

Excel?! Ernsthaft?

Wie die Toolbox for Research eine standortübergreifende Datenerhebung in kleinen Forschungsprojekten vereinfacht

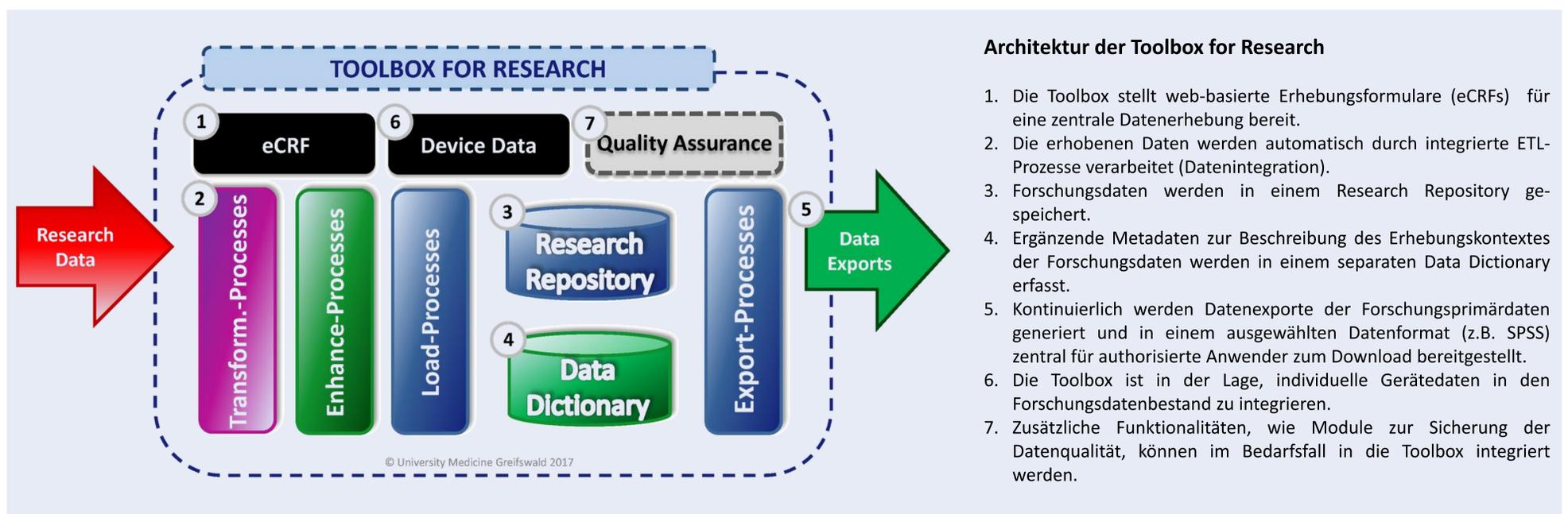
Rau H¹, Blumentritt A¹, Schuldt R¹, Penndorf P¹, Piegsa J¹, Thamm OC², Bahls T¹, Hoffmann W¹, Bialke M¹

¹Institut für Community Medicine, Abteilung Versorgungsepidemiologie und Community Health, Universitätsmedizin Greifswald
²Sana-Krankenhaus Gerresheim

Hintergrund

Limitierte Ressourcen bezüglich Budget, Personal und IT-Infrastrukturen gehören zu verbreiteten Kennzeichen der Epidemiologie und Versorgungsforschung. Insbesondere kleineren Registern und Kohortenstudien fehlt es oft an Personal mit Programmierkenntnissen. Daher behelfen sich derartige Studien häufig mit vermeintlich einfachen Erhebungs- und Studienteilnehmerverwaltungsverfahren anstelle eines IT-gestützten Datenmanagements inklusive Studiendatenbanken. Oftmals werden für kleinere Forschungsprojekte dezentrale Excel-Tabellen genutzt, in die die zu erhebenden Daten eingetragen und die in regelmäßigen Abständen manuell zusammengeführt werden. Automatisierte Pseudonymisierungsverfahren sind nicht die Regel, was die datenschutzgerechte Trennung von identifizierenden (IDAT) und medizinischen Daten (MDAT) erschwert. Im Rahmen des MOSAIC-Projektes [1] wurde eine flexible Software-Lösung für das Datenmanagement in kleineren Forschungsvorhaben kostenfrei bereitgestellt. Diese „Toolbox for

Research“ (kurz: Toolbox) ist für eine Vielzahl von Anwendungsszenarien geeignet und unterstützt bei der standortübergreifenden Erhebung, Verarbeitung und Speicherung von Forschungsdaten. Die automatische Installation der entstandenen Open Source Lösung wurde wesentlich durch den Einsatz von Docker und eine umfangreiche Begleitdokumentation vereinfacht. Da in kleineren Forschungsprojekten eine automatisierte Trennung von MDAT und IDAT ohne Treuhandstellenunterstützung [1] schwer umzusetzen ist, sollen innerhalb der Toolbox keine personenidentifizierenden Daten gespeichert werden. Um die medizinischen Forschungsdaten innerhalb der Toolbox und die IDAT aus dem klinischen Kontext dennoch nachvollziehbar in Beziehung setzen zu können, muss die Toolbox ein einheitliches und im Bedarfsfall transparentes Pseudonymisierungskonzept umsetzen.



Architektur der Toolbox for Research

1. Die Toolbox stellt web-basierte Erhebungsformulare (eCRFs) für eine zentrale Datenerhebung bereit.
2. Die erhobenen Daten werden automatisch durch integrierte ETL-Prozesse verarbeitet (Datenintegration).
3. Forschungsdaten werden in einem Research Repository gespeichert.
4. Ergänzende Metadaten zur Beschreibung des Erhebungskontextes der Forschungsdaten werden in einem separaten Data Dictionary erfasst.
5. Kontinuierlich werden Datenexporte der Forschungsprimärdaten generiert und in einem ausgewählten Datenformat (z.B. SPSS) zentral für autorisierte Anwender zum Download bereitgestellt.
6. Die Toolbox ist in der Lage, individuelle Gerätedaten in den Forschungsdatenbestand zu integrieren.
7. Zusätzliche Funktionalitäten, wie Module zur Sicherung der Datenqualität, können im Bedarfsfall in die Toolbox integriert werden.

Literatur: [1] Bialke M*, Bahls T, Havemann C, Piegsa J, Weitmann K, Wegner T, et al. MOSAIC. A modular approach to data management in epidemiological studies. METHODS OF INFORMATION IN MEDICINE. 2015; 54(4):364-371. doi: 10.3414/ME14-01-0133
[2] Thamm OC*, Perbix W, Kricheldorf J, Lefering R, Neugebauer EAM, Hartmann B, et al. Etablierung eines nationalen Verbrennungsregisters (Abstract). Deutsche Gesellschaft für Chirurgie. 131. Kongress der Deutschen Gesellschaft für Chirurgie. Berlin, 25.-28.03.2014., Düsseldorf, 2014.

Methoden

Innerhalb der implementierten Toolbox werden ausschließlich pseudonymisierte Forschungsdaten gespeichert. Die Pseudonymisierung erfolgt dabei vollkommen unabhängig vom Datenerhebungswerkzeug und macht die Angabe einer Einrichtung-spezifischen Patienten-ID und Fallnummer je Datensatz erforderlich. Nur die eingebende Einrichtung selbst, beispielsweise ein Krankenhaus, hat Kenntnis von den IDAT des Patienten. Eine darüber hinausgehende Trennung von IDAT und MDAT ist nicht erforderlich. Innerhalb der Toolbox wird die zentrale Pseudonymisierung mit Hilfe des *Pseudonymisierungsdienstes gPAS* [1] realisiert. Je Patient und Behandlungsfall werden konfigurierbare standort- und gerätespezifische Pseudonyme sowie ein übergeordnetes Projekt-Pseudonym (z. B. eine eindeutige Register-ID) erstellt und an zentraler Stelle verwaltet. Diese

Pseudonyme werden automatisch beim Anlegen des Patienten erstellt und gestatten keinen Rückschluss auf die im klinischen Kontext eindeutige Patienten-ID und Fallnummer. Gleichzeitig hat autorisiertes Personal die Möglichkeit, im Bedarfsfall,

z. B. zur Wiederkontaktierung, am Studienstandort eine Depseudonymisierung durchzuführen. Anhand der so ermittelten Patienten-ID und Fallnummer kann ein Patient im klinischen Kontext identifiziert werden. Dabei vereinfacht ein in die Toolbox integriertes Software-Modul (der sogenannte Dispatcher [1]) interne Prozesse, wie das Anlegen und Suchen von Patienten per Web-Oberfläche sowie die Steuerung der zentralen Vergabe von standort- und projektspezifischen Pseudonymen.

Ergebnisse und Diskussion

Als Pilotierung wurde die Toolbox inklusive des umgesetzten Pseudonymisierungskonzeptes für die technische Restrukturierung des nationalen Verbrennungsregisters (Leitung: Dr. Oliver C. Thamm) [2] genutzt. Auf diese Weise konnte eine zentrale und web-basierte Erfassung der Registerdaten über Standortgrenzen hinweg realisiert und gleichzeitig die strikte Trennung von IDAT und MDAT umgesetzt werden.

Die Toolbox gestattet derzeit ausschließlich die Dokumentation von Patientenfällen. Die eindeutige Identifikation von Patienten (Dublettenprüfung) ist aufgrund des nicht Vorhandenseins von IDAT innerhalb der Toolbox nicht möglich. Dennoch ist die

Nachverfolgung bzw. Depseudonymisierung für berechtigte Anwender am Studienstandort durchführbar, um z. B. eine Wiederkontaktierung von gespeicherten Teilnehmern durchzuführen. Aufgrund des modularen Aufbaus der Toolbox können zukünftig weitere Werkzeuge oder Arten von Gerätedaten integriert werden. Einzelne Komponenten sind im Bedarfsfall austauschbar.

Die Toolbox wird seit Frühjahr 2016 im Nationalen Verbrennungsregister eingesetzt. In rund 12 Monaten wurden 5.000 Fälle in mehr als 50 Standorten dokumentiert. Die Toolbox wird kontinuierlich weiterentwickelt und steht in der Version 1.2.3 über mosaic-greifswald.de zur Verfügung.



Dieses Werkzeug entstand im Rahmen des DFG-geförderten Projektes MOSAIC

DFG Deutsche
Forschungsgemeinschaft
Fördernummer: HO 1937/2-1

Kontakt

Henriette Rau, M. Sc. M. Sc.
Institut für Community Medicine
Abteilung Versorgungsepidemiologie und Community Health
Universitätsmedizin Greifswald
Ellernholzstrasse 1-2, 17487 Greifswald, Deutschland
Tel.: +49 3834 / 86-7774, Fax: +49 3834 / 86-7752
E-Mail: henriette.rau@uni-greifswald.de