

# Stellungnahme der DAGStat

## Daten und Statistik als Grundlage für Entscheidungen: Eine Diskussion am Beispiel der Corona-Pandemie

Version 22.03.2021

### Inhaltsverzeichnis

Einleitung.....	1
Daten.....	4
Modellierung und Simulation dynamischer Systeme.....	12
Decision Making.....	17
Statistical Literacy und visuelle Kommunikation in Zeiten von COVID-19.....	24
Internationale Initiativen.....	31
Zusammenfassung.....	32
Referenzen.....	33

### Einleitung

Eine Pandemie stellt in besonderem Maße national wie international Herausforderungen an den Prozess der Entscheidungsfindung. Virusinfektionen zeichnen sich durch einen potenziell exponentiellen Verlauf aus. Diese hohe zeitliche Dynamik macht vorausschauendes Handeln notwendig. Damit Entscheidungen jedoch evidenzbasiert erfolgen können, setzen diese eine zeitnahe Verfügbarkeit von Evidenz und Daten zur Analyse des Infektionsgeschehens voraus. Insbesondere bei der Aufarbeitung der Daten können hier Statistiker/innen helfen. Hier wurden auch international bereits Empfehlungen erarbeitet, siehe z.B. [72]. Am Beispiel der SARS-CoV-2 Pandemie hat sich eine Arbeitsgruppe der Deutschen Arbeitsgemeinschaft Statistik (DAGStat) mit der Frage beschäftigt, wie Daten und Statistiken politische Entscheidungen unterstützen können, und konkrete Empfehlungen formuliert. Diese Stellungnahme fasst sowohl die Grundlagen wie auch die abgeleiteten Empfehlungen zusammen.

Das Wissen über das Virus und den Krankheitsverlauf ist zu Beginn einer Pandemie grundsätzlich mit Unsicherheiten verbunden. Deshalb ist das systematische Sammeln von Daten auf regionaler, nationaler und internationaler Ebene von besonderer Bedeutung. Statistiker/innen können bei der richtigen Datenerhebung und beim richtigen Umgang mit Daten helfen. Verschiedene Aspekte sind dabei zu beachten: Welche Daten sollten erhoben werden? In welcher Qualität sind Daten verfügbar? Wie verändern sich die Daten im Laufe der Pandemie? Mit Hilfe dieser Daten können dann statistische und entscheidungsanalytische Modelle erstellt werden, die Vorhersagen (prediction) und die Simulation von Konsequenzen von Interventionen (counterfactual prediction) ermöglichen.

Das besondere an einer Pandemie ist, dass Entscheidungen und damit auch Modellvorhersagen einen direkten Einfluss auf die weitere Entwicklung der Prozesse haben. Mit der Kommunikation von Daten

und Entscheidungen wird das Verhalten der Menschen und damit der Pandemieverlauf bestimmt. Dies führt dazu, dass Modelle laufend angepasst und überarbeitet werden müssen.

Die in einer Pandemie erhobenen Daten werden aus sehr unterschiedlichen Perspektiven betrachtet: Virolog/innen, Epidemiolog/innen, Biolog/innen, Mediziner/innen, Soziolog/innen, Public-Health-Expert/innen, Ökonom/innen, Ethiker/innen sowie Bürger/innen und Politiker/innen stellen unterschiedliche Fragen. Statistische, mathematische und entscheidungsanalytische Modelle sind in vielen Fällen spezifisch für die gestellte Frage. Eine Konsequenz daraus ist, dass abhängig von der Fragestellung unterschiedliche Modelle und Modellansätze zur Anwendung kommen können.

Unsicherheiten im Wissen über das Virus, Unsicherheiten in den Daten und die Vielzahl von Fragestellungen bzw. Modellierungsansätzen motivieren u.a. eine disziplinübergreifende Zusammenarbeit von Statistiker/innen, Mathematiker/innen, Epidemiolog/innen, Mediziner/innen, Physiker/innen, Informatiker/innen, Soziolog/innen, Public-Health-Expert/innen und Entscheidungsanalytiker/innen.

Für jede Fragestellung an die Daten gibt es mehrere Herangehensweisen. Um schließlich zu einer Entscheidung zu kommen, müssen deshalb gesundheitsbezogene, ökonomische und ethische Konsequenzen von Interventionen gegeneinander abgewogen werden. Dieser Prozess erfordert unbedingt einen interdisziplinären Austausch.

Die Analyse von Daten und Konsequenzen von Maßnahmen dienen aber nicht nur zur Entscheidungsunterstützung von Politikern, sondern auch zur Information der Öffentlichkeit. Die Kommunikation von Unsicherheiten, Wahrscheinlichkeiten und Risiken (wie beispielsweise der Konsequenzen und Trade-Offs von Social Distancing) stellen eine Herausforderung dar. Hier kommt der Visualisierung von Daten, Vorhersagen und Modellierungsergebnissen eine besondere Bedeutung zu.

Ein Beispiel, an dem sich in der aktuellen Pandemie vielfältige Herausforderungen für modellgestützte Entscheidungen aufzeigen, ist die Frage der Übersterblichkeit: Zum einen spielen diese Daten eine Rolle bei der Diskussion über die Dringlichkeit von Handlungsentscheidungen / Maßnahmen. Zum anderen ist das zeitnahe und systematische Sammeln dieser Daten auch deshalb besonders wichtig, weil dies Vergleiche unterschiedlicher Strategien und Interventionen in verschiedenen Regionen und Ländern ermöglicht. Abbildung 1 zeigt das Zusammenspiel von Politik, Datenerhebung und theoretischen Erkenntnissen.

Eine Pandemie zeichnet sich dadurch aus, dass das Verhalten nahezu der gesamten Gesellschaft dazu beiträgt, wie sich die Krise entwickelt. Dabei muss jede/r Einzelne über sein persönliches Verhalten Entscheidungen treffen, und die Regierung muss Entscheidungen treffen, die das Verhalten der Bürger/innen beeinflussen.

Der statistischen und mathematischen Modellierung kommt in dieser Zeit eine besondere Bedeutung zu: Daten bilden die Grundlage für individuelle und gesellschaftliche Entscheidungen. Wissenschaft und Politik haben in solchen Zeiten eng verbundene und dennoch unterschiedliche Aufgaben: Vertrauenswürdige Wissenschaftler/innen informieren und erklären, während Politiker/innen überzeugen und – unter Berücksichtigung von Unsicherheiten über gesundheitsbezogene, ökonomische und ethische Konsequenzen – Entscheidungen treffen müssen.

Der bisherige Verlauf der aktuellen SARS-CoV-2 Pandemie hat gezeigt, dass die Grenzen zwischen Informieren und Überzeugen nicht immer klar gezogen wurden. Für einen Teil der Wissenschaftler/innen haben die Zahlen zu Neuinfektionen, zur Übersterblichkeit und die Belegung von Betten der Intensivstationen mit COVID-19-Patient/innen eine deutliche Sprache gesprochen und die Notwendigkeit für Maßnahmen zur Reduzierung des Infektionsgeschehens deutlich gemacht. Andere haben die gleichen Daten so interpretiert, dass aus ihrer Sicht weniger strikte Maßnahmen notwendig sind.

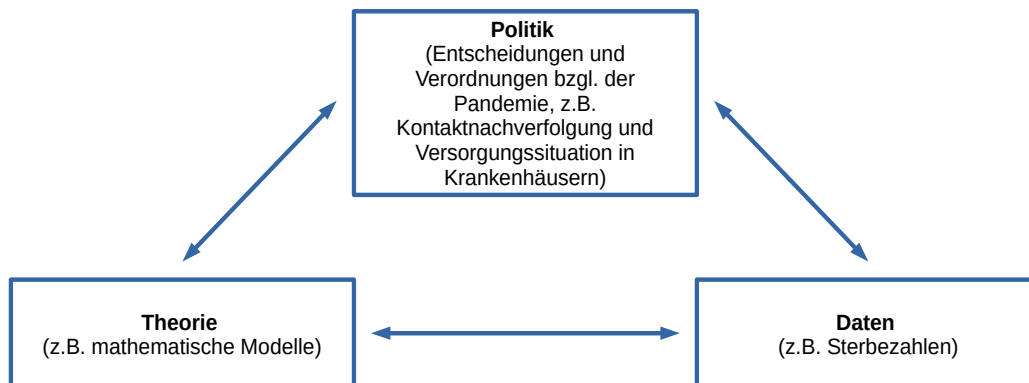


Abbildung 1: Zusammenspiel von Theorie, Daten und Politik. In Anlehnung an Schaubild 4 aus [68].

Dabei spielt die Qualität der Daten eine entscheidende Rolle: Die Übersterblichkeit beispielsweise kann nur im Rahmen einer nationalen Strategie und systematischen Sammlung von Daten zuverlässig ermittelt werden. Dann kann sie als Grundlage zur Evaluierung von Maßnahmen herangezogen werden, indem verschiedene Länder oder Regionen verglichen werden. In diesem Zusammenhang ist die Beteiligung an internationalen Initiativen zur standardisierten Datensammlung empfehlenswert bzw. notwendig.

### **Woran erkennt man vertrauenswürdige Analysen?**

Wie erkennt man vertrauenswürdige Analysen und Informationen? Vertrauenswürdigkeit entsteht durch Ehrlichkeit, Zuverlässigkeit, Kompetenz und Transparenz in der Analyse sowie Präsentation und Interpretation von Daten [1]. Eine vertrauenswürdige Kommunikation von Evidenz zeichnet sich z.B. durch folgende Aspekte aus:

- Die Herkunft und Generierung von Daten wird transparent dargestellt.
- Standards für die Qualität der Daten müssen nachvollzogen werden können bzw. zertifiziert sein.
- Die Vorgehensweise in der Analyse muss nachvollziehbar sein.
- Annahmen werden transparent gemacht und erläutert.
- Alternative Herangehensweisen werden mit ihren Vor- und Nachteilen verglichen.
- Unsicherheiten in Daten und Ergebnissen werden dargestellt.

- Einschränkungen der Analysen und von Vorhersagen werden deutlich gemacht.
- Ergebnisse von Analysen, Simulationen und anderen Forschungsergebnissen sollten inklusive Annahmen und Unsicherheiten verständlich kommuniziert werden.

Viele dieser Punkte sind nicht neu, sondern wurden z.B. bereits 1985 von dem Statistiker John Tukey im Rahmen seiner Abschiedsvorlesung zusammengestellt [2]. Desweiteren sind in der amtlichen Statistik unter dem Begriff Zuverlässigkeit die Grundsätze „Fachliche Unabhängigkeit“, „Unparteilichkeit“ und „Objektivität“ als wesentliche Kriterien genannt [10a].

In den folgenden Kapiteln gehen wir auf diese Punkte detaillierter ein und leiten Empfehlungen ab, wie Daten, Statistik und Entscheidungsanalyse zur Entscheidungsunterstützung in Krisensituationen genutzt werden können.

## Daten

Eine wesentliche Grundlage für eine evidenzbasierte Politik, aber auch für die Forschung, sind Daten. Entscheidend ist dabei nicht nur die reine Präsenz von Daten, sondern vielmehr der gesamte Prozess der Datengewinnung bis hin zur Analyse. Im Vordergrund steht hierbei die Frage, was eigentlich gemessen werden soll, d.h. was das Ziel ist und welche Indikatoren oder Modelle abgeleitet werden sollen. Beispielsweise werden für die Fragen „welche Faktoren beeinflussen die Infektionszahlen“ oder „ist es gerechtfertigt, bei einer 7-Tage-Inzidenz von über 200 Fällen pro 100.000 Einwohner/innen einen Lockdown anzuordnen“ sehr unterschiedliche Daten benötigt. Gerade in Bezug auf kleinräumige Aussagen, etwa in Landkreisen, werden die Anforderungen an die Qualität der Daten zunehmend größer.

Im Rahmen der Gewinnung bis hin zur Bereitstellung müssen folgende Fragen geklärt sein:

1. Wofür werden die Daten benötigt?
2. Wurden geeignete Daten gewonnen, enthalten sie die geeigneten Variablen für die Beantwortung der Fragestellung?
3. Wie wurden die Daten gewonnen?

Über die Ziele werden die Anforderungen an die Daten festgelegt. Hierzu gehören auch die räumliche und zeitliche Auflösung. Liegen z.B. Stichprobendaten vor, oder verfügt man über eine Vollerhebung bzw. Registerdaten?

Für politische Entscheidungen werden in der Regel Aggregat-Informationen auf verschiedenen Ebenen benötigt, also z.B. auf Bundesländer-, Kreis- oder gar Gemeindeebene. Man spricht auch von „abgeleiteten“ Statistiken. Zu diesen Statistiken gehören stets Meta-Informationen über die Qualität und Aussagekraft der Daten. Ebenso spielt gerade bei einem sich schnell ändernden Phänomen wie der SARS-CoV-2-Infektionsrate die Aktualität der Daten eine zentrale Rolle.

Für die Forschung und auch für die Gewinnung der zuvor genannten Indikatorwerte werden darüber hinaus oftmals Einzeldaten benötigt. Möchte man z.B. verstehen, welche Personengruppen einem höheren Risiko für einen schweren COVID-19-Verlauf ausgesetzt sind, muss man Individualmerkmale kennen. Aufgrund der hohen Anforderungen an den Datenschutz in Deutschland ist die Bereitstellung derartiger Daten keineswegs einfach, auch wenn im Bundesstatistikgesetz ein prioritärer Zugang der Wissenschaft vorgesehen ist. Gleichwohl erkennt man in Deutschland, dass im

Vergleich zu anderen Ländern relativ wenige Arbeiten zu COVID-19, basierend auf Mikrodaten, entstanden sind. Zweifelsohne ist der Datenschutz ein hohes Gut und sehr wichtig. Für die Zukunft wünscht man sich im Sinne eines schnelleren und sicheren Erkenntnisgewinns jedoch bessere und einfachere Zugangsmöglichkeiten, die den Erfordernissen des informellen Selbstbestimmungsrechts und des Gesundheitsdatenschutzes gerecht werden.

Im Weiteren geben wir einen Überblick über verschiedene relevante vorhandene Datenquellen im Zusammenhang mit COVID-19. Anschließend diskutieren wir Kriterien zur Beurteilung von Daten. Wir schließen dieses Kapitel mit einigen Empfehlungen.

## **Datenquellen**

Insgesamt stehen sehr viele unterschiedliche Daten aus unterschiedlichen Quellen bzw. von unterschiedlichen Datenproduzenten zur Verfügung. Ein wesentliches Merkmal der Daten ist hierbei die Qualität sowie ihre Verfügbarkeit. Im Kontext der Qualität ist vor allem auch der Zielkonflikt von Aktualität und Genauigkeit relevant. Erfahrungswerte zeigen, dass von allen Daten gerade diejenigen, die aktuell benötigt werden, oftmals nicht vorhanden sind; die COVID-19 Pandemie hat dies noch einmal deutlich gemacht.

Vorhandene Daten und Statistiken sind immer auch eine Reflektion der Vergangenheit. Arbeitsprogramme der amtlichen Statistik, vor allem im föderalen System, entstehen bei begrenzten Ressourcen in einem pluralistischen Diskurs. Amtliche Statistiken benötigen darüber hinaus eine gesetzliche Grundlage, die in der Regel nicht zeitnah umzusetzen ist. Auf der anderen Seite sind amtliche Statistiken zahlreich und auch regional bzw. für kleinere Gruppen leistungsfähig. Wissenschaftsgetragene Daten und Statistiken, wie z.B. das vom Deutschen Institut für Wirtschaftsforschung erhobene Sozioökonomische Panel (SOEP), können deutlich schneller auf neuen Datenbedarf reagieren. So arbeitet das SOEP gerade mit dem Robert-Koch-Institut (RKI) an einer gemeinsamen Befragung der SOEP-Haushalte [3]. Aber auch in diesem Bereich sind zum einen Ressourcen begrenzt, zum anderen lässt sich ein Fragenkatalog nicht ohne Weiteres kurzfristig anpassen. Ein weiterer Aspekt bei wissenschaftsgetragenen Daten ist, dass ihre Anzahl und Stichprobengröße deutlich hinter denen der amtlichen Statistik zurückbleibt.

Private Datenproduzenten können häufig schnell auf neuen Datenbedarf reagieren. Im Allgemeinen leiden sie jedoch unter geringen Rücklaufquoten (häufig in der Markt- und Meinungsforschung) oder unter unbekannter Selektionsverzerrung (insbesondere Web-basierte Datendienste). Erhebungen, die differenzierte (insbesondere auch regionale) Informationen für größere Regionen liefern, liegen daher kaum oder in unzureichender Qualität vor.

### Daten der amtlichen Statistik

Der Bereich der amtlichen Daten ist in Deutschland weit gefächert. Neben den Statistischen Ämtern des Bundes und der Länder sind hier vor allem auch die Bundesagentur für Arbeit und die Deutsche Bundesbank zu nennen. Diese beiden Institutionen betreiben neben ihren Hauptaufgaben im Bereich der Arbeitsmarkt- bzw. Geldpolitik auch umfangreich amtliche Statistik. Dies geschieht zum einen, indem sie eigene und weitere administrative Daten nutzen, zum anderen aber auch mit Hilfe der Durchführung eigener Befragungen.

Seitens der Statistischen Ämter sind seit Frühjahr 2020 eine Reihe neuer Angebote entstanden, die neben den direkten gesundheitsstatistischen Aspekten insbesondere auch die wirtschaftlichen und sozialen Auswirkungen betrachten. Hier ist zum einen das Dashboard Deutschland (<https://www.dashboard-deutschland.de/>) zu nennen, zum anderen die Rubrik EXDAT – Experimentelle Daten (<https://www.destatis.de/DE/Service/EXDAT/inhalt.html>, Abbildung 2). EXDAT-Publikationen des Statistischen Bundesamtes basieren auf nicht-traditionellen Daten wie z.B. Mobilfunksignalen und sind daher derzeit hinsichtlich einiger Qualitätsanforderungen noch weiter zu entwickeln, geben aber einen sehr schnellen Einblick in aktuelle Zusammenhänge.

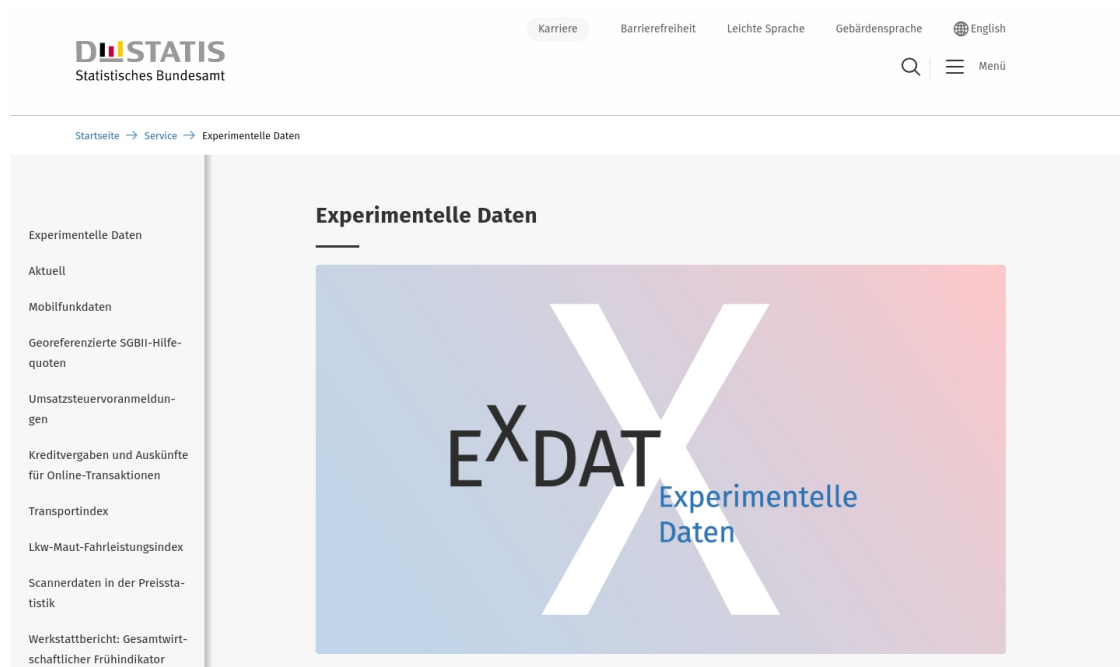


Abbildung 2: Die EXDAT-Rubrik des Statistischen Bundesamtes (<https://www.destatis.de/DE/Service/EXDAT/inhalt.html>).

Neben diesen Institutionen sind auch die Ministerien, insbesondere mit ihren nachgeordneten Behörden, als amtliche Datenproduzenten zu benennen. Zu diesen gehören z.B. das RKI als nachgeordnete Behörde des Bundesministeriums für Gesundheit oder das Kraftfahrtbundesamt als nachgeordnete Behörde des Bundesministeriums für Verkehr und digitale Infrastruktur. Kennzeichnend ist hier, dass die entsprechenden Daten bzw. Statistiken in der Regel gesetzlich angeordnet und öffentlich finanziert sind.

Eine gewisse Sonderstellung nehmen in der derzeitigen Pandemie die Daten von infas360 [72] ein, die gegenwärtig mit Hilfe einer Finanzierung durch das Bundesministerium für Wirtschaft und Energie (BMWi) zusammengetragen werden (<https://www.corona-datenplattform.de/>). Diese regional tiefer differenzierten Daten sollen der wissenschaftlichen Forschung zum Corona-Geschehen die notwendige empirische Infrastruktur liefern.

### Wissenschaftsgetragene Daten

Neben diesen amtlichen Daten im engeren Sinne werden eine Reihe von weiteren Daten produziert, die öffentlich finanziert sind, aber von wissenschaftlichen Instituten erhoben werden. Die Finanzierung erfolgt teilweise direkt durch Ministerien oder durch Wissenschaftsorganisationen wie z.B. die Leibniz- oder die Helmholtzgemeinschaft, die seitens Bundes- und Landesministerien finanziell gefördert werden. Aus der Fülle der wissenschaftsgetragenen Daten sind an dieser Stelle insbesondere zwei Datenbestände anzusprechen, die bereits jetzt als Datengrundlage zur weiteren Analyse der Corona-Pandemie-Auswirkungen genutzt werden:

In der empirischen Sozialforschung ist das Sozioökonomische Panel (SOEP) (<https://www.diw.de/soep>) eine auch international sehr intensiv genutzte Haushaltsstichprobe. Das SOEP wird jährlich durch das Deutsche Institut für Wirtschaftsforschung erhoben. Aktuell ist die bundesweite Antikörper-Studie „Leben in Deutschland – Corona-Monitoring“ zu benennen, die gemeinsam vom SOEP mit dem RKI durchgeführt wird.

In ökonomischen Analysen zu den konjunkturellen Auswirkungen der Corona-Krise werden vor allem Befragungen des IFO-Instituts genutzt. Bekanntester Datensatz ist hier der Geschäftsklimaindex (<https://www.ifo.de/umfrage/ifo-geschaeftsklimaindex>), aber auch die weiteren Erhebungen des IFO-Instituts zur wirtschaftlichen Entwicklung sind von Bedeutung (siehe <https://www.ifo.de/publikationen/2020/aufsatz-zeitschrift/ifo-konjunkturumfragen-ein-ueberblick-ueber-die-verfuegbaren>).

### Zugang zu öffentlich finanzierten Daten

Der Zugang zu amtlichen wie öffentlich finanzierten Daten, auch auf der Ebene der Einzeldaten, ist mittlerweile, nach einer rund dreißigjährigen Diskussion zwischen den öffentlichen Datenproduzenten und der empirischen Wissenschaft, im Grundsatz vollständig gegeben, auch wenn diese Zugänge derzeit noch nicht immer den technischen Möglichkeiten entsprechen.

Die Diskussion über den Zugang zu mit öffentlichen Mitteln produzierten Daten wurde in den letzten Jahren zunächst in einem Gründungsausschuss und im weiteren Verlauf im Rat für Sozial- und Wirtschaftsdaten (RatSWD) geführt. Finanziert durch das Bundesministerium für Bildung und Forschung (BMBF) hat der RatSWD ([www.ratswd.de](http://www.ratswd.de)), paritätisch besetzt mit empirisch arbeitenden Wissenschaftler/innen und Datenproduzenten, die sogenannte „informationelle Infrastruktur“ deutlich ausgebaut. Der Forschung steht heute eine hochentwickelte und nachhaltige Forschungsdateninfrastruktur zur Verfügung. Ein dezentrales Netzwerk von 38 vom RatSWD akkreditierten Forschungsdatenzentren (FDZ) ermöglicht einen kostengünstigen und einfachen Zugang zu einer Vielzahl von Forschungsdaten (<https://www.konsortswd.de/datenzentren/alle-datenzentren/>). Zu den erhältlichen Forschungsdaten gehören dabei auch umfangreiche Metadateninformationen, so dass der Generierungsprozess der Daten transparent nachvollzogen werden kann.

Gegenwärtig wird die o.g. Forschungsdateninfrastruktur in das Konsortium für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften ([www.KonsortSWD.de](http://www.KonsortSWD.de)) überführt. Die Wissenschaftskonferenz von Bund und Ländern hat im Juni 2020 die Förderung des KonsortSWD im Rahmen der Nationalen Forschungsdateninfrastruktur (NFDI) benannt [4].

## Neue digitale Daten

Neben den klassischen Befragungs- und administrativen Daten haben neue digitale Daten in den letzten Jahren deutlich an Bedeutung gewonnen. Es lässt sich durchaus schlussfolgern, dass sich mit der Corona-Krise diese Daten, die zu Beginn der Entwicklung mit „Big Data“ beschrieben wurden, etabliert haben. Nicht-traditionelle, digitale Daten unterscheiden sich wesentlich von Daten, die für einen Untersuchungszweck generiert werden sowie auch von administrativen Daten. Nicht-traditionelle Daten gehören, ebenso wie administrative Daten, zur Gruppe der sogenannten Sekundärdaten, da sie primär für einen anderen Zweck entstanden sind, anschließend aber sekundär für statistische Zwecke genutzt werden. Im Gegensatz zu administrativen Daten stehen nicht-traditionelle Daten zunächst nicht für statistische Zwecke zur Verfügung. Diese Daten fallen heute immer noch überwiegend als Nebenprodukt der eigentlichen Wertschöpfung der Unternehmen an. Ein Beispiel sind die Bewegungsdaten von Mobilfunknutzern, die als Nebenprodukt bei Unternehmen anfallen, welche am Kommunikationsmarkt tätig sind. Daneben gibt es auch Unternehmen, die gezielt nicht-traditionelle Daten für den Informationsmarkt produzieren [5, 6].

Nicht-traditionelle Daten, zu denen auch die Daten des sogenannten „Internet der Dinge“ gehören, stehen häufig über die Unternehmen nur sehr begrenzt zur Verfügung (und wenn, dann häufig zu vergleichsweise hohen Preisen). Darüber hinaus sind meist nur sehr eingeschränkt Informationen über den Generierungsprozess dieser Daten verfügbar. Somit lässt sich ihre Qualität oftmals nur beschränkt beurteilen. Zu den derzeit am häufigsten für statistische Zwecke genutzten nicht-traditionellen Daten zählen Mobilfunkdaten (siehe hierzu z.B. [7] oder [8]).

Im Gegensatz zu Mobilfunkdaten sind Satellitendaten eine in weiten Bereichen bereits etablierte Datenquelle. Hier stehen teilweise umfangreiche Datenbestände kostenfrei zur Verfügung. So können z.B. die Daten der Sentinel-Satelliten, die im Rahmen des von der European Space Agency geleiteten Copernicus-Programms entstehen, über das Copernicus Open Access Hub unter <https://scihub.copernicus.eu/> genutzt werden. Diese Daten finden bereits aktuell Verwendung, u.a. in den Landwirtschaftsstatistiken. Kommerzielle Satellitendaten, insbesondere in sehr hoher Auflösung und zeitlicher Frequenz, sind hingegen sehr kostenintensiv und wurden bisher nur in experimentellen Arbeiten, z.B. für Konjunkturanalysen untersucht (siehe hierzu z.B. <https://www.destatis.de/DE/Service/EXDAT/Datensaetze/satellitendaten.html>).

Daten aus dem Internet haben trotz ihrer Selektivität einen hohen statistischen Wert. In den Preisstatistiken werden diese Daten mittlerweile amtlich genutzt. In anderen Bereichen, wie z.B. bei der Nutzung von Unternehmensdaten aus dem Internet zur Ergänzung von Unternehmensstatistiken, steht die Entwicklung in Deutschland noch am Anfang. Technisch werden diese Daten mittels des sogenannten Web Scrapings bzw. Crawlings im Internet „gesammelt“ [9].

Nicht-traditionelle Daten sind derzeit ein Forschungsgebiet in den statistischen Ämtern weltweit. Im Rahmen des Europäischen Statistischen Systems wurde in zwei großen Projekten, an denen die Statistischen Ämter der EU sowie weitere europäische Ämter mitgearbeitet haben, neue digitale Daten hinsichtlich ihrer Potentiale, ihrer Qualität und ihrer Zugänglichkeit untersucht. Die Ergebnisse der ESSnet Big Data I und II finden sich unter [https://ec.europa.eu/eurostat/cros/content/essnet-big-data-1\\_en](https://ec.europa.eu/eurostat/cros/content/essnet-big-data-1_en).



## Kriterien zur Beurteilung von Daten

Nachdem in den vorigen Abschnitten ein Überblick über relevante, vorhandene und öffentlich zugängliche Datenquellen gegeben wurde, werden nun Kriterien zur Beurteilung von Daten und Statistiken diskutiert. Aus unserer Sicht sind hier insbesondere die folgenden Punkte zu nennen:

### 1. Eignung für ein Ziel

Daten an sich sind weder gut noch schlecht, sondern nur mehr oder weniger gut geeignet, ein bestimmtes Ziel zu erreichen. Um Daten zu beurteilen, muss zunächst das Ziel bzw. die Fragestellung beschrieben werden, zu dessen Erreichung die Daten einen Beitrag leisten sollen.

Das übergeordnete Ziel aller Daten, die im Laufe der Corona-Pandemie besprochen werden, ist es, als Gesellschaft eine größtmögliche Kontrolle über das Pandemie-Geschehen unter Berücksichtigung möglichst vieler Auswirkungen auf die Gesellschaft zu gewinnen. Dieses übergeordnete Ziel umfasst viele Teilziele, wobei dieselben Daten i.d.R. nicht für die Beantwortung aller Teil-Fragestellungen gleichermaßen gut geeignet sind.

Ein Beispiel stellt die Abhängigkeit der 7-Tage-Inzidenz des RKI pro 100.000 Einwohner vom Übermittlungs-Verzug dar: Die Infektionszahlen werden vom RKI nach Meldedatum der Fälle beim Gesundheitsamt aggregiert. Fälle mit einem Übermittlungs-Verzug von mehr als einer Woche werden der 7-Tage-Inzidenz der vorherigen Woche zugeschrieben. Dies führt zu einer systematischen Unterschätzung der jeweils aktuellen 7-Tage-Inzidenz. Diese Methode ist zwar gut für die Abbildung des bundesweiten Infektionsgeschehens, jedoch nicht optimal für die Entscheidung über zusätzliche Maßnahmen vor Ort. Dort sind lokale und aktuellere Zahlen vorzuziehen [<https://www.rki.de/SharedDocs/FAQ/NCOV2019/gesamt.html>].

### 2. Transparenz

Der Datengewinnungsprozess muss in Bezug auf Herkunft, Zeit und Art transparent gestaltet sein. Fehlen diese Informationen, sind die Daten in der Regel ungeeignet, um Ziele im politischen oder wissenschaftlichen Kontext zu erreichen, bei denen ein Konsens vieler erwünscht ist. Daher sind eine genaue Beschreibung der Datenerhebung sowie datenschutzkonformer Zugang zu den Daten (zur Überprüfung durch geeignet qualifizierte Dritte) eine Mindestanforderung an die Güte aller Daten zur Pandemie, die in den demokratischen Diskurs aufgenommen werden.

### 3. Qualitätsstandards

Daten sind dann gut geeignet für Gesamtübersichten und räumlich-zeitliche Trends, wenn in der Datenerhebung und in der Datenverarbeitung an verschiedenen Orten und über die Zeit klar und einheitlich definiert ist, was man wie erfasst. Zur Standardisierung gehört z.B. eine Harmonisierung von Datenprozessen (Fax vs. Computer, lokale vs. einheitliche Software), die adäquate Schulung der an der Erfassung beteiligten Personen, und ein Monitoring der Prozesse (siehe Grundsätze 4 & 14, Verhaltenskodex für europäische Statistiken [10,10a,10b]).

Streitpunkte bei der Erfassung und Definition von Todesfällen im Zusammenhang mit COVID-19 waren z.B. die Frage, ob man zwischen Patient/innen unterscheidet, die an oder mit Corona gestorben sind (und wie man dies gegebenenfalls feststellen will), ob man

„genesene“ Patient/innen mitzählt, die an den Spätfolgen von COVID-19 versterben, und auch, ob man Verstorbene mitzählt, die COVID-19-Symptome hatten, bei denen aber kein Test vorgenommen wurde. Die Klärung dieser Fragen erweist sich häufig als schwierig (und in der Praxis z.T. als unmöglich), jedoch bieten die von den Gesundheitsämtern gemeldeten Sterbezahlen zumindest ein gewisses Maß an Standardisierung.

#### 4. Wahrhaftigkeit

Daten müssen unabhängig, unparteilich und objektiv erhoben und verarbeitet werden. Insbesondere sollten Interessenskonflikte vermieden werden, um Glaubwürdigkeit und Wahrhaftigkeit nicht zu gefährden (siehe auch Grundsätze 1 und 6, Verhaltenskodex für europäische Statistiken [10,10a,10b]).

#### 5. Beschriebene Fehlerquellen

Die meisten Daten enthalten Fehler, wie z.B. Messfehler, Eingabefehler, Übertragungsfehler oder Fehler, welche aufgrund von Antwortverweigerung entstehen. Mit einer guten Beschreibung der Datenerhebung und der Datenprozessierung (siehe Transparenz oben) können mögliche Fehlerquellen eingeschätzt werden und in die Interpretation einfließen. Gutes Monitoring bekannter und möglicherweise neuer Fehlerquellen sollte etabliert werden. Wo notwendig können Verbesserungen der Prozesse vorgenommen werden (siehe auch Grundsatz 12, Verhaltenskodex für europäische Statistiken [10,10a,10b]). Ein Beispiel in diesem Zusammenhang sind die Wochenend-Effekte in der Übermittlung von Infektionszahlen und der Rückgang der Infektionszahlen in der Zeit „zwischen den Jahren“.

#### 6. Relevanz

Daten müssen relevante Informationen zum Erreichen eines Ziels liefern. Dafür müssen die Daten die für die Zielerreichung benötigten Charakteristika an den richtigen Untersuchungseinheiten messen. Ein Beispiel für die Schwierigkeit, die richtige Messung für ein bestimmtes Konzept zu finden, ist die Frage nach geeigneten Tests für die Bestimmung der Immunität der Bevölkerung gegenüber SARS-CoV-2.

Ein weiteres Beispiel, in diesem Fall für die notwendigen Überlegungen zu den richtigen Untersuchungseinheiten, ist die Wahl zwischen repräsentativen Stichproben und anders erhobenen Daten: Um systematische Verzerrungen bei der Generalisierung zu vermeiden, benötigt man repräsentative Stichproben. Generalisierung ist dann ein impliziter Schritt, wenn man mit vorhandenen Daten Entscheidungen für eine größere Gesamtheit und die Zukunft treffen will. Will man jedoch datenbasierte Erkenntnisse von einer großen Gemeinschaft auf kleinere Einheiten anwenden, die auch eine hohe Auflösung (z.B. auf Ebene von Landkreisen) erlauben, sind geeignete repräsentative Stichproben u.U. sehr aufwändig und teuer (siehe auch Grundsätze 11 und 7, Verhaltenskodex für europäische Statistiken [10,10a,10b]).

#### 7. Aktualität und Genauigkeit

Im Idealfall erfüllen Daten, die zur politischen Entscheidungsfindung genutzt werden, alle zuvor aufgeführten Qualitätskriterien. Aus Daten gewonnene Informationen müssen aber zusätzlich aktuell und von hoher Genauigkeit sein. Darüber hinaus sollte die Genauigkeit der aus Daten gewonnenen Informationen quantifiziert werden.

Jedoch kann man manche Entscheidungen nicht verschieben, um zu warten, bis standardisierte Prozesse definiert und implementiert sind und optimale Daten im Sinne der Kriterien 3-6 gesammelt wurden. Der hiermit verbundenen größeren Unsicherheit in den Daten muss mit Transparenz und mit einer großen Sorgfalt bei ihrer Interpretation begegnet werden.

Besonders zeitkritische Entscheidungen in der COVID-19 Pandemie sind diejenigen über Kontakt-Beschränkungen, da Zögerlichkeit hier direkt zu einer unwiederbringlichen Verschlimmerung des Pandemie-Geschehens führen könnte.

#### 8. Zugang zu Daten für die Wissenschaft

Um das übergeordnete Ziel zu erreichen, als Gesellschaft eine größtmögliche Kontrolle über das Pandemie-Geschehen zu gewinnen, ist es wichtig, gute Daten als Ressource einer weiten, wissenschaftlichen Öffentlichkeit zugänglich zu machen. Damit können die Daten in unterschiedlichen Kontexten und mit verschiedenen Methoden analysiert und die Interpretation der Daten aus der Perspektive verschiedener gesellschaftlicher Gruppen und wissenschaftlicher Disziplinen ermöglicht werden (siehe auch Grundsatz 15, Verhaltenskodex für europäische Statistiken [10,10a,10b]).

Neben der Verfügbarkeit von Daten und der damit verbundenen Qualität ist auch die Vollständigkeit zu überprüfen. Hierbei ist kritisch zu untersuchen, ob die Zielfragestellungen mit Hilfe der vorhandenen Datenquellen adäquat beantwortet werden können oder ob vielmehr weitere Variablen gewonnen werden müssen, etwa im Zusammenhang mit der Durchführung von Tests, oder gar neue Datenerhebungen stattfinden müssen, wie etwa zur Untersuchung der regionalen Inzidenzen auf Repräsentativität.

### **Empfehlungen:**

1.) Daten müssen Anforderungen der Relevanz, Transparenz, Wahrhaftigkeit, Aktualität und Genauigkeit genügen; potentielle Fehlerquellen sollen beschrieben und möglichst auch quantifiziert werden. Zusätzlich sollten diese Daten in Forschungsdatenzentren für die Wissenschaft zugänglich sein, um eine ausgewogene und offene wissenschaftliche Diskussion zu unterstützen, in der sehr unterschiedliche Methoden miteinander vergleichbar gemacht werden.

2.) Die Infektionszahlen hängen stark von der Verfügbarkeit von Tests, der Qualität der Tests und der Teststrategie ab. Trotzdem sind sie die besten frühen Indikatoren, um einzugreifen, wenn sich das Infektionsgeschehen ändert. Auf Dauer ist es gut, gemäß einer standardisierten Strategie regelmäßige Tests in gut definierten Zufallsstichproben durchzuführen, um das Infektionsgeschehen richtiger abbilden zu können. Die Prozesse hierfür können am besten außerhalb einer Pandemie etabliert werden, und z.B. zum Monitoring von anderen Infektionskrankheiten genutzt werden.

3.) Die Übersterblichkeit spielt(e) in der öffentlichen Diskussion eine wichtige Rolle. Um diese zuverlässig zu ermitteln, ist eine nationale Strategie zur Ermittlung und systematischen Sammlung solcher Daten sehr wichtig. Der Vergleich der Übersterblichkeit in verschiedenen Regionen und Ländern ist eine wichtige Grundlage zur Evaluierung von Maßnahmen. Eine Beteiligung an internationalen Initiativen zur standardisierten Datensammlung ist deshalb zu empfehlen.

4.) Zur besseren Bewältigung von zukünftigen Pandemien ist es sinnvoll, Standards und Prozesse zur Datenerhebung zu etablieren, die in den nächsten Jahren den Umgang mit Epidemien und Pandemien verbessern und die den Anforderungen der Relevanz, Transparenz, Wahrhaftigkeit, Aktualität und Genauigkeit genügen. Hierbei sollte aus der aktuellen Pandemie gelernt werden, welche Daten, Variablen und Informationen gewonnen werden sollten, um die wesentlichen politisch und gesellschaftlich relevanten Fragestellungen vollständig und mit der hinreichenden Qualität (auch in Bezug auf die erforderliche Granularität) beantworten zu können.

5.) Über die Corona-Pandemie und zukünftige Pandemien hinaus gibt es weitere Themen, die gemeinsame Anstrengungen von Gesellschaft, Politik und Wissenschaft erfordern. Als wichtige Beispiele seien hier die digitale Partizipation zur Teilhabe aller an politischen Entscheidungsprozessen und die globale Erwärmung genannt. In diesem Sinne ist es wünschenswert, dass fächerübergreifende Konsortien für Forschungsdateninfrastruktur gebildet werden, die gemeinsam mit der demokratischen Gesellschaft die Ziele definieren, von denen ausgehend geeignete Standards und Prozesse zur Datenerhebung für die Beantwortung dieser Fragestellungen etabliert werden können. Insbesondere wollen wir darauf hinweisen, dass bereits bei der initialen Formulierung solcher Fragestellungen Statistiker/innen mit eingebunden werden sollten. Eine partnerschaftliche Zusammenarbeit zwischen der Forschung und der amtlichen Statistik ist bei all diesen Zukunftsthemen wichtig.

## Modellierung und Simulation dynamischer Systeme

Für modellgestützte Entscheidungen lassen sich zwei *Zwecke* unterscheiden:

- i. Entscheidungen, die von einem „Ist-Zustand“ abhängig gemacht werden.
- ii. Entscheidungen, die von einer zukünftigen Entwicklung (und deren Konsequenzen) abhängig gemacht werden.

Ein Beispiel für die Ermittlung eines „Ist-Zustands“ sind Indikatoren (Richt- und Grenzwerte), wie der R-Wert, der während einer Pandemie abschätzt, wie viele Personen eine infizierte Person im Mittel ansteckt. Die Bestimmung solcher Indikatoren oder Kennzahlen geschieht auf der Basis der bisherigen Entwicklung und verfügbarer/historischer Daten. In vielen Fällen gibt es alternative Ansätze, die konkreten Werte für Kennzahlen zu ermitteln. Für Entscheidungen ist es deshalb wichtig, unterschiedliche Methoden zur Ermittlung der Kennzahlen miteinander zu vergleichen.

In Verbindung mit der Ermittlung von Daten zur Berechnung von Kennzahlen gibt es in vielen Fällen unvermeidbare Unsicherheiten. Zu den Quellen von Unsicherheiten gehören unvollständige Datensätze, ungenaue Daten sowie zufällige Faktoren. Eine Konsequenz für Entscheidungen ist, dass in den meisten Fällen nicht eine einzige Zahl, sondern ein Intervall oder eine Verteilung betrachtet werden müssen. Im Alltag erhält man die Vorhersage einer Temperatur für den nächsten Tag oder die Angabe einer Regenwahrscheinlichkeit in Prozent, und dennoch ist allgemein verstanden, dass diese Zahlen ein zu erwartendes Intervall abbilden.

### Vorhersagen zukünftiger Entwicklungen

Der zweite oben genannte Zweck von modellgestützten Entscheidungen verwendet Vorhersagen über die zeitliche und/oder räumliche Entwicklung von Indikatoren und Systemvariablen. Beispiele sind Klima- und Wettervorhersagen, bei denen zu erwartende Temperaturen als Grundlage für politische Entscheidungen eine Rolle spielen, und Windvorhersagen, die für Schifffahrt, Luftfahrt und Freizeitaktivitäten eine Entscheidungsgrundlage bilden.

Für die mathematische Modellierung und Simulation, mit dem Zweck solcher Vorhersagen, kann man weiterhin zwei *Zielsetzungen* unterscheiden:

- i. Präzise Vorhersagen.
- ii. Simulation zum Erklären von Zusammenhängen.

Für viele Fragen besteht der Wunsch nach einer zahlenmäßig akkuraten Vorhersage von Indikatoren, oder dem numerisch akkuraten Verlauf einer Variable. Während einer Pandemie besteht zum Beispiel der Wunsch, die Anzahl der zu erwartenden Neuinfektionen oder die Anzahl der benötigten Betten mit intensivmedizinischer Betreuung in den kommenden Tagen, Wochen und Monaten zahlenmäßig vorherzusagen. Die Zuverlässigkeit oder Genauigkeit solcher Vorhersagen ist stark von der Verfügbarkeit und Qualität der Daten abhängig, mit deren Hilfe die Werte von Parametern in den zugrundeliegenden mathematischen oder statistischen Modellen geschätzt werden.

Für viele Fragestellungen, und z.B. auch wenn die Datenlage keine zuverlässige Parametrisierung der Modelle zulässt, können Vorhersagemodelle zur Simulation von Szenarien eingesetzt werden. Hierbei wird kein Anspruch auf numerisch exakte Vorhersagen erhoben, sondern untersucht, welche Faktoren eine Rolle spielen und mit welchem Ausmaß zu rechnen ist.

## Modelle zur Simulation von Szenarien

Das Infektionsgeschehen in Pandemien kann mathematisch modelliert werden. Wie oben beschrieben gibt es zwei zu unterscheidende Zwecke für solche Modelle: (i) Die zahlenmäßig möglichst korrekte Vorhersage, z.B. von zu erwartenden Neuinfektionen; und (ii) die Untersuchung von Szenarien ohne Anspruch auf numerische Präzision. Im letzteren Fall liegt der Schwerpunkt auf dem Erklären und dem Verstehen von Vorgängen [11].

Das Modell des exponentiellen Wachstums kann als Beispiel dienen, die beiden Zielsetzungen zu unterscheiden. Zu Beginn der Pandemie, und z.B. zu Beginn der zweiten Welle von Mitte September bis Ende Oktober 2020 ließ sich die Zahl von intensivmedizinisch behandelten COVID-19-Patient/innen auf Intensivstationen mit dem Modell des exponentiellen Wachstums vorhersagen. Der Zweck eines solchen Modells war jedoch vor allem ein Hilfsmittel zur Unterstützung der Entscheidungsfindung. Der explosive Anstieg verdeutlichte die Notwendigkeit zum Handeln.

Andere Modelle, für Vorhersagen über größere Zeiträume, halfen im Frühjahr 2020 das Auftreten einer zweiten Welle zu erklären (und z.B. dass zu erwarten sei, dass diese größer ausfällt als die erste). Die dafür bekanntesten Modelle (SIR, SEIR) teilen die Bevölkerung in Gruppen ein (z.B. Suszeptible, Exponierte, Infizierte, Gesundete/Verstorbene) und simulieren dann zeitliche Veränderungen in diesen Gruppen in unterschiedlichen Szenarien.

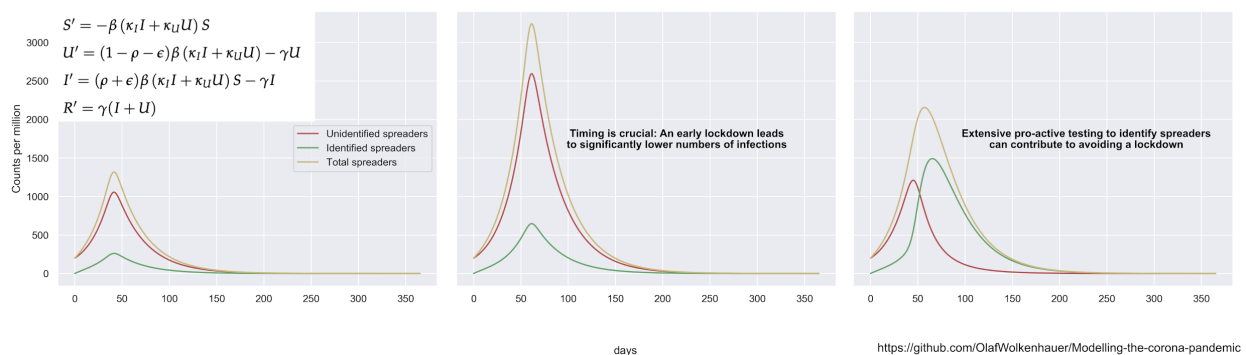


Abbildung 3: Der Einfluss verschiedener Interventionen auf den weiteren Pandemieverlauf anhand eines einfachen SIR-Modells. Quelle: <https://github.com/OlafWolkenhauer/Modelling-the-corona-pandemic>.

Der exponentielle Anstieg von Neuinfektionen und COVID-19-Patient/innen auf Intensivstationen in der zweiten Welle, aber auch die Vorhersage des Auftretens einer zweiten Welle wurden in der aktuellen Pandemie lange bezweifelt. Vorhersagemodelle leisteten hier einen Beitrag zum besseren Verständnis der Zusammenhänge und der Dynamik des Infektionsgeschehens. Ein Beispiel für solche Modelle zeigt Abbildung 3.

Ein sehr weit verbreiteter Ansatz für die Vorhersage zeitlicher Verläufe sind Differenzialgleichungsmodelle. Die diesen Modellen zu Grunde liegenden Gleichungen beschreiben die Geschwindigkeiten, mit denen sich die Werte der Variablen eines dynamischen Systems über die Zeit hinweg ändern. Solche Modelle sind per Definition eine reduzierte Abbildung, d.h. es werden bewusst Details in der Beschreibung ausgelassen. Die pauschale Kritik, solche Modelle seien unrealistisch, weil viele Faktoren nicht berücksichtigt würden, geht deshalb an der Zielsetzung der

Modelle vorbei. Relativ einfache epidemiologische Differenzialgleichungsmodelle können z.B. sehr gut darstellen, dass ein frühes Handeln notwendig ist. Solche Modelle können auch aufzeigen, in welchen Zeiträumen bei der Implementierung von Maßnahmen zu denken ist.

Viele Entscheidungen erfordern ein Abwägen von Evidenz und ein Verständnis biologischer, medizinischer oder physikalischer Zusammenhänge oder Mechanismen [12]. Hier spielt die Simulation von Szenarien eine wichtige Rolle, z.B. in der Entwicklung und Validierung von Hypothesen über Maßnahmen zur Reduktion von Kontakten oder zu Impf- und Teststrategien, die im nächsten Kapitel weiter diskutiert werden.

Beispiele von „simplen“ Differenzialgleichungsmodellen, die zwar nicht numerisch realistisch sind, aber dennoch wertvolle Beiträge zum Verständnis geliefert haben, sind (i) das meteorologische Modell von Edward Lorenz („Schmetterlingseffekt“, drei Gleichungen); (ii) das Lotka-Volterra Modell zur Erklärung zyklischer Änderungen in Tierpopulationen (zwei Gleichungen); oder (iii) das Hodgkin-Huxley Modell zu Potenzialen in Neuronen (drei Gleichungen; verantwortlich für einen Nobelpreis in der Medizin).

Wie auch bei statistischen Modellen oder Modellen des maschinellen Lernens müssen Vorhersagen solcher Modelle in den vorliegenden Kontext eingeordnet werden. Insbesondere muss deutlich gemacht werden, dass die Vorhersagen nicht darauf abzielen, zahlenmäßig genau zu sein, sondern Trends, Relationen und dynamische Muster zu verdeutlichen. Weil das Verhalten der Bevölkerung ebenso wie Reglementierungen einen starken Einfluss auf den Verlauf des Infektionsgeschehens haben, sind Simulationen solcher Szenarien über längere Zeiträume z.B nicht mit Wettervorhersagen zu vergleichen. Die Simulation von Handlungsszenarien untersucht die Frage „Was könnte passieren?“ und nicht die Frage „Was wird passieren?“.

#### Validierung von Modellen dynamischer Systeme

Die Frage nach der „Qualität“ eines Modells dynamischer Systeme muss sich an der Zielsetzung des Modells orientieren. Modelle komplexer Systeme sind per Definition eine vereinfachte Darstellung. Modelle, deren Ziel eine zahlenmäßige Vorhersage ist, können ihre Genauigkeit an historischen Daten demonstrieren. Dies trifft insbesondere für retrospektive Analysen (Now-casting) zu.

Beim Forecasting beeinflussen die Vorhersagen in vielen Fällen das Verhalten der Menschen, so dass eine rückblickende Validierung schwierig ist. Modelle zur Simulation von Szenarien sind Gedankenexperimente, mit denen man qualitative Vorhersagen macht. Ein Beispiel ist die Untersuchung unterschiedlicher Handlungsszenarien für einen Lockdown oder Teststrategien. Bei Maßnahmen zur Einschränkung von Kontakten stellen sich Fragen zum Zeitpunkt der Maßnahmen und deren Dauer.

In der Statistik sowie Entscheidungsanalyse hat der Begriff „Validierung“ umfassende Definitionen [13]. Man unterscheidet zwischen

- a) „Face validity“ (interdisziplinäre Fachexperten prüfen Modellstruktur, Parameterannahmen, Ergebnisse etc.)
- b) Verifizierung oder interne Validität (Prüfung der korrekten Implementierung und Programmierung von Algorithmen)

- c) modellübergreifende Validität (Vergleich der Modellierungsergebnissen unterschiedlicher Modelle mit ähnlichen Annahmen und Datengrundlage)
- d) externe Validität (Vergleich von Modellergebnissen oder Teilergebnissen mit „real world data“)
- e) prädiktive Validität (Vergleich von zukünftigen „real world data“, die durch Modelle vorhergesagt werden).

Bei der Validierung mechanistischer Modelle dynamischer Systeme liegt der Schwerpunkt in der Reflektion des Modellbildungsprozesses und weniger in dem Modell als Ergebnis. Die Wahl der im Modell berücksichtigten Variablen und die Bestimmung von Parameterwerten werden dabei ständig angepasst, um die Vertrauenswürdigkeit der Modelle zu erhöhen. Ein Beispiel sind Modelle, die das Infektionsgeschehen in unterschiedlichen Altersgruppen berücksichtigen. Solche Modelle können Entscheidungen für Schulschließungen unterstützen. Das Infektionsgeschehen bei Kindern und Jugendlichen wird in klinischen Studien ermittelt, die jedoch mehrere Monate in Anspruch nehmen. Daher sind die Parameterwerte in solchen Modellen stark davon abhängig, was über das Virus und das Infektionsgeschehen bekannt ist. Auch hier wird deutlich, dass es sich bei der Modellierung um einen Prozess handelt, der in Zeiten von Pandemien kontinuierlich und in Kooperation unterschiedlicher Expertisen realisiert werden muss.



### **Empfehlungen:**

6.) Da sich während Pandemien die Datenlage und die Fragestellungen kontinuierlich entwickeln, sollten Modellierungsprojekte als Prozess aufgefasst und kommuniziert werden. Die Ausgangslage für die Modellierung und Simulation dynamischer Systeme entwickelt sich in einer Pandemie nahezu täglich. Solche Projekte benötigen deshalb die volle Aufmerksamkeit der beteiligten Modellierer/innen als Vollzeitbeschäftigung. Die Bildung von multidisziplinären Arbeitsgruppen ist zu empfehlen.

7.) Die Vielfalt von Modellen ist durch die Vielzahl von Fragen begründet. Modelle sind kontextspezifisch oder kontextgebunden. Die Vielfalt von Fragestellungen bedingt eine Vielzahl von Modellen, die mit sehr unterschiedlichen Ansätzen untersucht werden können. Syntaktisch ähnliche Modelle können sehr unterschiedliche Zwecke erfüllen. Aus diesem Grund ist die Kommunikation der Zielsetzung für ein Modell sehr wichtig. Eine fächerübergreifende Zusammenarbeit von Modellierern/innen und ein kritischer Austausch in institutsübergreifenden Arbeitsgruppen ist deshalb dringend zu empfehlen.

8.) Da politische Entscheidungen und die Wahrnehmung der Bevölkerung zum Infektionsgeschehen dieses selbst durch das Verhalten der Menschen beeinflussen, ist es besonders wichtig, solche unvermeidbaren Unsicherheiten bei Vorhersagen zu kommunizieren. Eine Validierung im üblichen Sinne ist nur schwer möglich, weil sich die Prozesse, die modelliert werden, durch das Verhalten der Menschen ständig verändern. Um dennoch Vertrauen in solche Analysen zu schaffen, sollten Daten, Modelle, Programm-Code sowie die Wahl der Methoden und Annahmen transparent und reproduzierbar kommuniziert werden.

9.) Es lassen sich unterschiedliche Zielsetzungen für die Modellierung unterscheiden. Die jeweilige Zielsetzung muss kommuniziert werden, um eine Desinformation und Um-Interpretation zu vermeiden bzw. zu verhindern. Bei Modellen mit dem Ziel der Vorhersage numerischer Werte in der Zukunft werden Modellstruktur und Parameterwahl als angemessene Abbildung der Realität angenommen. Die Struktur, Parameterwerte und Annahmen müssen deshalb erklärt werden. Bei der Modellierung als Prozess, um Zusammenhänge oder Mechanismen zu verstehen, ist die Etablierung der Struktur des Modells das Ziel. Simulationen sollten in diesem Zusammenhang als Zwischenschritt in einem Prozess des "Verstehens" von zugrundeliegenden Prinzipien deklariert werden. Neben dem Ziel zahlgenauer Vorhersagen haben Modelle raumzeitlicher Prozesse eine Bedeutung, um Szenarien zu simulieren ("Gedankenexperimente"). Die oft fehlende Nähe zu konkreten Zahlen sollte transparent gemacht werden. Aufgrund der Komplexität solcher Modellierungsvorhaben kommt der Kommunikation eine besondere Bedeutung zu. Weiterbildungen im Bereich der Wissenschaftskommunikation sind deshalb sehr zu empfehlen.

## Decision Making

Nun wenden wir uns der Frage zu, wie basierend auf statistischen Kennzahlen politische Entscheidungen abgeleitet werden können. Basierend auf statistischen Kennzahlen müssen zunächst mögliche Maßnahmen definiert werden. Diese Maßnahmen sollten hinsichtlich ihrer kurz- und langfristigen Konsequenzen mittels Methoden der Health Decision Science evaluiert werden. Hierbei sind die Dynamik nach der Implementierung von Maßnahmen und entsprechende Auswirkungen auf das Umfeld und auf weitere Maßnahmen zu berücksichtigen. Kurz- und langfristige Auswirkungen von Maßnahmen oder Interventionen können abhängig von der Forschungsfrage mit unterschiedlichen Modellierungs-/ Simulationsmethoden ermittelt werden. Die Evaluationen und die Berücksichtigung zusätzlicher Aspekte führen in weiterer Folge zu staatlichem Handeln und Festlegung von Maßnahmen. Hierbei ist insbesondere die Akzeptanz in der Bevölkerung zu berücksichtigen.

### **Evidenzbasierte Politik: Der Nutzen statistischer Analysen bei politischer Entscheidungsfindung**

Politische Entscheidungsprozesse reagieren auf statistische Kennzahlen. Dabei ist zu beachten, dass die statistischen Analysen und Kennzahlen keine direkten Handlungsempfehlungen liefern. Aus keinen noch so hohen Infektionszahlen, Todeszahlen oder Reproduktionszahlen kann unmittelbar abgeleitet werden, was zu tun ist bzw. wie das Verhältnis zwischen Nutzen und Risiko sowie ökonomischen und ethischen Aspekten zu bewerten ist. Man benötigt vielmehr Kenntnis darüber, welchen Effekt welche Intervention auf das Gesamtgeschehen hat. Dazu muss man die kausalen Zusammenhänge zwischen den Maßnahmen und die durch sie bewirkten „physikalischen“ Effekte (z.B. Kontaktreduktion) und psychologisch bedingten Verhaltensänderungen kennen. Letztere beruhen allerdings auf komplexen Handlungsmodellen: bestimmte Informationen und Einstellungen (Vertrauen in die Regierung, Bereitschaft auf Anordnungen der Regierung zu hören (compliance) etc.), die wiederum mit situativen Perzeptionen interagieren (Variation der Wahrnehmung des eigenen Infektionsrisikos bei unterschiedlichen Kontaktpersonen, z.B. Familienmitgliedern und guten Bekannten vs. Fremden) führen zur Herausbildung bestimmter Verhaltensdispositionen und damit letztendlich zu diesem Verhalten selbst.

Politische Entscheidungen sind grundsätzlich darauf angewiesen, Akzeptanz bei der Bevölkerung zu finden. Dies berührt zwei wesentliche Aspekte: Die Legitimation und die politische Verantwortlichkeit (Accountability), die zentral für das Funktionieren eines demokratischen Staates ist.

### **Accountability und Kennzahlen**

Ein wesentliches Element von Demokratien besteht darin, dass Regierungen gewählt werden. Idealerweise sollte „gute“ Regierungsarbeit zu einer Wiederwahl und „schlechte“ Regierungsarbeit entsprechend zu einer Abwahl der jeweiligen Regierung führen. Die vorhandenen statistischen Analysen zur Pandemie erfüllen damit zweierlei Zwecke. Erstens verwendet die Regierung statistische Analysen und wissenschaftliche Modelle, um abzuschätzen, welche Art von Intervention welchen Effekt auf die Kennzahlen haben. Zweitens sind die Analysen von fundamentaler Bedeutung für Bürger/innen, da sie Daten und Kennzahlen (z.B. Infektionsraten, Todeszahlen, Ausschöpfung von Intensivbehandlungskapazitäten, Exzess-Todesrate in nicht Corona-Kontexten durch Corona-bedingte Ressourcenknappheit) zur Verfügung stellen. Anhand dieser Daten kann die Regierungsleistung bewertet und somit eine Wahlentscheidung getroffen werden. In Vorwegnahme dieser Rechenschaftsablegung gegenüber den Bürger/innen wiederum trifft die Regierung ihre

Entscheidungen. Die statistischen Informationen ermöglichen so die demokratische politische Steuerung. Dabei können zur Unterstützung evidenzbasierter politischer Entscheidungen mit Hilfe der Methoden der Health Decision Science kurz- und langfristige Konsequenzen (Outcomes) von Maßnahmen evaluiert werden.

### **Health Decision Science Framework**

Das Health Decision Science Framework bietet eine Grundlage für die Evaluation von Maßnahmen und das Zusammenführen von Daten und Evidenz für die medizinische und gesundheitspolitische Entscheidungsfindung. Wie in anderen Bereichen (Statistik, Epidemiologie, Public Health) wird in den (gesundheitsbezogenen) Entscheidungswissenschaften das PICO Schema (P-Population, I-Intervention, C-Comparator, O-Outcomes) verwendet und durch die Perspektive der Analyse (z.B. gesamtgesellschaftlich, Versicherungsträger) und den Zeithorizont (PICOST; S-Setting, T-Timing) weiter spezifiziert.

Als Instrument der Analyse wird dabei auf die sogenannte Entscheidungsanalyse bzw. entscheidungsanalytische Modellierung zurückgegriffen. Eine Entscheidungsanalyse verfolgt das Ziel, mittels systematischer, expliziter und quantitativer Methoden Entscheidungen unter Unsicherheit zu unterstützen. Hierbei kommen insbesondere Computersimulationen zum Einsatz, um den kurz- und langfristigen Nutzen und Schaden (sowie die Kosten) alternativer Interventionen, Technologien oder Maßnahmen in der Gesundheitsversorgung zu berechnen [14]. Im Vergleich zu Prognosemodellen [15], die für die Vorhersage epidemiologischer Kennzahlen des COVID-19-Infektionsgeschehens (z.B. der Vorhersage zukünftiger Neuinfektionen, von Todesfällen oder Hospitalisierungen, siehe vorheriges Kapitel) verwendet werden oder prognostischen Modellen [16, 17], bei denen basierend auf Kombination mehrerer Prädiktoren Risiken für einen bestimmten Endpunkt (z.B. schwere COVID-19-Krankheitsverläufe oder Tod) für einzelne Patienten berechnet werden, werden mittels Entscheidungsanalyse die Konsequenzen von alternativen Maßnahmen evaluiert.

Das entscheidungsanalytische Framework beinhaltet neben dem gewählten Modelltyp unter anderem die berücksichtigten relevanten Gesundheitszustände und Ereignisse zur Beschreibung möglicher Krankheitsverläufe, die Art der Analyse (z.B. Benefit-Harm, Kosten-Nutzen, Budget-Impact Analysen [18]) und die Simulationsmethode (kohorten- oder individuenbasiert). Neben der Basisfallanalyse (Verwendung der wahrscheinlichsten Parameter) sind Szenario- und Sensitivitätsanalysen [19] durchzuführen, um die Robustheit bzw. Unsicherheit der Ergebnisse darzustellen.

Ein zentraler Gedanke bei der Entscheidungsanalyse ist, dass Trade-Offs formalisiert und wenn möglich quantifiziert werden. Im Rahmen einer Benefit-Harm-Analyse betrifft dies beispielsweise die Quantifizierung des Nutzens einer COVID-19 Impfung in Form von vermiedenen Todesfällen sowie den Schaden einer Impfung in Form von möglichen Nebenwirkungen. Generell können zwei oder mehrere Interventionen in einer schrittweise inkrementellen Betrachtung verglichen werden. Als Entscheidungsgrundlage sollten vordergründig die Unterschiede in den Outcomes, die im Vergleich (z.B. Impfen versus nicht Impfen oder Impfen Gruppe A versus Impfen Gruppe B) von Maßnahmen erzielt werden können und nicht die geschätzten absoluten Outcomes dienen. Eine grafische Darstellungsmöglichkeit sind „Efficiency Frontiers“ [20], welche den inkrementellen Schaden dem inkrementellen Nutzen gegenüberstellen. Wichtig ist hierbei, dass man zunächst dominierte Maßnahmen identifiziert und ausschließt. Dominierte Maßnahmen weisen einen größeren Schaden auf

und bieten einen geringeren Nutzen als andere Maßnahmen (dominiert) oder haben ein höheres inkrementelles Schaden-Nutzen-Verhältnis als alternative Maßnahmen (schwach dominiert). Im Gegensatz dazu führen dominante Maßnahmen zu einem größeren Nutzen bei geringerem Schaden im Vergleich zu anderen Maßnahmen und sollten daher umgesetzt werden. Oftmals führen alternative Maßnahmen aber mit zunehmendem Nutzen auch zu einer Verschlechterung im Sinne eines Schadens. Beispielsweise können geringere gesundheitliche Spätfolgen aufgrund von vermiedenen SARS-CoV-2-Infektionen (Nutzen) bei Schulschließungen verlorenen Bildungsjahren (Schaden) gegenübergestellt werden. Welche Maßnahmen generell dargestellt und diskutiert werden, beeinflusst hierbei auch das Entscheidungsverhalten [21]. Gleiches gilt für Veränderung der Entscheidung durch Alternativen, die dargestellt werden. Hinsichtlich der Optimierung von Impfinterventionen können zeitliche Aspekte der Verfügbarkeit und Effektivität von Impfungen berücksichtigt werden. Alternative Strategien (beispielsweise die sofortige Impfung mit einem geringeren Impfschutz im Vergleich zur späteren Impfung mit einer erwarteten höheren Effektivität, aber dem Risiko einer zwischenzeitlichen Infektion) können evaluiert werden.

Eine Besonderheit bei der Beurteilung von Pandemien ist der dynamische Verlauf des Geschehens, der einen dynamischen Entscheidungsprozess nach sich zieht.

## Evaluation von Impfstrategien

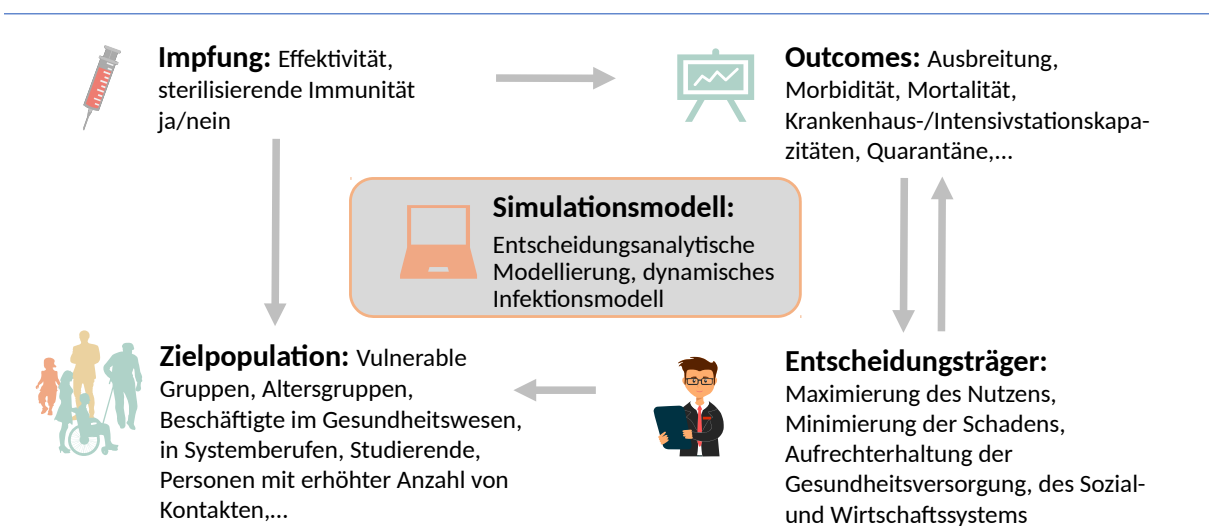


Abbildung 4: Health Decision Science Framework am Beispiel der Evaluation von Impfstrategien.

### Dynamischer Entscheidungsprozess

Die Entscheidung „Lockdown“ vs. „Ungebremschte wirtschaftliche Tätigkeit“ ist keine Einzelentscheidung, bei der statische Auszahlungen gegeneinander abgewogen werden können. Vielmehr handelt es sich auch um Entscheidungen, die andere und neue Entscheidungswege öffnen bzw. bestimmte Entscheidungswege schließen. Insofern sind z.B. die Einschränkungen eines

Lockdowns im guten Fall „nur“ der Preis, den man dafür zahlt, mit den Inzidenzwerten so weit herunter zu kommen, dass man danach das Niveau mit weit weniger schwerwiegenden Maßnahmen stabil halten kann. Die Entscheidung für solche Maßnahmen kann vor allem zum Ziel haben, Situationen zu verhindern, in denen Entscheidungen zu moralischen Dilemmata (wie Triage-Entscheidungen, siehe z.B. [69]) führen.

## **Modellierungsmethoden**

Für die entscheidungsanalytische Modellierung und Entwicklung von Computersimulationen kommen abhängig von der Forschungsfrage unterschiedliche Modellansätze zur Anwendung. Hierzu zählen: Entscheidungsbaummodelle, State-Transition-Models (Zustandsübergangsmodelle, Markov-Modelle), Discrete Event Simulation (diskrete Ereignissimulationsmodelle), agentenbasierte Modelle und dynamische Transmissionsmodelle.

### Entscheidungsbaummodelle

In einem Entscheidungsbaummodell werden die Konsequenzen von alternativen Interventionen oder Gesundheitstechnologien durch mögliche Pfade beschrieben. Entscheidungsbäume beginnen mit Entscheidungsknoten, gefolgt von alternativen Wahlmöglichkeiten (Interventionen, Technologien, etc.) des Entscheidungsträgers. Für jede Alternative werden danach die Pfade der Patient/innen, die durch den Zufall bestimmt werden und außerhalb der Kontrolle des Entscheidungsträgers liegen, durch Zufallsknoten beschrieben. Am Ende der Pfade werden die jeweiligen Folgen jedes Pfades angezeigt. Zu den Folgen oder Ergebnissen können Symptome, Überleben, Überleben unter Berücksichtigung der Lebensqualität, Todesfälle oder Kosten gehören. Abschließend werden die erwarteten Ergebnisse jeder alternativen Wahl durch Mittelwertbildung (d.h. der gewichtete Durchschnitt eines gesamten Pfades) berechnet [22, 23].

### Zustandsübergangsmodelle

Ein Zustandsübergangsmodell wird in Form einer Menge von (Gesundheits-)Zuständen und Übergängen zwischen diesen Zuständen konzeptualisiert. Die Zeit wird in Zeitintervallen abgebildet. Übergangswahrscheinlichkeiten, Zeitzykluslänge, Zustandswerte ("Rewards") und Abbruchkriterien werden im Vorfeld definiert. Während der Simulationen können sich Individuen in jedem Zyklus nur in einem Zustand befinden. Pfade von Individuen, die durch Ereignisse während eines Zyklus bestimmt werden, werden mit einem Markov-Zyklusbaum modelliert, der eine Reihe von Zufallsknoten verwendet. Die durchschnittliche Anzahl von Zyklen, in denen sich Individuen in jedem Zustand befinden, kann in Verbindung mit den Rewards (z.B. Lebensjahre, gesundheitsbezogene Lebensqualität oder Kosten) verwendet werden, um die Konsequenzen hinsichtlich Lebenserwartung, Lebenserwartung unter Berücksichtigung der Lebensqualität und erwarteten Kosten alternativer Interventionen oder Gesundheitstechnologien zu schätzen. Es gibt zwei gängige Arten von Analysen von Zustandsübergangsmodellen: Kohortenmodelle ("Markov") [24, 25] und Modelle auf Individualebene ("first-order Monte Carlo", "Mikrosimulations"-Modelle) [26-28].

### Diskrete Ereignissimulation (DES)

Die diskrete Ereignissimulation ist eine Simulation auf individueller Ebene, die in den 1960er Jahren in der Industrietechnik entwickelt wurde, aber in den letzten Jahren vermehrt im Gesundheitswesen

Anwendung findet [29-32]. Die Kernkonzepte der DES sind Entitäten (z.B. Patient/innen), Attribute (z. B. Patientencharakteristika), Ereignisse, Ressourcen (d.h. physische Ressourcen wie ärztliches Personal und medizinische Geräte), Warteschlangen und Zeit [29, 33, 34]. Ähnlich wie bei den Entscheidungsbäumen und Zustandsübergangsmodellen können gesundheitliche Outcomes und Kosten alternativer Gesundheitstechnologien bewertet werden. Zusätzlich zu diesen Ergebnissen können Leistungsmaße wie Ressourcennutzung oder Wartezeiten berechnet werden, da physische Ressourcen explizit modelliert werden können [34, 35]. Der Begriff diskret bezieht sich auf die Tatsache, dass DES sich in diskreten Intervallen in der Zeit vorwärts bewegt (d.h. das Modell springt vom Zeitpunkt eines Ereignisses zum Zeitpunkt des nächsten) und dass die Ereignisse diskret (sich gegenseitig ausschließend) sind [30].

### Agentenbasierte Modelle (ABM)

Die agentenbasierte Modellierung ist eine Form der Simulation auf individueller Ebene [30]. ABMs wurden zur Modellierung von biologischen Prozessen, Verkehrsmanagement, Kundenflußmanagement oder Aktienmärkten eingesetzt, und in den letzten Jahren vermehrt auch für Kosteneffektivitätsanalysen im Gesundheitswesen [36-38]. ABM bilden komplexe Systeme ab, bei denen einzelne 'Agenten' autonom handeln und zu Interaktionen fähig sind [39]. Diese Agenten können die Heterogenität von Individuen abbilden, und das Verhalten der Individuen kann durch einfache Regeln beschrieben werden. Solche Regeln beinhalten, wie Agenten interagieren, sich zwischen geographischen Zonen bewegen, Haushalte bilden oder konsumieren [40-42]. ABM werden oft angewendet, um „Emergent Behaviour“ als Ergebnis dieser vordefinierten Regeln zu untersuchen. Bei der Modellierung von Infektionskrankheiten führen die Verhaltensweisen der Agenten in Kombination mit den Übertragungsmustern und dem Krankheitsverlauf zu einer populationsweiten Dynamik, wie z. B. dem Ausbruch einer Krankheit [43]. Einen Vergleich von ABM, DES und Systemdynamik findet man in [36, 44, 45].

### Dynamische Transmissionsmodelle

Für die Modellierung von Infektionskrankheiten können abhängig von der Fragestellung unterschiedliche Modellierungsansätze verwendet werden [46-48]. Hierzu zählen mathematische Modellierungen (deterministische und stochastische Kompartimentmodelle basierend auf Differentialgleichungssystemen) wie im Kapitel Modellierung und Simulation dynamischer Systeme beschrieben sowie Systemdynamik und agentenbasierte Modellierungen. Deterministische Kompartimentmodelle (Kohortensimulationen) können für die Modellierung des durchschnittlichen Verhaltens von Krankheitsepidemien in großen Populationen unter Berücksichtigung einer kleinen Anzahl von Gruppen, wenig Heterogenität zwischen den Individuen sowie einer hohen Wahrscheinlichkeit der Übertragung eingesetzt werden. Stochastische Kompartimentmodelle sind insbesondere geeignet bei niedriger Prävalenz, kleinen Populationen, einer geringen Übertragungswahrscheinlichkeit und Epidemien in frühen Phasen oder nahe der Eliminierung. Wenn stochastische Effekte (z.B. das Aussterben von Krankheiten in kleinen Populationen), komplexe Interaktionen oder nicht-zufällige Verteilungen (z.B. die Ausbreitung von Krankheiten in Netzwerken) wichtig sind, sind stochastische agentenbasierte Ansätze (individuenbasierte Simulationen) zu bevorzugen [45].

Einen Überblick zu bestehenden COVID-19-Simulationsmodellen bietet die Datenbank der Society for Medical Decision Making (SMDM [49]). Eine Plattform zum wissenschaftlichen Austausch zu Modellierungsstudien bietet das Kompetenznetz Public Health COVID-19 [50].

### **Konkretes staatliches Handeln**

Mittels Entscheidungsanalysen können die Konsequenzen von verschiedenen Maßnahmen (mehr oder weniger genau) abgeschätzt und für eine Abwägung gegenübergestellt werden.

Um die Akzeptanz staatlichen Handelns abzuschätzen, kann man sich an den modelltypischen Abwägungen eines vernünftigen Individuums orientieren. Die relevante Entscheidung ist daher nicht eine banale Maximierungsentscheidung (Maximierung der Lebenserwartung oder Minimierung der Todeszahlen), sondern sie muss Abwägungen (*Trade-Offs*) vornehmen bezüglich des Ausmaßes, in dem sie Möglichkeiten und Freiheiten einschränken will, um z.B. eine Erhöhung der Überlebenswahrscheinlichkeit zu erzielen. Die Trade-Offs sind im entscheidungsanalytischen Modell abzubilden und zu evaluieren.

Eine elegante Form, diese Trade-Offs sichtbar zu machen, besteht im Rahmen der stochastischen Spieltheorie im Konzept des erwarteten Nutzens. Dieses beruht darauf, dass für die Bewertung einer Handlungsoption nicht nur die möglichen Ergebnisse der Handlungen berücksichtigt werden sondern auch die jeweiligen Wahrscheinlichkeiten. Zudem müssen die Ergebnisse mit einem Nutzen versehen werden.

Im Rahmen einer Pandemie sind die Trade-Offs oft sehr komplex und die Entscheidungen deshalb komplizierter. Es geht nicht nur um den Trade-Off zwischen gewonnenen Lebensjahren und wirtschaftlichen Nachteilen eines Lockdowns, sondern auch um viele andere Aspekte wie Lebensqualität und Aspekte von moralisch signifikantem Wert (z.B. Besuchsmöglichkeiten unserer älteren Verwandten, sozialer Wert von zwischenmenschlichen Kontakten an sich, die Freiheit der Personen, Wert des Erhalts einer Kulturszene, etc.).

Diese letztgenannten Trade-Offs können oftmals nur teilweise in Simulationsmodellen abgebildet werden. Die Politik muss deshalb zusätzlich zu den Modellergebnissen in die politische Gesamtabwägung soziale, ethische/moralische, rechtliche und weiterführende Aspekte wie Akzeptanz und Umsetzbarkeit in die Entscheidung einbeziehen.

### **Psychologie der Wahrnehmungen**

Für die Politik besteht ein Problem darin, dass sie die Akzeptanz der Bevölkerung für diejenigen Maßnahmen erhält, die aus Sicht der Bevölkerung vernünftig erscheinen und nicht unbedingt für die, die vernünftig *sind*. Typische Verzerrungen betreffen zum Beispiel die Wahrnehmung von Wahrscheinlichkeiten von Ereignissen, die wir nicht persönlich erfahren haben. Während des Frühjahrs z.B. bestand das Problem, dass auch bei relativ hohen Inzidenzen von 50 auf 100.000 Personen die Wahrscheinlichkeit, persönlich einen betroffenen Fall zu kennen, immer noch sehr gering war.

Hier steht die Politik also vor der Herausforderung, der Bevölkerung ein angemessenes Verständnis von Wahrscheinlichkeiten und dem dynamischen Verlauf einer Pandemie zu vermitteln („Flatten the curve“, exponentielles Wachstum). Gleichzeitig muss die Politik nicht nur diese harten, objektiven

Daten verständlich darstellen, sondern auch die Werte der Güter, um die es bei den Trade-Off-Entscheidungen geht, auf anschauliche Weise vermitteln.

Eine erfolgreiche Entscheidungsunterstützung und eine erfolgreiche Umsetzung von staatlichen Maßnahmen erfordert deshalb eine verständliche Kommunikation von Maßzahlen und zu erwartenden Ergebnissen und Trade-Offs. Dies wird im anschließenden Kapitel genauer diskutiert.

#### **Empfehlungen:**

10.) Um Nachvollziehbarkeit und Akzeptanz der Entscheidungen zu erreichen, müssen alle Schritte des Entscheidungsfindungsprozesses offen gelegt werden. Insbesondere müssen die Gründe für die jeweiligen Entscheidungen genannt werden, warum also z.B. welche Daten erhoben werden bzw. in die Modellierung eingehen.

11.) Insbesondere sollte offen gelegt werden, dass Entscheidungen von bestimmten Annahmen oder Zielsetzungen abhängen. So ist vor allem bei der Auswahl des jeweiligen Modelltyps, dessen Ergebnisse als Grundlage für Empfehlungen dienen sollen, die Abhängigkeit von der damit modellierten Forschungsfrage bzw. Problemstellung und damit die bedingte Aussagekraft der Ergebnisse klar zu machen.

12.) Es ist darauf zu achten, dass Simulationen mit und ohne Maßnahme "in der gleichen Welt" simuliert werden (unit homogeneity [67] oder ceteris paribus Effekt). Veränderbare Interventionen und vorgegebene Parameter in der Natur bestimmen dabei, was eine Maßnahme ist und womit eine "gleiche Welt" beschrieben wird.

13.) Um Trade-Offs zwischen verschiedenen Aspekten der Ergebnisse sinnvoll gegeneinander abwägen zu können, müssen die betrachteten Szenarien möglichst vollständig sein, d.h. alle für die politischen Entscheidungen relevanten Aspekte und Ergebnisparameter umfassen.



## Statistical Literacy und visuelle Kommunikation in Zeiten von COVID-19

In diesem Kapitel werden weiterführende Anforderungen an die Kommunikation von Maßzahlen, zu erwarteten Ergebnissen und Trade-Offs mit Entscheidungsträgern und der Öffentlichkeit diskutiert.

### Pandemie und Öffentlichkeit

Im Zuge der COVID-19-Pandemie erlebte die Berichterstattung über Infektionszahlen und abgeleitete epidemiologische Kennzahlen einen wahren Boom und belegte mit dramatischer Deutlichkeit die Notwendigkeit statistischer Bildung (Statistical Literacy oder Data Literacy – im Rahmen dieser Stellungnahme unterscheiden wir nicht zwischen diesen Begriffen) auf allen Ebenen der Gesellschaft, von einzelnen Bürger/innen über Verbände und Organisationen bis hin zu politischen Entscheidungsträgern. Die Corona-Krise hat eindrücklich ins allgemeine Bewusstsein gebracht, dass unser soziales Zusammenleben und politische Entscheidungen essentiell auf Daten, dem Abwägen von Risiken und damit auf Wahrscheinlichkeitseinschätzungen beruhen.

Das Virus kann nur durch gemeinsame Anstrengungen, sowohl auf internationaler Ebene wie auch durch das Mitwirken aller Mitglieder der Gesellschaft, überwunden werden. Die Bereitschaft zur Einhaltung beschlossener Maßnahmen verlangt von Bürger/innen ein Verstehen von Zusammenhängen, Fakten und Regeln. Damit Maßnahmen wirksam sind, müssen die Entscheidungsträger transparente und überzeugende Erklärungen für ihre Entscheidungen abgeben. Dies stellt hohe Anforderungen an Fähigkeiten zur Kommunikation statistischer Informationen seitens der Gesundheitsbehörden und Medien sowie an die Fähigkeit der Bürger/innen, diese Botschaften zu verstehen. Um die Dynamik der Pandemie in Grundzügen nachvollziehen und die Wirksamkeit bestimmter Maßnahmen (Social Distancing, Lockdown, AHA-Regeln u. v. a. m.) verstehen zu können, bedarf es grundlegender Kenntnisse im statistischen Denken.

Die Medien sind voll von Begriffen wie exponentielles Wachstum, Sieben-Tage-Inzidenz, Übersterblichkeit und Reproduktionszahl  $R_0$ . Wer z.B. die Merkmale des exponentiellen Wachstums kennt, kann auch Argumente widerlegen, dass COVID-19-Todesfälle in den frühen Phasen einer Epidemie kein Grund zur Sorge sind, wenn sie z.B. im Vergleich zu Todesfällen im Straßenverkehr gering sind. Zu einer Bewertung statistischer Informationen, wie sie in den Medien vermittelt werden, gehört auch die Fähigkeit, mit kritischem Verstand die Plausibilität von Informationen zu bewerten, Argumente zu analysieren und zu beurteilen, inwieweit ein Argument mit Belegen gestützt ist sowie die Bereitschaft, über mehrere unterschiedliche Erklärungen nachzudenken.

Informationen und Daten zu COVID-19 sind im Internet reichlich allgemein zugänglich, aber oft auch unvollständig, unklar und manchmal auch falsch. Für die Allgemeinheit gibt es zwei Hauptprobleme bei der Suche nach Informationen über COVID-19: eines ist mit Des- und Fehlinformationen verbunden; das andere ist eine Informationsüberflutung (oder die „Infodemie“). Bürger/innen benötigen Fähigkeiten, um vertrauenswürdige Quellen zu identifizieren und Websites zu finden, die prägnante Zusammenfassungen und klare Ratschläge enthalten.

### Statistical Literacy und Demokratie

Demokratie lebt von Argumenten, die auf nachweisbaren Fakten basieren. Lebendige Demokratien brauchen gut informierte Bürger/innen, die relevante gesellschaftliche Fragen verstehen, diskutieren und sich in öffentliche Entscheidungsprozesse einbringen können. Das gilt für das grundlegende

Verstehen der Verbreitung von Epidemien und möglicher Maßnahmen zu deren Einschränkung ebenso wie für viele andere zum Teil intensiv debattierte gesellschaftliche Themen (z.B. Klimawandel, Migration, Armut oder Kriminalität). Kenntnisse und Fähigkeiten, mit Daten angemessen zu argumentieren, sind eine wichtige Voraussetzung für das Funktionieren der Demokratie in unseren modernen Gesellschaften.

Ein grundlegendes Problem bei der Einschätzung von Wahrscheinlichkeiten liegt im intuitiven Vermischen von subjektiven Risiken („wie wahrscheinlich ist es, dass ich mich anstecke“) und allgemeinen Risiken („wie wahrscheinlich ist es, dass sich eine beliebige Person infiziert“). Diese aus dem Geburtstagsparadoxon wohlbekannte Diskrepanz der Risikoeinschätzung führt in Pandemiezeiten zu einer unterschiedlichen Bewertung und daraus resultierender Variabilität in der Akzeptanz beschlossener Maßnahmen.

In einer zunehmend komplexeren Welt ist das Engagement von Bürger/innen eine grundlegende Ressource bei öffentlichen Entscheidungen auf internationalem, nationalem und lokalem Niveau. Das Verständnis solcher Themen ist für das bürgerschaftliche Engagement in modernen Gesellschaften von großer Bedeutung, basiert aber oft auf komplexen multivariaten Daten, deren Interpretation und Erschließung Kenntnisse voraussetzt, die in der gewöhnlichen Grundbildung in Mathematik und Statistik nicht vermittelt werden – weder an Schulen noch an Hochschulen –, geschweige denn in schulischen Fächern wie Politik oder Gemeinschaftskunde. Statistische Daten zu Public Health und anderen gesellschaftlich brennenden Themen haben oft spezifische Merkmale, und ihr Verstehen erfordert Kompetenzen, auf die unsere Bildungseinrichtungen und Curricula nur ungenügend vorbereiten [51- 53].

Während das menschliche Denken zur Mustervereinfachung tendiert und auch die politische Kommunikation eine einfache Ursachen-Wirkungs-Relation bevorzugt, sind die realen Phänomene oft *multivariat*. So gilt es bei der Untersuchung von COVID-19 und Vorhersagen ihrer Ausbreitung deren Symptomatik, die Inzidenz und geographische Verbreitung von Krankheiten, Verhaltensmuster der Bevölkerung, staatliche Maßnahmen und Auswirkungen auf die Wirtschaft, auf Schulen, Menschen in Pflegeheimen und das soziale Leben insgesamt zu berücksichtigen und in die Datenanalysen zu integrieren. In den Daten beobachtete Zusammenhänge können oft durch Drittvariablen (Confounder) verursacht sein. Zudem stammen viele Daten aus Beobachtungsstudien, was eine belastbare Kausalattribution i.d.R. problematisch macht.

Darüber hinaus werden Entscheidungen über *Messungen* und *Operationalisierungen* getroffen, deren Hintergründe zu hinterfragen und zu diskutieren sind, ohne jedoch deren grundsätzliche Eignung in Frage zu stellen: Wie können Fälle oder Todesfälle genau genug gemessen werden, um Entscheidungen zu treffen? Was misst die 7-Tage-Inzidenz? Welche Bedeutung hat die Reproduktionszahl  $R_0$ ?

Daten werden häufig in *aggregierter Form* präsentiert, und die zugrunde liegenden räumlichen Strukturen ergeben sich aus verwaltungsrechtlichen Rahmenbedingungen, die in keinem direkten Zusammenhang mit dem Infektionsgeschehen stehen sollten. Für welche Fragestellungen und unter welchen Umständen ist es angebracht, Inzidenzen aus verschiedenen Regionen hinzuzufügen? Wie

valide sind Vergleiche der Inzidenzwerte unterschiedlicher Gruppen wie Schulkindern und Menschen in Pflegeheimen, vulnerable Personenkreise und Menschen ohne relevante Vorerkrankungen?

Eine neu auftretende Krankheit, die sich in rasantem Tempo global ausbreitet, führt zu einer Heterogenität der Daten und Erhebungsmethoden und somit zu *zeitabhängigen und geographisch gestreuten Daten* aus unterschiedlichen Datenquellen: Daten zum Verlauf von Epidemien werden über längere Zeiträume und an unterschiedlichen Orten erhoben. Dabei werden nicht überall und zu allen Zeiten dieselben Methoden eingesetzt. Die Modellierung von Krankheiten erfordert Schätzungen der Anzahl anfälliger Personen in der Bevölkerung und der Heilungschancen für diejenigen, die sich mit der Krankheit infiziert haben. Beide Parameter ändern sich im Laufe der Zeit, wenn mehr Menschen immun werden und sich die Behandlungsmethoden verbessern.

In Testsituationen ist nicht immer der Unterschied zwischen der Sensitivität eines diagnostischen Tests und dem positiv-prädiktiven Wert präsent [54-56, 71]. Insbesondere wird dabei die Prävalenz (oder Basisrate) oft vernachlässigt. Gehen wir beim Antigen-Schnelltest von einer Sensitivität von 80% und einer Spezifität von 98% aus und nehmen eine Prävalenz von 0,05% an (was gerade einer Anzahl von 50 Infektiösen pro 100.000 Menschen entspricht), so sind nur 1,9% aller Menschen mit positivem Testresultat tatsächlich infektiös. Entscheidend für die Bewertung der Ergebnisse solcher Tests ist jedoch das Verständnis der Testsituation. Denn je nach Anlass und Setting ändert sich die Aussagekraft der Tests. Die obige Rechnung mag bei einer Reihenuntersuchung angemessen sein. Bei einer anderen Basisrate von 10% (es werden nur Personen mit Erkältungssymptomen getestet) errechnet sich eine Wahrscheinlichkeit von 81,6% für eine positiv getestete Person, auch tatsächlich das Virus zu haben.

### **Visuelle Kommunikation**

Neue Entwicklungen in Computergraphik und Animation prägen die Art und Weise, wie Daten in Medien als Abbildungen oder als animierte Simulationen dargestellt werden. Visuelle Darstellungen nehmen eine zentrale Stellung in der öffentlichen Kommunikation ein und zielen darauf ab, die entsprechenden Dynamiken und Inhalte schnell verständlich zu repräsentieren. Dabei werden in der Regel entweder zeitabhängige Kenngrößen oder Daten mit Raumbezug visualisiert. Für die räumlich verteilten Daten werden bevorzugt Choroplethenkarten eingesetzt, bei denen idealerweise die für Gesundheitsfragen zuständigen Gebietskörperschaften gemäß der Verteilungsdichte der Infektionszahlen oder daraus abgeleiteter Größen eingefärbt werden (siehe Abbildung 5). Fast ausschließlich werden in der medialen Praxis Verordnungsgrenzwerte als Grundlage für die Farbskalierung verwendet, so dass bei flächendeckenden Über- und Unterschreitungen der gültigen Warnwerte keine Farbunterscheidung zwischen den Gebieten realisiert wird. Dadurch werden die realen räumlichen Verteilungsunterschiede ignoriert und der Informationsgehalt der Choroplethenkarte wird minimiert. Für eine sinnvolle visuelle Darstellung wären hier Farbskalen, die grenzwertbasiert unterschiedliche Grundfarben verwenden und diese dann intensitätsabhängig variieren, ein probates Mittel zur besseren Informationskommunikation.

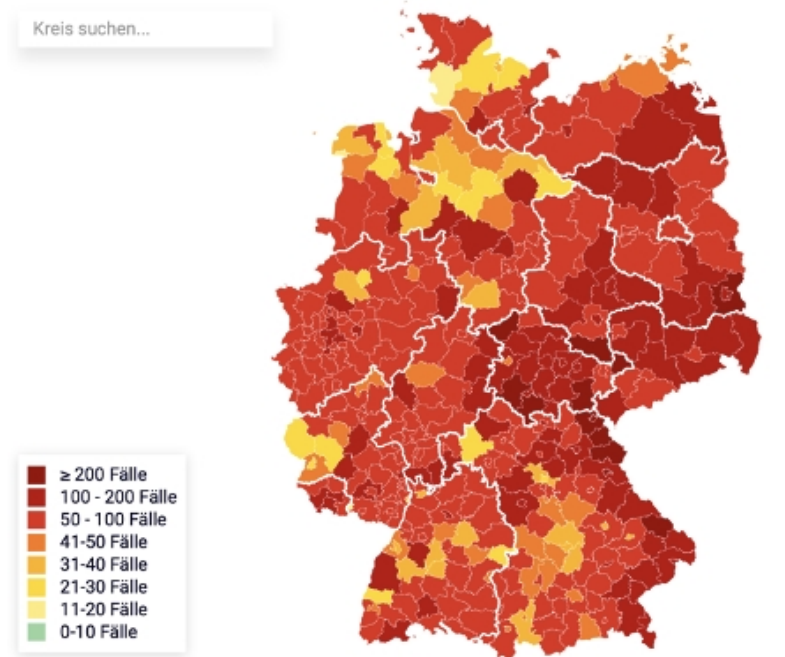


Abbildung 5: Choroplethenkarte der Inzidenzzahlen für Deutschland nach Landkreisen. Quelle: RKI <https://app.23degrees.io/export/oCRP768wQ3mCswE7-choro-corona-faelle-pro-100-000/image>

Bei den zeitabhängigen Kenngrößen kommen verschiedene Varianten von Zeitreihendiagrammen zum Einsatz, überwiegend Linien- und Säulendiagramme. Klassische Fehler der grafischen Darstellung wie etwa die Überbetonung der zeitlichen Variabilität durch Reduktion der Werteachse auf einen kleinen Ausschnitt sind mittlerweile in den Medien weitgehend verschwunden. Der Wechsel zwischen Linien- und Säulendiagramm aus rein designtechnischen Überlegungen, um entsprechende grafische Diversität zu produzieren, erscheint dennoch fragwürdig. Mit Vorsicht ist bei den Zeitreihendiagrammen die Verwendung logarithmischer Skalen zu bewerten. Zum einen verleiten sie beim oberflächlichen Lesen zur Unterschätzung dynamischer Wachstumsprozesse, zum anderen erhöhen sie die Anforderungen an die mathematische und statistische *Literacy* der Leserschaft ohne entsprechende Vorteile der visuellen Repräsentation. Abbildung 6 zeigt den zeitlichen Verlauf der 7-Tage Inzidenz pro 100.000 Einwohner/innen zwischen dem 24. Januar und dem 4. Februar 2021 für einige ausgewählte Länder. Während auf der logarithmischen Skala die Unterschiede relativ gering erscheinen, zeigt die lineare Skala erhebliche Differenzen.

Einige Medien nutzen zur Illustration innovative Visualisierungen. In der Financial Times [57] findet sich z.B. unter der Überschrift „COVID-19’s soaring death toll dwarfs figures from the first wave“ eine Grafik, die einen Überblick über den Verlauf der auf Corona zurückzuführenden Todesfälle getrennt nach Weltregionen gibt.

Besonders illustrativ wurden im Zuge der Pandemie in den Medien Simulationen eingesetzt. Ein inspirierendes Beispiel zur Illustration zur Verbreitung der Epidemie erschien schon am 14. März 2020 in der Washington Post [58] mit dem Titel „Why outbreaks like coronavirus spread exponentially, and how to flatten the curve“. Die Washington Post stellte diese Simulation kostenlos und in allen wichtigen Sprachen zur Verfügung, was dazu führte, dass sie weltweit verbreitet wurde, auch wiederholt im deutschen Fernsehen [59]. Die New York Times [60] hat eine dynamische Grafik mit dem Titel „How the Virus Won“ veröffentlicht, die die Ausbreitung von COVID-19-Fällen von Februar bis Juni 2020 in den USA abbildet. Es zeigt, wie eine Analyse der Assoziationen zwischen verschiedenen COVID-19-Stämmen und Reismustern helfen kann, die Ausbreitung der Krankheit zu verstehen.

Illustrativ ist auch eine Simulation aus ZEIT Online [61], die – basierend auf Modellen einer Forschergruppe des Max Planck Instituts für Chemie – Wahrscheinlichkeiten in verschiedenen Szenarien abschätzen lässt, mit der eine infizierte Person andere Menschen in geschlossenen Räumen ansteckt. Während die Visualisierungen die simulierten Ansteckungsprozesse eingängig darstellen, wird die Abhängigkeit der Simulationen von Parameterannahmen und -einstellungen meist nicht thematisiert. Simulationen sollten auch stets transparent machen, auf welchen Modellannahmen und welcher Datenbasis die Simulationen erstellt wurden.

Zur Darstellung von Testergebnissen empfehlen sich Faktenboxen, die mit Icon Arrays kombiniert werden. Beide Darstellungen beruhen auf natürlichen Häufigkeiten [62, 70] und stellen Fallzahlen möglichst einfach und konkret dar. Ein Beispiel aus der Deutschen Ärztezeitung (<https://www.aerztezeitung.de/Wirtschaft/Corona-Schnelltests-die-Praevalenz-machts-414743.html>) zeigt beispielsweise richtig positive, falsch positive, richtig negative und falsch negative Testergebnisse. Viele wissenschaftliche Studien zeigen, dass Icon Arrays Menschen helfen, Zahlen und Risiken leichter zu verstehen (z.B. [56]). Das Harding Center für Risikokompetenz zeigt viele weitere Beispiele einer transparenten Kommunikation von Risiken, auch im Bereich COVID-19 [63].

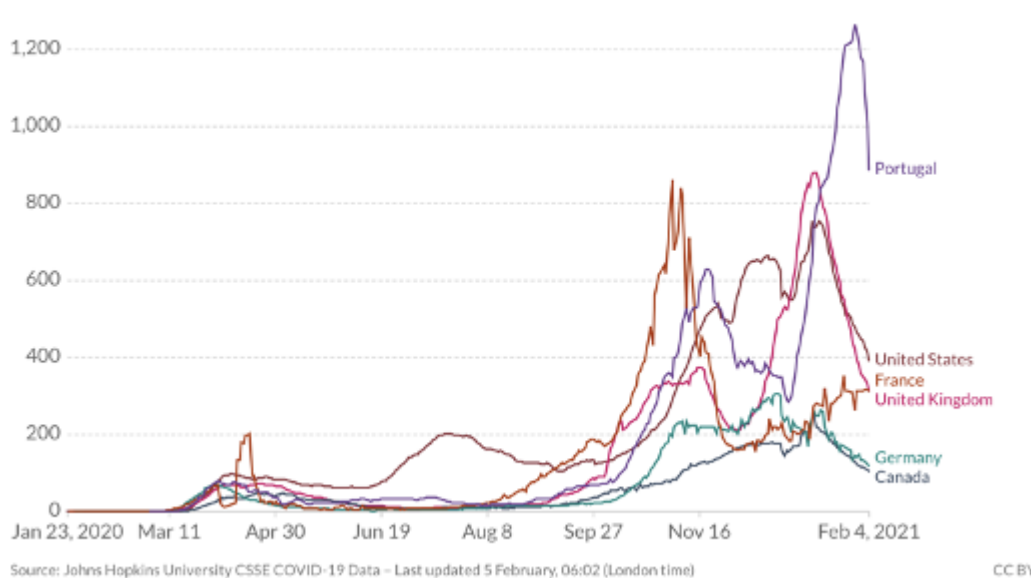


Abbildung 6: Der zeitliche Verlauf der 7-Tage-Inzidenz für verschiedene Länder. Auf der logarithmischen Skala (obere Grafik) wirken Unterschiede geringer. Die lineare Skala (untere Grafik) zeigt erhebliche Differenzen. Quelle: Our World in Data, <https://ourworldindata.org/covid-cases?country=IND~USA~GBR~CAN~DEU~FRA>.

### Statistical Literacy als Aufgabe und Herausforderung für Bildungseinrichtungen

Die Fähigkeit, Daten kritisch zu beurteilen und Sinn aus Daten zu erschließen, dient der Mündigkeit in einer modernen digitalisierten Welt und sollte daher in allen Bildungsbereichen verankert und als Teil

der Allgemeinbildung etabliert werden (siehe Data Literacy Charta [64]). Das europäische Kooperationsprojekt ProCivicStat hat neben der Bereitstellung von konkreten Lehr- und Lernmaterialien sechs Empfehlungen zur Förderung von Bildungsmaßnahmen und systemischen Veränderungen formuliert. Ziel ist es, durch die Förderung von Statistical Literacy zivilgesellschaftliches Engagement von Bürgern qualitativ zu unterstützen [52,53].

Auf der universitären Ebene gibt es in jüngster Zeit ähnliche Initiativen, in einem allgemeinbildenden Sinne Komponenten der Data Literacy einzuführen (vgl. die Initiativen des Stifterverbandes <https://www.stifterverband.org/data-literacy-education> sowie [65, 66]).

Die Herausforderung, Menschen zu befähigen, Sinn aus diesen Daten zu erschließen, richtet sich gleichermaßen an Bildungseinrichtungen (z.B. Schulen, Hochschulen, Volkshochschulen) wie auch an Medien und Produzenten von öffentlichen Statistiken. Der kompetente Umgang mit Daten, einschließlich der Fähigkeit, Daten auf kritische Art und Weise zu sammeln, zu managen, zu bewerten und anzuwenden, ist eine Schlüsselkompetenz des 21. Jahrhunderts. Dadurch wird die Urteilsfähigkeit, Selbstbestimmtheit und das Verantwortungsbewusstsein gestärkt und die gesellschaftliche und wirtschaftliche Teilhabe von uns allen in einer durch Digitalisierung geprägten Welt gefördert (siehe Data Literacy Charta [64]).

#### **Empfehlungen:**

14.) Bestehende Unsicherheiten aufgrund der neuartigen Situation und der noch unzureichenden Datengrundlage sollten konsequent kommuniziert werden.

15.) Zahlen sollten zu einer vereinheitlichenden Bezugsgröße ins Verhältnis gesetzt werden (z.B. X Todesfälle pro 100.000 Einwohner/innen). Zur besseren Einordnung sollten sie mit anderen Gesundheitsrisiken in Vergleich gesetzt und durch Grafiken veranschaulicht werden. Insbesondere bei Prozentzahlen ist wichtig, dass die Bezugsgröße genannt wird und dass diese relevant ist für die Fragestellung.

16.) Eine transparente Kommunikation von Entscheidungsmodellen sowie Risiken und Unsicherheiten kann eine aktive Zustimmung zu den empfohlenen Präventionsmaßnahmen und verantwortungsvolles gesundheitsbezogenes Verhalten fördern. Die Datengrundlage einschließlich Herkunft der Daten, Umfang sowie Erhebungsmethoden sollten transparent sein.

17.) Simulationen sind ein sehr probates Mittel, um die Dynamik von Infektionsgeschehen zu illustrieren. Bei ihrem Einsatz sollten zugrunde liegende Modellannahmen transparent gemacht werden.

18.) Zur Förderung von Statistical Literacy auf allen Ebenen ist die Zusammenarbeit von Statistiker/innen mit allen Interessensgruppen nötig, die sich mit statistischer Bildung im Hochschul- oder Sekundarschulbereich befassen oder die statistische Kompetenz der Bürger/innen fördern. Zu diesen Akteuren gehören Pädagog/innen auf allen Ebenen, Schul- und Universitätsadministrator/innen, politische Entscheidungsträger, offizielle Statistikanbieter, Forscher/innen, Medienfachleute, Lehramtsausbilder/innen, Softwareentwickler/innen und viele andere.

## Internationale Initiativen

International gibt es verschiedene ähnliche Initiativen. Wir nennen hier beispielhaft die kürzlich veröffentlichten zehn Punkte der britischen Royal Statistical Society [72]. Ins Deutsche übertragen lauten diese:

1. **In öffentliche Gesundheitsdaten investieren** - diese sollten als kritische nationale Infrastruktur betrachtet werden, und es sollte eine vollständige Überprüfung der Gesundheitsdaten durchgeführt werden.
2. **Evidenz veröffentlichen** - alle von der Regierung und ihren Beratern berücksichtigten Beweise müssen zeitnah und zugänglich veröffentlicht werden.
3. **Klar und offen über Daten berichten** - die Regierung sollte in ein zentrales Portal investieren, auf dem die verschiedenen Quellen offizieller Daten, Analyseprotokolle und aktuelle Ergebnisse zu finden sind.
4. **Den Missbrauch von Statistiken hinterfragen** - das Office for Statistics Regulation sollte finanziell aufgestockt werden, damit es die Regierung besser in die Pflicht nehmen kann.
5. **Die Medien müssen mehr Verantwortung übernehmen** - die Regierung sollte Medieninstitutionen unterstützen, die in spezialisierte wissenschaftliche und medizinische Berichterstattung investieren.
6. **Ausbau der statistischen Fähigkeiten von Entscheidungsträgern** - Politiker und hochrangige Beamte sollten sich in Statistik schulen lassen.
7. **Aufbau eines effektiven Überwachungssystems für Infektionskrankheiten**, um die Ausbreitung von Krankheiten zu überwachen - die Regierung sollte sicherstellen, dass ein Echtzeit-Überwachungssystem für zukünftige Pandemien bereit steht.
8. **Mehr Kontrolle und Offenheit für neue diagnostische Tests** - ähnliche Schritte wie bei der Bewertung von Impfstoffen und Medikamenten sollten auch bei diagnostischen Tests befolgt werden.
9. **Gesundheitsdaten sind ohne Sozialdaten unvollständig** - die Verbesserung der Sozialdaten sollte ein zentraler Bestandteil jeder Überprüfung der britischen Gesundheitsdaten sein.
10. **Evaluierung sollte in den Mittelpunkt der Politik gestellt werden** - effiziente Evaluierungen oder Experimente sollten von Anfang an in jede Intervention einbezogen werden.



## Zusammenfassung

Eine Pandemie stellt in besonderem Maße Herausforderungen an den Prozess der Entscheidungsfindung. Damit Entscheidungen evidenzbasiert erfolgen können, ist eine zeitnahe Verfügbarkeit von Evidenz und Daten notwendig. COVID-19 hat gezeigt, dass zu vielen Themen des Krisenmanagements keine ausreichende Vorbereitung vorlag. Dies traf insbesondere auch auf Daten, Statistiken und Modelle zu. Die DAGStat empfiehlt deshalb die Erarbeitung eines statistischen Fragenkatalogs, der für künftige Pandemien als Grundlage genutzt werden kann.

Die Statistik als breit angelegte wissenschaftsübergreifende Fachrichtung spielt hier eine besondere und wichtige Rolle. In dieser Stellungnahme haben wir Entscheidungsfindung am Beispiel der Corona-Pandemie diskutiert. Viele der abgeleiteten Maßnahmen und Empfehlungen lassen sich aber auch auf andere Themen verallgemeinern. Insbesondere lassen sich aus statistischer Sicht folgende Empfehlungen für Entscheidungsfindungsprozesse ableiten:

1. **Verfügbarkeit von Daten:** Eine nationale Strategie und systematische Sammlung von Daten ist wichtig, um statistische und entscheidungsanalytische Modelle erstellen zu können.
2. **Transparenz:** Um die Akzeptanz der Entscheidungen und damit einhergehender Maßnahmen zu erhöhen, müssen alle Schritte des Entscheidungsfindungsprozesses offen gelegt werden. Das beginnt mit den Daten und setzt sich in der Wahl der Modelle und der betrachteten Trade-Offs fort.
3. **Kommunikation:** Eine umfassende Kommunikation mit der Bevölkerung ist unabdingbar. Dies beinhaltet insbesondere die Kommunikation von Risiken und Unsicherheiten, der Zielsetzung eines Modells sowie der Wahl der Modellparameter und Annahmen. In diesem Zusammenhang kommt auch den Medien eine entscheidende Rolle zu. Die DAGStat als Dachverband verschiedenster Fachgesellschaften bietet sich hier als Netzwerk an.
4. **Interdisziplinäre Zusammenarbeit:** Eine Pandemie stellt besondere Herausforderungen an die gesamte Gesellschaft. Um diese möglichst effizient anzugehen, ist eine fächerübergreifende Zusammenarbeit unabdingbar. Darüber hinaus ist es empfehlenswert, als Fachgruppe statt als Individuum, breit aufgestellt und mediensensitiv aufzutreten.
5. **Statistical Literacy/ Data Literacy:** Die Fähigkeit, Daten kritisch zu beurteilen und Sinn aus Daten zu erschließen, dient der Mündigkeit in einer modernen digitalisierten Welt und sollte daher als Teil der Allgemeinbildung etabliert werden. Zur Förderung von Statistical Literacy auf allen Ebenen ist die Zusammenarbeit von Statistiker/innen mit allen Interessensgruppen, die sich mit statistischer Bildung befassen, nötig.

Die DAGStat stellt sich zur Verfügung, Anfragen zu statistischen Fragestellungen, z.B. Interpretation und Bewertung von Modellen sowie Analyse von Daten zu koordinieren und fachliche Ansprechpartner/innen zu benennen.

## Referenzen

- [1] Blastland, M, Freeman, A. L. J., van der Linden, S, Marteau, T. M. & Spiegelhalter, D: Five rules for evidence communication. *Nature*, 18 November 2020, <https://www.nature.com/articles/d41586-020-03189-1>.
- [2] Tukey, J. W. (1986). Sunset salvo. *The American Statistician*, 40(1), 72-76.
- [3] Bundesweite Antikörper-Studie „Leben in Deutschland – Corona-Monitoring“ des DIW/RKI, [https://www.rki.de/DE/Content/Service/Presse/Pressemitteilungen/2020/08\\_2020.html](https://www.rki.de/DE/Content/Service/Presse/Pressemitteilungen/2020/08_2020.html), accessed 12.02.21
- [4] <https://www.konsortswd.de/aktuelles/pressemitteilungen/26062020/>, accessed 12.02.21
- [5] Wiengarten L., Zwick M., (2017) Neue digitale Daten in der amtlichen Statistik, Statistisches Bundesamt, Wirtschaft und Statistik, Heft 5 <https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2017/05/neue-digitale-daten-052017.html>
- [6] Münnich, R.T., Zwick, M. (2016) (Hrsg) Big Data und was nun? Neue Datenbestände und ihre Auswirkungen. Sonderheft AStA Wirtsch Sozialstat Arch 10, <https://doi.org/10.1007/s11943-016-0196-y>
- [7] <https://www.telefonica.de/analytics/anonymisierte-daten/studie-so-bewegt-sich-deutschland.html>
- [8] <https://www.destatis.de/DE/Service/EXDAT/Datensaetze/mobilitaetsindikatoren-mobilfunkdaten.html>
- [9] Kühnemann, H (2021). Anwendungen des Web Scraping in der amtlichen Statistik, AStA Wirtsch Sozialstat Arch, im Erscheinen.
- [10] Verhaltenskodex für europäische Statistiken, 2017 [https://www.destatis.de/DE/Methoden/Qualitaet/verhaltenskodex.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Qualitaet/verhaltenskodex.pdf?__blob=publicationFile)
- [10a] Verordnung (EG) Nr. 223/2009 über europäische Statistiken, Fassung 2015 [https://www.destatis.de/DE/Methoden/Rechtsgrundlagen/Statistikbereiche/Inhalte/027\\_VO\\_europStatistiken.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Methoden/Rechtsgrundlagen/Statistikbereiche/Inhalte/027_VO_europStatistiken.pdf?__blob=publicationFile)
- [10b] <https://ec.europa.eu/eurostat/web/products-catalogues/-/KS-02-18-142>
- [11] Flaxman, S., Mishra, S., Gandy, A. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 584, 257–261 (2020). <https://doi.org/10.1038/s41586-020-2405-7>
- [12] Rippinger, C., Bicher, M., Urach, C. et al. Evaluation of undetected cases during the COVID-19 epidemic in Austria. *BMC Infect Dis* 21, 70 (2021). <https://doi.org/10.1186/s12879-020-05737-6>.
- [13] Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Med Decis Making* 2012;32(5):733-43.
- [14] Schöffski, Oliver, Graf von der Schulenburg, Johann-Matthias (Hrsg.). Gesundheitsökonomische Evaluationen. Springer-Verlag, 2012.

- [15] Holmdahl, Inga and Buckee, Caroline. Wrong but Useful — What COVID-19 Epidemiologic Models Can and Cannot Tell Us. *New England Journal of Medicine* 383, 303-305, 2020. DOI: 10.1056/NEJMp2016822.
- [16] Steyerberg, E. W., Moons, K. G., van der Windt, D. A., Hayden, J. A., Perel, P., Schroter, S., ... & PROGRESS Group. (2013). Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med*, 10(2), e1001381. doi: 10.1371/journal.pmed.1001381.
- [17] Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., ... & van Smeden, M. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *bmj*, 369. doi: <https://doi.org/10.1136/bmj.m1328>
- [18] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddard GL. Chapter 2: Basic types of economic evaluation. *Methods for the economic evaluation of health care programmes*. 3rd ed. New York: Oxford University Press, 2005;6-33.
- [19] Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Med Decis Making* 2012;32(5):722-32.
- [20] Neumann PJ, Ganiats TG, Russell LB, Sanders GD, Siegel JE. *Cost-Effectiveness in Health and Medicine*. United States of Amerika: Oxford, 2016.
- [21] Ariely, D., & Jones, S. (2008). *Predictably irrational*. New York, NY: Harper Audio.
- [22] Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, Weinstein M. Managing uncertainty. *Decision Making in Health and Medicine: Integrating Evidence and Values*. New York, USA: Cambridge University Press, 2001;33-60.
- [23] Rochau U, Jahn B, Qerimi V, Burger EA, Kurzthaler C, Kluibenschaedl M, Willenbacher E, Gastl G, Willenbacher W, Siebert U. Decision-analytic modeling studies: An overview for clinicians using multiple myeloma as an example. *Crit Rev Oncol Hematol* 2015;94(2):164-78.
- [24] Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making* 1983;3(4):419-458.
- [25] Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;13(4):322-38
- [26] Spielauer M. Dynamic microsimulation of health care demand, health care finance and the economic impact of health behaviours: survey and review. *Int J Microsimulation* 2007.
- [27] Groot Koerkamp B, Weinstein MC, Stijnen T, Heijnenbroek-Kal MH, Hunink MG. Uncertainty and patient heterogeneity in medical decision models. *Med Decis Making* 2010;30(2):194-205.
- [28] Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. *Pharmacoeconomics* 2006;24(11):1043-53.
- [29] Pidd M. *Computer simulation in management science*. England: Wiley, 2004.
- [30] Karnon J, Stahl J, Brennan A, Caro JJ, Javier M, Moeller J, on behalf of the ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling using Discrete Event Simulation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Freely available from: <http://>

[mdm.sagepub.com/content/32/5/701.full](http://mdm.sagepub.com/content/32/5/701.full) and <http://www.ispor.org/workpaper/Modeling-Using-Discrete-Event-Simulation.asp>. Medical Decision Making/ Value in Health 2012;32/ 15:701-11/ 821-27.

[31] Jun JB, Jacobson SH, Swisher JR. Application of discrete-event simulation in health care clinics: A survey. Journal of the Operational Research Society 1999;50(2):109-123.

[32] Zhang X. Application of discrete event simulation in health care: a systematic review. BMC health services research 2018;18(1):687-687

[33] Banks J, Carson JS, Nelson BL, Nicol DM. Discrete-Event System Simulation. USA: Pearson Prentice Hall, 2005.

[34] Jahn B, Theurl E, Siebert U, Pfeiffer KP. Tutorial in medical decision modeling incorporating waiting lines and queues using discrete event simulation. Value Health 2010;13(4):501-6.

[35] Jahn B, Pfeiffer KP, Theurl E, Tarride JE, Goeree R. Capacity constraints and cost-effectiveness: a discrete event simulation for drug-eluting stents. Med Decis Making 2010;30(1):16-28.

[36] Marshall DA, Burgos-Liz L, MJ IJ, Crown W, Padula WV, Wong PK, Pasupathy KS, Higashi MK, Osgood ND. Selecting a dynamic simulation modeling method for health care delivery research-part 2: report of the ISPOR Dynamic Simulation Modeling Emerging Good Practices Task Force. Value Health 2015;18(2):147-60.

[37] Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences 2002;99(suppl 3):7280.

[38] Macal CM, North MJ. Agent-based modeling and simulation: ABMS examples. Proceedings of the 40th Conference on Winter Simulation. Miami, Florida: Winter Simulation Conference, 2008;101-112.

[39] Miksch F, Jahn B, Espinosa KJ, Chhatwal J, Siebert U, Popper N. Why should we apply ABM for decision analysis for infectious diseases?—An example for dengue interventions. PLOS ONE 2019;14(8):e0221564.

[40] Chhatwal J, He T. Economic Evaluations with Agent-Based Modelling: An Introduction. PharmacoEconomics 2015;33(5):423-433.

[41] Bruch E, Atwell J. AGENT-BASED MODELS IN EMPIRICAL SOCIAL RESEARCH. Sociological methods & research 2015;44(2):186-221.

[42] Hunter E, Mac Namee B, Kelleher JD. A Taxonomy for Agent-Based Models in Human Infectious Disease Epidemiology. Journal of Artificial Societies and Social Simulation 2017;20(3):2.

[43] Macal CM, North MJ. Tutorial on agent-based modelling and simulation. Journal of Simulation 2010;4(3):151-162.

[44] Marshall DA, Burgos-Liz L, MJ IJ, Osgood ND, Padula WV, Higashi MK, Wong PK, Pasupathy KS, Crown W. Applying dynamic simulation modeling methods in health care delivery research-the

SIMULATE checklist: report of the ISPOR simulation modeling emerging good practices task force. *Value Health* 2015;18(1):5-16.

[45] Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, Brisson M. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-5. *Med Decis Making* 2012;32(5):712-21.

[46] Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes : a focused review of modelling approaches. *Pharmacoeconomics* 2008;26(3):191-215.

[47] Welte R, Postma M, Leidl R, Kretzschmar M. Costs and effects of chlamydial screening: dynamic versus static modeling. *Sex Transm Dis* 2005;32(8):474-83.

[48] Jit M, Brisson M. Modelling the epidemiology of infectious diseases for decision analysis: a primer. *Pharmacoeconomics* 2011;29(5):371-86.

[49] SMDM. COVID-19 Decision Model Repository. Available from <https://app.smartsheet.com/b/publish?EQBCT=1a3bc6acad99475f99acfd55a04a1564> [Accessed 08 May 2020]

[50] Kompetenznetz Public Health COVID-19. Available from <https://www.public-health-covid19.de/> [Accessed 08 May 2020]

[51] Engel, J.; Biehler, R.; Frischemeier, D.; Podworny, S.; Schiller, A.; Martignon, L. (2019). *Zivilstatistik:Konzept einer neuen Perspektive auf Data Literacy und Statistical Literacy*. *AStA Wirtsch Sozialstat Arch* 13 (13-244). <https://link.springer.com/article/10.1007/s11943-019-00260-w>.

[52] ProCivicStat Partners (2018). Engaging civic statistics: a call for action and recommendations. A product of the ProCivicStat project. <http://IASE-web.org/ISLP/PCS>

[53] Ridgway, J. (Ed.) (Forthcoming 2021). *Statistics for Empowerment and Social Engagement – teaching Civic Statistics to develop informed citizens*. Springer.

[54] Eddy, D. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 249-267). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511809477.019

[55] Gigerenzer, G., Gaissmaier, W. Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics, *Psychological Science in the Public Interest*, vol. 8, no. 2, pp. 53–96. <https://journals.sagepub.com/doi/10.1111/j.1539-6053.2008.00033.x>

[56] McDowell, M. E., Gigerenzer, G., Wegwarth, O., & Rebitschek, F. G. (2019). Effect of tabular and icon fact box formats on comprehension of benefits and harms of prostate cancer screening: A randomized trial. *Medical Decision Making*, 39, 41-56. <https://doi.org/10.1177/0272989X18818166>.

[57] <https://www.ft.com/content/a2901ce8-5eb7-4633-b89c-cbdf5b386938>.

[58] <https://www.washingtonpost.com/graphics/2020/world/corona-simulator/>

[59] <https://web.br.de/interaktiv/corona-simulation/>

- [60] <https://www.nytimes.com/interactive/2020/us/coronavirus-spread.html>
- [61] <https://www.zeit.de/wissen/gesundheit/2020-11/coronavirus-aerosole-ansteckungsgefahr-infektion-hotspot-innenraeume>
- [62] Gigerenzer, G. (2011). What are natural frequencies? *BMJ*, 343, d6386. <https://doi.org/10.1136/bmj.d6386>
- [63] <https://www.hardingcenter.de/de/mrna-schutzimpfung-gegen-covid-19-fuer-aeltere-menschen>
- [64] Schüller K., Koch, H. & Rampelt, FD. (2021). Data-Literacy-Charta. Berlin: Stifterverband, <https://www.stifterverband.org/charta-data-literacy>
- [65] Ridsdale C, Rothwell J, Smit M, Ali-Hassan H, Bliemel M, Irvine D, Wuetherick B (2015). Strategies and best practices for data literacy education: knowledge synthesis report. <http://hdl.handle.net/10222/64578>.
- [66] Heidrich J, Bauer P, Krupka D (2018). Strukturen und Kollaborationsformen zur Vermittlung von Data Literacy-Kompetenzen – Stand der Forschung. Arbeitspapier Nr. 32. [https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD\\_AP\\_Nr32\\_Data\\_Literacy\\_Kompetenzen\\_Literatur.pdf](https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr32_Data_Literacy_Kompetenzen_Literatur.pdf)
- [67] King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton university press.
- [68] Radermacher, W., Weisbrod, J., & Asef, D. (2004). Bedarf, Qualität, Belastung: Optimierung als Interessenausgleich. *Wirtschaft und Statistik*, 10, 1237-1244.
- [69] Gelinsky, K. (2021). Wer beatmet wird, und wer nicht. *Frankfurter Allgemeine Zeitung*, 13.01.2021.
- [70] Krauss, S., Weber, P., Binder, K., & Bruckmaier, G. (2020). Natürliche Häufigkeiten als numerische Darstellungsart von Anteilen und Unsicherheit – Forschungsdesiderate und einige Antworten. *Journal für Mathematikdidaktik*, 41, 485-521.
- [71] Binder, K., Krauss, S., & Wiesner, P. (2020). A new visualization for probabilistic situations containing two binary events: The frequency net. *Frontiers in Psychology*, 11(750).
- [72] Royal Statistical Society (2020). <https://rss.org.uk/policy-campaigns/policy/covid-19-task-force/statistics,-data-and-covid/>
- [73] <https://infas360.com/>

Autoren:

Behnke, Joachim	Zeppelin Universität Friedrichshafen	Sektion Methoden der DVPW
Engel, Joachim	Pädagogische Hochschule Ludwigsburg	Verein zur Förderung des schulischen Stochastikunterrichts
Friede, Tim	Universitätsmedizin Göttingen	Vorsitzender der DAGStat
Friedrich, Sarah	Universitätsmedizin Göttingen	Schriftführerin der DAGStat
Garczarek, Ursula	Cytel Inc	GfKI
Jahn, Beate	UMIT Tirol	Society for Medical Decision Making (SMDM)
Kaderali, Lars	Universitätsmedizin Greifswald	GMDS
Kestler, Hans	Universität Ulm	GfKI
Münnich, Ralf	Universität Trier	DStatG
Pauly, Markus	Technische Universität Dortmund	DMV Fachgruppe Stochastik
Siebert, Uwe	UMIT Tirol	DGEpi
Wilhelm, Adalbert	Jacobs University Bremen	GfKI
Wolkenhauer, Olaf	Universität Rostock	GMDS
Zwick, Markus	Destatis	Destatis

Redaktionelle Bearbeitung: Reiner Latsch