

# Integration von Bestandsdaten aus Kohorten- und Registerstudien in ein existierendes Forschungsnetzwerk: Nationales Pandemie Kohorten Netz (NAPKON)

## Integration of Inventory Data from Cohort and Registry Studies into an Existing Research Network: National Pandemic Cohort Network (NAPKON)

Autorinnen/Autoren

Anna-Lena Hofmann<sup>1</sup> , Jörg Janne Vehreschild<sup>3, 4, 5</sup>, Martin Witzenrath<sup>6, 7</sup>, Wolfgang Hoffmann<sup>8</sup>, Thomas Illig<sup>9</sup>, Stefan Schreiber<sup>10</sup>, Gabriele Anton<sup>11, 12</sup>, Johannes Christian Hellmuth<sup>13, 14</sup>, Maximilian Muenchhoff<sup>14, 15</sup>, Clemens Scherer<sup>14, 16</sup>, Christina Pley<sup>17</sup>, Charlotte Thibeault<sup>6</sup>, Florian Kurth<sup>18</sup>, Sarah Berger<sup>19</sup>, Michael Hummel<sup>20</sup>, Sina Marie Hopff<sup>21</sup>, Melanie Stecher<sup>22, 23</sup>, Katharina Appel<sup>24</sup>, Dana Stahl<sup>25, 26</sup>, Monika Kraus<sup>11, 27</sup>, Bettina Lorenz-Depiereux<sup>27, 28</sup>, Sabine Hanß<sup>29, 30</sup>, Sebastian von Kielmansegg<sup>31</sup>, Irene Schlünder<sup>32</sup>, Anna Niemeier<sup>33</sup>, Peter Heuschmann<sup>2, 34, 35</sup>, Michael Krawczak<sup>36</sup>, Jens-Peter Reese<sup>2, 34</sup>

### Institute

- 1 Universität Würzburg, Institut für klinische Epidemiologie und Biometrie, Würzburg, Germany
- 2 Universitätsklinikum Würzburg, Institut für medizinische Datenwissenschaften, Würzburg, Germany
- 3 Goethe-Universität Frankfurt am Main, Abteilung für Innere Medizin 2, Hämatologie/Onkologie, Frankfurt am Main, Germany
- 4 Universität zu Köln, Medizinische Fakultät und Universitätsklinikum Köln, Germany
- 5 Deutsches Zentrum für Infektionsforschung (DZIF), Partnerstandort Bonn-Köln, Köln, Germany
- 6 Charité Universitätsmedizin Berlin, Klinik für Pneumologie, Beatmungsmedizin und Intensivmedizin mit dem Arbeitsbereich Schlafmedizin, Berlin, Germany
- 7 Deutsches Zentrum für Lungenforschung (DZL), Standort Berlin, Berlin, Germany
- 8 Universität Greifswald, Institut für Community Medicine, Greifswald, Germany
- 9 Medizinische Hochschule Hannover, Hannover Unified Biobank, Hannover, Germany
- 10 Universitätsklinikum Schleswig-Holstein, Klinik Innere Medizin I, Kiel, Germany
- 11 Helmholtz Zentrum München Institute of Epidemiology, Institut für Epidemiologie, Neuherberg, Germany
- 12 Deutsches Zentrum für Infektionsforschung eV Standort München, Standort München, Germany
- 13 Universitätsklinikum LMU München, Medizinische Klinik III, München, Germany
- 14 Universitätsklinikum, LMU München, COVID-19 Register der LMU München (CORKUM), München, Germany
- 15 Nationales Referenzzentrum für Retroviren, LMU München, Max von Pettenkofer Institut & Genzentrum, Virologie, München, Germany
- 16 Universitätsklinikum, LMU München, Medizinische Klinik und Poliklinik I, München, Germany
- 17 Charité Universitätsmedizin Berlin, Clinical Trial Office, Berlin, Germany
- 18 Charité Universitätsmedizin Berlin, Klinik für Pneumologie, Beatmungsmedizin und Intensivmedizin mit dem Arbeitsbereich Schlafmedizin, Berlin, Germany
- 19 Charité Universitätsmedizin Berlin, Department of Infectious Diseases, Respiratory Medicine and Critical Care, Berlin, Germany
- 20 Charité Universitätsmedizin Berlin, Zentrale Biobank (ZeBanC), Berlin, Germany
- 21 Universität zu Köln, Medizinische Fakultät und Universitätsklinikum Köln, Abteilung I für Innere Medizin, Zentrum für Integrierte Onkologie Aachen Bonn Köln Düsseldorf, Köln, Germany
- 22 Uniklinik Köln Klinik I für Innere Medizin, Onkologie, Hämatologie, Klinische Infektiologie, Klinische Immunologie, Hämostaseologie, Internistische Intensivmedizin, Köln, Germany
- 23 Deutsches Zentrum für Infektionsforschung (DZIF), Standort Köln-Bonn, Köln, Germany
- 24 Goethe-Universität Frankfurt, Frankfurt am Main, Abteilung für Innere Medizin 2, Hämatologie/Onkologie, Frankfurt, Germany
- 25 Universitätsmedizin Greifswald, Unabhängige Treuhandstelle, Greifswald, Germany
- 26 DZHK e.V. (Deutsches Zentrum für Herz-Kreislauf-Forschung), DZHK e.V. (Deutsches Zentrum für Herz-Kreislauf-Forschung), Greifswald, Germany
- 27 DZHK e.V. (Deutsches Zentrum für Herz-Kreislauf-Forschung e.V.), Standort München, München, Germany
- 28 Institut für Epidemiologie, Helmholtz Zentrum München, München, Germany

- 29 Universitätsmedizin Göttingen, Institut für Medizinische Informatik, Göttingen, Germany
- 30 DZHK e.V. (Deutsches Zentrum für Herz-Kreislauf-Forschung e.V.), Standort Göttingen, Göttingen, Germany
- 31 Christian-Albrechts-Universität zu Kiel, Institut für Öffentliches Wirtschaftsrecht, Rechtswissenschaftliche Fakultät, Kiel, Germany
- 32 TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V, Berlin, Berlin, Germany
- 33 TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., Berlin, Berlin, Germany
- 34 Julius-Maximilians-Universität Würzburg, Institut für Klinische Epidemiologie und Biometrie, Würzburg, Germany
- 35 Universitätsklinikum Würzburg, Zentrale für Klinische Studien Würzburg, Würzburg, Germany
- 36 Christian-Albrechts-Universität zu Kiel, Universitätsklinikum Schleswig-Holstein Campus Kiel, Institut für Medizinische Informatik und Statistik, Kiel, Germany

### Schlüsselwörter

Datenintegration, Bestandsdaten, Covid-Pandemie

### Keywords

Data integration, inventory data, covid pandemic

Artikel online veröffentlicht 22.08.2024

### Bibliografie

Gesundheitswesen 2024; 86: 647–654

DOI 10.1055/a-2346-9680

ISSN 0941-3790

© 2024. Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14,  
70469 Stuttgart, Germany

### Korrespondenzadresse

Anna-Lena Hofmann  
Institut für klinische Epidemiologie und Biometrie  
Universität Würzburg  
Josef-Schneider-Straße 2  
97080 Würzburg  
Germany  
Hofmann\_A9@ukw.de

### ZUSAMMENFASSUNG

In der frühen Phase der COVID-19-Pandemie wurden in Deutschland viele lokale Sammlungen klinischer Daten mit SARS-CoV-2 infizierter Patient:innen initiiert. Im Rahmen des Nationalen Pandemie-Kohorten-Netzes (NAPKON) des Netzwerkes Universitätsmedizin wurde der „Integrationskern“ etabliert, um die rechtlichen, technischen und organisatorischen Voraussetzungen für eine Integration von Bestandsdaten in laufende prospektive Datensammlungen zu konzipieren und die Machbarkeit der entwickelten Lösungen mittels Use Cases (UCs) zu prüfen. De-

taillierte Studienunterlagen der Datensammlungen wurden eingeholt. Nach strukturierter Dokumentenanalyse, bewertete ein Review Board, gemäß definierter Kriterien die Integrierbarkeit der Daten in NAPKON. Von 30 kontaktierten Universitätskliniken hatten 20 auf die Anfrage geantwortet. Die Patient:inneninformationen und Einwilligungen zeigten ein heterogenes Bild bezüglich der pseudonymen Weitergabe der Daten an Dritte und des Re-Kontakts. Ein Großteil der Datensammlungen (n = 13) erfüllte die Kriterien für eine Integration in NAPKON, bei vier Studien wären Anpassungen der regulatorischen Dokumente erforderlich. Drei Kohorten waren nicht für einen Einschluss in NAPKON geeignet. Die rechtlichen Rahmenbedingungen einer retrospektiven Datenintegration und einer einwilligungsfreien Datennutzung über Forschungsklauseln (§27 BDSG) wurde durch ein Rechtsgutachten der TMF – Technologie- und Methodenplattform, für die vernetzte medizinische Forschung e. V., Berlin erarbeitet. Anhand zweier vom NAPKON-Lenkungsausschuss ausgewählter UCs (CORKUM, LMU München; Pa-COVID-19, Charité- Universitätsmedizin Berlin) wurde die Machbarkeit einer Datenintegration bis Ende 2021 in NAPKON gezeigt. Es erfolgte gemäß den Vorgaben die Qualitätssicherung und die aufwandsgenaue Abrechnung der übertragenen Fälle. Basierend auf den Ergebnissen können Empfehlungen für verschiedene Kontexte formuliert werden, um technisch-operative Voraussetzungen wie Interoperabilität, Schnittstellen und Datenmodelle für die Datenintegration zu schaffen sowie regulatorische Anforderungen an Ethik, Datenschutz, ärztliche Schweigepflichtsentbindung und den Datenzugang bei der Integration bestehender Kohortendaten zu erfüllen. Die mögliche Integration von Daten in Forschungsnetzwerke und deren Sekundärnutzung sollte bereits in der Planungsphase einer Studie – insbesondere beim Informed Consent – berücksichtigt werden, um den größtmöglichen Nutzen aus den erhobenen Daten zu ziehen.

### ABSTRACT

In the early phase of the COVID-19 pandemic, many local collections of clinical data on patients infected with SARS-CoV-2 were initiated in Germany. As part of the National Pandemic Cohort Network (NAPKON) of the University Medicine Network, the "Integration Core" was established to design the legal, technical and organisational requirements for the integration of inventory data into ongoing prospective data collections and to test the feasibility of the newly developed solutions using use cases (UCs). Detailed study documents of the data collections were obtained. After structured document analysis, a review board evaluated the integrability of the data in NAPKON according to defined criteria. Of 30 university hospitals contacted, 20 responded to the request. Patient information and consent showed a heterogeneous picture with regard to the pseudonymised transfer of data to third parties and re-contact. The majority of the data collections (n = 13) met the criteria for integration into NAPKON; four studies would require adjustments to the regulatory documents. Three cohorts were not suitable for inclusion in NAPKON. The legal framework for retrospective data integration and consent-free data use via research clauses (§27 BDSG) was elaborated by a legal opinion by TMF – Tech-

nology, Methods and Infrastructure for Networked Medical Research, Berlin. Two UCs selected by the NAPKON steering committee (CORKUM, LMU Munich; Pa-COVID-19, Charité-Universitätsmedizin Berlin) were used to demonstrate the feasibility of data integration in NAPKON by the end of 2021. Quality assurance and performance-based reimbursement of the cases were carried out according to the specifications. Based on the results, recommendations can be formulated for various contexts in order to create technical-operational prerequisites

such as interoperability, interfaces and data models for data integration and to fulfil regulatory requirements on ethics, data protection, medical confidentiality and data access when integrating existing cohort data. The possible integration of data into research networks and their secondary use should be taken into account as early as the planning phase of a study – particularly with regard to informed consent – in order to maximise the benefits of the data collected.

## Hintergrund und Zielsetzung

Zu Beginn der COVID-19-Pandemie wurden an vielen deutschen Universitätskliniken zeitnah Sammlungen von klinischen Daten, Bilddaten und Bioproben von COVID-19-Patient:innen gestartet [1–3]. Eine Harmonisierung fand aus unterschiedlichen Gründen (z. B. fehlender zentraler Koordination oder Finanzierung) nur begrenzt statt. Im April 2020 erfolgte die Gründung des Netzwerks Universitätsmedizin (NUM), das vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wird (Förderkennzeichen 01KX2021) und in dem alle 36 Universitätsklinika bei der Bekämpfung der COVID-19-Pandemie zusammenarbeiteten [4]. Im Rahmen von NUM wurden in der ersten Antragsphase 13 standortübergreifende Forschungsprojekte initiiert, darunter das Nationale Pandemie Kohorten Netz (NAPKON). Im Rahmen von NAPKON wurden Strukturen für die harmonisierte Erfassung qualitativ hochwertiger Daten und Bioproben aufgebaut [5–6].

Für ein effektives Management der COVID-19-Pandemie war eine zeitnahe und standardisierte Datenbasis zur Analyse der Ausbreitung, des Schweregrades und der Behandlung von SARS-CoV-2-Infizierten von großer Bedeutung. Die Integration bestehender Kohortenstudien bot die Möglichkeit, Prozesse zur Qualitätssicherung und regulatorischen Überprüfung zu etablieren, die zu einer verbesserten Datenverfügbarkeit sowie verbesserter Vergleichsgruppenanalysen beitragen könnten.

Für viele in der Frühphase einer Pandemie aufgebauten, proprietären Sammlungen erschien es daher wissenschaftlich sinnvoll, die jeweiligen Daten und Bioproben standardisiert in größere Sammlungen wie z. B. NAPKON zu integrieren und so einer breiteren wissenschaftlichen Community verfügbar zu machen. Neben Unterschieden hinsichtlich Datenstruktur, Studiendesign und inhaltlicher Schwerpunktsetzung ergeben sich hierbei jedoch auch erhebliche regulatorische Herausforderungen bezüglich der Möglichkeit der pseudonymisierten Datenweitergabe und der Re-Kontaktierung von Patient:innen, der Vergleichbarkeit der Patient:inneninformationen und -einwilligungen sowie der Vorkehrungen zum Datenschutz und zur Wahrung der ärztlichen Schweigepflicht.

Um auch für spätere Pandemien geeignete Lösungsansätze zur Integration von Bestandsdatensammlungen zu erarbeiten, wurde im Rahmen von NAPKON der sogenannte „Integrationskern“ etabliert. Dessen Aufgabe war neben der Recherche geeigneter Daten- und Bioprobensammlungen zu deutschen COVID-19-Patient:innen – außerhalb von NAPKON – auch die Identifizierung möglicher ethischer und rechtlicher Hindernisse für deren Integration in prospektive Datensammlungen.

Im Folgenden werden das Vorgehen im NAPKON-Integrationskern, die erzielten Ergebnisse und die daraus entwickelten Lösungsansätze beschrieben. Die praktische Machbarkeit der Integration von Bestandsdaten und -proben in ein existierendes Forschungsnetz konnte im Fall von NAPKON anhand zweier exemplarischer Use Cases gezeigt werden.

## Methodik

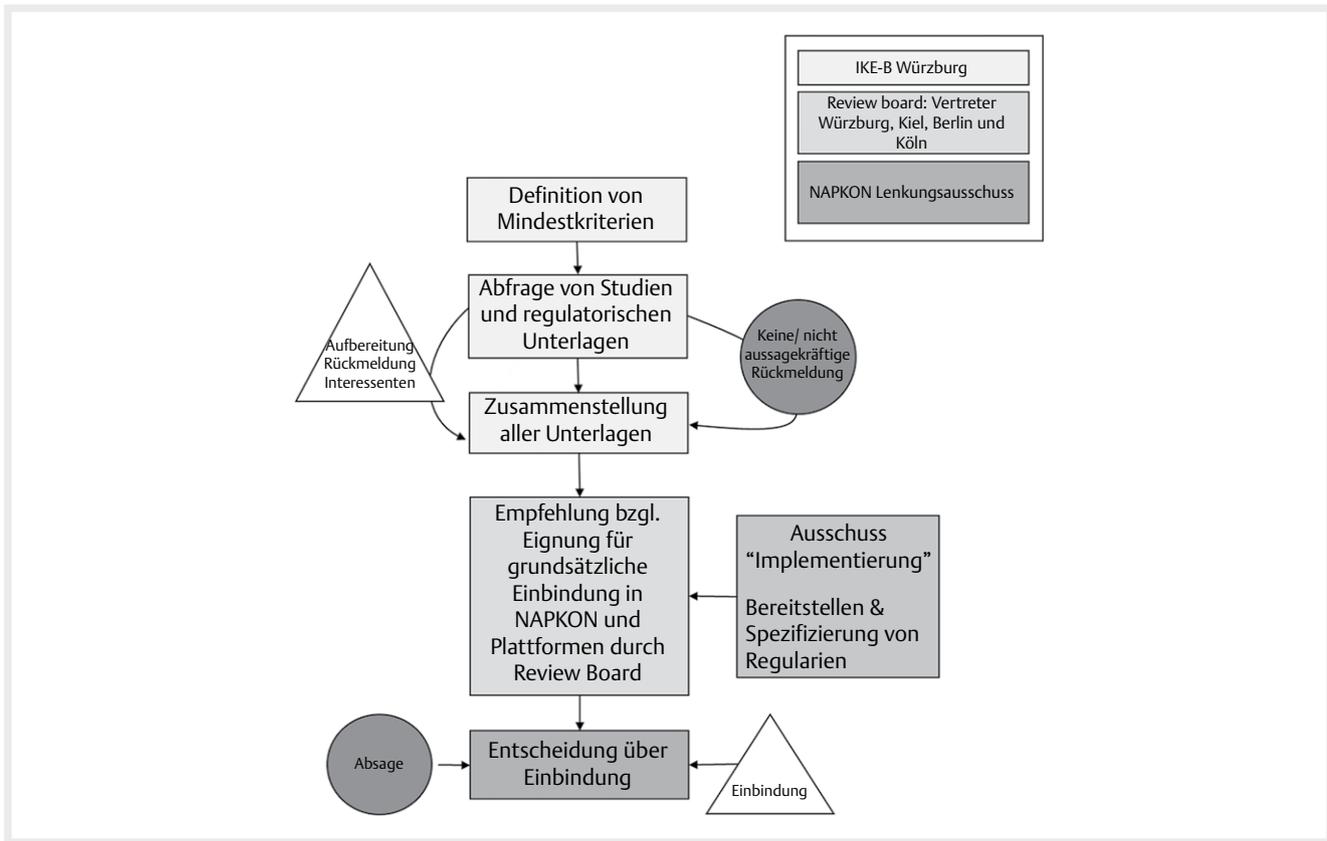
### Organisatorischer Rahmen

Das **Netzwerk Universitätsmedizin (NUM)** wurde im April 2020 gegründet. Ziel des NUM ist es, die COVID-19-Forschung der 36 Universitätskliniken in Deutschland zu koordinieren und innovative Strukturen und Prozesse für ein besseres Verständnis des Virus und der hervorgerufenen Erkrankung zu erlangen.

Das **Nationale Pandemie Kohorten Netz (NAPKON)** ist ein Teilprojekt des NUM mit dem langfristigen Ziel, ein bundesweites Netz für die Planung und Durchführung klinischer Studien zu COVID-19 aufzubauen. Derzeit umfasst NAPKON drei komplementäre prospektive Kohortenstudien („Plattformen“) [5]: die sektorenübergreifende Plattform (SÜP), die hochauflösende Plattform (HAP) und die populationsbasierte Plattform (POP) [6].

- Die SÜP erfasst soziodemografische und klinische Daten von SARS-CoV-2-infizierten Patient:innen aller Altersgruppen, die stationär oder ambulant behandelt werden.
- Die HAP fokussiert sich auf stationäre SARS-CoV-2-positive Patient:innen im Alter  $\geq 18$  Jahre, mit einem schweren Krankheitsverlauf von COVID-19, welcher eine intensivpflichtige Behandlung erfordert.
- Die POP konzentriert sich auf die Untersuchung der gesundheitlichen (Langzeit-)Folgen einer SARS-CoV-2-Infektion in der erwachsenen Allgemeinbevölkerung [6].

Zur Unterstützung wurden den drei Kohortenstudien vier übergreifende „Infrastrukturkerne“ (Interaktionskern, Epidemiologiekern, Bioprobekern, Integrationskern) mit jeweils unterschiedlichen Aufgaben und Kernkompetenzen zur Seite gestellt. Der Integrationskern wurde bis Ende 2021 gefördert, um Prozesse für die grundsätzliche regulatorische Prüfung der schnellen Aufnahme von Bestandsdaten in prospektive Datensammlungen wie NAPKON zu konzipieren sowie die Machbarkeit einer Integration anhand definierter Use Cases zu überprüfen.



► **Abb. 1** Vorgehen des Integrationskerns.

## Methodisches Vorgehen

### Erfassung geeigneter Datensammlungen

Die Prozesse einer Integration von Bestandsdaten in laufende Kohortenstudien wurden von dem NAPKON-Integrationskern schrittweise entwickelt (► **Abb. 1**). Am Anfang stand hierbei die Definition allgemeiner und plattformspezifischer Mindestkriterien für die Einbringung in NAPKON (s. Ergebnisse).

Universitätskliniken, die ein grundsätzliches Interesse an einer Mitarbeit und einer möglichen Integration Ihrer Bestandsdaten in NAPKON hatten, wurden kontaktiert und deren regulatorischen und Studienunterlagen abgefragt. Hierzu gehörten die verwendete Patient:inneninformation und -einwilligung, das Ethikvotum, das Studienprotokoll mit Visitenplan und soweit vorhanden, ein Data Dictionary, die Use & Access Regularien sowie eine datenschutzrechtliche Stellungnahme. Diese Unterlagen wurden strukturiert erfasst und bewertet.

### Beurteilung der Integrationsmöglichkeit

Für die Einschätzung der Integrierbarkeit in eine der drei NAPKON-Plattformen wurde ein Review Board aus fünf Experten gebildet. Die Einschätzungen wurden in einem Konsensmeeting des gesamten Review Boards diskutiert und die Integrierbarkeit jeder einzelnen Datensammlung gemäß einer der folgenden drei Kategorien eingestuft: (1) Eine Integration ist möglich, alle allgemeinen und plattformspezifischen Mindestkriterien sind erfüllt. (2) Eine Integration ist möglich, die Erfüllung einiger Mindestkriterien ist abzu-

klären. (3) Eine Integration ist derzeit nicht möglich, einige Mindestkriterien sind nicht erfüllt.

Der NAPKON-Implementierungsausschuss wurde etabliert, bestehend aus Mitgliedern der NAPKON-Plattformen und -Methodenkernen sowie der Datentreuhandstelle und weiteren Vertreter:innen des NUM. Aufgabe des Implementierungsausschusses war das Abwägen der Rahmenbedingungen einer Bestandsdatenintegration, verbunden mit der Spezifizierung von Anforderungen an Nutzungsverträge, Vergütung und Eigentumsverhältnisse. Außerdem wurde die Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V. (TMF) mit einem Rechtsgutachten beauftragt, um die rechtlichen Aspekte einer Integration von Bestandsdaten in NAPKON herauszuarbeiten. Ein besonderes Augenmerk lag dabei auf der Möglichkeit eines einwilligungsunabhängigen Vorgehens [7].

### Überprüfung der Machbarkeit an geeigneten Use Cases

Zum Nachweis der entwickelten Verfahren wurde die Integration in NAPKON an zwei Use Cases exemplarisch umgesetzt. Die Auswahl der Use Cases erfolgte durch den NAPKON- Lenkungs-ausschuss primär aufgrund der Kompatibilität der jeweiligen Datenstruktur mit den jeweiligen NAPKON-Plattformen SÜP und HAP. Im Austausch mit den Use Cases mit Unterstützung einer Anwaltskanzlei die jeweils passenden Einzelverträge und Konzepte für die Aufwandsentschädigung, technischer Übertragung der Daten in das elektronische Datenerfassungssystem von NAPKON (Secu-Trial) und zur Sicherstellung der Datenqualität ausgearbeitet.

► **Tab. 1** Allgemeine und plattformspezifische Mindestkriterien zur Integration von Bestandsdaten in NAPKON.

Allgemeine Mindestkriterien	Plattformspezifische Mindestkriterien
<ul style="list-style-type: none"> <li>▪ Positives Ethikvotum</li> <li>▪ Einverständniserklärung mit der Möglichkeit zur pseudonymen Weitergabe klinischer Daten</li> <li>▪ Re-Kontaktierung der Patient:innen</li> <li>▪ Vorhandensein von Bioproben</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>SÜP:</b> keine Erstvorstellungen nach Genesung, Erlaubnis einer Weitergabe von Daten Verstorbener Patient:innen</li> <li>▪ <b>HAP:</b> Einschluss in der Akutphase, hochauflösende Datenerhebung zu Beginn mindestens zweimal pro Woche, Gewährleistung einer studienspezifischen Nachverfolgung</li> <li>▪ <b>POP:</b> populationsrepräsentative Sammlung mit mindestens 2000 Proband:innen in Einzugsbereich, Nachuntersuchung 6–9 Monate nach Diagnose</li> </ul>

## Ergebnisse

### Erfassung geeigneter Datensammlungen

Die durchgeführte Abfrage ergab bei 30 der 33 kontaktierten Universitätskliniken eine grundsätzliche Bereitschaft zur Mitarbeit.

Für die Prüfung der regulatorischen Anforderungen wurden Mindestkriterien definiert, die sich in allgemeine und plattformspezifische Kriterien unterteilen (► **Tab. 1**). Zu den allgemeinen Mindestkriterien gehören das positive Ethikvotum, die Erlaubnis der Patient:innen, (a) ihre klinischen Daten in pseudonymisierter Form an externe Forscher:innen weiterzugeben und (b) sie ggf. für Follow-up-Untersuchungen oder eine Erweiterung ihrer Einwilligung erneut zu kontaktieren. Die plattformspezifischen Mindestkriterien wurden in Zusammenarbeit mit den Koordinatoren von SÜP, HAP und POP erarbeitet.

### Beurteilung der Integrationsmöglichkeit

Die 30 Universitätskliniken wurden kontaktiert, die Interesse an einer Mitarbeit zeigten. Davon übermittelten 20 die angefragten Studienunterlagen. Die Unterlagen wurden anhand der Mindestkriterien aufbereitet.

Hierbei zeigten sich Unterschiede zwischen den Standorten. Die Aufklärungs- und Einwilligungsdokumenten führten nicht durchgängig die COVID-19-bezogene Forschung als möglichen Zweck der späteren Datenverarbeitung auf. Einige Kohortenstudien starteten bereits vor der Pandemie und fügten ein Ergänzungsmodul zur COVID-19-Forschung ein. Die Weitergabe klinischer Daten an externe Forscher:innen war häufig unpräzise beschrieben. Zwei Standorte hatten eine solche Weitergabe nur in anonymisierter Form vorgesehen. Da NAPKON auch internationale Forschung unterstützen soll, setzte eine Integrierbarkeit von Daten die explizite Möglichkeit zur Übermittlung an Drittstaaten voraus. Die Erlaubnis zur Re-Kontaktierung war in vielen Fällen (n = 14) als Option enthalten;

Zwischen den Studienprotokollen ergab sich eine Reihe formaler und inhaltlicher Unterschiede. Für zwei Standorte gab es keine spezifischen Protokolle. Bei den Übrigen variierte der Umfang. Data Dictionaries oder Auszüge der CRFs wurden von 16 Standorten bereitgestellt. Die Visitenpläne unterschieden sich hinsichtlich der durchgeführten Untersuchungen ebenso wie in den jeweiligen Visiten-Zeitpunkten.

Zusammenfassend ergab die Prüfung durch das Review Board, dass mit 13 von 20 Datensammlungen ein Großteil der Standorte die definierten Mindestkriterien erfüllte. Vier weitere Sammlungen waren nur eingeschränkt für die Datenintegration geeignet. Entweder war eine Weitergabe der Daten ausschließlich in anonymi-

zierter Form möglich oder es bestand keine Möglichkeit zum Re-Kontakt. Drei Kohorten eigneten sich aufgrund ihres Studienaufbaus oder Durchführung nicht für eine Integration in NAPKON. Es wurden z. B. keine Bioproben entnommen oder das Einzugsgebiet war nicht groß genug.

Aus dem TMF-Rechtsgutachten ergab sich, dass eine unabhängige Datentreuhandstelle immer dann explizit in den Dokumenten zu erwähnen ist, wenn eine solche Einrichtung in das Datenmanagement involviert ist und personenidentifizierbare Daten erhält und speichert. Mit Ausnahme zweier Standorte wurde zwar auf die Weitergabe der Daten an Forscher:innen eingegangen, der explizite Verweis auf eine Weitergabe personenbezogener Daten an eine Datentreuhandstelle fehlte jedoch. Zudem sei laut Gutachten eine Verschwiegenheitsentbindung notwendig. Aus juristischer Sicht handelt es sich hierbei nicht um ein datenschutzrechtliches Erfordernis, vielmehr um die Verletzung der ärztlichen Schweigepflicht im strafrechtlichen Sinne (§ 203 StGB; § 9 MBO-Ä). Bereits die Information einer Behandlung gilt als geschütztes Patient:innengeheimnis.

Als weitere mögliche Rechtsgrundlage werden die Forschungsklauseln nach § 27 BDSG genannt. Diese setzen voraus, dass das öffentliche Interesse an der Forschungsnutzung die Risiken für die Betroffenen überwiegt. Dies erfordert umfassende Maßnahmen des organisatorischen und technischen Datenschutzes. Die einwilligungsfreie Einbringung in NAPKON wäre nur dann zulässig, wenn nicht von bereits eingeholten Einwilligungen abgewichen würde, insbesondere mit Blick auf die Rolle einer unabhängigen Datentreuhandstelle. Angesichts der pandemischen Lage ließe sich das Betreten rechtlichen Neulandes mit dem großen gesellschaftlichen Interesse an einer verbesserten Versorgung der von COVID-19 Betroffenen begründen.

### Umsetzung an geeigneten Use Cases

Unter Berücksichtigung des TMF-Rechtsgutachtens wurden Verfahrensweisen zur Datenintegration in prospektive Kohortenstudien entwickelt. Diese Vorgaben wurden an zwei Use Cases erprobt. Diese wurden wegen ihrer Kompatibilität zu SÜP und HAP ausgewählt. Im Fall der SÜP handelte es sich um die CORKUM Studie der Ludwig-Maximilians-Universität München (LMU), bei der HAP um die Pa-COVID-19 Studie der Charité Berlin [8–9].

CORKUM ist ein COVID-19-Register des LMU Klinikums, in dem klinische Daten und Biomaterialien zur Erforschung der Immunantwort und des Krankheitsverlaufs gesammelt werden. Bis Ende 2021 konnten exemplarisch 120 Fälle aus CORKUM in SÜP integriert werden [9]. Die longitudinale Registerstudie Pa-COVID-19 ist auf die klinische und molekulare Phänotypisierung von COVID-19 Patient:innen fokussiert mit dem Ziel, frühzeitig Risikofaktoren zu

erkennen und eine Prognose zum Krankheitsverlauf abzugeben. Aus Pa-COVID-19 konnten bis Ende 2021 exemplarisch 42 Fälle in HAP integriert werden [9].

Die klinischen Bestandsdaten wurden in die CRFs der jeweiligen NAPKON-Plattform übertragen und deren Vollständigkeit und Qualität auf gleiche Weise geprüft, wie bei prospektiv eingegebenen Daten. Es kamen die NAPKON entwickelten automatisierten Verfahren für Qualitätssicherung und Abrechnung zur Anwendung. Die geschlossenen Einzelverträge regeln die Kostenvergütung der Studiendurchführung und -dokumentation und berücksichtigen die lokalen Use and Access Prozesse, wie Mitsprache-, Veto- oder Publikationsrechte.

Der Datentransfer wurde mit den Standorten manuell oder mit einer programmierten Export-Pipeline umgesetzt. In Abstimmung mit der Datentreuhandstelle Greifswald wurde für den Transfer eine Lösung unter Wahrung der ärztlichen Schweigepflicht entwickelt. Personenidentifizierbare Daten und vergebene Pseudonyme wurden von den Datenhaltern selbst verwaltet. Diese waren auch weiterhin für die lokale Durchsetzung von Widerrufen oder Anträgen auf Datenlöschung verantwortlich. Die Datentreuhandstelle erhielt nur die lokalen Pseudonyme und ordnete diesen ein NAPKON-spezifisches Pseudonym zu. In den Plattformen wurden die retrospektiven klinischen Daten gekennzeichnet.

## Diskussion

Schon zu Beginn der COVID-19 Pandemie war ein national und international erklärtes Ziel der Wissenschaft, durch die zeitnahe Analyse empirischer Daten einen Beitrag zur Eindämmung des Infektionsgeschehen zu leisten. In Deutschland sollte möglichst schnell mit dem Aufbau großer Datensammlungen an den Universitätsklinikum begonnen werden [10]. Deren Verfügbarkeit würde Wissenschaftler:innen erlauben, innerhalb kürzester Zeit Antworten auf klinische und epidemiologische Fragen zur Pandemie abzuleiten und in Empfehlungen für den weiteren Umgang mit SARS-CoV-2 zu überführen.

Wegen des erheblichen Zeitdrucks kam es zu vielen nicht-harmonisierten Aktivitäten sowie einer erheblichen Variabilität hinsichtlich Art, Umfang und Qualität der lokalen Datenerhebung. Viele wertvolle Datenbestände standen zunächst nicht für eine koordinierte wissenschaftliche Nutzung zur Verfügung. Deshalb wurde der NAPKON-Integrationskern ins Leben gerufen, um grundsätzliche Konzepte für die Einbringung von Datenbeständen in laufende Datenerhebungen zu entwickeln. Das erarbeitete Vorgehen zur institutionsübergreifenden Datenintegration sollte als strategische Blaupause für die Datenbereitstellung z. B. in zukünftigen Pandemiesituationen dienen.

Die vorgestellten Arbeitsergebnisse offenbarten hinsichtlich Datenstruktur, Rechtsrahmen und Governance eine erhebliche Heterogenität der geprüften Datensammlungen. Obwohl diese Unterschiede eine hohe Hürde für eine Integration in NAPKON darzustellen schienen, konnte die Praktikabilität entwickelten Konzepts zur Bestandsdatenintegration anhand zweier Use Cases gezeigt werden. Eine derartige Überführung von Forschungsdaten in ein bereits existierendes Netzwerk von Standorten mit unterschiedlicher Ausgangslage hinsichtlich z. B. Studienprotokoll und Aufklärung/Einwilligung hat es in diesem Umfang nach unserem Wissensstand

in Deutschland bislang noch nicht gegeben. Die erfolgreiche Umsetzung zweier Use Cases lieferte nicht nur einen „proof-of-concept“, sie ergänzte die theoretischen Vorarbeiten um wertvolle praktische Erfahrungen im Umgang mit möglichen Hindernissen.

Die wichtigsten „*lessons learned*“ im Kontext des NAPKON-Integrationskern werden im Folgenden zusammengefasst.

- Grundsätzliche Herausforderungen ergeben sich bei der Integration von Bestandsdaten in laufende Kohortenstudien aus lokalen Unterschieden hinsichtlich Studienziel und -design, der Notwendigkeit einer studienübergreifend hohen Datenqualität und den verankerten Regeln zur Datenweitergabe an Dritte,
- Für die Bewältigung dieser Herausforderungen erwies sich die Etablierung zweier interdisziplinär Expertengremien, Review Board und Implementierungsausschuss, als äußerst hilfreich. Durch die enge Zusammenarbeit von Jurist:innen, Mediziner:innen, Ethiker:innen und Methodenwissenschaftler:innen sowie die Einbindung von Vertreter:innen der drei NAPKON-Plattformen konnten Lösungsansätze zeitnah und effizient entwickelt an die individuellen Belange einzelner Datensammlungen angepasst werden.
- Wegen der Unterschiede hinsichtlich Studiendesign und -inhalt erwies sich für jede potenzielle Integration in NAPKON ein vorheriges „Mapping“ von Bestandsdaten (d. h. die Verknüpfung von Datenpunkten aus unterschiedlichen Systemen) als unumgänglich. Da dabei in der Regel keine vollständige 1:1 Beziehung herstellbar ist, verdient die Datenübertragung besondere Aufmerksamkeit.
- Datensammlungen unterscheiden sich oft deutlich hinsichtlich der internen Qualitätskriterien. Das macht eine Qualitätsprüfung nach den Qualitätskriterien und -prozessen der „Aufnehmenden“ Studie unumgänglich. Unterschiede müssen dokumentiert und späteren Nutzer:innen zugänglich gemacht werden.
- Es empfiehlt sich, bei der Formulierung der lokalen Zugriffs- und Nutzungsregeln auch die Bedingungen einer potenziellen Datenweitergabe festzulegen bzw. zu ermöglichen. Das Teilen pseudonymisierter Daten wurde meistens erwähnt, während die Abgabe in Drittländer oder die Weitergabe personenbezogener Daten an eine Datentreuhandstelle nur selten Erwähnung fand. Viele unscharfe bzw. fehlende Formulierungen konnten durch Rücksprache mit den zuständigen Ethikkommissionen oder Amendments angepasst werden.

Die fehlende explizite Nennung einer Datentreuhandstelle im Consent stellt ein weit verbreitetes Problem zur Forschungsnutzung von Patient:innendaten dar. Die Aufgabe solcher Datentreuhandstellen ist die unabhängige Verwaltung personenbezogener Daten zum Zwecke der Pseudonymisierung oder Anonymisierung medizinischer Daten vor ihrer Weitergabe an Dritte [11]. Datentreuhandstellen ermöglichen auch die Re-Kontaktierung von Patient:innen. Eine Möglichkeit wäre die Verwendung des sogenannten „Broad Consents“ der Medizininformatik-Initiative. Dabei handelt es sich um den Mustertext einer Patient:innenaufklärung und -einwilligung in die Forschungsnutzung medizinischer Daten, dessen Inhalte von den Universitätsklinikum in Deutschland gemein-

sam erarbeitet und verabschiedet wurde. Er sieht u. a. die pseudonymisierte Datenweitergabe an Forschende und in Drittländer ebenso vor wie die optionale Zustimmung der Patient:innen zum Re-Kontakt. Durch die Erwähnung einer Datentreuhandstelle vermeidet der Text zudem das strafrechtliche Risiko einer Verletzung der ärztlichen Schweigepflicht [12].

Zusätzlich wurde die rechtliche Situation einer Datenintegration in einem Gutachten beleuchtet: In dem Rechtsgutachten wurde insbesondere die einwilligungsfreie Integration von Bestandsdaten in ein Forschungsnetz gemäß §27 BDSG thematisiert. Es böte die Möglichkeit einer Zusammenführung von Bestandsdaten in prospektive Datensammlungen, keine geeignete Zustimmung der Patient:innen zum Data Sharing eingeholt wurde. Eine Abweichung von bereits eingeholten Einwilligungen ist nicht möglich. Laut Gutachten wären bei der Berufung auf ein öffentliches Interesse eine Reihe von Einschränkungen zu beachten, wie ein umfassendes Datenschutzkonzept mit Zugriffskontrolle.

Im Verlauf der erfolgreichen Datenintegration in zwei Use Cases wurde die Grundlage geschaffen, bestehende Kohorten in Forschungsnetzwerke zu überführen, wobei bereits ein regulatorischer (Consente) und technischer Rahmen (Standardisierung, Interoperabilität) auch im Sinne der "Pandemic Preparedness" geschaffen wurde. Die genannten Herausforderungen, wie die heterogene Datenstruktur, die einwilligungsabhängige Datenfreigabe und der technische Datentransfer nach der Verknüpfung verschiedener Kohorten, führen zu Limitationen, die trotz der entwickelten rechtlichen, technischen und organisatorischen Maßnahmen nicht vollständig aufgelöst werden können. Eine Übertragung aller verfügbaren Daten ist aufgrund der Datenstruktur immer möglich. Bei der Anpassung von Daten für Integrationszwecke sind die Auswirkungen solcher Manipulationen auf den Informationsgehalt und die Datenqualität stets zu berücksichtigen. Eine sorgfältige Dokumentation aller Anpassungen muss bei einem solchen Vorgehen selbstverständlich sein.

Im internationalen Kontext wurden im Verlauf der COVID-19 Pandemie diverse nationale oder internationale Forschungsverbände oder Studien etabliert (z. B. *International Severe Acute Respiratory and emerging Infection Consortium (ISARIC)*, *SARS-Brazil*, *FrenchCOVID*) um die Bewältigung der Pandemie voranzutreiben [5, 13]. Unseres Wissens befassen sich diese Netzwerke jedoch nicht mit der Problematik der Datenintegration in vergleichbarer Weise wie der NAPKON Integrationskern. Andere Initiativen versuchen auf bestehende Datensysteme zurückzugreifen. Die Herausforderungen lassen sich mit denen des Integrationskerns vergleichen, wie die rechtlichen und sicherheitstechnische Fragen zum Datenaustausch (Governance, Datenschutz, Vertraulichkeit und Einverständniserklärung) sowie die Verfügbarkeit von digitalen Daten [14, 15]. Eine Integration von bereits vor der Pandemie existierenden Kohorten wird in anderen Forschungsprojekten nicht thematisiert. Zusammenfassend lässt sich feststellen, dass eine Integration von Daten aus laufenden Kohortenstudien mit ggf. abweichenden Zielen, Design und Voraussetzungen in ein Forschungsnetzwerk wie NAPKON grundsätzlich möglich ist. Trotz der dabei erwartbaren organisatorischen, technischen und ethisch-rechtlichen Hürden ergeben sich damit Möglichkeiten, große Datenmengen aus verschiedenen Ressourcen kurzfristig über das jeweilige Netzwerk verfügbar zu machen. In der praktischen Umset-

zung sollte ein besonderes Augenmerk auf die Verwendung von einheitlichen Datenqualitätsstandards mit Durchführung einer Qualitätssicherung, Verknüpfung der unterschiedlichen Datenmodelle mittels Mapping, Einbeziehung verschiedener Fachdisziplinen in einem Expertengremium und klare Regeln zur Datenweitergabe und Sicherstellung des Datenschutzes.

Das durch den NAPKON-Integrationskern erarbeitete Vorgehen lässt sich in weiten Teilen auf Datensammlungen zu anderen Krankheitsgebieten und ohne pandemische Situation übertragen. Idealerweise sollten Kohortenstudien allerdings künftig so geplant werden, dass eine Weitergabe der Daten an Dritte berücksichtigt und vorbereitet wird.

## Fördermittel

Bundesministerium für Bildung und Forschung (BMBF) – „NaFoUni-MedCovid19“ FKZ: 01KX2021

## Interessenkonflikt

Prof. Heuschmann berichtet über Zuschüsse vom deutschen Bundesministerium für Forschung und Bildung, während der Durchführung der Studie; weiterhin erhielt er Forschungszuschüsse des Bundesministeriums für Forschung und Bildung, Europäischen Union, Charité-Universitätsmedizin Berlin, Ärztekammer Berlin, Ärztekammer Berlin, Deutsche Parkinson Gesellschaft, Universitätsklinikum Würzburg, Robert-Koch-Institut, Deutsche Herzstiftung, Gemeinsamer Bundesausschuss Gemeinsamer Bundesausschuss (G-BA) im Innovationsfond, Deutsche Forschungsgemeinschaft Forschungsgemeinschaft, Bayerischer Staat (Ministerium für Wissenschaft und Kunst), Deutsche Krebshilfe, Charité-Universitätsmedizin Berlin (im Rahmen von Mondafis; unterstützt durch einen uneingeschränkten Forschungszuschuss von Bayer an die Charité), Universität Göttingen (im Rahmen von FIND-AF randomisiert; unterstützt durch eine Universität Göttingen von Boehringer-Ingelheim), Universitätsklinikum Heidelberg (im Rahmen von RASUNOAprime; unterstützt durch einen uneingeschränkten Forschungszuschuss an die Universität Universitätsklinikum Heidelberg von Bayer, BMS, Boehringer-Ingelheim, Daiichi Sankyo), außerhalb der eingereichten Arbeit. Dr. Scherer berichtet ein Referentenhonorar von AstraZeneca, unabhängig von dieser Arbeit. Prof. Jens-Peter Reese berichtet über Forschungszuschüsse des Bundesministeriums für Forschung und Bildung, des Bayerischen Staates (Ministerium für Wissenschaft und Kunst), des Gemeinsamen Bundesausschusses (G-BA) im Rahmen des Innovationsfonds, des Deutschen Zentrums für Lungenforschung; Honorar für Gutachten des Bundesministeriums für Gesundheit (BMG); Honorar für EBM-Fortbildungsvorlesung durch die Landesärztekammer Hessen, außerhalb der eingereichten Arbeit. Alle anderen Ko-Autor:innen geben an, keine Interessenskonflikte zu haben.

## Literatur

- [1] Eichner F, Gelbrich G, Weißbrich B et al. Seroprävalenz von COVID-19 und psychosoziale Auswirkungen in der Allgemeinbevölkerung: Ergebnisse des STAAB-COVID-One Programms. Das Gesundheitswesen 2021. DOI: 10.1055/a-1630-7601

- [2] Gaitzsch E et al. COVID-19 in Patients Receiving CD20-depleting Immunochemotherapy for B-cell Lymphoma. *HemaSphere* 2021; 5: e603. DOI: 10.1097/HS9.0000000000000603
- [3] Cornely OA et al. "Improving diagnosis of severe infections in immunocompromised patients (ISI)". Cologne, Germany: Uniklinik Köln; 2008
- [4] NUM Coordination Office, NUKLEUS Study Group Heyder R et al. Das Netzwerk Universitätsmedizin: Technisch-organisatorische Ansätze für Forschungsdatenplattformen. *Bundesgesundheitsbl* 2023; 66: 114–125. DOI: 10.1007/s00103-022-03649-1
- [5] Schons M, Pilgram L, Reese JP et al. The German National Pandemic Cohort Network (NAPKON): rationale, study design and baseline characteristics. *Eur J Epidemiol* 2022. DOI: 10.1007/s10654-022-00896-z
- [6] Horn A, Krist L, Lieb W et al. Long-term health sequelae and quality of life at least 6 months after infection with SARS-CoV-2: design and rationale of the COVIDOM-study as part of the NAPKON population-based cohort platform (POP). *Infection* 2021; 49: 1277–1287. DOI: 10.1007/s15010-021-01707-5
- [7] Lesch W, Richter G, Semler SC. Daten teilen für die Forschung: Einstellungen und Perspektiven zur Datenspende in Deutschland. In: Richter G, Loh W, Buyx A, Graf von Kielmansegg S (eds) *Datenreiche Medizin und das Problem der Einwilligung*. Springer; Berlin, Heidelberg: 2022. DOI: 10.1007/978-3-662-62987-1\_11
- [8] Bieber S, Kraechan A, Hellmuth JC et al. Left and right ventricular dysfunction in patients with COVID-19-associated myocardial injury. *Infection* 2021; 49: 491–500. DOI: 10.1007/s15010-020-01572-8
- [9] Kurth F, Roennefarth M, Thibeault C et al. Studying the pathophysiology of coronavirus disease 2019: a protocol for the Berlin prospective COVID-19 patient cohort (Pa.COVID-19). *Infection* 2020; 48: 619–626. DOI: 10.1007/s15010-020-01464
- [10] Schmidt CO et al. COVID-19-Forschungsdaten leichter zugänglich machen – Aufbau einer bundesweiten Informationsinfrastruktur [Making COVID-19 research data more accessible—building a nationwide information infrastructure]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 2021; 64: 1084–1092. DOI: 10.1007/s00103-021-03386-x
- [11] Kraus M et al. Studienübergreifende Harmonisierung datenschutzrechtlicher und ethischer Gesichtspunkte in Patientenunterlagen: Eine Praxisanalyse. In: Richter G, Loh W, Buyx A, Graf von Kielmansegg S (eds) *Datenreiche Medizin und das Problem der Einwilligung*. Springer; Berlin, Heidelberg: 2022. DOI: 10.1007/978-3-662-62987-1\_9
- [12] Hofmann S. Forschungsklausel statt Broad Consent. *Datenschutz Datensich* 2022; 46: 756–761. DOI: 10.1007/s11623-022-1691-7
- [13] Tilch K, Hopff SM, Appel K et al. Ethical and coordinative challenges in setting up a national cohort study during the COVID-19 pandemic in Germany. *BMC Med Ethics* 2023; 24: 84. DOI: 10.1186/s12910-023-00959-0
- [14] Frank G et al. Management and Data Sharing of COVID-19 Pandemic Information. *Biopreservation and Biobanking* 2020; 18: 570–580. DOI: 10.1089/bio.2020.0134
- [15] Pita Costa J et al. "Meaningful Big Data Integration for a Global COVID-19 Strategy," in *IEEE Computational Intelligence Magazine*. 2020; 15: pp 51–61. DOI: 10.1109/MCI.2020.3019898.