Clinical Communication Platform (CCP-IT)

Datenschutzkonzept

Projekt	Clinical Communication Platform (CCP-IT) im Deutschen Konsortium für Translationale Krebsforschung (DKTK)
Autoren	Andreas Borg[1] und Dr. Martin Lablans[2]
unter Mitarbeit von	Daniel Büttner, Dr. Corinna Eichelser, Kristina Ihrig, Dr. Sabrina Meister, Dr. Esther Schmidt, Diane Schneider, Barbara Uhl und Susanne Zabka
Träger	Deutsches Krebsforschungszentrum Im Neuenheimer Feld 280 69120 Heidelberg
Version	30. Juni 2017

[1]Universitätsmedizin der Johannes Gutenberg-Universität Mainz, Institut für Medizinische Biometrie, Epidemiologie und Informatik, 55101 Mainz

[2]Deutsches Krebsforschungszentrum, AG Verbundforschung, Medizinische Informatik in der Translationalen Onkologie, Im Neuenheimer Feld 280, 69120 Heidelberg

Inhalt

1. Einleitung

1.1 Zielsetzung

Vor dem Hintergrund des Wissens um das komplexe Zusammenwirken von individueller genetischer Disposition, Lebensstil und Umweltfaktoren für die Entstehung und den Verlauf von Krebserkrankungen verlangt die heutige Krebsforschung die Beobachtung von Krankheitsverläufen, individuellen Lebensgewohnheiten, und Umweltbedingungen über lange Zeiträume. Nur durch die Forschung in standortübergreifenden Verbünden mit ausreichenden Fallzahlen kann auch die Ursächlichkeit individueller genetischer Dispositionen erforscht werden.

Das DKTK hat sich zur Aufgabe gemacht, durch den Aufbau von effizienten translationalen Forschungseinheiten für anwendungsnahe Krebsforschung an bundesweit vernetzten Partnerstandorten die Erforschung der Entstehungsmechanismen von Krebserkrankungen sowie die Entwicklung optimierter Instrumente für eine gezielte Behandlung und für eine frühzeitige Erkennung von Krebserkrankungen voranzutreiben.

Die CCP-IT stellt dafür eine IT-Infrastruktur zur Verfügung, die es ermöglicht, vor allem Fallzahlen zu definierten Krankheitsbildern standortübergreifend zu ermitteln, aber auch Probanden für klinische Studien zu rekrutieren oder bereits vorliegende Daten für Forschungsfragestellungen zur Auswertung anzufordern. Dazu werden an den beteiligten Standorten Daten aus der klinischen Krebsdokumentation sowie zu dort vorhandenen Biomaterialproben erhoben und je nach Möglichkeit in einer zentralen Datensammlung oder lokal an den Standorten gesammelt. Um diese Daten für die Forschung nutzbar zu machen, stellt die CCP-IT eine zentrale Suchschnittstelle bereit, die es Forschern ermöglicht

- abzuschätzen, ob für ein Forschungsvorhaben im DKTK genügend Probanden mit den betrachteten Eigenschaften auffindbar sind,
- zu ermitteln, welche Institutionen passende Patienten behandeln oder behandelt haben und
- Anfragen zur Nutzung von medizinischen Daten und Biomaterialproben dieser Patienten für Forschungsvorhaben zu stellen.

Eine Besonderheit gegenüber anderen Forschungsverbünden besteht darin, dass auch Bestandsdaten, die vor der Einrichtung des DKTK im Behandlungskontext erhoben wurden, für die Forschung nutzbar gemacht werden sollen. Dieser Nutzung von Bestandsdaten sind enge datenschutzrechtliche Grenzen gesetzt, wenn sie nicht durch eine Patienteneinwilligung gedeckt ist. Andererseits könnte dadurch die sonst nötige Wartezeit zwischen Beginn der Datenerhebung und dem Erreichen eines für Forschungszwecke hinreichend großen Datenbestandes erheblich reduziert werden. Darüber hinaus ermöglicht der Einbezug von Bestandsdaten die Evaluation von Therapiefortschritten. Das Ziel dieses Datenschutzkonzepts ist, in dieser Situation, in der Anforderungen bezüglich des Datenschutzes besonders stark mit Wünschen hinsichtlich der Nutzbarmachung für die medizinische Forschung konkurrieren, einen wirksamen und den rechtlichen Anforderungen genügenden Datenschutz sicherzustellen.

1.2 Überblick über die Datenverarbeitung

In der CCP-IT werden Daten von Tumorpatienten, die an den teilnehmenden Kliniken (auch als „Standorte“ bezeichnet) behandelt werden, erhoben und verarbeitet. Sie werden zum größten Teil aus vorhandenen Datenverarbeitungssystemen (z.B. Krankenhausinformationssysteme und Software zur Tumordokumentation), in die zentralen Komponenten der CCP-IT eingebracht. Darüber hinaus erlaubt die Komponente „Lokales Datenmanagement“ (siehe Abschnitt 2.1) die manuelle Eingabe von Daten zu Biomaterialproben.

Grundsätzlich teilen sich die erhobenen datenschutzrelevanten Daten in medizinische und identifizierende Daten auf, die im Folgenden in Anlehnung an die TMF[1]-Datenschutzkonzepte als MDAT und IDAT bezeichnet werden. Die erhobenen MDAT umfassen

- klinische Daten, wie zum Beispiel codierte Diagnosen und Tumor-Klassifikationen (MDS-K),
- Daten zu Biomaterialproben (MDS-B),
- Hinreichend vergröberte demographische Daten (z.B. Geschlecht und Geburtsjahr).

Weiterhin ist vermerkt, an welchen Studien des DKTK der Patient[2] teilnimmt, sowie welche Experimente mit seinen Biomaterialproben bereits durchgeführt wurden.[3] Die IDAT enthalten demografische Daten, die eine eindeutige Identifikation des Patienten erlauben. Genauere Informationen zum Umfang dieser Daten finden sich in den Darstellungen der folgenden Abschnitte (insbesondere Kapitel 2.2, Abschnitt „Kontrollnummern-Erzeuger“), sowie im Anhang.

1.3 Rechtsgrundlage

In Hinblick auf die Rechtsgrundlage sowie die Prozesse der Datenverarbeitung ist zwischen folgenden zwei Patientengruppen zu unterscheiden:

1. Patienten, die der Verwendung ihrer Daten und/oder Biomaterialproben im DKTK explizit zugestimmt haben (im Folgenden *explizit eingewilligte Patienten*). Die dafür vorgesehene Einwilligung (siehe Anhang 1) deckt neben der Weitergabe von Daten an die zentrale Suche auch die Weitergabe von IDAT zur Erzeugung eines zentralen Pseudonyms ab (Kapitel 2.2, Abschnitt „Kontrollnummern-Erzeuger“). Rechtsgrundlage ist hier also die genannte informierte Einwilligung des Patienten (siehe Abschnitt 6.1). Auf der Grundlage dieser Einwilligung können MDAT in eine zentrale Komponente (die sogenannte *MDS-Datenbank*) in pseudonymisierter Form exportiert werden und können dort über die *zentrale Suche* von DKTK-Forschern durchsucht werden.
2. Patienten, die der Verwendung ihrer Daten im DKTK nicht explizit zugestimmt haben (im Folgenden *nicht explizit eingewilligte Patienten*). Dieser Fall liegt insbesondere bei Bestandsdaten vor, die aus der Zeit vor Errichtung des DKTK stammen und damit wesentlich zahlreicher als Daten von *explizit eingewilligten Patienten* sind. Rechtsgrundlage sind hier die am Standort anwendbaren landes- und bundesrechtlichen Datenschutzbestimmungen sowie die lokale Einwilligung in die Verwendung und Weitergabe klinischer Daten und/oder Biomaterialproben für standortübergreifende Forschungsprojekte in anonymisierter bzw. pseudonymisierter Form (vgl. Abschnitt 6.2).
   In Hinblick auf die Rechtsgrundlage für den Export von Daten in die MDS-Datenbank ist zu erwähnen, dass für diesen Zweck keine IDAT den Standort verlassen, sondern nur ein festgelegter, sparsamer Satz an MDAT; weiteres siehe Abschnitt 6.2 und Anhang 2.

Die MDAT beider Patientengruppen werden in eine lokale Komponente, den *Brückenkopf*, in pseudonymisierter Form importiert und dort gespeichert. Der Brückenkopf steht unter lokaler Kontrolle der behandelnden Einheit des jeweiligen Standorts, und auch nur dort kann (mit erheblichem technischen Aufwand) mithilfe des Pseudonyms auf die Identität des Patienten geschlossen werden. Auf diesen lokal verbleibenden MDAT können zwar im Rahmen einer sogenannten *dezentralen Suche* Suchanfragen von außerhalb des Standorts durchgeführt werden, die Ergebnisse dieser Suchanfragen sind aber im Rahmen der CCP-IT nur nach manueller Freigabe durch den Standort, der

[1] Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.
[2] „Patient“, „Arzt“ und ähnliche Begriffe bezeichnen in diesem Dokument Funktionsrollen und meinen Personen jeglichen Geschlechts. Auf eine gendergerechte Formulierung wurde aus Gründen der Lesbarkeit verzichtet.
[3] Letzteres dient dazu, die mehrfache Erhebung von Daten zu vermeiden.

die Daten besitzt, für den Anfragenden sichtbar. Rechtsgrundlage sind hier die am Standort anwendbaren landes- und bundesrechtlichen Datenschutzbestimmungen. (vgl. Abschnitt 6.2).

In beiden Fällen werden IDAT durch ein geeignetes Verfahren so verschlüsselt, dass eine Reidentifizierung des Patienten durch einen Dritten praktisch unmöglich ist, und nur in dieser faktisch anonymen Form in einer zentralen Datenbank gespeichert. Im Fall *nicht explizit eingewilligter Patienten* ist darüber hinaus der verwendete Schlüssel nur dem jeweiligen Standort bekannt, um bestimmte Angriffsszenarien zusätzlich zu erschweren. Eine umfassende Beschreibung der Prozesse zur Verarbeitung der IDAT findet sich im Abschnitt 3.2.

Nutzer bzw. Empfänger von Daten sind Forscher des DKTK[4]. Aus deren Sicht gibt es zwei verschiedene Zugriffswege (eine detaillierte Beschreibung der Komponenten und Prozesse wird in den folgenden Abschnitten gegeben):

- Datensätze, die in der zentralen MDS-Datenbank gespeichert sind, können von Forschern nach vorgegebenen Kriterien direkt durchsucht werden (*Zentrale Suche*, vgl. Abschnitt 3.4). Hier wird direkt ein Ergebnis in Form eines zusammenfassenden Überblicks über die gefundenen Datensätze zurückgeliefert. Die zentrale Suche erlaubt also mit ihrer Abfrage eine erste Einschätzung von im DKTK vorliegenden Daten und Biomaterialproben. Die Datensätze werden aus der MDS-Datenbank nicht weitergegeben, sondern die Anzahl geeigneter Patienten und/oder Proben aufgrund der ausgewählten Suchkriterien angezeigt.
- Nur lokal (d.h. im Brückenkopf) gespeicherte Datensätze können über die *Dezentrale Suche* angefragt werden (vgl. Abschnitt 3.5). Dabei wird keine zentrale Datenbank durchsucht, sondern die Suchkriterien werden an die Standorte übermittelt. Der Anfragende sieht zunächst kein Suchergebnis, sondern die auf die Suchkriterien passenden Datensätze werden nur dem Dateninhaber am Standort angezeigt. Am Standort wird die Anfrage auf inhaltliche Kriterien und rechtliche Zulässigkeit geprüft und, falls beides positiv ausfällt, manuell durch den Dateninhaber beantwortet. Diese Art der Suche kommt am ehesten einer klassischen schriftlichen Anfrage gleich, nur dass ihre Begutachtung und Beantwortung durch technische Hilfsmittel unterstützt werden.

1.4 Träger

Träger des Vorhabens ist das Deutsche Krebsforschungzentrum (DKFZ), Heidelberg.

2. Datenverarbeitende Komponenten

2.1 Brückenkopf

Der Brückenkopf dient dazu, die Daten eines Standorts in ein DKTK-kompatibles Format zu überführen und für die anderen Komponenten nutzbar zu machen. Seine Aufgaben sind:

- *Datenharmonisierung*: Daten werden im Brückenkopf so harmonisiert und dupliziert gespeichert, dass sie von den übrigen Komponenten der CCP-IT verstanden werden.
- *Sichtbarmachung für das DKTK*: z.B. für zentrale und dezentrale Suche.
- *Einhaltung von Datenhoheit*: Der Brückenkopf erlaubt dem Standort eine Teilnahme an der CCP-IT auch ohne „Upload auf Verdacht“ seiner patientenbezogenen Daten an eine externe Stelle, was Datenschutz und Datenhoheit fördert.

Der Brückenkopf besteht aus den folgenden, lokal in der Klinik installierten, Softwarekomponenten:

- *Lokales Datenmanagement*: Bereitet die in den lokalen Primärsystemen prinzipiell vorliegenden, aber unterschiedlich strukturierten Datenbestände für eine Nutzung im DKTK auf. Bietet für Biomaterialbanken an

[4] Siehe Abschnitt 4.2, „Teilnehmende Forscher“, für die genaue Bestimmung dieses Personenkreises.

den Standorten eine rudimentäre Proben- und Materialverwaltung samt Nutzer- und Gruppenverwaltung und Formularhandling. Diese Komponente entspricht funktional und technisch weitgehend einem Clinical Data Warehouse.

- *Teiler*: Leistet eine kontrollierte Freigabe der Datenbestände des lokalen Datenmanagements zur Nutzung durch Projekte des DKTK. Dabei kommen zwei sich ergänzende „Teilmethoden" zum Einsatz: Eine zentrale Suche gibt sofort erste Ergebnisse aus, dann folgt eine langsamere, aber dafür umfassendere dezentrale Suche.
- *Lokales Identitätsmanagement*: Stellt für die Pseudonymisierung sowohl von *explizit eingewilligten* als auch von *nicht explizit eingewilligten Patienten* eine einheitliche Schnittstelle bereit.

Die im Brückenkopf gespeicherten MDAT können prinzipiell alle Elemente des einheitlichen onkologischen Basisdatensatzes der Arbeitsgemeinschaft Deutscher Tumorzentren (http://www.tumorzentren.de/onkol-basisdatensatz.html) sowie Daten zu Biomaterialproben umfassen.

Diese Komponenten stehen unter Kontrolle des jeweiligen Zentrums, das heißt, die in diesen Komponenten gespeicherten Daten stehen weiter unter der Hoheit der Institution, in der sie erhoben wurden. Gegebenenfalls ist der Zugriff auf die behandelnde Einheit, z.B. die Fachabteilung, einzuschränken (vgl. auch Abschnitt 7).

2.2 Zentrales Identitätsmanagement

Pseudonymisierung ist ein zur Aufrechterhaltung eines hohen Datenschutzniveaus notwendiger Schritt, um den Patienten vor Rückidentifizierung zu schützen. Anstelle seiner identifizierenden Daten (IDAT) treten Pseudonyme. In der CCP-IT kommen folgende Arten von Pseudonymen zum Einsatz:

- *$S_{\#}ID$*: Ein lokaler Identifikator, der nur für den Standort # eindeutig ist und keine standortübergreifende Verknüpfung von Daten eines Patienten erlaubt.
- *$DKTK_{\#}ID$*: Ein lokaler Identifikator, der nur für den Standort # eindeutig ist. Im Gegensatz zur $S_{\#}$ID können $DKTK_{\#}$IDs, die an verschiedenen Standorten zu einem Patienten existieren, in der zentralen Patientenliste einander zugeordnet werden.
- *MDS-ID*: Ein Identifikator, der nur innerhalb der zentralen MDS-Datenbank eindeutig ist. In der zentralen Patientenliste kann die MDS-ID anderen IDs eines Patienten zugeordnet werden.

Im Falle eines *explizit eingewilligten Patienten* wird eine jeweils standortspezifische $DKTK_{\#}$ID erzeugt, welche auch (mithilfe der zentralen Patientenliste) eine Verfolgung des Patienten über Institutionsgrenzen hinweg erlaubt. Im Fall eines *nicht explizit eingewilligten Patienten* dürfen Klarnamen den Behandlungskontext des Standorts nicht verlassen; es wird dann ein lediglich die nur lokal vergleichbare $S_{\#}$ID, generiert.

Das zentrale Identitätsmanagement ermöglicht eine datenschutzgerechte Zusammenführung („Record Linkage") der von mehreren Standorten gesendeten patientenbeziehbaren Daten. Dazu werden drei Methoden/Werkzeuge kombiniert, vgl. folgende Unterabschnitte und **Abbildung 1**.

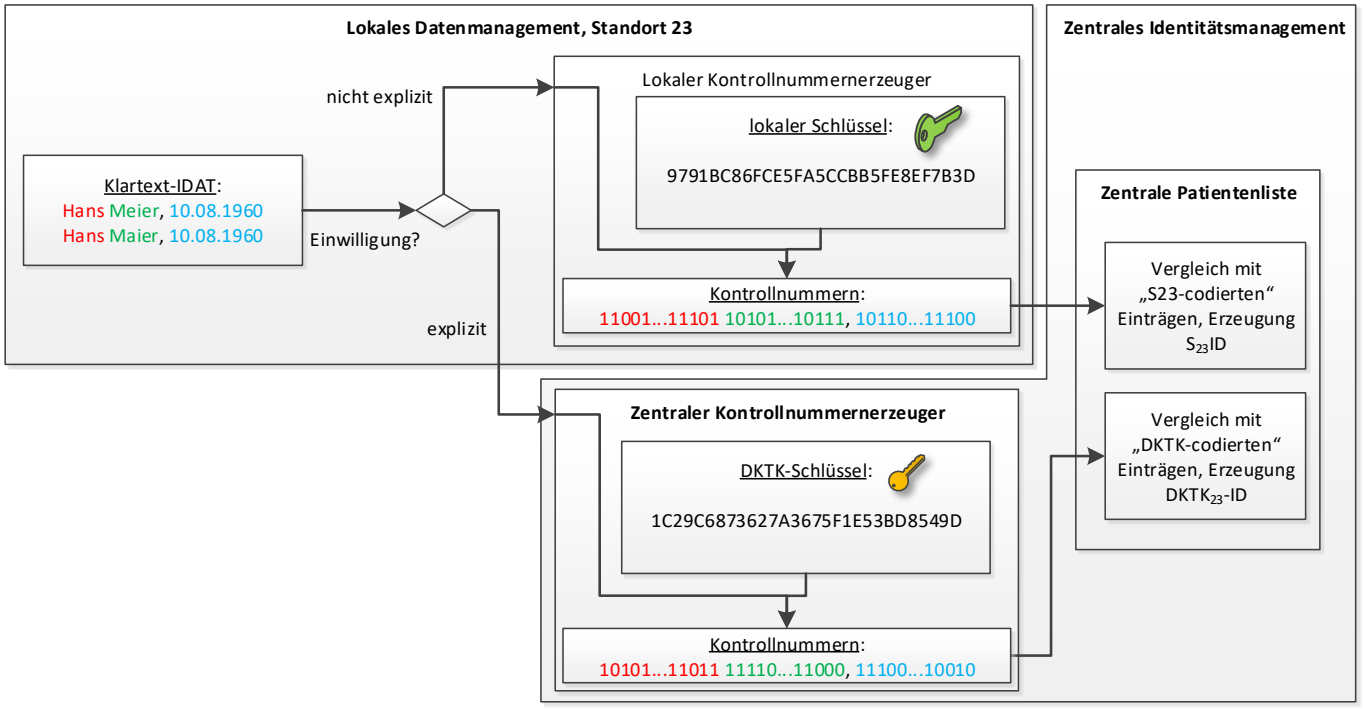

Abbildung 1 – Beispielhaftes Record Linkage-Verfahren.

Kontrollnummern-Erzeuger

Wandelt Identifikationsdaten (Klartext-IDAT) durch eine spezielle Einwegverschlüsselung in unlesbare und nicht rückführbare, aber immer noch gewichtet vergleichbare Zeichenketten („Kontrollnummern“[5]) um. Um Wörterbuchattacken zu vermeiden, wird dabei ein Keyed-Hash Message Authentication Code verwendet, wobei der verwendete Schlüssel (im folgenden „Geheimnis“) für jede Instanz des Kontrollnummern-Erzeugers eindeutig und nur dort bekannt ist.[6] Daraus folgt die Voraussetzung eines organisatorisch unabhängigen Betriebs dieses Moduls.

Die an den Kontrollnummern-Erzeuger gesendeten IDAT bestehen aus folgenden Attributen:

- Vorname
- Nachname
- Frühere Namen (z.B. Geburtsname bei Namensänderung durch Heirat)
- Geburtsdatum, aufgetrennt in die Komponenten Tag, Monat und Jahr
- Staatsangehörigkeit
- Geschlecht

Für die Pseudonymisierung *explizit eingewilligter Patienten* kommt ein zentraler Kontrollnummern-Erzeuger zum Einsatz. Die empfangenen Daten (IDAT) werden vom Kontrollnummernerzeuger nicht gespeichert. Als weitere Sicherheitsmaßnahme (z.B. gegen Angriffe mit Probeverschlüsselungen), werden die dort erzeugten Kontrollnummern nicht an den Brückenkopf zurückgegeben, sondern direkt an die zentrale Patientenliste übermittelt. Die $DKTK_{\#}ID$ wiederum wird nie dem zentralen Kontrollnummern-Erzeuger bekannt gemacht (Details siehe Abschnitt

[5] Streng genommen sind das also keine Nummern. Idee und Name wurden (in angepasster Form) aus dem Bereich epidemiologischer Krebsregister übernommen.
[6] Für technische Details siehe: Schnell R, Bachteler T, Reiher J: Privacy-preserving record linkage using Bloom filters. BMC Medical Informatics and Decision Making 2009, 9:41. http://www.biomedcentral.com/1472-6947/9/41

3.2). Bei *nicht explizit eingewilligten Patienten* erfolgt die Kontrollnummernerzeugung im lokalen Identitätsmanagement.

Patientenliste

Vergibt für eine Gruppe von Kontrollnummern eines Patienten einen Personenidentifikator (DKTK#ID, S#ID, MDS-ID). Ihr Kontrollnummern-Matcher kann dabei Patienten selbst bei abweichender Schreibweise und im Fall der DKTK#ID auch aus unterschiedlichen Standorten wiedererkennen.

Manuelles Linken

Eine Schnittstelle erlaubt einem Menschen, Ergebnisse des automatischen Matching zu überprüfen und ggfls. zu korrigieren, d.h. Duplikate zusammenzuführen oder fälschlicherweise zusammengeführte Datensätze zu trennen. Hierfür werden die Matchgewichte (Vergleichswerte zwischen den einzelnen Attributen von zu prüfenden Patienten) angezeigt und es besteht die Möglichkeit, zur Entscheidungsfindung auf medizinische Daten zuzugreifen. Diese Komponente ist deshalb beim Betreiber der zentralen MDS-Datenbank, wo ohnehin MDAT gespeichert sind, angesiedelt.

2.3 Zentrale MDS-Datenbank

Die zentrale MDS-Datenbank nimmt mithilfe einer Webschnittstelle die von den Teilern verschickten MDAT entgegen und verwahrt sie in einer Datenbank. Die Speicherung erfolgt zusammen mit der MDS-ID, um die Zuordnung verschiedener Datensätze eines Patienten zueinander möglich zu machen. Die MDAT umfassen Meldedatensätze zweier Klassen:

- *MDS-K*: Klinische Daten aus der Tumordokumentation. Diese beinhalten Daten zum Patienten, zum Primärtumor, zur Primärtherapie, zum Ansprechen und zum Vitalstatus und basieren auf dem onkologischen Basisdatensatz der Arbeitsgemeinschaft deutscher Tumorzentren (ADT und GEKID)[7].
- *MDS-B*: Metadaten zu Biomaterialproben. Der MDS-B umfasst Auskunft über das Vorhandensein von Biomaterial von Patienten und allgemeine Informationen zur Beschreibung des Biomaterials (z.B. Gewebetyp, Probenart).

Autorisierte Nutzer (siehe auch Abschnitt 5.2, „Authentifizierung“) können in einer Suchmaske mittels Suchkriterien aus MDS-K und MDS-B Abfragen auf diesem Datenbestand ausführen; diese Funktion wird im Abschnitt 3.4 („Zentrale Suche“) beschrieben. Im Rahmen der Protokollierung von Suchvorgängen können identifizierende Daten des zugreifenden Forschers und seine Eingaben in das System, bspw. Suchabfragen, gespeichert werden.

2.4 Suchbroker für die dezentrale Suche

Der Suchbroker für die dezentrale Suche stellt eine Schnittstelle zur Formulierung von Anfragen zur Verfügung und verwaltet diese Anfragen. Er verarbeitet keine personenbezogenen Daten von Patienten. Personenbezogene Daten von zugreifenden Benutzern können im Rahmen der Projektverwaltung (vgl. Abschnitt 3.5) und Protokollierung gespeichert werden.

2.5 Metadata Repository

Das Metadata Repository (MDR) speichert die Bedeutung (Semantik) sämtlicher im DKTK verwendeten (Nutz-) Datenelemente. Es bietet ein kontrolliertes Vokabular (Syntax) und kann maschinenlesbare, strukturierte Aussagen

[7] www.tumorzentren.de/onkol-basisdatensatz.html

über Datenelemente machen, bspw. konzeptuelle Domänen oder Wertebereiche. Hier sind auch die Meldedatensätze definiert. Da das MDR keine personenbezogenen Daten verarbeitet, wird innerhalb dieses Datenschutzkonzepts nicht weiter darauf eingegangen.

3. Datenverarbeitende Prozesse

3.1 Import in Brückenkopf

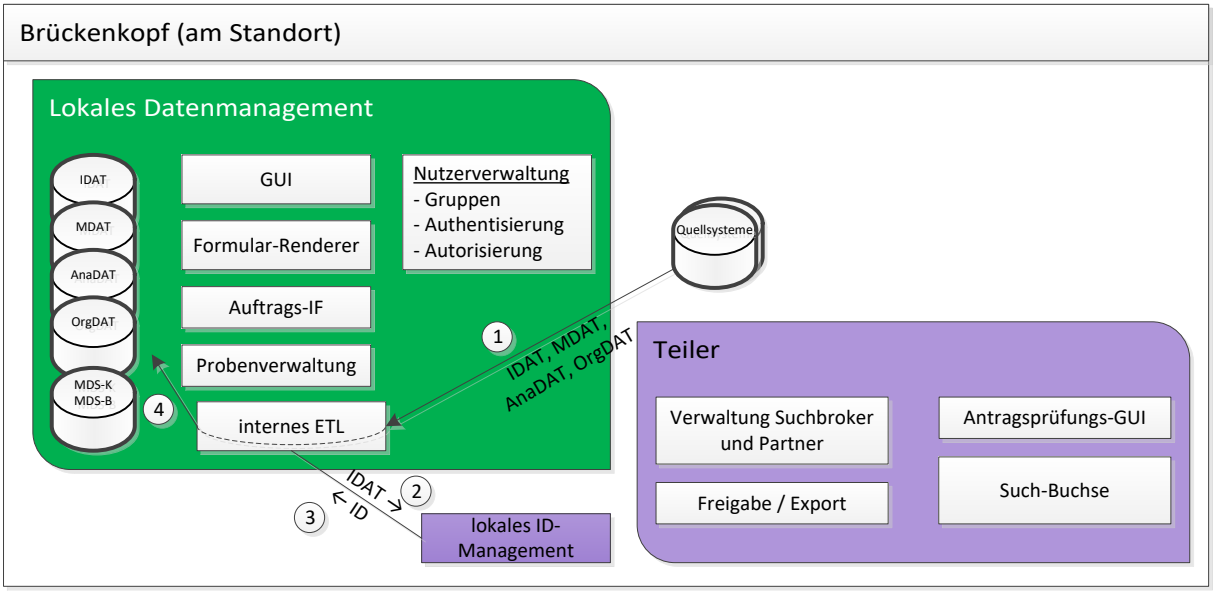

Abbildung 2 - Import von Daten in den Brückenkopf.

Abbildung 2 zeigt, wie Daten aus Quellsystemen eines Standorts in den Brückenkopf importiert werden:

1. Identifizierende, medizinische und Probendaten werden aus mehreren Quellsystemen durch einen ETL-Prozess extrahiert.
2. Identifizierende Daten werden mithilfe des lokalen Identitätsmanagements mit einem primären Patientenidentifikator erster Stufe versehen (Details siehe Abschnitt 3.2, „Pseudonymisierung“).
3. Der ETL-Prozess ordnet den Identifikator dem Datensatz zu.
4. Daten werden im lokalen Datenmanagement abgelegt.

3.2 Pseudonymisierung

Aus Datenschutzgründen entstehen zwei verschiedene Datenflüsse, je nachdem ob für den zu pseudonymisierenden Patienten eine DKTK-spezifische Patienteneinwilligung vorliegt (*explizite Einwilligung*) oder nicht (*ohne explizite Einwilligung*). Die Fälle unterscheiden sich wie folgt:

- Fall *explizite Einwilligung*: IDAT dürfen den Standort verlassen. Es wird ein lokales, aber mithilfe der zentralen Patientenliste verbundweit verknüpfbares Pseudonym ($DKTK_{\#}ID$) erzeugt, das eine Zuordnung des Patienten auch über Standortgrenzen hinweg erlaubt.
- Fall *ohne explizite Einwilligung*: Es dürfen keine IDAT den Standort verlassen. Es wird ein lokales Pseudonym ($S_{\#}ID$) erzeugt, d.h. eines, das nur innerhalb des Standorts zugeordnet werden kann. Man kann es etwa dazu verwenden, um passendes Biomaterial zu einem klinischen Datensatz zu finden. Es ist aber nicht möglich, einen Patienten über Standortgrenzen hinweg zuzuordnen.

Abbildung 3 zeigt, wie die Pseudonymisierung im Rahmen des ETL-Prozesses genutzt wird um ein Pseudonym zu erhalten.

1. Das lokale Identitätsmanagement erhält die Klartext-IDAT eines Patienten.
2. Klartext-IDAT werden übertragen an den Kontrollnummern-Erzeuger…
   a. im Fall *explizite Einwilligung*: …des zentralen Identitätsmanagements.
   b. im Fall *ohne explizite Einwilligung*: …des lokalen Identitätsmanagements.
3. Der jeweilige empfangende Kontrollnummern-Erzeuger errechnet aus den Klartext-IDAT und seinem Geheimnis Kontrollnummern (KN) und verwirft die IDAT.
4. Nur im Fall *explizite Einwilligung*: Der zentrale KN-Erzeuger übermittelt die (DKTK-weit vergleichbaren) Kontrollnummern an die Patientenliste. Diese erstellt ein temporäres Kontrollnummern-Ticket (KNTKT) und gibt dies an das lokale Identitätsmanagement zurück.
5. Das lokale Identitätsmanagement übermittelt die erzeugten Kontrollnummern, oder, im Fall *explizite Einwilligung* das eben erhaltene Ticket KNTKT, an die Patientenliste des zentralen Identitätsmanagements. Durch die Verwendung des Tickets bleiben die DKTK-weit vergleichbaren KN dem lokalen Identitätsmanagement verborgen.
6. Das zentrale Identitätsmanagement gleicht die erhaltenen Kontrollnummern mit den bestehenden ab (KN-Matcher). Im Falle eines Treffers wird eine bestehende ID ($DKTK_{\#}ID$ bzw. $S_{\#}ID$) zurückgegeben; falls kein passender Datensatz gefunden wird, wird mithilfe des Pseudonym-Generators eine neue ID erzeugt und zusammen mit den Kontrollnummern in der Datenbank gespeichert.

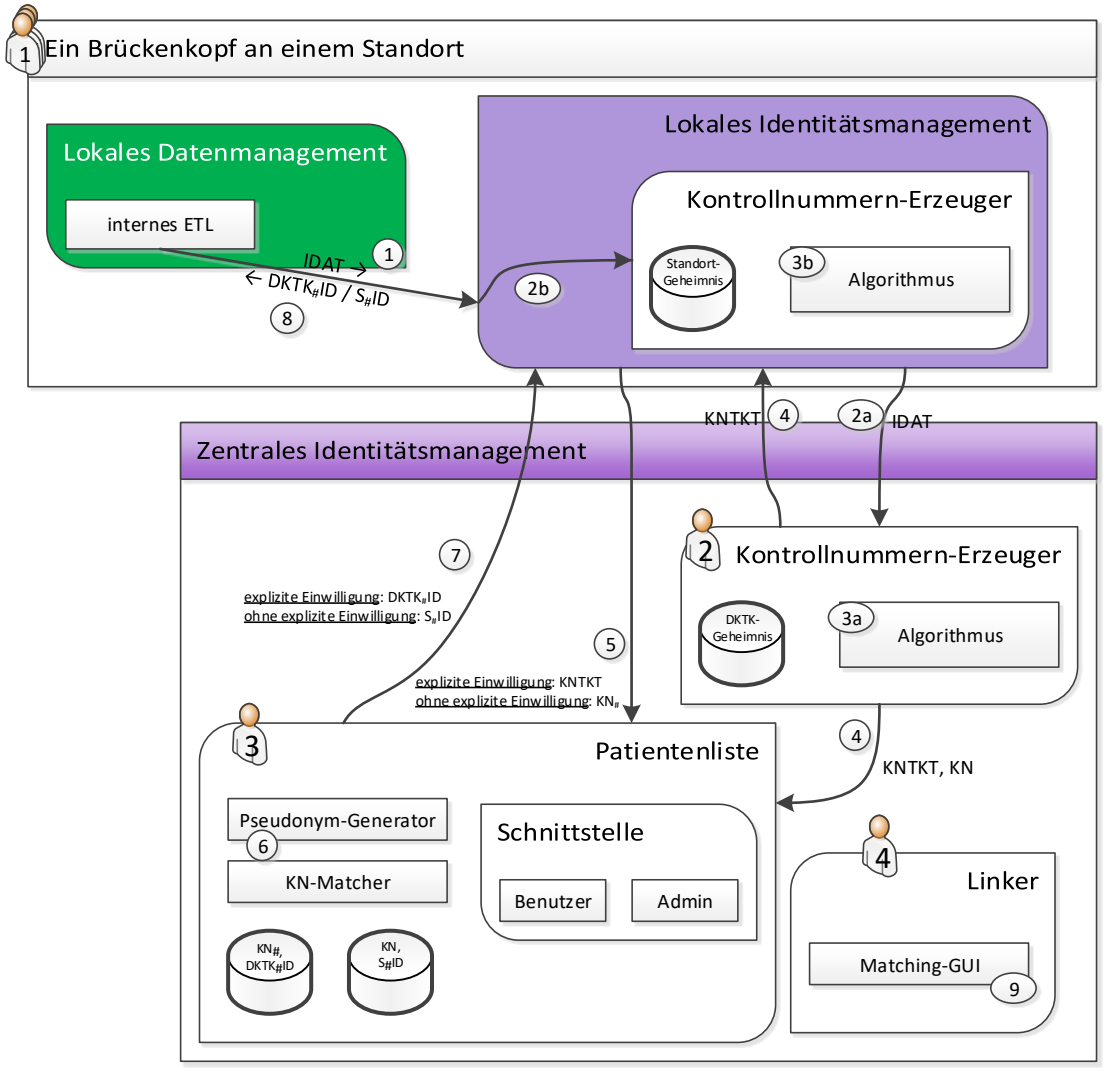

Abbildung 3 – Erzeugung lokaler und verbundweiter Pseudonyme.

7. Das lokale Identitätsmanagement erhält als Antwort auf seine Anfrage die ID. Es erfährt nicht, ob der Patient bereits bekannt war.
8. Die ID wird als Antwort auf die ursprüngliche Anfrage zurückgegeben. Es ist nicht erkennbar, ob der Patient bereits bekannt war. Es ist aber erkennbar, ob es sich um eine lokale S#ID oder um eine verbundweit verknüpfbare DKTK#ID handelt.
9. (optional) Sollte die Patientenliste sich nicht sicher sein, ob ein Patient schon in der Liste steht oder nicht, erfährt das der für das Record Linkage verantwortliche Administrator (4). Er entscheidet manuell auf Basis von Matchgewichten (aber keinen IDAT!) und – falls vorhanden – den in der zentralen MDS-Datenbank gespeicherten MDAT, ob es sich um denselben Patienten handelt.

Erläuterung: Geheimnisse und Gültigkeit der Pseudonyme

Die Fälle *explizite Einwilligung* und *ohne explizite Einwilligung* unterscheiden sich genau in diesem Geheimnis: Das DKTK-weite Geheimnis, das genau dem Kontrollnummern-Erzeuger des zentralen Identitätsmanagements bekannt ist (Fall *explizite Einwilligung*), erlaubt beim Matchen durch die zentrale Patientenliste eine Zuordnung des Patienten über Standortsgrenzen hinweg. Die lokalen Geheimnisse hingegen, die genau den Kontrollnummern-Erzeugern der lokalen Identitätsmanagement-Instanzen in allen Brückenköpfen desselben Standorts bekannt sind (Fall *ohne explizite Einwilligung*), erlauben nur einen Abgleich der Patienten desselben Standorts. Das ist nötig um z.B. zu einem gegebenen klinischen Datensatz das passende Biomaterial zu finden. Lokale Kontrollnummern werden, obgleich sie nur lokal Sinn ergeben, durch die zentrale Patientenliste abgeglichen. Das erspart die lokale Installation einer Patientenliste, wahrt aber weiterhin den Datenschutz, da die Kontrollnummern per Konstruktion eine Reidentifikation außerhalb des Standorts nahezu unmöglich machen.

3.3 Upload in zentrale MDS-Datenbank

Upload von MDAT *explizit eingewilligter Patienten*

Der Upload in die zentrale MDS-Datenbank erfolgt automatisch nach einem festen Zeitintervall (in der Regel täglich). Dabei werden aktuelle (d.h. seit dem letzten Upload hinzugekommene) MDAT von *explizit eingewilligten Patienten* gemäß der Meldedatensätze MDS-K und MDS-B vom Teiler aus dem lokalen Datenmanagement ausgelesen und über eine sichere HTTPS-Verbindung in die zentrale MDS-Datenbank exportiert.

Um die richtige Zuordnung der Daten in der MDS-Datenbank zu gewährleisten, werden diese beim Export mit Hilfe der zentralen Patientenliste umpseudonymisiert, d.h. die DKTK#ID wird durch die zugehörige MDS-ID ersetzt. Dazu kommt ein asymmetrisches Verschlüsselungsverfahren zum Einsatz, wobei der private Schlüssel nur der zentralen MDS-Datenbank bekannt ist. Der Ablauf ist wie folgt (vgl. das Sequenzdiagramm in Abbildung 4):

1. Der Teiler übermittelt die DKTK#ID des betreffenden Patienten an die zentrale Patientenliste.
2. Die zentrale Patientenliste bestimmt die zugehörige MDS-ID.
3. Die zentrale Patientenliste verschlüsselt die MDS-ID asymmetrisch mit dem öffentlichen Schlüssel (*MDS-DB-Pub*) der zentralen MDS-Datenbank.
4. Die verschlüsselte MDS-ID (*MDS-ID**) wird an das lokale Datenmanagement zurückgegeben.
5. Verschlüsselte MDS-ID und MDAT werden vom lokalen Datenmanagement an die zentrale MDS-Datenbank übermittelt.
6. Die zentrale MDS-Datenbank entschlüsselt die MDS-ID mit ihrem privaten Schlüssel (*MDS-DB-Priv)*.
7. Die zentrale MDS-Datenbank speichert MDS-ID und MDAT.
8. Die erfolgreiche Bearbeitung wird dem lokalen Datenmanagement zurückgemeldet.

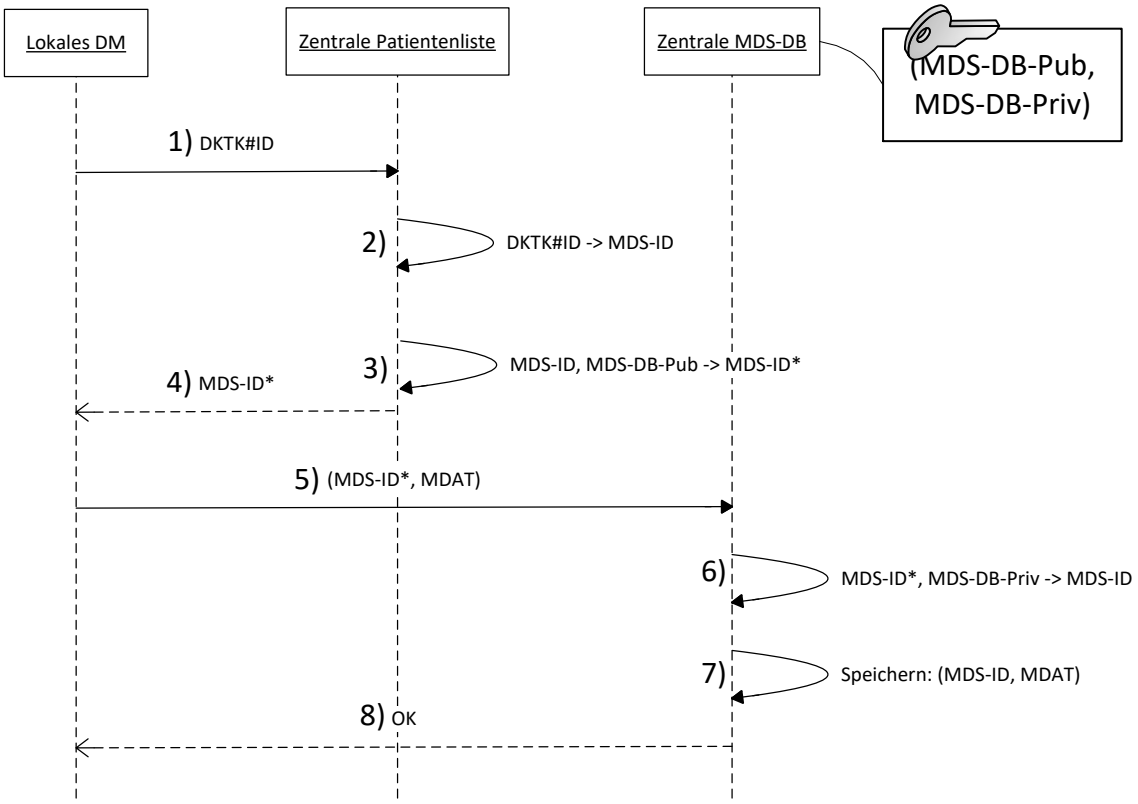

Abbildung 4 – Umpseudonymisierung beim Upload in die zentrale MDS-Datenbank

Um Wörterbuchattacken zu vermeiden, wird durch Randomisierung im Verschlüsselungsalgorithmus sichergestellt, dass die wiederholte (d.h. zeitlich hintereinander erfolgende) Verschlüsselung der gleichen MDS-ID verschiedene Chiffrate erzeugt.

Die übermittelten Daten können von den Administratoren von Brückenkopf und MDS-Datenbank im Rahmen von Wartungsarbeiten (vgl. Abschnitt 4.3 „Zugriff durch Systemadministratoren“) eingesehen werden. Ansonsten erfolgt durch den Upload selbst keine Sichtbarmachung, die Daten stehen mit dem Upload aber für die Zentrale Suche (Abschnitt 3.4) zur Verfügung.

Upload von MDAT *nicht explizit eingewilligter Patienten*

Falls lokale Datenschutzvoraussetzungen dies erlauben, kann im Brückenkopf auch ein Upload von Datensätzen *nicht explizit eingewilligter Patienten* an die zentrale MDS-Datenbank erfolgen. Dies kann zum Beispiel möglich sein, wenn am Standort eine lokale Patienteneinwilligung die Weitergabe pseudonymisierter oder anonymisierter Daten zu Forschungszwecken erlaubt und die Daten anonym weitergeben werden (näheres zu den datenschutzrechtlichen Aspekten siehe, Abschnitt 6.2 und Anlage 2). Diese Funktion erfordert eine zweistufige Freischaltung:

1. Generelle Freischaltung in der Konfiguration des Brückenkopfs. Dies erfordert den administrativen Zugriff auf Betriebssystemebene.
2. Aktivierung und Bestätigung durch den Administrator in der Benutzeroberfläche des Brückenkopfs (vgl. Abbildung 5). Das entsprechende Bedienelement erscheint nur, wenn die Freischaltung auf Systemebene (Punkt 1) erfolgt ist.

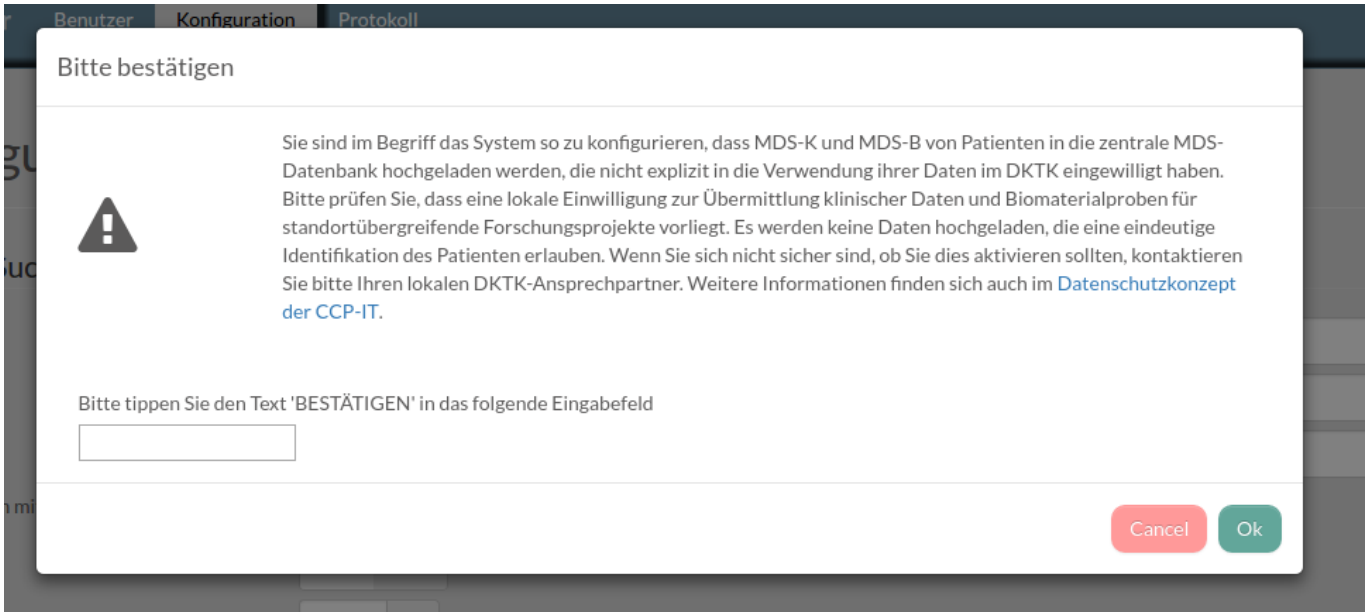

Abbildung 5 – Aktivierung des Uploads lokal eingewilligter Patienten im Brückenkopf

Der Upload erfolgt technisch wie im Fall explizit eingewilligter Patienten, an Stelle der DKTK#ID wird bei lokal eingewilligten Patienten aber die S#ID übermittelt, um die verschlüsselte MDS-ID abzufragen. In diesem Fall ist die MDS-ID nicht standortübergreifend vergleichbar. Durch einen im Teiler konfigurierbaren Filter wird außerdem sichergestellt, dass nur für das DKTK relevante Fälle hochgeladen werden.

3.4 Zentrale Suche

Mit der zentralen Suche können DKTK-Forscher den Datenbestand der zentralen MDS-Datenbank durchsuchen und im Sinne einer Machbarkeitsanalyse abfragen, ob Daten und Proben vorhanden sind, die für ein Forschungsvorhaben relevant sein könnten. Die zentrale MDS-Datenbank stellt dafür ein Webformular bereit, in dem Attribute der Meldedatensätze nach vorgegebenen Werten oder per Freitextsuche durchsucht werden können. Mehrere Suchattribute können durch logische Operatoren frei kombiniert werden.

Nach Ausführen der Suchabfrage erhält der anfragende Forscher maximal folgende Informationen[8]:

- Die Anzahl der Patienten bzw. Proben, die die Suchkriterien erfüllen.
- Die Altersverteilung der gefundenen Patienten in 10-Jahres-Intervallen.
- Die Geschlechtsverteilung (Anzahl Patienten männlichen / weiblichen / unbekannten Geschlechts).
- Die Verteilung auf die Standorte.
- Kontaktdaten der Standorte, an denen die Daten erhoben wurden.

Die Datensätze selbst werden nicht übermittelt. Ein eventueller Zugriff auf Daten durch den anfragenden Forscher sowie die datenschutzrechtliche Grundlage dafür wird zwischen diesem und dem jeweiligen Dateneigentümer (=Standort) verhandelt und findet außerhalb der zentralen MDS-Datenbank statt.

3.5 Dezentrale Suche

Wie die zentrale Suche dient die dezentrale Suche dem Auffinden von „passenden“ Patienten und Proben für ein Forschungsvorhaben, berücksichtigt im Gegensatz dazu aber auch Patienten, deren Daten nicht an die zentrale

[8] Derzeit sind die ersten drei Punkte umgesetzt, die weiteren Punkte geben den maximalen Umfang eventueller Erweiterungen wieder.

MDS-Datenbank übermittelt wurden. Da bei diesen angenommen werden muss, dass keine Rechtsgrundlage für die Übermittlung von Daten aus dem zuständigen Standort heraus gegeben ist, erfolgt hier keine automatische Übermittlung von Suchergebnissen ohne Freigabe des datenhaltenden Standorts.

Das Suchformular der dezentralen Suche wird vom Suchbroker für die dezentrale Suche bereitgestellt. Sie erlaubt nicht nur die Suche nach den Meldedatensätzen, sondern allen im Metadata Repository (vgl. 2.5) abgelegten Begriffen; zusätzlich sind Ergänzungen im Freitext (Prosa) möglich. Der anzufragende Datensatz ist hier also prinzipiell unbeschränkt. Die Anfrage wird zunächst gespeichert und dem anfragenden Forscher lediglich ihre Speicherung mitgeteilt. Die Teiler der Standorte rufen in regelmäßigen Abständen neu hinzugekommene Anfragen vom Suchbroker ab und ermitteln, welche Datensätze im lokalen Datenmanagement den Suchkriterien entsprechen. Der Inhalt der Anfrage sowie die gefundenen Datensätze können an jedem Standort von einer dazu berechtigten Person eingesehen werden. Diese kann nun den anfragenden Forscher kontaktieren, um eine mögliche Weitergabe von Daten oder Proben zu vereinbaren. Dieser Vorgang erfolgt wiederum außerhalb der CCP-IT und die damit verbundenen datenschutzrechtlichen Fragen müssen im Einzelfall von den beteiligten Personen geklärt werden.

4. Organisatorische Rahmenbedingungen

Die datenverarbeitenden Personen und Institutionen sowie Datenempfänger in der CCP-IT verteilen sich auf die Betreiber der zentralen Komponenten sowie die am DKTK teilnehmenden Standorte.

4.1 Betrieb der Komponenten

Der Betrieb der Brückenköpfe erfolgt durch die im DKTK vertretenen Partnerstandorte:

- Berlin: Charité Comprehensive Cancer Center
- Dresden: Universitätsklinikum Dresden, Technische Universität Dresden
- Essen / Düsseldorf: Westdeutsches Tumorzentrum am Universitätsklinikum Essen, Heinrich-Heine-Universität Düsseldorf
- Frankfurt / Mainz: Universitäres Centrum für Tumorerkrankungen Frankfurt, Johann Wolfgang Goethe-Universität Frankfurt, Georg-Speyer-Haus - Chemotherapeutisches Forschungsinstitut in Frankfurt am Main, Krankenhaus Nordwest Frankfurt, Universitätsmedizin Mainz
- Freiburg: Tumorzentrum - Comprehensive Cancer Center Freiburg, Albert-Ludwigs-Universität Freiburg, Universitätsklinikum Freiburg, Max-Planck-Institut Freiburg
- Heidelberg: Deutsches Krebsforschungszentrum (Kernzentrum), Universitätsklinikum Heidelberg, Ruprecht-Karls-Universität Heidelberg, Nationales Centrum für Tumorerkrankungen, Paul-Ehrlich-Institut (assoziierter Partner)
- München: Klinikum der Universität München, Klinikum rechts der Isar der TU München
- Tübingen: Universitätsklinikum Tübingen, Eberhard-Karls-Universität

Der Betrieb der zentralen Komponenten erfolgt an folgenden Stellen:

- Zentrale Patientenliste: Abteilung Medizininformatik, Institut für medizinische Biometrie, Epidemiologie und Informatik, Universitätsmedizin der Johannes Gutenberg-Universität Mainz.
- Zentraler Kontrollnummernerzeuger: Dezernat 7 – Informations- und Kommunikationstechnologie (DICT), Universitätsklinikum Frankfurt.
- Zentrale MDS-Datenbank und zentrale Suche: Abteilung Theoretische Bioinformatik, Deutsches Krebsforschungszentrum Heidelberg.

- Metadata Repository, Authentifizierungsdienst, Suchbroker für die dezentrale Suche und Wartungsdienste für Brückenköpfe: Arbeitsgruppe Verbundforschung, Abteilung Medizinische Informatik in der Translationalen Onkologie, Deutsches Krebsforschungszentrum Heidelberg.

4.2 Teilnehmende Forscher

Teilnehmende Forscher sind die Personen, die über die zentrale und die dezentrale Suche Anfragen an das System stellen können. Generell können alle Mitglieder der DKTK-Standorte als teilnehmende Forscher die CCP-IT nutzen, wobei jeder Standort selbst entscheidet, welche seiner Mitglieder eine Zugangsberechtigung erhalten (siehe auch Abschnitt 5.2, „Authentifizierung“).

Wissenschaftler, die nicht Mitglieder eines DKTK-Standorts sind, können auf Antrag vom Ausschuss für Datenschutz eine Zugangsberechtigung erhalten. Diese ist angemessen zu befristen.

4.3 Zugriff durch Systemadministratoren

Die in der CCP-IT gespeicherten Daten können prinzipiell von den Administratoren der verwendeten IT-Infrastruktur eingesehen werden. Zugriffe auf die Daten durch Administratoren dürfen nur erfolgen, wenn dies zur Erfüllung ihrer Aufgaben zwingend erforderlich ist. Alle Administratoren sind auf diesen Grundsatz und auf ihre Pflicht zur Verschwiegenheit hinzuweisen[9].

4.4 Ausschuss für Datenschutz

Vom Lenkungsausschuss des DKTK wird ein Ausschuss für Datenschutz eingesetzt. Dieser erfüllt insbesondere folgende Aufgaben:

- Prüfung und Bewilligung von Anträgen externer Forscher[10] für die Nutzung der CCP-IT (zentrale und dezentrale Suche).
- Prüfung und Bewilligung von Anträgen auf Export medizinischer Daten und/oder Bereitstellung von Biomaterial für externe Forschungsprojekte.
- Prüfung und Bewilligung von Anträgen auf die Benachrichtigung betroffener Patienten über Forschungsergebnisse.

Darüber hinaus ist der Ausschuss für Datenschutz erster Ansprechpartner für datenschutzrechtliche Angelegenheiten.

Der Ausschuss für Datenschutz wird so besetzt, dass jeder der DKTK-Standorte darin vertreten ist. Zu den Mitgliedern zählen mindestens:

- Ein Arzt, der vorwiegend in der Behandlung betroffener Patienten tätig ist.
- Ein Wissenschaftler, der mit den in der CCP-IT verwalteten Daten (oder Daten ähnlichen Typs) forscht.
- Ein Datenschutzbeauftragter oder ein mit dem Thema Datenschutz vertrauter Jurist.

Zusätzlich kann ein Vertreter der Entwickler der CCP-IT in beratender Funktion hinzugezogen werden.

[9] Dies sollte in der Regel im Rahmen des Arbeitsverhältnisses an der zuständigen Institution ohnehin geschehen sein.
[10] D.h. Personen, die nicht Mitglied eines DKTK-Standorts sind.

5. Maßnahmen zum Datenschutz

5.1 Informationelle Gewaltenteilung

Konsequent durchgeführt wird eine informationelle Gewaltenteilung, angedeutet durch die nummerierten Personenbilder (in Abbildung 3). Das bedeutet, dass die Komponenten mit unterschiedlichen Nummern logisch, physikalisch und organisatorisch getrennt voneinander laufen, was die Gefahr eines Datenlecks verringert:

- Das zentrale Identitätsmanagement (,) wird getrennt betrieben von den übrigen zentralen Komponenten der CCP-IT ( ). So kann jemand, der in der CCP-IT auf klinische oder Biomaterialdaten Zugriff hat, diese keinen realen Patienten zuordnen.
- Innerhalb des zentralen Identitätsmanagements wird der Kontrollnummernerzeuger (; zu schützen ist hier das verbundweite Geheimnis) getrennt betrieben von der Patientenliste ( ). So ist sichergestellt, dass die in der Patientenliste gespeicherten Kontrollnummern keinen direkten Rückschluss auf die Identität des Patienten zulassen.
- Selbiges gilt natürlich für lokale Geheimnisse.

Für die konkreten Betreiber der zentralen Komponenten vgl. Abschnitt 4.1.

5.2 Authentifizierung

Authentifizierung von Benutzern

Die Authentifizierung von teilnehmenden Forschern (im Folgenden auch „Benutzer“) gegenüber der CCP-IT erfolgt über Benutzername und Passwort gegenüber einem zentralen Authentifizierungsdienst, der vom DKFZ betrieben wird. Die Prüfung von Identität und Berechtigung von Benutzern erfolgt dabei auf Standort- oder Projektebene. Dazu wird pro DKTK-Standort bzw. -Projekt eine zuständige Person ernannt, welche die Anträge zum Zugriff auf die CCP-IT entgegennimmt, Identität und Berechtigung prüft und dann dem DKFZ die freizuschaltenden Personen mitteilt.

Im Fall von externen Forschern prüft der Ausschuss für Datenschutz Identität und Berechtigung und veranlasst die Freischaltung durch das DKFZ.

Die Freischaltung von Benutzern des Brückenkopfs erfolgt direkt durch den jeweiligen Standort. Dabei sind lokale Regelungen des Datenschutzes (zum Beispiel Sichtbarkeit bestimmter Patienten in bestimmten Abteilungen) zu berücksichtigen.

Authentifizierung von Komponenten

Zugriffe einer CCP-IT-Komponente auf eine andere über das Internet erfolgen nur nach erfolgreicher Authentifizierung, d.h. nicht nur die Berechtigung (Autorisierung), sondern auch die Identität der zugreifenden Komponente wird geprüft.

5.3 Maßnahmen in der IT-Infrastruktur

Sicherheit der gespeicherten Daten

Alle in den zentralen Komponenten der CCP-IT erhobenen Daten werden auf verschlüsselten Festplattenpartitionen gespeichert. Der zugehörige Schlüssel befindet sich jeweils auf einem getrennten Medium pro Server (z.B. Papier, USB-Stick). Dieses Medium wird nur während des Mount- bzw. Bootvorgangs benötigt und wird ansonsten sicher verwahrt. Nur der Administrator des jeweiligen Servers hat Zugriff auf ‚sein‘ Schlüsselmedium. Der Schlüssel kann nicht errechnet werden. Alle Server befinden sich in Rechenzentren, die über eine personengebundene Zugangskontrolle (zum Beispiel per Chipkarte) verfügen.

Sicherheit der Kommunikation

Die Vertraulichkeit der Kommunikation zwischen den Komponenten wird durch folgende Maßnahmen sichergestellt:

- Die Kommunikation zwischen den Komponenten erfolgt grundsätzlich über verschlüsselte Verbindungen (HTTPS). Die dafür eingesetzten Schlüssel und Zertifikate sind so zu erstellen, dass sie den aktuell anerkannten Anforderungen entsprechen (z.B. Schlüssellänge).
- Durch Firewalls ist sichergestellt, dass die Server, auf denen die zentralen Komponenten laufen, nur über diejenigen Protokolle und Ports erreichbar sind, die für die Kommunikation mit Benutzern oder anderen Komponenten erforderlich sind (in der Regel HTTPS-Verbindungen). Der administrative Zugang ist auf das Intranet des Betreibers beschränkt.
- Alle Kommunikationsvorgänge zwischen dem Brückenkopf und zentralen Komponenten werden vom Brückenkopf initiiert. Der Brückenkopf kann dadurch hinter einer Firewall oder einem Proxyserver betrieben werden, ohne über eine öffentliche Adresse aus dem Internet erreichbar zu sein.

Protokollierung

Es erfolgt eine Protokollierung der Zugriffe von Forschern auf die Komponenten sowie zwischen den Komponenten untereinander. Das Protokoll enthält mindestens:

- Die Identität der zugreifenden Person oder Komponente.
- Datum und Uhrzeit des Zugriffs.
- Den Inhalt des Zugriffs (die übermittelten Daten, ggfls. aggregiert) oder Informationen, aus denen dieser rekonstruiert werden kann (z.B. Verweis auf einen Datenbankeintrag o.ä.). Davon ausgenommen ist die Übertragung von IDAT an den Kontrollnummerngenerator.

Das Protokoll wird zusammen mit den Nutzdaten des entsprechenden Servers gespeichert und zwischen einem und sechs Monaten aufbewahrt. Die aufgezeichneten Daten werden nur für folgende Zwecke verarbeitet und eingesehen:

- Im Rahmen der technischen Administration (insbesondere zur Fehlersuche).
- Zur Aufdeckung möglicher Missbrauchsfälle.
- Zur Erstellung anonymisierter Nutzungsstatistiken.

6. Wahrung von Betroffenenrechten

6.1 Aufklärung und Einwilligung

Im Falle von *explizit eingewilligten Patienten* ist die informierte Einwilligung (Volltext siehe Anhang 1) Rechtsgrundlage der Datenverarbeitung. Mit der Einwilligung erklärt sich der Patient insbesondere dazu bereit, dass

- seine identifizierenden Daten an das zentrale Identitätsmanagement übermittelt werden,
- medizinische Daten gemäß MDS-K und MDS-B an die zentrale MDS-Datenbank übermittelt werden und
- diese Daten von Forschern des DKTK gemäß der Funktionsweise der zentralen Suche durchsucht werden können.

Sofern möglich, sollte eine solche Einwilligung von jedem *nicht explizit eingewilligten Patienten*, dessen Daten im Brückenkopf gespeichert werden, in der Folge eingeholt werden (in der Regel beim nächsten Patientenkontakt). Mit Einholen der Einwilligung wird der Patient außerdem über sein Recht auf Auskunft und Widerruf informiert.

6.2 Rechtsgrundlage bei *nicht explizit eingewilligten Patienten*

Daten von Patienten *ohne explizite Einwilligung* für die Verwendung der Daten und/oder Proben durch das DKTK liegen unter der Hoheit des behandelnden Standorts. In Bezug auf die Patientenrechte sind also die lokalen Regelungen des Standorts zu berücksichtigen.

Es ist zwischen a) der Speicherung im Brückenkopf und b) dem Upload von Daten in die zentrale MDS-Datenbank zu unterscheiden.

a) Die Erhebung und Speicherung der Daten von Patienten von *nicht explizit eingewilligten Patienten* erfolgt nur in der behandelnden Institution. Dennoch reicht der Behandlungsvertrag nicht als Rechtsgrundlage für diese Datenverarbeitung aus. Ähnlich wie zum Beispiel bei einem Clinical Data Warehouse verfolgt die Speicherung in diesem Fall nämlich nicht mehr den ursprünglichen Zweck der medizinischen Versorgung, sondern dient der medizinischen Forschung. Hier ist zunächst zu prüfen, ob eine lokale Einwilligung in die Verwendung und Weitergabe von klinischen Daten und/oder Biomaterialproben (i.d.R. nach Ethik-Votum für standortübergreifende Forschungsprojekte) als Rechtsgrundlage für diese Datenverarbeitung vorliegt. Falls dies nicht zutrifft, sind die jeweiligen landesrechtlichen Regelungen mit den entsprechenden Ausnahmetatbeständen zu prüfen. Sollten diese im Einzelfall eine Verwendung von Bestandsdaten aus dem Behandlungskontext für die medizinische Forschung zulassen, können diese auch ohne Einwilligung verwendet werden. Ebenso ist die Speicherung von Daten von *nicht explizit eingewilligten Patienten* im Brückenkopf unbedenklich, sofern sichergestellt ist, dass diese Daten auch nach der Speicherung im Brückenkopf nur von der behandelnden Einheit eingesehen werden können (vgl. Abschnitt 7, „Lokale Umgebung“).
   Falls es für einen der beteiligten Standorte keine spezialgesetzlichen Ermächtigungsregelungen für die Verwendbarkeit der Daten aus dem Behandlungskontext zu Forschungszwecken (z.B. Landeskrankenhausregelungen) gibt, so kann eine Erhebung der Daten von *nicht explizit eingewilligten Patienten* auf Basis der Forschungsklauseln des jeweiligen Landesdatenschutzrechts (für öffentliche Stellen der Länder) oder des Bundesdatenschutzgesetzes (für Stellen in privater Trägerschaft) möglich sein. Die Regelungen sehen eine Verarbeitung personenbezogener Daten auch ohne Einwilligung vor, wenn „dies zur Durchführung wissenschaftlicher Forschung erforderlich ist, das wissenschaftliche Interesse an der Durchführung des Forschungsvorhabens das Interesse des Betroffenen an dem Ausschluss der Erhebung, Verarbeitung und Nutzung erheblich überwiegt und der Zweck der Forschung auf andere Weise nicht oder nur mit unverhältnismäßigem Aufwand erreicht werden kann“[11]. Im Fall der *nicht explizit eingewilligten Patienten* ist aus folgenden Gründen davon auszugehen, dass diese Anforderungen erfüllt werden:
   - Die Wichtigkeit der translationalen Krebsforschung ist in der Fachwelt anerkannt. Ihre Förderung durch das BMBF und andere öffentliche Stellen belegt das große öffentliche Interesse an den Forschungszielen des DKTK. Im Sinne des Gesetzes ist das „wissenschaftliche Interesse an der Durchführung des Forschungsvorhabens“ hoch anzusetzen.
   - Die Daten dieser Patienten verlassen nicht die Institution, die sie ohnehin erhoben und gespeichert hat. Der Personenkreis, der diese Daten einsehen kann, ändert sich durch die Speicherung im Brückenkopf also nicht, lediglich die Zweckbindung an die Behandlung entfällt. Die Pseudonymisierung der Daten erschwert zudem die Reidentifizierung des Patienten aus den Daten im Brückenkopf. Das Interesse des Patienten daran, diese Datenverarbeitung zu verhindern, kann deshalb im Vergleich
   - zum wissenschaftlichen Interesse als gering angesehen werden.

[11] §13, Abs. 2, Bundesdatenschutzgesetz. Diese Stelle ist beispielhaft erwähnt, im konkreten Fall gelten ggfls. entsprechende Regelungen der Landesdatenschutzgesetze. Da an dieser Stelle nicht alle möglichen lokal gültigen Ausnahmetatbestände erfasst werden können, ist eine Prüfung vor Ort erforderlich.

- Wegen der immer stärkeren Zergliederung der Forschung im Bereich der Onkologie, mit zahlreichen molekular definierten Subgruppen, wird in Zukunft weder ein Standort alleine über die für Forschungsprojekte ausreichende Menge an Daten verfügen, noch wird die Anzahl der *explizit eingewilligten Patienten* in absehbarer Zeit groß genug für die wissenschaftlichen Ziele des DKTK sein. Die Bedingung, dass „der Zweck der Forschung auf andere Weise nicht oder nur mit unverhältnismäßigem Aufwand erreicht werden kann" ist damit ebenso als erfüllt anzusehen.
- Eine Beschränkung auf Daten *explizit eingewilligter Patienten* würde zu einem Bias in den Suchergebnissen der Plattform führen, da auch innerhalb jedes Standorts manche Kliniken eine Patienteneinwilligung erfolgreich ausrollen und andere nicht. Es ist also im Sinne einer repräsentativen Datenbasis notwendig, eine solche Selektion zu vermeiden.

b) Beim Upload von Daten in die zentrale MDS-Datenbank verlassen MDAT die behandelnde Institution. Daher muss, zusätzlich zur Rechtsgrundlage für die Speicherung von Daten im Brückenkopf (a), geprüft werden, ob der Upload des datensparsamen MDS (siehe Anhang 2) durch eine lokale Patienteneinwilligung oder durch landesrechtliche Regelungen datenschutzrechtlich legitimiert ist. Dies kann zum Beispiel dann gegeben sein, wenn die lokale Patienteneinwilligung eine pseudonymisierte Weitergabe der erhobenen Daten erlaubt, oder wenn der zu übermittelnde Datensatz als anonym betrachtet wird. In letzterem Fall sind die übermittelten Daten nicht mehr als personenbezogen anzusehen und unterliegen dadurch nicht den Datenschutzgesetzen (siehe Anhang 2).

Sofern möglich, sollte von *nicht explizit eingewilligten Patienten*, deren Daten im Brückenkopf gespeichert werden, in der Folge eine explizite Einwilligung eingeholt werden (in der Regel beim nächsten Patientenkontakt). Falls der Patient diese ablehnt, sind die bereits gespeicherten Daten zu löschen. Daten von Patienten, die der Verwendung ihrer Daten zu Forschungszwecken generell widersprochen haben, dürfen nicht in den Brückenkopf importiert werden.

6.3 Auskunft über gespeicherte Daten

Alle Patienten, deren Daten in den technischen Komponenten verwendet werden, haben das Recht, Auskunft über die in der CCP-IT über sie gespeicherten Daten zu erhalten. Der Antrag auf Auskunft ist schriftlich an die behandelnde Klinik zu stellen. Diese wendet sich zunächst unter Nennung der DKTK#ID (bei *explizit eingewilligten Patienten*) bzw. der S#ID (bei *nicht explizit eingewilligten Patienten*) an den Betreiber der zentralen Patientenliste. Dieser vergibt für den Auskunftsvorgang eine eindeutige, nicht-sprechende[12] Fallnummer und meldet diese an die anfragende Klinik zurück. Die Anfrage wird nun, unter Nennung der Fallnummer und der MDS-ID, aber ohne Nennung der DKTK#ID, an den Betreiber der zentralen MDS-Datenbank weitergeleitet. Dieser erstellt einen menschenlesbaren Ausdruck der Daten (mit Ausnahme der MDS-ID) und stellt diese unter Nennung der Fallnummer in einem versiegelten Umschlag der zuständigen Klinik zu. Diese kann nun anhand der Fallnummer den Patienten identifizieren und ihm den Ausdruck aushändigen.

Die in der zentralen Patientenliste gespeicherten Kontrollnummern liegen nicht in menschenlesbarer Form vor und können prinzipiell nicht in die ursprünglich eingegebenen Klartextdaten zurücktransformiert werden (vgl. Abschnitt 3.2). Eine Auskunft über diese Daten findet deshalb nur auf ausdrücklichen Wunsch des Patienten statt. In Antwort auf den allgemeinen Antrag auf Auskunft wird der Patient darüber informiert, dass eine Mitteilung der IDAT aufgrund der Chiffrierung impraktikabel ist und auf sein Recht, dennoch einen Ausdruck dieser Daten anzufordern, hingewiesen.

[12] D.h. für Dritte ist kein Rückschluss auf Daten oder Pseudonyme des Patienten möglich.

6.4 Widerruf, Löschung, Anonymisierung

Sämtliche Patienten, deren Daten in den technischen Komponenten verwendet werden, haben das Recht, die Einwilligung in die Verarbeitung ihrer Daten in der CCP-IT zu widerrufen. Der Widerruf ist schriftlich an einen der Standorte (in der Regel die behandelnde Klinik) zu richten. Der betroffene Patient kann mit dem Widerruf zusätzlich die vollständige Löschung seiner Daten beantragen. Fehlt dieser Antrag, so erfolgt eine Anonymisierung.

Nach Prüfung des Widerrufs wird der Antrag auf Löschung oder Anonymisierung zunächst an den Betreiber der Patientenliste weitergeleitet. Der Patient wird dabei über die $DKTK_{\#}ID$ (bei *explizit eingewilligten Patienten*) bzw. die $S_{\#}ID$ (bei *nicht explizit eingewilligten Patienten*) des betreffenden Standorts identifiziert. Zwecks Löschung oder Anonymisierung der Daten in der zentralen MDS-Datenbank wird deren Betreiber vom Betreiber der zentralen Patientenliste die MDS-ID des Patienten mitgeteilt. Im Falle einer Löschung werden alle dem Patienten zugeordneten Datensätze in Patientenliste und zentraler MDS-Datenbank gelöscht. Im Falle der Anonymisierung werden die Datensätze in der zentralen Patientenliste gelöscht und in den Datensätzen in der MDS-Datenbank die MDS-ID des Patienten durch ein zufälliges Pseudonym ersetzt. Durch den Algorithmus zur Pseudonymerzeugung ist sichergestellt, dass die Pseudonyme eines gelöschten oder anonymisierten Patienten am jeweiligen Standort nicht mehr für neue Patienten verwendet werden.

Die Löschung bzw. Anonymisierung ist von den zuständigen Betreibern zeitnah, maximal innerhalb von 14 Werktagen, vorzunehmen[13]. Der Betreiber der Patientenliste informiert wiederum alle Standorte über die Löschung oder Anonymisierung. In den Standorten werden eventuell lokal gespeicherte Daten des Patienten gelöscht, oder (bei Anonymisierung) die $DKTK_{\#}IDs$ durch Standortpseudonyme ersetzt[14]. Dieser Vorgang wird protokolliert und dem Betreiber der Patientenliste bestätigt. Der Abschluss der Löschung oder Anonymisierung wird von den Betreibern von MDS-Datenbank und Patientenliste dem Standort, an dem der Widerruf eingegangen ist, mitgeteilt und von diesem dem Patienten schriftlich bestätigt.

6.5 Dauer der Speicherung

Die erhobenen Daten bleiben in den zentralen Komponenten der CCP-IT gespeichert, so lange es für sie eine sinnvolle wissenschaftliche Verwendung im Rahmen der Patienteneinwilligung gibt. Falls die Daten nicht mehr in der vorgesehenen Form genutzt werden können (z.B. falls die CCP-IT außer Betrieb genommen wird), prüft der Ausschuss für Datenschutz, ob eine Rechtsgrundlage für eine anderweitige Verwendung der Daten, gegebenenfalls in anonymisierter Form, besteht. Falls diese Prüfung negativ ausfällt, sind die zentralen Daten zu löschen. Hinsichtlich der nur lokal in den Brückenköpfen gespeicherten Daten ist die Entscheidung über den weiteren Umgang mit den Daten von jedem Standort individuell zu treffen, da diese Daten möglicherweise weiter für Behandlungszwecke oder eigene Forschungsvorhaben genutzt werden dürfen.

[13] Die meist impraktikable Löschung oder Anonymisierung in Datensicherungen ist verzichtbar, sofern die Sicherungen nur durch den zuständigen Systemadministrator eingesehen werden können und alte Sicherungen regelmäßig gelöscht werden.
[14] Der Patient wird also im Sinne der CCP-IT von einem *explizit eingewilligten* zu einem *nicht explizit eingewilligten* Patienten.

7. Lokale Umgebung

Bei Installation und Betrieb der Brückenköpfe sind die lokalen Bestimmungen des Standorts hinsichtlich Datenschutz und -sicherheit zu beachten. Generell werden die für den Betrieb der zentralen Komponenten genannten Maßnahmen (siehe Abschnitt 5.3, „Maßnahmen in der IT-Infrastruktur“) empfohlen. Da die Komponenten des Brückenkopfes über das Intranet der jeweiligen Einrichtungen kommunizieren, kann hier auf die Verwendung verschlüsselter Verbindungen verzichtet werden.

Sofern der Brückenkopf nicht von der behandelnden Einheit selbst betrieben wird, ist durch Zugriffsrechte sicherzustellen, dass deren Datenhoheit gewahrt bleibt. Die Regelungen hinsichtlich administrativer Zugriffe (Abschnitt 4.3) gelten entsprechend.

7.1 Lokale Bestimmungen für den Standort [Standort]

[Sollte es abweichende Bestimmungen für die lokal an Ihrem Standort laufenden Komponenten oder Prozesse geben, können Sie diese hier einsetzen. Andernfalls löschen Sie den Abschnitt.]

Anhang

1. Patienteneinwilligung DKTK

Die angehängte Mustererklärung wurde vom CCP Office erstellt und dient als Vorlage, die von jedem Standort an lokale Erfordernisse angepasst wird. Diese tatsächlich eingesetzte Einwilligungserklärung erhalten Sie beim Standortvertreter in der AG CCP-IT.

Autor der hier eingebundenen Mustererklärung ist das CCP Office; zuständiger Ansprechpartner ist:

Barbara Uhl
Wissenschaftliche Projektkoordination

Deutsches Konsortium für Translationale Krebsforschung (DKTK)
CCP-Office

Universitätsklinikum Frankfurt
Med. Klinik II
Haus 33, Zi. 208
Theodor-Stern-Kai 7
60590 Frankfurt am Main

Tel.: +49 (0)69 / 6301 84237
Fax: +49 (0)69 / 6301 7463
E-Mail: b.uhl@dkfz.de

2. Anonyme Weitergabe festgelegter sparsamer Datensätze in die zentrale MDS-Datenbank zum Zweck der Machbarkeitsanalysen für Forschungsprojekte

Die angehängte Sachlage zur Anonymität der Datensätze in der zentralen MDS-Datenbank wurde federführend vom CCP Office erstellt; zuständiger Ansprechpartner ist:

Kristina Ihrig
Wissenschaftliche Projektkoordination

Deutsches Konsortium für Translationale Krebsforschung (DKTK)
CCP-Office

Universitätsklinikum Frankfurt
Med. Klinik II
Haus 33, Zi. 208
Theodor-Stern-Kai 7
60590 Frankfurt am Main

Tel.: +49 (0)69 / 6301 84237
Fax: +49 (0)69 / 6301 7463
E-Mail: k.ihrig@dkfz.de

3. Votum der AG Datenschutz der TMF

Die Arbeitsgruppe Datenschutz der Technologie- und Methodenplattform für die vernetzte medizinische Forschung (TMF e.V.) hat zu diesem Datenschutzkonzept in seiner ursprünglichen Fassung vom 16. Januar 2014 das angehängte Votum ausgesprochen.

4. Aktueller Meldedatensatz (MDS)

„Patienteninformation
zur Sammlung von Körpermaterialien und Freigabe von klinischen Daten und Körpermaterialien für die medizinische Forschung
Biobank und Forschung“

Inhalt

1 Was ist der Zweck dieser Einwilligung?

Sehr geehrte Patientin,
sehr geehrter Patient,

wir möchten Sie mit dieser Patienteninformation um eine Teilnahme an einem Vorhaben bitten, das mit unserem Forschungsauftrag als Universitätsklinikum zusammenhängt.

Wir bieten Ihnen modernste Diagnostik und umfassende Therapie auf höchstem nationalem und internationalem Niveau und versuchen dies durch intensive Forschung und Entwicklung ständig zu verbessern. Eine der notwendigen Voraussetzungen dafür ist die systematische Dokumentation von Patientendaten und eine Biobank.
Unter einer Biobank verstehen wir eine Sammlung von Proben menschlichen Gewebes und Flüssigkeiten (z.B. Proben von Organen und anderen festen Geweben, sowie Blut, Urin, Speichel), die mit personen- und krankheitsbezogenen Daten von Patienten (z.B. Alter, Geschlecht, Art der Erkrankung und Untersuchungsbefunde) in Zusammenhang stehen. Proben fallen als überzähliges, nicht weiter benötigtes Untersuchungsmaterial an. Anstatt das Material zu vernichten, kann es in einer Biobank aufbewahrt und dann für die medizinische Forschung genutzt werden. Es kann auch dann auch für Methoden genutzt werden, die erst in der Zukunft entwickelt werden.

2 Wie werden die Biomaterialien gewonnen?

Im Rahmen Ihrer Behandlung, werden zu diagnostischen und therapeutischen Zwecken Körpergewebe und -flüssigkeiten entnommen. Nach Abschluss der Untersuchungen bleiben nicht verwendetes Körpergewebe und -flüssigkeiten als sogenanntes Restgewebe übrig. Dies sind beispielsweise Biopsiematerial, operativ entnommenes Gewebe oder Flüssigkeiten, wie Blut, Urin, Rückenmarksflüssigkeit, Mund- oder Lungenspülflüssigkeit. Wir möchten Sie um Ihre Zustimmung dafür bitten, dass dieses Restgewebe in einer Biobank gesammelt wird.
Darüber hinaus möchten wir Sie bitten max. 25 ml Blut, das im Rahmen einer routinemäßigen Blutentnahme zusätzlich entnommen wird, der Biobank zu spenden. Das sind etwa 2 Esslöffel voll Blut. Bei dieser zusätzlichen Blutentnahme wird keine erneute Punktion der Vene durchgeführt. Sie können auch um die zusätzliche Spende von Urin, Speichel, Rückenmarksflüssigkeit oder Abstrichen gebeten werden. Ihr behandelnder Arzt wird Sie je nach Erkrankung bzw. Studie, an der Sie teilnehmen, dazu befragen.
Bei manchen Erkrankungen ist der Verlauf der Erkrankung von besonderer Bedeutung für die Forschung. Deshalb werden wir, wenn Sie zustimmen, in Einzelfällen in Abständen bei Routineblutuntersuchungen zusätzlich Blut für die Biobank entnehmen. Das ist für Sie mit keinen zusätzlichen apparativen (wie Röntgen-Aufnahmen) oder sonstigen klinischen Untersuchungen verbunden.

3 Wie werden die Daten erhoben?

Ihre personen- und krankheitsbezogenen Daten gehen direkt aus der Krankenakte bzw. dem Klinikinformationssystem hervor und werden in einer lokalen Datenbank des Krankenhauses dokumentiert und gespeichert.

Wie im vorherigen Abschnitt erwähnt, ist der Verlauf der Erkrankung oftmals von entscheidender Bedeutung. Daher werden wir Sie, wenn Sie zustimmen, in der Zukunft kontaktieren, um Ihren Gesundheitszustand zu erfragen und somit Ihre krankheitsbezogenen Daten zu ergänzen.

4 Was wird mit dem Biomaterial und den entsprechenden Patientendaten generell gemacht?

Die weitere Verwendung der Biomaterialien und der zugehörigen personen- und krankheitsbezogenen Daten findet ausschließlich im Rahmen von biomedizinischen Forschungsvorhaben statt. Diese Forschung schließt unter anderem die Untersuchung

- des Gewebes/des Blutes/des Urins selbst bzw. der darin enthaltenen Zellen, deren Erscheinungsbild, Beschaffenheit sowie biochemischer Eigenschaften
- von Zellbestandteilen (z.B. Fett- und Eiweißstoffe, Stoffwechselprodukte) einschließlich der Erbsubstanz (DNS)
- und eine vollständige Sequenzierung Ihrer Erbsubstanz mit ein.

Die gewonnenen Erkenntnisse sollen zukünftig eine Vorhersage z.B. der Wahrscheinlichkeit des Ansprechens auf eine spezifische Therapie ermöglichen und es ermöglichen Grundlagen für individualisierte – also genau auf den einzelnen Patienten ausgerichtete – Therapieentscheidungen zu entwickeln.

Dadurch, dass Ihre Biomaterialien auf unbegrenzte Zeit *(ggf. kann auch ein Zeitraum z.B. 30 Jahre angegeben werden)* gelagert werden, können sie auch zu einem späteren Zeitpunkt für Untersuchungen nach dem dann neuesten Stand der Wissenschaft herangezogen werden und Basis für die Entwicklung von neuen Therapien sein.

Ein weiterer wichtiger Aspekt der medizinischen Forschung ist die Durchführung von klinischen Studien. Anhand Ihrer personen- und krankheitsbezogenen Daten können Ärzte und Wissenschaftler ermitteln, ob es eventuell eine Studie gibt, welche sich mit Ihrer Erkrankung befasst.

Die Verwendung Ihrer personen- und krankheitsbezogenen Daten dient auch als Grundlage für epidemiologische Forschung, d.h. für die Auswertung von z.B. lokalen oder zeitlichen Auftretungswahrscheinlichkeiten von bestimmten Tumorerkrankungen in der Bevölkerung. Diese Informationen können als Grundlage zur Erforschung der Ursachen für bestimmte Erkrankungen dienen.

5 Ist die Vertraulichkeit gewährleistet?

Die elektronische Speicherung und Verarbeitung von Daten (personen- und krankheitsbezogenen Daten) erfüllt alle Anforderungen des Datenschutzes. Es wird sichergestellt, dass eine Rückverfolgung der Daten auf Ihre Person oder eine Verknüpfung mit Ihrer Krankengeschichte durch Dritte nicht möglich ist. Dazu werden Ihre identifizierenden Daten (z.B. Name, Geburtsdatum) mit Hilfe eines Pseudonyms in verschlüsselter, nicht rückführbarer Version gespeichert. Die gesammelten Proben und Daten werden diesem Pseudonym zugeordnet und somit ebenfalls verschlüsselt aufbewahrt. Dies bedeutet, dass Wissenschaftler oder Personen, die mit Ihren Proben und Daten zukünftig arbeiten, nicht wissen werden, von wem diese Proben oder Daten stammen. Um eine korrekte Dokumentation aller Ihrer wichtigen Daten zu gewährleisten, müssen speziell dafür beauftragte Fachleute Einblick in Ihre Krankheitsdaten nehmen. Die mit der Dokumentation beauftragten Personen sind zur strengen Vertraulichkeit und zur Beachtung des Datenschutzes verpflichtet. Indem Sie die beiliegende Einverständniserklärung unterschreiben, geben Sie Ihre Zustimmung zur oben beschriebenen Handhabung Ihrer Proben und Daten.

6 Vorteile?

Welcher Nutzen ergibt sich für Sie persönlich?
Für Sie ergeben sich durch die Sammlung und Lagerung bzw. Nutzung von Biomaterialien/Daten keine unmittelbaren Vorteile. Eine finanzielle Vergütung für entnommene Biomaterialproben kann und darf aus ethischen Gründen nicht erfolgen.

Welcher Nutzen ergibt sich für unsere Gesellschaft?
Alle derzeit durchgeführten, wie auch künftige medizinisch-wissenschaftliche Forschungsvorhaben zielen auf eine Verbesserung unseres Verständnisses der Krankheitsentstehung und der Diagnosestellung und auf dieser Basis auf die Neuentwicklung von verbesserten Behandlungsansätzen.

7 Nachteile oder Risiken?

Eine Teilnahme ist für Sie mit keinerlei Nachteilen oder Risiken verbunden. Ihre Behandlung wird durch die Teilnahme nicht beeinflusst.

Die genetischen Daten enthalten sehr viele Informationen einer Person. Es ist prinzipiell möglich, von spezifischen genetischen Veränderungsmustern im Erbgut auf den Einzelnen zurückzuschließen, auch wenn der genetische Datensatz von Ihren persönlichen Daten getrennt ist. Dies gelingt in der Regel aber nur, wenn von Ihnen Vergleichsproben vorliegen (Haare, DNA, Speichel) oder besondere Merkmale (Haarfarbe, Größe, Erkrankungen etc.) bekannt sind und mit Ihrer Person in Bezug gebracht werden können. Wir versichern Ihnen aber, dass wir alles tun werden, um Ihre Daten zu schützen und so die Identifizierung Ihrer Person zu verhindern. Der Gesetzgeber hat den Missbrauch Ihrer Daten unter Strafe gestellt.

Es wird ausdrücklich darauf hingewiesen, dass Ihre Teilnahme an der Biomaterialbank freiwillig ist.

8 Vorzeitige Beendigung/Widerruf?

Die einmal gegebene Einwilligungserklärung kann von Ihnen jederzeit ohne Angabe von Gründen bei Ihrem behandelnden Arzt am Klinikum bzw. der Biobank widerrufen werden. Es entstehen Ihnen damit keine Vor- oder Nachteile. Ab dem Zeitpunkt des Widerrufs werden die Proben anonymisiert, d.h. der Bezug zu Ihrer Person kann nicht mehr hergestellt werden. Falls Sie es wünschen, können die Proben auch vernichtet werden. Ihre klinischen Daten im Krankenhaus werden nicht gelöscht, weil diese Daten Bestandteil der Patientenakte sind und zu Ihrem Nutzen aus gesetzlichen Gründen aufbewahrt werden müssen.

9 Information über Ergebnisse der Forschung?

Wir werden Sie nicht über eine Nutzung Ihrer Proben und Daten und die damit erzielten Forschungsergebnisse unterrichten. Es werden in der Regel nicht einzelne, sondern eine große Vielzahl verschiedener Proben untersucht. Wir werden Sie auch nicht über Zufallsbefunde unterrichten, die von Forschern im Rahmen ihrer Projekte erhoben wurden. Zufallsbefunde können Erkenntnisse über weitere Erkrankungen oder Veranlagungen zu Erkrankungen sein.

10 Was mache ich, wenn ich noch weitere Fragen habe?

Für weitere Fragen bezüglich der Biobank wenden Sie sich bitte an XXX (Kontaktdaten Biobank) und für Fragen bezüglich der Dokumentation Ihrer klinischen Daten wenden sie sich bitte an XXX (Kontaktdaten Tumordokumentation). Sie können sich auch jederzeit an Ihren behandelnden Arzt wenden.

11 Weitergabe an Dritte

Bei manchen Forschungsaktivitäten ist es notwendig, mit anderen öffentlichen oder privaten Forschungseinrichtungen im In- und Ausland zu kooperieren. Auch Forschungskooperationen mit Partnern aus der Industrie sind möglich, um beispielsweise bessere Diagnostika und Therapeutika gezielt zu entwickeln. Auf Antrag und nach Bewilligung entsprechender Projekte durch die zuständige Ethikkommission erhalten diese Kooperationspartner dann Ihre pseudonymisierten Proben und Daten. Wir werden in diesen Fällen einen Vertrag (Materialübertragungsvereinbarung) mit den externen Forschern machen, der diese zur Einhaltung des Daten- und Persönlichkeitsschutzes verpflichtet. Ihre Einwilligung gilt auch für diese Forschungsprojekte. Forschungsergebnisse, die mit Ihren Proben und Daten erzielt wurden, dürfen anonymisiert veröffentlicht werden.

12 Deutsches Konsortium für Translationale Krebsforschung

Sollten Sie aufgrund einer bösartigen Tumorerkrankung am Uniklinikum xxx behandelt werden, möchten wie Sie mit dieser Patienteninformation auch um die Teilnahme an einem Vorhaben bitten, das mit unserem Forschungsauftrag als Mitglied des Deutschen Konsortiums für Translationale Krebsforschung (DKTK; http://www.dkfz.de/de/dktk/) zusammenhängt. In diesem Konsortium sind deutschlandweit 8 onkologische Spitzenzentren vertreten, die mit ihren regionalen Kooperationspartner zusammenarbeiten, um Fortschritte in der Vorbeugung, der Früherkennung, der Diagnostik und in der Behandlung von Krebserkrankungen zu erreichen.

Um wichtige wissenschaftliche Fragestellungen beantworten zu können, ist es von besonderer Bedeutung, Biomaterial und klinische Daten den Partner- und Kooperationsstandorten zur Verfügung zu stellen. Dazu werden Ihre pseudonymisierten personen- und krankheitsbezogenen Daten sowie Angaben zu Ihrem Biomaterial in einer gemeinsamen Datenbank aller DKTK-Partner für autorisierte Wissenschaftler zur Verfügung gestellt. Durch diese Bereitstellung der Daten wird es Ärzten und Wissenschaftlern des DKTK ermöglicht, im Rahmen einer Datenbanksuche, die Anzahl geeigneter Patienten oder vorhandenen Biomaterials für ein Forschungsvorhaben zu ermitteln.

Zusätzlich zur lokalen Pseudonymisierung Ihrer Daten, wird für dieses Projekt ein zentrales Pseudonym erstellt. Hierfür werden Ihre identifizierenden Daten (z.B. Name, Geburtsdatum) an eine zentrale Einheit des DKTK elektronisch übermittelt und dort in einen doppelt verschlüsselten Kode überführt. Ein optimaler Datenschutz ist dadurch gewährleistet, da Ihre identifizierenden Daten und die Pseudonyme nicht in der gleichen Datenbank vorliegen und somit eine Reidentifikation nur autorisiertem Personal für Forschungszwecke möglich ist.

Jegliche Verarbeitung und Nutzung von Daten und Biomaterial im Rahmen eines Forschungsprojektes, muss beim DKTK und Ihrer behandelnden Klinik (Name der Klinik….) formal beantragt werden. Dabei können weiterführende Projekte auch mit externen Partnern im In- und Ausland sowie privaten Unternehmen durchgeführt werden. Neben der Erteilung dieser Nutzungsgenehmigung durch das DKTK und Ihrer behandelnden Klinik, ist zusätzlich im Vorfeld ein zustimmendes Votum der für den Antragsteller zuständigen Ethikkommission einzuholen. Die Ethikkommissionen schützen im Rahmen von klinischen Forschungsvorhaben die Interessen der teilnehmenden Patienten und vertreten deren Rechte gegenüber den forschenden Ärzten und Wissenschaftlern im DKTK.

Einwilligungserklärung

zur Teilnahme an der Biobank

- Hiermit erkläre ich (Patientenname) ……., dass ich durch Herrn/Frau Dr. (Name Arzt) ..... verständlich über die Biobank und die Dokumentation meiner Daten, sowie deren Tragweite aufgeklärt worden bin.
  Ich habe darüber hinaus den Text der Patienteninformation erhalten und sowohl diesen als auch die Einverständniserklärung gelesen und verstanden. Aufgetretene Fragen wurden mir vom aufklärenden Arzt verständlich und ausreichend beantwortet.

  Diese Einwilligung ist freiwillig und ich habe jederzeit das Recht, die Einwilligung zur Verarbeitung und Nutzung meines Materials auch ohne Angabe von Gründen zu widerrufen. Meine Proben werden daraufhin anonymisiert oder auf Wunsch vernichtet. Bis zum Widerruf bereits ausgewertete Daten werden in anonymisierter Form (d.h. ohne Bezug zu meiner Person) archiviert, entstandene Forschungsergebnisse bleiben somit erhalten. Mir entstehen keine Vor- oder Nachteile wenn ich widerrufe, insbesondere nicht für meine Behandlung.

- Ich stimme der Sammlung, Lagerung und Nutzung von meinen Körpermaterialen, die im Rahmen meiner medizinischen Behandlung entnommen wurden und nicht weiter benötigt werden, durch die Biobank zu. Ich bin mit der Erhebung, -verarbeitung und -nutzung meiner Daten einverstanden. Meine Proben und Daten werden auf unbefristete Zeit *(ggf. angegebener Zeitraum siehe S. 3)* bzw. bis zu meinem Widerruf genutzt. Ich bin mir darüber im Klaren, dass meine Materialien und Daten damit sehr lange verwendet werden dürfen. Meine Einwilligungserklärung gilt über meinen Tod hinaus.

- Ich spende und übereigne der Biobank Biomaterialien zu verschiedenen Zeitpunkten des Krankheitsverlaufs wie in der Patienteninformation beschrieben. Mein behandelnder Arzt wird dies jeweils mit mir absprechen. Die Abstände der Blut- und Flüssigkeitsspenden können unterschiedlich (Wochen, Monate oder Jahre) sein und erfolgen im Rahmen von routinemäßigen Blut- und Flüssigkeitsentnahmen.

- Um eine korrekte Dokumentation der Daten zu gewährleisten, dürfen zur Verschwiegenheit verpflichtete Personen Einblick in meine personenbezogenen Krankheitsdaten nehmen, soweit diese in der Klinik vorliegen und soweit dies erforderlich ist. Ich entbinde die behandelnden Ärzte insoweit von der ärztlichen Schweigepflicht.

- Ich stimme der Nutzung von meinen Proben und Daten, die an meinem behandelnden Zentrum genommen und erhoben wurden, jeweils in kodierter, d.h. pseudonymisierter Form für verschiedene biomedizinische Forschungszwecke einschließlich der genetischen Forschung zu. Ich erkläre mich auch damit einverstanden, dass meine Proben pseudonymisiert an externe Forschergruppen im In- und Ausland, sowie private Unternehmen weitergegeben werden. Ich stimme auch Forschungsprojekten zu, deren Methodik erst in der Zukunft entwickelt wird. Ich stimme zu, dass die aus meinen Biomaterialien und Daten gewonnenen Ergebnisse in

pseudonymisierter Form gespeichert und weiter genutzt werden können und dass Forschungsergebnisse, die mit meinen Proben und Daten erzielt werden, anonymisiert veröffentlicht werden dürfen.

- Ich willige in die Verwendung meiner identifizierenden Daten für die Erzeugung eines Pseudonyms im Rahmen der DKTK-Forschung ein. Für die Verwendung meiner pseudonymisierten personen- und krankheitsbezogenen Daten sowie der Angaben über mein Biomaterial im Rahmen der zentralen Suche des DKTK willige ich ein. Ich bin mir darüber klar, dass meine Daten hierfür in einer zentralen Datenbank gespeichert werden.

- Ich bin damit einverstanden, dass ich **nicht** über Forschungsergebnisse und Zufallsbefunde informiert werde auch wenn diese mit meiner Erkrankung in Zusammenhang stehen sollten.
  Hiervon unberührt bleibt jedoch das mir zustehende Auskunftsrecht über meine gespeicherten Daten nach Bundesdatenschutzgesetz.

- Ich willige ein, dass meine Kontaktdaten in der Klinik gespeichert werden. Mit einer Kontaktaufnahme per Brief zu einem späteren Zeitpunkt für weitere Forschungsvorhaben und/oder zur Erhebung weiterer Daten und Proben erkläre ich mich einverstanden.

- Im Falle einer Auflösung der Biobank erkläre ich mich damit einverstanden, dass meine Proben und ausgewählte klinische Daten anonymisiert in eine andere Biobank übertragen werden können.

Ich stimme den oben genannten Punkten zu.
Eine unterschriebene Einwilligungserklärung wurde mir in Kopie ausgehändigt.

……	…….	……
(Ort/Datum)	*(Unterschrift Patient)*	*(Unterschrift Arzt)*

Quellen:

- Patienteninformation des Universitätsklinikums Dresden
- Patienteninformation zur Biobank und EW des Universitätsklinikums Essen; Version 1.5, 10/2012
- Patienteninformation und Einverständniserklärung der UCT Biomaterialbank Frankfurt; Version allgemein 1.2, 20.01.2009
- Patienteninformation der Gewebebank des NCT Heidelberg; Stand 01.03, Juli 2008
- Mustertext zur Patienteninformation des m4 München; 13.08.2012
- Patienteninformation für Gewebe-/Blutproben des Universitätsklinikums Tübingen
- Mustertext zur Spende, Einlagerung und Nutzung von Biomaterialien sowie zur Erhebung, Verarbeitung und Nutzung von Daten in Biobanken AKEK; 09.11.2013
- Patienteninformation aus der Stellungnahme „Eckpunkte für eine Heidelberger Praxis der Ganzgenomsequenzierung“ Heidelberg; Juni 2013
- Patienteninformation zur INFORM-Register Studie (älter 18 Jahre); Version 06.09.2013
- Checkliste und Leitfaden zur Patienteneinwilligung TMF; 2006

DKTK: Anonyme Weitergabe festgelegter sparsamer Datensätze in die zentrale MDS-Datenbank zum Zweck der Machbarkeitsanalysen für Forschungsprojekte

Hintergrund:

Die folgenden Erläuterungen dienen der Einschätzung, ob lokal vorliegende Patientendaten (festgelegter Meldedatensatz (MDS[1])) in die zentrale MDS-Datenbank (DB) auf Basis **der Datenschutzgesetze** weitergegeben werden dürfen.

Der Zweck der zentralen MDS-DB ist die Machbarkeitsprüfung von Forschungsprojekten mit Hilfe der Anzeige einer Fallzahl von Suchergebnissen.

Das Bundesdatenschutzgesetz (BDSG) und die LDSG bzw. BlnDSG greifen nicht, wenn Daten anonymisiert worden sind[2]. Gemäß BDSG[3]/anderer DSG[4] besteht Anonymität, wenn „Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können".

Für den am teilnehmenden Standort stattfindenden datenverarbeitenden Schritt der Pseudonymisierung oder Anonymisierung liegt eine Einwilligung des Patienten vor.

Daraus resultiert, dass die Weitergabe von Daten in die zentrale MDS-DB für die zentrale Suche zum Zweck der Machbarkeitsprüfung von Forschungsprojekten zulässig ist, da die Anonymität gewährleistet ist.

Informationen zur Anonymität der zentralen MDS-DB:

- Die Inhalte eines Datensatzes sind reduziert (datensparsam).
- Ein Datensatz enthält keine personenbezogenen Daten (z.B. Klarnamen).
- Die Daten zu Biomaterial eines Patienten beziehen sich auf verwaltende Angaben (z.B. Tumorgewebe, Normalgewebe).
- Die klinischen Daten eines Patienten sind eine Teilgruppe des ADT[5]-Basisdatensatzes, der im Rahmen des KFRG[6] an andere Datenbanken weitergegeben werden darf.
- Die zentrale MDS-DB enthält ein Pseudonym mit einer Standortinformation als ein doppelt verschlüsseltes Pseudonym (MDS-Pseudonym), mit dem Ziel, wiederholte Datenimporte von den Standorten inkrementell durchzuführen. Dadurch werden höchste technische Anforderungen beachtet und faktisch Anonymität erreicht (siehe Datenschutzkonzept, Abschnitt 3.4 und 3.1, 10.10.2014). Dies ist durch die TMF e.V. bereits positiv begutachtet worden.

Zuständige Ansprechperson im DKTK:

Kristina Ihrig, Dipl.-Biol. (CCP Office)
E-Mail: k.ihrig@dkfz.de
Tel.: 069-6301-84237

[1] Meldedatensätze
[2] Metschke, Wellenbrock (2002): Datenschutz in Wissenschaft und Forschung
[3] BDSG, §3, Abs. 6
[4] Bayern, Berlin, Baden-Württemberg, Hessen, Rheinland-Pfalz; SächsDSG modifiziert
[5] Arbeitsgemeinschaft Deutscher Tumorzentren e.V. und GEKID
[6] Krebsfrüherkennungs- und Registergesetz

TMF – Technologie- und Methodenplattform
für die vernetzte medizinische Forschung e.V.

TMF e. V. · Charlottenstraße 42/Dorotheenstraße · 10117 Berlin

Berlin, 16. Januar 2014

Stellungnahme zum Datenschutzkonzept für die Clinical Communication Platform (CCP-IT) im Deutschen Konsortium für Translationale Krebsforschung (DKTK)

Das Datenschutzkonzept für die CCP-IT wurde von der TMF-Arbeitsgruppe „Datenschutz", zuletzt auf der Sitzung am 13. November 2013, beraten. Inzwischen liegt der AG das Konzept in der aufgrund der Beratungen überarbeiteten Version vom 8. Januar 2014 vor.

Das vorliegende Konzept beschreibt ein Portal, das Anfragen nach Fallzahlen und Machbarkeit von Forschungsprojekten ermöglichen soll. Zu diesem Zweck sammelt es medizinische Daten von Krebspatienten der teilnehmenden Kliniken pseudonymisiert in einer zentralen Datenbank (einschließlich Hinweisen auf vorhandene Biomaterialproben), beruhend auf einer entsprechenden Einwilligung der betroffenen Patienten („DKTK-Patienten"). Da die Datenbank ein ansonsten nur in der Patientenliste bekanntes Pseudonym verwendet (als MDS-ID bezeichnet), erfüllt das Konzept die Anforderungen des Modells A des generischen Datenschutzkonzepts der TMF. Anstatt der TempID des TMF-Konzepts wird eine äquivalente Lösung verwendet, bei der das Pseudonym asymmetrisch verschlüsselt an die zentrale Datenbank durchgereicht wird. Insofern, als überhaupt kein Online-Zugriff außer statistischen Abfragen auf die zentrale Datenbank vorgesehen ist, ist das Konzept sogar deutlich strenger als das generische TMF-Konzept. Die konkrete Ansiedlung der Datenbanken zur Erreichung einer wirksamen informationellen Gewaltenteilung ist noch nicht definiert. Die Besetzung des Ausschusses Datenschutz ist bisher nur in Form der beteiligten Rollen beschrieben.

Parallel dazu sollen aber auch die Daten von Patienten nutzbar gemacht werden, für die keine Einwilligung vorliegt. Hierbei handelt es sich um „Altfälle". Zunächst sind dies sehr viele, die Anzahl wird aber im Laufe der Zeit abnehmen, da neue Fälle grundsätzlich nur mit Einwilligung aufgenommen werden, und bei Altfällen die Einwilligung nach Möglichkeit (bei Weiterbehandlung) nachgeholt werden soll. Die Daten dieser Patienten, als „Nicht-DKTK-Patienten" bezeichnet, werden nur in den behandelnden Einrichtungen in einem sogenannten Brückenkopf gespeichert. Dieser ist nach dem Modell eines klinischen Data-Warehouse konzipiert und muss natürlich den entsprechenden lokal geltenden Datenschutzregeln genügen. An eine zentrale Patientenliste wird nur ein Satz von „Kontrollnummern" weitergegeben, der ein lokal erzeugtes Pseudonym darstellt. Diese Kontrollnummern erlauben keinen einrichtungsübergreifenden Abgleich. Aufgrund der lokalen Erzeugung sind sie in der Patientenliste auch vor Angriffen durch Probeverschlüsselung sicher. Daher ist es gerechtfertigt, die Kontrollnummern als faktisch anonym anzusehen.

Darüber hinaus erlaubt das Portal keine direkten Abfragen auf die zugehörigen Datenbestände von Nicht-DKTK-Patienten, sondern dient nur zur Vermittlung von externen Abfragen nach Fallzahlen, die vor Ort außerhalb dieses Systems im Rahmen der rechtlich erlaubten Möglichkeiten bearbeitet werden. Die hierfür anzuwendenden Anonymitätskriterien sowie Anonymisierungsmaßnahmen für

TMF e. V.
Charlottenstraße 42/Dorotheenstraße
D-10117 Berlin
Tel.: +49 (30) 22 00 24 70
Fax: +49 (30) 22 00 24 799
E-mail: info@tmf-ev.de
Internet: www.tmf-ev.de

Commerzbank AG
BLZ: 100 400 00
Kto.: 737 811 000
IBAN: DE60 1004 0000 0737 8110 00
BIC: C0BADEFFXXX
St-Nr.:27/640/56509
USt-IdNr.: DE244871253

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

eventuell folgende Datenexporte betreffen nicht das Portal und sind daher nicht Gegenstand des vorliegenden Datenschutzkonzepts.

Verantwortliche Stelle für die Datenverarbeitung ist das Deutsche Krebsforschungzentrum (DKFZ) in Heidelberg.

Die AG Datenschutz sieht in dem Teil des Konzepts, der DKTK-Patienten betrifft, eine Umsetzung einer strengeren Variante des TMF-Datenschutzkonzepts, Modell A. Der Teil, der Nicht-DKTK-Patienten betrifft, sieht nur die Weitergabe eines faktisch anonymen Satzes von Kontrollnummern vor. Daher befürwortet die AG das vorgelegte Konzept. Die zugehörige Patientenaufklärung und Einwilligungserklärung sowie sonstige rechtlich bindende Abmachungen zwischen den Projektbeteiligten liegen bisher nicht vor und sind daher nicht Gegenstand dieser Stellungnahme.

Prof. Dr. Klaus Pommerening
Sprecher der AG Datenschutz

Meldedatensätze - MDS-K und MDS-B

Der **MDS-K** (klinische Daten) enthält Attribute, wodurch klinische Daten zum Patienten, zum Primärtumor, zur Primärtherapie, zum Ansprechen und zum Vitalstatus für die zentrale Suche der zentralen MDS-Datenbank zur Verfügung stehen.

Der **MDS-B** (Biomaterialdaten) enthält 5 Attribute, welche eine eindeutige Identifizierung des Biomaterials erlauben.

Nr.	Merkmal
Allgemeine Daten	
A-0	Standortpseudonym
A-1	Geburtsmonat/-jahr (nicht durchsuchbar; Berechnung des Alters)
A-2	Geschlecht
Klassifikation von Primärtumoren	
K-1	Datum der TNM-Dokumentation/Datum Befund (nicht durchsuchbar)
K-2	Diagnosejahr/-datum (Jahr suchbar; Diganosedatum nicht durchsuchbar, Berechnung des Alters, Qualitätssicherung)
K-3	Alter bei Erstdiagnose
K-4	Diagnose
K-5	ICD-Katalog (Version)
K-6	Lokalisation
K-7	ICD-O Katalog Topographie (Version)
K-8	Seitenlokalisation
K-9	Morphologie
K-10	ICD-O Katalog Morphologie (Version)
K-11	Grading
K-12	UICC Stadium
K-13	TNM-T
K-14	TNM-m-Symbol
K-15	TNM-N
K-16	TNM-M
K-17	c/p/u-Präfix T
K-18	c/p/u-Präfix N
K-19	c/p/u-Präfix M
K-20	TNM-y-Symbol
K-21	TNM-r-Symbol
K-22	TNM-Version
K-23	Lokale Beurteilung Resttumor
K-24	Gesamtbeurteilung Resttumor
K-25	Fernmetastasen [ja/nein]
K-26	Datum diagnostische Sicherung (nicht durchsuchbar)
K-27	Lokalisation Fernmetastasen

Nr.	Merkmal
Generierter Suchkatalog	
K-31	Klinische-relevante Tumorentitäten: DKTK-Katalog: Übersetzung Tumorcodes für Diagnose & Morphologie in klinisch-relevante Gruppen]
Therapie des Primärtumors	
K-32	OP [ja/nein]
K-33	Intention OP
K-34	Strahlentherapie [ja/nein]
K-35	Intention Strahlentherapie
K-36	Stellung zur Operation
K-37	Chemotherapie [ja/nein]
K-38	Intention Chemotherapie
K-39	Stellung zur Operation
K-40	Immuntherapie [ja/nein]
K-41	Hormontherapie [ja/nein]
K-42	Knochenmarktransplantationen [ja/nein]
K-43	Weitere Therapie(n) [ja/nein]
Ansprechen auf Primärtherapie	
K-45	Ansprechen Primärtherapie
K-46	Datum des ersten Verlaufs (nicht durchsuchbar)
K-47	Lokales/regionäres Rezidiv
K-48	Datum (lokales/regionäres) Rezidiv (entspricht Verlaufsdatum, nicht durchsuchbar)
K-49	Lymphknoten-Rezidiv
K-50	Datum Lymphknoten-Rezidiv (entspricht Verlaufsdatum, nicht durchsuchbar)
K-51	Fernmetastasen
K-52	Datum Fernmetastasen (entspricht Verlaufsdatum, nicht durchsuchbar)
Aktueller Tumorstatus	
K-53	Ansprechen innerhalb der letzten 3 Monate
K-54	Monat.Jahr des letztbekannten Verlaufs (nicht durchsuchbar)
Vitalstatus	
K-55	Monat.Jahr des letztbekannten Vitalstatus (nicht durchsuchbar)
K-56	Vitalstatus [lebend/verstorben]
Biomaterial	
B-1	Patienten mit Biomaterial
B-2	Probentyp (z.B. Gewebeprobe)
B-3	Probenart (z.B. Tumorgewebe, Normalgewebe)
B-4	Entnahmedatum (nicht durchsuchbar)
B-5	Fixierungsart (z.B. Paraffin, Kryo/Frisch)