

Kennzahlen rund um die Corona-Pandemie – eine Betrachtung aus Sicht der Statistik

Dr. Heiko Bergmann M.Sc., Dorothee Engelhardt, M.A., Dipl.Soz.Univ. Nina Storfinger¹

Selten erfahren statistische Kennzahlen und Verfahren eine solch hohe Aufmerksamkeit wie im Zuge der gegenwärtigen Corona-Pandemie. Ob Verdopplungszahl, Reproduktionszahl, falsch-positive Testergebnisse oder Dunkelziffer – Wissenschaft und Gesundheitsbehörden müssen unter sehr großem Zeitdruck neue Statistiken liefern und die Öffentlichkeit diskutiert rege über die zum Teil im Stundentakt aktualisierten Kennzahlen.

Doch auch wenn es begrüßenswert ist, dass politische Entscheidungen stets auf aktuellen Zahlen und neuen wissenschaftlichen Erkenntnissen beruhen, musste zum einen die Öffentlichkeit mit dieser Flut an Informationen zunächst lernen umzugehen. Zum anderen darf dabei die Aussagekraft beziehungsweise die korrekte Interpretation der verwendeten Statistiken nicht außer Betracht gelassen werden.

Der vorliegende Beitrag beleuchtet daher aus statistischer Sicht einige der in der Corona-Krise geläufigsten Kennzahlen näher: Die Reproduktionszahl, die Genauigkeit von Tests und die Dunkelziffer. Wie werden sie berechnet? Welche Aspekte gilt es bei deren Betrachtung zu berücksichtigen? Was sind mögliche Fallstricke bei ihrer Verwendung?

Einleitung

Selten waren seitens der Öffentlichkeit Interesse an und Forderung nach sofort verfügbaren Daten, Zahlen und Statistiken derart groß wie in der derzeitigen Corona-Krise. Pressekonferenzen des Robert Koch-Instituts (RKI) wurden live im TV übertragen, Tageszeitungen, Online-Medien sowie Fernsehsender berichteten und berichten zum Teil im Stundentakt über die aktuellsten Entwicklungen und Zahlen. Das COVID-19-Dashboard der Johns Hopkins University in Baltimore, das unter anderem die Anzahl der Infektionen, Gestorbenen und Genesenen für zahlreiche Staaten mehrfach täglich aktualisiert, hatte zum Teil mehr als eine Milliarde Aufrufe – jeden Tag. Dass dabei im Spannungsfeld zwischen Genauigkeit auf der einen und Geschwindigkeit auf der anderen Seite die Sorgfalt – insbesondere im Umgang mit Zahlen und Statistiken – mitunter auf der Strecke bleiben kann, dürfte dabei wenig überraschen.

Auch sehnen sich die Öffentlichkeit und die Politik gerne nach „der einen Zahl“, anhand derer wir wissen und die Politik weiß, wann welche Maßnahmen zur Eindämmung der Pandemie getroffen werden müssen und ab welchem Zeitpunkt Maßnahmen sukzessive gelockert werden können. Komplexe und vielschichtige Fragen lassen sich hingegen äußerst selten auf nur eine Zahl reduzieren.

Der Fokus dieses Beitrags ist daher vielmehr auf folgende Fragen gerichtet: Wie sind die verschiedenen Zahlen und Statistiken rund um die Corona-Pandemie statistisch einzuordnen? Wurden in der Berichterstattung gewisse Sachverhalte – aus statistischer Sicht – vereinfacht oder verkürzt? Wurden komplexe Themen auf nur eine Zahl reduziert? Wurden Zahlen als präzise ausgegeben, die aber unter Berücksichtigung eines statistischen Schätzfehlers mit gewisser Vorsicht behandelt werden

¹ Das Autorenteam bedankt sich bei Prof. Dr. Manfred Wildner, MPH, und seinen Kolleginnen und Kollegen des Bayerischen Landesamts für Gesundheit und Lebensmittelsicherheit für die hilfreichen Anregungen und die konstruktive Kritik, die zum Gelingen dieses Beitrags beigetragen haben.

sollten? Oder allgemein: Welche Aspekte gilt es bei der Betrachtung statistischer Größen zu berücksichtigen?

Aus statistischer Sicht näher betrachtet werden sollen daher drei Kenngrößen, die in den Medien vielfach genutzt und zitiert werden: die Reproduktionszahl, die Genauigkeit von Tests sowie die Dunkelziffer.

1. Erfassung der Infektionszahlen

Unabhängig davon, ob nun die Reproduktionszahl, die Genauigkeit von Tests oder die Dunkelziffer betrachtet werden sollen: Zentral für alle diese Kennzahlen und Konstrukte ist die zuverlässige Erfassung von Personen, die mit dem Virus SARS-CoV-2 infiziert sind. Diese „Fallzahl der Infektionen“ bildet die Grundlage dieser Kennzahlen. Sie ist definiert als die aufaddierte (kumulierte) Anzahl der bisher bestätigten Fälle von Infektionen mit SARS-CoV-2 (RKI 2020h). Die Anzahl aller – einschließlich unentdeckter – Infektionen in der Bevölkerung ist hingegen unbekannt und kann im besten Fall mit statistischen Verfahren hochgerechnet werden. Eine andere wichtige Zahl im Rahmen der Erfassung von Infektionszahlen ist die Anzahl der Neuinfektionen pro Tag, welche zusätzlich regelmäßig veröffentlicht wird.

Bis diese Zahlen für Deutschland vom RKI veröffentlicht werden, durchlaufen sie eine Vielzahl von Schritten, wodurch Verzögerungen entstehen können. Dieser Verzug wird primär durch die folgenden drei Faktoren verursacht (Dehning et al. 2020):

1. Inkubationszeit, das heißt die Zeit von der Ansteckung mit dem Virus bis zur Entwicklung von ersten Symptomen, mit einem Median von circa fünf bis sechs Tagen.
2. Verzug von bis zu drei Tagen bis der PCR-Test (siehe Infokasten) durchgeführt wird.
3. Verzögerung um weitere ein bis vier Tage bis das Testergebnis im Labor vorliegt und dem Gesundheitsamt vor Ort gemeldet wird.

Somit dauert es in der Regel mehrere Tage ab der Ansteckung, bis das Gesundheitsamt vor Ort Kenntnis über den Fall erlangt und dieser elektronisch erfasst wird. Dieser Tag gilt als Meldedatum.

Tab. 1 Fallzahlen der Infektionen mit SARS-CoV-2 (Stand 14.08.2020)

Bundesland	An das RKI elektronisch übermittelte Fälle (Stand 14.08.2020, 00:00 Uhr)			
	Anzahl	Fälle in den letzten 7 Tagen	7-Tage-Inzidenz	Todesfälle
Baden-Württemberg	38 270	522	4,7	1 859
Bayern	52 645	861	6,6	2 631
Berlin	9 976	383	10,2	224
Brandenburg	3 683	61	2,4	169
Bremen	1 828	30	4,4	56
Hamburg	5 821	176	9,6	264
Hessen	13 187	653	10,4	526
Mecklenburg-Vorpommern	959	33	2,1	20
Niedersachsen	15 318	424	5,3	656
Nordrhein-Westfalen	53 661	2 571	14,3	1 772
Rheinland-Pfalz	8 059	343	8,4	241
Saarland	2 976	57	5,8	174
Sachsen	5 705	72	1,8	225
Sachsen-Anhalt	2 099	47	2,1	64
Schleswig-Holstein	3 744	173	6,0	158
Thüringen	3 482	66	3,1	186
Gesamt	221 413	6 472	7,8	9 225

Quelle: Robert Koch-Institut (RKI 2020g)

Zusätzlicher Verzug entsteht bei der Übermittlung der Zahlen von Gesundheitsämtern an das Landesamt für Gesundheit und Lebensmittelsicherheit (LGL) und schließlich der Weitergabe der Zahlen an das RKI. Durch diesen Zeitverzug können die Zahlen von unterschiedlichen Quellen (zum Beispiel zwischen RKI und LGL) zeitweise voneinander abweichen (LGL 2020). Detaillierte Informationen zur Datenerfassung, des Meldewegs und des -verzugs sowie möglicher Differenzen zwischen Zahlen von unterschiedlichen Quellen finden sich in der Internetpräsenz des RKI (RKI 2020h).

Obgleich die vorliegenden Zahlen die Realität nicht vollständig abbilden können, es zu Meldeverzögerungen von 10 bis 14 Tagen kommen kann und unterschiedliche Quellen zeitweise voneinander abweichende Zahlen ausweisen, so bilden doch die offiziell vom RKI veröffentlichten, kumulierten Fallzahlen die zuverlässigsten und die objektiv am besten nachvollziehbaren Daten zu den Fallzahlen in Deutschland ab. Die in diesem Artikel verwendeten Zahlen beruhen daher stets auf den dem RKI vorliegenden beziehungsweise von diesem veröffentlichten Zahlen. Einzig für die Zahlen von Ende

Februar 2020 greifen wir auf die Daten der Johns Hopkins University zurück, da für diesen Zeitraum keine Zahlen vom RKI vorlagen.

Zur besseren Vergleichbarkeit der Infektionszahl zwischen Regionen kann die Fallzahl ins Verhältnis zur Einwohnerzahl gesetzt werden (Fallzahl pro 100 000 Einwohner). So können die Belastung des Gesundheitssystems und lokale Ausbruchsgeschehen schneller erkannt werden, insbesondere wenn hierbei Neuinfektionen betrachtet werden. Eine für die Politik wichtige Größe zur Beurteilung der aktuellen Infektionslage ist die „7-Tage-Inzidenz pro 100 000 Einwohner“. Dabei werden die Fälle mit Meldedatum der letzten sieben Tage aufaddiert und ins Verhältnis zu 100 000 Einwohnern gesetzt. Ist diese Zahl höher als „35“ (Bayern) beziehungsweise „50“ (Deutschland), werden gegebenenfalls Maßnahmen vor Ort ergriffen. Zum deutschlandweiten Vergleich der bis zum Redaktionsschluss veröffentlichten Infektionszahlen siehe Tabelle 1 (vgl. RKI 2020g).

2. Reproduktionszahl

Die Reproduktionszahl wird als eine der wichtigsten Kennzahlen in der Infektionsepidemiologie bezeichnet und wird dort unter anderem zur Prognose der Entwicklung von Infektionszahlen sowie zur Bestimmung von benötigten Impfquoten zur erfolgreichen Bekämpfung einer übertragbaren Krankheit herangezogen. Ihren Ursprung hat die Reproduktionszahl zwischen den 1880er- und 1920er-Jahren im Bereich der Demographie zum Zweck der Bestimmung der Geburten- und Bevölkerungsentwicklung (Heesterbeek 2002).

In der Epidemiologie beschreibt die Reproduktionszahl, wie viele Menschen eine infizierte Person im Mittel ansteckt. Dabei wird zwischen der Basisreproduktionszahl (R_0) sowie der effektiven Reproduktionszahl (R_{eff}) unterschieden. Erstere beschreibt die Anzahl der Personen, die eine infizierte Person im Mittel ansteckt, wenn ein Erreger auf eine Bevölkerung trifft, in der noch keine Immunität gegen den Erreger besteht, noch kein Impfstoff verfügbar ist sowie (noch) keine Maßnahmen zur Eindämmung der Ausbreitung einer Infektionskrankheit getroffen wurden (Diekmann et al. 1990). Die effektive Reproduktionszahl hingegen

beschreibt die durchschnittliche Anzahl der Personen, die eine infizierte Person im Mittel ansteckt unter Berücksichtigung von möglichen (Teil-)Immunitäten in der Bevölkerung (zum Beispiel da Patientinnen und Patienten nach einer Infektion temporär immun sein können) sowie getroffener Maßnahmen zur Eindämmung. Die effektive Reproduktionszahl unterliegt somit einer zeitabhängigen Variabilität aufgrund sinkender Empfänglichkeit gegen das Virus und aufgrund der Implementierung von Eindämmungsmaßnahmen (Nishiura & Chowell 2009). Da es sich bei dem Virus SARS-CoV-2 um ein neuartiges Virus handelt, kann angenommen werden, dass es in der Bevölkerung zu Beginn keine Immunität gegen das Virus gab².

Die vielfach zitierte „kritische“ Reproduktionszahl (R) beträgt 1. Ist diese im Durchschnitt kleiner als 1, steckt jede infizierte Person im Mittel weniger als eine Person an und die Anzahl der Neuinfektionen nimmt ab; bei einem R-Wert größer als 1 steigt die Anzahl der Neuinfektionen, bei einem R-Wert von exakt 1,0 bleibt die Anzahl der Neuinfektionen im Verlauf der Zeit konstant.

Grundsätzlich wird die Reproduktionszahl (R_0 und R_{eff}) primär durch die folgenden Faktoren³ beeinflusst:

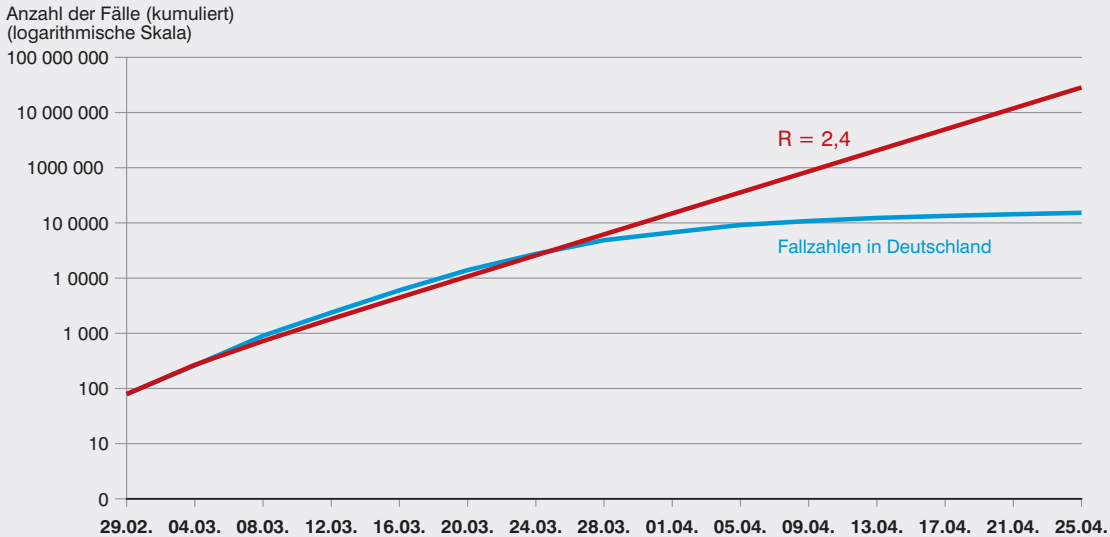
- die Anzahl der Kontakte eines Infizierten innerhalb einer bestimmten Zeiteinheit (k)
- die Ansteckungswahrscheinlichkeit bei Kontakt (b)
- die mittlere Dauer der Ansteckungsfähigkeit/Infektiosität (D)

Zumindest in der Theorie lässt sich die Reproduktionszahl als das Produkt dieser drei Einflussgrößen berechnen: $R = k \times b \times D$ (Lipsitch et al. 2003). Anhand dieser Formel lässt sich gut erkennen, an welchen Stellschrauben zum Zweck einer Eindämmung von Infektionskrankheiten angesetzt werden kann: Die Anzahl der Kontakte lässt sich zum Beispiel durch Kontaktbeschränkungen reduzieren, die Ansteckungswahrscheinlichkeit je nach Übertragungsart durch verschiedene Hygienemaßnahmen und die mittlere Dauer der (effektiven) Infektiosität im besten Fall indirekt durch eine frühzeitige Isolation/Absonderung der (wahrscheinlich) Infizierten.

² Wissenschaftlich ist dies aber nicht belegt; möglich wäre zum Beispiel eine sogenannte Kreuzimmunität (siehe Infokasten) auf Grund anderer zirkulierender Corona-Viren wie zum Beispiel der gewöhnlichen Erkältungsviren (Mateus et al. 2020).

³ Andere Einflussfaktoren sind zum Beispiel die Bevölkerungsdichte, soziokulturelle, klimatische oder Wetterbedingungen (Delamater et al. 2019).

Abb. 1
Entwicklung der kumulierten Fallzahlen in Deutschland vom 29. Februar bis 25. April 2020



Während die blaue Linie die gemeldeten Fallzahlen in Deutschland darstellt, zeigt die rote Linie die (theoretische) exponentielle Entwicklung der Fallzahlen für eine angenommene Basisreproduktionszahl von 2,4. Der Startwert für beide Linien beträgt hier 79, was der Anzahl der in Deutschland positiv gemeldeten SARS-CoV-2-Fälle vom 29.02.2020 entspricht (laut Johns Hopkins University; ab diesem Datum wurden die offiziell gemeldeten Fallzahlen laut RKI verwendet). Eigene Berechnungen.

Da für eine exakte Berechnung der Reproduktionszahl insbesondere zu Beginn der Infektionsausbreitung die Daten meist nicht in der erforderlichen Tiefe vorliegen, wird die Reproduktionszahl in der Regel mithilfe des vorhandenen Datenmaterials, komplexer mathematischer Modelle sowie einiger Annahmen geschätzt (Delamater et al. 2019). Auf diese Weise wurde die Basisreproduktionszahl für SARS-CoV-2 auf Grundlage verschiedener wissenschaftlicher Studien auf einen Wert zwischen 2,4 und 3,3 geschätzt (RKI 2020a)⁴. Dies bedeutet, dass 100 infizierte Personen im Durchschnitt 240 bis 330 weitere Personen anstecken würden und somit kumuliert 340 bis 430 Personen infiziert wären. Geht man nun davon aus, dass die ursprünglich 100 infizierten Personen nicht mehr infektiös sind, sondern lediglich die neu infizierten 240 bis 330 Personen die Erkrankung weitergeben können, wären eine Generation später (siehe Infokasten: Generationszeit; die Generationszeit für SARS-CoV-2 wird aktuell auf circa vier Tage geschätzt) kumuliert bereits zwischen 916 und 1 519 Personen infiziert, wovon wiederum nur ein Teil infektiös ist⁵.

Eine solche exponentielle Steigung mit dem konstanten Faktor R_0 ist jedoch nur zu Beginn einer

Infektionsausbreitung zu erwarten (siehe Abschnitt 2.2). Abbildung 1 stellt die Entwicklung der kumulierten Fallzahlen der (theoretisch) exponentiellen Entwicklung mit einer Reproduktionszahl von 2,4 den laut RKI gemeldeten tatsächlichen Fallzahlen in Deutschland gegenüber. Erkennbar ist, dass die Infektionsentwicklung in Deutschland zu Beginn der Ausbreitung von SARS-CoV-2 leicht über der exponentiellen Entwicklung beziehungsweise der Reproduktionszahl von 2,4 liegt. Ab circa der letzten März-Woche 2020 flacht die Kurve der gemeldeten Fallzahlen deutlich ab.

Auch wenn die Reproduktionszahl eine zentrale Rolle für die Bestimmung der Infektionsdynamik spielt, so gibt es aus Sicht der Statistik auch diverse Schwierigkeiten, die bei deren Interpretation beachtet werden sollten.

2.1 Die effektive Reproduktionszahl ist ein Schätzwert

Basisreproduktionszahl, effektive Reproduktionszahl, 4-Tages-R und 7-Tages-R – es gibt nicht „die eine“ Reproduktionszahl. Jeder Berechnung der unterschiedlichen Varianten der Reproduktionszahl liegen gewisse Annahmen und (vorläufige)

4 Zum Vergleich: hoch ansteckende Krankheiten, wie zum Beispiel die Masern, weisen eine Reproduktionszahl von > 10 auf; Grippe- beziehungsweise Influenzaviren in der Regel zwischen 1 und 2 (Aronson et al. 2020).

5 Berechnet über $\sum_{t=0}^2 N \cdot R_t^i$, wobei t = Generation und $N = 100$ (Anzahl Infizierte bei $t = 0$). Die Ansteckungsfähigkeit bei SARS-CoV-2 beginnt zum Teil schon vor dem Auftreten der Krankheitssymptome und nimmt mit Abklingen der Krankheitssymptome wieder ab. Bei anderen Erregern kann die Infektiosität Jahre oder Jahrzehnte anhalten.

wissenschaftliche Erkenntnisse zu Grunde, die relativ starke Auswirkungen auf ein solch sensibles Maß wie die Reproduktionszahl haben können⁶. Dies erschwert, wie im Folgenden ersichtlich wird, insbesondere die Interpretation von Schwankungen der Reproduktionszahl und sollte zur Vorsicht mahnen.

Zunächst ist es wichtig, herauszustellen, dass die Reproduktionszahl stets nur mit einem zeitlichen Verzug berechnet werden kann. Für die Berechnung der 4-Tages-Reproduktionszahl (vier Tage entsprechen der Generationszeit) benötigt man grundsätzlich die Anzahl der Neuinfektionen über einen Zeitraum von insgesamt acht Tagen: die durchschnittliche Anzahl der Neuinfizierten der Tage 1 bis 4 (erste Generation) und die durchschnittliche Anzahl der Neuinfizierten der Tage 5 bis 8 (zweite Generation). Dividiert man die durchschnittliche Anzahl der Neuinfizierten der zweiten Generation durch die durchschnittliche Anzahl der Neuinfizierten der ersten Generation, erhält man die (effektive) Reproduktionszahl.

Aufgrund der unsicheren Datenlage für die letzten Tage lässt das RKI bei der Berechnung der Reproduktionszahl die vorangegangenen drei Tage stets außer Betracht. Soll beispielsweise die Reproduktionszahl am 12. August 2020 berechnet werden, finden die Daten für den 9. bis 11. August keine Berücksichtigung (vorangegangene drei Tage). Datengrundlage für die Berechnung der Reproduktionszahl wären stattdessen der 1. bis einschließlich 4. August (erste Generation) sowie der 5. bis einschließlich 8. August 2020 (zweite Generation). Hierfür benötigt man für die in die Rechnung einfließenden Neuinfizierten dieses Zeitraums das Datum des Krankheitsbeginns (beziehungsweise Symptombeginns). Dieses liegt jedoch für circa 30 % der Patientinnen und Patienten nicht vor, so dass es in diesen Fällen seitens des RKI geschätzt beziehungsweise mithilfe des sogenannten Nowcasting-Verfahrens (siehe Infokasten) imputiert werden muss⁷. Soll zum Beispiel die Anzahl der Neuinfektionen für den 8. August 2020 (in unserem Beispiel der vierte und damit letzte Tag der zweiten Generation) bestimmt werden, so liegt dem RKI ein Großteil der hierfür erforderlichen Daten, einschließlich

des Datums des Symptombeginns, (noch) nicht vor und muss geschätzt werden. Dementsprechend weist das RKI darauf hin, dass die Werte von wenig zurückreichenden Tagen mit relativ hohen Unsicherheiten behaftet sind (RKI 2020i). Wird dann am nächsten Tag (13. August) erneut die Reproduktionszahl berechnet, liegen dem RKI in der Zwischenzeit einige aktualisierte Meldungen der Landesgesundheitsbehörden vor und es werden die „alten“ imputierten Werte auf dieser Grundlage neu geschätzt. Mit anderen Worten: Die Anzahl der Neuinfektionen der vergangenen Tage wird auf der Grundlage der sukzessive eingehenden Meldungen der Gesundheitsbehörden vom RKI täglich korrigiert. Dies resultiert folgerichtig in Aktualisierungen beziehungsweise Änderungen der Reproduktionszahl für ein und denselben Beobachtungszeitraum (in unserem Beispiel der 1. bis 8. August).

Darüber hinaus unterscheiden sich der „4-Tages-R-Wert“, der vom RKI vom Beginn der Epidemie an täglich berichtet wird und insbesondere in Phasen einer Niedrigprävalenz sehr sensitiv auf lokale Ausbruchsgeschehen reagiert, teilweise relativ deutlich vom seit dem 14.05.2020 parallel berichteten „7-Tage-R-Wert“, bei dem derartige Schwankungen (auch wochentagbedingte Schwankungen) bis zu einem gewissen Punkt ausgeglichen werden können, da dieser sich auf einen längeren Zeitraum bezieht (gemäß RKI bildet er das Infektionsgeschehen von vor etwa acht bis 16 Tagen ab). So betrug der Unterschied zwischen diesen beiden Reproduktionszahlen beispielsweise laut RKI-Lagebericht vom 21. Juni 0,85 Prozentpunkte (4-Tage-R: 2,88, 95%-Konfidenzintervall: [2,16, 3,73], 7-Tage-R: 2,03, 95%-Konfidenzintervall: [1,60, 2,49]) (RKI 2020k)⁸.

Schließlich können unterschiedliche Forschungsgruppen, bedingt durch unterschiedliche Annahmen oder abweichende Methodiken, zu voneinander abweichenden Reproduktionszahlen kommen. Wie der Wissenschaftsjournalist David Adam in einem „News Feature“ für die renommierte wissenschaftliche Fachzeitschrift *Nature* berichtet, präsentieren beispielsweise in Großbritannien circa zehn Forschergruppen ihre Ergebnisse für die Reproduktionszahl einem eingesetzten Kom-

6 Welche große Auswirkungen selbst kleine Änderungen der Reproduktionszahl haben können, zeigte auch Frau Bundeskanzlerin Dr. Merkel in der Bundespresskonferenz vom 15.04.2020 auf, als sie von Modellrechnungen berichtete, wonach bei einer Reproduktionszahl von 1,1 das Gesundheitssystem im Oktober an der Belastungsgrenze sei. Bereits ein geringfügig höherer Wert von 1,3 würde zu einer Überlastung im Juni – zwei Monate nach der Pressekonferenz – führen.

7 Gründe für das fehlende Datum des Symptombeginns sind in der Regel ein symptomfreier Verlauf, aber auch fehlende Angaben bei der Übermittlung an das RKI.

8 Üblicherweise ist der Unterschied zwischen diesen beiden R-Werten kleiner. Der Grund für den großen Unterschied liegt hier zu einem Großteil an einem lokalen Ausbruchsgeschehen im Raum Gütersloh in diesem Zeitraum.

tee der britischen Regierung, das sich folglich auf einen gewissen Wertebereich der Reproduktionszahl einigt (Adam 2020).

Wichtig zu erwähnen ist in diesem Zusammenhang, dass keine der oben genannten Methoden der Berechnung der Reproduktionszahl falsch ist; es werden zum Teil lediglich unterschiedliche Schwerpunkte gesetzt und geringfügig unterschiedliche Annahmen getroffen. Vielmehr sollten die variierenden Ergebnisse daran erinnern, dass es sich bei der Reproduktionszahl lediglich um einen Schätzwert handelt und – wie wiederholt vom RKI betont – zwar eine wichtige, bei weitem aber nicht die einzige Kennzahl für die Bestimmung der Infektionsdynamik ist. Sie kann nicht sensitiv genug sein, um sie im Bereich von wenigen Dezimalpunkten zu interpretieren – geringe Schwankungen und Abweichungen sind in der Regel statistisch unerheblich, liegen in der Regel innerhalb des statistischen Schätzfehlers und sind – entgegen mancher Schlagzeilen – kein Grund zur Besorgnis. Insbesondere ist sie zudem weniger geeignet, lokale Ausbruchsgeschehen zu identifizieren beziehungsweise treiben lokale Ausbruchsgeschehen den deutschlandweiten Mittelwert stark

in die Höhe (zudem sollten lokale Ausbrüche im Regelfall auffallen bevor sich diese auf die zeitverzögerte Berechnung der Reproduktionszahl auswirken). So ist es laut RKI wichtig, neben der Reproduktionszahl unter anderem die absolute Zahl der täglichen Neuinfektionen sowie die Schwere der Erkrankungen zu berücksichtigen (RKI 2020i). Adam (2020) ergänzt dies noch um die Anzahl der Todesfälle und Krankenhausaufnahmen. So müsse die absolute Zahl der Neuinfektionen klein genug sein, um eine effektive Kontaktpersonennachverfolgung zu ermöglichen und die Kapazitäten von Intensivbetten nicht zu überlasten (RKI 2020i).

2.2 Die (effektive) Reproduktionszahl unterliegt einer zeitlichen Dynamik

Die Reproduktionszahl entspricht ab einem bestimmten Zeitpunkt nicht mehr der Basisreproduktionszahl, da – unabhängig von eventuell eingeführten Eindämmungsmaßnahmen – (Teil-)Immunitäten auf Grund bereits überstandener Erkrankungen eine zunehmend stärkere Rolle bei der weiteren Infektionsentwicklung spielen (vorausgesetzt, genesene Patientinnen und Patienten sind zumindest temporär gegen eine weitere Ansteckung immun). Ab dann entspricht der – auch in

Abb. 2

Die Entwicklung der kumulierten Fallzahlen und der Reproduktionszahl laut Modellrechnung

Abb. 2a
Anzahl der Fälle (kumuliert)
in Tausend

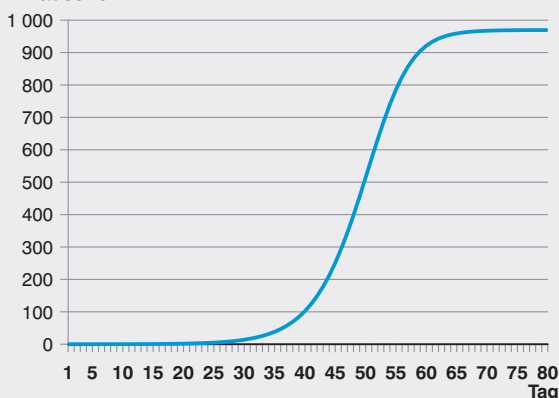
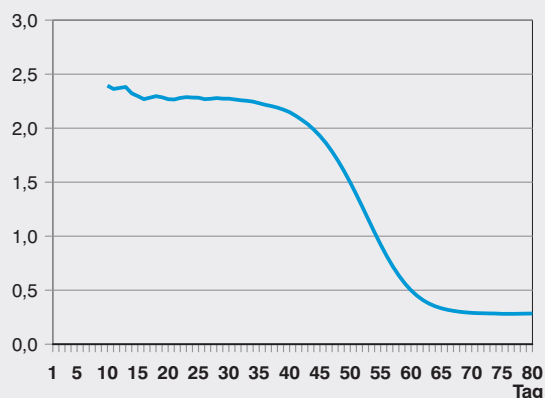


Abb. 2b
Reproduktionszahl



Die beiden Abbildungen zeigen das Ergebnis eines sog. SEIR-Modells, gemäß dem Mitglieder einer Population entweder (1) empfänglich für das Virus sind, (2) angesteckt wurden, (3) ansteckend sind oder (4) genesen sind (im Englischen: Susceptible, Exposed, Infectious, Recovered) und diese vier Stufen in dieser Reihenfolge durchlaufen. Die Größe der Ausgangspopulation beträgt hier 1 Million Personen und die Anzahl der anfänglich Infizierten 20. Es werden hier keine Maßnahmen zur Eindämmung der Infektionskrankheit getroffen und es wird davon ausgegangen, dass alle Mitglieder die gleiche Wahrscheinlichkeit einer Ansteckung haben und sich alle Mitglieder der Population zufällig vermischen (homogene Vermischung).

den Medien berichtete – „R-Wert“ nicht mehr der Basisreproduktionszahl R_0 . In Abbildung 1 lässt sich erkennen, dass ein ungebremst exponentieller Anstieg quasi ausgeschlossen ist. Steigt die Anzahl der Infizierten in der Bevölkerung an, steigt auch die Wahrscheinlichkeit, dass Kontaktpersonen von infizierten Personen bereits infiziert waren und daher (vermutlich) gegen eine weitere Ansteckung (temporär) immun sind. Anstatt eines exponentiellen Anstiegs der Infektionszahlen ist somit eine logistische Funktion anzunehmen, wie sie in Abbildung 2a dargestellt ist. Die dazugehörige Entwicklung der Reproduktionszahl, ausgehend von einer Basisreproduktionszahl von 2,4, ist in Abbildung 2b abgebildet (unter der Annahme einer Entwicklung einer Immunität nach überstandener Infektion über mindestens die Dauer des Beobachtungszeitraums von 80 Tagen).

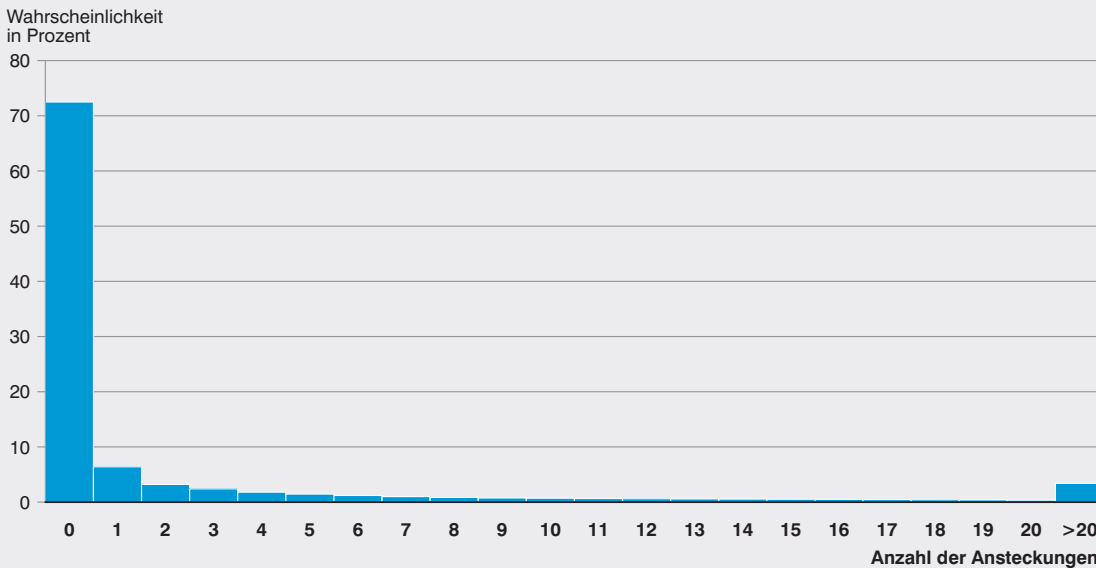
Die in Abbildung 2a und 2b dargestellten Ergebnisse des Simulationsmodells könnten suggerieren, dass die Strategie der Erreichung eines sogenannten Herdenschutzes beziehungsweise der Herdenimmunität (siehe Infokasten), die zumindest zeitweise von Großbritannien und Schweden verfolgt wurde, bereits zu einem relativ frühzeitigen Zeitpunkt erreicht sein könnte. Bereits nach etwas mehr als 50 Tagen wären laut dem Modell knapp 60% der Bevölkerung infiziert (womit der Herdenschutz bei einer Basisreproduktionszahl von 2,4 erreicht wäre). Die Ergebnisse des Modells können aber aus mehreren Gründen nicht auf die realen Bedingungen eines Infektionsverlaufs übertragen werden. Zunächst beträgt in diesem Modell die Populationsgröße lediglich eine Million Mitglieder; in einer deutlich größeren Bevölkerung wäre der Herdenschutz später erreicht. Viel wichtiger ist jedoch die dem Modell zugrunde liegende Annahme, dass keinerlei pharmakologischen (insbesondere Schutzimpfungen) oder nicht-pharmakologischen Interventionen, wie zum Beispiel Maskenpflicht oder Mindestabstand, eingeführt werden. Hierzu gehört auch, dass – selbst wenn es von staatlicher Seite keinerlei Interventionen geben sollte – die Bevölkerung ihr Verhalten nicht anpasst. Diese Annahme ist im Zeichen einer drohenden Infektion mit einem potenziell lebensgefährlichen Virus unrealistisch. Doch auch von staatlicher Seite

gab und gibt es in quasi allen Staaten einige Interventionen zur Eindämmung oder Milderung des Infektionsgeschehens – auch dort, wo mehr oder weniger explizit die Strategie einer Herdenimmunität verfolgt wurde. Diese Maßnahmen sind nicht zuletzt auch zum Zweck der Vermeidung der Überlastung des Gesundheitswesens essenziell. Bei einer Hospitalisationsrate von etwa 10% wäre dieses – ohne jegliche Interventionen – schon frühzeitig überlastet. So errechnet der Infektionsforscher Professor Meyer-Hermann vom Helmholtz-Zentrum für Infektionsforschung, dass die Herdenimmunität unter Berücksichtigung und Einhaltung der Kapazitäten des Gesundheitssystems erst in vielen Jahren erreicht würde; in Deutschland wäre demnach die Herdenimmunität bei täglich 2 500 Neuinfektionen in circa 25 Jahren erreicht (Der Tagesspiegel, 15.04.2020). Doch auch diese Rechnung gilt lediglich unter der Annahme einer anhaltenden Immunität von Personen, die die Krankheit durchlaufen haben. Die grundsätzliche Überlegung einer zeitlichen Dynamik der Entwicklung der Reproduktionszahl aufgrund von (temporären) Immunitäten bleibt hiervon aber unberührt; dieser Effekt trägt – neben den Maßnahmen zur Eindämmung des Infektionsgeschehens – dazu bei, dass die Kurve abflacht und eben kein unbegrenztes exponentielles Anwachsen der Infektionszahlen möglich ist.

2.3 Mittelwert, Streufaktor und Superspreader

Wie schon an der Definition der Reproduktionszahl ersichtlich, stellt diese einen Mittelwert über alle infizierten Personen dar. Zum einen bedeutet dies, dass die Reproduktionszahl typischerweise für Deutschland insgesamt ausgewiesen wird und somit in der Regel lokale Ausbruchsgeschehen nicht zuverlässig abgebildet werden. Doch gibt es noch ein weiteres Problem bei der Fokussierung auf den Mittelwert: Steckt beispielsweise jede infizierte Person zwei Personen an, beträgt die Reproduktionszahl 2,0. Eine identische Reproduktionszahl ergibt sich aber auch, wenn 90% der Infizierten lediglich eine Person anstecken und die anderen 10% jeweils 11 Personen. Durch diese sehr ungleiche Verteilung ist jedoch eine andere Infektionsdynamik zu erwarten, als wenn jede infizierte Person exakt zwei weitere Personen ansteckt. In letzterem Fall wäre keinerlei Streuung in der Anzahl der infi-

Abb. 3
Mögliche Wahrscheinlichkeitsverteilung der Anzahl der Ansteckungen



Bei einem angenommenen R-Wert von 2,5 und einem Dispersionsfaktor von $\kappa = 0,1$ (ca. 10 % der Infizierten sind für 80 % der Ansteckungen verantwortlich) stecken etwas mehr als 70 % der infizierten Personen keine weitere Person an.

zierten Personen, die eine Person ansteckt (die sogenannten sekundären Fälle), zu beobachten. Dieser Streufaktor beziehungsweise im wissenschaftlichen Diskurs Dispersionsfaktor genannt, ist in der Virologie und Epidemiologie eine relevante Größe. Für viele Infektionskrankheiten scheint eine 20/80-Regel zu gelten, die besagt, dass typischerweise 20% der Infizierten für circa 80% der verursachten Infektionen verantwortlich sind (Woolhouse et al. 1997). Für das SARS-CoV-1-Virus wurde sogar geschätzt, dass circa 10% der infizierten Personen für circa 80% der Infektionen verantwortlich waren (Lloyd-Smith et al. 2005); bei der Spanischen Grippe 1918 wird hingegen geschätzt, dass die Überdispersion nur eine geringe Rolle gespielt hat, das heißt, jede infizierte Person hat in etwa gleich viele Personen angesteckt (Fraser et al. 2011). Für SARS-CoV-2 weisen erste Forschungsergebnisse noch relativ stark voneinander abweichende Werte für den Dispersionsfaktor aus.

Was bedeutet diese Überdispersion beziehungsweise die Streuung für die Praxis? Endo et al. (2020) zeigen die theoretische Wahrscheinlichkeitsverteilung der Anzahl der sekundären Fälle für eine angenommene Basisreproduktionszahl von

$R_0 = 2,5$ und einem Dispersionsfaktor von $\kappa = 0,1$ wie dieser laut Endo et al. für SARS-CoV-2 angenommen werden kann. Wie Abbildung 3 zu entnehmen ist, würden entsprechend dieser negativen Binomialverteilung in diesem Fall etwas mehr als 70% der infizierten Personen keine weitere Person anstecken und weniger als 10% lediglich eine weitere Person. Etwa 3% der infizierten Personen würden bei dieser Wahrscheinlichkeitsverteilung hingegen mehr als 20 weitere Personen anstecken.

Das Infektionsgeschehen würde bei dieser ungleichmäßigen Verteilung also primär von einigen wenigen sogenannten Superspreadern beziehungsweise Superspreader-Ereignissen getrieben, bei denen mehr als 10 oder gar über 20 Personen von einer einzelnen Person angesteckt werden. Diese Ereignisse wären bei dieser schiefen Ansteckungsverteilung folglich für einen beträchtlichen Teil der Infektionsausbreitung und somit für den hohen Mittelwert der Reproduktionszahl verantwortlich.

Allerdings gilt es hier zu bedenken, dass in vielen Ländern initial eher kleinere Infektionsketten vorherrschten und es nur in ein paar wenigen Ländern

zu extensiven Ausbrüchen auf Grund großer Infektionsketten kam (Anderson et al. 2020). Weiter warnen Anderson et al. vor einer Überinterpretation von Superspreader-Ereignissen, die bei allen Infektionskrankheiten eine wichtige, aber nicht zwangsläufig überragende Rolle spielen.

Dennoch gibt es zwischenzeitlich international wie national zahlreiche Hinweise auf derlei Superspreader-Ereignisse. Sie treten insbesondere dort auf, wo Menschen in geschlossenen Räumen in engem Kontakt stehen beziehungsweise nicht die empfohlenen Hygieneabstände einhalten (können). Diese Vorfälle – beispielsweise Chorproben (Deutschlandfunk Kultur, 10.07.2020) oder beengtes Arbeiten in Schlachtbetrieben (Zeit Online, 23.07.2020) – suggerieren, dass wenige Personen viele andere Personen angesteckt haben. Überproportional häufig scheinen derlei Ereignisse von 20-39-jährigen prä- oder asymptomatischen Personen auszugehen (Furuse et al. 2020). Unklar ist, ob die Ursache für derlei Ereignisse bestimmten Eigenschaften der Person (zum Beispiel zahlreiche Kontakte, lautes Sprechen oder ein Ausscheiden einer hohen Viruslast) oder der Situation (zum Beispiel viele Personen in einem engen, geschlossenen Raum) geschuldet ist. Im Allgemeinen hängen aber diese persönlichen und situativen Aspekte eng zusammen (Asadi et al. 2019).

Eine stark ungleiche Infektionsverteilung bietet jedenfalls gewisse Chancen im Hinblick auf die Möglichkeiten zur Eindämmung der Infektionsdynamik. So kann mit zum Teil wenig einschneidenden Maßnahmen die Anzahl der Neuinfektionen relativ stark reduziert werden. So sollten insbesondere die Ereignisse vermieden werden, bei denen einzelne Personen eine Vielzahl an weiteren Personen anstecken, das heißt sogenannte Superspreader-Ereignisse müssen deutlich reduziert beziehungsweise im Idealfall vermieden werden (mit Blick auf Abbildung 3 wird diese Strategie im Englischen auch „chopping the tail“ oder „cutting the tail“ genannt, das heißt ein „Abschneiden“ der Ausläufer im rechten Teil der Abbildung 3; Kain et al. 2020). Endo et al. (2020) haben in einer statistischen Modellsimulation gezeigt, dass, wenn es gelingt, Superspreader-Ereignisse, bei denen mehr

als zehn Personen angesteckt werden, zu verhindern, die Reproduktionszahl von einem (hypothetischen) Wert von 2,0 auf 1,0 abgesenkt würde – ohne weitere Maßnahmen oder Beschränkungen zur Eindämmung (siehe auch Kain et al. 2020).

3. Genauigkeit von Tests

Seit Februar 2020 wurde in Deutschland eine Vielzahl von Testungen durchgeführt. Dabei sind folgende drei Testverfahren grundsätzlich voneinander zu unterscheiden: PCR-Testungen, Antikörper-Testungen und Antigen-Testungen.

- PCR-Testungen (siehe Infokasten) weisen mittels Abstriche aus den oberen Atemwegen das Virus direkt nach. Ein positives Testergebnis bedeutet demnach, dass der Patient mit einer bestimmten Wahrscheinlichkeit mit SARS-CoV-2 infiziert ist. Die Zeitspanne bis zum Testergebnis beträgt in etwa ein bis zwei Tage. Aufgrund einer hohen Laborauslastung kann es dabei aber natürlich zu Verzögerungen kommen. Der Nachweis des Virus ist allerdings nur bei einer aktiven Erkrankung möglich.
- Antikörpertestungen weisen eine Infektion dagegen indirekt nach. Hier wird im Blut des Patienten nicht nach dem Virus gesucht, sondern danach, ob der Patient Antikörper gegen SARS-CoV-2 gebildet hat. Das bedeutet, dass dieser Test erst nach einer überstandenen COVID-19-Erkrankung angewendet werden sollte, da sich erst dann Antikörper (möglicherweise) gegen das Virus gebildet haben. Fällt der Test positiv aus, so kann man mit einer bestimmten Wahrscheinlichkeit darauf schließen, dass der Patient mit SARS-CoV-2 infiziert war und damit gegebenenfalls temporär immun sein könnte. Antikörper-Testungen werden derzeit vor allem auch dafür verwendet, festzustellen, wie stark sich das Virus in der Bevölkerung tatsächlich ausgebreitet hat. Viele Personen haben aufgrund sehr schwacher (oder keinerlei) Symptome nicht bemerkt, dass sie sich mit dem Virus infiziert haben und haben sich daher auch keinem PCR-Test unterzogen. Derzeit laufen verschiedene Antikörper-Studien, deren Ergebnisse abzuwarten sind (siehe Abschnitt

4.2). Es ist allerdings darauf zu achten, dass der Patient erst nach einer gewissen Zeit Antikörper bildet. Der Test darf daher nicht zu früh eingesetzt werden, da dessen Ergebnis somit unzuverlässig wäre.

- Der Vorteil von Antigen-(Schnell-)Tests ist wiederum, dass sie bereits vor Einsetzen der ersten Symptome eingesetzt werden können. Des Weiteren liegt das Ergebnis innerhalb einer sehr kurzen Zeit vor (in etwa 15 Minuten). Ähnlich wie bei PCR-Tests wird mittels eines Nasen-Rachen-Abstrichs direkt nach den SARS-CoV-2 Viren gesucht und damit auf eine aktive Infektion getestet. Ein positives Ergebnis bedeutet, dass der Patient mit SARS-CoV-2 infiziert ist. Der Nachteil ist, dass Antigen-Testungen nicht sehr genau sind und eventuell mittels eines PCR-Tests abgesichert werden müssen (RKI 2020j). Im Folgenden wird vorrangig auf PCR-Tests Bezug genommen.

Die Testkapazitäten für PCR-Testungen der deutschen Labore wurden bereits im März 2020 stark ausgebaut. Von 7 115 Tests pro Tag Anfang März (KW 11) stieg diese Zahl auf 103 515 pro Tag Ende März (in KW 14) (Seifried & Hamouda 2020). Ende Mai (KW 21) erreichte die Testkapazität sogar 159 418 Tests pro Tag. Zuletzt stieg die Testkapazität sogar auf 202 761 Tests pro Tag in KW 36. Die tatsächlich durchgeführten Tests in Deutschland stiegen ebenfalls rasant an. Bereits in KW 10 wurden 124 716 Tests durchgeführt. Diese Zahl steigerte sich in KW 20 auf 430 882 Tests und in KW 35 sogar auf 1 101 299 Tests (RKI 2020d). Über diese Zahlen und auch die veröffentlichten Ergebnisse der Tests wurde in den Medien ausführlich berichtet. Aus Sicht der Statistik gilt es bei der Interpretation von Testergebnissen allerdings einiges mehr zu beachten. Zwar sollten Tests idealerweise zuverlässig aussagen, ob eine Person tatsächlich mit SARS-CoV-2 infiziert ist oder nicht; Tests mit hundertprozentiger Gewissheit gibt es jedoch nicht.

3.1 Sensitivität und Spezifität

Eine ausschließliche Betrachtung der als positiv getesteten Personen ist für eine umfassende Beurteilung der Testgüte nicht ausreichend. Um einen

Tab. 2 Erwartete Vierfeldertafel (Sensitivität 90%, Spezifität 95,5%, Prävalenz 1%)

		Person		
		erkrankt	nicht erkrankt	
Test- ergebnis	positiv	900 <i>(richtig positiv)</i>	4 455 <i>(falsch positiv)</i>	5 355 <i>(alle positiv Getesteten)</i>
	negativ	100 <i>(falsch negativ)</i>	94 545 <i>(richtig negativ)</i>	94 645 <i>(alle negativ Getesteten)</i>
		1 000 <i>(alle Erkrankten)</i>	99 000 <i>(alle Gesunden)</i>	100 000

umfassenden Überblick der Güte von Tests zu erhalten, sollten die Ergebnisse im Detail betrachtet werden. In Tabelle 2 sind beispielhaft die Testergebnisse abgebildet, die sich ergeben könnten, wenn 100 000 asymptomatische Personen ungezielt getestet werden. Wie man in Tabelle 2 erkennen kann, können bei derartigen Testungen vier Fälle beziehungsweise Kombinationen auftreten. Zwei Fälle werden auch in der Presse oft dargestellt. Dies sind zum einen die Personen, die ein positives Testergebnis aufweisen und die auch tatsächlich erkrankt sind ($n = 900$), und zum anderen die Personen, die negativ getestet werden und die tatsächlich nicht erkrankt sind ($n = 94 545$). Die Fälle in der ersten Gruppe werden als richtig Positive (RP) bezeichnet und die Personen in der zweiten Gruppe als richtig Negative (RN). In unserem Beispiel werden fast alle erkrankten Personen auch tatsächlich positiv getestet sowie ein sehr hoher Anteil der Gesunden als negativ eingestuft. Der Test liefert daher auf den ersten Blick gute Ergebnisse.

Aus diesen Häufigkeiten werden in der Statistik spezielle Kennzahlen berechnet: die Sensitivität und die Spezifität (Fahrmeir et al. 2016). Die Sensitivität gibt den Anteil der vom Test als positiv klassifizierten Personen an der Anzahl der tatsächlich erkrankten Personen an. In unserem Beispiel wird ein Großteil aller erkrankten Personen vom Test auch als positiv klassifiziert, jedoch nicht alle. Die Sensitivität beträgt damit 90%.

$$SENS = \frac{RP}{\text{alle tatsächlich Erkrankten}} = \frac{900}{1\,000} = 0,9$$

Die Spezifität zeigt analog den Anteil an Personen, die vom Test als negativ, das heißt als gesund, eingestuft wurden, gemessen an der Anzahl aller tatsächlich gesunden Personen. Da hier die Mehrheit aller Gesunden auch als negativ klassifiziert wird, reagiert der Test sehr spezifisch und die Spezifität beträgt damit 95,5%.

$$\text{SPEZ} = \frac{RN}{\text{alle tatsächlich Gesunden}} = \frac{94\,545}{99\,000} = 0,955$$

Das WHO Collaborating Centres in Genf hat eine Auswahl an aktuell existierenden PCR-Tests unter anderem bezüglich der Sensitivität und Spezifität evaluiert. Die letzte Aktualisierung vom 3. Juli 2020 zeigt, dass die aufgelisteten Werte hier zwischen 90% und 100% für die Sensitivität sowie zwischen 95% und 100% für die Spezifität variieren (FIND 2020). Es ist allerdings zu berücksichtigen, dass die klinische Sensitivität stark in Abhängigkeit vom Erkrankungsverlauf schwankt (siehe Abschnitt 3.2).⁹

Um nun die Güte des Tests beziehungsweise die Trennschärfe umfassend zu beurteilen, muss gleichzeitig betrachtet werden, wie oft der Test ein falsches Ergebnis liefert. Wie bereits oben beschrieben, ist eine hundertprozentig richtige Zuordnung zu den beiden Gruppen (gesund/krank) meist nicht möglich. Zuordnungen können in zweierlei Weise falsch sein. Einerseits könnte eine Person fälschlicherweise negativ getestet werden, obwohl sie eigentlich mit SARS-CoV-2 infiziert ist (in unserem Beispiel $n = 100$). Dies kann zum Beispiel passieren, wenn die Probenentnahme nicht sachgerecht durchgeführt oder transportiert wurde (Woloshin et al. 2020). Diese Ergebnisse entsprechen dann den falsch Negativen (FN). Andererseits wäre es aber genauso möglich, dass eine gesunde Person fälschlicherweise als positiv, das heißt als erkrankt getestet wird (in unserem Beispiel $n = 4\,455$). Dies entspricht den falsch Positiven (FP).

Entweder werden also Erkrankte übersehen und damit im Falle von SARS-CoV-2 die vorgesehenen Quarantänemaßnahmen nicht eingehalten, was zu weiteren unentdeckten Neuansteckungen führen kann, oder es werden Personen in Quarantäne geschickt, die eigentlich gesund sind. Beide Fälle sollten nach Möglichkeit nahe Null liegen.

3.2 Vorhersagewahrscheinlichkeiten

In der Statistik lassen sich diese falschen Zuordnungen mit weiteren Kennzahlen bestimmen, die allerdings nur selten berichtet werden, obwohl dies gerade bei den falsch Positiven von hoher Wichtigkeit wäre. So lässt sich in unserem Beispiel der Anteil der korrekt positiv getesteten Personen gemessen an allen positiv getesteten Personen, der sogenannte positive Vorhersagewert, berechnen. Wie Tabelle 2 zeigt, wird die Mehrheit der tatsächlich infizierten Personen auch als krank erkannt ($n = 900$), doch werden darüber hinaus viele Personen ebenfalls als krank eingestuft, obwohl sie eigentlich gesund sind ($n = 4\,455$). Damit beträgt der positive Vorhersagewert lediglich rund 17%.

$$\text{pos. Vorhersagewert} = \frac{RP}{(RP + FP)} = \frac{900}{5\,355} = 0,167$$

Gemäß dem positiven Vorhersagewert bedeutet ein positives Testergebnis in Tabelle 2 also, dass man lediglich mit einer Wahrscheinlichkeit von rund 17% auch tatsächlich erkrankt ist. Anders gesprochen bedeutet dies aber auch, dass man mit einer Wahrscheinlichkeit von 83% eigentlich gesund sein könnte, obwohl man positiv getestet wurde. Testergebnisse, die eine derartige niedrige Vorhersagewahrscheinlichkeit haben, führen unweigerlich zu weiteren Untersuchungen, die das Ergebnis spezifizieren müssen (siehe Drost 2020). Im Fall von SARS-CoV-2 würde ein derart hoher Anteil an falsch Positiven ohne weitere Spezifizierung fälschlicherweise zu Quarantänemaßnahmen führen, die in diesem Umfang nicht notwendig wären.

Analog kann auch der negative Vorhersagewert berechnet werden, welcher für das Infektionsgeschehen eine weitaus wichtigere Rolle spielt, als die Rate der falsch Positiven. In Tabelle 2 liegt der negative Vorhersagewert bei 99,9%, da fast alle gesunden Personen als negativ klassifiziert werden sowie nur eine geringe Anzahl an erkrankten Personen als falsch negativ getestet wird.

$$\text{neg. Vorhersagewert} = \frac{RN}{(RN + FN)} = \frac{94\,545}{94\,645} = 0,999$$

Tabelle 2 zeigt allerdings die Fallzahlen bei einer falsch Negativen-Rate von etwa 10%. In der Wissenschaft wird der Anteil der falsch Negativen in

⁹ Diagnostische/ klinische Sensitivität und Spezifität beschreiben, wie gut der Test Kranke und Gesunde unterscheiden kann. Die Werte werden durch tatsächliche Testungen mit Probanden erhoben. Die klinische Sensitivität gibt hier das Vermögen des Tests an, möglichst alle Erkrankten zu erfassen. Die klinische Spezifität zeigt das Vermögen des Tests, ein Krankheitsbild gezielt zu erfassen und somit Fehlzusammenhänge gering zu halten. Davon zu unterscheiden sind die analytische Sensitivität und Spezifität. Erstere beschreibt die Nachweisstärke des Labortests und zweite inwieweit der Test das misst, was er vorgibt zu messen. Die Werte können durch technische Eigenschaften des Tests variiert werden (Bundesgesundheitsblatt 2008).

Zusammenhang mit SARS-CoV-2 teilweise aber höher eingeschätzt. Wie bereits erwähnt, kann er genau genommen sogar erheblich schwanken, je nachdem an welchem Tag die Probenentnahme erfolgt ist (Kucirka et al. 2020). Kucirka et al. zeigen, dass am ersten Tag einer vermuteten Infektion in keinem Abstrich das Virus festgestellt wurde. Alle vermeintlich Kranken wurden daher als gesund klassifiziert. Auch am vierten Tag der Infektion sinkt die mittlere Rate der falsch Negativen nur auf 67%. Erst am Tag acht (beziehungsweise drei Tage nach dem Auftreten der ersten Symptome) fällt die mittlere falsch Negativen-Rate auf akzeptable 20%. Schließlich steigt die Rate der falsch Negativen sogar ab Tag 21 wieder, das heißt, das Virus ist dann nicht mehr sicher nachweisbar, obwohl der Infizierte das Virus bereits weitergegeben haben könnte. Es wird damit deutlich, dass hohe Werte der Sensitivität und Spezifität alleine nicht ausreichen, um die Trennschärfe eines Tests zu beurteilen, sie aber unabdingbar für eine Anwendung von Testungen sind.

3.3 Prävalenz

Neben der Trennschärfe ist aus statistischer Sicht ein weiterer Aspekt für die Aussagekraft von Testergebnissen unabdingbar: die Prävalenz, das heißt der Grad der Verbreitung der Krankheit in der Population (siehe auch Infokasten). Die Berücksichtigung der Prävalenz lässt die Testergebnisse in einem anderen Licht erscheinen. In unserem ersten Beispiel (siehe Tabelle 2) haben wir eine Prävalenz von 1% angenommen. Das bedeutet, dass nur ein geringer Anteil der Bevölkerung erkrankt ist. Anhand des niedrigen positiven Vorhersagewertes haben wir gesehen, dass die Aussagekraft des Tests, hinsichtlich der Erkennung von tatsächlich erkrankten Personen, stark eingeschränkt ist. In Tabelle 3 nehmen wir nun eine Prävalenz von 20% an, das heißt eine stärkere Verbreitung der Krankheit in der Bevölkerung. Diese Prävalenz könnte auch bei einer Testung in einem Risikosetting (zum Beispiel in einem Altersheim) auftreten (Horvath 2020).

Tabelle 3 zeigt, dass die Anzahl der richtig Positiven ($n = 18\,000$) und richtig Negativen ($n = 76\,400$) relativ hoch ist. Zusätzlich geht die Anzahl an falsch Positiven nun zurück ($n = 3\,600$) und der positive

Tab. 3 Erwartete Vierfeldertafel (Sensitivität 90%, Spezifität 95,5%, Prävalenz 20%)

		Person		
		erkrankt	nicht erkrankt	
Test- ergebnis	positiv	18 000 <i>(richtig positiv)</i>	3 600 <i>(falsch positiv)</i>	21 600 <i>(alle positiv Getesteten)</i>
	negativ	2 000 <i>(falsch negativ)</i>	76 400 <i>(richtig negativ)</i>	78 400 <i>(alle negativ Getesteten)</i>
		20 000 <i>(alle Erkrankten)</i>	80 000 <i>(alle Gesunden)</i>	100 000

Vorhersagewert steigt damit auf 83%. Das bedeutet, dass prozentual erheblich weniger Personen, die eigentlich gesund sind, fälschlicherweise als erkrankt klassifiziert werden und ein positives Testergebnis aussagekräftiger wird. Je höher die Prävalenz in der Gesamtpopulation ist, desto mehr Gesunde werden richtig klassifiziert und desto weniger fallen die falsch Positiven ins Gewicht. Gleichzeitig ist aber auch zu beachten, dass die Anzahl der falsch Negativen stark ansteigt. Es werden nun bedeutend mehr Infizierte übersehen. Der negative Vorhersagewert fällt damit von 99,9% auf 97,5%. Damit wird auch deutlich, dass sich die beiden Möglichkeiten einer falschen Klassifizierung gegenseitig beeinflussen. Es ist wichtig, zu entscheiden, welche Fehlklassifikation als schwerwiegender betrachtet wird. Im Fall von SARS-CoV-2 beziehungsweise bei PCR-Testungen wird es vermutlich die Rate der falsch Negativen sein, da diese Personen das Virus unwissentlich weiterverbreiten, obwohl sie eigentlich negativ getestet wurden. Es ist daher von hoher Wichtigkeit, die Anzahl der falsch Negativen so gering wie möglich zu halten. Bei der Testung ist allerdings der bereits im vorherigen Abschnitt erwähnte Aspekt des Zeitpunktes der Probenentnahme zu beachten. Letztlich ist im Zusammenhang mit der Prävalenz noch zu berücksichtigen, dass diese durchaus auch gruppenspezifisch unterschiedlich hoch sein kann, beispielsweise bezüglich Alter und Geschlecht. Für die Berechnung der statistischen Kennzahlen müsste dies weiter berücksichtigt werden (Gigerenzer 2009).

Wie wir bisher gezeigt haben, sollte bei einer PCR-Testung der Anteil der falsch Negativen möglichst

gering sein. Dies resultiert daraus, dass keine Infizierten übersehen werden sollten, da sie das Virus weiter verbreiten könnten. Der Anteil der falsch Positiven könnte idealerweise mittels einer hohen Prävalenz auf einem relativ niedrigen Niveau gehalten werden. Bei Antikörper-Testungen bedeutet dagegen ein positives Ergebnis, dass der Patient Antikörper gebildet hat und folglich bereits mit dem SARS-CoV-2 infiziert war. Demnach könnten auch in der Praxis alle im Antikörper-Test positiv getesteten, also auch die falsch positiv getesteten, Personen zunächst davon ausgehen, dass sie bereits an COVID-19 erkrankt waren und unter Umständen sogar immun sind. Folglich würden diese Personen möglicherweise weniger auf Hygienemaßnahmen und Abstandsregeln achten, was bei den falsch Positiven dazu führt, dass sie das Infektionsgeschehen weiter fördern, indem sie sich selbst infizieren. Gleichzeitig muss allerdings erwähnt werden, dass derzeit wissenschaftlich nicht geklärt ist, ob jede Infektion mit SARS-CoV-2 auch zur Bildung von entsprechenden Antikörpern führt oder der Nachweis von Antikörpern auch bedeutet, dass die Person eine schützende Immunität ausgebildet hat und wenn ja, wie lange diese Immunität anhalten könnte (Solbach et al. 2020). Daher gilt gerade bei Antikörpertests, dass der positive Vorhersagewert möglichst hoch sein sollte (Horvath 2020). Das heißt der Anteil an falsch Positiven, also an Personen, die bisher nicht mit SARS-CoV-2 infiziert waren, aber fälschlicherweise positiv auf Antikörper getestet wurden, sollte möglichst gering sein. Bei den derzeit eingesetzten ELISA-Verfahren („Enzyme Linked Immunosorbent Assays“, siehe Infokasten) wird die Sensitivität zwischen 72,2% und 94,4% beziehungsweise die Spezifität zwischen 96,7% und 99,5% angegeben (Horvath 2020, LMU 2020).

Der Effekt der Prävalenz darf natürlich nicht dazu verleiten, Tests erst nach einer entsprechend hohen Durchseuchungsrate durchzuführen. Besser sollten Tests frühzeitig und gezielt durchgeführt werden.¹⁰ Dies erklärt sich daraus, dass die Testung einer Population von symptomatischen Personen eine hohe Prävalenz aufweist und damit der positive Vorhersagewert relativ hoch sein wird. Der richtige Zeitpunkt der Probenentnahme im

Verlauf der Krankheit kann zusätzlich den negativen Vorhersagewert erhöhen. „Die Erkrankungsrate in der Bezugspopulation bestimmt also, mit welcher Wahrscheinlichkeit ein positives Testergebnis korrekt ist.“ (Gigerenzer et al. 2009). Auch das RKI rät auf seiner Webseite von einer ungezielten Testung von asymptomatischen Personen ab (RKI 2020b).

Die aktuellen Werte bezüglich klinischer Sensitivität und Spezifität für PCR-Tests werden derzeit recht unterschiedlich berichtet. Der Medizinreport des Deutschen Ärzteblatts (Schlenger 2020) gibt für die derzeitigen PCR-Tests eine Sensitivität zwischen 71% und 98% an. Wir wollen im Folgenden kurz aufzeigen, welchen Effekt derartige unterschiedliche Werte der Sensitivität auf den positiven sowie den negativen Vorhersagewert haben. Abbildungen 4 und 5 zeigen die Werte des positiven beziehungsweise negativen Vorhersagewertes für unterschiedliche Prävalenzen. Die Prävalenzen variieren hier zwischen sehr niedrigen Werten, wie sie beispielsweise für Deutschland angenommen werden könnten (beispielsweise 1%) und sehr hohen Werten, wie sie etwa in einer Isolierstation zu erwarten wären (80%) (Schlenger 2020).

Abbildung 4 zeigt, was wir bereits erläutert haben: Ein hoher Sensitivitätswert begünstigt (mit steigender Prävalenz) auch einen hohen positiven Vorhersagewert, das heißt, positive Testergebnisse sind zuverlässiger. Abbildung 5 zeigt aber nun den Effekt für den negativen Vorhersagewert. Es wird deutlich, dass bei einem niedrigen Sensitivitätswert und trotz sehr hoher Prävalenz, der negative Vorhersagewert nur 44% erreicht. Ein negatives Testergebnis ist nun nur noch mit einer Wahrscheinlichkeit von 44% auch tatsächlich negativ. 56% der eigentlich negativ getesteten Personen tragen das Virus also in sich und könnten es in der Bevölkerung unwissentlich weiter verbreiten. Allerdings erreicht der negative Vorhersagewert auch bei einer hohen Sensitivität und einer hohen Prävalenz nur einen Wert von etwa 89%. Ein stark erhöhtes Auftreten der Krankheit führt damit einerseits dazu, dass weniger Gesunde als infiziert klassifiziert werden, gleichzeitig jedoch erhöht sich die Anzahl der erkrankten Personen, die fälschlicher-

¹⁰ Das Robert Koch-Institut (RKI) gibt Empfehlungen und Leitlinien vor, unter welchen Bedingungen ein Test auf SARS-CoV-2 durchgeführt werden soll. Zu Beginn der Pandemie empfahl das RKI nur dann einen Test auf eine Corona-Infektion durchzuführen, wenn ein Kontakt zu einem Infizierten bekannt war. Dadurch wurden Personen mit typischen COVID-19-Symptomen, aber ohne Kontakt zu Infizierten nur im Ausnahmefall (oder aber in Abhängigkeit von Testkapazitäten) getestet. Später genügte spezifische Symptome und es war nicht mehr notwendig, einen Kontakt nachzuweisen. Allerdings wurde erst im Verlauf der Erforschung des Virus erkannt, dass beispielsweise auch der Verlust des Geruchssinns als typisches Symptom gelten kann. Wenn dies als alleiniges Symptom auftrat, konnte so die Infektion übersehen werden. Hinzu kommt, dass die Kapazitäten der Testlabore erst ausgebaut werden mussten und so die Umsetzung der Empfehlungen möglicherweise zunächst regional unterschiedlich war.

Abb. 4
Positiver Vorhersagewert, differenziert nach Prävalenz und Sensitivität

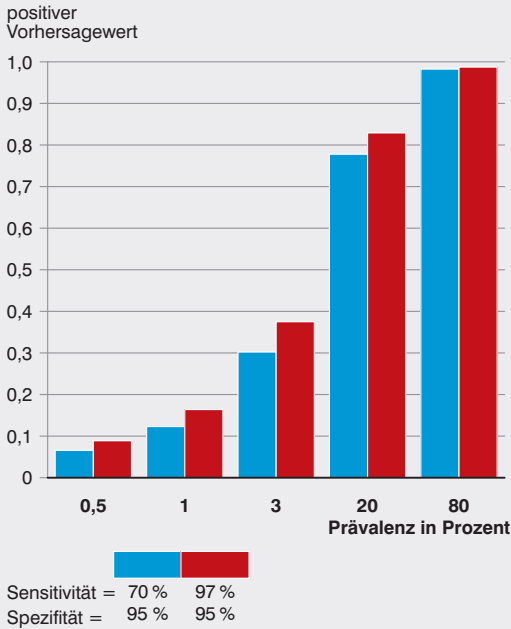
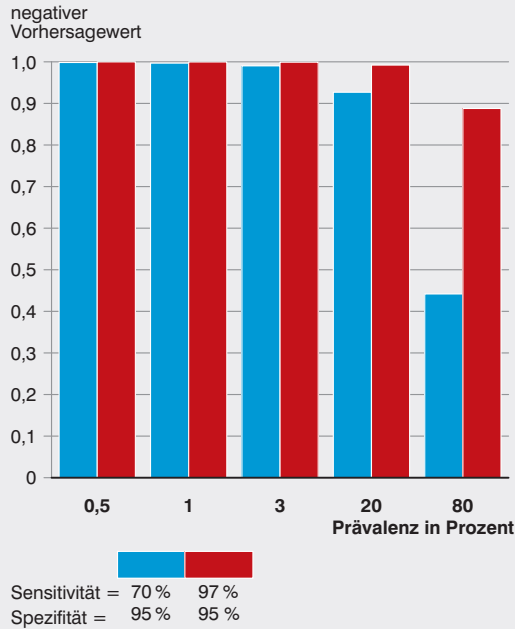


Abb. 5
Negativer Vorhersagewert, differenziert nach Prävalenz und Sensitivität



weise übersehen wurden. Es gilt also einerseits die Entwicklung der Prävalenz sowie den richtigen Zeitpunkt bei einer Testung auf SARS-CoV-2 zu finden oder durch andere Teststrategien die Anzahl der falsch negativen Tests zu minimieren. Nur so können die Testergebnisse den bestmöglichen Wert annehmen.

Sensitivität, Spezifität und Prävalenz spielen bei der Interpretation der Testergebnisse somit eine große Rolle. Problematisch ist bei neuartigen Krankheiten, wie COVID-19, dass die tatsächliche Prävalenz unbekannt beziehungsweise zu wenig erforscht ist und sich natürlich weiterhin dynamisch entwickelt. Hier ist darauf hinzuweisen, dass zunächst zuverlässige Testmethoden entwickelt und diese fortlaufend im Rahmen von geeigneten Teststrategien zum Einsatz gebracht werden müssen, um dem dynamischen Verlauf gerecht zu werden. Hinzu kommen noch Differenzierungen nach Subpopulationen mit unterschiedlichen Erkrankungsrisiken. Aus statistischer Sicht wäre die Bestimmung der Prävalenz zu einem bestimmten Zeitpunkt nur mittels Testung von repräsentativen Stichproben oder durch Vollerhebungen zuverlässig

ermittelbar. Da dies einen erheblichen Aufwand mit sich bringt, muss die Entwicklung der Prävalenz oder auch die Dunkelziffer auf andere Art und Weise bestimmt werden.

4. Dunkelziffer

Die tatsächliche Prävalenz und damit die Dunkelziffer¹¹ von Infektionen mit SARS-CoV-2 sind derzeit unbekannt. Die durch offizielle Stellen wie das RKI oder das Bayerische Landesamt für Gesundheit und Lebensmittelsicherheit (LGL) veröffentlichten Fallzahlen der (Neu-)Infektionen mit SARS-CoV-2 geben nur Auskunft über die positiv getesteten Fälle und nicht darüber, wie viele Infizierte es tatsächlich in der Bevölkerung gibt.

Zunächst lässt sich feststellen, dass es einen direkten Zusammenhang zwischen der Dunkelziffer und der verfolgten Teststrategie eines Landes beziehungsweise einer Region gibt. Umfassende Tests auf Infektionen mit SARS-CoV-2 können die Dunkelziffer verringern. Der Nachteil von umfangreichen Tests besteht allerdings darin, dass insbesondere bei einer niedrigen Prävalenz die falsch-positiven Ergebnisse zunehmen (siehe Abschnitt 3).

11 „Dunkelziffer“ bezeichnet im Folgenden die Zahl der Personen, die mit SARS-CoV-2 infiziert waren, jedoch nicht als offiziell bestätigte Fälle erfasst wurden.

Die unterschiedlichen Teststrategien müssen insbesondere auch bei internationalen Vergleichen von Fallzahlen in die Interpretation miteinbezogen werden.

Zur Einordnung der internationalen Fallzahlen kann die „Positivrate“ verwendet werden: Wie viele der durchgeführten Tests haben ein positives Ergebnis (Nachweis der Infektion)? Eine niedrige Positivrate kann laut WHO unter bestimmten Voraussetzungen ein Indiz für eine geringe Verbreitungsgeschwindigkeit des Virus sein. So empfiehlt die WHO, dass die Positivrate für mindestens 14 Tage geringer als 5% sein sollte, bevor Lockerungen von Eindämmungsmaßnahmen beschlossen werden. Allerdings kann dieses Kriterium zur Beurteilung nur herangezogen werden, wenn es im Land eine umfangreiche Kontaktnachverfolgung und Testung von Verdachtsfällen gibt und im Beurteilungszeitraum keine Änderung der Teststrategie erfolgte (WHO 2020). Die Positivrate kann also nicht ohne Kenntnis der Teststrategie eines Landes interpretiert werden. Die Positivrate kann sinken, weil Testkapazitäten ausgebaut werden und auch weniger eindeutige Verdachtsfälle getestet werden; sie sinkt aber auch, wenn die Verbreitung des Virus in der Bevölkerung abnimmt – ohne dass Testkapazitäten ausgebaut werden. Die bisher höchste Positivrate wurde für Deutschland in der Kalenderwoche 14 (Ende März 2020) mit 9,0% festgestellt, Anfang Juli betrug sie nur noch 0,6% und in KW 35 0,74% (RKI 2020f, RKI 2020d).

Durch den Zusammenhang mit der Teststrategie ist in aller Regel auch ein zeitlicher Gradient im Hinblick auf die Dunkelziffer zu beobachten: Da gegebenenfalls zunächst die ersten Infektionsfälle übersehen werden, Tests entwickelt oder Testkapazitäten aufgebaut werden müssen, kann man davon ausgehen, dass vor allem zu Beginn einer sich entfaltenden Epidemie ein besonders hoher Anteil der neu mit SARS-CoV-2 Infizierten unentdeckt bleibt. Mittels eines massiven Ausbaus der Testkapazitäten beziehungsweise einer konsequenten Nachverfolgung und Testung aller Kontaktpersonen der Infizierten können prozentual mehr Neuinfizierte entdeckt werden (RWI 2020, FiveThirtyEight 2020).

Doch auch hier gilt es wieder, die oben erläuterten Besonderheiten bei statistischen Tests zu berücksichtigen, wie zum Beispiel die falsch Positiven oder falsch Negativen. Des Weiteren wird bei COVID-19 vermutet, dass ein beachtenswerter Anteil der Infektionen symptomfrei oder symptomarm verläuft. Ihr Anteil schwankt je nach wissenschaftlicher Studie zwischen 4% und 41% (Byambasuren et al. 2020). Durch einen symptomfreien Verlauf wurden möglicherweise infizierte Personen übersehen, entsprechend auch nicht getestet und fanden daher keinen Eingang in die Fallzahlenstatistik.

Es gibt mehrere Methoden zur Berechnung der Dunkelziffer, zum Beispiel

1. Berechnung anhand der Zahl der Toten und der Infektionssterblichkeit (siehe Infokasten) durch Modellrechnungen.
2. (Repräsentative) Antikörperstudien.

Bei beiden Methoden werden Annahmen verwendet, die zur Vorsicht bei der Interpretation mahnen. Zusätzlich können statistische Zufallsfehler und systematische Fehler die Ergebnisse verzerren (siehe Ende des Abschnitts 4.3).

4.1 Berechnung anhand der Zahl der Toten und der Infektionssterblichkeit

Eine relativ unkomplizierte Methode zur Schätzung der Dunkelziffer ist die Verwendung der Todesfälle, der Infektionssterblichkeit (siehe Infokasten) und der (angenommenen) Dauer vom Zeitpunkt der Infektion bis zum Tod. Die Fallsterblichkeitsrate lässt sich nicht verwenden, da diese anhand der Todesfälle pro „Fall“ errechnet wird. Sie hängt davon ab, was als „Fall“ definiert wird. In der Regel sind dies die offiziell gemeldeten Infizierten, weswegen sie höher ist als die Infektionssterblichkeitsrate, die alle tatsächlich Infizierten mit einbezieht.

Nimmt man beispielsweise an, dass die (bisher unbekannte) Infektionssterblichkeitsrate bei 1% liegt und die Dauer vom Zeitpunkt der Infektion bis zum Tod durchschnittlich 24 Tage beträgt¹², kann die Anzahl der Infizierten vor 24 Tagen anhand der Anzahl der Verstorbenen am heutigen Tag hochgerechnet werden. Wird am heutigen Tag beispielsweise ein Toter gemeldet, betrug die Anzahl der

¹² Laut Verity et al. (2020) vergehen durchschnittlich 17,8 Tage (95% Konfidenzintervall: [16,9, 19,2]) vom Einsetzen der Symptome bis zum Tod. Wird noch die Inkubationszeit von circa 5 bis 6 Tagen addiert (RKI 2020a), ergibt dies einen Zeitraum von durchschnittlich 24 Tagen.

Neuinfektionen vor 24 Tagen 100. Wenn man nun noch eine Annahme über die Generationszeit und Reproduktionszahl trifft, kann auch die Anzahl der Neuinfektionen für alle Tage dazwischen berechnet werden. Die Dunkelziffer ergibt sich dann aus der Differenz zwischen den auf Basis der Sterblichkeitsrate hochgerechneten Fällen und den gemeldeten Fallzahlen. Bei allen vier Größen (Infektionssterblichkeit, durchschnittliche Dauer bis zum Tod, Generationszeit, Reproduktionszahl) der Berechnung müssen Annahmen getroffen werden. Die Zahl der Personen, die mit oder an COVID-19 gestorben sind, ist relativ genau bekannt, aber selbst hier muss die Annahme getroffen werden, dass keine Fälle übersehen wurden. Möglicherweise wurden frühe COVID-19-Todesfälle nicht mit der neuen Viruserkrankung in Verbindung gebracht und auch Meldeverzögerungen können die Modellierung beeinträchtigen (Flaxman et al. 2020).

Ein Beispiel für eine epidemiologische Modellierung, die mit diesen Größen arbeitet und die anhand der beobachteten Todesfälle die Zahl der Infizierten in elf europäischen Ländern (darunter Deutschland) zurückrechnet, ist der Ansatz von Flaxman et al. (2020). In allen untersuchten Ländern wird anhand des Modells eine Untererfassung festgestellt. Anhand des Modells wird geschätzt, dass am 4. Mai 2020 in Deutschland 0,85% [95%-Konfidenzintervall: 0,66%, 1,1%] der Bevölkerung infiziert waren, was 710 000 [95%-Konfidenzintervall: 550 000, 930 000] Personen entspricht. Offiziell wurden vom RKI für diesen Tag 163 175 bestätigte Fälle gemeldet (RKI 2020j), sodass die Zahl der tatsächlich Infizierten 4,4-mal höher gewesen wäre als bekannt.

Darüber hinaus gibt es weitere Faktoren, die bei derartigen Berechnungen berücksichtigt werden müssen. Das „demografische Skalierungsmodell“, das vom Max-Planck-Institut für Demografieforschung in Rostock und der Universität Helsinki entwickelt wurde, greift sparsam auf nur wenige Informationen zurück, bezieht allerdings auch demografische Faktoren mit ein. Benötigt werden Zahlen zu den COVID-19 assoziierten Todesfällen pro Altersgruppe, zur Infektionssterblichkeit sowie zur Restlebenserwartung. Im Modell wird die alters-

spezifische Infektionssterblichkeit eines Referenzlandes (in diesem Fall die Provinz Hubei, China) unter Berücksichtigung der Restlebenserwartung auf andere Länder übertragen (Bohk-Ewalda et al. 2020). Durch Einbeziehung der Restlebenserwartung sollen Unterschiede in der Altersstruktur, Vorerkrankungen und im Gesundheitssystem der Länder berücksichtigt werden. Die Übertragung der Infektionssterblichkeit erfolgt auf diese Weise: Für eine Altersgruppe eines Landes wird die Infektionssterblichkeit derjenigen Altersgruppe des Referenzlandes verwendet, die eine möglichst gleiche demografisch errechnete Anzahl an verbleibenden Lebensjahren hat.

Für das Modell müssen jedoch zwei Annahmen getroffen werden, auf deren geringe Belastbarkeit die Forschenden hinweisen:

1. Die Anzahl der Personen, die an beziehungsweise mit COVID-19 gestorben sind, wird im Land, für das die Dunkelziffer berechnet werden soll, korrekt erfasst.
2. Die Infektionssterblichkeitsrate eines Referenzlandes kann auf andere Länder übertragen werden.

Die Forschenden gehen allerdings davon aus, dass die Zahl der COVID-19-Todesfälle belastbarer ist als die Zahl der bestätigten Infektionsfälle. Eine zusätzliche Schwierigkeit besteht allerdings darin, dass die Todesfälle pro Altersgruppe vorliegen müssen.

Als Ergebnis stellen Bohk-Ewalda et al. (2020) bei Heranziehung der Todeszahlen bis 13. Mai 2020 eine große Variation der Dunkelziffer über Länder hinweg fest. So schätzen sie mittels des entwickelten Modells die tatsächlichen Fallzahlen für Italien sechs Mal höher und für Deutschland lediglich 1,8-mal höher als die bestätigten Fälle. Für Italien wurden in einer Antikörperstudie (siehe Abschnitt 4.2) vergleichbare Ergebnisse festgestellt (Istat/Ministero della Salute 2020). Allerdings weisen Bohk-Ewalda et al. darauf hin, dass die Ergebnisse mit hoher Unsicherheit behaftet sind, was sich auch in der großen Spanne des 95%-Konfidenzintervalls zeigt (dieses liegt für Italien zwischen dem Faktor 2,7 bis 16,5 und für Deutschland zwischen 0,8 und 4).

4.2 (Repräsentative) Antikörperstudien

Eine gute Möglichkeit zur Klärung der epidemiologischen Fragestellung, wie viele Personen die Infektion mit SARS-CoV-2 unentdeckt durchlebt haben, sind Antikörperstudien. Dabei kommen Tests von Blutproben auf Antikörper bei einer idealerweise repräsentativen Auswahl der Testpersonen zum Einsatz. Ein Ziel dieser seroepidemiologischen Studien, bei denen Blutserum untersucht wird, ist auch die Ermittlung der Infektionssterblichkeitsrate (siehe Infokasten).

In mehreren europäischen Staaten (zum Beispiel Österreich, Italien, Spanien und Frankreich) wurden Antikörperstudien mit teilweise bis zu 150 000 Teilnehmern bereits durchgeführt oder sind geplant – auch mit Unterstützung der nationalen Statistikämter. In Deutschland werden epidemiologische Studien durch das RKI sowie durch Universitätskliniken und Forschungsinstitute durchgeführt. Das RKI verfolgt drei Ansätze, um die Verbreitung des Coronavirus in Deutschland zu erforschen (RKI 2020c):

1. Untersuchung von Blutspenden auf Antikörper aus ganz Deutschland.
2. Lokale Antikörperstudien (CORONA-MONITORING lokal) an vier besonders betroffenen Orten (siehe Ende des Abschnitts für erste Ergebnisse für die Gemeinde Kupferzell).
3. Eine für Deutschland repräsentative Untersuchung mit circa 30 000 Personen ab September 2020. Es werden sowohl Querschnitt- als auch Längsschnittstudien durchgeführt, bei denen entweder die Allgemeinbevölkerung oder auch spezielle Bevölkerungsgruppen untersucht werden (siehe Poethko-Müller et al. 2020 für eine Übersicht).

Bei den oben genannten Studienansätzen soll mit einem Test einer Blutprobe auf spezifische Antikörper nachgewiesen werden, ob die Person Kontakt mit dem SARS-CoV-2-Virus hatte und daraufhin eine Immunantwort entwickelt hat. Die Anzahl der Studienteilnehmer, bei denen entweder ein positiver PCR-Test (siehe Infokasten) oder ein positiver Antikörpertest vorliegt, werden addiert; der ermittelte Anteil an (repräsentativ ausgewählten) Studienteilnehmern wird dann auf die Grundgesamtheit hoch-

gerechnet. Das Ergebnis ist eine statistische Schätzung mit entsprechenden Konfidenzintervallen.

Neben der Komplexität der Stichprobenauswahl führt auch eine mangelnde Testgenauigkeit (siehe Abschnitt 3) dazu, dass diese Studien keine ganz exakten Ergebnisse liefern können. Zudem muss der Stichprobenumfang möglichst groß sein, wenn von einer (bisher) niedrigen Prävalenz ausgegangen wird, um trotzdem möglichst verlässliche statistische Schätzungen vornehmen zu können. Je seltener ein gesuchtes Merkmal in einer Grundgesamtheit auftritt, umso größer sollte die Stichprobe sein, mit der das Vorhandensein in der Grundgesamtheit geschätzt werden soll. Hinzu kommt, dass zwar der Nachweis von Antikörpern ein Hinweis auf eine durchlebte Infektion ist. Allerdings weist eine erste vorläufige Studie aus Lübeck darauf hin, dass möglicherweise nicht alle Infizierten Antikörper entwickeln (Solbach et al. 2020). Dies würde bedeuten, dass selbst repräsentative Studien die Dunkelziffer nicht verlässlich ermitteln könnten beziehungsweise, dass dieser Faktor in die Hochrechnung mit einfließen muss.

Für die Übertragbarkeit der Ergebnisse auf eine Grundgesamtheit ist die repräsentative und zufällige Auswahl der Untersuchungspersonen zentral, um etwa die Verbreitung von Antikörpern in einer Grundgesamtheit statistisch schätzen zu können. Die Auswahl der Studienteilnehmer und die methodischen Rahmenbedingungen der Studie müssen zur Einordnung der Ergebnisse miteinbezogen werden. Eine Studie mit Freiwilligen ist nicht für die Gesamtbevölkerung repräsentativ, insbesondere da sich wahrscheinlich überproportional viele Personen mit Verdacht auf eine (überstandene) Infektion zur Studienteilnahme melden würden. Dadurch würden potenziell mehr Infizierte in die Stichprobe gelangen als in der Grundgesamtheit existieren und die Antikörperverbreitung würde überschätzt. Bei der Kommunikation der Ergebnisse ist es wichtig, auf die Unsicherheiten hinzuweisen, die mit den statistischen Schätzungen von Prävalenzen einhergehen. Die Ergebnisse bewegen sich aufgrund von Zufallsfehlern in einem interpretationsbedürftigen Konfidenzintervall. Zusätzlich können die Daten von als repräsentativ geplanten Studien aufgrund

von systematischen, nicht zufälligen Fehlern verzerrt sein. Folgende Ursachen für systematische Verzerrungen sind beispielsweise denkbar:

1. **Selektionsbias:** Die Teilnahmebereitschaft von repräsentativ ausgewählten Studienteilnehmern kann abhängig von Eigenschaften der potenziellen Teilnehmer wie Alter und/oder Bildungshintergrund sein oder sie ist durch das Studiendesign beeinflusst (zum Beispiel Smartphone-Nutzung, Sprachbarrieren). Ist die Teilnahmerauswahl nicht zufällig erfolgt, weil beispielsweise Freiwillige gesucht werden, ist die Verzerrung durch den Selektionsbias besonders groß.
2. **Messfehler:** Eine fehlerhafte Messung kann zum Beispiel durch teilweise verunreinigtes Labormaterial oder durch falsche Durchführung der Tests auftreten. Takahashi et al. (2020) weisen darauf hin, dass die SARS-CoV-2 Antikörpertests bei Personen mit niedrigeren Antikörperkonzentrationen möglicherweise nur unzureichende Nachweise einer durchlebten Infektion erbringen. Als möglichen Grund nennen sie, dass die Sensitivität der Tests anhand schwer erkrankter Personen validiert wurde, die eventuell höhere Antikörperkonzentrationen aufweisen.

Nicht zuletzt sind bei der Studieninterpretation die Bevölkerungsstruktur und die Situation im untersuchten Gebiet zu berücksichtigen: Handelt es sich beispielsweise um ein besonders betroffenes Gebiet? Waren bereits Maßnahmen in Kraft, die die Möglichkeit zur Ansteckung verringert haben? Leben im Vergleich zum deutschen Durchschnitt sehr viele Ältere/Jüngere im betreffenden Gebiet? Fehlt eine solche Einordnung in den Kontext beziehungsweise in die Erklärungsreichweite einer Studie, wenn Ergebnisse kommuniziert werden, kann das zu öffentlichen Kontroversen führen, wie dies beispielsweise bei der „Heinsberg-Studie“ geschehen ist. So kann der Anteil der infizierten Personen, der durch Antikörperstudien an besonders betroffenen Gebieten ermittelt wurde, nur schwerlich auf die Gesamtbevölkerung übertragen werden. Dies gilt sowohl für die „Heinsberg-Studie“ in Gangelt in Nordrhein-Westfalen (Streeck et al. 2020) als auch für Studien in Ischgl, Österreich oder in Kup-

ferzell. In Gangelt wurden Anfang April bei 15,5% (95%-Konfidenzintervall: [12,3%, 19,0%]) der etwa 1 000 Einwohner aus zufällig ausgewählten Haushalten Antikörper nachgewiesen, in Ischgl Ende April sogar bei 42,4% der untersuchten Einwohner (79% der Einwohner nahmen teil, APA-Science 2020). In der lokalen Antikörperstudie des RKI in der Gemeinde Kupferzell (Hohenlohekreis, Baden-Württemberg) wurden bei 7,7% der 2 203 Teilnehmenden Antikörper nachgewiesen, was bedeutet, dass 3,9 mal so viele Infektionen nachgewiesen wurden als bislang bekannt. Diese Dunkelziffer ist aber spezifisch für Kupferzell. Bei 28,2% derjenigen mit bereits bekannter zurückliegender Corona-Infektion wurden keine Antikörper nachgewiesen (RKI 2020). Was dies für die Immunität bedeutet, muss Gegenstand weiterer Studien sein.

Schlussbetrachtung

In weiten Teilen Bayerns wütete im Jahr 1633 die Pest. Das Dorf Oberammergau war bis dahin dank strenger Wachen an den Stadtgrenzen weitestgehend verschont geblieben – bis ein Tagelöhner Sehnsucht nach Frau und Kindern hatte und sich zum Kirchweihfest heimlich ins Dorf schlich. Kurz darauf verstarben er und weitere 84 Bewohnerinnen und Bewohner binnen weniger Wochen – an der vom Tagelöhner eingeschleppten Pest. In ihrer Verzweiflung gelobten die Oberammergauer, regelmäßig die Geschichte vom Leiden und Sterben Jesu aufzuführen, wenn Gott sie nur von ihren Leiden erlöse. Die Gebete wurden anscheinend erhört, denn hiernach infizierte sich der Legende nach niemand mehr in dem Ort (Spiegel, 27.06.2009). Bis heute finden die Oberammergauer Passionsspiele alle zehn Jahre statt – doch wie viele andere Veranstaltungen fielen sie in diesem Jahr ebenfalls der Corona-Pandemie zum Opfer. So mussten sie zum dritten Mal in ihrer Geschichte verschoben werden – auf das Jahr 2022 (zwei Mal wurden sie abgesagt) (Zeit Online, 19.03.2020).

Aus heutiger Sicht können zwei Lehren aus dem Wunder von Oberammergau gezogen werden: Selbst wenn korrekt sein sollte, dass nach den Gebeten niemand in Oberammergau mehr an der Pest erkrankte, so ist ein kausaler Zusammenhang doch sehr unwahrscheinlich. Heutzutage würde

man – hoffentlich – Korrelation und Kausalität nicht mehr so einfach miteinander vermischen. Vielmehr ist der Umgang mit der Corona-Pandemie ein gutes Beispiel dafür, dass politische Entscheidungen in heutiger Zeit oftmals auf wissenschaftlichen Erkenntnissen basieren. Nicht nur auf Bundesebene, sondern auch in Bayern wurden Expertenräte – bestehend aus Wissenschaftlerinnen und Wissenschaftlern verschiedener Disziplinen – gebildet, die Bundes- und Landes-Regierung wissenschaftlich beraten, um die Grundlage für evidenzbasierte Entscheidungen zu schaffen.

Auch wenn manche in der Öffentlichkeit kursierenden Zahlen und Statistiken rund um die Corona-Pandemie mit Vorsicht zu genießen sind und es aus statistischer Sicht einige Fallstricke zu beachten gibt, weist Deutschland ein im internationalen Vergleich (bisher) recht mildes Infektionsgeschehen auf. Dies deutet darauf hin, dass in Deutschland die auf wissenschaftlichen Empfehlungen basierenden Infektionsschutzmaßnahmen größtenteils zum richtigen Zeitpunkt getroffen wurden. So gibt es inzwischen erste wissenschaftliche Studien, die die Auswirkungen der getroffenen Maßnahmen auf das weitere Infektionsgeschehen in Deutschland evaluieren. Eine Gruppe des Max-Planck-Instituts für Dynamik und Selbstorganisation in Göttingen untersuchte anhand von Computermodellen beispielsweise, wie sich die einzelnen Maßnahmen mit zeitlicher Verzögerung auf die Infektionsdynamik in Deutschland ausgewirkt haben (Dehning et al. 2020; siehe auch Küchenhoff et al. 2020 mit einem sehr ähnlichen Ansatz auf Basis von Infektionszahlen aus Bayern).

Mithilfe bayesscher Regressionsmodelle mit Bruchpunkten konnten die Wissenschaftlerinnen und Wissenschaftler zeigen, dass die Bremsung der Ausbreitung sehr stark mit den Zeitpunkten der getroffenen Maßnahmen korrelierte und erkennbar in drei Stufen erfolgte: Sowohl das Verbot von Großveranstaltungen am 7. März 2020 als auch die Schließungen vieler Geschäfte, Schulen und Kindertagesstätten in der Woche vom 16. März und die allgemeinen Kontaktbeschränkungen sowie Schließungen sämtlicher nicht-systemrelevanter Geschäfte ab der Woche vom 22. März hatten deutliche Rückgänge

der Ausbreitungsrate zur Folge. So sank die Ausbreitungsrate durch die Maßnahmen vom 7. März von vorher 30% auf nur noch 12%. Durch die weiteren Maßnahmen von Mitte März sank die Rate weiter von 12% auf 2%. Eine positive Wachstumsrate hätte aber nach wie vor ein potenziell exponentielles Wachstum der Fallzahlen bedeutet. Erst durch die Beschlüsse vom 22. März konnte die Ausbreitungsrate schließlich auf -3% (95%-Konfidenzintervall: $[-5\%, 2\%]$) gesenkt werden.

Die Autorinnen und Autoren weisen jedoch in ihrem Artikel explizit darauf hin, dass dieser Wert nur geringfügig unter null liegt und sich die Infektionsdynamik bei kleineren Änderungen auch leicht wieder umkehren kann. Ein weiteres zentrales Ergebnis der Studie ist, dass die positiven Auswirkungen der Maßnahmen stets erst nach circa zwei Wochen in den Daten sichtbar wurden – was im Hinblick auf die bekannten Verzögerungen (Inkubationszeit, Testabnahme, Laboruntersuchung und Ergebnisübermittlung) mehr als plausibel erscheint. Dies bedeutet auch für zukünftige Maßnahmen, seien es Verschärfungen oder Lockerungen, dass deren Auswirkungen immer erst entsprechend diesem zeitlichen Verzug evaluiert werden können.

Zusammenfassend weist die Göttinger Forschungsgruppe darauf hin, dass sich das Infektionsgeschehen durchaus auch wieder ausbreiten könne. So schreibt die Leiterin der Forschungsgruppe Dr. Viola Priesemann: „Wenn alle Personen weiterhin sehr vorsichtig sind, und die Kontaktnachverfolgung durch die Gesundheitsämter effektiv greift, und gleichzeitig alle neuen Infektionsherde früh aufgespürt und eingedämmt werden, dann können die Fallzahlen weiterhin sinken. Wie genau sich die Zahlen in Zukunft entwickeln, hängt also entscheidend von unserem Verhalten, dem Einhalten von Abstandsempfehlungen und den Hygienemaßnahmen ab.“ (Max-Planck-Gesellschaft 2020).



Falls nicht anders angegeben, beruhen die folgenden Definitionen auf dem „Fachwörterbuch Infektionsschutz und Infektionsepidemiologie“ des RKI (RKI 2015).

ELISA-Test (enzyme-linked immuno sorbent assay): Standardisierter Reaktionsablauf zum Nachweis einer Substanz (zum Beispiel Antikörper).

Generationszeit: Die Generationszeit entspricht der mittleren Zeitspanne von der Infektion einer Person bis zur Infektion der von ihr angesteckten Folgefälle. Sie beträgt bei SARS-CoV-2 nach aktuellen Erkenntnissen circa vier Tage (RKI 2020i).

Herdennimmunität/Herdenschutz: Effekt, dass ein gewisser Anteil immuner Individuen innerhalb einer Population (entstanden durch Impfung oder abgelaufene Infektionen) auch nichtimmunen Personen einen relativen Schutz bietet. Um den Anteil der Bevölkerung zu errechnen, der infiziert sein muss, um eine Herdennimmunität zu erreichen, ist die einfachste Möglichkeit eine Berechnung über die Basisreproduktionszahl: $H = 1 - 1/R_0$ (Porta 2008). Bei einem R_0 von circa 3 also: $1 - 1/3 = 2/3$ beziehungsweise 66,67% der Bevölkerung. Bei dieser simplen Berechnungsmethode geht man von der (unrealistischen) Annahme aus, dass alle Mitglieder der Population gleichermaßen am gesellschaftlichen Leben teilhaben und die gleichen Chancen auf eine Ansteckung haben (Spiegel, 04.07.2020; Quanta, 30.06.2020).

Infektionssterblichkeit (Infektionssterblichkeitsrate, engl. „Infection Fatality Rate“): Die Infektionssterblichkeitsrate gibt die Wahrscheinlichkeit an, an einer Infektion zu versterben unter der Bedingung, dass eine Infektion erfolgt ist. Für die Berechnung werden alle an und mit einer Infektion Verstorbenen durch alle Infizierten, auch die nicht gemeldeten Fälle, dividiert. Da die Anzahl aller Infizierten nicht bekannt ist, muss diese Größe in wissenschaftlichen Modellen geschätzt werden. Die Infektionssterblichkeit ist deswegen keine feste Größe. Die Fallsterblichkeitsrate (engl. „Case Fatality Rate“) hingegen berücksichtigt nur laborbestätigte, gemeldete Fälle und wird berechnet, indem die Anzahl der an und mit einer Infektion Verstorbenen durch die Anzahl aller gemeldeten Fällen in einer Population dividiert wird (Streeck et al. 2020, S. 3).

Inzidenz: Eine Rate als Maß der Häufigkeit des Auftretens neuer Ereignisse (Infektionen, Erkrankungen und andere) in einer bestimmten Bevölkerung als geschlossener Grundgesamtheit in einem bestimmten Zeitraum. Die Inzidenz bezeichnet zum Beispiel das absolute Risiko, unter bestimmten Bedingungen zu erkranken. Zur besseren Vergleichbarkeit wird sie auf einen als Bezugsgröße gewählten Teil der gesamten zu untersuchenden Population (zum Beispiel 100 000 Einwohnerinnen und Einwohner) bezogen.

Kreuzimmunität: Wechselseitige Immunität, bei der der Kontakt mit einem Erregerstamm eine (Teil-)Immunität gegen einen ähnlichen Erreger bewirkt (Kamo & Sasaki 2002).

Nowcasting: Methode des RKI, mit deren Hilfe eine Schätzung des Verlaufs der Anzahl von bereits erfolgten SARS-CoV-2-Erkrankungsfällen in Deutschland unter Berücksichtigung des Diagnose-, Melde- und Übermittlungsverzugs vorgenommen wird.

PCR-Test (Polymerase-Kettenreaktion oder Polymerase-Chain-Reaction): Methode zur exponentiellen Vervielfältigung (in vitro) spezifischer DNA-Sequenzen (Saiki et al. 1988).



Prävalenz: Eine epidemiologische Maßzahl zur Angabe der Bestandshäufigkeit, des Maßes einer Verbreitung, des Bestands an Erkrankten, Infizierten (Antikörperprävalenz) oder von Merkmalsträgerinnen und -trägern in einer bestimmten Bevölkerung zu einem bestimmten Zeitpunkt oder innerhalb einer bestimmten Zeitperiode. Angaben zur Prävalenz können sich auf bestimmte Befunde oder Ereignisse beziehen.

Sensitivität: Ein Maß der Wahrscheinlichkeit, mit einem Test eine bestimmte Krankheit (beziehungsweise einen gesuchten Erreger oder mit ihm assoziierte Veränderungen) sicher zu erkennen. Bei hoher Sensitivität ist allerdings mit mehr falsch positiven Ergebnissen zu rechnen.

Spezifität: Ein Maß der Wahrscheinlichkeit, eine bestimmte gesuchte Krankheit (beziehungsweise einen gesuchten Erreger oder mit ihm assoziierte Veränderungen) bei Gesunden sicher auszuschließen beziehungsweise bei Gesunden einen negativen Befund zu erhalten.

Übertragungszeit/Serienintervall: Zeitdauer, während der eine Übertragung eines Erregers tatsächlich stattfinden kann oder in einem konkreten Fall stattfinden konnte.

Vorhersagewert: Der prädiktive Wert eines Tests im Rahmen der Diagnostik oder des Screenings gestattet Aussagen zur Wahrscheinlichkeit, dass eine Person mit einem positiven Testergebnis auch wirklich Trägerin oder Träger des nachzuweisenden Merkmals ist, dass zum Beispiel eine gesuchte Infektion tatsächlich vorliegt. Der Vorhersagewert eines Tests im Rahmen eines Screenings wird bestimmt von der Sensitivität und der Spezifität des Tests sowie von der Prävalenz des nachzuweisenden Merkmals.

Literatur

Adam, David (2020): A guide to R – the pandemic’s misunderstood metric. *Nature*, 583, S. 346–348 (<https://doi.org/10.1038/d41586-020-02009-w>).

Anderson, Roy M./Heesterbeek, Hans/Klinkenberg, Don/Hollingsworth, T. Déirdre (2020): How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, 395, S. 931–934 ([https://doi.org/10.1016/S0140-6736\(20\)30567-5](https://doi.org/10.1016/S0140-6736(20)30567-5)).

Aronson, Jeffrey K./Brassey, Jon B./Mahtani, Kamal R. (2020): „When will it be over?“ An introduction to viral reproduction numbers, R_0 and R_e . In: Centre for Evidence-Based Medicine (CEBM), University of Oxford: www.cebm.net/covid-19/when-will-it-be-over-an-introduction-to-viral-reproduction-numbers-r0-and-re/, Beitrag vom 14.04.2020, abgerufen am 16.06.2020.

Asadi, Sima/Wexler, Anthony S./Cappa, Christopher D./Barreda, Santiago/Bouvier, Nicole M./Ristenpart, William D. (2019): Aerosol emission and superemission during human speech increase with voice loudness. *Scientific Reports*, 9, S. 1–9 (<https://doi.org/10.1038/s41598-019-38808-z>).

Austria Presse Agentur (APA-Science 2020): Ischgl-Studie: 42,4 Prozent sind Antikörperpositiv: https://science.apa.at/rubrik/medizin_und_biotech/Ischgl-Studie_42_4_Prozent_sind_Antikoerper-positiv/SCI_20200625_SCI39451352255218286, Beitrag vom 25.06.2020, abgerufen am 08.07.2020.

- Bohk-Ewalda, Christina/Dudel, Christian/Myrskylä, Mikko (2020): A demographic scaling model for estimating the total number of COVID-19 infections (<https://doi.org/10.1101/2020.04.23.20077719>).
- Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz (2008): Leitlinien Diagnostische Validität, 51, S. 1353–1356 ([10.1007/s00103-008-0726-z](https://doi.org/10.1007/s00103-008-0726-z)).
- Byambasuren, Oyungerel/Cardona, Magnolia/Bell, Katy/Clark, Justin/McLaws, Mary-Louise/Glasziou, Paul (2020): Estimating the extent of true asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. (<https://doi.org/10.1101/2020.05.10.20097543>).
- Dehning, Jonas/Zierenberg, Johannes/Spitzner, F. Paul/Wibrall, Michael/Neto, Joao P./Wilczek, Michael, Priesemann, Viola (2020): Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*, eabb9789. (<https://doi.org/10.1126/science.abb9789>).
- Delamater, Paul L./Street, Erica J./Leslie, Timothy F./Yang, Y. Tony/Jacobsen, Kathryn H. (2019): Complexity of the Basic Reproduction Number (R₀). *Emerging Infectious Diseases*, 25, S. 1–4. (<https://doi.org/10.3201/eid2501.171901>).
- Der Tagesspiegel (15.04.2020). Helmholtz-Forscher erklärt an einer Zahl, warum strenge Corona-Regeln bleiben müssen: www.tagesspiegel.de/wissen/herdenimmunitaet-waere-erst-in-25-jahren-erreicht-helmholtz-forscher-erklaert-an-einer-zahl-warum-strenge-corona-regeln-bleiben-muessen/25742256.html, abgerufen am 16.08.2020.
- Deutschlandfunk Kultur (10.07.2020): Charité-Studie zum Singen in Corona-Zeiten – Chöre sind Aerosol-Schleudern. www.deutschlandfunkkultur.de/charite-studie-zum-singen-in-corona-zeiten-choere-sind.1013.de.html?dram:article_id=480346, abgerufen am 24.07.2020.
- Diekmann, O./Heesterbeek, J. A. P./Metz, J. A. J. (1990): On the Definition and the Computation of the Basic Reproduction Ratio R₀ in Models For Infectious-Diseases in Heterogeneous Populations. *Journal of Mathematical Biology*, 28, S. 365–382 (<https://doi.org/10.1007/BF00178324>).
- Drosten Christian (2020): Coronavirus-Update (39): Welche Chancen neue Tests bieten: www.ndr.de/nachrichten/info/39-Coronavirus-Update-Welche-Chancen-neue-Tests-bieten,podcastcoronavirus206.html#Spezifitaet
- Endo, Akira/Abbot, Sam/Kucharski, Adam J./Funk, Sebastian (2020): Estimating the over-dispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research*, 5: 67, S. 1–17 (<https://doi.org/10.12688/wellcomeopenres.15842.3>).
- Fahrmeir, Ludwig/Künstler, Rita/Pigeot, Iris/Tutz, Gerhard (2016): *Statistik. Der Weg zur Datenanalyse*. 8., überarb. und erg. Auflage. Springer Spektrum, Berlin/ Heidelberg.
- FiveThirtyEight (2020): Coronavirus Case Counts Are Meaningless – Unless you know something about testing. And even then, it gets complicated. <https://fivethirtyeight.com/features/coronavirus-case-counts-are-meaningless/>, abgerufen am 16.06.2020.
- Flaxman, Seth/Mishra, Swapnil/Gandy, Axel/Unwin, H. Juliette T./Mellan, Thomas A./Coupland, Helen/Whittaker, Charles/Zhu, Harrison/Berah, Tresnia/Eaton, Jeffrey W./Monod, Mélodie/Imperial College COVID-19 Response Team/Ghani, Azra C./Donnelly, Christl A./Riley, Steven M./Vollmer, Michaela A. C./Ferguson, Neil M./Okell, Lucy C./Bhatt, Samir (2020): Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* (<https://doi.org/10.1038/s41586-020-2405-7>).

- Foundation for Innovative New Diagnostics (FIND) (03.07.2020): SARS-CoV-2 molecular assay evaluation: results. www.finddx.org/wp-content/uploads/2020/07/FIND_SARS-COV2_molecular-assay-evaluation-results_03Jul2020.pdf, abgerufen am 14.08.2020.
- Fraser, Christophe/Cummings, Derek A. T./Klinkenberg, Don/Burke, Donald S./Ferguson, Neil M. (2011): Influenza Transmission in Households During the 1918 Pandemic. *Journal of Epidemiology*, 174, S. 505–514 (<https://doi.org/10.1093/aje/kwr122>).
- Furuse, Yuki/Sando, Eiichiro/Tsuchiya, Naho/Miyahara, Reiko/Yasuda, Ikkoh/Ko, Yura K./Saito, Mayuko/Morimoto, Konosuke/Imamura, Takeaki/Shobugawa, Yugo/Nagata, Shohei/Jindai, Kazuaki/Imamura, Tadatsugu/Sunagawa, Tomimasa/Suzuki, Motoi/Nishiura, Hiroshi/Oshitani, Hitoshi (2020): Clusters of Coronavirus Disease in Communities, Japan, January–April 2020. *Emerging Infectious Diseases*. wwwnc.cdc.gov/eid/article/26/9/20-2272_article (<https://doi.org/10.3201/eid2609.202272>).
- Gigerenzer, Gerd (2009): Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken, Berliner Taschenbuchverlag.
- Gigerenzer, Gerd/Gaissmaier, Wolfgang/Kurz-Milcke, Elke/Schwartz, Lisa M./Woloshin, Steven (2009): Glaub keiner Statistik, die du nicht verstanden hast. *Gehirn & Geist*, 10, S. 34–39 (www.spektrum.de/magazin/glaub-keiner-statistik-die-du-nicht-verstanden-hast/1006395).
- Heesterbeek, J. A. P. (2002): A Brief History of R_0 and a Recipe for its Calculation. *Acta Biotheoretica*, 50, S. 189–204 (<https://10.1023/a:1016599411804>).
- Horvath, Karl/Semlitsch, Thomas/Jeitler, Klaus/Krause, Robert/Siebenhofer, Andrea (2020): Antikörpertests bei COVID-19 – Was uns die Ergebnisse sagen. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* (<https://doi.org/10.1016/j.zefq.2020.05.005>).
- Istituto Nazionale di Statistica/Ministero della Salute (Istat/Ministero della Salute 2020): Primi risultati dell'indagine di sieroprevalenza sul SARS-COV-2. Veröffentlicht am 3. August 2020. www.istat.it/it/files//2020/08/ReportPrimiRisultati-IndagineSiero.pdf, abgerufen am 07.08.2020.
- Kain, Morgan P./Childs Marissa L./Becker, Alexander D./Mordecai, Erin A. (2020): Chopping the tail: how preventing super-spreading can help to maintain COVID-19 control. www.ncbi.nlm.nih.gov/pmc/articles/PMC7340192/, abgerufen am 16.08.2020 ([dx.doi.org/10.1101/2020.06.30.20143115](https://doi.org/10.1101/2020.06.30.20143115)).
- Kamo, Masashi/Sasaki, Akira (2002): The effect of cross-immunity and seasonal forcing in a multi-strain epidemic model. *Physica D*, 165, S. 228–241 ([https://doi.org/10.1016/S0167-2789\(02\)00389-5](https://doi.org/10.1016/S0167-2789(02)00389-5)).
- Kucirka, Lauren M./Lauer, Stephan A./Laeyendecker, Oliver/Boon, Denali/Lessler, Justin (2020): Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction – Based SARS-CoV-2 Tests by Time Since Exposure. *Annals of Internal Medicine* (<https://doi.org/10.7326/M20-1495>).
- Küchenhoff, Helmut/Günther, Felix/Bender, Andreas/Höhle, Michael (2020): Analyse der Epidemischen Covid-19 Kurve in Bayern durch Regressionsmodelle mit Bruchpunkten. www.stablab.stat.uni-muenchen.de/_assets/docs/analyse_covid19_bayern.pdf, abgerufen am 03.08.2020.
- Landesamt für Gesundheit und Lebensmittelsicherheit (LGL 2020): Übersicht der Fallzahlen von Coronavirusinfektionen in Bayern. www.lgl.bayern.de/gesundheits/infektionsschutz/infektionskrankheiten_a_z/coronavirus/karte_coronavirus/index.htm, abgerufen am 24.07.2020.
- Lipsitch, Marc/Cohen, Ted/Cooper, Ben/Robins James M./Ma, Stefan/Lyn, James/Gowri, Gopalakrishna/Chew, SuokKai/Tan, Chorh C./Samore, Matthew H./Fisman, David/

- Murray, Megan (2003). Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. *Science*, 300, S. 1966–1970 (<https://doi.org/10.1126/science.1086616>).
- Lloyd-Smith, James O./Schreiber, Sebastian J./Kopp, P. Ekkehard/Getz, Wayne M. (2005): Superspreading and the effect of individual variation on disease emergence. *Nature*, 438, 355–359 (<https://doi.org/10.1038/nature04153>).
- LMU Klinikum München (LMU 2020): Neuartiges Coronavirus (SARS-CoV-2): Aktivitäten des Tropeninstituts zur Bewältigung der COVID-19 Pandemie: www.klinikum.uni-muenchen.de/Abteilung-fuer-Infektions-und-Tropenmedizin/de/COVID-19/KoCo19/FAQ/index.html, abgerufen am 29.06.2020.
- Mateus, Jose/Grifoni, Alba/Tarke, Alison/Sidney, John/Ramirez, Sydney I./Dan, Jennifer M./Burger, Zoe C./Rawlings, Stephen A./Smith, Davey M./Phillips, Elizabeth/Mallal, Simon/Lammers, Marshall/Rubiro, Paul/Quiambao, Lorenzo/Sutherland, Aaron/Dawen Yu, Esther/da Silva Antunes, Ricardo/Greenbaum, Jason/Frazier, April/Markmann, Alena J./Premkumar, Lakshmanane/de Silva, Aravinda/Peters, Bjoern/Crotty, Shane/Sette, Alessandro/Weiskopf, Daniela (2020). Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science*. (2020). Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science*, eabd3871 ([dx.doi.org/10.1126/science.abd3871](https://doi.org/10.1126/science.abd3871)).
- Max-Planck-Gesellschaft (2020): Die Effekte der Corona-Maßnahmen. www.mpg.de/14836584/was-bringen-die-massnahmen, abgerufen am 29.06.2020.
- Nishiura, Hiroshi/Chowell, Gerardo (2009): The Effective Reproduction Number as a Prelude to Statistical Estimation of Time-Dependent Epidemic Trends. In: Chowell G., Hyman J.M., Bettencourt L.M.A., Castillo-Chavez C. (Hrsg.), *Mathematical and Statistical Estimation Approaches in Epidemiology*. Springer, Dordrecht, 2009. S. 103–121 (https://doi.org/10.1007/978-90-481-2313-1_5).
- Poethko-Müller, Christina/Prütz, Franziska/Buttmann-Schweiger, Nina/Fiebig, Julia/Sarganas, Giselle/Seeling, Stefanie/Thamm, Roma/Baumann, Jan/Hamouda, Osamah/Offergeld, Ruth/Schaade, Lars/Lampert/Thomas/Neuhauser/Hannelore (2020): Studien zur Seroprävalenz von SARS-CoV-2 in Deutschland und international. *Journal of Health Monitoring* 5 (S4): 2–16 (<https://doi.org/10.25646/7023>).
- Porta, Miquel (2008): *Dictionary of Epidemiology*. New York: Oxford University Press.
- Quanta Magazine (30.06.2020): The Tricky Math of Herd Immunity for COVID-19: www.quantamagazine.org/the-tricky-math-of-covid-19-herd-immunity-20200630/, abgerufen am 04.07.2020.
- Robert Koch-Institut (RKI 2015): *RKI-Fachwörterbuch Infektionsschutz und Infektionsepidemiologie*. Berlin: RKI.
- Robert Koch-Institut (RKI 2020a): SARS-CoV-2 Steckbrief zur Coronavirus-Krankheit-2019 (COVID-19). www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html#doc13776792bodyText6, abgerufen am 16.06.2020.
- Robert Koch-Institut (RKI 2020b): Hinweise zur Testung von Patienten auf Infektion mit dem neuartigen Coronavirus SARS-CoV-2: www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Vorl_Testung_nCoV.html, abgerufen am 07.07.2020.
- Robert Koch-Institut (RKI 2020c): Pressemitteilung vom 09.04.2020. Wie viele Menschen sind immun gegen das neue Coronavirus? Robert Koch-Institut startet bundesweite Antikörper-Studien: www.rki.de/DE/Content/Service/Presse/Pressemitteilungen/2020/05_2020.html, abgerufen am 29.06.2020.

- Robert Koch-Institut (RKI 2020d): COVID-19-Lagebericht vom 02.09.2020. www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Sept_2020/2020-09-02-de.pdf?__blob=publicationFile, abgerufen am 04.09.2020.
- Robert Koch-Institut (RKI 2020e): COVID-19-Lagebericht vom 31.05.2020. www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-05-31-de.pdf?__blob=publicationFile, abgerufen am 09.07.2020.
- Robert Koch-Institut (RKI 2020f): Erfassung der SARS-CoV-2-Testzahlen in Deutschland (Update vom 06.08.2020). *Epidemiologisches Bulletin*, 32/33, S. 15–16 (<https://doi.org/10.25646/7075>).
- Robert Koch-Institut (RKI 2020g): COVID-19-Lagebericht vom 14.08.2020. www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-08-14-de.pdf?__blob=publicationFile, abgerufen am 14.08.2020.
- Robert Koch-Institut (RKI 2020h): Falldefinitionen des Robert Koch-Instituts zur Übermittlung von Erkrankungs- oder Todesfällen und Nachweisen von Krankheitserregern. Stand 29.05.2020. www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Falldefinition.pdf, abgerufen am 29.07.2020.
- Robert Koch-Institut (RKI 2020i): Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting. *Epidemiologisches Bulletin*, 17 (<https://doi.org/10.25646/6692>).
- Robert Koch-Institut (2020j): Bericht zur Optimierung der Laborkapazitäten zum direkten und indirekten Nachweis von SARS-CoV-2 im Rahmen der Steuerung von Maßnahmen, Berlin 07.07.2020.
- Robert Koch-Institut (RKI 2020k): COVID-19-Lagebericht vom 21.06.2020. www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-06-21-de.pdf?__blob=publicationFile, abgerufen am 09.08.2020.
- Robert Koch-Institut (RKI 2020l): Corona-Monitoring lokal. Erste Eckdaten für Kupferzell. Stand 14.08.2020. www.rki.de/DE/Content/Gesundheitsmonitoring/Studien/cml-studie/Factsheet_Kupferzell.html, abgerufen am 17.08.2020.
- Saiki, R.K./Gelfand, D.H./Stoffel, S./Scharf, S.J./Higuchi, R./Horn, G.T./Mullis, K.B./Erllich, H.A. (1988): Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239, S. 487–491 (<https://doi.org/10.1126/science.239.4839.487>).
- Schlenger, Ralf (2020): PCR-Tests auf SARS-CoV-2: Ergebnisse richtig interpretieren. In: *Deutsches Ärzteblatt* 2020, 117 (24): S. A-1194–1195: www.aerzteblatt.de/archiv/214370/PCR-Tests-auf-SARS-CoV-2-Ergebnisse-richtig-interpretieren, abgerufen am 08.07.2020.
- Seifried, Janna/Hamouda, Osamah (2020): Erfassung der SARS-CoV-2-Testzahlen in Deutschland. *Epidemiologisches Bulletin*, 15, S. 3–4 (<https://doi.org/10.25646/6634.2>).
- Solbach, Werner/Schiffner, Julia/Backhaus, Insa/Burger, David/Staiger, Ralf/Tiemer, Bettina/Bobrowski, Andreas/Hutchings, Timothy/Mischnik, Alexander (2020): Antibody profiling of COVID-19 patients in an urban low-incidence region in Northern Germany (<https://doi.org/10.1101/2020.05.30.20111393>).
- Spiegel (27.06.2009): „Die Pest“: Wie der Schwarze Tod nach Oberammergau kam: www.spiegel.de/kultur/gesellschaft/die-pest-wieder-schwarze-tod-nach-oberammergau-kam-a-632622.html, abgerufen am 08.07.2020.

- Spiegel (04.07.2020): Sind wir der Herdenimmunität schon näher? www.spiegel.de/wissenschaft/medizin/coronavirus-sind-wir-der-herdenimmunitaet-schon-naeher-a-c3e9d501-7c97-4436-9342-ae38360b9b2d, abgerufen am 04.07.2020.
- Streeck, Hendrik/Schulte, Bianca/Kuemmerer, Beate/Richter, Enrico/Hoeller, Tobias/Fuhrmann, Christine/Bartok, Eva/Dolscheid, Ramona/Berger, Moritz/Wessendorf, Lukas/Eschbach-Bludau Monika/Kellings, Angelika/Schwaiger, Astrid/Coenen, Martin/Hoffmann, Per/Noethen, Markus/Eis-Huebinger, Anna-Maria/Exner, Martin/Schmithausen, Ricarda/Schmid, Matthias/Hartmann, Gunther (2020): Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event (<https://doi.org/10.1101/2020.05.04.20090076>).
- Takahashi, Saki/Greenhouse, Bryan/Rodríguez-Barraquer (2020): Are SARS-CoV-2 seroprevalence estimates biased? OSF Preprints (<https://doi.org/10.31219/osf.io/y3fxt>).
- Verity, Robert/Okell, Lucy C./Dorigatti, Ilaria/Winskill, Peter/Whittaker, Charles/Imai, Natsuko/Cuomo-Dannenburg, Gina/Thompson, Hayley/Walker, Patrick G. T./Fu, Han/Dighe, Amy/Griffin, Jamie T./Baguelin, Marc/Bhatia, Sangeeta/Boonyasiri, Adhiratha/Cori, Anne/Cucinubá, Zulma/John, Rich Fitz/Gaythorpe, Katy/Green, Will/Hamlet, Arran/Hinsley, Wes/Laydon, Daniel/Nedjati-Gilani, Gemma/Riley, Steven/van Elsland, Sabine/Volz, Erik/Wang, Haowei/Wang, Yuanrong/Xi, Xiaoyue/Donnelly, Christl A./Ghani, Azra C./Ferguson, Neil M. (2020): Estimates of the severity of coronavirus disease 2019: a model-based analysis. In: *The Lancet Infectious Diseases* (20), S. 669–77 ([https://doi.org/10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7)).
- Woloshin, Steven/Patel, Neeraj/Kesselheim, Aaron (2020): False Negative Tests for SARS-CoV-2 Infection – Challenges and Implications. *New England Journal of Medicine*; 383:e38 (<https://doi.org/10.1056/NEJMp2015897>).
- Woolhouse, Mark J./Dye, Christopher/Etard, Jean-Francois/Smith Thomas/Charlwood Jacques D./Garnett, G. P./Hagan, P./Hii, Jeffrey L./Ndhlovu, P. D./Quinnell, Rupert J./Watts, C. H./Chandiwana, Stephen K./Anderson Roy M. (1997): Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proceedings of the National Academy of Sciences of the United States of America*, 94, S. 338–342 (<https://doi.org/10.1073/pnas.94.1.338>).
- World Health Organization (WHO 2020): Public health criteria to adjust public health and social measures in the context of COVID-19. Annex to Considerations in adjusting public health and social measures in the context of COVID-19. Veröffentlicht am 12.05.2020. www.who.int/publications/i/item/public-health-criteria-to-adjust-public-health-and-social-measures-in-the-context-of-covid-19, abgerufen am 07.08.2020.
- Zeit Online (19.03.2020): Oberammergauer Passionsspiele: Not und Spiele. www.zeit.de/gesellschaft/2020-03/oberammergauer-passionsspiele-coronavirus-pest-epidemie, abgerufen am 08.07.2020.
- Zeit Online (23.07.2020): Mitarbeiter bei Tönnies war offenbar Superspreader. www.zeit.de/wissen/gesundheit/2020-07/studie-toennies-coronavirus-superspreader-arbeitsbedingungen, abgerufen am 24.07.2020.