



# Arzneimittel mit KI entwickeln: Von der Idee bis zur Zulassung

Anwendungen, Potenziale und Herausforderungen

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

 **acatech**  
DEUTSCHE AKADEMIE DER  
TECHNIKWISSENSCHAFTEN

WHITEPAPER

Budde, K. et al.  
AG Gesundheit,  
Medizintechnik, Pflege

# Inhalt

---

Zusammenfassung .....	3
1 Ausgangslage.....	4
2 Klassische Verfahren in der Wirkstoffentwicklung.....	8
3 KI in der Wirkstoffentwicklung .....	14
3.1 Anwendungsbereiche.....	14
3.2 Vergleich mit klassischen Verfahren .....	23
4 Anwendungsbeispiele mit KI .....	26
4.1 KI in der personalisierten Krebstherapie.....	26
4.2 Nutzung von KI zur Lead-Identifikation und -Optimierung.....	26
4.3 Nutzung von KI in retrospektiven Real-World-Studien .....	27
5 Potenziale und Herausforderungen.....	29
6 Offene Fragen und Gestaltungsoptionen .....	32
Literatur.....	36
Über dieses Whitepaper.....	40

# Zusammenfassung

---

Die Entwicklung eines Arzneimittels ist ein komplexes, langwieriges Unterfangen: Bis zur Markteinführung dauert es an die zwölf Jahre mit durchschnittlichen Gesamtkosten von etwa 2,8 Milliarden US-Dollar. Dieser Prozess von der ersten Idee bis zur Zulassung kann mit dem Einsatz von Künstlicher Intelligenz (KI) effizienter gestaltet werden und bietet die Möglichkeit, jahrelange Arbeit und kostspielige Investitionen einzusparen: durch systematische Analysen bei der Datenverarbeitung, beispielsweise beim Erkennen relevanter Muster aus riesigen Datenmengen, dem Auffinden geeigneter Wirkstoffkandidaten oder auch bei der Umsetzung klinischer Studien. Die durch KI bereitgestellten Freiräume lassen sich in der Arzneimittelforschung produktiv nutzen, um mithilfe von KI-Tools medizinisch bessere und individuellere Medikamente schneller zu entwickeln: zum eigenen Wettbewerbsvorteil wie zum Wohle von Patientinnen und Patienten.

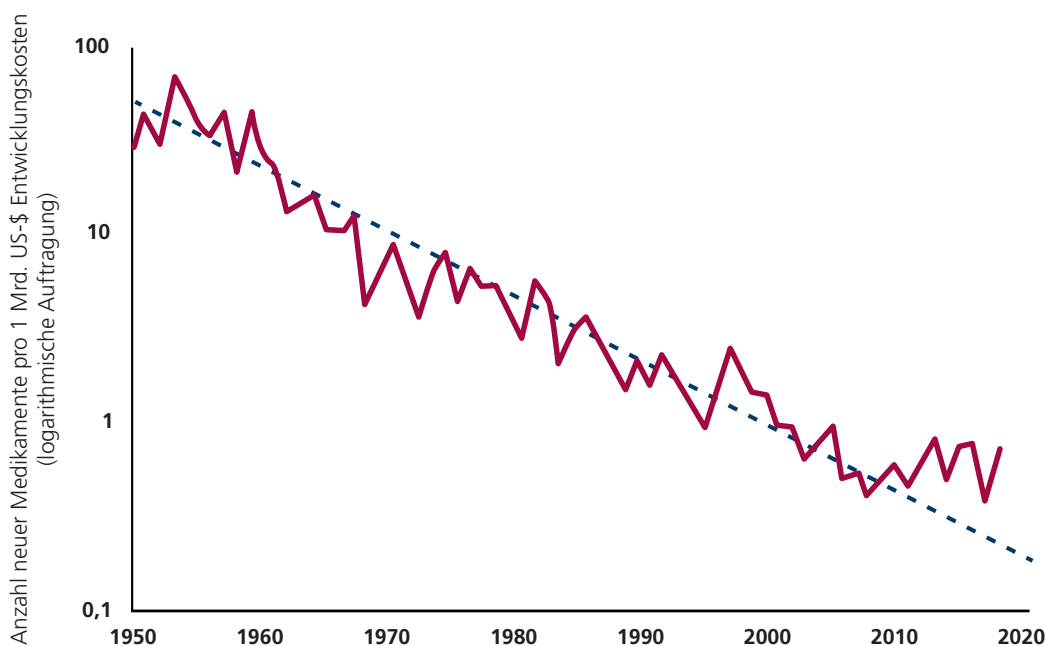
Die Autorinnen und Autoren der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme zeigen in dem Whitepaper auf, dass KI-Technologien bei der Entwicklung neuer Wirkstoffe einen wesentlichen Beitrag hinsichtlich einer gezielteren und kosteneffizienteren Wirkstoffsuche und -prüfung sowie auch darüber hinaus leisten können. Die Potenziale sowie Anwendungsgebiete von KI über die gesamte Arzneimittelentwicklung werden im Vergleich zum klassischen Entwicklungsverfahren anhand konkreter erfolgreicher Praxisbeispiele veranschaulicht. Daneben identifiziert das Papier aber auch bestehende Herausforderungen wie das Fehlen regulatorischer Regeln und Richtlinien, Datenqualität und -verfügbarkeit sowie Interpretierbarkeit und Transparenz, die für einen zuverlässigen KI-Einsatz als fester Bestandteil im Alltag der Arzneimittelforschung unerlässlich sind. Die daraus folgenden Gestaltungsoptionen zeigen Wege für eine KI-unterstützte Arzneimittelentwicklung auf und richten sich an verschiedene Akteure aus Politik, dem Gesundheitsbereich, Unternehmen im Pharma- und Biotechbereich sowie dem KI-Technologiebereich und an Entwicklerinnen und Entwickler.

# 1 Ausgangslage

Auf Arzneimittel entfielen 2022 etwa 17,8 Prozent der deutschen Gesundheitsausgaben beziehungsweise rund 48,8 Milliarden Euro. Damit bilden sie den drittgrößten Ausgabenposten hinter stationären sowie ambulanten ärztlichen und pflegerischen/therapeutischen Leistungen im Gesundheitswesen (GKV-SV, 2023). Für die moderne medizinische Versorgung sind sie als Heil- und Hilfsmittel unverzichtbar; in Deutschland wird bei jedem Arztbesuch durchschnittlich ein Arzneimittel verordnet (BPB, 2017). Die nationalen Pharma- und Biotech-Branchen bilden nach wie vor relativ zum Umsatz die forschungsstärksten Branchen mit einem FuE-Investment von 8,7 Milliarden Euro in 2021 (vfa, 2021). Trotzdem konzentrieren sich die vermehrt entstehenden Partnerschaften zwischen den etwa 700 KI-Unternehmen weltweit im Bereich der Wirkstoffentwicklung und der Pharma-Branche sowie relevanter Investoren größtenteils auf die USA, in Europa auf Großbritannien (Deep Pharma Intelligence, 2022). Das Papier beschäftigt sich daher mit den Potenzialen von KI in der Wirkstoffentwicklung und wie diese in Deutschland und Europa gefördert werden können. Dazu erfolgt zuerst eine kurze Darstellung, warum KI in diesem Bereich immer relevanter wird.

In der Arzneimittelentwicklung zeigt sich seit den 1950er-Jahren ein klarer Trend: Die Entwicklung neuer Arzneimittel wird immer teurer. Im Durchschnitt entstehen für ein neues Arzneimittel Entwicklungskosten von 2,8 Milliarden US-Dollar<sup>1</sup>. In Umkehr zum Mooreschen Gesetz halbiert sich somit die Produktivität im Bereich der Wirkstoffentwicklung mittlerweile alle fünf Jahre, was entsprechend als Eroomsches Gesetz (Abbildung 1) festgehalten wird (Scannell et al., 2012). Die Gründe dafür sind vor allem die immer komplexer werdenden Produkte und Studiendesigns, die steigenden Anforderungen für Dokumentation und Sicherheit während der Entwicklung und die Rekrutierung von Teilnehmenden für klinische Studien.

**Abbildung 1: Eroomsches Gesetz: Entwicklung neu zugelassener Medikamente pro einer Milliarde US-\$ Investitionen in Forschung und Entwicklung**



Quelle: Scannell et al., 2012.

<sup>1</sup> Diese Zahl spiegelt nicht nur die tatsächlichen Entwicklungskosten eines Arzneimittels wider, sondern beinhaltet auch die damit verbundenen Fehlschläge und Opportunitätskosten, das heißt, die indirekten Finanzierungskosten für derart lange und kostenintensive Entwicklungsprojekte.

Infolge steigender Entwicklungskosten sank der durchschnittliche Spitzenumsatz pro Arzneimittel von 816 Millionen US-Dollar 2010 auf 407 Millionen US-Dollar im Jahr 2018 (Deloitte, 2019). Da eine Patentanmeldung für einen neuen Wirkstoff meist bereits am Anfang der Entwicklung erfolgt, um die Idee zu schützen, verbleiben bei einer durchschnittlichen Entwicklungszeit von zwölf Jahren effektiv nur acht Jahre am Markt mit Alleinstellungsmerkmal: Die Patentschutzlaufzeit beläuft sich auf maximal 20 Jahre. Verlängerte Entwicklungszeiten reduzieren so auch die Anzahl der Jahre zur Umsatzgenerierung.

Neben diesen wirtschaftlichen Herausforderungen für Arzneimittelhersteller entstehen durch die steigenden Aufwände in der Arzneimittelentwicklung auch unmittelbare Nachteile für die Bevölkerung und das Gesundheitswesen. Der Kostendruck führt dazu, dass in der Entwicklung nur einige wenige, vermeintlich hochrentable Krankheitsbilder fokussiert werden. Die US-amerikanische Food and Drug Administration (FDA) hat in ihrer Geschichte bisher 1.578 Arzneimittel zugelassen, die 819 Krankheitsbilder betreffen und damit nur einen Bruchteil der insgesamt geschätzten 20.000 bis 30.000 bekannten Krankheiten (Deloitte, 2019). In Deutschland entfielen 2021 fast 50 Prozent der klinischen Studien zur Erprobung von neuen Arzneimitteln auf Krebserkrankungen und immunologische Erkrankungen, während sich die anderen 50 Prozent auf 19 Krankheitsgruppen aufteilten (vfa, 2022).

Krebserkrankungen sind einerseits weitverbreitet in westlichen und damit zahlungskräftigen Ländern und bieten aufgrund ihrer diversen Arten und Untergruppen mehr Möglichkeiten für sogenannte First-in-class-Arzneimittel<sup>2</sup> oder Arzneimittel für seltene Erkrankungen<sup>3</sup> (orphan drug). Der Prozess der Preisbildung für neue Arzneimittel ist komplex und die rechtlichen Rahmenbedingungen dazu verändern sich stetig aufgrund des Spannungsfelds zwischen der Schaffung von Anreizen für Innovationen und der Sicherstellung der Finanzierbarkeit des Gesundheitswesens. Diese Mechanismen werden in der [Infobox in Kapitel 2](#) genauer betrachtet; trotzdem sollen hier kurz die Rahmenbedingungen angesprochen werden, welche die Arzneimittelentwicklung für bestimmte Krankheitsbilder wirtschaftlich attraktiver machen.

Arzneimittel für seltene Erkrankungen profitieren von einer vergünstigten Zulassung und einer verstärkten Marktexklusivität (BfArM, 2023). Zudem gibt es nationale Sonderregelungen in der Preisbildung und Erstattung, die die Besonderheiten von orphan drugs berücksichtigen. Sie sollen die Rentabilität der Entwicklung sicherstellen. Die Anreize haben Wirkung gezeigt: Arzneimittel für seltene Erkrankungen machten 55 Prozent der Neuzulassungen in Deutschland 2019 aus, 2009 waren es dagegen nur 19 Prozent. Der Großteil davon dient zur Behandlung von Krebserkrankungen.

Der Fokus auf diese Art von Arzneimitteln zeigt sich auch in der Preisentwicklung für Neuzulassungen im Vergleich zu bestehenden Arzneimitteln. Der durchschnittliche Packungspreis bei Ersteren steigt seit 2010 deutlich stärker als der Gesamtmarkt für Arzneimittel (WIdO, 2020). Im Gegensatz dazu bilden Antibiotika ein konkretes Beispiel für fehlende Neuentwicklung von Wirkstoffen aufgrund von Unrentabilität, obwohl diese aber eine wesentliche Säule der Gesundheitssicherung sind, da sie Infektionskrankheiten entgegenwirken.<sup>4</sup>

2 First-in-class-Arzneimittel bezeichnet Wirkstoffe mit einem neuen, einzigartigen und damit bisher nicht bekannten Wirkmechanismus.

3 Als Arzneimittel für seltene Erkrankungen (orphan drugs) werden Wirkstoffe mit einem kleinen Anwendungsgebiet aufgrund der Seltenheit der Krankheit bezeichnet. Der Status als Arzneimittel für seltene Erkrankungen kann sich auch auf bestimmte Unterarten von Krebserkrankungen beziehen.

4 Steigende Antibiotikaresistenzen bei Erregern führen zur verminderten oder fehlenden Wirksamkeit bestehender Wirkstoffe, was 2050 zu zehn Millionen Todesfällen durch Antibiotikaresistenzen führen könnte (de Kraker et al., 2016).

In der mittlerweile langjährigen Geschichte der industriellen Wirkstoffentwicklung sind bereits viele Wirkstoffe entdeckt und entwickelt worden; zu Beginn standen dabei, im Vergleich zu heute, strukturell einfachere Moleküle (vgl. Acetylsalicylsäure, Markteinführung als Aspirin 1899 in Deutschland vs. Adalimumab, Markteinführung als Humira 2003 in Deutschland) oder Krankheiten mit einfachen Wirkmechanismen als Entwicklungsziele im Vordergrund. Zudem zielten frühere Arzneimittel oft auf die Behandlung von weit verbreiteten Symptomen ab und waren teilweise frei verkäuflich, wodurch ein größerer Absatzmarkt gegeben war als im Fall der aktuellen Entwicklungen für spezialisierte Krankheitsbilder und die verschreibungspflichtige Anwendung (Johnson, 2023). Als Beispiel für ein frei verkäufliches Arzneimittel mit breitem Anwendungsgebiet ist das 1976 eingeführte Arzneimittel Tagamet (Wirkstoff: Cimetidin) zu nennen, das vor allem für die Behandlung von Sodbrennen verwendet wird; dieses war 1986 das erste sogenannte Blockbuster-Arzneimittel<sup>5</sup>.

Eine weitere Herausforderung stellt die Rekrutierung von Probandinnen und Probanden für klinische Studien dar. Klinische Studien sind notwendig, um die Sicherheit und Wirksamkeit von Arzneimitteln zu demonstrieren. Vor allem bei neueren Therapieformen, wie der Gen-Therapie, und Krankheitsbildern mit wenig Betroffenen wird der Wirksamkeitsnachweis erschwert und die Entwicklung potenziell zeit- und damit kostenintensiver, da Untergrenzen (zum Beispiel > 300 Teilnehmende in Phase III) für die Mindestanzahl an Teilnehmenden erfüllt werden müssen (SKC, 2023a). Deutschland liegt bei der Bereitschaft zur Teilnahme an klinischen Studien im internationalen Vergleich im hinteren Drittel (vfa, 2023a). Deutschland ist bei der Anzahl durchgeführter klinischer Studien von Platz zwei 2012 auf mittlerweile Platz sechs im internationalen Vergleich abgerutscht und auch die Beteiligung an klinischen Studien pro eine Million Einwohnerinnen und Einwohner liegt international nur im Mittelfeld. In einer vfa-Umfrage unter den Mitgliedsunternehmen gaben deswegen 69 Prozent der Befragten an, dass Deutschland als Pharma-Innovationsstandort in diesem Bereich in den nächsten fünf Jahren an Bedeutung verlieren wird (vfa & Kearney, 2023).

Die vorhergehenden Abschnitte zeigen deutlich die aktuellen Herausforderungen in der Wirkstoffentwicklung als Zielkonflikt zwischen Wirtschaftlichkeit und Bedarf. Der Einsatz von Künstlicher Intelligenz (KI) bietet hier einen Ansatzpunkt zur Abmilderung dieses Konflikts und findet mittlerweile auch in der Pharmaindustrie immer breitere Anwendung. Dies liegt an den vor allem seit 2017 zahlreichen technologischen Entwicklungen, die eindrucksvoll den Mehrwert von KI in diesem Bereich zeigen (Deep Pharma Intelligence, 2022; BMBF, 2023). Laut einer Umfrage nutzen mittlerweile bereits 23 Prozent der Pharmaunternehmen KI in der Wirkstoffentwicklung (Schulte, 2022).

## KURZINFO

### Potenziale von KI in der Arzneimittelentwicklung

Konkrete **Einsparpotenziale durch KI** sind abhängig von der jeweiligen Entwicklung sowie dem Wissensstand zur jeweiligen Krankheit und daher nur bedingt abschätzbar. Einer Marktanalyse von Bekryl Market Analysis nach hat KI bis 2028 das Potenzial, weltweit 70 Milliarden US-Dollar in der Arzneimittelentwicklung einzusparen. Laut einer Analyse von BCG erlaubt KI, je nach verfügbarem Vorwissen zum Entwicklungsziel, **zeitliche und finanzielle Einsparungen von 25 bis 50 Prozent in der Arzneimittelentwicklung** (BCG, 2023).

<sup>5</sup> Bezeichnung für Arzneimittel mit einem Jahresumsatz von mehr als eine Milliarde Dollar.

## Potenziale von KI in der Arzneimittelentwicklung

Aktuelle Beispiele verdeutlichen dieses Potenzial und zeigen auf, welcher Mehrwert durch den KI-Einsatz bereits heute realisiert wird:

**DeepMind (Großbritannien):** Die von DeepMind entwickelte Software AlphaFold, auch der Nachfolger AlphaFold2, ermöglicht die KI-basierte Vorhersage von Proteinstrukturen, deren Kenntnis als Angriffsziele von Arzneimitteln für deren Entwicklung essenziell ist, innerhalb weniger Stunden mit hoher Genauigkeit. Um eine vergleichbare Genauigkeit und Auflösung zu erzielen (Jumper et al., 2021), mussten diese Strukturen bisher experimentell in teils monatelanger Arbeit erforscht werden.

**Insilico Medicine (USA):** Das Biotechnologieunternehmen konnte die Entwicklung eines Wirkstoffkandidaten gegen Fibrose bis zur präklinischen Phase durch KI-Unterstützung für weniger als 850.000 US-Dollar durchführen, im Vergleich zu durchschnittlichen Kosten von etwa 664 Millionen US-Dollar in der klassischen Entwicklung (siehe Kapitel 4.1).

**Standigm (Südkorea):** Das Pharmatechunternehmen hat eine KI-basierte Plattform zur Identifikation von Arzneimitteln mit neuen Wirkmechanismen entwickelt, die eine Identifikation dieser Strukturen innerhalb von durchschnittlich sieben Monaten im Vergleich zu typischerweise 30 Monaten ermöglicht (Deep Pharma Intelligence, 2022).

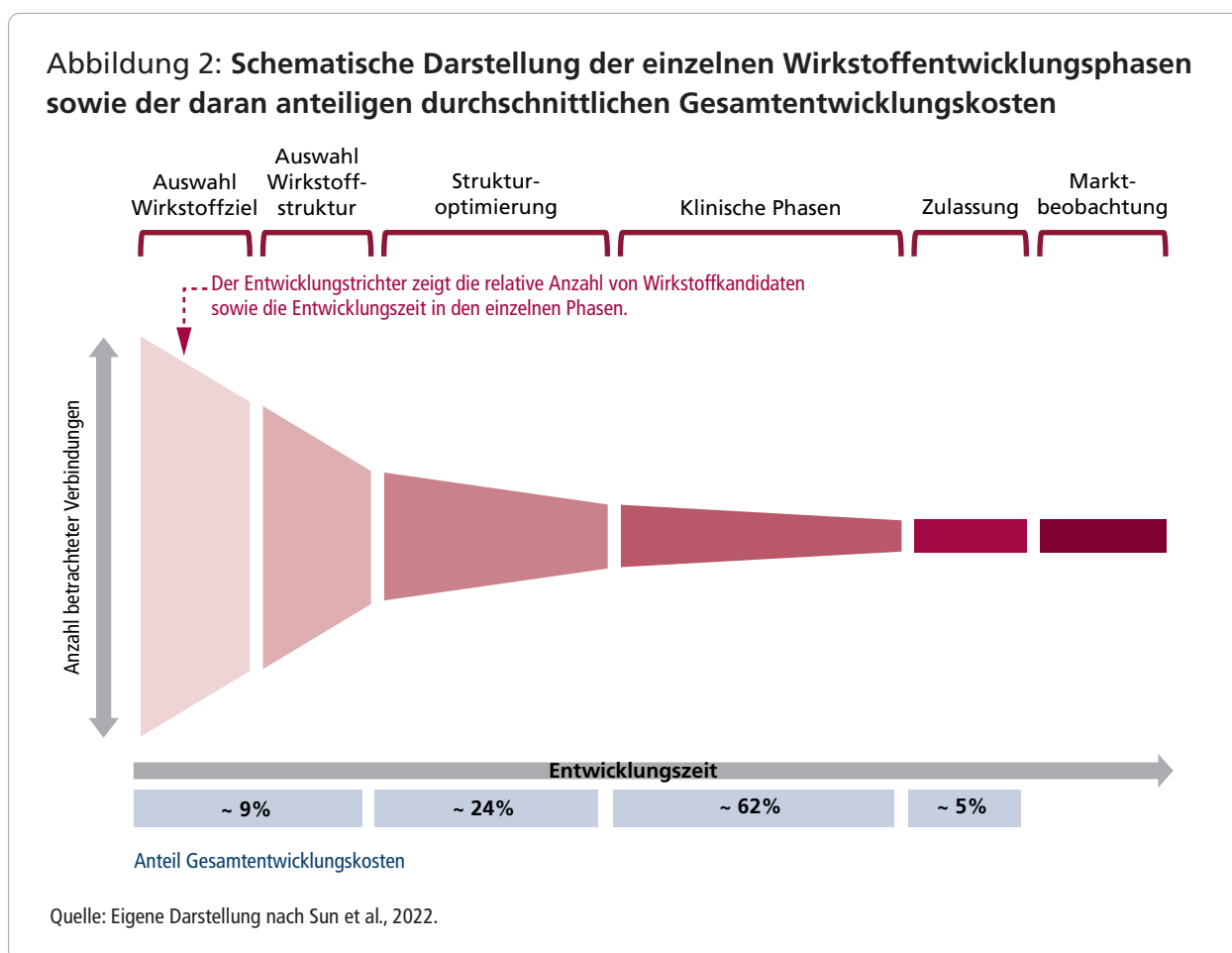
Erkennbarer Mehrwert entsteht in vier (Kern-)Bereichen durch KI in der Wirkstoffentwicklung (BCG, 2022a):

1. Zugang zu neuer Biologie in Form von mehr Biomolekülen und mehr Informationen zu Krankheitsmechanismen und -ursachen
2. Verbesserte oder neue Chemie zur Herstellung von Wirkstoffen
3. Höhere Erfolgsraten in der Entwicklung durch Prognosen und gezielteres Studiendesign in der Klinik
4. Schnellerer und damit günstigerer Entwicklungsprozess

In diesem Papier sollen daher die diversen Einsatzmöglichkeiten von KI in der Wirkstoffentwicklung sowie ihre Vorteile im Vergleich zu klassischen Verfahren dargestellt werden. Gleichzeitig werden Entwicklungsbedarfe und Gestaltungsoptionen identifiziert, wie KI in diesem Bereich mehr Verbreitung finden kann, um Deutschland als Standort für forschungsstarke Pharmaindustrie zu stärken und damit eine resilientere, günstigere und umfassendere Versorgung von Patientinnen und Patienten mit Arzneimitteln sicherzustellen. Der Fokus liegt dabei auf sogenannten niedermolekularen beziehungsweise chemischen Wirkstoffen, da diese immer noch einen erheblichen Teil der weltweiten Arzneimittel-Umsätze ausmachen (Statista, 2023) und der KI-Einsatz bei deren Entwicklung bereits etablierter ist; Biopharmazeutika, also große Moleküle wie Proteine oder Antikörper, die mithilfe von Mikroorganismen produziert werden, dagegen nicht. Denn aufgrund der gesteigerten Komplexität bei der Entwicklung dieser Wirkstoffklasse ist hier der KI-Einsatz noch wenig verbreitet.

## 2 Klassische Verfahren in der Wirkstoffentwicklung

In diesem Kapitel wird zuerst schematisch der Prozess der klassischen Wirkstoffentwicklung als Grundlage für die Darstellung der Einsatzmöglichkeiten von KI in der Wirkstoffentwicklung erläutert. Abbildung 2 sowie die im Folgenden beschriebenen Entwicklungsphasen beziehen sich dabei, wie bereits dargelegt, auf die Entwicklung niedermolekularer beziehungsweise chemischer Wirkstoffe. Der Prozess gliedert sich grob in sechs Phasen (Abbildung 2). Ausgehend von einem großen Pool an möglichen Wirkstoffen wird die Auswahl im voranschreitenden Prozess immer weiter eingeeignet, bis am Ende ein Wirkstoff tatsächlich zugelassen wird.



- **Auswahl des Wirkstoffziels:** Identifikation eines Angriffspunkts (Targets), um den Krankheitsverlauf zu beeinflussen

Ausgehend von einem Krankheitsbild, das bisher nicht oder nur bedingt behandelbar ist und worüber ausreichend Informationen über den zugrunde liegenden Mechanismus der Erkrankung vorliegen, steht am Anfang die Identifikation eines Targets. Target meint hier ein Molekül, welches eine entscheidende Rolle im jeweiligen Krankheitsverlauf spielt und somit einen potenziellen Angriffspunkt für die Beeinflussung des Krankheitsbildes darstellt. Targets sind typischerweise Proteine innerhalb von Zellen oder Rezeptor-Proteine an der Außenseite von Zellen ([siehe Infobox: Wie wirkt ein Arzneimittel?](#)).



## KURZINFO

**Wie wirkt ein Arzneimittel?**

Unabhängig von der Art der zu behandelnden Krankheit oder des Wirkstoffs basieren Arzneimittel darauf, dass der darin enthaltene **molekulare Wirkstoff** selektiv **mit einem bestimmten Biomolekül – seinem Target – interagiert** und seine Funktion beeinflusst.

**Biomoleküle:** Relevante Biomoleküle sind im Allgemeinen Enzyme sowie im Speziellen Rezeptoren, Ionenkanäle oder Transportsysteme an der Zelloberfläche. Die Interaktion mit dem Wirkstoff führt zu Strukturänderungen im Biomolekül, die dessen Aktivität steigern oder mindern; alternativ wird die Interaktion mit anderen Molekülen beeinflusst, sodass diese besser oder schlechter an das Target binden können. Targets können menschliche Proteine sein, aber auch bakterielle Proteine, wie im Fall von Antibiotika.

Zwei konkrete Beispiele verdeutlichen den hier allgemein beschriebenen Wirkmechanismus von Arzneimitteln.

**Acetylsalicylsäure (Aspirin)**

Aspirin wirkt unter anderem schmerzstillend; dazu bindet der Wirkstoff (Acetylsalicylsäure, ASS) an sogenannte Cyclooxygenasen als Target. Das sind Proteine, welche auch an der Herstellung von Prostaglandinen beteiligt sind. Diese sind wiederum Botenstoffe im Körper und interagieren unter anderem mit Schmerzrezeptoren. Die Bindung von ASS an Cyclooxygenasen hemmt die Herstellung von Prostaglandinen und in der Folge auch das Schmerzempfinden, da weniger Botenstoffe zur Aktivierung der Schmerzrezeptoren zur Verfügung stehen (DAZ, 2019).

**Betablocker**

Als Beispiel für Rezeptorproteine als Targets dienen Beta-Rezeptoren an Herzzellen. Diese binden Botenstoffe wie Adrenalin und Noradrenalin, wodurch der Blutdruck steigt. Zur Senkung des Blutdrucks werden sogenannte Betablocker eingesetzt. Diese binden an die beschriebenen Rezeptoren und verhindern dadurch die Bindung und aktivierende Wirkung von Adrenalin und Noradrenalin (DHS, 2021).

**■ Auswahl der Wirkstoffstruktur: Identifikation von Molekülstrukturen als Grundlage für den Wirkstoff**

Ausgehend von einem definierten Target geht es im nächsten Schritt darum, Moleküle zu finden, die mit dem Target interagieren, sodass der für die Behandlung der Krankheit erwünschte Effekt entsteht. Hier werden verschiedene Verfahren verwendet. Wenn noch keine Anhaltspunkte für mögliche Zielstrukturen vorliegen, erfolgt ein sogenanntes Hochdurchsatz-Screening. Dabei werden in automatisierten Verfahren zehntausende unterschiedliche Verbindungen aus einer Molekülbibliothek hinsichtlich ihrer Interaktion mit dem Target getestet. Gibt es bereits Information über die Struktur des Targets oder sogenannter Liganden, also körpereigene Stoffe, die an das Target binden, können diese als Ausgangspunkt für die Entwicklung der Lead-Struktur dienen. Diese Entwicklung erfolgt dann im ersten Schritt simulativ im Rahmen eines Ligandenbasierten Moleküldesigns ausgehend von der Target-Struktur.

## ■ **Struktur-Optimierung: Optimierung der Molekülstruktur zur Verwendung als Wirkstoff**

Auf Basis der verschiedenen Verfahren zur Lead-Identifizierung werden mehrere Strukturen generiert, die an das Target binden und die gewünschte Wirkung erzielen. Diese werden verglichen, um einheitliche Strukturmerkmale zu identifizieren und eine Leitstruktur (Lead) für das weitere Wirkstoffdesign festzulegen. Um als Wirkstoff einsetzbar zu sein, müssen diese Moleküle allerdings neben der Interaktion mit dem Target bezüglich weiterer Eigenschaften optimiert werden; das betrifft zum Beispiel die Toxizität, die Wechselwirkung mit anderen Molekülen im Körper, die Stabilität unter physiologischen Bedingungen oder die Wasserlöslichkeit. Im nächsten Schritt erfolgt daher die chemische Synthese von Derivaten der ursprünglichen Strukturen. Derivate bezeichnen hier Moleküle mit leicht veränderter Struktur im Vergleich zur Ausgangsstruktur. Dabei soll die Bindungswirkung an das Target erhalten bleiben, während gleichzeitig die anderen angesprochenen Eigenschaften der Verbindung optimiert werden. Dieser Prozess verläuft iterativ und erfordert die wiederholte Synthese und Testung der unterschiedlichen Substanzen. Das schließt auch die Untersuchung der Moleküle hinsichtlich ihrer Wirkung auf menschliche Zellen in einer Laborumgebung mit ein sowie an Versuchstieren, wie zum Beispiel Mäusen, im Rahmen der vorklinischen Entwicklung.

## ■ **Klinische Phasen: Klinische Testung zum Ausschluss von Toxizität und Nachweis der Wirksamkeit**

Ausgehend von einer optimierten Lead-Struktur, welche die vorklinische Phase erfolgreich durchlaufen hat, erfolgt im Rahmen von klinischen Studien am Menschen der Wirksamkeitsnachweis als Basis für die Zulassung des Arzneimittels. Diese gliedert sich in drei Phasen: In Phase I erfolgt die Erprobung des Wirkstoffs an wenigen (< 100) gesunden Probandinnen und Probanden, um Toxizität oder unerwartete Nebeneffekte im Menschen auszuschließen. In Phase II wird der Wirkstoff an 100 bis 500 Kranken getestet, um die Wirksamkeit zu bestätigen und die Dosierung zu optimieren. In Phase III wird die Testung auf 300 bis 3.000 Kranke erweitert (auch abhängig von der Inzidenz der Erkrankung). Neben dem Nachweis der Wirksamkeit wird dabei parallel die Darreichungsform des Arzneimittels optimiert. Gleichzeitig erfolgt die Analyse der probandenspezifischen Variabilität und Wirksamkeit des Wirkstoffs. Dabei sollen sogenannte Biomarker, wie Blutwerte oder genetische Eigenschaften, identifiziert werden, die als Grundlage für eine individualisierte Verabreichung beachtet werden müssen.

## ■ **Zulassung: Behördliche Prüfungen für den Zulassungsprozess**

Wenn die klinischen Prüfungen erfolgreich waren, kann für das Arzneimittel eine Zulassung beantragt werden. Dafür sind umfangreiche Dokumente mit teilweise mehreren tausend Seiten zu sämtlichen vorklinischen und klinischen Studienergebnisse erforderlich, die von der Zulassungsbehörde geprüft werden. In Deutschland sind das Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) sowie das Paul-Ehrlich-Institut (PEI) für die Zulassung von Arzneimitteln verantwortlich. Ihre Zuständigkeit grenzt sich anhand der jeweiligen Arzneimittel nach §77 Arzneimittelgesetz ab.

Auf europäischer Ebene bietet die European Medicines Agency (EMA) für alle EU-Mitgliedstaaten die Möglichkeit für ein zentrales Zulassungsverfahren; bestimmte Arzneimittel müssen zentral in der EU zugelassen werden. Die wissenschaftliche Bewertung der Antragsunterlagen wird vom Ausschuss für Humanarzneimittel (CHMP) vorgenommen. Über diesen Ausschuss sind Wissenschaftlerinnen und Wissenschaftler aus allen europäischen Zulassungsbehörden, somit auch das BfArM und das PEI, an der Bearbeitung und Bewertung der Anträge beteiligt. Durchschnittlich dauert die Bearbeitung eines Zulassungsantrags – inklusive Rückfragen an die bzw. den Antragsstellenden und Beantwortung durch diese bzw. diesen – dabei auf europäischer Ebene 13 Monate (EU-Kommission, 2023a).

## ■ Marktbeobachtung: Vertrieb sowie Pharmakovigilanz

Nach erfolgter Zulassung kann das Arzneimittel vertrieben werden und es beginnt die Nachmarktbeobachtung beziehungsweise Pharmakovigilanz. Pharmakovigilanz bezeichnet die Erkennung, Bewertung, das Verständnis und die Vorbeugung von unerwünschten Wirkungen von Arzneimitteln oder anderen arzneimittelbezogenen Problemen. Die Kenntnisse über die Sicherheit von Arzneimitteln sind zum Zeitpunkt ihrer erstmaligen Zulassung nicht vollständig. Dies ergibt sich vor allem daraus, dass die klinische Erprobung eines Arzneimittels an einer relativ geringen Zahl von Probandinnen und Probanden durchgeführt wird. Diese sind zudem unter verschiedenen Aspekten für die klinische Prüfung besonders ausgewählt worden und repräsentieren in der Regel nicht die Durchschnittspatientinnen und -patienten in ärztlichen Praxen. Erst nach Markteinführung können einerseits besonders seltene Nebenwirkungen mit Häufigkeiten unter 0,01 Prozent belastbar analysiert werden. Andererseits können dann erst unerwartete Nebenwirkungen bei Gruppen charakterisiert werden, die in klinischen Studien schwer abbildbar sind: Dies betrifft beispielsweise ältere Menschen oder Schwangere. Dieser Zeitraum wird daher als klinische Phase IV (Wirksamkeitsnachweis) bezeichnet, wobei er außerhalb des klinischen Rahmens abläuft (vfa, 2023b).

### KURZINFO

#### Preisbildung und Nutzenbewertung von Arzneimitteln mit neuen bzw. neuartigen Wirkstoffen

##### Preisbildung

Grundsätzlich entsteht der Preis für ein neues Arzneimittel in Deutschland in zwei Schritten:

**Preisbildung – Schritt 1:** Nach Markteinführung kann der Hersteller für einen Zeitraum von sechs Monaten den Preis seines Arzneimittels frei festlegen. Dieser wird vollumfänglich durch die Kostenträger erstattet. In dieser Zeit erfolgt eine Bewertung des medizinischen Zusatznutzens im Vergleich zu einer zweckmäßigen Vergleichstherapie (zVT) im Rahmen der Nutzenbewertung durch den Gemeinsamen Bundesausschuss (G-BA).

**Preisbildung – Schritt 2:** Anhand dieser Ergebnisse der Nutzenbewertung wird im nächsten Schritt ausgehend von staatlich fixierten Zuschlägen auf Arzneimittelpreise für Apotheken und den Großhandel sowie dem vom Hersteller festgelegten Verkaufspreis der Erstattungsbetrag für das Arzneimittel zwischen Hersteller und dem Spitzenverband der gesetzlichen Krankenkassen (GKV-SV) verhandelt (BMG, 2023).

##### Gesetzliche Vorgabe

Generell herrscht bei der Preisbildung von Arzneimitteln immer ein Spannungsfeld zwischen der Anreizschaffung für pharmazeutische Innovation sowie der Begrenzung von öffentlichen Gesundheitsausgaben (GLI, 2023). Das GKV-Finanzstabilisierungsgesetz (GKV-FinStG) von 2022 legt unter anderem fest, dass für eine höhere Bepreisung gegenüber einer patentgeschützten, bereits hochpreisigen zVT der Nachweis eines mindestens beträchtlichen Zusatznutzens notwendig ist. Zudem wurde die Periode der freien Preisbildung nach Markteinführung von zwölf auf sechs Monate reduziert sowie weitere Änderungen an den gesetzlichen Rahmenbedingungen wie die verpflichtende Berücksichtigung von mengenbezogenen Aspekten vorgenommen.

## Preisbildung und Nutzenbewertung von Arzneimitteln mit neuen bzw. neuartigen Wirkstoffen

### Verfahren der Nutzenbewertung

Arzneimittel für seltene Erkrankungen bis zu einem Umsatz von 30 Millionen Euro pro Jahr (bis 2022 50 Millionen Euro pro Jahr) profitierten bisher von einer vereinfachten Bewertung des Zusatznutzens im Rahmen der Nutzenbewertung. Der Zusatznutzen für diese Arzneimittel gilt bereits mit der Zulassung auf europäischer Ebene und dem darauffolgenden Marktzugang als belegt, lediglich das Ausmaß des Zusatznutzens ist im Rahmen der Nutzenbewertung durch den G-BA festzustellen. Wird dieser Jahresumsatz überschritten, ist ein vollumfänglicher klinischer Nachweis des Zusatznutzens in Bezug auf die zVT sowie eine erneute Verhandlung des Erstattungsbetrags notwendig (SKC, 2023b).

### Reform des EU-Arzneimittelrechts

Auf EU-Ebene wurde ein Kommissionsentwurf zur Überarbeitung des Europäischen Arzneimittelrechts (vom 26.04.2023) vorgestellt, der unter anderem die Versorgungssicherheit von Arzneimitteln innerhalb der EU durch eine engere Überwachung und die verbindliche Einführung von Engpassvermeidungsplänen steigern soll. Zudem sollen regulatorische Prozesse wie z. B. das Zulassungsverfahren beschleunigt und vereinfacht werden; ferner sollen insbesondere Arzneimittel, die einen ungedeckten medizinischen Versorgungsbedarf adressieren, besondere wissenschaftliche und regulatorische Unterstützung erhalten können. Sogenannte „Regulatory Sandboxes“ sollen – ohne Einschränkungen bei Sicherheit, Qualität und Wirksamkeit – die Entwicklung von neuen Arzneimitteln ermöglichen, auf die der bestehende regulatorische Rahmen nicht zugeschnitten ist (EU-Kommission, 2023).

### Förderung des europäischen Arzneimittelmarkts

Eine Kürzung von acht auf sechs Jahre sieht der Kommissionsvorschlag beim Unterlagenschutz für neue Arzneimittel vor; zwei zusätzliche Jahre sollen wiederum gewährt werden, wenn das Arzneimittel innerhalb eines bestimmten Zeitraums in allen Mitgliedstaaten in ausreichender Zahl in Verkehr gebracht wird, für die die Zulassung gilt (EU-Kommission, 2023). Da die gesamteuropäische Vermarktung allerdings nicht allein in der Hand der Unternehmen liegt (die Markteinführung ist regelmäßig von entsprechenden Entscheidungen zu Preisfestsetzung und Erstattung abhängig), besteht die Gefahr, dass sich eine solche Regelung negativ auf die Attraktivität des europäischen Marktes auswirken könnte. Die Anreizschaffung für einen Fokus auf den gesamteuropäischen Arzneimittelmarkt und eine verstärkte lokale und damit resilientere Produktion sollen den vermehrt auftretenden Arzneimittelengpässen entgegenwirken (KAS, 2023). Gleichzeitig entsteht dadurch ein zusätzlicher Kostendruck in der Wirkstoffentwicklung, da diese aktuell vor allem in Asien zu wirtschaftlich günstigeren Bedingungen stattfindet. Dadurch steigt das Risiko, dass sich Pharma-Unternehmen vermehrt auf den bezüglich der Nutzenbewertung weniger regulierten US-Markt konzentrieren und Arzneimittel auf den EU-Markt daher verspätet oder erst gar nicht gebracht werden.

Die Vielzahl an Schritten, die bei der Wirkstoffentwicklung durchlaufen werden, sowie die zunehmenden regulatorischen Vorgaben zur Gewährleistung der Sicherheit der Patientinnen und Patienten zeigen, warum der Gesamtprozess so zeit- und kostenintensiv ist (siehe Infobox: Preisbildung). Gleichzeitig ist es jederzeit möglich, dass die Entwicklung scheitert. Beispielsweise kann es sein, dass im Screening-Verfahren keine Hits mit ausreichender Aktivität gegen das Target identifiziert werden. Bei der Lead-Optimierung kann ein Konflikt entstehen zwischen der Aktivität der Verbindung gegen das Target und Strukturmerkmalen, die für die weiteren pharmakologischen Eigenschaften entscheidend sind. Die präklinischen Verfahren können eine verringerte Aktivität der Verbindung im Organismus zeigen und trotz aller Voruntersuchungen ist die Wirksamkeit im Menschen beim Start der klinischen Phasen wegen fehlender beziehungsweise geringer Übertragbarkeit der Daten aus Untersuchungen in Zell- und Tiermodellen auf den Menschen nicht gesichert.

Gleichzeitig stellt sich hier die Herausforderung, dass ausreichend freiwillige Probandinnen und Probanden gefunden werden müssen und dass diese über die komplette Studiendauer hinweg dabeibleiben. Für die verschiedenen klinischen Phasen zeigt eine Studie, die den Zeitraum von 2010 bis 2017 betrachtet, folgende Erfolgsquoten: 66,4 Prozent: Phase I, 48,6 Prozent: Phase II, 59 Prozent: Phase III und damit nur eine 19-prozentige Erfolgsquote über alle klinischen Phasen hinweg. Aus all den genannten Gründen ist die Erfolgsquote bei der Arzneimittelentwicklung so gering und typischerweise wird nur eine von ursprünglich 10.000 untersuchten Verbindungen tatsächlich zugelassen. Das Entwicklungsrisiko steigt zusätzlich für Arzneimittel mit einem gänzlich neuen Wirkmechanismus (Sun et al., 2022). Gerade diese werden allerdings aktuell aufgrund der flexibleren Preisgestaltung in der Entwicklung forciert (siehe Kapitel 1), wodurch die Entwicklungskosten steigen; gleichzeitig ist diese Klasse von Wirkstoffen aber auch entscheidend, um der großen Anzahl an aktuell nicht behandelbaren Krankheitsbildern entgegenzuwirken.

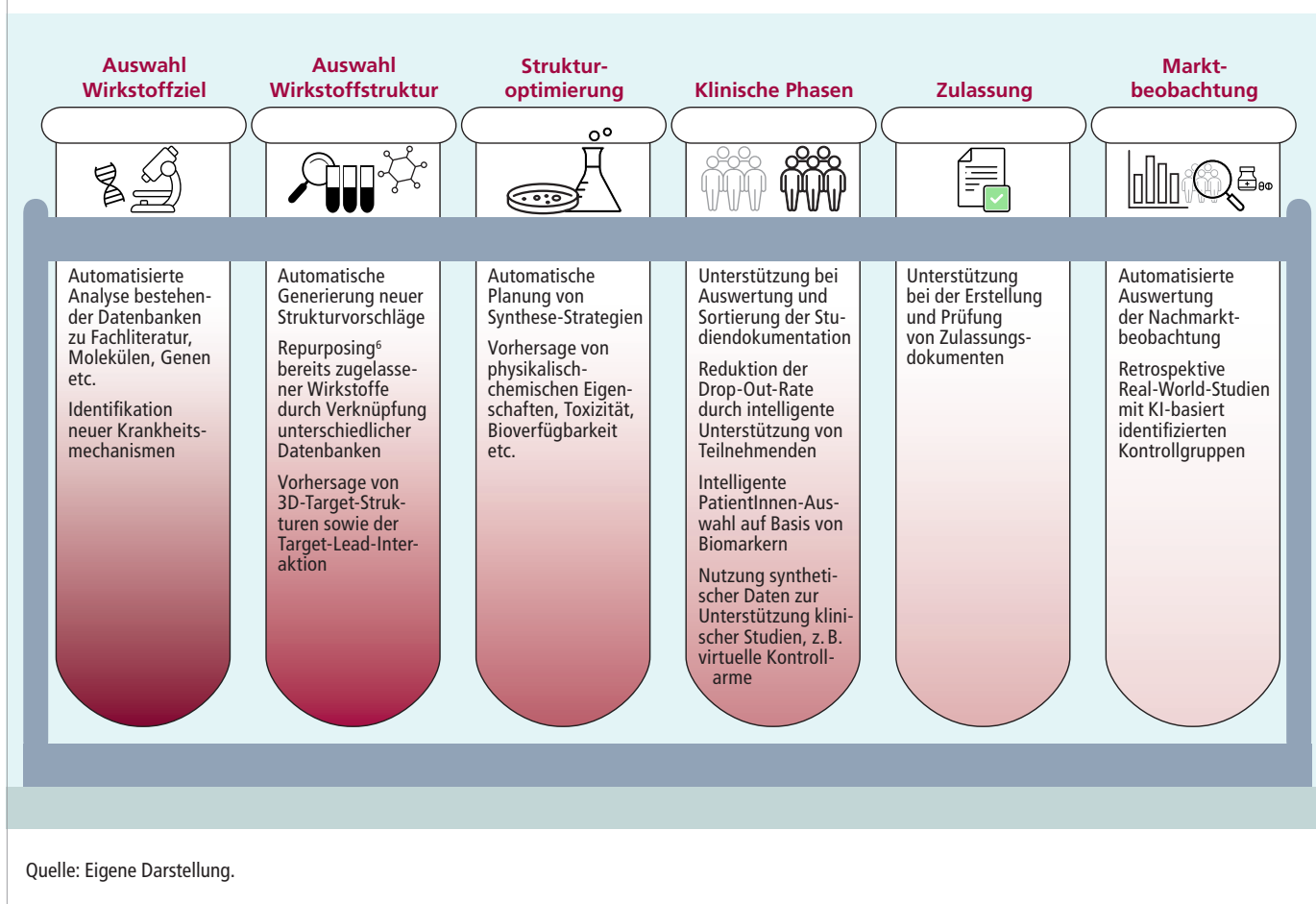
## 3 KI in der Wirkstoffentwicklung

Ausgehend von der Beschreibung der klassischen Verfahren bei der Wirkstoffentwicklung werden im Folgenden mögliche Anwendungsbereiche von KI hierbei aufgezeigt. Zudem erfolgt ein Vergleich zwischen KI-basierten und klassischen Verfahren, um Potenziale beim KI-Einsatz auszuloten.

### 3.1 Anwendungsbereiche

Es gibt mittlerweile für alle Phasen der Wirkstoffentwicklung Beispiele aus der Forschung (Abbildung 3), wo KI effektiv unterstützen kann und die teilweise bereits auch in der Praxis konkrete Anwendungen finden. Der Einsatz von KI ändert am grundlegenden Ablauf der Wirkstoff-Entwicklung nichts, aber dieser kann Prozesse beschleunigen und die Chance auf eine erfolgreiche Entwicklung steigern.

Abbildung 3: KI-Anwendungsmöglichkeiten in den einzelnen Phasen der Wirkstoffentwicklung



<sup>6</sup> Repurposing bezeichnet die Umnutzung von bereits zugelassenen Arzneimitteln für andere Krankheitsbilder; dadurch wird die Entwicklung deutlich schneller und günstiger, da diese Stoffe direkt in die klinische Phase II eingehen können.

Firmen im Bereich der KI-basierten Wirkstoffentwicklung konzentrieren sich aktuell vor allem auf die frühen Phasen der Wirkstoffentwicklung, also Target- und Lead-Identifizierung. Weltweit bearbeiten 40 Prozent der KI-Start-ups in diesem Bereich das Screening von Molekülbibliotheken und 28 Prozent konzentrieren sich auf die Entdeckung neuer Targets, die Optimierung von Lead-Strukturen oder sogenanntes Repurposing<sup>6</sup> (Deloitte, 2019). Ebenso existieren bereits auch Anwendungen zur Nutzung von KI in der präklinischen und klinischen Phase sowie komplette End-to-End-Plattformen, welche den ganzen Entwicklungsprozess abdecken (Deep Pharma Intelligence, 2022).

Im Folgenden werden schematisch KI-Anwendungen für die einzelnen Entwicklungsphasen dargestellt. Dabei wird auch auf das Spektrum der Möglichkeiten zur KI-Nutzung aus Sicht der Forschung eingegangen. Soweit diese in den unterschiedlichen Phasen der Wirkstoffentwicklung bereits vorhanden sind, werden zusätzlich konkrete Beispiele für bereits in Anwendung befindliche KI-Systeme in der Wirkstoffentwicklung aufgezeigt.

### KI-Anwendung bei Auswahl des Wirkstoffziels

In der ersten Phase der Wirkstoffentwicklung, der Identifikation eines Targets, kann KI helfen, bestehendes Wissen zu sortieren und umfassender zu nutzen. Dies wird insbesondere relevant, da die Menge an verfügbarem medizinischem Wissen zunehmend wächst und für Menschen kaum mehr im Detail nachverfolgbar ist. PubMed, eine der größten Datenbanken für wissenschaftliche Literatur im Bereich Medizin, umfasst aktuell über 35 Millionen Artikel und wächst jeden Tag um mehr als 2.500 weitere Einträge (Van Meenen et al., 2022).

Im Rahmen des sogenannten Computer-aided drug design (CADD) wird KI in Form von Sprachmodellen und daraus abgeleiteten Chatbots (siehe Infobox: Sprachmodelle in der Medizin) genutzt, um automatisiert Informationen aus verfügbaren Datenbanken zum gewünschten Krankheitsbild zu sammeln und zu organisieren. Dies betrifft nicht nur Literatur-Datenbanken, sondern auch zum Beispiel Gen- oder Proteindatenbanken sowie die steigende Menge an Multiomics-Daten (Deep Pharma Intelligence, 2022). Omics-Daten beziehen sich dabei auf die Analyse der bei einer Krankheit beteiligten Gene (Genomics) und deren Übersetzung in Proteine (Transkriptomics), die daraus entstehenden Proteine (Proteomics) und die von diesen Proteinen verarbeiteten Stoffwechselprodukte (Metabolomics).

#### KURZINFO

##### Sprachmodelle in der Medizin

Sprachmodelle und Chatbots sind seit ChatGPT auch in der Öffentlichkeit angekommen und sie werden zunehmend für immer diversere Aufgaben eingesetzt. Dieses Potenzial wird auch in der Wirkstoffentwicklung realisiert, wobei hier, wie generell für Anwendungen im Gesundheitssektor, erhebliche Anpassungen der bekannten Modelle (ChatGPT, BARD, BLOOM) beziehungsweise Eigenentwicklungen notwendig sind, um die spezifischen Aufgaben bewältigen zu können (Löser et al., 2023).

##### Anwendungsbeispiele

##### Medizinisches Sprachmodell Med-PalM

Med-PalM ist ein von Google entwickeltes Sprachmodell speziell für medizinische Fragen. Dieses ermöglicht unter anderem die intuitive, textbasierte Abfrage von relevanten Genen



## KURZINFO

für bestimmte Krankheiten auf Basis von Informationen, die in Wissensgraphen organisiert sind. In ähnlicher Weise nutzt Insilico Medicine einen Chatbot als intuitive Schnittstelle zur firmeneigenen Forschungsplattform, wodurch abstrakte oder komplexe Datensätze intuitiver erfasst werden können.

**Chatbot des Pharmatechkonzerns Exscientia**

Als Gegenstück dazu hat Exscientia (Großbritannien) einen Chatbot zur Erstellung von Wissensgraphen ausgehend von Spracheingaben und der generativen Funktion von KI entwickelt. Die KI kann ebenso genutzt werden, um neue Moleküle oder Proteine zu erstellen, wenn das Sprachmodell auf Basis von Molekül- oder Proteindatenbanken trainiert wurde. Chemische Formeln oder die Aminosäuresequenz von Proteinen werden dazu in maschinenlesbare Formen überführt und können dann von der KI analog zu Text verarbeitet und als Trainingsgrundlage genutzt werden.

**Herausforderungen medizinischer Sprachmodelle: Halluzination**

Das Problem der Halluzination<sup>7</sup> von Sprachmodellen wird im medizinischen Kontext teilweise durch sogenanntes Evidence Surfacing kompensiert. Dabei wird das Statement der KI durch reale Zitate aus Publikationen oder Verweise auf Datenbankeinträge als Basis für eine Validitätsprüfung durch den Menschen erweitert (Savage, 2023).

Ausgehend davon können automatisiert Wissensgraphen für Krankheiten erstellt werden. Diese sind netzartige Darstellungsweisen aller Ursache-Wirkungs-Zusammenhänge und Wechselwirkungen im Körper auf molekularer Ebene im Zusammenhang mit einer Krankheit. Die Graphen können entweder von Menschen oder der KI zur Target-Identifikation genutzt werden (Fleming, 2018). Der dadurch entstehende Mehrwert wird vor allem bei komplexen Krankheitsbildern wie Krebs und neurodegenerativen Krankheiten relevant; diese betreffen oft eine Vielzahl biologischer Prozesse im Körper und sind nur mit Kombinationstherapien mehrerer Wirkstoffe effektiv behandelbar, was die intuitive Ableitung von Krankheitsmechanismen und Targets durch Menschen erschwert (Pun et al., 2023). Hier bietet KI großes Potenzial durch die gleichzeitige Betrachtung einer Vielzahl von potenziell auch multimodalen Parametern, welche die Krankheit beeinflussen.

In einem weiteren Schritt ist die KI-basierte Erstellung der 3D-Struktur des identifizierten Targets möglich. Diese 3D-Struktur ist entscheidend für die folgenden Entwicklungsschritte, um mögliche Angriffspunkte für Wirkstoffe bestimmen zu können. In diesem Kontext erlaubt das von DeepMind entwickelte Tool AlphaFold die 3D-Strukturvorhersage von Proteinen ausgehend von der in Datenbanken direkt zugänglichen zugrunde liegenden Aminosäuresequenz mit hoher Übereinstimmung zu experimentell bestimmten Proteinstrukturen. Die daraus abgeleitete AlphaFold-Datenbank enthält mittlerweile über 200 Millionen Proteinstrukturen (Jumper et al., 2021).

Neben der Nutzung und Aufbereitung von Daten aus bestehender Literatur bieten datengetriebene Ansätze zur Auswertung neu generierter experimenteller Daten ebenfalls die Möglichkeit zur Identifikation neuer Krankheitsmechanismen und damit potenzieller Targets. Dazu erfolgt beispielsweise eine vergleichende Analyse der Daten aus einer großen Anzahl von gesunden und von Krebs befallenen Zellproben, um neue Hypothesen zu Krebsmechanismen zu generieren (Fleming, 2018).

<sup>7</sup> Halluzination bezeichnet bei Sprachmodellen die Erzeugung falscher beziehungsweise erfundener Angaben infolge fehlender oder unvollständiger Informationen zu einem Thema in der Datenbasis des Sprachmodells. Diese Halluzinationen erscheinen trotzdem teilweise plausibel und sind daher je nach Kontext schwer als solche erkennbar.



## KI-Einsatz zur Auswahl des Wirkstoffziels

### Automatisierte Analyse von Datenbanken

Der luxemburgische Pharmakonzern Benevolent AI hat die Plattform BenAI Engine entwickelt, die Informationen aus mehr als 85 unterschiedlichen Arten von Datenquellen KI-basiert vereint und verknüpft. In Partnerschaft mit AstraZeneca konnten darauf aufbauend bereits Wirkstoffkandidaten gegen Lungenfibrose und eine chronische Nierenerkrankung identifiziert werden, die sich aktuell in der klinischen Untersuchung befinden (Deep Pharma Intelligence, 2022).

### Identifikation neuer Krankheitsmechanismen

Das in Oxford ansässige Pharmatechunternehmen Exscientia nutzt KI zur kombinierten Auswertung von genetischen Daten und wissenschaftlicher Literatur zur Vorhersage möglicher Wirkstoff-Targets. Der Einsatz dieser Technologie trug in Zusammenarbeit mit Evotec dazu bei, einen neuen Wirkstoffkandidaten gegen Krebserkrankungen mit soliden, fortgeschrittenen Tumoren innerhalb von nur acht Monaten zu identifizieren (Evotec, 2021). Dieser war 2021 der erste durch KI entwickelte Wirkstoff, der für klinische Studien zugelassen wurde.

Das amerikanische Unternehmen Genomenon für genomische Intelligenzforschung hat eine Software namens Prodigy entwickelt, die die KI-basierte Analyse und Verknüpfung von Fachliteratur, Bevölkerungsdatenbanken, theoretischen Vorhersagen und anderen Datenquellen bezogen auf Inhalte zu genetischen Ursachen von Krankheiten ermöglicht. Die Daten aus unterschiedlichen Quellen werden dazu mittels eines auf Genomdaten spezialisierten KI-Sprachmodells automatisch standardisiert und organisiert, um anschließend von Fachleuten begutachtet werden zu können. Für die Wilson-Krankheit, eine erbliche Stoffwechselstörung, konnte die Firma dadurch beispielsweise 869 für das Krankheitsbild relevante Genvarianten identifizieren und damit deutlich mehr als die bisher bekannten 235 Varianten (Deep Pharma Intelligence, 2022).

## KI-Anwendung bei Auswahl der Wirkstoffstruktur inklusive Repurposing

Ähnlich zur Target-Identifizierung unterstützt KI bei der Lead-Identifikation, indem sie verfügbare Datenbanken auswertet und damit nutzbar macht. Vor allem zu niedermolekularen Strukturen, die nach wie vor den Großteil der Wirkstoffe bilden, existieren umfangreiche Datenbanken zu relevanten Stoffeigenschaften wie Absorption, Diffusion, Verstoffwechslung oder Proteininteraktion (Gupta et al., 2021). PubChem, eine der größten Stoffdatenbanken, enthält aktuell Informationen zu mehr als 115 Millionen Verbindungen (NIH, 2023).

Alle diese Informationen können im virtuellen Screening genutzt werden, um potenzielle Wirkstoff-Kandidaten zu identifizieren. Die KI-Technologie hilft dabei, aus den bekannten Stoffeigenschaften Modelle für die Interaktion mit dem geplanten Target vorherzusagen. Dieses Verfahren wird als quantitative Struktur-Wirkungs-Beziehung (englischer Begriff: Quantitative Structure-Activity Relationship – QSAR) bezeichnet und wurde bereits ohne KI genutzt; die mittlerweile verfügbaren Datenmengen in Kombination mit der Auswertung durch KI erlauben allerdings deutlich präzisere Angaben zur Target-Lead-Interaktion und auch die beschleunigte Analyse einer größeren Menge von Verbindungen (Bhattamisra et al., 2023).

Ist die Target-Struktur bekannt und ausreichend charakterisiert, ermöglicht QSAR mit KI-Unterstützung die Vorhersage von idealen Wirkstoffstrukturen; diese Vorschläge können dann im nächsten Schritt zum Beispiel hinsichtlich Ähnlichkeit zu den bereits verfügbaren, kommerziellen Verbindungen oder Synthetisierbarkeit sortiert werden (Gupta et al., 2021). Mittels der beschriebenen Ansätze wird das Repurposing von bereits zugelassenen Wirkstoffen ebenfalls deutlich vereinfacht. Die bekannten Strukturen können dadurch virtuell und somit nicht experimentell auf Aktivität gegen neue Targets getestet und hinsichtlich notwendiger Strukturpassungen analysiert werden (Yang et al., 2021).

Neben der Nutzung verfügbarer Stoffdatenbanken im Rahmen von Screening-Ansätzen erlaubt generative KI auch die Erstellung komplett neuer Strukturen, deren Eigenschaften wiederum KI-basiert vorhergesagt werden können. Dadurch werden im Rahmen von sogenannten virtuellen Screening-Bibliotheken immer mehr Verbindungen zugänglich. Diese Bibliotheken bestehen aus Verbindungen, die noch nicht hergestellt wurden, aber synthetisch zugänglich sind und teilweise bereits bis zu 36 Milliarden Verbindungen umfassen (Lyu et al., 2023). Die hier genutzten generativen KI-Modelle arbeiten dabei, ähnlich zu generativen Sprachmodellen wie ChatGPT, auf Basis von Auto-Encodern. Als Input dienen bekannte chemische Strukturen in der SMILES-Darstellung (simplified molecular-input line-entry system). Dadurch können auch komplexe Strukturen in Form von Textzeilen dargestellt werden und sind damit maschinell lesbar (Bian & Xie, 2021).

## PRAXISBEISPIEL

### KI-Einsatz zur Auswahl der Wirkstoffstruktur

#### Generierung neuer Strukturvorschläge

Die Software Moler wurde von Novartis in Zusammenarbeit mit Microsoft entwickelt. Hier erfolgt die KI-basierte Generierung neuer möglicher Wirkstoffstrukturen auf Basis einer Auto-Encoder-Struktur, welche eine von Expertinnen und Experten definierte Bibliothek von möglichen Strukturbausteinen verwendet. Im Gegensatz zur Strukturgenerierung auf atomarer Ebene sind dadurch deutlich schnellere Vorhersagen möglich, welche gleichzeitig mit höherer Wahrscheinlichkeit einen funktionalen Wirkstoff darstellen (Microsoft, 2022).

#### Vorhersage von 3D-Strukturen

Die Software Molecue Pro von MindRank, einem asiatischen Pharmatechunternehmen, erlaubt die KI-basierte Generierung von Wirkstoffstrukturen ausgehend von der 3D-Struktur des Targets, wobei neben der Strukturvorhersage auch Eigenschaften zu Absorption, Distribution, Metabolisierung, Exkretion und Toxizität vorhergesagt werden. Mittlerweile umfasst die Entwicklungs-Pipeline der Firma auf Basis dieser Software bereits zwölf Verbindungen. Ein Wirkstoffkandidat gegen Typ-2-Diabetes konnte so in 18 Monaten bis zur Klinik (vs. circa 8 Jahre mit klassischen Verfahren) entwickelt werden und befindet sich aktuell in der klinischen Phase II (MindRank, 2023).

#### Repurposing

Die Firma Oncocross (Seoul) nutzt in ihrer Analyse-Plattform mit dem Namen Raptor KI-basierte Analyse von Transkriptom-Datenbanken ([siehe S. 15 Transkriptomics](#)) unter anderem, um neue Anwendungsmöglichkeiten für bestehende Arzneimittel zu identifizieren (Repurposing). Dadurch konnte die Firma mittlerweile eine Pipeline von 16 Wirkstoffkandidaten aufbauen, von denen sich mittlerweile vier Verbindungen bereits in klinischen Studien befinden (Deep Pharma Intelligence, 2022).

## KI-Anwendung zur Struktur-Optimierung

Die beschriebenen Methoden zur KI-basierten Analyse und Vorhersage von Stoffeigenschaften in der Lead-Identifizierung werden in der Phase der Lead-Struktur-Optimierung erweitert. Hier liefern *in silico*-Vorhersagen<sup>8</sup> weitere Angaben über Wirkstoffkandidaten zum Beispiel zur Bioaktivität und -verfügbarkeit, Verstoffwechslung und Ausscheidungswegen bezogen auf den menschlichen Stoffwechsel (Paul et al., 2020; Patel & Shah, 2022). Genauso existieren bereits Tools zur KI-gestützten Vorhersage von Nebenwirkungen durch Natural Language Processing-basierte Literaturanalysen (Gupta et al., 2021).

Neben der Vorhersage von physiologischen und pharmakologischen Eigenschaften ist bei der Struktur-Optimierung die Synthetisierbarkeit, also die Herstellbarkeit eines neuen Wirkstoffs mit bekannten Methoden der chemischen Synthese, entscheidend für die Umsetzbarkeit. Hier unterstützt KI ebenfalls durch die bereits erwähnte Analyse verfügbarer Datenbanken, aber auch durch die Planung möglicher Synthesewege. Dabei können beispielsweise auch unterschiedliche Ausgangsstoffe in Abhängigkeit von Preis und Verfügbarkeit miteingeplant werden. Die beschriebenen KI-Tools zur Lead-Optimierung beziehen sich vor allem auf die Analyse sogenannter niedermolekularer Verbindungen. Wie bereits in der Einleitung erwähnt, existieren aktuell für die Entwicklung hochmolekularer Verbindungen als Wirkstoffe aufgrund der höheren Komplexität ihrer Strukturen noch wenige KI-basierte Anwendungen. Derzeit machen niedermolekulare Verbindungen weltweit immer noch mehr als 50 Prozent des Arzneimittelmarkts aus (Statista, 2023). Allerdings wird der Bereich hochmolekularer Wirkstoffe in Zukunft mehr an Bedeutung gewinnen, unter anderem durch die Zunahme von personalisierten Therapien auf Basis von Antikörper-, Zell- und Gentherapien sowie einem weiteren Ausbau der Biotechnologie-Industrie (BCG, 2022b).

Die KI-unterstützte Identifikation von Wirkstoffziel und Wirkstoffstruktur sowie die intelligente Struktur-Optimierung beschleunigen einerseits die Entwicklung, da experimentelle Arbeiten reduziert und zielgerichteter geplant werden können. Andererseits steigen die Erfolgchancen der Entwicklung, da potenzielle Ausschlussgründe für Wirkstoffe, wie Off-Target-Aktivität oder Toxizität, durch Nutzung einer breiten Datenbasis bereits frühzeitig und umfassender als mit klassischen Methoden miteinbezogen werden können.

### PRAXISBEISPIEL

#### KI-Einsatz zur Struktur-Optimierung

##### Automatisierte Synthesepanung

Merck hat die Software Synthia entwickelt, um KI-gestützt Synthesestrategien für neue Wirkstoffkandidaten entwickeln und vergleichen zu können. Die Grundlage dazu bilden mehr als 100.000 handkodierte Reaktionsregeln. Ausgehend davon optimiert die KI durch Nutzung eines Scoring-Systems den Syntheseweg zum Beispiel hinsichtlich der notwendigen Anzahl an Syntheseschritten oder der Darstellbarkeit aus verfügbaren Ausgangsstoffen (Merck, 2023).

<sup>8</sup> *In silico* bezieht sich auf Experimente, die im Computer simuliert werden. Der Begriff ist an die in der Laborforschung gebräuchlichen Bezeichnungen *in vivo* (Versuche in lebenden Organismen, zum Beispiel in Mäusen) und *in vitro* (Versuche im Reagenzglas, zum Beispiel in isolierten Zellkulturen) angelehnt.

## KI-Einsatz zur Struktur-Optimierung

### Vorhersage von Stoffeigenschaften

Peptilogics (amerikanisches Biotechnologieunternehmen) nutzt ein generatives KI-Modell zur Identifizierung neuer Peptid-Wirkstoffe und im nächsten Schritt Vorhersage-Modelle zur Ableitung von Stoffeigenschaften der generierten Peptide hinsichtlich physikochemischer Eigenschaften, Pharmakokinetik, Entzündungshemmung etc. Die Entwicklungs-Pipeline umfasst mittlerweile sechs Wirkstoffkandidaten; ein Wirkstoff zur Behandlung bakterieller Infektionen befindet sich aktuell in der klinischen Phase II (Peptilogics, 2023). Bayer Pharma entwickelt und nutzt bereits seit 20 Jahren eine KI-unterstützte Plattform zur Vorhersage von Eigenschaften zu Pharmakokinetik und physikochemischen Eigenschaften neuer Wirkstoffkandidaten (Göller, A. H. et al., 2020).

## KI-Anwendung in den klinischen Phasen

Wie bereits erwähnt, bilden die klinischen Studien für neue Wirkstoffe eine der größten Herausforderungen und gleichzeitig den größten Kostenpunkt in der Entwicklung. Klinische Studien sind durchschnittlich für etwa 50 Prozent der gesamten Entwicklungskosten eines neuen Arzneimittels verantwortlich (Harrer, S. et al., 2019). Bereits die Rekrutierung einer ausreichend großen Gruppe von Probandinnen und Probanden ist oft herausfordernd; eine Auswertung von 114 Studien in Großbritannien zeigt, dass nur 31 Prozent der Studien den geplanten Anmeldestand erreichten. Nach Beginn der Studie kann eine unzureichende Beteiligung der Betroffenen ebenfalls zur Verlängerung beziehungsweise zum Scheitern der Studie führen (Fogel, 2018).

KI kann hier an mehreren Punkten eingesetzt werden, um die Erfolgchancen und die Aussagekraft von klinischen Studien zu steigern. Dies beginnt bereits bei der intelligenten Auswahl von Teilnehmenden vor Beginn der Studie. Im Sinne einer präziseren und personalisierten Medizin können KI-basiert die Genomprofile oder andere Biomarker analysiert werden, um sicherzustellen, dass potenzielle Probandinnen und Probanden dem Entwicklungsprofil des Wirkstoffs entsprechen. Ähnliches gilt für die Analyse der individuellen Stoffwechselmerkmale zur Anpassung der individuellen Dosierung. Durch die verbesserte Homogenität der Gruppenteilnehmenden steigen die Erfolgchancen der klinischen Studie und das Arzneimittel kann in der Folge besser auf die individuellen Bedarfe der Betroffenen abgestimmt werden (Sharma & Manchikanti, 2020).

Während der Studie kann KI-unterstütztes Monitoring von Probandinnen und Probanden dazu dienen, die Ausfallrate zu reduzieren und damit die Erfolgchancen der Studie zu steigern. Dazu können beispielsweise Wearables genutzt werden, um automatisch relevante Daten zu erfassen und KI-basiert in Echtzeit und vorausschauend auszuwerten. Dadurch können eventuelle Probleme hinsichtlich der für die Studie notwendigen Qualitätskriterien bereits frühzeitig erkannt und kompensiert werden (Bhattamisra et al., 2023). Diese Datenerfassung wird, wie bereits für die Erfassung innerhalb der Kliniken etabliert, der betroffenen Person offengelegt werden und erfolgt zustimmungsbasiert. Auf diese Art kann KI-basiert eine Vorhersage des potenziellen Drop-Out-Risikos von Teilnehmenden für ein bestimmtes Studiendesign erstellt werden, um darauf aufbauend Anpassungen im Design zum Beispiel hinsichtlich der Einnahmemenge oder -häufigkeit vorzunehmen (Harrer, S. et al., 2019).

Neben diesen Ansätzen zur Verbesserung von klassischen klinischen Studien könnte KI auch dazu genutzt werden, neue Studiendesigns zu ermöglichen. Im Kern steht dabei die Idee zur Nutzung von synthetischen also KI-generierten Daten von Probandinnen und Probanden im Rahmen klinischer Studien. Dabei werden ausgehend von verfügbaren medizinischen Daten klinische digitale Zwillinge der Teilnehmenden erstellt. Diese können dann zum Beispiel als virtueller Kontrollarm einer Studie dienen. Klassischerweise bestehen klinische Studien aus einer realen Test- und Kontrollgruppe als die beiden Arme einer Studie. Für eine Studie mit virtuellem Kontrollarm ist nur eine Gruppe von Probandinnen und Probanden notwendig, welche den Wirkstoff erhält. Die Ergebnisse der Kontrollgruppe werden KI-basiert ausgehend von Untersuchungsdaten der Teilnehmenden vorhergesagt, wodurch die notwendige Anzahl von Teilnehmenden deutlich reduziert werden kann (Strayhorn Jr., 2021). Dieser Ansatz kann potenziell auf komplett virtuelle klinische Untersuchungen erweitert werden. Dazu kann ein bestehender Datensatz verwendet oder künstlich erweitert werden, indem fehlende Informationen simuliert oder durch Variation definierter Merkmale mittels generativer KI zusätzliche, synthetische Probanden erzeugt werden. Hierzu gibt es bereits technische Ansätze, allerdings noch keine Vorgaben zur Validität solcher Daten von Seiten der Zulassungsbehörden (D'Amico et al., 2023). Gleichzeitig ist noch nicht absehbar, mit welcher Genauigkeit physiologische Prozesse in einem digitalen Zwilling abgebildet werden können.

## PRAXISBEISPIEL

### KI-Einsatz in den klinischen Phasen

#### Automatisierte Datenauswertung

Das Technologieunternehmen Saama (USA) bietet eine Software-as-a-Service-Plattform zur KI-gestützten Auswertung von Daten aus klinischen Studien an. Dabei können Daten aus unterschiedlichen Quellen (Labor, Klinik, Bildgebung, elektronische Patientenakten ...) verarbeitet, standardisiert und ausgewertet werden. Dadurch sind bei der Auswertung Zeiteinsparungen von mehr als 35 Prozent möglich (Saama, 2023).

Pfizer nutzte in der Entwicklung des Corona-Impfstoffs KI zur beschleunigten Datenauswertung einer klinischen Studie mit 44.000 Personen. Dadurch konnten innerhalb der generierten Datenmenge schneller und effizienter Signale von Nebenwirkungen und zum Nachweis der Wirksamkeit identifiziert werden (Sharma et al., 2022).

#### Reduktion der Drop-Out-Rate

Das amerikanische Medizintechnikunternehmen Strados Labs bietet einen bereits von der FDA zertifizierten Biosensor an. Dieser erfasst kontinuierlich Daten zu Husten- sowie Lungen- und Atemgeräuschen, wobei KI genutzt wird, um Störgeräusche aus den Daten zu entfernen und relevante Daten für die klinische Studie zu identifizieren. Für Erkrankungen der Lunge ist dadurch in klinischen Studien eine kontinuierliche Datenerfassung an Probandinnen und Probanden möglich sowie eine schnellere Auswertung der erhobenen Daten (Deep Pharma Intelligence, 2022).

## KI-Anwendungen bei der Zulassung

Klinische Studien werden von einer umfassenden Dokumentation begleitet, die für die Zulassung des Arzneimittels unabdingbar ist, um den medizinischen Nutzen sowie eventuelle Nebenwirkungen nachzuweisen. Die Konsolidierung dieser Dokumente für die Zulassung ist sehr zeitaufwändig; hier kann KI unterstützen und Teile des Dokumentationsprozesses übernehmen, indem beispielsweise Medical Writer bei der Erstellung von Studienberichten unterstützt werden. Dabei generiert eine KI auf Basis der verfügbaren Ergebnisse einen ersten Entwurf, der dann durch die Fachkraft geprüft und verfeinert werden kann (O’Shea, 2023; Spinner, 2022).

Ausgehend von der Zulassung erfolgt die Preisgestaltung und -verhandlung neuer Arzneimittel; dieser Prozess hängt, wie bereits beschrieben, von unterschiedlichen Variablen ab, weswegen es hier bereits Forschung zur Nutzung von KI für die beschleunigte Preisgestaltung durch automatische Generierung und Optimierung von Preisvorschlägen gibt (Mirza, 2020). Gleichzeitig ist hier noch unklar, wie der beidseitige Einsatz von KI durch Hersteller und Krankenkassen die Preisverhandlungen beeinflussen wird und ob hier zusätzliche Steuerungsinstrumente notwendig werden.

In der Produktion bietet KI, ähnlich zu anderen Industriezweigen, auch in der Pharmaindustrie breite Einsatzmöglichkeiten zur Effizienzsteigerung durch Optimierung des Produktionsdesigns und der Prozesskontrolle, zur vorausschauenden Wartung und Trendbeobachtung der Produktqualität auf Basis von Kundenbeschwerden oder bei Abweichungen im Produktionsablauf (FDA, 2023b). Hinzu kommen bei der Wirkstoffproduktion die ausführlichen Dokumentationspflichten im Rahmen des Qualitätsmanagements, die ebenfalls von der Automatisierung durch den KI-Einsatz profitieren. Auch in der Vermarktung von Arzneimitteln kann KI unterstützen, indem beispielsweise automatisch eine Roadmap zur Vorhersage der zentralen Treiber für Verkäufe in pharmazeutischen Firmen erstellt wird (Paul et al., 2020).

### PRAXISBEISPIEL

#### KI-Einsatz bei der Zulassung

KI kann vielerlei Unterstützungsmöglichkeiten liefern bei Prüfung und Erstellung von Zulassungsdokumenten:

##### **Unterstützung bei der Prüfung von Zulassungsdokumenten**

Die Europäische Arzneimittel-Agentur (EMA) hat Ende 2023 einen Arbeitsplan veröffentlicht, um die Möglichkeiten der KI für die eigene Produktivität zu nutzen, Prozesse und Systeme zu automatisieren, den Einblick in Daten zu verbessern und die Entscheidungsfindung zum Nutzen der Gesundheit von Mensch und Tier zu unterstützen (EMA, 2023a).

Die schweizerische Zulassungs- und Aufsichtsbehörde für Heilmittel Swissmedic hat ein Pilotprojekt zur Nutzung von KI, speziell zu maschinellem Lernen und Natural Language Processing (NLP), gestartet. Ziel ist es, die Anwendbarkeit dieser Technologien in der automatisierten Prüfung der Dokumentation zu klinischen Studien zu testen. Eine erste Auswertung zeigt, dass KI im Vergleich zur manuellen Prüfung mehr sicherheitsrelevante Signale in den Dokumenten in einem Bruchteil der Zeit identifiziert; gleichzeitig stellten sich 100 Prozent der durch KI identifizierten Signale in einer manuellen Nachprüfung als korrekt heraus (ICMRA, 2021).

## KI-Einsatz bei der Zulassung

### Unterstützung bei der Erstellung von Zulassungsdokumenten

Das französische Softwareunternehmen Yseop bietet eine KI-Lösung auf Basis von NLP und generativer KI an. Dadurch kann automatisiert der gesamte Inhalt von Dokumenten aus klinischen Studien erfasst werden. Im nächsten Schritt wird dieser Inhalt genutzt, um Entwürfe für Zulassungsdokumente und Berichte an Behörden zu erstellen und diese oft im Voraus automatisiert auszufüllen. Diese Berichte erfordern ohne KI oft mehrere Wochen in der Erarbeitung; entsprechend groß ist hier die mögliche Zeitersparnis durch KI (Yseop, 2023).

## KI-Anwendung für die Marktbeobachtung

Die im Kontext der Zulassung bereits angesprochenen Anwendungen von KI bei der Unterstützung und Auswertung großer Datenmengen bieten auch bei der Nachmarktbeobachtung beziehungsweise Pharmakovigilanz von Arzneimitteln große Möglichkeiten. Diese beziehen sich einerseits auf die Effizienzsteigerung bei bestehenden Prozessen zur Erhebung von unerwarteten Nebenwirkungen, andererseits in der verbesserten Auswertung der bestehenden Daten. Die Bearbeitung von Fallberichten im Rahmen der Nachmarktbeobachtung kostete große Pharma-Firmen 2021 durchschnittlich 86 Dollar pro Fall und allein in Europa wurden 2021 in der EudraVigilance-Datenbank der EMA 3,5 Millionen individuelle Fallberichte registriert (Murali et al., 2019; EMA, 2021). Dies zeigt, welches ökonomische Potenzial auch in diesem Bereich der Wirkstoffentwicklung durch Automatisierung von Prozessen mittels KI liegt, aber auch welche Datenmengen in diesem Kontext bereits jetzt schon erhoben werden. Deren intelligente Auswertung und Kombination mit anderen Datenquellen ermöglicht eine frühzeitigere Detektion oder Vorhersage von Mustern im Auftreten von Risiken, wovon die Sicherheit von Patientinnen und Patienten entscheidend profitiert (Bate & Stegmann, 2023). Über die Charakterisierung von Nebenwirkungen hinaus bieten sich weitere Möglichkeiten; in Kanada wird KI beispielsweise eingesetzt, um Gesundheitsprodukte mit falschen Werbeversprechen zu detektieren (ICMRA, 2021).

## 3.2 Vergleich mit klassischen Verfahren

Ausgehend von der Darstellung der klassischen Wirkstoffentwicklung in Kapitel 2 sowie der KI-unterstützten Wirkstoffentwicklung in Kapitel 3.1 werden im Folgenden diese beiden Vorgehensweisen miteinander verglichen, um die vielfältigen Potenziale von KI bei der Wirkstoffentwicklung aufzuzeigen (Abbildung 4).

**Anzahl zugänglicher Moleküle:** Grundsätzlich ermöglicht die Nutzung von KI den Zugang zu deutlich mehr Molekülen in den ersten drei Anfangsphasen der Entwicklung, sowohl bezogen auf Targets als auch auf mögliche Wirkstoffkandidaten (Abbildung 4). Bisher wurden erst knapp 700 nutzbare Protein-Targets für die Wirkstoffentwicklung im Menschen identifiziert, während ausgehend von KI-gestützten Vorhersagen potenziell mehr als 10.000 krankheitsrelevante Proteine durch Wirkstoffe beeinflussbar sind (Kana & Brylinski, 2019). Bezogen auf mögliche Wirkstoffstrukturen enthalten klassische Screening-Bibliotheken in pharmazeutischen Unternehmen nur einige Millionen Verbindungen, während KI teilweise auf die Daten mehrerer hundert Millionen Verbindungen für die Entwicklung zugreifen kann (Lyu et al., 2019) und virtuelle Screening-Bibliotheken sogar mehrere Milliarden Verbindungen enthalten können. Dabei ist allerdings zu bedenken,

dass diese Verbindungen teilweise nicht so umfassend charakterisiert sind wie in realen Bibliotheken. Wie bereits angesprochen können viele relevante Eigenschaften, wie Toxizität oder Wasserlöslichkeit, allerdings KI-basiert analysiert und mit einer gewissen Wahrscheinlichkeit vorhergesagt werden.

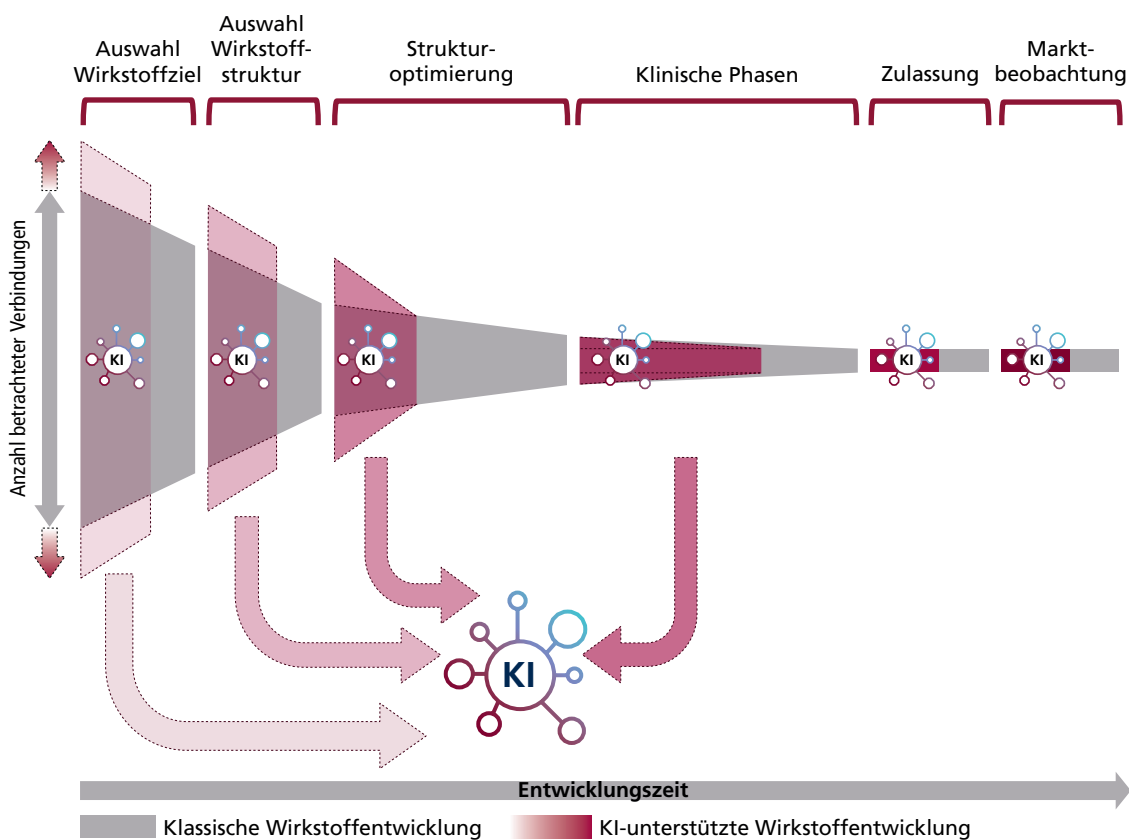
**Effizienz hinsichtlich Entwicklungsgeschwindigkeit und -kosten:** Neben den zusätzlich zugänglichen Molekülen erlaubt die Nutzung von KI eine deutlich effizientere und potenziell schnellere Wirkstoffentwicklung. Dies gilt für alle Entwicklungsphasen und wird in aktuellen Beispielen aus der Praxis deutlich (Abbildung 4). Viele im realen Experiment zeit- und arbeitsintensive Schritte, wie das Screening von Molekülbibliotheken, können simulativ ersetzt oder zumindest unterstützt werden. Die KI-basierte Vorhersage von Stoffeigenschaften erlaubt eine schnellere Eingrenzung auf relevante Zielstrukturen und bezieht gleichzeitig bereits für spätere Phasen relevante Eigenschaften mit ein, wie zum Beispiel die Synthetisierbarkeit oder pharmakologische Eigenschaften. Die KI-Unterstützung in der Entwicklung steigert durch die umfassende Vorhersage der Stoffeigenschaften die Chance der Übertragbarkeit von präklinischen Daten in die Klinik und senkt damit die Ausfallraten von Wirkstoffkandidaten in der Entwicklung. Zusammen mit einer intelligenten und gezielten Auswahl von Probandinnen und Probanden in den klinischen Phasen erlaubt KI dadurch eine effizientere und damit kostengünstigere Entwicklung mit höheren Erfolgchancen.

Ähnlich zum zugänglichen Raum an Molekülen ist hier allerdings wieder zu beachten, dass die reale Datengrundlage entscheidend ist für die Qualität der KI-basierten Vorhersagen. Dies bezieht sich sowohl auf die Daten zu Molekülen als auch auf die in Probandinnen und Probanden gemessenen Biomarker im Rahmen der klinischen Studien. Daneben ergeben sich beim Einsatz von KI neue Herausforderungen. Zwar sind im virtuellen Screening deutlich mehr Verbindungen schneller zugänglich, aber dadurch gestaltet sich auch das Design der Auswahlkriterien als Basis für die KI-basierte Optimierung für die gewünschte Zielstruktur schwieriger. Teilweise stehen Stoffeigenschaften von Wirkstoffstrukturen im Gegensatz zueinander, wie zum Beispiel Aktivität gegen das Target versus Wasserlöslichkeit. Beide sind allerdings essenziell für die Funktionalität des Wirkstoffs. Entsprechend liegt bei einem KI-basierten virtuellen Screening immer eine Abwägung vor; der Fokus der Optimierung auf wenige Kriterien birgt die Gefahr von nur bedingt nutzbaren Strukturen als Ergebnis; eine parallele Optimierung mehrerer Kriterien übergeht potenziell alternative oder neue Strukturmerkmale (Lowe, 2023).

**Datennutzung:** In der klassischen Wirkstoffentwicklung und dem bereits vorgestellten, sich verengenden Entwicklungstrichter entsteht in der Entwicklung im Schnitt aus 10.000 Verbindungen zu Beginn ein zugelassenes Arzneimittel. Die anderen Verbindungen, die Entwicklungskriterien wie zum Beispiel Aktivität gegen das Target oder Wasserlöslichkeit nicht erfüllen und daher in der Entwicklung nicht weiterverfolgt werden, erzeugen ebenfalls valide Daten (Abbildung 4, gestrichelt umrahmte rote Bereiche). Durch Einbindung von KI in den Entwicklungsprozess könnten die Daten dieser Verbindungen nutzenbringend für die Verbesserung des Modells durch eine Erweiterung der Datenbasis zum Training genutzt werden. Dies gilt generell für alle experimentell erhobenen Daten, auch aus erfolglosen Entwicklungsprojekten.



**Abbildung 4: Schematische Darstellung der Vorteile des KI-Einsatzes in der Wirkstoffentwicklung über alle Entwicklungsphasen hinweg: kürzere Entwicklungsphasen, mehr zugängliche Moleküle sowie Nutzbarmachung aller generierten Daten zum Training der KI**



Quelle: Eigene Darstellung.

**Personalisierung von Therapien:** Neben der bereits angesprochenen intelligenten Auswahl von Probandinnen und Probanden in klinischen Studien durch KI-basierte Analyse von Biomarkern bildet diese Analyse zugleich auch die Basis für die Entwicklung personalisierter Therapien. Diese können dadurch konkreter auf das Krankheitsbild der betroffenen Person abgestimmt werden, dies auch im Hinblick auf gleichzeitig vorliegende Erkrankungen oder die medizinische Historie der Person. Zusätzlich kann KI einen Beitrag dazu leisten, die für die personalisierte Medizin aufwändigere Produktion und Entwicklung effizienter zu gestalten und damit wirtschaftlich in der Breite leichter umsetzbar zu machen. Aktuelle Ansätze von personalisierter Medizin beziehen sich im Gegensatz dazu noch meist auf Einzelfälle und sind mit sehr hohen Kosten für die jeweilige Behandlung verbunden.

**Infrastruktur:** Die klassische Wirkstoffentwicklung ist vor allem auf ausreichende Laborkapazitäten zur Durchführung von Screenings und praktischen Untersuchungen angewiesen. Durch eine verstärkte KI-Nutzung verschiebt sich der Fokus vermehrt auf die digitale Infrastruktur und Rechenkapazitäten sowie entsprechend rares Fachpersonal, wie zum Beispiel Data Scientists. Die Virtualisierung der Entwicklungsschritte erlaubt so potenziell die Reduktion notwendiger Laborflächen und bietet mehr Möglichkeiten für eine dezentrale Wirkstoffentwicklung. Gleichzeitig wird dadurch die Zusammenarbeit zwischen Unternehmen gefördert: Denn im Gegensatz zu realen Molekülen können Daten leichter ausgetauscht und bezüglich des gewünschten Detailgrades angepasst werden. Der Fokus auf Daten zu Moleküleigenschaften ermöglicht kooperierenden Unternehmen eine erweiterte Datenbasis. Dadurch kann das Training der internen KI-Modelle verbessert werden, ohne notwendigerweise wettbewerbsrelevante Informationen teilen zu müssen (siehe hierzu Kapitel 5).

## 4 Anwendungsbeispiele mit KI

---

Die Vorteile von KI in der Wirkstoffentwicklung werden im Folgenden an drei konkreten Beispielen verdeutlicht, um ihren Einsatz, aber auch die Anwendungsbreite abzubilden. Das erste Beispiel behandelt KI in der Krebstherapie zur Personalisierung von Behandlungen, da Krebs eine der häufigsten Erkrankungen und Todesursachen in Deutschland ist. Das zweite Beispiel zeigt die Entwicklung eines Wirkstoffs bis zu Phase II auf, wobei erstmals umfassend in allen Entwicklungsschritten KI eingesetzt wurde. Das dritte Beispiel verdeutlicht, welchen Mehrwert KI aus bereits verfügbaren Daten generieren kann, hier beim Einsatz in retrospektiven klinischen Studien.

### 4.1 KI in der personalisierten Krebstherapie

Die Entwicklung von Krebstherapien gestaltet sich aufgrund der Diversität von Krebserkrankungen äußerst schwierig. Diese Diversität und damit Individualität der Erkrankung entsteht aus den unterschiedlichen Umwelteinflüssen und genetischen Veranlagungen sowie dem Wechselspiel des Immunsystems der jeweiligen Person mit der Krebserkrankung. Dieses kann aber auch dazu genutzt werden, um individuelle und damit hochwirksame Therapien zu entwickeln. Einen möglichen Ansatz bildet hier die Neoantigen-basierte individualisierte Immuntherapie. Dabei wird das körpereigene Immunsystem dahingehend unterstützt, Krebszellen gezielt zu erkennen und abzutöten. Die Erkennung als Basis für eine Immunantwort funktioniert über sogenannte Neoantigene; damit werden Moleküle bezeichnet, die bei Tumorzellen im Gegensatz zu gesunden Zellen verändert sind, an der Zelloberfläche präsentiert werden und in der Lage sind eine Immunreaktion auszulösen. Die Präsentation erfolgt in Form von Proteinkomplexen, sogenannten pMHC-Strukturen (peptide-major histocompatibility complex). Die in Krebszellen vorhandenen Mutationen können zur Bildung von für die Krebszellen spezifischen Neoantigenen führen (Xie et al., 2023).

Auf Basis von Sequenzierdaten und Software, die unter anderem trainierte KI-Modelle nutzt, können final Neoantigenkandidaten als Basis für die Immuntherapie vorhergesagt werden. Dazu forscht unter anderem die Firma Biontech an diversen KI-Modellen zur Bewertung der Neoantigenkandidaten. Eine Möglichkeit könnte hier die Miteinbeziehung von KI-basierten Vorhersagen zu pMHC-Strukturen sein. Im Vergleich zu großen vortrainierten Modellen wie alphaFold können für diesen Zweck speziell entwickelte KI-Modelle pMHC-Strukturen schneller mit ähnlicher Genauigkeit vorhersagen und benötigen deshalb nur einen Bruchteil der Anzahl von Parametern für die Entwicklung. Ausgehend von den Informationen zu Neoantigenkandidaten können für die Immuntherapie beispielsweise Impfungen gegen die Krebserkrankung entwickelt werden. Krebsimpfstoffe zielen darauf ab, die Antwort des Immunsystems auf ein bestimmtes Neoantigen zu verstärken. Dazu werden der betroffenen Person Moleküle (z. B. RNA oder Peptide) injiziert, welche die Mutationen enthalten, die im Tumorgewebe als Neoantigenkandidaten vorhergesagt wurden. Der personalisierte Behandlungsansatz steigert durch den angepassten Wirkstoff die Chancen auf eine erfolgreiche Behandlung deutlich. Der KI-Einsatz trägt durch die Beschleunigung der Analyse-Prozesse auch dazu bei, personalisierte Wirkstoffe in der Breite wirtschaftlich tragfähig zu machen.

### 4.2 Nutzung von KI zur Lead-Identifikation und -Optimierung

Als einer der ersten erfolgreichen KI-Einsätze in der Wirkstoffentwicklung wird hier die Entwicklung eines Wirkstoffkandidaten der Firma Insilico Medicine gegen Fibrose vorgestellt, bei dem der Prozess von der Target- und Lead-Identifizierung über Lead-Optimierung bis hin zur Vorklinik in nur 46 Tagen durchlaufen wur-

de; in der Regel dauert dieser Prozess zwei bis drei Jahre (Knapp, 2019). Ziel war es, ein Arzneimittel gegen Lungenfibrose zu entwickeln, eine Erkrankung, die zu einer krankhaften Vermehrung des Bindegewebes in menschlichem Gewebe und Organen führt. Für diese Erkrankung existieren bereits bekannte Targets, wodurch automatisiert mittels Natural Language Processing von wissenschaftlicher Literatur ein Rezeptorprotein als Target identifiziert werden konnte: der Discoidin Domänen Rezeptor 1 (DDR1) (Chace, 2022).

Ausgehend davon wurde KI-basiert eine chemische Struktur hinsichtlich folgender Merkmale entwickelt und optimiert: Synthetisierbarkeit mit bekannten Methoden der chemischen Laborsynthese, Effektivität gegen das Target sowie Unterscheidung zu bestehenden Molekülen in Literatur und Patenten. Als Datengrundlage für die Moleküsentwicklung dient eine Datenbank zu Strukturen kommerziell erhältlicher chemischer Moleküle, Strukturen bekannter DDR1-Inhibitoren, Inhibitoren gegen die gleiche Klasse von Enzymen als Positivkontrolle, Inhibitoren gegen andere Enzymklassen als Negativkontrolle sowie Daten von Patientinnen und Patienten zu biologisch aktiven Molekülen. Dadurch entstanden innerhalb von nur 46 Tagen sechs Verbindungen als Basis für vorklinische Untersuchungen. Die Syntheseplanung erfolgte zwar manuell, konnte aber beschleunigt erfolgen, da die Moleküle bereits von der KI hinsichtlich Synthetisierbarkeit vorgefiltert wurden (Zhavoronkov et al., 2019). Damit wurde erstmals mit einem durch den beschriebenen KI-Einsatz entwickelten Wirkstoff die klinische Phase I erfolgreich beendet und die klinische Phase II soll bis Mitte 2024 abgeschlossen sein (Colangelo, 2023).

Die Firma Insilico Medicine bietet zudem unter dem Namen Pharma.AI eine umfassende End-to-End-Plattform zur KI-basierten Wirkstoffentwicklung an. Dabei liefern umfassende Omics-Daten die Grundlage für die Identifikation relevanter Targets. Ausgehend davon erfolgt KI-basiert eine Rekonstruktion des Wirkmechanismus der Krankheit und eine Priorisierung und Filterung relevanter Targets. Im nächsten Schritt können dann potenzielle Leadstrukturen entwickelt und optimiert werden. Bei den klinischen Studien unterstützt KI die Dokumentation und Auswertung der klinischen Daten (Insilico Medicine, 2023).

### 4.3 Nutzung von KI in retrospektiven Real-World-Studien

Das dritte Anwendungsbeispiel zeigt eine Anwendung im Bereich der klinischen Bewertung von Arzneimitteln mittels retrospektiver Studien auf Basis von realen Gesundheitsdaten, die in großen Mengen vorhanden sind, sogenannte Real-World-Daten, auf. Im Unterschied zu prospektiven Studien, bei denen neue Daten zu einem Wirkstoff unter kontrollierten klinischen Bedingungen erzeugt werden, nutzen retrospektive Studien bereits verfügbare Daten. Diese werden in der Rückschau ausgewertet, um die Wirkung und/oder die Nebenwirkungen eines Wirkstoffs zu bewerten. Im Gegensatz zu prospektiven Studien können so unterschiedliche, bereits verfügbare Datenquellen (wie GKV-Abrechnungsdaten und Registerdaten) verwendet werden. Auf diese Weise können mehr und auch solche Daten verwendet werden, die die Realversorgung abbilden. Der Einsatz retrospektiver Studien bietet sich vor allem im Rahmen der Marktbeobachtung (Phase IV) oder bei der Analyse von Versorgungsinnovationen auf spezielle Bevölkerungsgruppen an, die in klinischen Studien schwer erfassbar sind, wie beispielsweise hochaltrige, minderjährige oder kognitiv eingeschränkte Personen.

Eine Herausforderung bei retrospektiven Studien bildet allerdings die Auswahl der Vergleichsgruppe (sog. Kontrollgruppe), welche im Unterschied zur Interventionsgruppe keine Behandlung mit dem zu bewertenden Wirkstoff erhalten hat. Diese ist notwendig, um die Effekte der Behandlung im Vergleich zur Standard-Therapie messbar zu machen. Die beiden Gruppen sollen dabei so ähnlich wie möglich sein. Typischerweise werden dafür Ähnlichkeits-(Propensity)-Modelle mit relevanten demographischen und Morbiditäts-Parametern genutzt, die durch Expertinnen und Experten oder mittels Regressionsanalyse ausgewählt werden. Auf

Grundlage des modellierten Ähnlichkeitsscores wird dabei die Kontrollgruppe so festgelegt, dass möglichst vergleichbare Paare (Matched Pairs) aus Interventionsteilnehmenden und Kontrollpartnern entstehen. KI kommt hier zum Einsatz, um das Propensity-Modell und damit auch die Vergleichbarkeit hinsichtlich des zu bewertenden Zielparameters (Outcome-Risiko) zu verbessern und so die Aussagekraft von retrospektiven Real-World-Studien zu erhöhen. Bezogen auf die Wirkstoffentwicklung ergeben sich hier große Möglichkeiten, da beispielsweise das Problem der geringen Studienpartizipation in Deutschland oder die bürokratischen Aufwände im Zusammenhang mit der Studienbeantragung reduziert werden können. Gleichzeitig können kontinuierlich Real-World-Daten aus der Realversorgung mit ausreichender Qualität als Basis für die Gewinnung von Langzeit-Real-World-Evidence (RWE) genutzt werden. Unter RWE versteht man den klinischen Nachweis der Wirksamkeit oder Nebenwirkungen einer Behandlung ausgehend von Real-World-Daten, die nicht im Rahmen von kontrollierten klinischen Studien, sondern in der realen Gesundheitsversorgung erhoben werden.

Im vorliegenden Beispiel wurde im deutsch-kanadischen [AIR\\_PTE-Projekt](#) ein sogenannter Rapid Evidence Generator (REG) entwickelt, um durch KI-Methoden die relevanten Risikoparameter für ein optimales Ähnlichkeitsmodell zu identifizieren. Im Unterschied zu klassischen Verfahren wird dabei sowohl die Wahrscheinlichkeit der Behandlung als auch die Wahrscheinlichkeit der zu bewertenden Behandlungsergebnisse (Outcome) modelliert. Im nächsten Schritt wird dann durch Anwendung des Modells ein risikoadjustiertes Ähnlichkeitsmaß für jede Teilnehmenden-ID erzeugt und mittels Matching eine risikoadjustierte Kontrollgruppe identifiziert. Der REG wurde unter anderem genutzt, um retrospektiv das Blutungsrisiko bei der Verwendung von zwei unterschiedlichen Gerinnungshemmern im Rahmen der Behandlung von tiefer Venenthrombose (VTE) zu vergleichen. Dabei kamen regelhaft verfügbare Abrechnungsdaten zum Einsatz und es konnte eine Studienpopulation von 38.627 Versicherten ausgewertet werden. In das KI-basierte risikoadjustierte Propensity Score Matching (RAPSM) wurde neben Demografie, Morbidität und Therapie auch das Blutungsrisiko einbezogen. Als Ergebnis konnte eine signifikant niedrigere Häufigkeit von Blutungen im Vergleich zur Heparin-Therapie als Folge der VTE-Behandlung für die Nutzung von Plättchenhemmern ermittelt werden (Zahn et al., 2023).

## 5 Potenziale und Herausforderungen

---

Die vorhergehenden Abschnitte machen die Notwendigkeit von Innovationen – auch mit KI – in der Wirkstoffentwicklung und die vielfältigen Anwendungsmöglichkeiten von KI in diesem Bereich deutlich. Die daraus entstehenden Potenziale, aber auch die Herausforderungen für den KI-Einsatz in der Wirkstoffforschung werden im Folgenden diskutiert.

### Wirtschaftliche Umsetzbarkeit

Die großen Potenziale von KI in der Wirkstoffentwicklung, die teilweise bereits realisiert werden, betreffen vor allem die Reduktion der Zykluszeiten sowie Entwicklungskosten für Arzneimittel, die Generierung neuer Hypothesen durch erweiterte Datennutzung und damit einhergehende Veränderungen in der Forschungskultur (Schneider et al., 2020). In Kombination profitiert davon die Bevölkerung von mehr und schnelleren Innovationen im Bereich der Wirkstoffe, günstigeren Arzneimittelpreisen sowie mehr Wirkstoffentwicklungen für seltene Erkrankungen.

Für Unternehmen bieten sich Möglichkeiten für neue Geschäftsmodelle sowie zum Aufbau eines digitalen Ökosystems zur gemeinschaftlichen Datennutzung. Dies setzt allerdings die Bereitschaft und Möglichkeit zur Datenteilung voraus im Rahmen eines Systems mit aktiver Koopetition<sup>9</sup>. Der verstärkte Einsatz von KI macht Daten aber potenziell auch zum Wettbewerbsvorteil, was der Datenteilung entgegensteht. Hier könnte der Austausch von Metadaten einen Mittelweg darstellen, indem beispielsweise nicht die tatsächlichen Strukturen der jeweiligen Wirkstoffe geteilt werden, sondern nur die chemischen Unterschiede zwischen Molekülen (Schneider et al., 2020). In klinischen Studien könnten beispielsweise die Daten der Placebo-Gruppe in anonymisierter und aggregierter Form als Basis für die Erstellung digitaler Zwillinge frei geteilt werden, da es sich hier um qualitativ hochwertige Daten handelt, die keine Informationen über den untersuchten Wirkstoff enthalten.

Insgesamt profitieren die beteiligten Unternehmen dabei von Effizienzgewinnen in der Entwicklung aufgrund einer erweiterten Datenbasis zum Training der firmeneigenen KI-Modelle beziehungsweise durch den Zugang zu besseren KI-Modellen, ohne wettbewerbsrelevante Daten weitergeben zu müssen. Ein Beispiel für eine solche vorwettbewerbliche Zusammenarbeit fand beispielsweise bereits im [MELLODDY-Projekt](#) statt, bei dem zehn Pharmaunternehmen in Zusammenarbeit mit Zulassungsbehörden über drei Jahre hinweg ein Machine-Learning-Modell zur verbesserten Vorhersage der Eigenschaften von chemischen Verbindungen trainiert haben (Mock et al., 2023).

Neben den beschriebenen Herausforderungen bei der Datenteilung im Kontext des wirtschaftlichen Wettbewerbs stellt die Implementierung von KI in den Betrieben eine Herausforderung dar, die auch in anderen Bereichen beobachtet wird (Stowasser et al., 2020). Dies betrifft die Verfügbarkeit von qualifiziertem Personal, mangelnde Flexibilität zur Einbringung neuer Technologien in etablierte Abläufe oder ethische Bedenken in Bezug auf den Verlust von Arbeitsplätzen (Paul et al., 2020; Patel & Shah, 2022; Savage, 2021).

---

<sup>9</sup> Koopetition meint die Kooperation von Unternehmen in bestimmten Bereichen wie zum Beispiel Forschung, obwohl diese im direkten Wettbewerb zueinander stehen.

## Datenverfügbarkeit und -qualität

Die Datenteilung zwischen Unternehmen wurde bereits angesprochen, aber die Bereitstellung von qualitativ hochwertigen Daten durch die Bevölkerung zum Beispiel aus der elektronischen Patientenakte (ePA) ist noch wichtiger, um den Mehrwert für KI in der Wirkstoffentwicklung auszuschöpfen. Während Daten zu Moleküleigenschaften bereits heute gut erfassbar und in entsprechenden Datenbanken (PubChem, AlphaFold Protein Structure Database ...) frei zugänglich sind, gibt es aufgrund der damit einhergehenden Komplexität noch Lücken in der Datenbasis zur menschlichen Biologie, zum Beispiel in Bezug auf Krankheitsmechanismen und die Wirkung von Arzneimitteln (Schneider et al., 2020). Diese können durch mehr Daten von Patientinnen und Patienten und KI-unterstützte Auswertung potenziell geschlossen werden, indem beispielsweise Genomanalysen ausgeweitet werden oder bei der Nachmarktbeobachtung von Arzneimitteln KI-basiert Real-World-Evidence (RWE) generiert wird.

Genomanalysen erlauben Vorhersagen zur personenspezifischen Wirksamkeit und Metabolisierung von Arzneimitteln. Hier bestehen auch große Potenziale zur kooperativen Datennutzung, zum Beispiel indem Genomprofile von den Krankenkassen erhoben und kassenübergreifend genutzt werden. Dadurch könnten Effekte wie in den USA vermieden werden, wo Genomanalysen bereits verbreiteter sind, aber nur wenige Unternehmen über die Genomdaten in ihren firmeneigenen Datenbanken verfügen.

Durch die Nutzung von RWE können ein selektiver Mehrwert für bestimmte Untergruppen der Bevölkerung und Nebenwirkungen von Arzneimitteln erfasst und dadurch effektivere Behandlungsempfehlungen abgeleitet werden. Je breiter und kontinuierlicher die verfügbare Datenbasis ist, umso mehr Möglichkeiten bieten sich für retrospektive Studien oder Prädiktionsmodelle. Allerdings braucht es dazu auch entsprechende Standards und Protokolle für diese Arten von Daten zur Sicherstellung der Datenqualität.

Genauso ist eine aktive Beteiligung der gesamten Bevölkerung entscheidend, um Unterschiede in der menschlichen Biologie abzubilden und dadurch eine ausgewogene Datenbasis sicherzustellen. Verzerrungen in der Datenbasis könnten anderenfalls zu ungleichem Zugang zu Behandlungen für bestimmte Bevölkerungsgruppen führen, wenn die Wirksamkeit nur für eine Teilmenge der Bevölkerung optimiert wird (Gupta et al., 2021).

So wurden in der Vergangenheit bei der Wirkstoffentwicklung klinische Studien vor allem mit männlichen Probanden durchgeführt, wodurch bestimmte für Frauen spezifische Nebenwirkungen nicht oder nicht ausreichend charakterisiert wurden. KI kann hier zwar durch Vorhersagen von Wirkungen und Nebenwirkungen auf bestimmte Gruppen unterstützen, braucht dazu aber eine ausgewogene und für alle Gruppen ausreichend große Datenbasis (Chandak & Tatonetti, 2020). Die bereits angesprochenen Möglichkeiten für individualisierte und dadurch optimierte Therapieansätze erfordern daher die Bereitstellung individueller Daten der Patientinnen und Patienten. Dies ist insbesondere bei seltenen Erkrankungen entscheidend, wo ohnehin nur eine beschränkte Datenbasis vorhanden ist.

## Regulatorischer Rahmen

Im Kern der regulatorischen Rahmenbedingungen für Arzneimittel steht die Gewährleistung der Sicherheit im Verkehr mit Arzneimitteln. KI bietet hier potenziell große Vorteile in der Wirkstoffentwicklung; einerseits werden ab Beginn des Entwicklungsprozesses neben bereits genutzten Informationen auch KI-basierte Prognosen beispielsweise zu Toxizität und Off-Target-Aktivität ausgehend von einer breiten Datenbasis miteinbezogen, andererseits kann die intelligente Auswahl von geeigneten Probandinnen und Probanden durch ein

erweitertes Set an Biomarkern sowohl in klinischen Studien als auch in der späteren Anwendung zu einem geringeren Risiko von unerwarteten Nebenwirkungen führen aufgrund des spezifischeren Anwendungsprofils für neue Wirkstoffe.

Unabhängig von der Art der Regulierung ist allerdings zu beachten, dass Algorithmen keine kausalen Zusammenhänge aufzeigen, was die Fehleranalyse und auch die Bewertung von Forschungsdaten durch die Zulassungsbehörden erschwert (Panch et al., 2022). Aufgrund des in der KI-Entwicklung bekannten Blackbox-Problems herrscht teilweise auch noch Skepsis innerhalb der Pharmaindustrie, was den umfassenden Einsatz von KI in der Wirkstoffentwicklung betrifft (Paul et al., 2020).

KI kann unabhängig davon dazu beitragen, die Entwicklungszyklen für Wirkstoffe stark zu beschleunigen; dadurch entstehen potenziell mehr Zulassungsanträge für neue Arzneimittel in kürzerer Zeit als bisher. Im Gegensatz zur klassischen Wirkstoffentwicklung mit definiertem Studiendesign können sich KI-Systeme darüber hinaus im Entwicklungsprozess verändern, wenn beispielsweise die Daten aus den frühen klinischen Phasen genutzt werden, um das Modell für die Auswahl von Probandinnen und Probanden in der klinischen Phase III zu verbessern. Diese Veränderung der KI-Modelle während der Nutzung stellt eine Herausforderung für Regulatorik und Zulassungsbehörden dar. KI kann hier aber auch auf Seiten der Behörden dazu beitragen, Prozesse zu beschleunigen und zu automatisieren, um mit dem gesteigerten Entwicklungstempo Schritt zu halten.

In Analogie zu den in der Pharmaindustrie bereits etablierten Regeln zur guten Herstellungspraxis (GMP) wurden von der FDA bereits Regeln zur guten Nutzung von Machine Learning aufgestellt, die auf den Einsatz von KI in der Wirkstoffentwicklung ausgeweitet werden können. Sowohl die FDA als auch die EMA betonen in aktuellen Papieren die Chancen durch KI in der Wirkstoffentwicklung, weisen aber auch auf die Herausforderungen aufgrund der hohen Entwicklungsgeschwindigkeit in diesem Feld hin. Beide Organisationen haben daher 2023 einen Prozess zum Austausch mit relevanten Stakeholdern auf Basis von Diskussionspapieren initiiert, um auszuloten, wie aktuelle Entwicklungen im Bereich KI genutzt und im Rahmen der Zulassung adäquat berücksichtigt werden können (EMA, 2023b; FDA, 2023b).

Die EMA nennt die Nutzung von KI in der Entwicklung und Zulassung von Arzneimitteln als eines der Ziele in ihren strategischen Betrachtungen zur Regulierungswissenschaft bis 2025 (EMA, 2020). Bezüglich der Bewertung von KI orientiert sich die EMA grundsätzlich an dem risikobasierten Ansatz aus der Verordnung der EU-Kommission zum AI Act; gleichzeitig werden auch Anwendungsszenarien aufgezeigt, welche die Nutzung synthetischer Daten oder von nicht oder nur bedingt erklärbaren KI-Modellen rechtfertigen können. Ein Konsultationsprozess im Jahr 2023 diente als Basis für die Erarbeitung neuer Leitlinien und eventueller Anpassungen aktueller Regelungen (EMA, 2023b). Die Europäische Arzneimittel-Agentur (EMA) hat Ende 2023 einen Arbeitsplan veröffentlicht, um die Möglichkeiten von KI für die eigene Produktivität zu nutzen, Prozesse und Systeme zu automatisieren, den Einblick in Daten zu verbessern und die Entscheidungsfindung zum Nutzen der Gesundheit von Mensch und Tier zu unterstützen (EMA, 2023a).

## 6 Offene Fragen und Gestaltungsoptionen

---

Vielfältige Potenziale von KI in der Wirkstoffentwicklung sind zu benennen, die bereits in konkreten Anwendungen realisiert werden. Zwar ergeben sich noch offene Fragen, aber zugleich auch konkrete Gestaltungsoptionen zur Förderung des KI-Einsatzes in der deutschen Pharma- und Biotechnologie-Industrie.

### Was bedeutet die Nutzung von KI in der Wirkstoffentwicklung für die Zulassung und Erstattung von Arzneimitteln?

KI kann potenziell die Entwicklungszyklen von Arzneimitteln stark verkürzen und die Anwendung von personalisierten Therapien vor allem in der Breite fördern. Die zugrunde liegenden Daten entstehen dabei nicht mehr nur durch klassische Studien, sondern durch KI-basierte Vorhersagen oder durch die Auswertung synthetisch generierter Daten. Beispielsweise ermöglicht KI die Verkürzung klinischer Phasen, indem Surrogat-Parameter zur KI-basierten Vorhersage der klinischen Endpunkte genutzt werden und so deutlich schneller Aussagen zur Wirksamkeit und zu Nebenwirkungen getroffen werden können. Zusätzlich könnten die aktuell noch im Rahmen der Präklinik vorgeschriebenen Tierversuche durch die Kombination von Organ-on-a-Chip-Ansätzen<sup>10</sup> mit KI-basierter Auswertung reduziert oder sogar vermieden werden.

Die Entwicklungen bezüglich der Kombination von KI mit neuartigen Ansätzen aus der Labortechnik als Surrogat-Systeme für klinische Untersuchungen müssen sowohl in der Zulassung als auch in der Erstattung berücksichtigt werden. KI-Modelle, die zur Datenauswertung von Surrogat-Ansätzen genutzt werden, sollten daher idealerweise nachvollziehbar sein, um die im Rahmen der Entwicklung gewonnenen Erkenntnisse sowie die Evidenzbasis nachvollziehbar und prüfbar zu machen. Einen wichtigen Faktor bildet dabei eine differenzierte Risikobewertung, zum Beispiel bezogen auf die Datenauswertung in unterschiedlichen Entwicklungsphasen: Während die KI hinter Strukturvorschlägen nur bedingt nachvollziehbar sein muss, sollten KI-basierte Auswertungen und Vorhersagen zu medizinischen Aspekten klar belegbar sein.

Gleichzeitig sind von Seiten der Behörden verbindliche Standards für die Prüfung und Gültigkeit von KI-basierten Daten im Rahmen der Zulassung von Wirkstoffen gefordert. Diese Regelungen sind notwendig, um den Mehrwert durch KI zu maximieren und KI-basierte Erkenntnisse als eigenständige Datengrundlage, die nicht pauschal durch reale Experimente gestützt werden muss, zu etablieren.

Alternative Modelle zur Regulierung können zudem Antragsverfahren entzerren, indem die medizinische Bewertung der Wirkstoffe zentral bei der jeweiligen nationalen Zulassungsbehörde erfolgt, dagegen die technische Bewertung der eingesetzten und sich verändernden KI durch unterschiedliche zertifizierte Akteure erfolgen kann (zum Beispiel technische Prüfdienstleister). Eine weitere Möglichkeit wäre die Errichtung einer eigenen Abteilung innerhalb der Zulassungsbehörden für KI zu medizinischen Fragestellungen (Panch et al., 2022). Diese Abteilung könnte alternativ auch auf europäischer Ebene in der EMA angesiedelt sein, um eine europaweite Vereinheitlichung und Zusammenführung von Zulassungsprozessen im Kontext von KI sicherzustellen. Dadurch könnten Daten aus der Marktbeobachtung länderübergreifend zugänglich werden.

---

<sup>10</sup> Organ-on-a-Chip bezeichnet eine Technologie zur miniaturisierten Nachbildung der physiologischen Abläufe im menschlichen Organismus beziehungsweise in einzelnen Organen durch Kombination unterschiedlicher Zelltypen in einem mikrofluidischen Chip.



## Wie kann die Sicherheit von Wirkstoffen in der Anwendung bei einer KI-unterstützten Entwicklung gewährleistet oder sogar im Vergleich zu bisherigen Verfahren verbessert werden?

Wie bereits angesprochen stellen sich beim KI-Einsatz in der Wirkstoffentwicklung auch in Bezug auf die Sicherstellung der Sicherheit von Studienteilnehmenden neue Herausforderungen. Die Sicherstellung beginnt bei der für das Training der KI-Modelle zugrunde liegenden Datenbasis. Einerseits bietet KI hier die Möglichkeit für mehr Sicherheit in der Entwicklung, da sowohl eine größere Menge an Daten als auch mehr unterschiedliche Datenarten verknüpft werden können (siehe Kapitel 3.1, Anwendungsbereiche). Andererseits birgt die KI-basierte Datenauswertung immer das Risiko einer Propagierung von Unausgewogenheiten aus der zugrunde liegenden Datenbasis (siehe Kapitel 5). Hinzu kommt die Frage, wie synthetische Daten<sup>11</sup> oder virtuelle klinische Studien<sup>12</sup> im Rahmen der Arzneimittelsicherheit bewertet werden sollen. Hierzu gibt es aktuell noch keine Positionen der Zulassungsbehörden.

Aufgrund des großen Potenzials, vor allem mit Blick auf die Herausforderungen bei der Durchführung klinischer Studien, sollten diese Möglichkeiten in Zukunft verstärkt ausgeschöpft werden. Dazu sind klare Rahmenbedingungen zur Validierung der KI-Modelle sowie zu den zugrunde liegenden Datensätzen notwendig, da Unsicherheiten bezüglich des Technologieeinsatzes eine Anwendung in der Breite bremsen.

Dabei ergibt sich ein Spannungsfeld zwischen der Allgemeingültigkeit und Anwendbarkeit von Modellen und Datensätzen. Idealerweise sollten die KI-generierten Ergebnisse für alle Bevölkerungsgruppen auf Basis einer ausgeglichenen Datengrundlage anwendbar sein. Allerdings existieren solche umfassenden Datensätze aktuell noch nicht; verfügbare Datensätze enthalten aktuell noch immer bekannte und potenziell unbekannt Unausgewogenheiten (Levi & Gorenstein, 2023). Daher sollten Methoden zur Identifikation von Bias kontinuierlich weiterentwickelt werden und ausgehend davon die Datenerhebung für bisher unterrepräsentierte Gruppen verstärkt werden. Trotz alledem sollten verfügbare Datensätze bereits heute nutzenbringend eingesetzt werden, allerdings nur in dem Kontext, den sie ausreichend abbilden. Ähnlich zu Beipackzetteln sollten daher der Anwendungskontext und die zugrunde liegende Datenbasis transparent nachvollziehbar sein; die Kombination unterschiedlicher Datenbanken mit definierten Schwerpunkten ermöglicht so auch mit der aktuell verfügbaren Datenbasis größer angelegte Studien.

## Wie kann die Datenverfügbarkeit für die KI-unterstützte Wirkstoffentwicklung verbessert werden?

Die Ausführungen in Kapitel 3.1, speziell zur Target- und Lead-Identifikation, zeigen die Möglichkeiten von KI bei ausreichend großer Datengrundlage. Hier erlauben Moleküldatenbanken und wissenschaftliche Literatur die datengetriebene Erforschung und Erprobung unzähliger Verbindungen sowie die Vorhersage von Eigenschaften bisher unbekannter Verbindungen.

In Bezug auf die Daten von Studienteilnehmenden beziehungsweise auf generelle Daten zur menschlichen Biologie ist die Verfügbarkeit und Allgemeingültigkeit von Datensätzen aktuell noch sehr eingeschränkt. Gleichzeitig gibt es in diesem Kontext sowohl auf europäischer als auch auf nationaler Ebene aktive Vorhaben

10 Synthetische Daten bezeichnen hier künstlich durch KI erzeugte Daten zu Probandinnen und Probanden, die nicht real existieren, aber die statistischen Eigenschaften der Daten ahmen reale Gruppen von Probandinnen und Probanden nach.

11 Virtuelle klinische Studien bezeichnen Studien, die im Computer durchgeführt werden, indem KI-basiert die Folgen einer Behandlung simuliert werden.

zur Steigerung der Datenverfügbarkeit in Form des Verordnungsentwurfs zum [European Health Data Space \(EHDS\)](#) sowie dem [Gesundheitsdatennutzungsgesetz \(GDNG\)](#) und dem [Digitalgesetz](#) in Deutschland. Dadurch sollen Gesundheitsdaten für Forschungszwecke umfassend und auch für die forschende Industrie zugänglich gemacht werden. Zusätzlich wird unter anderem die Dateninfrastruktur des Modellvorhabens Genomsequenzierung nach § 64e des Fünften Sozialgesetzbuches (SGB V) geändert, um auch hier einen zweckbasierten Datenzugang zu ermöglichen.

Der EHDS bietet wichtige Impulse für die nationale und grenzüberschreitende Datennutzung im Gesundheitswesen. Über den EHDS können Gesundheitsdaten in sicheren Verarbeitungsumgebungen zugänglich gemacht werden, die auch zum Training von KI-Systemen genutzt werden können. Die im Forschungsdatenzentrum Gesundheit (FDZ) beim Bundesinstitut für Arzneimittel und Medizinprodukte zukünftig verfügbaren Abrechnungsdaten aller gesetzlich Versicherten sollen entsprechend des mit dem GDNG angepassten §303e (2) Nr. 10 SGB V auch zum Zweck der Entwicklung, Weiterentwicklung, Nutzenbewertung und Überwachung der Sicherheit von Arzneimitteln und damit auch für Untersuchungen in der Marktbeobachtung nutzbar sein.

Die FDZ-Daten sollen dabei durch die neue Datenzugangs- und Koordinierungsstelle für Gesundheitsdaten auch in Verknüpfung mit den Daten der klinischen Krebsregister genutzt werden können, vgl. § 4 GDNG. Durch den erleichterten und erweiterten Datenzugang ergeben sich mehr Möglichkeiten in Bezug auf die Nutzung von Real-World-Data und den daraus abgeleiteten Real-World-Evidence, zum Beispiel im Kontext von retrospektiven Studien (siehe Kapitel 4.3).

Diese neuen Möglichkeiten sollten entsprechend genutzt werden, sowohl zur Forschung als auch bei der Bewertung der medizinischen Wirksamkeit oder dem Auftreten von Nebenwirkungen. Dazu können zukünftig potenziell auch multinationale Daten zum Beispiel aus Meldedatenbanken anderer EU-Staaten oder der EudraVigilance-Datenbank verknüpft und genutzt werden. Für den Bereich der Forschung und Innovation sind hier alle verfügbaren Daten relevant, weswegen die gesteigerte Datenverfügbarkeit nicht durch Einschränkungen bei der KI-gestützten Forschung torpediert werden sollte. Diese notwendige Freiheit beim Einsatz von KI findet sich daher auch in der Bereichsausnahme für Forschung und Innovation wieder hinsichtlich des Anwendungsbereichs der KI-Verordnung der Europäischen Union.

Für die Nutzung multinationaler Datensätze im Rahmen der KI-unterstützten Überwachung der Arzneimittelsicherheit sollte allerdings geprüft werden, inwieweit diese Datensätze verwertbar sind aufgrund unterschiedlicher nationaler Vorgaben zur Arzneimittelsicherheit. Hier müssen rechtliche Unsicherheiten vermieden werden, um einerseits die Nutzung von KI in diesem Bereich nicht einzuschränken und andererseits die rechtskonforme Pharmakovigilanz sicherzustellen.

Bezüglich der Daten aus der ePA von gesetzlich Versicherten ist aktuell keine aktive Zustimmung zur Datenweiterverwendung, sondern nur die Möglichkeit zum Widerspruch gegen die Datenweiterverwendung vorgesehen. Eine aktuelle Civey-Umfrage zeigt allerdings bei 51 Prozent der Befragten eine Bereitschaft zum Widerspruch gegen die Einrichtung einer ePA (Trappe, 2023), und vor allem die Datenweitergabe an die Industrie wird deutlich kritischer betrachtet als beispielsweise an öffentliche Forschungseinrichtungen (PWC, 2023). Die Datennutzung zur KI-basierten Wirkstoffentwicklung kann hier dazu dienen, Vertrauen in die Datenweitergabe zu schaffen. Krebserkrankungen werden beispielsweise immer häufiger und die Therapie profitiert gleichzeitig stark von datengetriebenen und personalisierten Ansätzen mit KI. Entsprechend besteht hier die Chance, den Mehrwert der Datenteilung für Betroffene unmittelbar greifbar zu machen (Dössel & Lenarz, 2023). In einer aktuellen deutsch-israelischen Vergleichsstudie zur Bereitschaft, Daten zu teilen, zeigte

sich trotz deutlich höheren Nutzungszahlen in Israel eine in Deutschland vergleichbare Bereitschaft zum Teilen von Gesundheitsdaten mit Leistungserbringern (ca. 80 Prozent) und zur Forschung durch Arzneimittelhersteller (ca. 30 Prozent) (Weisband et al., 2024).

### **Wie kann eine ausreichende Qualität und internationale Anschlussfähigkeit von Datensätzen für KI in der Wirkstoffentwicklung sichergestellt werden?**

Neben der Verfügbarkeit muss gleichzeitig trotz diverser unterschiedlicher Datenquellen eine ausreichende Datenqualität sichergestellt werden. Dies betrifft beispielsweise Real-World-Daten, welche große Möglichkeiten bieten, aber aufgrund der Vielzahl an möglichen Datenquellen aktuell kaum standardisiert erfasst werden. Hier braucht es verbindliche Standards, die gleichzeitig als Qualitätskriterium für Hersteller in der Bewerbung und für Kundinnen und Kunden in der Bewertung dienen können.

Bei der Standardisierung sollte gleichzeitig die internationale Anschlussfähigkeit und Interoperabilität der Datensätze mitbedacht werden, speziell in Bezug auf den Aufbau des EHDS. In diesem Kontext sollte daher der Fokus auf internationalen Standards und Ontologien als Klassifikationssysteme für Krankheitsbilder oder Diagnosen liegen. Auch hier könnte der KI-Einsatz dazu beitragen, deutschspezifische Ontologien in internationale Standardformate zu übertragen.

# Literatur

---

**Bate, A. & Stegmann, J.-U. (2023):** Artificial intelligence and pharmacovigilance: What is happening, what could happen and what should happen? Health Policy and Technology 12, 100743. <https://doi.org/10.1016/j.hlpt.2023.100743>

**BCG – Boston Consulting Group (2022a):** Adopting AI in Drug Discovery.

Online unter: <https://www.bcg.com/publications/2022/adopting-ai-in-pharmaceutical-discovery>

**BCG – Boston Consulting Group (2022b):** AI in Biologics Discovery: An Emerging Frontier.

Online unter: <https://www.bcg.com/publications/2022/ai-in-biologics-discovery-an-emerging-frontier>

**BCG – Boston Consulting Group (2023):** Unlocking the Potential of AI in Drug Discovery.

Online unter: <https://www.bcg.com/publications/2023/unlocking-the-potential-of-ai-in-drug-discovery>

**BfArM – Bundesinstitut für Arzneimittel und Medizinprodukte (2023):** Arzneimittel für seltene Erkrankungen.

Online unter: [https://www.bfarm.de/DE/Arzneimittel/Zulassung/Arzneimittel-fuer-seltene-Erkrankungen/\\_node.html;jsessionid=10BF4D7106509DA02F6CB7F101B009E9.intranet671](https://www.bfarm.de/DE/Arzneimittel/Zulassung/Arzneimittel-fuer-seltene-Erkrankungen/_node.html;jsessionid=10BF4D7106509DA02F6CB7F101B009E9.intranet671)

**Bhattamisra, S.K. et al. (2023):** Artificial Intelligence in Pharmaceutical and Healthcare Research. Big Data and Cognitive Computing 7(10). Online unter: <https://doi.org/10.3390/bdcc7010010>

**Bian, Y. & Xie X.-Q. (2021):** Generative chemistry: drug discovery with deep learning generative models.

Journal of Molecular Modeling 27(71). Online unter: <https://doi.org/10.1007/s00894-021-04674-8>

**BMBF – Bundesministerium für Bildung und Forschung (2023):** BMBF-Aktionsplan Künstliche Intelligenz.

Neue Herausforderungen chancenorientiert angehen. Online unter: <https://www.bmbf.de/bmbf/de/forschung/digitale-wirtschaft-und-gesellschaft/kuenstliche-intelligenz/ki-aktionsplan.html>

**BMG – Bundesministerium für Gesundheit (2023):** Gentherapien – Erfolgreicher Market Access, strategische

Herausforderungen und Lösungen. Online unter: <https://www.bundesgesundheitsministerium.de/arzneimittelpreise.html>

**BPB (2016):** Verordnung und Verbrauch von Arzneimitteln. Online unter: <https://www.bpb.de/themen/gesundheit/gesundheitspolitik/253121/verordnung-und-verbrauch-von-arzneimitteln/>

**Chace, C. (2022):** First Wholly AI-Developed Drug Enters Phase 1 Trials. Online unter: <https://www.forbes.com/sites/calumchace/2022/02/25/first-wholly-ai-developed-drug-enters-phase-1-trials/?sh=7906e38d2680>

**Chandak, P. & Tatonetti N. P. (2020):** Using Machine Learning to identify adverse Drug Effects posing increased Risk

to Women. Pattern 1(7), 100108. Online unter: <https://doi.org/10.1016/j.patter.2020.100108>

**Colangelo, M. (2023):** FDA Grants Orphan Drug Status For Drug Discovered Using Generative AI.

Online unter: <https://www.linkedin.com/pulse/fda-grants-orphan-drug-status-discovered-using-ai-colangelo/>

**D'Amico, S. et al. (2023):** Synthetic Data Generation by Artificial Intelligence to accelerate Research and Precision Medicine in Hematology. JCO Clinical Cancer Informatics 7, e2300021.

Online unter: <https://ascopubs.org/doi/abs/10.1200/cci.23.00071>

**DAZ – Deutsche Apothekerzeitung (2019):** Acetylsalicylsäure (ASS). Online unter: <https://www.deutsche-apotheker-zeitung.de/news/artikel/2019/05/20/wirkstoff-lexikon-acetylsalicylsaure-ass>

**de Kraker, M. E. A. et al. (2016):** Will 10 Million People die a Year due to Antimicrobial Resistance by 2050?

PLoS Medicine 13(11), e1002184. Online unter: <https://doi.org/10.1371/journal.pmed.1002184>

**Deep Pharma Intelligence (2022):** Artificial Intelligence for Drug Discovery Landscape Overview Q4 2022.

Online unter: <https://www.deep-pharma.tech/ai-in-dd-q4-2022>

**Deloitte (2019):** Intelligent Drug Discovery powered by AI. Online unter: <https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/life-sciences-health-care/deloitte-ch-en-intelligent-drug-discovery.pdf>

**DHS – Deutsche Herzstiftung e.V. (2021):** Betablocker: Wirkung, Nebenwirkungen und Wechselwirkungen.

Online unter: [https://herzstiftung.de/infos-zu-herzerkrankungen/gerinnungshemmung-und-medikamente/betablocker#:~:text=Betablocker%20oder%20genauer%20Betarezeptorenblocker%20sind,chronischer%20Herzinsuffizienz%20\(Herzschw%C3%A4che\)%20verordnet](https://herzstiftung.de/infos-zu-herzerkrankungen/gerinnungshemmung-und-medikamente/betablocker#:~:text=Betablocker%20oder%20genauer%20Betarezeptorenblocker%20sind,chronischer%20Herzinsuffizienz%20(Herzschw%C3%A4che)%20verordnet)

**Dössel, O. & Lenarz T. (2023):** Gesundheitsdatennutzung – sicher und souverän (acatech IMPULS), München 2023.

Online unter: [https://doi.org/10.48669/aca\\_2023-10](https://doi.org/10.48669/aca_2023-10)

**EMA – European Medicines Agency (2020):** EMA Regulatory Science to 2025 Strategic reflection.

Online unter: [https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection\\_en.pdf](https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection_en.pdf)

- EMA – European Medicines Agency (2021):** 2021 Annual Report on EudraVigilance for the European Parliament, the Council and the Commission. Online unter: [https://www.ema.europa.eu/en/documents/report/2021-annual-report-eudravigilance-european-parliament-council-commission\\_en.pdf](https://www.ema.europa.eu/en/documents/report/2021-annual-report-eudravigilance-european-parliament-council-commission_en.pdf)
- EMA – European Medicines Agency (2023a):** Multi-annual AI workplane. Online unter: [https://www.ema.europa.eu/en/documents/work-programme/multi-annual-artificial-intelligence-workplan-2023-2028-hma-ema-joint-big-data-steering-group\\_en.pdf](https://www.ema.europa.eu/en/documents/work-programme/multi-annual-artificial-intelligence-workplan-2023-2028-hma-ema-joint-big-data-steering-group_en.pdf)
- EMA – European Medicines Agency (2023b):** Reflection paper on the use of artificial intelligence in the lifecycle of medicines. Online unter: <https://www.ema.europa.eu/en/news/reflection-paper-use-artificial-intelligence-lifecycle-medicines>
- EU-Kommission (2023a):** Fragen und Antworten: Reform des Arzneimittelrechts. Online unter: [https://ec.europa.eu/commission/presscorner/detail/de/qanda\\_23\\_1844](https://ec.europa.eu/commission/presscorner/detail/de/qanda_23_1844)
- EU-Kommission (2023b):** Europäische Gesundheitsunion: Kommission schlägt Arzneimittelreform für leichter zugängliche, erschwinglichere und innovativere Arzneimittel vor. Online unter: [https://ec.europa.eu/commission/presscorner/detail/de/IP\\_23\\_1843](https://ec.europa.eu/commission/presscorner/detail/de/IP_23_1843)
- Evotec (2021):** Evotec und Exscientia geben klinische Studie mit neuartigem Immunonkologie-Wirkstoff bekannt. Online unter: <https://www.evotec.com/f/eb3a903bce483c6fd3b422071b2be2e8.pdf>
- FDA – Food and Drug Administration (2023a):** Artificial Intelligence and Machine Learning (AI/ML) for Drug Development. Online unter: <https://www.fda.gov/science-research/science-and-research-special-topics/artificial-intelligence-and-machine-learning-aiml-drug-development>
- FDA – Food and Drug Administration (2023b):** Artificial Intelligence in Drug Manufacturing. Online unter: <https://www.fda.gov/media/165743/download>
- Fleming, N. (2018):** Computer-calculated compounds. *Nature* 557, S. 55–57. Online unter: <https://www.nature.com/articles/d41586-018-05267-x>
- Fogel, D. B. (2018):** Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. *Contemporary Clinical Trials Communication* 11, S. 156–165. Online unter: <https://doi.org/10.1016/j.conctc.2018.08.001>
- GLI – Global Legal Insights (2023):** Pricing & Reimbursement Laws and Regulations 2023 | Germany. Online unter: <https://www.globallegalinsights.com/practice-areas/pricing-and-reimbursement-laws-and-regulations/germany>
- GKV-SV – Spitzenverband Bund der Krankenkassen (2023):** GKV-Kennzahlen. Online unter: [https://www.gkv-spitzenverband.de/service/zahlen\\_und\\_grafiken/gkv\\_kennzahlen/gkv\\_kennzahlen.jsp](https://www.gkv-spitzenverband.de/service/zahlen_und_grafiken/gkv_kennzahlen/gkv_kennzahlen.jsp)
- Göller, A. H. et al. (2020):** Bayer's in silico ADMET platform: a journey of machine learning over the past two decades. *Drug Discovery Today* 25(9), S. 1702–1709. Online unter: <https://doi.org/10.1016/j.drudis.2020.07.001>
- Gupta, R. et al. (2021):** Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular Diversity* 25, S. 1315–1360. Online unter: <https://doi.org/10.1007/s11030-021-10217-3>
- Harrer, S. et al. (2019):** Artificial Intelligence for clinical Trail Design. *Trends in Pharmacological Sciences* 40(8), S. 577–591. Online unter: <https://doi.org/10.1016/j.tips.2019.05.005>
- ICMRA – International Coalition of Medicines regulatory Authorities (2021):** Horizon Scanning Assessment Report – Artificial Intelligence. Online unter: [https://www.icmra.info/drupal/sites/default/files/2021-08/horizon\\_scanning\\_report\\_artificial\\_intelligence.pdf](https://www.icmra.info/drupal/sites/default/files/2021-08/horizon_scanning_report_artificial_intelligence.pdf)
- Insilico Medicine (2023):** Pharma.AI. Online unter: <https://insilico.com/platform>
- Johnson, J. (2023):** Pharma's pricing problem. Online unter: <https://www.investorchronicle.co.uk/news/2023/03/16/pharma-s-pricing-problem/>
- Jumper, J. et al. (2021):** Highly accurate protein structure prediction with AlphaFold. *Nature* 596, S. 583–589. Online unter: <https://www.nature.com/articles/s41586-021-03819-2>
- Kana, O. & Brylinski, M. (2019):** Elucidating the druggability of the human proteome with eFindSite. *Journal of Computer-Aided Molecular Designs* 33(5), S. 509–519. <https://doi.org/10.1007%2Fs10822-019-00197-w>
- KAS – Konrad-Adenauer-Stiftung e.V. (2023):** Pharmastandort Europa. Stärkung der Resilienz und Wettbewerbsfähigkeit. Online unter: <https://www.kas.de/documents/252038/22161843/Pharmastandort+Europa+%E2%80%93+St%C3%A4rkung+der+Resilienz+und+Wettbewerbsf%C3%A4higkeit.pdf/9023f14f-e41d-7ea9-73d0-c8eea1577e6c>
- Knapp, A. (2019):** This Startup Used AI To Design A Drug In 21 Days. Online unter: <https://www.forbes.com/sites/alexknapp/2019/09/02/this-startup-used-ai-to-design-a-drug-in-21-days/?sh=23978fc92594>
- Levi, R. & Gorenstein D. (2023):** AI in medicine needs to be carefully deployed to counter bias – and not entrench it. Online unter: <https://www.npr.org/sections/health-shots/2023/06/06/1180314219/artificial-intelligence-racial-bias-health-care>

- Lowe, D. (2023): Screening the Big Libraries – How's It Going? Online unter: <https://www.science.org/content/blog-post/screening-big-libraries-how-s-it-going>
- Löser, A., Tresp, V. et al. (2023): Große Sprachmodelle – Grundlagen, Potenziale und Herausforderungen für die Forschung. Whitepaper aus der Plattform Lernende Systeme, München. Online unter: [https://doi.org/10.48669/pls\\_2023-3](https://doi.org/10.48669/pls_2023-3)
- Lyu, J. et al. (2019): Ultra-large library docking for discovering new Chemotypes. Nature 566, S. 224–229. Online unter: <https://doi.org/10.1038/s41586-019-0917-9>
- Lyu, J. et al. (2023): Modeling the expansion of virtual screening libraries. Nature Chemical Biology 19, S. 712–718. Online unter: <https://doi.org/10.1038/s41589-022-01234-w>
- Merck (2023): Synthia Retrosynthese-Software. Online unter: <https://www.synthiaonline.com/de/>
- Microsoft (2022): MoLeR: Creating a path to more efficient drug design. Online unter: <https://www.microsoft.com/en-us/research/blog/moler-creating-a-path-to-more-efficient-drug-design/>
- Mindrak (2023): Pipeline. Online unter: <https://www.mindrak.ai/en/pipeline>
- Mirza, M. (2020): Künstliche Intelligenz: Wie hilft KI Pharma wettbewerbsfähiger zu sein? Online unter: <https://www.health-relations.de/wie-hilft-kuenstliche-intelligenz-bei-der-wettbewerbsbeobachtung/>
- Mock, M. et al. (2023): AI can help to speed up drug discovery — but only if we give it the right data. Nature 621, S. 467–470. Online unter: <https://doi.org/10.1038/d41586-023-02896-9>
- Murali, K. et al. (2019): Artificial intelligence in pharmacovigilance: Practical utility. Indian Journal of Pharmacology 51(6), S. 373. Online unter: [https://journals.lww.com/iphr/fulltext/2019/51060/artificial\\_intelligence\\_in\\_pharmacovigilance\\_.1.aspx](https://journals.lww.com/iphr/fulltext/2019/51060/artificial_intelligence_in_pharmacovigilance_.1.aspx)
- NIH – National Institute of Health (2023): PubChem Data Counts. Online unter: <https://pubchem.ncbi.nlm.nih.gov/docs/statistics>
- O'Shea, D. (2023): Generative AI speeds tedious aspects of drug discovery process. Online unter: <https://www.fierceelectronics.com/embedded/generative-ai-speeds-tedious-aspects-drug-discovery-process>
- Panch, T. et al. (2022): A distributed approach to the regulation of clinical AI. PLoS Digital Health 1(5), e0000040. Online unter: <https://doi.org/10.1371/journal.pdig.0000040>
- Patel, V. & Shah, M. (2022): Artificial intelligence and machine learning in drug discovery and development. Intelligent Medicine 2, S. 134–140. Online unter: <https://doi.org/10.1016/j.imed.2021.10.001>
- Paul, D. et al. (2020): Artificial intelligence in drug discovery and development. Drug Discovery Today 26(1). Online unter: <https://doi.org/10.1016/j.drudis.2020.10.010>
- Peptilogics (2023): Pipeline. Online unter: <https://peptilogics.com/pipeline/#plg0206>
- Pun, F. W. et al. (2023): AI-powered therapeutic target discovery. Trends in Pharmacological Sciences 44(9), S. 561–572. Online unter: <https://doi.org/10.1016/j.tips.2023.06.010>
- PwC – PricewaterhouseCoopers (2023): Healthcare-Barometer 2023. Online unter: <https://www.pwc.de/de/gesundheitswesen-und-pharma/healthcare-barometer.html>
- Saama (2023): Saama Launches Industry's First AI-driven Data Platform to Accelerate Clinical Development. Online unter: <https://www.saama.com/news/saama-launches-industrys-first-ai-driven-data-platform-to-accelerate-clinical-development/>
- Savage, N. (2021): Tapping into the drug discovery potential of AI. Online unter: <https://www.nature.com/articles/d43747-021-00045-7>
- Savage, N. (2023): Drug discovery companies are customizing ChatGPT: here's how. Nature Biotechnology 41, S. 585–586. Online unter: <https://doi.org/10.1038/s41587-023-01788-7>
- Scannell, J. W. et al. (2012): Diagnosing the decline in pharmaceutical R&D efficiency. Nature Reviews Drug Discovery 11, S. 191–200. Online unter: <https://doi.org/10.1038/nrd3681>
- Schneider, P. et al. (2020): Rethinking drug design in the artificial intelligence era. Nature Reviews Drug Discovery 19, S. 353–364. Online unter: <https://doi.org/10.1038/s41573-019-0050-3>
- Schönemark Kielhorn Collegen (2023): Whitepaper GKV-Finanzstabilisierungsgesetz - Strategische Implikationen für den Market Access in Deutschland. Online unter: [https://skc-beratung.de/de/insights/strategy\\_reports/2023/gkv-finstg.php](https://skc-beratung.de/de/insights/strategy_reports/2023/gkv-finstg.php)
- Schulte, A. (2022): Pharmaindustrie nutzt Künstliche Intelligenz zur Arzneimittelforschung. Online unter: <https://www.handelsblatt.com/technik/medizin/neue-medikamente-pharmaindustrie-nutzt-kuenstliche-intelligenz-zur-arzneimittelforschung/28161478.html>
- Sharma, K. & Manchikanti, P. (2020): Regulation of Artificial Intelligence in Drug Discovery and Health Care. Biotechnology Law Report 39(5). Online unter: <https://doi.org/10.1089/blr.2020.29183.ks>

- Sharma, A. et al. (2022):** Artificial Intelligence-based Data-driven Strategy to accelerate Research, Development, and Clinical Trials of COVID Vaccine. *BioMed Research International* 7205241. Online unter: <https://doi.org/10.1155/2022/7205241>
- Spinner, J. (2022):** Lilly deploys AI-powered tech to expedite drug approval process. Online unter: <https://www.outsourcing-pharma.com/Article/2022/05/18/lilly-deploys-ai-powered-tech-to-expedite-drug-approval-process>
- Statista (2023):** Size of the global chemical drugs and biologics pharmaceutical market from 2014 to 2023. Online unter: <https://www.statista.com/statistics/1085563/revenue-chemical-drugs-and-biologics-global-pharmaceuticals/>
- Stowasser, S. et al. (2020):** Einführung von KI-Systemen in Unternehmen. Gestaltungsansätze für das Change Management. Whitepaper aus der Plattform Lernende Systeme. Online unter: [https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG2\\_Whitepaper\\_Change\\_Management.pdf](https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG2_Whitepaper_Change_Management.pdf)
- Strayhorn Jr., J. M. (2021):** Virtual controls as an alternative to randomized controlled trials for assessing efficacy of interventions. *BMC Medical Research Methodology* 21(3). Online unter: <https://doi.org/10.1186/s12874-020-01191-9>
- Sun, D. et al. (2022):** Why 90% of clinical drug development fails and how to improve it? *Acta Pharmaceutica Sinica B* 12(7), S. 3049–3062. Online unter: <https://doi.org/10.1016%2Fj.apsb.2022.02.002>
- Trappe, T. (2023):** Jeder Zweite offen für Datenweitergabe. Online unter: <https://background.tagesspiegel.de/gesundheit/jeder-zweite-offen-fuer-datenweitergabe>
- Van Meenen, J. et al. (2022):** Making Biomedical Sciences publications more accessible for machines. *Medicine, Health Care and Philosophy* 25, S. 179–190. Online unter: <https://doi.org/10.1007/s11019-022-10069-0>
- vfa – Verband forschender Arzneimittelhersteller (2021):** Die forschungsintensivste Branche Deutschlands? Pharma! Online unter: <https://www.vfa.de/de/wirtschaft-politik/wirtschaft/pharma-forschungsintensivste-branche-deutschlands>
- vfa – Verband forschender Arzneimittelhersteller (2022):** Deutschland verliert bei klinischen Studien weiter an Boden. Online unter: <https://www.vfa.de/de/arzneimittel-forschung/forschungsstandort-deutschland/klinische-studien-deutschland.html>
- vfa – Verband forschender Arzneimittelhersteller (2023a):** Was den Studienstandort Deutschland bremst. Online unter: <https://www.vfa.de/de/arzneimittel-forschung/klinische-studien/was-den-studienstandort-deutschland-bremst>
- vfa – Verband forschender Arzneimittelhersteller (2023b):** So entsteht ein neues Medikament. Online unter: <https://www.vfa.de/de/arzneimittel-forschung/so-funktioniert-pharmaforschung/so-entsteht-ein-medikament.html>
- vfa – Verband forschender Arzneimittelhersteller & Kearney (2023):** Pharma-Innovationsstandort Deutschland. Online unter: <https://www.vfa.de/download/pg-004-2023.pdf>
- Weisband, Y. et al. (2024):** GIHF-AI Studie 2023. Vertrauen in die Nutzung von Gesundheitsdaten – ein Vergleich zwischen Deutschland und Israel. Online unter: [https://gihf-ai.eu/wp-content/uploads/2024/01/20240118\\_GIHF-AI\\_Survey\\_DE\\_WEB.pdf](https://gihf-ai.eu/wp-content/uploads/2024/01/20240118_GIHF-AI_Survey_DE_WEB.pdf)
- WIdO – Wissenschaftliches Institut der AOK (2020):** Der GKV-Arzneimittelmarkt Bericht 2020. Online unter: [https://www.wido.de/fileadmin/Dateien/Dokumente/Forschung\\_Projekte/Arzneimittel/wido\\_arz\\_gkv-arzneimittelmarkt\\_2020.pdf](https://www.wido.de/fileadmin/Dateien/Dokumente/Forschung_Projekte/Arzneimittel/wido_arz_gkv-arzneimittelmarkt_2020.pdf)
- Xie, N. et al. (2023):** Neoantigens: promising targets for cancer therapy, *Signal Transduction and Targeted Therapy* 8, 9. Online unter: <https://doi.org/10.1038/s41392-022-01270-x>
- Yang, F. et al. (2021):** Machine Learning Applications in Drug Repurposing. *Interdisciplinary Sciences: Computational Life Sciences* 14, S. 15–21. Online unter: <https://doi.org/10.1007/s12539-021-00487-8>
- Yseop (2023):** Clinical Document Automation. Online unter: <https://yseop.com/solutions/life-sciences/clinical/>
- Zahn, T. P. et al. (2023):** KI basierter Rapid Evidence Generator mit risikoadjustiertem Propensity Score Matching. *Prävention und Gesundheitsförderung*. Online unter: <https://doi.org/10.1007/s11553-023-01048-1>
- Zhavoronkov, A. et al. (2019):** Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology* 37, S. 1038–1040. Online unter: <https://doi.org/10.1038/s41587-019-0224-x>

# Über dieses Whitepaper

---

Die Autorinnen und Autoren des Whitepapers sind Mitglieder der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme. Als eine von insgesamt sieben Arbeitsgruppen untersucht die Arbeitsgruppe die Chancen und Herausforderungen, die Lernende Systeme für Prävention, Diagnose und Therapie in der Medizin sowie in der Pflege und Rehabilitation bieten. Sie behandelt damit auch Fragen der gesellschaftlichen Akzeptanz sowie des Datenschutzes in diesen Anwendungsbereichen. Die Arbeitsgruppe Gesundheit, Medizintechnik, Pflege repräsentiert die verschiedensten Bereiche des Gesundheitswesens und arbeitet eng mit den anderen Arbeitsgruppen der Plattform Lernende Systeme zusammen.

Die vorliegende Publikation wurde von folgenden Mitgliedern der Arbeitsgruppe Gesundheit, Medizintechnik, Pflege der Plattform Lernende Systeme erstellt:

## **Autorinnen und Autoren**

Prof. Dr. med. Klemens Budde, Charité – Universitätsmedizin Berlin (Hauptautor & AG-Leiter)

Prof. Dr.-Ing. Dr. med. Karsten Hiltawsky, Drägerwerk AG & Co. KGaA (AG-Leiter)

Dr. Jean-Enno Charton, Merck KGaA

Prof. Dr. Björn Eskofier, Friedrich-Alexander-Universität Erlangen-Nürnberg

Prof. Dr. Dagmar Krefting, Institut für Medizinische Informatik/Universitätsmedizin Göttingen

Dr. Maren Lang, BioNTech SE

Prof. Dr. Wolfgang Nejd, Leibniz Universität Hannover, Forschungszentrum L3S

Dr.-Ing. Matthieu-P. Schapranow, Hasso-Plattner-Institut für Digital Engineering gGmbH

Prof. Dr. Thomas Zahn, bbw Hochschule

## **Redaktion**

Dr. Thomas Steiner, Geschäftsstelle der Plattform Lernende Systeme

Christine Wirth, Geschäftsstelle der Plattform Lernende Systeme



## Impressum

### Herausgeber

Lernende Systeme –  
Die Plattform für Künstliche Intelligenz  
Geschäftsstelle | c/o acatech  
Karolinenplatz 4 | 80333 München  
[www.plattform-lernende-systeme.de](http://www.plattform-lernende-systeme.de)

### Gestaltung und Produktion

PRpetuum GmbH, München

### Stand

März 2024

### Bildnachweis

gorodenkoff/iStock/Titel  
freepik/S. 14

### Empfohlene Zitierweise

Budde, K. et al. (2024): Arzneimittel mit KI entwickeln: Von der Idee bis zur Zulassung. Anwendungen, Potenziale und Herausforderungen. Whitepaper aus der Plattform Lernende Systeme, München.

DOI: [https://doi.org/10.48669/pls\\_2024-1](https://doi.org/10.48669/pls_2024-1)

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, der Entnahme von Abbildungen, der Wiedergabe auf fotomechanischem oder ähnlichem Wege und der Speicherung in Datenverarbeitungsanlagen, bleiben – auch bei nur auszugsweiser Verwendung – vorbehalten.

Bei Fragen oder Anmerkungen zu dieser Publikation kontaktieren Sie bitte Dr. Thomas Schmidt (Leiter der Geschäftsstelle):  
[kontakt@plattform-lernende-systeme.de](mailto:kontakt@plattform-lernende-systeme.de)



## Über die Plattform Lernende Systeme

Die Plattform Lernende Systeme ist ein Netzwerk von Expertinnen und Experten zum Thema Künstliche Intelligenz (KI). Sie bündelt vorhandenes Fachwissen und fördert als unabhängiger Makler den interdisziplinären Austausch und gesellschaftlichen Dialog. Die knapp 200 Mitglieder aus Wissenschaft, Wirtschaft und Gesellschaft entwickeln in Arbeitsgruppen Positionen zu Chancen und Herausforderungen von KI und benennen Handlungsoptionen für ihre verantwortliche Gestaltung. Damit unterstützen sie den Weg Deutschlands zu einem führenden Anbieter von vertrauenswürdiger KI sowie den Einsatz der Schlüsseltechnologie in Wirtschaft und Gesellschaft. Die Plattform Lernende Systeme wurde 2017 vom Bundesministerium für Bildung und Forschung (BMBF) auf Anregung des Hightech-Forums und acatech – Deutsche Akademie der Technikwissenschaften gegründet und wird von einem Lenkungskreis gesteuert.