Das Datenschutzkonzept der TMF für Biomaterialbanken

The TMF Data Protection Scheme for Biobanks

Klaus Pommerening, Johannes-Gutenberg-Universität, Mainz

Zusammenfassung Das Datenschutzkonzept der TMF für Biomaterialbanken beschreibt organisatorische und technische Maßnahmen, die einen rechtlich einwandfreien Umgang mit Biomaterialien und den Aufbau von Biomaterialbanken ermöglichen und die im Sinne der medizinischen Forschung notwendige Abschwächungen der Einwilligung erlauben. ►►► **Summary** The TMF data protection scheme defines organisatorial and technical measures for building biobanks in a legally and ethically sound way. It allows weaker forms of the informed consent that leave open future extensions of the original research purposes.

KEYWORDS K.4.1 [Computing Milieux: Computers and Society: Public Policy Issues] Ethics, Privacy; J.3 [Computer Applications: Life and Medical Sciences]; Biomaterialbanken, Datenschutz, Informationstrennung, Pseudonymisierung / Biobanks, data protection, separation of information, pseudonymisation

1 Einleitung

Biomaterialbanken sind eine wesentliche Grundlage gegenwärtiger und künftiger medizinischer Forschung. Nutzen bringt Biomaterial für die meisten Fragestellungen nur, wenn auch zugehörige qualitativ hochwertige Diagnose- und Behandlungsdaten zur Verfügung stehen; für Projekte der genetischen Epidemiologie sind darüber hinaus auch soziodemographische Daten unverzichtbar.

Biomaterialbanken bringen eine erhöhte Gefährdung der Persönlichkeitsrechte mit sich, da die Proben und Probenextrakte, wie klein sie auch sein mögen, praktisch stets die vollständige genetische Information über ihren Spender enthalten. Diese Informationsdichte ist nicht teilbar und – im Sinne der Datensparsamkeit – nicht beschränkbar. Darüber hinaus ist durch die stets aus Körpermaterial ableitbaren „genetischen Fingerabdrücke“ über eine Vergleichsprobe der Spender eindeutig zu identifizieren. Sind die Proben mit weiteren Daten, z. B. Gesundheitsdaten, der betroffenen Person verknüpft, lassen sich über eine solche Zuordnung die Kenntnisse über diese Person beträchtlich erweitern.

Für die Nutzung einer Biomaterialbank ist oft entscheidend, dass Aufbewahrungsdauer, Zweckbestimmung und möglicher Nutzerkreis möglichst offen bleiben. Das schließt eine Regelung aus, die allein auf der Einwilligung der Betroffenen beruht; diese Problematik wird im Beitrag von P. Ihle [1] vertieft. Die dort vorgeschlagene Erweiterung der Einwilligungserklärung ist rechtlich und ethisch nur vertretbar, wenn die gesamte Biomaterialbank in eine strikte, transparente Organisation eingebunden und durch starke technische Maßnahmen abgesichert ist.

Das Datenschutzkonzept für Biomaterialbanken [2] der TMF (*Telematikplattform für Medizinische Forschungsnetze*) [3] beschreibt solche Maßnahmen. Es wurde in Anlehnung an das generische Datenschutzkonzept für medizinische Forschungsnetze [4] entwickelt und in den wesentlichen datenschutzrelevanten Fragen mit den Datenschutzbeauftragten des Bundes und der Länder intensiv diskutiert und abgestimmt. Folgende Vorarbeiten wurden u. a. wesentlich berücksichtigt:

- Die Stellungnahme des nationalen Ethikrates „Biobanken für die Forschung“ [5].
- Das Konsenspapier „Epidemiologie und Datenschutz“ der Datenschutzbeauftragten mit der DAE [6].
- Aufsätze von R. Wellbrock über Biobanken für Forschungszwecke [7; 8].
- Bereits existierende Konzepte für Biomaterialbanken, vor al-

lem: GEPARD im Kompetenznetz Parkinson, GENOmatch, Brain-Net.

2 Abzuwehrende Risiken

Gesundheitsdaten sind sehr sensible Informationen, deren bekannt werden dem Betroffenen großen Schaden, z. B. durch Stigmatisierung, zufügen kann. Daher sind alle in einer Datenbank enthaltenen medizinischen Daten grundsätzlich bestmöglich davor zu schützen, dass sie einer bestimmten Person zugeordnet werden können.

Bei den Risiken ist zu unterscheiden zwischen Angriffen auf die Proben selbst und Angriffen auf die Analysedaten und andere Daten, die zum Gesamtkomplex der Biomaterialbank gehören. Das Risiko für die Proben – aus dem Gesichtspunkt des Probanden – besteht darin, dass sie entwendet und unbefugt analysiert oder für Experimente genutzt werden; dieses Risiko wird durch physische Schutzmaßnahmen eingedämmt. Dazu kommt natürlich das Risiko der Vernichtung, Beschädigung oder anderer Sabotageversuche, die nicht im Fokus des Datenschutzkonzepts stehen.

Abzuwehrende Risiken für die Daten sind:

1. Unbefugter Zugriff auf Informationen, die bei den beteiligten Stellen der Biomaterialbank gespeichert und verarbeitet werden.
2. Unbefugte Rückidentifizierung eines Probanden unter Verwendung von berechtigt oder unberechtigt erlangten Informationen, insbesondere Abgleich mit anderen Datenbeständen.

Hauptziel des Datenschutzkonzepts ist die Minimierung und Kontrolle des Rückidentifizierungsrisikos. Das schließt auch geeignete Zugriffsregelungen ein.

3 Methoden

Im Zentrum der organisatorischen Maßnahmen des Datenschutzkonzepts steht ein Regelwerk für die Biomaterialbank, das den Umgang mit den Daten und Proben detailliert regelt. Zu diesem Regelwerk gehören:

- Die Satzung der Trägerorganisation oder vergleichbare verbindliche Dokumente.
- Aufteilung der Verantwortlichkeit für verschiedene Prozesse, Einrichtung entsprechender Aufgabenbereiche und Gremien.
- Die Musterformulierung der Einwilligungserklärung.
- Ein formalisiertes Verfahren zur Prüfung und Genehmigung von Forschungsprojekten.
- Sicherheitsrichtlinien („Policies“) und Richtlinien für Verfahrensabläufe (SOPs = „Standard Operating Procedures“).
- Verpflichtungserklärungen für das eigene Personal.
- Vertragliche Regelungen mit externen Kooperationspartnern.

Die organisatorischen Regelungen werden durch informatische Werkzeuge unterstützt:

- Eine Netzarchitektur mit einer Zerlegung in unabhängige Komponenten, die die verteilte Verantwortung und die Informationsteilung widerspiegelt und absichert.
- Netzdienste, die über Trusted Third Parties (TTP) abgewickelt werden.
- Kommunikationskontrolle und -absicherung, insbesondere auch durch kryptographische Verfahren.
- Zugangs- und Zugriffskontrolle.
- Sicherheitsmaßnahmen nach dem Stand der Technik für Datenbankserver und Netzdienste.
- Anonymisierung und Pseudonymisierung von Proben und Daten.

3.1 Getrennte Speicherung von Daten

Die Trennung von Identitätsdaten und medizinischen Daten in der Forschung ist ein wichtiger Grundsatz des Datenschutzkonzepts, denn die getrennte Datenspeicherung erschwert die Zuordnung von Personen und Daten sowohl intern als auch extern. Ferner empfiehlt es sich, auch die Zuständigkeit für die Probenbank und die zentrale Datenbank, die die medizinischen Daten enthält, zu trennen. Daher sollte die Verwaltung von medizinischen Daten, Proben und identifizierenden Daten in jedem Fall unter getrennter Verantwortung erfolgen. Es ist zu erwägen, inwieweit medizinische oder soziodemographische Daten mit erhöhtem Identifizierungspotenzial sogar noch unter getrennter Verantwortung gespeichert werden müssen. Ein Beispiel hierfür können feingliedrige epidemiologische Daten sein. Auch Bilddaten können mitunter eine Rückidentifizierung erleichtern.

Die aus den Proben gewonnenen molekularbiologischen Analysedaten bergen im Vergleich zu den Proben sogar ein höheres Rückidentifizierungsrisiko, da sie leichter kommunizierbar sind. Probenanalysedaten sollten daher grundsätzlich separat gespeichert werden.

Aufgrund dieser Überlegungen sind im Grundmodell zumindest die vier folgenden Datenarten unter getrennter Verantwortung zu speichern:

- IDAT (Identitätsdaten): in der Patientenliste
- MDAT (medizinische Daten): in der Forschungsdatenbank
- Probe mit zugehörigen OrgDAT (organisatorischen Daten): in der Probenbank
- ProbDAT (Probenanalysedaten): in der Analysendatenbank.

Auch die weiteren Datenarten SozDAT (soziodemographische Daten), BildDAT (Bilder, wie z. B. Röntgenbilder) und andere sind, wenn die Analyse des Rückidentifizierungspotenzials dies erfordert, jeweils getrennt zu speichern.

3.2 Anonymisierung und Pseudonymisierung

Um den Personenbezug der Daten zu beseitigen, sind Daten und Pro-

ben zu anonymisieren, sobald dies nach dem Forschungszweck möglich ist. Es besteht vielfach aber ein hohes Interesse der Forschung, manchmal auch des Probanden selbst, unter bestimmten Bedingungen den Personenbezug wieder herstellen zu können. Die Daten und Proben können dann pseudonymisiert verwendet werden, vorausgesetzt, das Regelwerk der Biomaterialbank und die Einwilligungserklärung sind entsprechend gestaltet.

Eine Rückidentifizierung (Depseudonymisierung) darf nur gemäß den in Regelwerk und Einwilligungserklärung festgelegten Bedingungen und Verfahren stattfinden. Gründe sind beispielsweise Maßnahmen der Qualitätssicherung oder die Kontaktaufnahme mit dem Patienten im Fall einer Information oder zur Erhebung weiterer Daten.

Das Rückidentifizierungsrisiko für Probanden soll so weit minimiert werden, dass eine unbefugte Rückidentifizierung – wie bei einer faktischen Anonymisierung – nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft möglich ist. Zu einer wirksamen Anonymisierung oder Pseudonymisierung gehört auch die klassische Methode, Einzelfälle, die anhand extremer Einzelmerkmale oder Merkmalskombinationen leicht zu identifizieren sind, vor Weitergabe durch Weglassen, Runden, Verändern oder Aggregieren von Merkmalen unkenntlich zu machen [9].

3.3 Technische Sicherheitsmaßnahmen

Auf technischer Ebene sind kryptographische Verfahren für die Sicherung der Datenübermittlung und die Dokumentensicherheit vorzusehen. Für die Datenbanken sind Datenbankserver nach Industrie-Standard mit ihren originären Sicherheitstechnologien zu verwenden, um unberechtigte Zugriffe zu verhindern. Alle Datenbanken müssen in einem durch Firewall-Technik geschützten lokalen Netz auf gehärteten Servern [10] liegen. Es sind rollenbasierte Zugriffsrechte vorzusehen. Unerlaubte Zugriffsversuche auf Daten sind aus den Log-Protokollen zu ermitteln.

4 Organisation der Biomaterialbank

4.1 Organisationsformen

Eine Biomaterialbank (BMB) kann auf verschiedene Weise organisiert sein,

- integriert in eine Klinik oder medizinische Einrichtung,
- als eigenständige BMB,
- als Komponente in einem medizinischen Forschungsnetz.

Für Daten- und Materialflüsse sowie Verarbeitungs- und Analyseprozesse gibt es zusätzlich unterschiedliche Formen:

- Zentrale BMB mit zentraler Verarbeitung, bei der nur Analyseergebnisse weitergegeben werden
- Zentrale BMB mit dezentraler Verarbeitung, bei der Proben weitergegeben werden
- Kooperative BMB in Zusammenarbeit unabhängiger Einrichtungen mit einer zentralen Datenbank
- Vermittelnde BMB, bei der Daten, Proben und Ergebnisse nur in einer zentralen Verweisdatei registriert und an Forschungsprojekte vermittelt werden.

Diese unterschiedlichen Organisationsformen wirken sich auf die Architektur des Netzes und damit an einigen Stellen auf das Datenschutzkonzept aus, insbesondere bei Überlegungen zur Verhältnismäßigkeit.

4.2 Rollen

Aus der Organisationsform ergibt sich ein Rollenmodell, in dem die Beteiligten nach ihrer institutionellen Zugehörigkeit und ihren Aufgaben klassifiziert werden; dieses Rollenmodell ist die Basis zur Gewährung von Zugriffs- und Funktionsrechten. Wichtige generische Rollen sind behandelnder Arzt und Forscher sowie die jeweiligen Verantwortlichen und Systemadministratoren der verschiedenen Datenbanken und Dienste.

4.3 Weitergabe an Forscher

Forschungsprojekte können innerhalb der Biomaterialbank, d. h. an einer der beteiligten Institutionen, oder extern durchgeführt werden. Zugang zu Proben und Daten sowie Richtlinien zu deren Verwendung sind im Regelwerk festzulegen; über die Freigabe entscheidet ein internes Gremium, das im generischen Konzept als „Ausschuss Datenschutz“ bezeichnet wird.

Die nötigen Daten sind aus dem Gesamtbestand auf den Teilbestand zu selektieren, der im einzelnen Forschungsprojekt faktisch erforderlich wird. Das entspricht dem Prinzip der Datensparsamkeit.

Die Weitergabe von Proben sollte sogar noch restriktiver gehandhabt werden. In der Regel sollten Analysen in der Probenbank oder einer speziell damit beauftragten Institution durchgeführt und an Forschungsprojekte nur die benötigten Ergebnisse exportiert werden.

Die Verpflichtung der Trägers von Biomaterialbanken, die Weitergabe von Proben und medizinischen Daten an Dritte zu kontrollieren, ergibt sich aus der Zweckbindung der Gewinnung von Proben und der Erhebung von Daten, wie sie in der Einwilligungserklärung des Probanden definiert ist, und die sicherzustellen der Träger der BMB in angemessenem Ausmaß auch bei der Weitergabe verpflichtet ist. Jede Weitergabe muss beim Betreiber der BMB nachvollziehbar dokumentiert werden.

Ist dem Spender der Probe zugesichert worden, dass er eine Vernichtung der Probe veranlassen kann oder dass er die Verwendung der Probe für bestimmte Forschungsbereiche auch später noch einschränken kann, ist eine anonymisierte Weitergabe von Material an Forschungsprojekte nur zulässig,

wenn dort zeitnah ein vollständiger Verbrauch vorgesehen ist. In jedem Fall müssen die Möglichkeit hierzu und der potenzielle Empfängerkreis in der Einwilligungserklärung benannt werden. Soll eine Probe, deren Nutzungsrechte durch den Probanden geändert werden können, auf längere Zeit genutzt werden, ist nur eine pseudonymisierte Weitergabe möglich. Grundsätzlich sollte immer der Empfänger vertraglich verpflichtet werden, Analysen zu unterlassen, die auf eine Rückidentifizierung der Probanden ausgerichtet sind.

4.4 Einrichtungen der Biomaterialbank

Ausgangspunkt ist das Standardmodell einer BMB, in dem die Einrichtungen

- Proben- und Datenquellen,
- Probenbank,
- Analysedatenbank,
- Forschungsdatenbank,
- Forschungsprojekt, sowie die Treuhänderdienste
- Identitätsmanagement/Patientenliste,
- Pseudonymisierungsdienst

im Sinne einer möglichst differenzierten informationellen Gewaltenteilung getrennt sind. Es handelt sich hierbei um ein „Maximalmodell“, das zur Kompensation der Offenheit bezüglich Zweck und Verwendung der Biomaterialien zusätzliche Sicherheitsmaßnahmen enthält, die zu einer erwünschten (teilweisen) Redundanz führen. Prinzipiell werden Proben und Daten an verschiedenen Stellen durch unterschiedliche Pseudonyme gekennzeichnet. Dieses Maximalmodell wird im Bild 1 grafisch dargestellt.

5 Pseudonyme

Proben und Daten sind in ihrem Personenbezug spätestens in dem Schritt zu pseudonymisieren, der einen Übergang von der Behandlung zur Forschung darstellt. Behandelnde Ärzte dürfen über die Kombination von personenidentifizierenden Daten und medizinischen Daten verfügen, Forscher dürfen auf die medizinischen Daten nur unter einem Pseudonym zugreifen. Ein direkter Kontakt zwischen Proband und Forscher ist nicht vorgesehen. Führen wissenschaftliche Untersuchungen zu Ergebnissen, die für den individuellen Patienten relevant sind, und ist die Rückmeldung mit ihm vereinbart, so muss der behandelnde Arzt in die Lage versetzt werden, mit diesem Patienten Kontakt aufzunehmen.

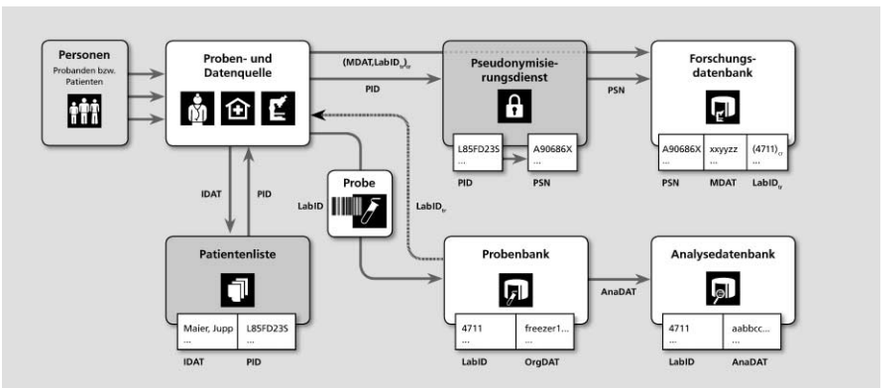

Bild 1 Informationsfluss im Maximalmodell.

Für die Zuordnung der getrennten Teildatenbestände werden die Kennungen

- PID
- PSN
- PSN_i
- LabID
- $LabID_{tr}$

als Pseudonyme verwendet, die jeweils nur unter genau definierten Bedingungen miteinander verknüpfbar sind.

5.1 Der PID

Der PID (Patientenidentifikator) ist ein eindeutiger, aber nichtsprechender Ordnungsparameter für einen in die Biomaterialbank eingeschlossenen Probanden. Die Erzeugung des PID wird durch die anmeldende Stelle veranlasst. Der PID wird gemeinsam mit den Identitätsdaten IDAT in der Patientenliste gespeichert und nur im Bereich dieser einen Biomaterialbank verwendet.

5.2 Das PSN

Das PSN (Pseudonym) ist ein nichtsprechender Identifikator eines Probanden. Das PSN wird aus dem PID durch kryptographische Transformation erzeugt und in der Forschungsdatenbank als Ordnungsparameter verwendet.

5.3 PSN_i, $i = 1, 2, 3, \ldots$

Ist ein Rückidentifizierungspotenzial aus dem Gesamtdatenbestand nicht auszuschließen, muss für die Herausgabe von Daten an ein Forschungsprojekt außerdem ein weiteres Pseudonym PSN_i erzeugt werden, das sich von dem in der Forschungsdatenbank verwendeten unterscheidet. Bei jedem Export ist ein anderes Pseudonym zu verwenden, damit keine externen Datenbestände aufgebaut werden können, die über einen kritischen Informationsbestand verfügen.

5.4 Die LabID

Die LabID dient zur eindeutigen Kennzeichnung einer Probe, etwa als Barcode-Aufkleber. Darüber hinaus wird sie auch als Teil der organisatorischen Daten OrgDAT in der Datenverwaltung der Probenbank mitgeführt. Sie soll nichtsprechend, also pseudonym, sein, um keinen Hinweis auf die Herkunft der Probe zu geben. Sie wird bei der probengewinnenden Stelle erzeugt, und zwar dezentral oder über ein zentrales Online-Verfahren analog zur PID-Erzeugung. Sie wird zusammen mit der Probe – aber ohne identifizierende Daten IDAT

und medizinische Daten MDAT – an die Probenbank weitergegeben. Die LabID dient auch zur Identifizierung der Probenanalysedaten, die ja ohnehin jederzeit wieder aus der Probe gewonnen werden könnten.

5.5 Die $LabID_{tr(ans)}$

Es empfiehlt sich, in der Forschungsdatenbank bei den MDAT eine kryptographisch transformierte $LabID_{tr}$ als Ordnungsparameter zu verwenden. Diese wird von der Probenbank erzeugt und kann auch nur dort der LabID zugeordnet werden. Als Folge davon ist auch die Anforderung von Analysedaten ProbDAT nur über die Probenbank möglich.

6 Allgemeines Datenmanagement und Probenmanagement

6.1 Informationsverteilung und Informationsfluss

Eine Übersicht über die an den verschiedenen Einrichtungen der BMB jeweils vorhandenen Daten gibt Tabelle 1. Ausgehend von diesem Modell sind Vereinfachungen möglich, für deren Zulässigkeit möglichst explizite Kriterien aufzustellen sind; bei PID und LabID ist auch zu analysieren, ob sie an verschiedenen Stellen nur zeitweise gespeichert zu werden brauchen. Der Informationsfluss ist in Bild 1 beschrieben.

6.2 Daten- und Informationsfluss bei Anforderung aus einem Forschungsvorhaben

Ein Forschungsprojekt rekrutiert die zu berücksichtigenden Fälle – nach einem entsprechenden, vom Ausschuss Datenschutz genehmigten Antrag – entweder

- aus der Forschungsdatenbank durch eine Standard-Datenbankabfrage über die MDAT, oder
- aus der Analysendatenbank durch eine Standard-Datenbankabfrage über die ProbDAT;

beides wird in der Regel als Auftrag von einem zuständigen Mitarbeiter der jeweiligen Datenbank ausgeführt, der die gefundenen Fälle mit den benötigten Daten an das Forschungsprojekt exportiert. Bei einer Recherche mit einer Kombination aus MDAT und ProbDAT werden beide Abfragen separat durchgeführt und bei der Probenbank über die $LabID_{tr}$ zusammengeführt; dabei werden die MDAT und ProbDAT der gefundenen Fälle durch Verschlüsselung mit dem Schlüssel des Forschungsprojekts vor den Mitarbeitern der Probenbank verborgen.

Werden Analysen oder Probenanteile selbst von einem Forschungsprojekt unter Angabe des in der FDB vorhandenen Ordnungsparameters $LabID_{tr}$ angefordert, kann die Probenbank mithilfe des Schlüssels hieraus wieder LabID bestimmen und mithilfe der probenbankinternen Verwaltung auf die entsprechende Probe zugreifen. Werden Analysedaten zurückgemeldet, so wird als zugehöriger Ordnungsparameter $LabID_{tr}$ gewählt.

6.3 Auskunft an den Probanden

Wie weit ein Proband Anrecht auf Auskunft an die BMB hat, ist in der Einwilligungserklärung zu regeln. Rein datenschutzrechtlich besteht ein Anspruch nur an die Stellen, wo personenbezogene Daten vorliegen; im Normalfall einer BMB wäre dies die Patientenliste.

Wünscht ein Proband Auskunft über die von ihm in der BMB vorhandenen Daten, wendet er sich an den behandelnden Arzt, der im Kontakt mit der BMB steht oder die auf der Einwilligungserklärung genannte Kontaktperson. Der Arzt löst mit den identifizierenden Stammdaten IDAT des Probanden über die Patientenliste die Auskunftserteilung aus. Der Pseudonymisierungsdienst erhält von der Patientenliste den Personenidentifikator PID des Probanden mit einer Anweisung zur Auskunftserteilung. Diese wird in geeigneter Form geprüft und nach Transformation des PID in das Pseudonym PSN an die Forschungsdatenbank weitergeleitet.

Die Forschungsdatenbank selektiert mithilfe des Ordnungsparameters PSN die zum Probanden gehörenden Daten. Da der interne Ordnungsparameter der zentralen Datenbank nicht in den Behandlungskontext zurückgespiegelt werden darf, werden die Daten mit einem Ad-hoc-Merkmal versehen. Soll die eindeutige Zugehörigkeit zum Probanden sichergestellt werden, wird dasselbe Merkmal zusammen mit dem internen Ordnungsparameter PSN von der Forschungsdatenbank an den Pseudonymisierungsdienst übermittelt, der die PSN in den PID retransformiert und die Referenz PID/Merkmal an den informierenden Arzt sendet.

6.4 Widerruf

Der Proband wendet sich an seinen Ansprechpartner, das ist der betreuende Arzt oder der in der Aufklärung und Einwilligung genannte Verantwortliche der BMB. Dieser

Tabelle 1 Informationsverteilung in einer BMB. „+" bedeutet „vorhanden (zumindest zeitweise)", „(+)" bedeutet „teilweise vorhanden (oder Auswahl)", „?" bedeutet „eventuell".

	IDAT	PID	PSN	PSN_i	LabID	$LabID_{tr}$	Probe	OrgDAT	ProbDAT	MDAT
Quelle	+	+			+		+			+
Probenbank					+	+	+	+		
Analysedatenbank					+				+	
Forschungs-DB			+	+		+		(+)		+
Projekt				+			?	(+)	(+)	(+)
Patientenliste	+	+								
Pseud.-Dienst		+	+							

veranlasst über Patientenliste und Pseudonymisierungsdienst die Löschung bzw. Anonymisierung in der Forschungsdatenbank. Von dort aus wird über die $LabID_{tr}$ gegebenenfalls die Vernichtung der Probe und die Löschung in der Analysendatenbank veranlasst.

6.5 Depseudonymisierung

Die Depseudonymisierung ist zweistufig angelegt: Die erste Stufe wird – technisch gesehen – auf dem inversen Weg der Pseudonymisierung durch die Transformation eines Pseudonyms PSN in einen Patientenidentifikator PID geleistet. In der zweiten Stufe wird der PID an die Patientenliste übersandt, um ihn dort um die Identifikationsdaten zu ergänzen. Beide Stufen können nur von autorisierten Personen nach dem Regelwerk der BMB durchgeführt werden.

7 Dienste

7.1 Patientenliste und Identitätsmanagement für Probanden

In der Patientenliste werden die Identitätsdaten IDAT der Probanden zentral verwaltet und jedem Probanden ein eindeutiger Patientenidentifikator (PID) zugewiesen; weitere Daten werden nicht gespeichert oder übermittelt. Da hier der Rückbezug auf die Identität der Betroffenen unmittelbar möglich ist, muss die Patientenliste an einer vertrauenswürdigen Einrichtung angesiedelt und als TTP-Dienst aufgesetzt sein.

Der PID-Generator ist ein Dienst innerhalb der zentralen Patientenliste. Er erzeugt für jeden Probanden einen Patientenidentifikator PID und übermittelt diesen der datenerhebenden Stelle zurück. Über die identifizierenden Daten IDAT wird im Bestand dieser Liste geprüft, ob der Patient bereits erfasst und ein PID vergeben ist. Im negativen Fall wird ein neuer PID erzeugt und mit den IDAT in den Bestand der Patientenliste übernommen.

Die Patientenliste soll in der Regel zentral für die BMB geführt werden, da die Grundannahme gilt, dass Proben und Daten für einen Probanden von unterschiedlichen Einrichtungen zu verschiedenen Zeitpunkten geliefert werden, auch um z. B. Längsschnittuntersuchungen zu ermöglichen. Bei zentraler Anordnung wird sichergestellt, dass ein Proband trotz unterschiedlicher Schreibweise der IDAT korrekt identifiziert wird, während eine dezentrale Anordnung leicht zu Homonymfehlern (mehr als eine PID pro Person) führen kann.

Die Patientenliste soll auf einem dedizierten Rechner geführt und in einem lokalen Netzwerk durch Firewalls geschützt angeordnet werden. Die Kommunikation mit der Außenwelt erfolgt über einen kontrollierten Kanal unter Nutzung des SSL-Protokolls oder gleichwertiger Lösungen.

7.2 Qualitätssicherung

Qualitätssicherungsaspekte können es erforderlich machen, die Daten auch noch nach ihrer Pseudonymisierung zu prüfen. In diesem Fall wird gemäß dem Modell B des generischen Datenschutzkonzepts [4] eine temporäre Datenbank TempDB eingerichtet. In dieser werden die Daten: PID, MDAT, OrgDAT, LabID oder $LabID_{tr}$, soweit nötig, kurzzeitig gespeichert. Hier können sie zur Qualitätssicherung genutzt werden, bevor eine Überführung in die zentrale Datenbank stattfindet. Bei der Pseudonymisierung und Übertragung in die zentrale Datenbank werden die Daten in der TempDB gelöscht, damit keine Zuordnung zwischen PID und PSN über gleiche Datensätze möglich ist.

7.3 Pseudonymisierungsdienst

Der Pseudonymisierungsdienst ist eine unabhängige Einrichtung zur Erzeugung des Pseudonyms PSN, unter dem die Daten in der Forschungsdatenbank geführt werden. Nur unter Mithilfe dieser Einrichtung kann auch eine Depseudonymisierung vorgenommen werden.

Als Pseudonym PSN wird eine symmetrische kryptographische Transformation des PID verwendet, die vom Pseudonymisierungsdienst ausgeführt wird. Zum Zwecke der Depseudonymisierung ist die Funktion umkehrbar. Der Schlüssel für die Transformation sollte unauslesbar auf einer Smart-Card gespeichert sein, sodass er sicher als Geheimnis bewahrt werden kann.

Der Sender übermittelt medizinische Daten MDAT zusammen mit dem PID an den Pseudonymisierungsdienst. Dabei werden die MDAT im Pseudonymisierungsdienst verschlüsselt durchgeschleust: Sie sind mit dem öffentlichen Schlüssel des Endempfängers, also des Betreibers der Datenbank verschlüsselt und können daher vom Pseudonymisierungsdienst nicht gelesen werden. Der Pseudonymisierungsdienst sendet das PSN mit den verschlüsselten MDAT an den Betreiber der zentralen Datenbank. Dort werden die MDAT entschlüsselt und mit dem PSN in die Forschungsdatenbank übertragen.

Um eine unberechtigte Nutzung des Dienstes auszuschließen, werden die Daten nur von zugelassenen Absendern übernommen. Als Instrument wird die gegenseitige Authentisierung mithilfe des SSL-Protokolls empfohlen.

7.4 Ausschuss Datenschutz

Die Biomaterialbank hat ein zentrales verantwortungstragendes Gremium als Ausschuss Datenschutz zu benennen. Dieser beschließt, überwacht (in Kooperation mit dem gesetzlich geforderten Datenschutzbeauftragten) und pflegt das gesamte datenschutzrelevante Regelwerk und prüft und entscheidet über den Zugang zu den Forschungsdaten sowie die Anträge auf Rückidentifizierung. Der Zugang zu nicht anonymisierten Daten darf nur nach Prüfung eines entsprechenden Antrags erfolgen. Eine eventuell erforderliche Zu-

stimmung einer Ethikkommission zur Beurteilung von Forschungsvorhaben wird dadurch nicht überflüssig gemacht.

7.5 Verwaltung von Zugriffsrechten (Authentifizierung und Autorisierung)

Für die Biomaterialbank ist ein netzweiter zentraler Verzeichnisdienst einzurichten, in dem Rollen und Rechte verwaltet werden.

Bis zu dem Zeitpunkt, an dem die in der Gesundheitstelematik zu schaffende Public Key Infrastructure (PKI) bereitsteht, wird für die Daten erhebenden Stellen eine Authentifizierung über Passwort als ausreichend angesehen, sofern die Übertragung durch SSL kryptographisch gesichert ist.

7.6 Sicherheit in der Datenübermittlung und Dokumentensicherheit

Für die kanalorientierte Sicherheit, die telematische Kommunikation zwischen allen Teilnehmern, ist die Verwendung von SSL vorgesehen. Dokumentenorientierte Sicherheit ist zwingend erforderlich bei der Übermittlung der MDAT von der Klinik bzw. vom Qualitätssicherungsdienst zur zentralen Datenbank, da diese beim Pseudonymisierungsdienst verschlüsselt durchgereicht werden.

Die kanalorientierte Sicherheit wird dadurch erreicht, dass die Daten beim Eintritt in den Transportweg verschlüsselt und beim Austritt wieder entschlüsselt werden. Der Vorgang wird durch das Transportprotokoll ausgelöst und abgeschlossen.

Dokumentenorientierte Sicherheit wird dadurch erreicht, dass ein Dokument verschlüsselt wird, bevor es dem Transportprotokoll übergeben wird. Dies geschieht mit dem öffentlichen Schlüssel des Empfängers. Keine Station, die das Dokument auf dem Wege zum Empfänger erreicht, ist in der Lage, das Dokument zu entschlüsseln. Erst der Besitzer des privaten Schlüssels, der dem öffentlichen Schlüssel zugeordnet ist, vermag dies.

Für die kanalorientierte Sicherheit soll das in Browsern verfügbare Tool SSL – vorläufig mit softwarebasierten Schlüsseln und Zertifikaten für Nutzer und Server – eingesetzt werden.

8 Überlegungen zur Verhältnismäßigkeit

8.1 Redundanz von Sicherheitsmaßnahmen

Datenschutzmaßnahmen sind unter der Maßgabe der Verhältnismäßigkeit zu beurteilen. Verhältnismäßigkeit bedeutet nicht, dass ein kontinuierlicher Sicherheitsparameter mehr oder weniger hoch angesetzt wird, sondern dass Redundanz erhöht oder abgebaut wird. Redundanz ist ein wichtiger Aspekt in Sicherheitskonzepten: Wenn eine Sicherheitsmaßnahme unwirksam wird, soll eine „sichere Rückfallposition“ erreicht werden; unwirksam kann eine Sicherheitsmaßnahme aus verschiedenen Gründen werden:

- Nicht regelgemäßes Verhalten einzelner Beteiligter
- Unbefugte Kooperation verschiedener Beteiligter oder eines Beteiligten mit einem Externen
- Ausfall oder Kompromittierung einer technischen Komponente.

Ein Beispiel für Redundanz ist die Kombination eines Verbots (z. B. in einer vertraglichen Regelung) mit einer technischen Erschwerung (z. B. durch Zugriffskontrolle) oder Überprüfung (z. B. durch Protokollierung von Handlungen).

Unter den für Biomaterialbanken vorgesehenen Maßnahmen stößt eine sehr feingliedrige Trennung der Verantwortung für einzelne Funktionen der Proben- und Datenverwaltung auf Grenzen der Durchführbarkeit, wo Biomaterialbanken von relativ kleinen Organisationseinheiten geführt werden. Eine „kleine“ Biomaterialbank kann mit wenig Redundanz in technischen und organisatorischen Datenschutzmaßnahmen betrieben werden, weil sie als Angriffsziel weniger attraktiv ist, weniger Angriffspunkte bietet, organisatorisch übersichtlich ist und sich daher im Zusammenspiel vieler Komponenten weniger Sicherheitslücken verbergen können; ist sie an der Probenquelle und somit im Bereich der ärztlichen Schweigepflicht angesiedelt, wird dadurch wieder anderweitige Redundanz gegeben.

Für verschiedene Modellvarianten und Organisationsformen sind im Konzept an verschiedenen Stellen Abweichungen von der grundsätzlich geforderten Informationstrennung beschrieben. Hierbei und bei weiteren Abweichungen ist eine Einzelfallprüfung mit Analyse der im folgenden Abschnitt vorgestellten Parameter immer erforderlich.

8.2 Parameter für die Risikoabschätzung

Die für die Risikoabschätzung relevanten Aspekte einer BMB werden in vier Dimensionen gegliedert, die nicht notwendig unabhängig voneinander sind. Es kann keine einfache Formel geben, die aus konkreten Werten für die Parameter die Höhe des Risikos berechnet. Manche Parameter können sich auch gegenläufig auswirken, indem sie an einer Stelle das Risiko erhöhen, es aber an anderer Stelle senken. Sinn dieser Parameterliste ist vielmehr, für eine konkrete BMB potenzielle Schwachstellen aufzudecken.

Zu beachtende Parameter sind:

- Größe der BMB, ausgedrückt durch Fallzahl, Einzugsbereich, Anzahl der Zulieferer, finanzielle Ausstattung und Zahl der Beschäftigten, Komplexität.
- Brisanz der BMB und Attraktivität für Angreifer; Gesichtspunkte hierfür sind Art der Erkrankung, Vollständigkeit der Erfassung, Umfang der Datenerhebung, Forschungsziele, Art des gelagerten Materials, Einzigartigkeit des Materials.
- Organisation der BMB mit den Aspekten Beschlagnahmesicherheit, Präzision der Aufklärung

und Einwilligung, Verteiltheit der Zulieferung, Verteiltheit der Speicherung und Probenlagerung, Umfang geplanter Nacherhebungen, Qualität der Policies und SOPs, vorgesehenes Monitoring-Verfahren.

- Verbindung mit externen Daten wie Abgleichmöglichkeit oder -pläne mit anderen Datenquellen oder Registern, Vorhandensein von Referenzdateien, etwa mit genetischen Fingerabdrücken.

9 Ausblick

Durch die im Datenschutzkonzept beschriebenen organisatorischen und technischen Maßnahmen ist ein rechtlich einwandfreier Umgang mit Biomaterialien und Aufbau von Biomaterialbanken möglich, der auch im Sinne der medizinischen Forschung notwendige Abschwächungen der Einwilligung erlaubt.

Das beschriebene Datenschutzkonzept ist generisch. Es dient als Vorlage, ein konkretes Datenschutzkonzept für eine konkrete Biomaterialbank zu entwickeln. Hierbei kann die AG Datenschutz der TMF Hilfestellung leisten; in einer Reihe von Fällen ist das schon geschehen.

Die erwünschte Verzahnung zwischen Versorgung und medizinischer Forschung erfordert künftig eine enge Integration von Forschungsnetzen und Biomaterialbanken in die Gesundheitstelematik. Die Forschung profitiert von besserem und systematisch gesammeltem Daten- und Probenmaterial und der dort aufzubauenden Sicherheitsinfrastruktur. Die im vorliegenden Konzept beschriebenen Maßnahmen sind auch auf dieses Integrationsszenario übertragbar.

Literatur

[1] P. Ihle, U. Harnischmacher: Patienten-Einwilligung für Biobanken – Investition für die Zukunft (in diesem Heft).

[2] K. Pommerening u. a.: Ein generisches Datenschutzkonzept für Biomaterialbanken. MMV München 2007 (im Druck).

[3] http://www.tmf-ev.de/ [Zugriff am 2. Mai 2007].

[4] C.-M. Reng u. a.: Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin. MMV München 2006.

[5] Nationaler Ethikrat: Biobanken für die Forschung. Stellungnahme 17. März 2004. Online unter http://www.ethikrat.org/themen/pdf/Stellungnahme_Biobanken.pdf [Zugriff am 2. Mai 2007].

[6] Arbeitskreis Wissenschaft der Konferenz der Datenschutzbeauftragten des Bundes und der Länder, Deutsche Arbeitsgemeinschaft für Epidemiologie (DAE): Epidemiologie und Datenschutz, 1998. Online unter http://www.datenschutz-bayern.de/verwaltung/epidem.htm [Zugriff am 2. Mai 2007].

[7] R. Wellbrock: Datenschutzrechtliche Aspekte des Aufbaus von Biobanken für Forschungszwecke. In: Medizinrecht 2003/2, S. 77–82.

[8] R. Wellbrock: Biobanken für die Forschung. Datenschutz und Datensicherheit 28 (2004), S. 561–565.

[9] U. Wauschkuhn, G. Paaß: Datenzugang, Datenschutz und Anonymisierung. Oldenbourg, München 1985.

[10] GMDS-Arbeitsgruppe Datenschutz in Gesundheitsinformationssystemen: Sicherheitsempfehlungen zum Betrieb von Servern und lokalen Netzen in Krankenhäusern. Online unter http://www.eh-cc.de/agdgi/Empfehlungen/Server.html [Zugriff am 2. Mai 2007].

Univ.-Prof. Dr. Klaus Pommerening,
ist Leiter der Arbeitsgruppe Medizinische Informatik am Institut für Medizinische Biometrie, Epidemiologie und Informatik der Universität Mainz sowie Leiter der AG Datenschutz der TMF.
Adresse: Institut für Medizinische Biometrie, Epidemiologie und Informatik der Johannes-Gutenberg-Universität, Obere Zahlbacher Straße 69, 55131 Mainz,
Deutschland,
Tel.: +49 (6131) 173106,
Fax: +49 (6131) 172968,
E-Mail: pommerening@imbei.uni-mainz.de