



Deutsche Gesellschaft für
Integrierte Versorgung
im Gesundheitswesen e.V.

Positionspapier

GenAI in Healthcare und Life Sciences

GenAI: Der neue Speed-Booster für das deutsche Gesundheitssystem und besonders Integrierte Versorgung?

Chancen und Risiken von **Daten-Demokratisierung** und Nutzung **generativer KI** am Beispiel der Entwicklung innovativer Medikamente

DGIV Arbeitskreis Digitalisierung

Dr. Markus Fuchslocher, Jens Dommel

Vorwort: Dr. Michael Meyer

Inhaltsverzeichnis

Vorwort	3
1 Executive Summary	3
2 Herausforderungen im Gesundheitsbereich	5
3 Künstliche Intelligenz und Datendemokratisierung	5
3.1. Abgrenzung Künstliche Intelligenz	5
3.2 Datendemokratisierung	7
4 Zielbild einer effektiveren intersektoralen Versorgung mit Medikamenten durch generative KI.....	7
4.1 Pharma-Forschung.....	8
4.2 Klinische Entwicklung	9
4.3 Zusätzliche Mehrwerte durch Datendemokratisierung.....	11
4.4. Technische Umsetzung mit Hilfe von Health-Datalakes	12
5 Anpassungsbedarf und Handlungsempfehlungen für GenAI in der intersektoralen Anwendung	13

Vorwort

Wird über die zügige und resiliente Umsetzung intersektoraler, interdisziplinärer und interprofessioneller Versorgungsmodelle diskutiert, so treten dabei immer stärker die unterstützenden Möglichkeiten digitaler Infrastrukturen und Anwendungen in den Vordergrund. Zusätzlich rücken aber auch die Optionen sowohl einer „Daten-Demokratisierung“, also eines Prozesses, digitale Daten für Anwender innerhalb eines Systems sinnvoll verfügbar zu machen, als auch die Nutzung generativer KI in den Vordergrund. Am Beispiel der Entwicklung innovativer Medikamente wird in diesem Papier dargelegt, welche Chancen und Risiken hiermit verbunden sind und – vielleicht noch wichtiger als diese einfache Abwägung – was nun ganz konkret zu tun ist, damit mögliche Gefahren beherrschbar bleiben und die Vorteile auch wirklich nutzbar sind.

Die Deutsche Gesellschaft für Integrierte Versorgung im Gesundheitswesen e.V. begrüßt diese Diskussion und ruft dazu auf, theoretische Denkmodelle schnell in die Verprobung zu bringen und – wo sinnvoll – einen nachhaltigen „Wirkbetrieb“ zu starten. Dies ist der beste Weg, zum Wohle der Patienten zu arbeiten und die Chancen moderner integrativer Technologien zu nutzen.

Dr. Michael Meyer, Generalsekretär DGIV e.V.

1 Executive Summary

Ein sicherer Health-Datalake in der Cloud, kombiniert mit Generative KI (GenAI), kann die Medikamentenentwicklung und Versorgungsqualität erheblich beschleunigen, Kosten optimieren und gleichzeitig tiefere Einsichten liefern und damit die Qualität und Patientenwohl verbessern.

Das Gesundheitswesen ist geprägt durch eine Vielzahl heterogener, komplexer und sensibler Daten aus verschiedenen Quellen. Diese Daten stammen aus beispielsweise elektronische Patientenakten, medizinischen Bildern, Laborergebnissen, genetischen Tests, klinischen Studien, Versicherungsdaten, regulatorische Daten, aber auch Wearables oder sozialen Medien. Diese Daten haben einen hohen Wert für die Verbesserung der Gesundheit und des Wohlbefindens der Menschen, die Optimierung der Gesundheitsversorgung, die Förderung der medizinischen Forschung und Entwicklung sowie der Schaffung neuer Geschäftsmodelle im Gesundheitswesen.

Die moderne Geschäftswelt erlebt einen signifikanten Wandel durch die Integration Künstlicher Intelligenz (KI) in verschiedenen Sektoren, insbesondere im Bereich der Gesundheitsversorgung und des Managements von Lieferketten. Der vorliegende Artikel beleuchtet das Potential von GenAI am Beispiel der Medikamentenversorgung – mit Fokus auf die Entwicklung neuer Medikamente – und zeigt auf, wie sich dabei Effizienz und Effektivität steigern lassen.

Dazu gibt es bereits viele Veröffentlichungen – allerdings stets mit einem Siloblick bzgl. der eigenen Unternehmensdaten oder innerhalb der Pharmaindustrie. Die „Demokratisierung“ von Daten bietet darüber hinaus enorme Chancen für eine verbesserte Patientenversorgung und effektivere und effizientere Forschung und Entwicklung von Medikamenten. Der Übergang zu einem datengetriebenen Ansatz erfordert jedoch erhebliche Anstrengungen in Datenqualität,

Datensicherheit und Verhinderung von Manipulierbarkeit von Daten sowie Interoperabilität von verschiedenen Datenquellen und IT-Systemen.

Mit dem im Mai 2024 verabschiedeten Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz, DVG) wurde ein wichtiger Meilenstein gesetzt, zum einen ein rechtssicherer Rahmen für die Digitalisierung geschaffen, zum anderen aber auch die Einführung der EPA, Einrichtung der E-Health Infrastruktur, DiGA, DiPA als auch der Aufbau und Betrieb des Forschungsdatenzentrums geregelt. Der rechtliche Rahmen ist also weitgehend gesetzt. Somit gibt es jetzt keine Argumente mehr zum entschleunigten Aussitzen von dringlich erforderlicher digitaler Transformation im Gesundheitswesen im Allgemeinen und bei der Entwicklung neuer Medikamente sowie der Versorgungssicherheit mit Medikamenten im Speziellen.

Durch den Paradigmenwechsel im Gesundheitsdaten-Management durch Cloud und den Aufbau und die Nutzung von sogenannten Health-Datalakes können die Stakeholder im Gesundheitswesen die Herausforderungen der Datenintegration und -analyse meistern und gleichzeitig die Compliance und Verlässlichkeit der Daten gewährleisten bzw. darauf vertrauen. Ein Health-Datalake ist eine Datenplattform, die eine umfassende Sammlung von Gesundheitsdaten aus verschiedensten Quellen wie Krankenakten, Sensordaten, klinischen Studien und vielem mehr integriert. Diese können zentral oder föderiert gespeichert werden. Damit können Gesundheitsdienstleister und aber auch die Pharmaforschung Erkenntnisse aus diesen Daten gewinnen, um die Patientenversorgung zu verbessern und neue Behandlungsmöglichkeiten zu entwickeln. Die Integration von Echtzeitdaten, neuen Datenquellen (z.B. Wearables oder genetischen Daten) und einer stetigen Weiterentwicklung der GenAI-Modelle eröffnet neue Möglichkeiten für personalisierte Medizin und verbesserte Prävention und Diagnose.

Neben den Vorteilen der GenAI-gestützten Transformation in der Gesundheits- und Pharmaindustrie, beleuchtet der Artikel auch bestehende Hindernisse und mögliche Handlungsempfehlungen zu einer erfolgreichen Vermeidung sowie für die Implementierung demokratisierter Daten und GenAI.

Konkret leiten sich für die DGIV folgende zentralen Forderungen ab, die im Kapitel 4 detailliert werden:

1. GenAI als Innovationsmotor sollte in (öffentlichen) Ausschreibungen nicht nur zuzulassen, sondern stärker gefördert beziehungsweise aktiv gefordert werden.
2. Die „Demokratisierung“ von Daten spielt eine entscheidende Rolle, um das Potential von GenAI noch stärker ausschöpfen zu können. Wir fordern deshalb, dass Daten innerhalb des Gesundheitssystems allen relevanten Stakeholdern einfach zur Verfügung stehen sollten, um fundiertere Entscheidungen zu treffen, Prozesse zu verbessern oder Innovationen voranzutreiben.
3. Den Zugang zu führenden Technologien wie beispielsweise Cloud und KI-Plattformen gilt es zu fördern, mit dem Ziel: beste Technologie für bestes Ergebnis.
4. Zur intersektoralen Nutzung von GenAI regen wir an, dass ähnlich dem Krankenhauszukunftsgesetz für Krankenhäuser öffentliche Mittel zum Anschub digitaler Vernetzung im Sinne von Health-Datalakes bereitgestellt werden. Damit sollen Erfahrungen gesammelt werden im Rahmen von Proof-of-Concepts (POCs) bei der Nutzung von GenAI zwischen Kliniken, in regionalen Gesundheitsnetzwerken sowie auch in Public-Private-Partnerships im Rahmen der Entwicklung neuer Medikamente zusammen mit der Pharmabranche.
5. Für einen vereinfachten, standardisierten, kosteneffizienten Vergabeprozess empfehlen wir ein Finanzierungsangebot für Pilotprojekte zu schaffen, um Spezifika zu erarbeiten für Referenzarchitekturen für GenAI-Modelle und use-case-spezifische Lösungen sowie Geschäftsmodelle - Schwerpunkte sollten dabei sein:

- GenAI in regionalen Versorgungsnetzen
- GenAI zur Verbesserung der medizinischen Versorgung und der Entlastung des insbesondere medizinischen Personals
- Aufbau von multimodalen Datenplattformen zur sektorübergreifenden Forschung und Nutzung von GenAI durch öffentliche und private Unternehmen. Vereinfachung des Zugangs zu Daten durch Data-Broker-Modelle

2 Herausforderungen im Gesundheitsbereich

Die COVID-19-Pandemie hat weltweit erhebliche Herausforderungen und Lücken in der Entwicklung neuer Medikamente und der Sicherstellung der Medikamentenversorgung aufgezeigt. Diese Herausforderungen lassen sich in mehrere Schlüsselbereiche unterteilen:

1. **Resilienz in der Gesundheitsversorgung:** Die Herausforderung der Resilienz von Gesundheitssystemen besteht darin, diese trotz wachsender Komplexität, knapper Ressourcen und unvorhersehbarer Krisen wie Pandemien, Naturkatastrophen oder Cyberangriffe funktionsfähig zu halten. Dies erfordert flexible Strukturen, hohe Anpassungsfähigkeit und die Fähigkeit, schnell auf Notfälle zu reagieren, um die Versorgung der Bevölkerung sicherzustellen.
2. **Unterbrechung der Lieferketten:** Die Pandemie führte zu erheblichen Störungen in den globalen Lieferketten, insbesondere für pharmazeutische Produkte. Ein Großteil der weltweiten Produktion von Vorprodukten und Wirkstoffen ist auf wenige Länder konzentriert, insbesondere China und Indien. Diese Konzentration führte zu erheblichen Engpässen, als die Pandemie die Produktion in diesen Ländern unterbrach und die globalen Lieferketten ins Stocken gerieten.
Um die Versorgungssicherheit zu erhöhen, haben Unternehmen und Regierungen erste Pläne zur Risikominderung in den Lieferketten angedacht. Dazu gehören beispielsweise die Erhöhung der Transparenz in den Lieferketten, die Antizipation von Marktnachfrage und Bedarfe, der Aufbau von Sicherheitsbeständen und die Implementierung von Notfallplänen bis hin zum Überdenken der Regionalisierung der Wertschöpfungskette.
3. **Beschleunigte Entwicklung und Zulassung neuer Medikamente:** Die Pandemie zeigte aber auch dass es möglich ist, die Effizienz bei der Entwicklung und Zulassung neuer Medikamente signifikant zu steigern. Der Einsatz moderner Technologien, wie Künstliche Intelligenz und „Big Data“, kann hier eine entscheidende Rolle spielen. Diese Technologien ermöglichen eine schnellere und präzisere Analyse von Daten, was die Entwicklung und Zulassung neuer Medikamente beschleunigen kann.

3 Künstliche Intelligenz und Datendemokratisierung

3.1. Abgrenzung Künstliche Intelligenz

Der Ansatz der Künstlichen Intelligenz ist nicht neu. Bereits 1950 veröffentlichte Alan Turing seinen bahnbrechenden Aufsatz "Computing Machinery and Intelligence", in dem er den heute als "Turing-Test" bekannten Ansatz zur Bewertung von KI vorschlug. 1956 wurde der Begriff "Künstliche Intelligenz" auf der Dartmouth Conference von John McCarthy geprägt. Erst durch

den enormen Fortschritt der Computertechnologie und durch die Verfügbarkeit elastischer Rechenleistung in der Cloud gab es seit 2010 enormen Fortschritt durch „Deep Learning und größeren neuronalen Netzen.

Künstliche Intelligenz ist ein weitreichendes Feld der Informatik, das sich mit der Entwicklung von Algorithmen und Systemen beschäftigt, die in der Lage sind, Aufgaben zu erledigen, die typischerweise menschliche Intelligenz erfordern. Dazu gehören Problemlösung, Lernen, Spracherkennung und Entscheidungsfindung. In den letzten Jahrzehnten hat KI erhebliche Fortschritte gemacht und umfasst nun verschiedene spezialisierte Bereiche wie Maschinelles Lernen (ML), Deep Learning, neuronale Netzwerke und Generative Künstliche Intelligenz.

Maschinelles Lernen

Maschinelles Lernen ist ein Teilbereich der KI, der sich darauf konzentriert, Computern die Fähigkeit zu geben, aus Daten zu lernen und Vorhersagen oder Entscheidungen zu treffen, ohne explizit programmiert zu sein. ML-Algorithmen verwenden statistische Methoden, um Muster in Daten zu erkennen. Diese Algorithmen werden mit Trainingsdaten „gefüttert“, um Modelle zu erstellen, die dann auf neue, unbekannte Daten angewendet werden können. Anwendungen von ML finden sich in vielen Bereichen, darunter Vorhersagemodelle wie Wettervorhersagen und Börsenprognosen, Klassifikation wie Spam-Erkennung und Diagnosen von Krankheiten, sowie Clustering wie Kundensegmentierung und Bildsegmentierung.

Neuronale Netzwerke

Neuronale Netzwerke sind die Grundbausteine vieler ML- und Deep Learning-Systeme. Inspiriert von der Struktur und Funktionsweise des menschlichen Gehirns bestehen sie aus Neuronen, die in Schichten organisiert sind. Ein neuronales Netzwerk nimmt Eingabedaten entgegen, verarbeitet sie durch verborgene Schichten und erzeugt eine Ausgabe. Diese Verarbeitung erfolgt durch gewichtete Verbindungen zwischen den Neuronen, die während des Trainings angepasst werden, um die Genauigkeit des Modells zu maximieren. Neuronale Netzwerke können für eine Vielzahl von Aufgaben eingesetzt werden, darunter Klassifikation, Regression und Mustererkennung.

Deep Learning

Deep Learning ist ein Teilbereich des maschinellen Lernens, der auf tiefen neuronalen Netzwerken basiert. Diese Netzwerke bestehen aus vielen Schichten (daher "tief"), die jeweils Daten durch eine Reihe von Transformationen verarbeiten und analysieren. Deep Learning-Modelle, wie Convolutional Neural Networks (CNNs) und Recurrent Neural Networks (RNNs), sind besonders gut geeignet für Aufgaben wie die Verarbeitung von Bildern, Videos und natürlicher Sprache. Die automatische Merkmalsextraktion ist ein herausragendes Merkmal des Deep Learnings, da es relevante Merkmale aus Rohdaten lernen kann, ohne dass diese manuell definiert werden müssen. Dies hat zu bahnbrechenden Fortschritten in Bereichen wie der Bild- und Spracherkennung geführt.

Generative Künstliche Intelligenz (GenAI)

Während konventionelle KI-Systeme ausgelegt sind, bestimmte Aufgaben, wie Klassifizierung, Vorhersagen oder Entscheidungsfindung zu lösen, konzentriert sich GenAI, als ein spezialisierter Bereich der KI, darauf **neue Inhalte zu erstellen**, wie z.B. Texte, Bilder, Musik und andere Medien. Die Neuartigkeit von GenAI liegt in der Verfügbarkeit der sogenannten Foundation Modelle (FM) und Anwendungen, die diese FMs nutzen. Eine spezifische Ausprägung der FMs sind die Modelle für die Verarbeitung und Generierung von natürlicher Sprache (Large Language Modelle).

Bei den FMs handelt es sich um großangelegte, vortrainierte Maschinen-Learning-Modelle.

Diese Modelle werden typischerweise auf riesigen (Milliarden oder Trilliarden Parameter) und vielfältigen Datensätzen trainiert, um die Feinheiten der Sprache und Muster in Daten zu erlernen. Nach dem Training werden die FMs für spezifische Anwendungen fein abgestimmt.

GenAI beschleunigt die KI-Modellerstellung, und es werden weniger Daten benötigt, um die nutzbare KI-Modelle zu entwickeln und zur Verfügung zu stellen. Das könnte insbesondere in der Entwicklung neuer Medikamente bahnbrechend sein, wie Kapitel vier zeigt.

3.2 Datendemokratisierung

Die „Demokratisierung“ von Daten spielt eine entscheidende Rolle, um das Potential von KI noch stärker ausschöpfen zu können. Datendemokratisierung bedeutet, dass Daten innerhalb einer Organisation oder eines Ökosystems – wie beispielsweise das Gesundheitssystems – allen relevanten Stakeholdern zur Verfügung stehen, um fundierte Entscheidungen zu treffen, Prozesse zu verbessern oder Innovationen voranzutreiben. Es stehen also weit mehr Daten zur Verfügung gegebenenfalls sogar Echtzeitdaten, die zuvor nicht zur Verfügung gestanden haben. Das könnte sich signifikant auf Qualität, Zeit und Kosten eines Prozesses auswirken.

Ein Großteil der heute produzierten Gesundheitsdaten kann jedoch für die klinische und Sekundärnutzung nicht verwendet werden, da sie primär für einen anderen Zweck erhoben wurden, unstrukturiert sind oder sich in voneinander getrennten Daten-Silos bei unterschiedlichen Endgeräten oder Institutionen befinden. Der Zugang zu Daten wird derzeit stark durch Anwendungen geprägt, wie beispielsweise Krankenhausinformationssysteme, Archivierungssysteme für medizinischen Bilddaten (sog. PACS-Systeme), Laborinformationsmanagementsysteme, bis hin zu Automatisierungs- und Verwaltungssystemen der gesamten Unternehmensprozesslandschaft (sog. ERP – Enterprise Resource Planning-Systeme) bzw. IT-Systemen bei niedergelassenen Ärzten sowie in Reha oder Pflege. Die Interoperabilität – also die Fähigkeit von verschiedenen IT-Systemen, Diensten oder Geräten eigenständig miteinander kommunizieren zu können und Daten auszutauschen – ist eine stetige Herausforderung. Es gibt unterschiedliche (klinische) Daten-Standards wie HL7 FHIR, openEHR, Observational Medical Outcomes (OMOP).

Durch neue Cloud-Technologien verändert sich die Art und Weise vom Datenmanagement, das man von einem Paradigmenwechsel im Gesundheitsdatenmanagement sprechen kann:

1. Datenmodelle werden neuartig aufgebaut in sogenannten Health-Datalakes und nicht wie bisher in statischen Datawarehouse-Systemen.

2. Der Zugriff auf Daten wird in zentralen Data-Governance-Modellen geregelt und überwacht.

3. Anwendungen werden zu den Daten gebracht und nicht umgekehrt wie bisher.

4. Zweckbasierte Nutzung der richtigen Technologie bei der Datenanalytik und Datenverarbeitung führt zu besseren Ergebnissen. Zudem erfordert „Big-Data“ Verarbeitung für Populationsforschung eine unterschiedliche Technologie als bei der Diagnoseunterstützung am „Point-of-Care“.

4 Zielbild einer effektiveren intersektoralen Versorgung mit Medikamenten durch generative KI

Traditionelle Methoden der Arzneimittelentwicklung sind oft zeitaufwendig und kostspielig. So dauert die Entwicklung eines neuen Medikaments durchschnittlich zehn Jahre und kostet mehr als eine Milliarde Euro. Der Einsatz von KI hat in der Pharmaindustrie Forschung und

Entwicklung ist kein Neuland. So werden bereits heute Deep Learning Modelle eingesetzt, um Krankheiten besser zu verstehen oder Targets bzw. Proteinstrukturen zu analysieren, mit denen dann ein potentieller Wirkstoff interagieren kann.

GenAI kann den gesamten Forschungs- und Entwicklungsprozess erheblich beschleunigen, indem sie automatisierte und optimierte Lösungen für verschiedene Phasen der Forschung und Entwicklung bietet. Von der Target-Identifizierung und -Validierung über die Wirkstoffsuche bis hin zu präklinischen Tests und klinischen Test sowie der Dokumentation im Rahmen des Zulassungsprozesses kann GenAI zeit- und ressourceneffiziente Alternativen bereitstellen.

4.1 Pharma-Forschung

GenAI vor allem unter Verwendung derzeit kaum oder nicht genutzter Datenquellen von Krankenversicherungen, Krankenhäusern, Apotheken, patientenbezogene Daten aus den elektronischen Gesundheitsakten etc. – und all das weltweit kann die gesamte Forschung beschleunigen.

Target-Identifizierung

GenAI kann insbesondere den manuellen Teilprozess im Rahmen der Target-Identifizierung und -Bewertung deutlich beschleunigen: Informationen aus Dokumenten wie Patenten, Veröffentlichungen in Fachmagazinen oder auch klinischen Studien zu extrahieren. Nicht selten sind diese Informationen unvollständig oder fehlerhaft, was durch zusätzliche intersektorale (auch unstrukturierte) Daten in Text-, Bild- oder Sprachform komplettiert und validiert werden kann.

Insbesondere bei der Evaluierung der vielversprechendsten Targets können „Real-World“-Daten herangezogen werden: Arztbesuche und Informationen aus Krankenhausinformationssystemen, Abrechnungsdaten von Versicherungen, elektronische Gesundheitsakten etc.

GenAI verbessert dabei deutlich Verständnis und Interpretation menschlicher Sprache – insbesondere im medizinischen Kontext – im Gegensatz zu klassischen KI-Ansätzen (z.B. Natural Language Processing, NLP).

Wirkstoffkandidaten bestimmen

Aufbauend auf dieser analytischen Target-Identifizierung kann nun GenAI den Prozess beschleunigen, mögliche Wirkstoffkandidaten bzw. chemische Verbindungen zu identifizieren und vor allem die geeignetsten auszuwählen. Dabei werden Millionen von Wirkstoffkombinationen mit bereits existierenden Ergebnissen getesteter Moleküle verglichen, um somit mögliche zusätzliche Sequenzen in der Struktur von „small molecules“ vor allem aber für komplexen „large molecules“ von Biologics vorauszusagen.

Im Bereich der Biowissenschaften bestehen durch GenAI und Cloud völlig neue Möglichkeiten in der Proteinentwicklung. Proteine sind der Programmiercode unseres Lebens. Mit neuen KI-Modellen können die Funktion von Biomolekülen besser entschlüsselt werden und gänzlich neue Proteine entwickelt werden. Die programmierbare Biologie beschleunigt die Arzneimittelentwicklung. Dazu werden bereits spezifische Foundation Modelle für Forscher zur Verfügung gestellt.

GenAI kann dazu beitragen, kontextuelle Informationen effizient zu verarbeiten und komplexe chemische Muster zu erkennen, was zu innovativen und potenziell effektiveren Medikamentenkandidaten führen kann. GenAI kann bestehende Molekülstrukturen optimieren, indem sie gezielte Modifikationen und verschiedene Varianten testen. Dies kann die Wirksamkeit und Sicherheit von Molekülen verbessern.

GenAI kann zudem zur Vorhersage von Wirksamkeit, Nebenwirkungen und Toxizität von Wirkstoffkandidaten eingesetzt werden. Durch das Training auf Basis umfangreicher biomedizinischer Daten können sie die Wirksamkeit sowie mögliche Nebenwirkungen und Toxizität vorhersagen. Dadurch werden die Sicherheit neuer Medikamente erhöht und zugleich die Entwicklungskosten dadurch gesenkt, dass problematische Kandidaten frühzeitig aus dem Entwicklungsprozess genommen und damit „Late-Stage-Flops“ vermieden werden können.

Design von personalisierten Medikamenten

Die Fähigkeit von GenAI, komplexe Muster in großen Datenmengen zu erkennen und zu generieren, kann auch im Bereich der personalisierten Medizin genutzt werden. Durch die Analyse genetischer, klinischer und biochemischer Daten einzelner Patienten kann GenAI individuell angepasste Medikamente oder Behandlungsstrategien vorschlagen, was zu präziseren und effektiveren Behandlungen führt.

Voraussetzung auch hierfür: sektorübergreifende Verfügbarkeit von Daten.

Zusammenfassung

GenAI kann die Integration und Analyse großer und komplexer Datensätze in der Pharmaforschung verbessern. Sie ist in der Lage, heterogene Datenquellen wie genomische Daten, klinische Studien, Patente und Literatur effizient zu verarbeiten und zu kombinieren. Darüber hinaus können weitere Daten herangezogen werden, die bis dato noch keine Berücksichtigung fanden wie beispielsweise patientenbezogene Daten von Krankenversicherungen, Krankenhäusern (zum Beispiel auf Krankenhausinformationssystemen) oder aus den elektronischen Patientenakten. Dies ermöglicht eine umfassendere und tiefere Analyse, die zu neuen Erkenntnissen und besseren Entscheidungsgrundlagen führt.

4.2 Klinische Entwicklung

Im klinischen Entwicklungsprozess wird ein mögliches neues Medikament an gesunden und kranken Menschen getestet. Klinische Studien sind intersektoral und dennoch wird das Potential, was die Verwendung sektorübergreifender Daten anbelangt, nur unzureichend ausgeschöpft.

Klinische Studien sind geprägt von Schwierigkeiten bei der Patientenrekrutierung, persönlichen Arzt-Patient-Gesprächen, manuellem Erfassen von Daten, sehr viel Dokumentation im Rahmen des Zulassungsprozesses, Incompliance von Patienten und Studienabbruchern. Das führt dazu, dass der Entwicklungsprozess durchschnittlich sieben Jahre dauert und knapp 60% der Forschungs- und Entwicklungsausgaben verschlingt.

GenAI bietet erhebliche Verbesserungspotentiale.

Patientenrekrutierung und -auswahl

GenAI kann die Identifikation und Rekrutierung geeigneter Patienten für klinische Studien erheblich verbessern. Durch die Analyse großer Mengen an Patientendaten, einschließlich elektronischer Patientenakten, genetischer Informationen und anderer relevanter Datenquellen, können diese Modelle potenzielle Studienteilnehmer präzise identifizieren. So können beispielsweise Patienten, die genetisch bedingt nicht auf das Medikament reagieren werden, gar nicht erst in eine Studie eingeschlossen werden. Dies erhöht die Rekrutierungsrate und beschleunigt die Patientenselektion, was zu effizienteren und

erfolgreicheren Studien führt – auch dadurch, dass die Anzahl potentieller Studienabbrecher deutlich, reduziert werden kann.

Optimierung der Studiendesigns

GenAI kann verwendet werden, um verschiedene Studiendesigns im Vorfeld einer Studie zu simulieren und die vielversprechendsten Ansätze zu identifizieren. Sie kann Ergebnisse basierend auf unterschiedlichen Studiendesignparametern vorhersagen, was es ermöglicht, die optimale Balance zwischen Studiengröße, Dauer und statistischer Power zu finden. Dies führt zu effizienteren Studien mit höherer Erfolgswahrscheinlichkeit.

Automatisierte und datengetriebene klinische Studien

GenAI ist besonders gut geeignet für die Echtzeitanalyse großer und komplexer Datensätze im Rahmen klinischer Studien. Ein enormer Effizienzgewinn im Vergleich zum heute noch stark händisch dominierten oder semi-automatisierten Prozess. In klinischen Studien können diese Modelle zudem kontinuierlich Studiendaten analysieren und Trends sowie kritische Erkenntnisse identifizieren und Alarme proaktiv auslösen.

GenAI ermöglicht es, Wissenschaftler und auch die Principal Investigators an verschiedenen Standorten automatisiert, proaktiv und personalisiert anzusprechen. Forschern, schneller auf Entwicklungen zu reagieren und fundierte Entscheidungen zu treffen, was die Qualität und Geschwindigkeit der klinischen Forschung erhöht.

Ein weiterer wichtiger Einsatzbereich von GenAI im klinischen Entwicklungsprozess ist das Echtzeit-Überwachungs- und Vorhersagemanagement von Nebenwirkungen. GenAI kann kontinuierlich Patientendaten überwachen und Anomalien oder frühe Anzeichen von Nebenwirkungen erkennen. Dadurch können potenzielle Risiken frühzeitig identifiziert und gemanagt werden, was die Sicherheit der Patienten erhöht und die Gesamteffizienz der klinischen Studien verbessert.

Proaktiver und automatisierter Dokumentationsprozess

GenAI bietet darüber hinaus signifikanten Effizienzgewinn im Rahmen des Zulassungsprozesses:

Das gilt zum einen für die sogenannten Health Authority Queries (HAQ) während der klinischen Studien. Diese können zum einen durch GenAI antizipiert und gegebenenfalls frühzeitig mitigiert werden, sodass die Anzahl der HAQs sinkt oder auch Antworten vorformuliert werden.

Die eigentlichen Zulassungsdokumente können mit Hilfe von GenAI deutlich schneller erstellt werden. Jeder einzelne Schritt, der heute mehrere Tage dauert ist mit GenAI in wenigen Minuten erledigt. Mit GenAI kann so ein qualitativ guter erster Entwurf entstehen, auf dem dann die cross-funktionalen Experten aufsetzen können und sich auf die komplexen Abschnitte fokussieren.

Zusammenfassung

GenAI kann nicht nur die pharmazeutische Forschung, sondern auch den klinischen Entwicklungsprozess erheblich verbessern. Von der Patientenrekrutierung und -auswahl über die Optimierung von Studiendesigns bis hin zur Echtzeitanalyse und dem Nebenwirkungsmanagement bietet GenAI innovative Lösungen, die die Effizienz und Effektivität der klinischen Entwicklung deutlich steigern. Wichtig hierbei ist, jedoch nicht weitere Datensilos zu schaffen,

sondern zum Beispiel bei den Krankenhäusern oder in regionalen Versorgungsstrukturen, Datalake-Architekturen zu schaffen, die sowohl den Bedarf an klinischen Studien aber auch in der weiteren Nutzung der Daten für Optimierung, Automatisierung oder klinische Entscheidungsunterstützung.

4.3 Zusätzliche Mehrwerte durch Datendemokratisierung

Unter Datendemokratisierung verstehen wir den Zugang zu und die Nutzung von Daten für alle Beteiligten eines Ökosystems wie beispielsweise des Gesundheitswesens. Datendemokratisierung bedeutet, dass Daten nicht mehr nur von Experten oder Spezialisten eines Bereichs (z.B. Krankenhaus) kontrolliert oder analysiert werden, sondern dass jeder, der an Daten interessiert ist, sie frei einsehen, teilen und anwenden kann. Das erfordert selbstverständlich die Einhaltung datenschutzrechtlicher sowie gesetzlicher und regulatorischer Rahmenbedingungen.

Die Integration von GenAI in die Medikamentenforschung und -entwicklung kann erheblich von der Nutzung weiterer Datenquellen außerhalb des traditionellen Pharmasektors profitieren. Diese zusätzlichen Datenquellen bieten wertvolle zusätzliche Einblicke und Informationen, die die Effizienz, Qualität und Kosteneffektivität der Medikamentenentwicklung verbessern können.

Versicherungsdaten

Versicherungsdaten bieten umfassende Informationen über Behandlungsmuster, Gesundheitszustände und die Nutzung von Gesundheitsdiensten. Diese Daten umfassen Behandlungshistorien, Kostenanalysen, die Häufigkeit und Art von Diagnosen sowie Patientenergebnisse. GenAI kann diese Daten nutzen, um besser zu verstehen, welche Therapien in der realen Welt am effektivsten sind und wo es Lücken in der Versorgung gibt. Dies kann zu verbesserten Vorhersagen der Wirksamkeit neuer Medikamente, zur Identifikation kosteneffizienter Behandlungsstrategien und zur Optimierung klinischer Studien durch zielgerichtete Patientenauswahl führen.

Elektronische Patientenakten

Elektronische Patientenakten enthalten detaillierte Informationen über die Gesundheitsgeschichte eines Patienten, einschließlich Diagnosen, Behandlungen, Laborergebnissen und Medikamentenverordnungen. Diese Daten ermöglichen die Entwicklung personalisierter Behandlungsansätze und die Überwachung der Wirksamkeit von Medikamenten in verschiedenen Populationen. GenAI kann diese Daten nutzen, um personalisierte Medizin zu fördern, die Patientenüberwachung zu verbessern und Nebenwirkungen sowie langfristige Ergebnisse zu identifizieren.

Wearables und Gesundheits-Apps

Daten von Wearables und Gesundheits-Apps bieten kontinuierliche und detaillierte Informationen über die Gesundheitszustände und Aktivitäten der Nutzer. Zu den generierten Daten gehören beispielsweise Herzfrequenz, Schlafmuster, körperliche Aktivität und Glukosewerte. Diese Daten sind besonders wertvoll für die Echtzeit-Überwachung der Medikamenteneffizienz und die Anpassung der Dosierungen. GenAI kann diese Informationen verwenden, um Behandlungspläne in Echtzeit anzupassen und Gesundheitsprobleme frühzeitig zu erkennen.

Sozial- und demografische Daten

Sozial- und demografische Daten bieten Einblicke in die Prävalenz bestimmter Krankheiten in verschiedenen Bevölkerungsgruppen. Diese Daten umfassen Altersverteilung, Geschlecht, Einkommensniveau, Bildungsgrad und Lebensgewohnheiten. GenAI kann diese Informationen nutzen, um klinische Studien besser zu gestalten, Medikamente zu entwickeln, die spezifische Bedürfnisse und Bedingungen adressieren, und die Patientenzentrierung zu verbessern.

Daten aus Patientengruppen und sozialen Medien

Patientendaten aus nicht-klinischen Quellen wie Patientenforen und sozialen Medien bieten wertvolle Einblicke in die Erfahrungen und Erwartungen der Patienten. Diese Daten umfassen Patientenberichte, Feedback zu Medikamenten und Diskussionen über Nebenwirkungen. GenAI kann diese Informationen auswerten, um seltene Nebenwirkungen schneller zu erkennen, die Patientenaufklärung zu verbessern und patientenfreundlichere Medikamente zu entwickeln.

4.4. Technische Umsetzung mit Hilfe von Health-Datalakes

Der Paradigmenwechsel im Gesundheitsdaten-Management besteht in einem Übergang von isolierten, zweckgebundenen Datenspeichern hin zu integrierten, skalierbaren Datenplattformen – den sogenannten Health-Datalakes.

Traditionell werden Gesundheitsdaten in getrennten Systemen für elektronische Patientenakten, Abrechnungen, Forschungsstudien etc. gespeichert. Das erschwert den ganzheitlichen Zugriff und die Nutzung dieser Informationen. Durch die Cloud-Technologie ist hier ein neuer technologischer Ansatz möglich.

Der Paradigmenwechsel besteht darin, dass er eine neue Flexibilisierung in die Datenerfassung, -zusammenführung und anonymisierte -bereitstellung ermöglicht. Ein Health-Datalake ist ein Datenspeicher-Ökosystem, das strukturierte und unstrukturierte Gesundheitsdaten in großen Mengen und den Ursprungsformaten speichert. Die Daten können aus klinischen Studien, Patientenakten, Laborergebnissen, Bilddaten (z.B. MRT etc.) und mehr stammen.

Dieser Wandel vom fragmentierten zum ganzheitlichen Datenmanagement ist eine Schlüsselvoraussetzung, um die Leistungsfähigkeit des Gesundheitssystems durch datengetriebene Innovationen zu steigern.

Durch Data-Governance werden die Zugriffsrechte bestimmt, so dass sie im Weiteren auf Unternehmens-, Instituts-, Regional- oder Länderebene festgelegt werden können. Ziel dabei ist die Datendemokratisierung, d.h. die Zugriffsmöglichkeit unabhängig von Anwendungen zu ermöglichen. Die Health-Datalakes in der Cloud ermöglichen zweckbasierte Datenmodelle. Dieses Konzept ist wichtig, da dadurch je nach Anwendungsbedarf der Nutzung der Daten die optimale Technologie genutzt werden, z.B. für Big Data Analysen für Populationsforschung oder transaktionsbasierten Entscheidungsunterstützung in der Diagnostik.

Der neue Ansatz der Datalakes zielt darauf ab, alle relevanten Daten aus verschiedensten Quellen in einer zentral oder dezentral gespeicherten, flexiblen Infrastruktur zu bündeln. So können Ärzte, Wissenschaftler und Entscheidungsträger die Daten umfassend analysieren, um Behandlungsprozesse zu optimieren, Präventions- und Versorgungskonzepte zu entwickeln oder zusätzliche Erkenntnisse bei der Entwicklung neuer Medikamente einfließen zu lassen.

GenAI kann nicht nur bei der Auswertung, sondern bereits bei der „Vorverarbeitung“ großer, heterogener Datenmengen genutzt werden. Dies umfasst das Normalisieren von Daten (z.B. Formatierung, Bereinigung von Duplikaten etc.) und die Standardisierung unterschiedlicher Quellen, sodass alle Daten im erforderlichen Format vorliegen und besser analysiert werden können. GenAI kann auch genutzt werden, um synthetische Daten zu generieren, die ähnlich realer Daten sind. Dies ist besonders wertvoll, wenn reale Daten rar sind. So könnten z.B. synthetische Patientendaten für Trainingszwecke genutzt werden. Diese Datenbasis ist die Grundlage für die Analyse, für die Entwicklung und Nutzung von KI-Modellen (GenAI), für Forschung, Entscheidungsunterstützung und Automatisierung. Das bedeutet, dass Modelle fortlaufend mit transformierten und bereinigten Daten trainiert werden können.

Dazu bedarf es aus unserer Sicht eines neutralen „Data-Brokers“ – eines vertrauensvollen Mediators, der die Daten Governance sicherstellt, die erforderlichen Daten und deren Qualität und Aktualität organisiert sowie ggf. auch vortrainierte GenAI Modelle für fallspezifische Anfragen anbietet.

Der Data-Broker ist verantwortlich für die Einhaltung von Datenschutz (v.a. bei Datenaustausch und -speicherung) und gesetzlichen Vorgaben.

5 Anpassungsbedarf und Handlungsempfehlungen für GenAI in der intersektoralen Anwendung

Aus den genannten notwendigen strukturellen Veränderungen in der Forschung und Entwicklung von neuen Medikamenten und den Verbesserungspotentialen, die intersektorale Zusammenarbeit im Allgemeinen und Datendemokratisierung und GenAI im Speziellen leisten kann, leitet die DGIV folgende politischen Grundforderungen als Leitplanken ab, um letztlich Digitalisierung im deutschen Gesundheitswesen nachhaltig voranzutreiben:

1. GenAI ist bahnbrechend und birgt enorme Potentiale in der Verbesserung intersektoraler Zusammenarbeit und bei der Entwicklung neuer Medikamente. Es ist wichtig eine offene Innovationskultur zuzulassen, d.h. zu ermöglichen Erfahrungen zu sammeln, um daraus zu lernen und auf Basis von Wissen und Erfahrung zu regulieren und nicht ex ante auf aufgrund theoretischer Risikobetrachtungen.
2. GenAI schafft einen neuen globalen Markt. Es ist wichtig hier den europäischen und deutschen Unternehmen ein positives Umfeld zu schaffen, Förderung zu ermöglichen, durch Datendemokratisierung die Grundlage für KI zu schaffen, als auch internationale Kooperationen mit globalen Cloudanbietern zu fördern, um sowohl die Umsetzungsgeschwindigkeit zu erhöhen als auch eine schnelle weltweite Skalierbarkeit von GenAI-Anwendungen zu ermöglichen.
3. Für einen vereinfachten, standardisierten, kosteneffizienten Vergabeprozess empfehlen wir ein Finanzierungsangebot für Pilotprojekte zu schaffen, um Spezifika zu erarbeiten für Referenzarchitekturen für GenAI-Modelle und use-case-spezifische Lösungen sowie Geschäftsmodelle - Schwerpunkte sollten dabei sein:
 - GenAI in regionalen Versorgungsnetzen
 - GenAI zur Verbesserung der medizinischen Versorgung und der Entlastung des insbesondere medizinischen Personals
 - Aufbau von multimodalen Datenplattformen zur sektorübergreifenden Forschung (beispielsweise bei der Entwicklung neuer Medikamente im Zusammenspiel mit der Pharmaindustrie) und Nutzung von GenAI durch öffentliche und private Unternehmen. Vereinfachung des Zugangs zu Daten durch Data-Broker-Modelle

4. Es sollten die Finanzierungsmodelle auch für Investitionen im öffentlichen Sektor angepasst werden (Variabilisierung von Kosten im Sinne Capex zu Opex), um transaktionsbasierte Services wie KI für Leistungserbringer und Kostenträger in der Nutzung zu ermöglichen.
5. Wir sehen die Notwendigkeit eines Mediators und Data Brokers, der die einzelnen Akteure und Datalakes zusammenführt, beziehungsweise einen sicheren Datenaustausch ermöglicht und die Umsetzung der Geschäftsmodelle (z.B. Monetarisierung) sicherstellt.
6. GenAI Technologieanbieter sollten den Responsible AI (verantwortungsvollen KI) Ansatz anwenden, um durch Transparenz Fairness und Nichtdiskriminierung, Sicherheit und Zuverlässigkeit, Datenschutz und Persönlichkeitsrechte bei der Entwicklung von KI-Modelle aber auch bei deren Nutzung zu gewährleisten.
7. Das GenAI Modell-Training benötigt Daten von Gesundheitsinstitutionen, Unternehmen und Organisationen. GenAI Technologie-Anbieter müssen gewährleisten, dass diese Daten in der Hoheit und der Kontrolle der einbringenden Organisation bleiben. Damit wird Datenschutz gewährleistet.
8. Bei der Modellerstellung ist darauf zu achten, dass GenAI Anbieter genutzt werden, die Wahlfreiheit von verschiedenen Modellen zulassen. Für das beste Ergebnis von KI-Modellen können unterschiedliche LLMs und FMs benötigt werden, damit wird der Login-Effekt vermieden.)
9. Die Intellectual Property (IP) der entstehenden KI-Modelle durch Nutzung von FM-Modellen sollten in der Hoheit des KI-Modellerzeugers bleiben.
10. Die meisten IT- und Transformationsprojekte scheitern nicht aus technischen Gründen, sondern aufgrund der Missachtung, die Veränderung und Transformation zu managen und zwar mit der erforderlichen Topmanagement-Unterstützung. Übertragen auf ein intersektorales Gesundheitsprojekt bedeutet das: einen Verantwortlichen und einen gemeinsamen Lenkungsausschuss aller beteiligten Akteure am entsprechenden Ökosystem.