Präv Gesundheitsf
https://doi.org/10.1007/s11553-022-00991-9
Eingegangen: 11. Juli 2022
Angenommen: 26. September 2022

Johannes Drepper

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

Datenschutzgerechte Wege zur Nutzung von Real World Data

Einerseits verspricht die Real World Data Analysis eine bessere Übertragbarkeit von Forschungsergebnissen in das alltägliche Versorgungsgeschehen. Andererseits sind die datenschutzrechtlichen Anforderungen in der Forschung mit der Einführung der europäischen Datenschutz-Grundverordnung spürbar gestiegen und die Berücksichtigung heterogener Rahmenbedingungen bei der Verarbeitung von Real World Data verkompliziert den datenschutzgerechten Umgang zusätzlich. Hier wird eine Übersicht über relevante Rahmenbedingungen und deren Konsequenzen gegeben sowie auf mögliche Wege einer datenschutzgerechten Nutzung von Real World Data hingewiesen.

Wann unterliegen Real World Data dem Datenschutzrecht?

Die Erhebung von Real World Data folgt keinen starren Vorgaben, wie sie z. B. in einem Studienprotokoll zu finden sind. Die Daten werden oft in sehr unterschiedlichen Situationen erhoben, für die nur schwerlich ein immer passendes, ordnendes Raster zu finden ist. In manchen Kontexten wird man schon den Begriff des Erhebens lieber durch „Anfallen" ersetzen wollen, da letzterer zu schwer planbaren Rahmenbedingungen besser zu passen scheint. Vor diesem Hintergrund ist es besonders wichtig zu verstehen, wann die Verarbeitung von Real World Data dem Datenschutzrecht unterliegt. Im Folgenden sollen daher zunächst die Kriterien herausgearbeitet werden, nach denen entschieden werden kann, ob das Datenschutzrecht bei der Verarbeitung zu beachten ist.

Personenbezug

Das Datenschutzrecht ist auf Daten anwendbar, die einen Bezug zu natürlichen Personen (im Unterschied zu juristischen Personen, wie etwa Firmen, Krankenhäusern, Vereinen etc.) aufweisen. Was mit „Bezug" genau gemeint ist, ist allerdings nicht immer ganz einfach zu verstehen. Hierzu ein Beispiel: Die Trägerin einer Smartwatch hat diese so eingestellt bzw. eine entsprechende Vorkonfiguration so belassen, dass die Uhr über das Internet in regelmäßigen Abständen Fitness- und Gesundheitsdaten an einen Server des Herstellers der Smartwatch überträgt. Wenn diese Daten mit einem Account der Trägerin auf der Website des Herstellers direkt in Verbindung gebracht werden können, und die Account-Informationen auch den Namen, das Alter, vielleicht sogar die Adressdaten der Trägerin umfassen, dann besteht tatsächlich ein Bezug der Daten zu der natürlichen Person – das Datenschutzrecht ist anzuwenden. Wie ist es aber, wenn statt der die Trägerin klar identifizierenden Daten nur eine Nummer (beispielsweise eine IP-Adresse) mit übertragen wird. Die Zuordnung der – vielleicht auch nur dynamisch vergebenen – IP-Adresse zu der Trägerin der Smartwatch wäre in unserem Beispiel zudem nur dem Internet-Provider der Trägerin der Smartwatch und nicht dem Hersteller bekannt.

Interessanterweise gehört diese Frage zu den ganz wenigen Beispielen zur Anwendung des Datenschutzrechts, die bereits einmal höchstrichterlich geklärt wurden. Der Europäische Gerichtshof (EuGH) hat 2016[1] festgestellt, dass speziell solche dynamisch vergebenen IP-Adressen tatsächlich einen Personenbezug herstellen können, da Internetzugangsprovider auf rechtlichem Wege zur Herausgabe der Zuordnung einer IP-Adresse zu einem Kunden gezwungen werden können. Gleichzeitig hat der EuGH zu diesem Anlass aber auch geklärt, dass alleine die Tatsache, dass bei einer bestimmten Stelle eine Zuordnung eines Datensatzes zu einer bestimmten natürlichen Person möglich ist, nicht dazu führt, dass dieser Datensatz für alle verarbeitenden Stellen einen Personenbezug aufweisen muss. Wenn die Zuordnungsinformation bei einer Stelle vor dem Zugriff anderer verarbeitenden Stellen besser geschützt ist als die Verknüpfung einer IP-Adresse mit einem Kunden beim Zugangsprovider, dann können unter bestimmten Voraussetzungen Daten bei einzelnen Verarbeitenden auch ohne Personenbezug und damit ohne Anwendbarkeit des Datenschutzrechts verarbeitet werden.

Pseudonymisierung

Die Methode, Informationen zur direkten Zuordnung von Daten zu einer natürlichen Person von sonstigen Daten abzutrennen und an anderer Stelle zu speichern und zu verarbeiten, ist in der EU-Datenschutz-Grundverordnung (DSGVO) unter dem Begriff der Pseudonymisierung aufgenommen worden (ausführliche Darstellung in [1] ab S. 174). Dementsprechend ist das eine Maßnahme zur Datenminimierung: Es werden nur noch die tatsächlich

[1] Aktenzeichen C-582/14.

Published online: 08 November 2022

notwendigen Daten an einer Stelle verarbeitet. Wenn die Namens- und ggf. Adressangaben von Probanden für eine wissenschaftliche Auswertung keine Rolle spielen, dann sollten sie einer Wissenschaftlerin für die Auswertung auch gar nicht erst zur Verfügung stehen. Werden diese Informationen aber z. B. für die Abwicklung eines Widerrufs einer Einwilligung oder für die Anfrage zu einer Nacherhebung benötigt, dann können sie für solche Zwecke an einer anderen Stelle gebündelt und geschützt vorgehalten werden. Für solche Stellen hat sich die Verwendung der Begriffe „Treuhandstelle", „Datentreuhänder" oder auch „Vertrauensstelle" etabliert [6]. Um Zuordnungen in solchen Fällen zu ermöglichen, werden Pseudonyme genutzt, also nicht-sprechende Zeichenketten oder Nummern, die als Verbindungsstück (Link) zwischen den identifizierenden Daten bei einer Vertrauensstelle und den wissenschaftlich genutzten und pseudonymen Daten bei der Wissenschaftlerin dienen. Solche Maßnahmen zur Datenminimierung, wie z. B. die in der Wissenschaft weit verbreitete Pseudonymisierung, werden von der DSGVO immer dann gefordert, wenn sie die eigentliche Datennutzung und z. B. die Erreichung wissenschaftlicher Ziele nicht hindern (vgl. Art. 89 Abs. 1 DSGVO). Daraus ergibt sich, dass die Nicht-Anwendung einer Pseudonymisierung zu begründen ist.

Zu der Frage, wie gut und weitgehend man die Zuordnungsdaten von den eigentlich wissenschaftlich zu nutzenden Daten trennen muss, gibt es auf Basis umfassender Erfahrung und Abstimmung mit den Datenschutzbehörden bereits einschlägige Empfehlungen (vgl. Kap. 6.7 in [7]). Wer mit einer besonders sicheren Trennung erreichen möchte, dass die Verarbeitung der wissenschaftlichen Daten ohne Berücksichtigung des Datenschutzrechts erfolgen kann, der stößt dann allerdings auf ein weiteres und in aller Regel deutlich größeres Problem: Die Personenbeziehbarkeit, die schon in den wissenschaftlich interessanten Daten enthalten ist und von diesen nicht so einfach getrennt werden kann. Wenn die im obigen Beispiel beschriebene Smartwatch zu den Aktivitäts- und Vitaldaten auch noch Angaben zur Geolokalisation und genaue Zeitstempel erhebt und überträgt, dann wird schnell klar, dass man aus diesen Daten sehr genaue Bewegungs- und Aktivitätsprofile ableiten kann, die wiederum in vielen Fällen wohl eine Identifizierung des Trägers der Smartwatch ganz ohne IP-Adresse oder Namensangabe ermöglichen würden.

Anonymisierung

In der DSGVO findet man bei den Begriffsbestimmungen in Art. 4 keinen Eintrag zur Anonymisierung oder Anonymität von Daten. In Erwägungsgrund Nr. 26 findet man dafür die Erklärung: Dort wird begründet, dass auf die Verarbeitung anonymer Daten eben nicht mehr das Datenschutzrecht anzuwenden ist. Insofern ist Anonymität das Gegenteil der Personenbeziehbarkeit. Wo letztere endet, beginnt erstere. Um die Personenbeziehbarkeit von Daten zu überprüfen, sollen demnach alle Mittel berücksichtigt werden, die nach objektiven Faktoren wie dem Zeitaufwand, den Kosten usw. sowie der verfügbaren Technologie und technologischen Entwicklungen wahrscheinlich zur Identifizierung von Personen in einem Datensatz angewandt werden. Demnach müssen für eine Anonymisierung reichhaltiger medizinischer Datensätze, wie sie gerade auch in der Real World Data Analysis eine Rolle spielen, oft Informationen entfernt werden, die gerade für wissenschaftliche Fragestellungen von großem Interesse sind. Einen einfachen Weg aus diesem Dilemma heraus gibt es zumeist nicht [8]. Schauen wir uns also an, welche Konsequenzen sich aus der Anwendbarkeit des Datenschutzrechts ergeben.

Wer ist verantwortlich für die Verarbeitung von Real World Data?

Wenn eine Verarbeitung personenbezogener Daten stattfindet, ist ein zentraler Dreh- und Angelpunkt die Feststellung der Verantwortlichkeit für diese Verarbeitung. In komplexen Projekten mit vielen Beteiligten und Verarbeitungsschritten ist dabei die Frage nach der Verantwortlichkeit für unterschiedliche Prozessschritte oft differenziert zu beantworten. Für jeden Verarbeitungsschritt ist dabei zu klären, wer über die Mittel und Zwecke der Verarbeitung bestimmt. Derjenige ist nach der DSGVO regelmäßig der Verantwortliche für diese konkrete Verarbeitung. Auch wenn die DSGVO juristisch knapp immer nur von einem Verantwortlichen spricht, ist im Regelfall bei den hier relevanten Verarbeitungen von Gesundheitsdaten bzw. Real World Data nicht eine einzelne Person als Verantwortlicher anzusprechen, sondern die für die Verarbeitung verantwortliche Stelle, also eine juristische Person und nicht eine natürliche Person.

In vielen Projekten zur Nutzung von Real World Data arbeiten heute mehrere Einrichtungen kooperativ miteinander. Wenn mehrere juristisch voneinander unabhängige Stellen gemeinsam die Mittel und Zwecke einer Datenverarbeitung festlegen, sind sie nach Art. 26 DSGVO gemeinsam verantwortlich. In diesem Fall müssen sie in einer Vereinbarung untereinander festlegen, wer für welche Aufgaben und Pflichten eines Verantwortlichen zuständig ist. Der wesentliche Teil dieser Regelung muss zudem den von der Datenverarbeitung betroffenen Personen gegenüber transparent gemacht werden.

Grundsätzlich kann eine Kooperation hingegen auch so aussehen, dass Real World Data von zwei unabhängigen Einrichtungen verarbeitet werden, diese aber jeweils eigenständig über Mittel und Zwecke der Datenverarbeitung bestimmen. In diesen Fällen kann von jeweils eigenständiger Verantwortlichkeit und im Überblick von getrennter Verantwortlichkeit ausgegangen werden.

Keine eigenständige Verantwortlichkeit wird begründet, wenn eine Einrichtung im direkten Auftrag einer anderen eine Verarbeitung von Real World Data durchführt, ohne dabei selbst die Zwecke der Verarbeitung mitbestimmen zu können. Dies wäre z. B. der Fall, wenn die Speicherung und Verarbeitung einer größeren Datenmenge in ein Rechenzentrum ausgelagert würden, ohne dass das Rechenzentrum an der Festlegung von Art und Zweck der Auswertung der Da-

ten beteiligt wäre. In solchen Fällen sieht die DSGVO eine Verarbeitung im Auftrag nach Art. 28 vor, die Verantwortlichkeit bleibt dann bei dem Auftraggeber.

Im Einzelfall kann die Bestimmung der Verantwortlichkeit für einzelne Prozesse der Datenverarbeitung und insbesondere auch für Übermittlungsprozesse überraschend komplex werden. Hierzu hat auch die Rechtsprechung des EuGH in den letzten Jahren beigetragen, die das Fundament für eine sehr breite Anwendbarkeit der gemeinsamen Verantwortlichkeit nach Art. 26 geschaffen hat. So kann z. B. eine Stelle auch verantwortlich sein, wenn sie selber gar keinen Zugriff auf die personenbezogenen Daten hat. Eine zusammenfassende Darstellung für den Kontext der medizinischen Forschung findet sich in einem aktuellen Gutachten von Weichert (Kap. 5 in [13]).

Welche Rechtsgrundlagen erlauben die Verarbeitung von Real World Data?

Ein zentrales Grundprinzip des Datenschutzrechts ist das grundsätzliche Verbot der Verarbeitung personenbezogener Daten. Nur in den Fällen, in denen sich eine Verarbeitung auf eine explizit geregelte Erlaubnis stützen kann, ist sie auch zulässig. Die Klärung, dass eine ausreichende Rechtsgrundlage für die Verarbeitung personenbezogener Daten besteht, ist somit die erste wichtige Aufgabe der verantwortlichen Einrichtung bzw. in bestimmten Fällen auch der verantwortlichen Person.

Die „Rechtmäßigkeit der Verarbeitung" ist zentral in Art. 6 der DSGVO geregelt. Etwas eingeschränktere Regeln in Bezug auf die Erlaubnistatbestände finden sich in Art. 9 DSGVO für besondere Kategorien personenbezogener Daten, deren Verarbeitung mit besonderen Risiken für die Betroffenen verbunden ist oder sein kann. Da zu diesen besonderen Kategorien personenbezogener Daten auch Gesundheitsdaten gehören, müssen diese Einschränkungen auch für die hier interessierende Real World Data Analysis immer mitberücksichtigt werden.

Zusammenfassung · Abstract

Präv Gesundheitsf https://doi.org/10.1007/s11553-022-00991-9

J. Drepper

Datenschutzgerechte Wege zur Nutzung von Real World Data

Zusammenfassung

Hintergrund. Die Nachnutzung vorhandener realweltlicher Daten wird als vielversprechende, die Durchführung klassischer Studien ergänzende Methode der medizinischen Forschung angesehen. Real World Data werden in sehr unterschiedlichen Situationen erhoben und unterliegen damit auch datenschutzrechtlich heterogenen Rahmenbedingungen.
Ziel der Arbeit. Ziel der Arbeit ist die Unterstützung der datenschutzgerechten Nutzung von Real World Data.
Material und Methoden. Neben dem allgemeinen Datenschutzrecht auf europäischer, nationaler und bundeslandspezifischer Ebene werden auch für Gesundheitsdaten spezifische Rechtsgebiete wie die ärztliche Schweigepflicht oder das Sozialrecht beleuchtet. Schutzmethoden wie die Pseudonymisierung und Anonymisierung werden untersucht und eingeordnet.
Ergebnisse. Die Verarbeitung von Real World Data führt im Regelfall zur Anwendung des Datenschutzrechts. Die Klärung der datenschutzrechtlichen Verantwortlichkeit kann bei komplexen Verbundvorhaben anspruchsvoll sein. Die Art der möglichen Rechtsgrundlage für die Verarbeitung hängt von spezifischen Rahmenbedingungen sowie der Art der Verarbeitung ab. Zudem sind die Daten während der Verarbeitung durch technische und organisatorische Maßnahmen zu schützen.
Schlussfolgerung. Die datenschutzrechtlichen Rahmenbedingungen für die Verarbeitung von Real World Data sind komplex. Eine Vereinfachung und Harmonisierung wurde mit der europäischen Datenschutz-Grundverordnung nicht einmal innerhalb Deutschlands erreicht. Bestimmte Wege zur Nutzung dieser Daten, z. B. auf Basis eines „broad consent" oder mit Hilfe einer abgestimmten Bewertung gemäß einer Forschungsklausel, sind mit viel Aufwand verbunden und stehen damit im Regelfall nur größeren Projekten oder Infrastrukturen zur Verfügung.

Schlüsselwörter

Datenschutzrecht · Informierte Einwilligung · Technische und Organisatorische Maßnahmen · Pseudonymisierung · Anonymisierung

Data protection-compliant ways to use real world data

Abstract

Background. The secondary use of existing real world data is seen as a promising method of medical research that complements the conduct of closely controlled studies. However, these real world data are collected in very different situations and are therefore subject to heterogeneous framework conditions in terms of data protection.
Objectives. Supporting the privacy-compliant use of real world data.
Materials and methods. In addition to general data protection laws at the European, national, and state levels, areas of law specific to health data, such as medical confidentiality or social law, are also examined. Protection methods such as pseudonymization and anonymization are examined and classified.
Results. The processing of real world data usually leads to the application of data protection law. Clarifying responsibility under data protection law can be challenging in complex collaborative projects. The type of possible legal basis for processing depends on specific framework conditions as well as the type of processing. In addition, the data must be protected during processing by technical and organizational measures.
Conclusions. The data protection legal framework for the processing of real world data is complex. Simplification and harmonization have not even been achieved within Germany with the European General Data Protection Regulation. Certain ways of using this data, e.g., on the basis of broad consent or with the help of an agreed assessment in accordance with a research clause, involve a great deal of effort and expense and are thus generally only available to larger projects or infrastructures.

Keywords

Data protection law · Informed consent · Technical and organisational measures · Pseudonymization · Anonymization

Informierte Einwilligung

In den allermeisten Anwendungsfällen besteht die gesetzlich notwendige Erlaubnis in einer ausdrücklichen und informierten Einwilligung der Betroffenen nach Art. 9 Abs. 2 a in Verbindung mit Art. 6 Abs. 1 a DSGVO. Wichtige Voraussetzungen der informierten Einwilligung sind zudem in Art. 7 DSGVO geregelt. Hierzu gehören insbesondere die Freiwilligkeit, die Verständlichkeit der Erläuterungen, die Unterscheidbarkeit von anderen Erklärungen sowie die Widerrufbarkeit. Wichtig dürfte für viele Anwendungsfälle sein, dass die DSGVO keine schriftliche Einwilligung mehr fordert. In Erwägungsgrund Nr. 32 DSGVO findet sich der Hinweis, dass die Einwilligung auch „durch Anklicken eines Kästchens beim Besuch einer Internetseite" ausgedrückt werden kann.

Zweckbindung und „broad consent"

Eine weitere wichtige Einschränkung besteht darin, dass jede Einwilligung nach Art. 9 Abs. 2 a „für einen oder mehrere festgelegte Zwecke" erfolgt. Das dahinterstehende Prinzip der Zweckbindung im Datenschutz findet sich auch in dem grundlegenden Art. 5 Abs. 1 b der DSGVO formuliert. Um zu verstehen, was der Gesetzgeber damit konkret meint und wie eng Zwecke demnach festzulegen sind, hilft ein Blick in die Erwägungsgründe der DSGVO, die gerade bei der Auslegung der einzelnen Regelungen heranzuziehen sind. In Erwägungsgrund Nr. 33 hat der Gesetzgeber hierzu festgestellt, dass sich in der wissenschaftlichen Forschung die Zwecke oftmals zum Zeitpunkt der Erhebung der Daten nicht vollständig angeben lassen. Demnach soll auch eine Einwilligung in bestimmte Bereiche wissenschaftlicher Forschung möglich sein. Somit ist immerhin klar, dass das Prinzip der Zweckbindung in der Forschung zwar nicht aufgehoben, aber etwas weiter zu verstehen ist und Zwecke breiter beschrieben werden dürfen als in anderen Zusammenhängen. Was der Gesetzgeber mit der Formulierung „bestimmte Bereiche wissenschaftlicher Forschung" genau meint, bleibt aber leider auch in Erwägungsgrund Nr. 33 DSGVO offen.

Die Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder (Datenschutzkonferenz, DSK) hat 2019, angeregt durch die vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Medizininformatikinitiative (MII; [11]), einen Beschluss zur Auslegung dieses Begriffs gefasst [3]. Zwar wird man nach Lektüre des Dokuments auch nicht schlauer in Bezug auf den Bedeutungsgehalt dieses Begriffs. Dafür beschreibt das Dokument aber konkret, unter welchen Rahmenbedingungen und Voraussetzungen man sich auf den in Erwägungsgrund Nr. 33 DSGVO formulierten Ausnahmetatbestand aus Sicht der Datenschutzbehörden stützen darf. Mit den Einwilligungsdokumenten der MII zur breiten Nachnutzung von klinischen Behandlungsdaten für die medizinische Forschung[2] („broad consent") liegt seit 2020 immerhin ein Beispiel für eine informierte Einwilligung vor, welche von der DSK vor dem Hintergrund des eigenen Beschlusses aus 2019 explizit anerkannt und für einsetzbar gehalten wurde.[3] Schaut man sich den Aufwand an, den die MII mit der Abstimmung und Etablierung dieses „broad consent" sowie des Aufbaus der dafür notwendigen Infrastruktur auf sich genommen hat [14], so wird schnell klar, dass dieses Modell eines „broad consent" nicht von jedem beliebigen Forschungsprojekt mit eng begrenztem Budget übernommen werden kann. Eher handelt es sich um ein Modell für wenige, große Infrastrukturen, die umfangreiche Datenbestände langfristig und für breite Zwecke zugänglich machen.

Zweckvereinbarkeit

Neben dem in Erwägungsgrund Nr. 33 DSGVO ausgeführten Auslegungshinweis zum „broad consent" kennt die DSGVO noch eine weitere wichtige Ergänzung zum Zweckbindungsgrundsatz: die Zweckvereinbarkeit. Art. 5 Abs. 1 b DSGVO stellt zur Zweckbindung fest, dass Daten nicht in einer mit dem ursprünglich erlaubten Zweck unvereinbaren Weise weiterverarbeitet werden dürfen. Dies impliziert, dass es auch vereinbare Zwecke der Weiterverarbeitung gibt. An derselben Stelle wird sogar der Grundsatz postuliert, dass eine Weiterverarbeitung für wissenschaftliche Zwecke unter Berücksichtigung der in Art. 89 Abs. 1 DSGVO formulierten Zusicherungen als vereinbar mit den ursprünglichen Zwecken gilt. So richtig verständlich wird die Zweckvereinbarkeit allerdings erst, wenn man sich die hierfür beschriebenen fünf Kriterien in Art. 6 Abs. 4 DSGVO anschaut. Hier wird deutlich, dass es nicht nur um eine Prüfung der Zwecke geht, sondern vielmehr um eine Vereinbarkeit der Verarbeitungen. So spielt beispielsweise auch das Verhältnis des Verantwortlichen für die Verarbeitung zu den Betroffenen eine Rolle oder die möglichen Folgen einer Weiterverarbeitung für die davon betroffenen Personen. Insofern geht es hier – etwas verkürzt ausgedrückt – um den Grundsatz, dass bestimmte Weiterverarbeitungen so wenige Risiken für die betroffenen Personen bergen, dass hierfür die Heranziehung immer neuer Rechtsgrundlagen nicht zu rechtfertigen wäre. Entsprechend hilft auch Erwägungsgrund Nr. 50 DSGVO bei der Auslegung der Zweckvereinbarkeit weiter, in dem er postuliert, dass im Falle der Vereinbarkeit keine neue Rechtsgrundlage notwendig ist. Als Beispiel einer solchen Weiterverarbeitung könnte man das Zählen von Datensätzen in einem Real-World-Dataset ansehen, um z. B. die Machbarkeit eines Auswertungsdesigns abzuschätzen. Das Zählen selbst ist noch als Verarbeitung eines personenbezogenen Datensatzes anzusehen, auch wenn das Ergebnis eine absolut anonyme Fallzahl ist. Die Art der Verarbeitung und die Nutzung einer anonymen Fallzahl haben jedoch so wenig Einfluss auf die davon betroffenen Personen, dass die Zweckvereinbarkeit in solchen Fällen die Suche nach einer neuen Rechtsgrundlage entbehrlich macht. Das Konstrukt der

[2] siehe https://www.medizininformatik-initiative.de/de/mustertext-zur-patienteneinwilligung.

[3] siehe https://www.datenschutzkonferenz-online.de/media/pm/20200427_Einwilligungsdokumente_der_Medizininformatik-Initiative.pdf.

Zweckvereinbarkeit ist in der rechtswissenschaftlichen Literatur immer noch ein wenig umstritten, initial wurde sogar von einem diesbezüglichen Redaktionsversehen in Erwägungsgrund Nr. 50 ausgegangen [9]. Allerdings hat auch eine zwischenzeitliche redaktionelle Korrektur der DSGVO in ihren Übersetzungen nicht zu einer Änderung der relevanten Aussagen in Erwägungsgrund Nr. 50 DSGVO geführt und die aktuelle Kommentierung folgt hier auch eher dem Gesetzestext und stellt diesen nicht mehr in Frage (z. B. Roßnagel, Art. 5 Rn. 99f in [12]). Die Privilegierung wissenschaftlicher Forschungszwecke wird in diesem Zusammenhang auch gerade damit erklärt, dass sich die Datenverarbeitung hier typischerweise nicht auf die einzelne Person bezieht, deren Daten verarbeitet werden (Roßnagel, Art. 5 Rn. 104 in [12]). Eine zusammenfassende Darstellung findet sich in dem bereits erwähnten Gutachten von Weichert (Kap. 4 in [13]).

Forschungsklauseln

Welche Möglichkeiten der Verarbeitung reichhaltiger Real World Data hat man aber, wenn die Einholung einer Einwilligung viel zu aufwändig oder gar unmöglich wäre und die Verarbeitung auch nicht so geringfügig ist, dass man eine Zweckvereinbarkeit mit ursprünglich gerechtfertigten Zwecken annehmen könnte? In solchen Fällen ist zu klären, ob es einen gesetzlich geregelten Erlaubnistatbestand unabhängig von einer Einwilligung gibt. Dies können konkrete Erlaubnisse sein, wie sie z. B. für die Erhebung onkologischer Daten in den Krebsregistern auf den Ebenen der Länder und des Bundes geregelt sind. Für die Forschung sind im Regelfall jedoch die Forschungsklauseln ergiebiger, die der verantwortlichen Stelle regelhaft eine Abwägung zwischen den (öffentlichen) Forschungsinteressen und den Interessen der Betroffenen auferlegen. Kommt diese Abwägung zu dem Schluss, dass die Interessen an der Durchführung der Forschung die Interessen der Betroffenen (erheblich) überwiegen, kann in bestimmten Fällen eine Verarbeitung der Daten auch ohne Einwilligung rechtmäßig sein. Das prominenteste Beispiel einer solchen Forschungsklausel findet sich in § 27 Abs. 1 des Bundesdatenschutzgesetzes (BDSG), der eine Datenverarbeitung auch von Gesundheitsdaten zu wissenschaftlichen Forschungszwecken ohne Einwilligung erlaubt, wenn die Verarbeitung dafür erforderlich ist und die Interessen des Verantwortlichen an der Verarbeitung die Interessen der betroffenen Person an einem Ausschluss der Verarbeitung erheblich überwiegen. Eine aktuelle Übersicht zu weiteren Regelungen dieser Art im nationalen Recht bietet nochmal das Gutachten von Weichert (Kap. 4.5 ab S. 37 in [13]).

Bei größeren Verbundprojekten mit der Nutzung von Daten aus mehreren Bundesländern wird man bei den Forschungsklauseln regelmäßig auf das Problem stoßen, dass diese in den Landesgesetzen alle etwas unterschiedlich formuliert und ausgeprägt sind. Zudem werden unterschiedliche Beteiligte und ggf. auch prüfende Datenschutzbehörden bei der notwendigen Abwägung der Interessen oft zu unterschiedlichen Ergebnissen kommen, was die Rechtssicherheit der Anwendung erheblich einschränken kann. Zwar hat der Bundesgesetzgeber im Rahmen der Gesetzgebung zur Coronapandemie im Frühjahr 2020 mit § 287a Sozialgesetzbuch Nr. 5 (SGB V) eine Regelung für die Verbundforschung erlassen, die die Anwendbarkeit von § 27 Abs. 1 BDSG bundeslandübergreifend sowie auch eine federführend zuständige Aufsicht regelt. Allerdings ist die verfassungsrechtliche Zuständigkeit des Bundes für eine solche Regelung sowie auch die Anwendbarkeit der Regelung außerhalb des Sozialrechts umstritten (vgl. hierzu [5]). Zudem bleibt auch bei einheitlicher Anwendung der Abwägung nach § 27 Abs. 1 BDSG das Problem, dass das Ergebnis der Abwägung unterschiedlich gewertet werden kann und ein verlässliches Ergebnis kaum zu erzielen ist.

Ärztliche Schweigepflicht

Bei der Verarbeitung von Real World Data ist weiter zu beachten, dass diese oft aus einem ärztlichen Behandlungsverhältnis stammen und hier die berufsrechtlich und strafrechtlich geregelte ärztliche Schweigepflicht eine zusätzliche Schranke vor der Weiterverwendung der Daten in anderen Kontexten darstellt. Patientinnen und Patienten sollen sich den behandelnden Personen uneingeschränkt anvertrauen dürfen. Insbesondere sollen schlechte oder gar gefährliche Behandlungen, die auf fehlenden Informationen beruhen, weitestgehend ausgeschlossen werden. Im Unterschied zum Datenschutzrecht stellt hier allerdings die Zweckbindung kein eigenständiges Ziel der ärztlichen Schweigepflicht dar [2]. Vielmehr geht es hier um den Kreis derjenigen, denen die Informationen aus der Behandlung offenbart werden. Diese müssen in die Behandlung eingebunden sein oder auch als berufsfremde Personen an der Behandlung mitwirken (vgl. § 203 Strafgesetzbuch). Anderen Personen dürfen die Daten aus der Behandlung nicht offenbart werden, es sei denn, die betroffene Patientin oder der betroffene Patient hat dies erlaubt und insoweit die behandelnde Person von der Schweigepflicht entbunden. Eine informierte Einwilligung (s. oben), die sich explizit auf die Daten aus dem geschützten Behandlungsverhältnis bezieht und zudem explizit darstellt, dass die Daten diesen geschützten Bereich verlassen werden, kann als implizite Schweigepflichtentbindung gewertet werden.[4] Ohne eine Entbindung von der Schweigepflicht ist eine Weiterverwendung der Daten aus dem Behandlungskontext nur möglich, wenn die Verarbeitung durch in den Behandlungskontext eingebundene Personen erfolgt (und eine datenschutzrechtliche Grundlage für die Zweckänderung bzw. Zweckvereinbarkeit vorliegt) oder eine gesetzliche Erlaubnis existiert, die sich spezifisch genug auf die Daten aus dem geschützten Behandlungsverhältnis bezieht und explizit eine Offenbarung der Daten gegenüber anderen Personen gestattet oder fordert (s. S. 76 in [10]). Die oben dargestellte Forschungsklausel in § 27 Abs. 1 BDSG stellt übrigens nach § 1 Abs. 2 BDSG keine Offenbarungsbefugnis dar. Auch andere datenschutzrechtliche Erlaubnis-

[4] So auch in den Einwilligungsdokumenten der MII, dort zudem von Ethikkommissionen und Datenschutzbehörden umfangreich geprüft.

normen scheiden somit im Regelfall als Offenbarungsbefugnis aus.

Sozialrecht

Zusätzliche zu berücksichtigende rechtliche Rahmenbedingungen können sich ergeben, wenn Sozialdaten in Real World Data Analysis genutzt werden. Diese Daten, wie sie z. B. von Krankenkassen genutzt und z. T. auch für die Forschung zur Verfügung gestellt werden können, unterliegen den Regelungen aus dem Sozialrecht und hier insbesondere den Regelungen der SGB V (Gesetzliche Krankenversicherung) und X (Sozialdatenschutz). Eine ausführliche Darstellung der sich daraus ergebenden Einschränkungen und Möglichkeiten verbietet sich hier aus Platzgründen. Eine vergleichsweise aktuelle Darstellung der Möglichkeiten und Rahmenbedingungen aus Sicht der medizinischen Forschung findet sich in einem Gutachten von Dierks und Roßnagel aus 2019 [1].

Welche Pflichten ergeben sich aus der Nutzung von Real World Data?

Transparenz

Aus der Verantwortlichkeit für eine Verarbeitung personenbezogener Daten folgt eine Reihe von Pflichten. Neben den allgemeinen Rechenschafts- (vgl. Art. 5 Abs. 2 DSGVO) und Dokumentationspflichten (vgl. Art. 30 DSGVO), denen oftmals durch die Erstellung und Abstimmung eines Datenschutzkonzepts nachgekommen wird, sind dies zuvörderst Transparenzpflichten nach Art. 12–14 DSGVO. Demnach sind die von der Verarbeitung betroffenen Personen namentlich über den Verantwortlichen oder auch die verantwortlichen Stellen sowie ggf. über Kontaktmöglichkeiten zu bestellten Datenschutzbeauftragten zu informieren. Zudem müssen Dauer und Zwecke der Datenverarbeitung sowie Empfänger oder Kategorien von Empfängern der Daten transparent gemacht werden. Und nicht zuletzt müssen die betroffenen Personen auch über ihre Rechte in Bezug auf die Daten, zu denen es noch einen Bezug zu ihrer Person gibt, ins Licht gesetzt werden.

Betroffenenrechte

Selbstverständlich ist über diese Rechte nicht nur zu informieren, diese müssen auch gewährt bzw. im Falle der Beanspruchung vom Verantwortlichen oder einem Beauftragten umgesetzt werden. Das sind die Rechte auf Auskunft (Art. 15), auf Berichtigung (Art. 16) und auf Löschung der von der Verarbeitung betroffenen personenbezogenen Daten (Art. 17) sowie auf Einschränkung der Verarbeitung (Art. 18), auf Übertragbarkeit selbst bereitgestellter Daten (Art. 20) sowie auf Widerspruch zu einer Verarbeitung auf gesetzlicher Grundlage (Art. 21 DSGVO).

Zu diesen Betroffenenrechten ist zu sagen, dass einige davon im Forschungskontext auch etwas eingeschränkt werden können. Das Recht auf Löschung gilt nach Art. 17 Abs. 3 d DSGVO nicht, wenn die Gewährung voraussichtlich die Verwirklichung der Ziele der Verarbeitung unmöglich machen oder ernsthaft beeinträchtigen würde. Nach Weichert (s. Antwort auf Frage 10.1 auf S. 203 in [13]) kann diese Ausnahme angewendet werden, wenn die Daten eines Forschungsprojekts nach guter wissenschaftlicher Praxis zum Nachvollziehen der Ergebnisse für 10 Jahre aufbewahrt werden. Erfolgt in diesem Zeitraum von einer betroffenen Person ein Widerruf einer Einwilligung oder ein Widerspruch zu einer gesetzlich gestützten Verarbeitung, so würde die Nachvollziehbarkeit des gesamten Projekts gefährdet, was die Anwendung dieser Ausnahme erlaubt. Für weitere Betroffenenrechte sind in der DSGVO lediglich Öffnungsklauseln formuliert (z. B. in Art. 89 Abs. 2), die von den Ländern in der EU zusätzlich im nationalen Recht umzusetzen und zu konkretisieren sind (vgl. § 27 Abs. 2 BDSG).

Im Umgang mit Real World Data ist zu berücksichtigen, dass sich aus den Daten ggf. auch sehr sensible Informationen über die betroffenen Personen gewinnen lassen, so z. B. Hinweise auf eine Anfälligkeit für eine sehr ernsthafte Erkrankung. Diese Anfälligkeit muss der betroffenen Person nicht in jedem Fall schon bekannt sein, könnte aber von einem Recht auf Auskunft betroffen sein. In solchen Fällen ist daher auch das Recht auf Nichtwissen der betroffenen Personen zu berücksichtigen und ggf. eine ethisch begründete Entscheidung im Einzelfall zu treffen, wenn die Gesetzeslage eine Einschränkung des Auskunftsrechts zulässt.

Technische und Organisatorische Maßnahmen

Der Verantwortliche hat mit geeigneten Mitteln für die Sicherheit der Verarbeitung der personenbezogenen Daten zu sorgen (vgl. Art. 32 DSGVO). Hierzu gehört zunächst, die Risiken für die Rechte und Freiheiten der betroffenen Personen auf ein vertretbares Maß zu beschränken. Das bedeutet insbesondere, dass die klassischen Gewährleistungsziele der Vertraulichkeit, Integrität und Verfügbarkeit der Datenverarbeitung sicherzustellen sind. Zur Gewährleistung der Integrität und Verfügbarkeit gehört typischerweise die Einhaltung von Standards der IT-Sicherheit, wie sie beispielsweise in den Vorgaben zum IT-Grundschutz vom Bundesamt für Sicherheit in der Informationstechnik (BSI) niedergelegt sind.[5]

Die Vertraulichkeit wird hingegen in der Forschung regelmäßig mit Mitteln wie der Pseudonymisierung unterstützt. Dazu werden Informationen für unterschiedliche Anwendungen und Anwender aufgeteilt, so dass beispielsweise eine Person, die die Auswertung der Daten vornimmt, keine direkt identifizierenden Daten wie etwa Namens- oder Adressangaben erhält. Für die Kontaktierung der betroffenen Personen, z. B. um zusätzliche Daten anzufragen, könnte aber eine hierfür eingebundene Vertrauensstelle die identifizierenden Daten selektiv nutzen. In der Vertrauensstelle wären dann dafür andere Daten, die für die Kontaktierung nicht benötigt werden, nicht vorhanden. Auch die Verschlüsselung von Daten bei der Speicherung oder Übertragung gehört in aller Regel

[5] https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/it-grundschutz_node.html.

zu den hierfür anzuwendenden Maßnahmen. Und organisatorische Maßnahmen, wie z. B. die Bestimmung eines kompetent besetzten Gremiums, welches über die Freigabe von Daten für spätere Forschungsprojekte entscheidet, gehören im Regelfall zentral zum Maßnahmenkatalog dazu.

Bei der Festlegung der Maßnahmen wäre es natürlich hilfreich zu wissen, welche Maßnahmen in der medizinischen Forschung typischerweise bei welcher Art von Forschungsprojekt anzuwenden sind. Hierfür gibt es den Leitfaden der TMF [7], der zwar schon etwas in die Jahre gekommen ist, aber in Bezug auf die Festlegung eines passenden Mix aus technischen und organisatorischen Maßnahmen für unterschiedliche Arten von Forschungsprojekten auch heute noch gut als Ausgangspunkt genutzt werden kann. Zudem zeichnet diesen Leitfaden aus, dass er im Jahr der Veröffentlichung auch von der Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder geprüft und für die Anwendung in der medizinischen Forschung empfohlen wurde.[6]

Datenschutz-Folgenabschätzung

Bei umfangreicher Verarbeitung von Gesundheitsdaten – bei Real World Data Analysis wohl oft anzunehmen – ist nach Art. 35 DSGVO ergänzend zur Festlegung der technischen und organisatorischen Maßnahmen eine Datenschutz-Folgenabschätzung durchzuführen, die eine umfassende Analyse der mit einer Verarbeitung verbundenen Risiken für die Rechte und Freiheiten der betroffenen Personen umfasst. Die Bestimmung der Risiken erfolgt dabei unter Berücksichtigung der geplanten technischen und organisatorischen Maßnahmen zum Schutz der Daten und zum Schutz der Rechte und Freiheiten der betroffenen Personen. Die Prüfung, ob eine Datenschutz-Folgenabschätzung durchgeführt werden muss, und dann ggf. die Durchführung derselben obliegen als Pflichten dem oder den Verantwortlichen (vgl. [4]). Kommt die Datenschutz-Folgenabschätzung zu dem Ergebnis, dass eine Verarbeitung trotz der getroffenen Sicherheitsmaßnahmen weiterhin mit hohen Risiken für die Betroffenen einhergeht, muss vor dem Start der Verarbeitung die zuständige Datenschutzbehörde konsultiert werden (Art. 36 Abs. 1 DSGVO).

Meldungen

Zu guter Letzt müssen Verantwortliche auch darauf vorbereitet sein, im Falle von Datenpannen, diese ggf. an Betroffene (Art. 34 DSGVO) oder Datenschutzbehörden (Art. 33 DSGVO) zu melden.

Wo bekomme ich welche Unterstützung zum Datenschutz bei der Nutzung von Real World Data?

Die Beantwortung der meisten hier bisher gestellten Fragen erfolgt idealerweise im Rahmen eines Datenschutzkonzepts. Somit bietet dieses eine gute Grundlage für eine datenschutzgerechte Verarbeitung von Real World Data und kann insbesondere auch für die Abstimmung mit allen Beteiligten zu datenschutzkritischen Aspekten genutzt werden.

Ein solches Datenschutzkonzept stellt zunächst die Anwendungsfälle und somit die Zwecke der Datenverarbeitung allgemeinverständlich dar. Weiter ist die verantwortliche Stelle oder sind die verantwortlichen Stellen aufzuführen und ihre jeweiligen Rollen bzw. Zuständigkeitsbereiche zu benennen. Dem schließt sich eine Darstellung der Rechtsgrundlagen der Verarbeitung personenbezogener Daten an. Ausführlich ist auf die technischen und organisatorischen Maßnahmen zum Schutz der Daten einzugehen. Ist die Durchführung einer Datenschutz-Folgenabschätzung nach Art. 35 Abs. 1 bzw. Abs. 3 DSGVO notwendig, kann sich dann die Dokumentation einer durchgeführten Risikoanalyse anschließen, in der je identifiziertem Risiko auf die passenden, bereits aufgeführten technischen und organisatorischen Schutzmaßnahmen verwiesen wird, die genau dieses Risiko mindern. Somit kann die Dokumentation zu einer durchgeführten Datenschutz-Folgenabschätzung in ein Datenschutzkonzept eingebettet werden (vgl. S. 137f in [13]). Abschließend ist auf die präzise Umsetzung der Betroffenenrechte sowie die Fristen für die Nutzung und Speicherung personenbezogener Daten einzugehen.

Für die Bestimmung passender technischer und organisatorischer Schutzmaßnahmen bei typischen Anwendungsfällen kann immer noch der TMF-Leitfaden aus 2014 genutzt werden [7]. Dazu bietet die Arbeitsgruppe Datenschutz der TMF nach wie vor ein umfassendes Beratungsangebot an, welches für Mitglieder der TMF sogar bis zu einem schriftlichen Votum reicht, welches die Einhaltung der von den zuständigen Behörden geprüften und empfohlenen Maßnahmen bestätigen kann.[7]

Ein reichhaltiges Sortiment an Tools, Unterlagen, Beratungs- und Schulungsangeboten bietet das im Aufbau befindliche Portal ToolPool-Gesundheitsforschung an.[8] So können hier Templates und Checklisten für Datenschutzkonzepte, frei verfügbare Pseudonymisierungs- und Anonymisierungswerkzeuge, Rechtsgutachten, Schulungsangebote zum Datenschutz allgemein oder zur Nutzung bestimmter Tools im Besonderen sowie viele weitere Hilfsmittel und Informationen gefunden und genutzt werden.

Fazit für die Praxis

- **Die sich aus dem Datenschutzrecht ergebenden Anforderungen an die Verarbeitung und Nutzung von Real World Data dürfen nicht unterschätzt werden.**
- **Ausreichende Ressourcen und Vorlaufzeiten für die Klärung relevanter Sachverhalte sind einzuplanen.**
- **In den meisten Fällen sind die beteiligten Einrichtungen als juristische Personen für die Verarbeitung der personenbezogenen Daten datenschutzrechtlich verantwortlich und nicht einzelne Forscherinnen oder**

[6] https://datenschutz.hessen.de/datenschutz/statistik-und-wissenschaft/wissenschaft/neues-rahmenkonzept-f%C3%BCr-die-vernetzte.

[7] siehe www.tmf-ev.de/datenschutz.

[8] siehe www.toolpool-gesundheitsforschung.de.

Forscher. Daher sind die Datenschutzbeauftragten der beteiligten Stellen frühzeitig anzusprechen und einzubinden.
- **Bestehenden Lösungsansätzen, Konzepten, Tools ist im Regelfall der Vorzug zu geben. In diesem komplexen Umfeld muss nicht alles neu oder selbst erfunden werden.**
- **Bestehende Austausch- und Beratungsangebote, wie sie z. B. die TMF mit ihrer Arbeitsgruppe Datenschutz bietet, sollten genutzt werden.**

Korrespondenzadresse

Dr. Johannes Drepper
TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V.
Berlin, Deutschland
johannes.drepper@tmf-ev.de

Funding. Open access funding was provided by Roche Pharma AG, Novartis Pharma GmbH, Pfizer Pharma GmbH, and Takeda Pharma Vertrieb GmbH & Co. KG.

Einhaltung ethischer Richtlinien

Interessenkonflikt. J. Drepper gibt an, dass kein Interessenkonflikt besteht.

Für diesen Beitrag wurden von den Autor/-innen keine Studien an Menschen oder Tieren durchgeführt.

Literatur

1. Dierks C, Roßnagel A (2019) Sekundärnutzung von Sozial- und Gesundheitsdaten – Rechtliche Rahmenbedingunge. Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin https://doi.org/10.32745/9783954665181
2. Dochow C (2019) Unterscheidung und Verhältnis von Gesundheitsdatenschutz und ärztlicher Schweigepflicht (Teil 1). MedR 2019:279–287
3. DSK (2019) Beschluss der 97. Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder zu Auslegung des Begriffs „bestimmte Bereiche wissenschaftlicher Forschung" im Erwägungsgrund 33 der DS-GVO. In:Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder – Datenschutzkonferenz (DSK), https://www.datenschutzkonferenz-online.de/media/dskb/20190405_auslegung_bestimmte_bereiche_wiss_forschung.pdf (Zugegriffen: 24. Jun. 2022)
4. DSK (2018) Kurzpapier Nr. 5: Datenschutz-Folgenabschätzung nach Art. 35 DS-GVO. In:Konferenz der unabhängigen Datenschutzbehörden des Bundes und der Länder – Datenschutzkonferenz (DSK). https://www.datenschutzkonferenz-online.de/media/kp/dsk_kpnr_5.pdf. Zugegriffen: 24. Juni 2022
5. Graf von Kielmansegg S (2021) Gesetzgebung im Windschatten der Pandemie: § 287a SGB V und der Datenschutz in der Gesundheitsforschung. VerwArch 2021:133–168
6. Metschke R, Wellbrock R (2002) Datenschutz in Wissenschaft und Forschung. In:Berliner Beauftragter für Datenschutz und Informationsfreiheit, Berlin. https://www.forschungsdaten-bildung.de/files/metschkewellbrock2002.pdf. Zugegriffen: 24. Juni 2022
7. Pommerening K, Drepper J, Helbing K et al (2014) Leitfaden zum Datenschutz in medizinischen Forschungsprojekten – Generische Lösungen der TMF 2.0. Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin https://doi.org/10.32745/9783954662951
8. Sariyar M, Schlünder I (2016) Reconsidering anonymization-related concepts and the term "identification" against the backdrop of the European legal framework. Biopreserv Biobank 14:367–374
9. Schantz P (2016) Die Datenschutz-Grundverordnung – Beginn einer neuen Zeitrechnung im Datenschutzrecht. NJW 2016:1841–1904
10. Schneider UK (2015) Sekundärnutzung klinischer Daten – Rechtliche Rahmenbedingungen. Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin https://doi.org/10.32745/9783954663224
11. Semler SC, Wissing F, Heyder R (2018) German Medical Informatics Initiative. A National Approach to Integrating Health Data from Patient Care and Medical Research. Methods Inf Med 57:e50–e56
12. Simitis S, Hornung G, Spiecker gen. Döhmann (Hrsg) (2019) Datenschutzrecht. DSGVO mit BDSG. Großkommentar. Nomos, Baden-Baden
13. Weichert T (2022) Datenschutzrechtliche Rahmenbedingungen medizinischer Forschung. Vorgaben der EU-Datenschutz-Grundverordnung und national geltender Gesetze. Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin https://doi.org/10.32745/9783954667000
14. Zenker S, Strech D, Ihrig K et al (2022) Data protection-compliant broad consent for secondary use of health care data and human biosamples for (bio)medical research: Towards a new German national standard. J Biomed Inform 131:104096