum Passwortgebrauch

Die CTC-A Verfahren sind nur mit einer gültigen Kombination aus Nutzerkennung (E-Mail-Adresse des Benutzers) und dem zugehörigen Passwort erreichbar. Folgende Rahmenbedingungen werden durch die Implementierung erzwungen:

- Das Passwort ist mindestens 8 Zeichen lang und enthält mindestens ein Zeichen, das kein Buchstabe ist. Die Wahl von Trivialpasswörtern ist somit nicht möglich.

- Für die Erstanmeldung neuer Benutzer werden Einmalpasswörter vom System zufällig vergeben, die ebenfalls diesen Richtlinien entsprechen.

- Jeder Benutzer kann sein eigenes Passwort jederzeit ändern.

- Bei der Authentifizierung wird das Passwort nicht auf dem Bildschirm angezeigt.

- Erfolglose Anmeldeversuche werden mit einer kurzen Fehlermeldung abgelehnt. Es ist nicht erkennbar, ob der Benutzername, das Passwort oder beides falsch waren.

- Nach drei aufeinander folgenden fehlerhaften Passworteingaben wird der Zugang gesperrt und eine automatische E-Mail an den Benutzer und die Gruppe der Master-User gesendet. Diese können dann den Zugang manuell reaktivieren.

– Die Passwörter werden nur verschlüsselt übertragen und gespeichert (Hashing). Auch System-Administratoren können Passwörter nicht im Klartext auslesen.

3.1.6. Benutzerinformationen und Einwilligungserklärungen

Im Hinblick auf die in § 10 Abs. 2 DSG NRW geforderte Transparenz ist es besonders wichtig, die an den Studien beteiligten Personen darüber aufzuklären, dass und wie ihre Daten gespeichert und verarbeitet werden, und wer hierfür verantwortlich ist. Ein entsprechenden Passus wird für jede Studie in die Patienteninformationsschrift integriert. Der Passus kann im Bedarfsfalle angepasst und aktualisiert werden. Darüber hinaus wird der Hinweis auf die Datenerfassung und Verarbeitung als expliziter Punkt in die Einwilligungserklärung der Patienten aufgenommen:

„Ich wurde darüber informiert und bin damit einverstanden, dass meine medizinischen Daten in pseudonymisierter Form gespeichert und analysiert werden. Ich wurde darüber informiert und bin damit einverstanden, dass meine persönlichen Kontaktdaten in Verantwortung des Clinical Trial Centers Aachen (CTC-A) treuhändisch vom Institut für Medizinische Informatik am Universitätsklinikum der RWTH Aachen gespeichert werden."

Personenbezogene Daten der Benutzer des CTC-A Study-Management-Tools, die über die üblichen Adressdaten hinausgehen, werden im Bereich der Aufwandserfassung erhoben und gespeichert. Das an den klinischen Studien arbeitende Personal ist angehalten, die für die jeweiligen Studien erbrachten Leistungen nach entsprechenden Leistungsklassen (standardisierte Arten von Aufwänden) stundenweise zu dokumentieren, damit eine Kostenanalyse der einzelnen Studien durchgeführt werden kann. Da die Personen nur ihre eigenen Zeiten ins System eingeben können und dies aus freien Stücken selbst tun, ist eine gesonderte Aufklärung darüber, dass diese Zeiten auch gespeichert und studienbezogen summiert werden, nicht erforderlich. Gemäß § 18 DSG NRW können entsprechende Übersichten (Reports) jederzeit aktuell erzeugt werden.

3.2. Paradigmen der Software-Eigenentwicklungen

Am IMI werden Softwarelösungen für das CTC-A anwenderspezifisch entwickelt, parametriert und betrieben [DD11]. Folgende Paradigmen sichern dabei die Qualität der Software hinsichtlich Verfügbarkeit, Integrität und Revisionsfähigkeit der Daten.

3.2.1. Einsatz eines Versionsmanagements

Alle Eigenentwicklungen für das CTC-A basieren auf der Programmiersprache Java. Zur strukturierten Ablage der am IMI erstellten Quellcodes wird das Quellcodeversionierungssystem Apache Subversion (SVN) verwendet. SVN ist eine freie Software zur Versionsverwaltung von Dateien und Verzeichnissen. Die Versionierung erfolgt in einem zentralen Projektarchiv (Repository) in Form einer linearen Revisionszählung. Für die Entwicklung der CTC-A Anwendungen speichert SVN alle vorgenommenen Ände-

rungen des Quellcodes und erlaubt so eine lückenlose Verfolgung der Entwicklungsschritte. Das Bestehen potentieller Sicherheitslücken und der Zeitpunkt der Behebung können ebenfalls zu jedem Zeitpunkt nachvollzogen werden. Weiterhin wird ein Anforderungsmanagement abgebildet, das die Nachverfolgbarkeit (Traceablity) von Funktionsänderungen ermöglicht, d.h. die Dokumentation der Anforderungen mit Rückwärtsverweis, welcher Code auf welche Anforderung hin erstellt wurde.

3.2.2. Systematische strukturierte Software-Tests

Weiterhin wird eine systematische Testumgebung in der Softwareentwicklung eingesetzt. Diese nutzt standardisierte Java-Technologien zur automatischen Prüfung der Programmfunktionalität. Hierzu wird JUnit-Framework mit entsprechenden Erweiterungen für die Entwicklungspakete GWT und ExtGWT eingesetzt. Durch die systematische Prüfung aller relevanten Programmbestandteile, die automatisiert vor jeder Veröffentlichung eines Updates stattfindet, sollen Fehler im Produktivsystem vermieden werden.

Darüber hinaus werden alle Neuentwicklungen im Rahmen eines Änderungsmanagements auch auf Pseudodaten getestet. Hierzu wird ein eigenes Testsystem vorgehalten mit einer modifizierte Spiegelung des Echtsystems auf einem weiteren virtuellen Server vorgehalten, auf dem dynamische Tests während der Programmausführung funktionsorientiert durchgeführt werden. Mit Positivtest wird zunächst versucht, die Anforderungen gemäß den Spezifikationen zu verifizieren. Negativtests werden eingesetzt, um die Robustheit der Anwendungen zu testen. Screenshots werden zur Dokumentation der Tests eingesetzt.

3.2.3. E-Mail-Services und Bug-Tracking

Für die multizentrische, dezentrale und translationale Struktur der an klinischen Studien beteiligten Personen in verschiedenen Organisationen bekommt die E-Mail als zentrales Kommunikationsmedium eine besondere Bedeutung. Dementsprechend unterstützt das CTC-A Study-Management-Tool die E-Mail-Kommunikation der registrierten Benutzer auf besondere Weise:

- *Automatische E-Mails*: Vom System werden automatische E-Mails mit standardisierten Texten an einzelne Benutzer oder Benutzergruppen versendet. Diese sind an bestimmte Events geknüpft. Beispielsweise wird ein definierter Benutzerkreis automatisch informiert, wenn im System eine neue Studie angelegt wurde. Weiterhin wird immer eine E-Mail versandt, wenn die Rechte eines Nutzers neu gesetzt oder modifiziert wurden.

- *Bug-Reports*: Das CTC-A bietet jedem Nutzer die Möglichkeit, auf Knopfdruck eine Fehlermeldung an die Systemadministratoren zu versenden. Neben dem individuellen Text der Fehlerbeschreibung werden automatisch Systemvariablen (Art der Netzanbindung, verwendetes Betriebssystem und Internet-Browser mit Versionsnummern, etc.) mit übertragen, die das Lokalisieren und Beheben des Fehlers vereinfachen und beschleunigen.

– *CTC-A Mitteilungen*: Die Benutzer mit Zugriffs-Level „Master“ können eine frei formulierbare E-Mail an alle Nutzer mit gleichem Zugriffslevel senden. Beispielsweise können alle aktiven Benutzer mit dieser Funktion über technische Probleme oder Ausfallzeiten des Systems informiert werden.

Trac ist ein Management-Tool für die Softwareentwicklung, mit dem sich u.a. auftretende Bugs erfassen und die Lösungswege managen und dokumentieren lassen. Dabei werden Bugs als Tickets formuliert, Entwicklern zugewiesen und der Fortschritt protokolliert. Trac ermöglicht, Bilder (Screenshots) an das Ticket zu heften. Tickets können entweder direkt aus dem CTC-A Study Managment Tool heraus, per Mail oder im Trac-System selbst angelegt werden.

3.2.4. Weitere Maßnahmen nach BSI

Die bislang präsentierten Methoden können im Hinblick auf Datenschutz und Datensicherheit als Kernparadigmen aufgefasst werden, die bereits zahlreiche Gefährdungspotentiale und die mit ihnen verbundenen Risiken verringern. Darüber hinaus stellt das BSI einen umfangreichen Maßnahmenkatalog zur Verfügung, um die Sicherheit und den Schutz personenbezogener Daten weiter zu verbessern. Gefährdungen und passende Maßnahmen sind dabei in sog. Kreuztabellen Baustein-spezifisch miteinander verknüpft. Diese Tabellen beinhalten zu jeder Maßnahme eine Qualifizierungsstufe (Siegel):

– *Einstieg (A)*: Diese Maßnahmen sind essentiell, vorrangig umzusetzen und bereits für das Auditor-Testat „IT-Grundschutz Einstiegsstufe“ erforderlich.

– *Aufbau (B)*: Diese Maßnahmen müssen für das Auditor-Testat „IT-Grundschutz Aufbaustufe“ umgesetzt sein. Eine zügige Realisierung ist anzustreben.

– *Zertifikat (C)*: Diese Maßnahmen müssen für das ISO 27001-Zertifikat umgesetzt sein und können zeitlich nachrangig umgesetzt werden.

– *Zusatz (Z)*: Diese Maßnahmen müssen weder für ein Auditor-Testat noch für das ISO 27001-Zertifikat auf Basis von IT-Grundschutz verbindlich umgesetzt werden.

– *Wissen (W)*: Diese Maßnahmen dienen der Vermittlung von Grundlagen und Kenntnissen im Datenschutz und in der IT-Sicherheit und sind ebenso optional.

Zur Gewährleistung des Datenschutzes und der Datensicherheit am CTC-A wurden alle Siegelstufen betrachtet und jede einzelne Maßnahme entsprechend durchlaufen. Mithilfe dieser Systematik konnten zahlreiche Schwachstellen (fehlende Türschlösser bzw. fehlende Vorgaben an Passwörter) bereits identifiziert und durch Änderung des Umsetzungskonzeptes schon im Vorfeld behoben werden. Diese Vorgehensweise ist somit auch bei ähnlichen Projekten unbedingt zu empfehlen.

3.3. Etablierung und Fortführung eines (Daten-)Sicherheits-Managementsystems

Datenschutz und IT-Sicherheit sind integrale Bestandteile des Qualitätsmanagementsystems des CTC-A. Etablierung und Fortführung der Sicherheitsrichtlinien werden im Qualitätshandbuch des CTC-A ausführlich beschrieben [D11]. Dort sind vor allem Organisation und Verantwortung von Technologien und Dienstleistungen geregelt.

Für das Qualitätsmanagements des CTC-A wird auf die kontinuierliche Verbesserung der Prozesse großer Wert gelegt. Erfahrungen daraus fließen wieder zurück in die Planung, so dass der bekannte Plan, Do, Check, Act (PDCA)-Regelkreis entsteht:

- *Qualitätsplanung (Plan)*: Der Ist-Zustand wird ermittelt. Die Rahmenbedingungen für das Qualitätsmanagement werden festgelegt und Konzepte sowie Abläufe erarbeitet. Dies ist im ersten Durchlauf mit den erstellten Lastenheften geschehen.

- *Qualitätslenkung (Do)*: Die in der Planphase gewonnenen Ergebnisse werden umgesetzt. Die Freigabe der Software erfolgt durch die Oberste Leitung des CTC-A, die von einem Sicherheitsmanagementbeauftragten für die elektronische Informationsverarbeitung unterstützt wird, der halbjährlich berichtet und ggf. notwendige Änderungen beantragt, umsetzt und überprüft. Allenfalls werden nur solche Verfahren in Betrieb genommen, bei denen keine Gefahren hinsichtlich des Datenschutzes bestehen oder für die geeignete organisatorische und technische Maßnahmen getroffen wurden, um diese Gefahren zu vermindern und die unerwünschten Situationen zu verhindern, vgl. § 10 (3) DSG NRW.

- *Qualitätssicherung (Check)*: Die Wirksamkeit der Maßnahmen zum Datenschutz wird unter Berücksichtigung sich verändernder Rahmenbedingungen und Entwicklungen der Technik regelmäßig überprüft. Die sich daraus ergebenden notwendigen Anpassungen werden zeitnah umgesetzt, vgl. § 10 (3) DSG NRW. Datenschutz- und IT-Sicherheitskonzepte unterliegen dem im Qualitätshandbuch beschriebenen Verteilungs- und Änderungsdiensten. Managementspezifische Sicherheitsaspekte werden regelmäßig bewertet, also alle Regelungen, Zuständigkeiten und die damit verbundenen Befugnisse und notwendigen Qualifikationen. Bei Migrationen innerhalb der fünf Realisierungsstufen der CTC-A Verfahren wird ein eigener PDCA-Zyklus durchlaufen und das Datenschutzkonzept aktualisiert.

- *Qualitätsgewinn (Act)*: Die aus der vorherigen Phase gewonnenen Informationen werden für Strukturverbesserungsmaßnahmen und Prozessoptimierung eingesetzt. Erfolge und Ergebnisse werden kommuniziert, das heißt, das so verbesserte Verfahren wird breitflächig etabliert. Bereits jetzt sind verbessernde und ergänzende Maßnahmen geplant. (I) Es sollen CTC-A Software-Benutzungsrichtlinien entworfen und verabschiedet werden. (II) In Zusammenarbeit mit dem Datenschutzbeauftragten werden Vorgehensweisen zur Auswertung der Protokolldateien erarbeitet und in Kraft gesetzt. (III) Als ergänzender Teil der CTC-A Hinweise zur IT-Nutzung wird ein Sicherheitsmeldeformular entwickelt und etabliert, dass zusätzlich zu dem E-Mail-Dienst von Benutzern, die auffällige Ereignisse beobachtet haben, strukturiert ausgefüllt und per Fax oder Hauspost direkt versendet werden kann.

4. Resümee und Ausblick

Die stufenweise Umsetzung von CTC-A eigenen IT-Verfahren zum administrativen und medizin-inhaltlichen Management klinischer Studien und der darin erhobenen Daten ist mit dem generischen Datenschutzkonzept der TMF datenschutzkonform umsetzbar. Damit ist der Einsatz des CTC-A Study-Management-Tools auch als gemeinsame IT-Plattform für einzelne dezentrale Studienzentren im Universitätsklinikum Aachen möglich und erscheint nach differenzierter Analyse des Datenschutzes und der Datensicherheit auch sinnvoll, denn er lässt eine erhebliche Verbesserung der Datenqualität und -sicherheit erwarten.

Eine Einführung in anderen Bereichen des Hauses wird vorab gründlich diskutiert werden. Auch die Umsetzung der einzelnen Phasen ist Work-in-Progress. Die abschließende Prüfung des IT-Sicherheitskonzeptes des CTC-A durch den Datenschutzbeauftragten der RWTH Aachen steht ebenfalls noch aus.

Literaturverzeichnis

[BS11] Bundesamt für Sicherheit in der Informationstechnik (BSI) Hrsg. IT-Grundschutz-Kataloge. 12. Ergänzungslieferung, Stand September 2011. Erhältlich unter: www.bsi.bund.de/grundschutz

[D11] Deserno V. Qualitätshandbuch. Clinical Trial Center Aachen (CTC-A), Aachen, 2011

[DD11] Deserno TM, Deserno V et al. IT-Unterstützung für translationales Management klinischer Studien auf Basis des Google Web Toolkits. German Medical Science, 2011.

[DG00] Der Innenminister des Landes Nordrhein-Westfalen (Hrsg). Gesetz zum Schutz personenbezogener Daten (Datenschutzgesetz Nordrhein-Westfalen - DSG NRW) in der Fassung der Bekanntmachung vom 9. Juni 2000, Düsseldorf, 2000.

[GG99] Der Innenminister des Landes Nordrhein-Westfalen (Hrsg). Gesetz zum Schutz personenbezogener Daten im Gesundheitswesen (Gesundheitsdatenschutzgesetz - GDSG NRW) in der Fassung vom 22. Februar 1994, §2 geändert durch Gesetz vom 17. Dezember 1999, Düsseldorf, 1999.

[R11] Der Rektor der RWTH Aachen (Hrsg). Satzung des Clinical Trial Center Aachen der Medizinischen Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen vom 20.05.2011. Mitteilung Nr. 2011/052, Dez. 1.0 der RWTH Aachen, 2011.

[RD06] Reng CM, Debold P, Specker C, Pommerening K. Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, 2006.

[SS00] Schulz JB, Spitzer K. Kooperationsvertrag zwischen dem Clinical Trial Center Aachen und dem Institut für Medizinische Informatik der RWTH Aachen. Aachen, 2010.

[W08] Wirth S. Hinweise zur Risikoanalyse und Vorabkontrolle nach dem Hamburgischen Datenschutzgesetz. Der Hamburgische Beauftragte für Datenschutz und Informationssicherheit, Hamburg, 2008.

[GV04] Der Bundesminister der Justiz (Hrsg.). Verordnung über die Anwendung der Guten Klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Arzneimitteln zur Anwendung am Menschen (GCP-Verordnung - GCP-V) in der Fassung vom 9. August 2004 (BGBl. I S. 2081), zuletzt geändert am 3. November 2006 (BGBl. I S. 2523).

[WG08] Wagner M, Glock J, Sariyar M, Borg A. PID Generator. Technical Manual. Version 1.2. Institut für Medizinische Biometrie, Epidemiologie und Informatik. Johannes-Gutenberg-Universität Mainz, 2008.