

Medizinische Fakultät  
der  
Universität Duisburg-Essen

Aus dem Institut für Medizinische Informatik, Biometrie und Epidemiologie (IMIBE)

Epidemiologische Charakteristika von Epi- und Pandemien

I n a u g u r a l d i s s e r t a t i o n  
zur  
Erlangung des Doktorgrades der Medizinwissenschaft  
Durch die Medizinische Fakultät  
Der Universität Duisburg-Essen

Vorgelegt von  
Fabian Tobias Eberhard, Standl  
aus München

2023

# DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT  
DUISBURG  
ESSEN

*Offen im Denken*

ub | universitäts  
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

**DOI:** 10.17185/duepublico/79244

**URN:** urn:nbn:de:hbz:465-20231212-074122-7

Alle Rechte vorbehalten.

Dekan: Herr Univ.-Prof. Dr. med. J. Buer

1. Gutachter/in: Herr Univ.-Prof. Dr. med. A. Stang

2. Gutachter/in: Herr Univ.-Prof. Dr. rer. nat. M. Trilling

Tag der mündlichen Prüfung: 17. Oktober 2023

## Publikationen, Vorträge und Preise

### Publikationen:

Standl F, Jöckel K-H, Kowall B, Schmidt B, Stang A. Subsequent waves of viral pandemics, a hint for the future course of the SARS-CoV-2 pandemic. medRxiv. 2020:2020.07.10.20150698.

### Vorträge und Preise:

Standl F. Prediction of the timing of the second wave peak in SARS-CoV-2 using the “Duty Cycle” values of historic (global) epidemics. 20. Tag der Forschung der Medizinischen Fakultät der Universität Duisburg-Essen. Präsentationspreis der Medizinischen Fakultät.

| Inhalt   | Seite |
|--|-------|
| A. Einleitung  | 6     |
| I.    Epidemie und Pandemie  | 6     |
| II.   Epidemische Kurve und epidemische Welle  | 8     |
| III.  Forschungsstand und historischer Hintergrund   | 9     |
| 1.  Forschungsstand der historischen Epidemiologie   | 10    |
| 2.  Forschungsstand zu epidemischen Wellen in viralen Pandemien                                    | 12    |
| 3.  Bestehende Influenzadatenbanken  | 12    |
| 4.  Physikalische Analogien in der Epidemiologie: Duty Cycle                                       | 13    |
| B. Fragestellungen   | 17    |
| C. Material und Methoden   | 17    |
| I.    Strukturierte Literaturrecherche zu Influenzapandemien mittels PubMed                        | 17    |
| II.   INZDS Daten zu COV1, COV2 und Ebola  | 19    |
| III.  Interraterreliabilität bei epidemischen Wellen   | 20    |
| IV.   Aufbereitung und Auswertung INZDS  | 21    |
| V.    Vorhersage des Zeitpunkts des zweiten Wellengipfels in COV2 auf Basis historischer Epidemien | 23    |
| 1.  Vorbereitung der Datensätze  | 23    |
| 2.  Gleichungssysteme  | 25    |
| 3.  Güte der Schätzungen   | 28    |
| VI.   Passende historische Epidemie  | 29    |
| D. Ergebnisse  | 30    |
| I.    Ergebnisse der Literaturrecherchen   | 30    |
| II.   Interratervergleich  | 30    |
| III.  Häufigkeit von Inzidenzwellen, Duty Cycle und Wellenhöhenprofile                             | 31    |
| 1.  Häufigkeit von Folgewellen   | 31    |
| 2.  Duty Cycle in den verschiedenen Epidemien  | 36    |
| 3.  Wellenhöhen  | 40    |

| Inhalt  | Seite |
|---|-------|
| IV. Vorhersage des zweiten Wellengipfels: Anteile der Streckenbezeichnungen, Güte und Präzision der Schätzung | 40    |
| 1. $R^2$ Adj. als Gütekriterium der Schätzung   | 41    |
| 2. Die Absolute Abweichung als Gütekriterium der Schätzung  | 42    |
| V. Passende historische Epidemie anhand quantitativer Merkmale  | 50    |
| E. Diskussion   | 54    |
| I. Zusammenfassung der Ergebnisse   | 54    |
| II. Vergleich zu anderen Untersuchungen   | 56    |
| III. Stärken und Limitationen   | 59    |
| IV. Implikationen   | 61    |
| V. Unbeantwortete Fragen und zukünftiger Forschungsbedarf   | 62    |
| F. Zusammenfassung  | 63    |
| G. Summary  | 64    |
| H. Literaturverzeichnis   | 65    |
| I. Anhang   | 74    |
| I. Tabellenverzeichnis  | 74    |
| II. Abbildungsverzeichnis   | 74    |
| III. Abkürzungsverzeichnis  | 75    |
| J. Danksagung   | 77    |
| K. Lebenslauf   | 78    |

## A. Einleitung

Das bisherige Bild von Epidemien kann inhomogen erscheinen, da es eine unterschiedliche Anzahl an Wellen gab, die Dauer einer Epidemie von Monaten bis Jahren reichte, einzelne Wellen Wochen oder Monate dauerten, wenige tausend Menschen erkrankten, etwa bei der ersten Coronapandemie von 2002/2003 (COV1), oder gar 1/18 der Weltbevölkerung verstarb, wie während der Spanischen Grippe (SG) (Johnson and Mueller, 2002; Roser Max, 2020; World Health Organization). Zudem gibt es mehrere, z. T. widersprüchliche Definitionen.

### I. Epidemie und Pandemie

Im Jahr 2008 veröffentlichte die „World Health Organization“ (WHO) einen neuen Entwurf des „Pandemic Influenza Preparedness Plan“. Während die Phasen eins und zwei ein zirkulierendes Virus im Tierreich thematisierten, so wurden in den Phasen drei und vier noch weitestgehend sporadische Infektionen von Menschen mit tierischen Viren beschrieben (World Health Organization (WHO), 2009). Wichtig für den Kontext dieser Promotionsarbeit sind die Phasen fünf und sechs, denn diese definierten Pandemien entweder als Übertragung von Mensch zu Mensch in zwei oder mehreren nichtbenachbarten Ländern in einer geographischen Region oder als Erkrankungskuster in mindestens zwei der geographischen Regionen (Afrika, Asien, Europa, Amerika und Ozeanien) (Center for Infectious Disease Research and Policy, 2008).

Laut „Centers for Disease Control and Prevention“ (CDC) befand man sich ab dem 11.06.2009 in einer durch das H1N1 Influenzavirus verursachten und durch die WHO erklärten Pandemie, bzw. in der sechsten Phase der Pandemiebeschreibung. Die CDC waren der Auffassung, dass die Ausrufung einer Pandemie nicht den Schweregrad einer Epidemie oder eine Veränderung des Virusverhaltens charakterisierte, sondern bedeutete, dass sich das Virus in mehreren Teilen der Welt ausbreitete (Centers for Disease Control and Prevention, 2009). 2011 stellte das „Fineberg Komitee“, das die Geschehnisse um die Influenzapandemie von 2009 aufarbeitete, fest, dass die Pandemiedefinition der WHO zwar den Grad der Verbreitung berücksichtigte, aber keine Merkmale, die die Schwere des epidemischen Geschehens einordneten. Es wurde empfohlen, adäquate Maße hierfür zu entwickeln und anzuwenden (European Centre for Disease Prevention and Control, 2011).

Am 11.03.2020 erklärte die WHO COVID-19 zur Pandemie (COV2), da in den zwei Wochen zuvor die Anzahl an COV2 Fällen außerhalb Chinas um das dreizehnfache gestiegen war, mehr als 118.000 Fälle in 114 Ländern gezählt worden waren, 4.291 Menschen ihr Leben verloren hatten und es viele infizierte hospitalisierte Personen gab. Man vertrat die Auffassung, dass die Zahl an Erkrankungen, Todesfällen und betroffenen Ländern weiter ansteigen würde und man deshalb alarmiert sei. Aus diesen Gründen kam man zu dem Schluss, dass COV2 als Pandemie charakterisiert werden konnte (World Health Organization, 2020).

Zur Definition von Pandemien verwendeten Potter und Jennings historische Aufzeichnungen. Eine Pandemie zeichnete sich hier dadurch aus, dass sie einen eindeutigen geografischen Ausgangspunkt hatte und dann eine große Zahl von Menschen in vielen Ländern infizierte. Ein weiteres Kriterium war, dass sich das Pandemievirus von dem Virus der saisonalen Grippe unterschied. Zwischen 1889 und 1968 unterschieden sich die Virussubtypen in vier von fünf Pandemien. Bislang waren ausschließlich die Subtypen H1, H2 und H3 als Verursacher von Pandemien bekannt. Zudem war für die Definition einer Pandemie eine hohe Zahl von Infizierten erforderlich. Dieser Punkt war strittig, da seit 1968 Influenzaausbrüche als Pandemien bezeichnet wurden, ohne dass weite Teile der Bevölkerung betroffen waren (Potter and Jennings, 2011).

Singer et al. vertraten die Ansicht, dass dem Begriff Pandemie eine quantitative Definition fehlte, und es sich, gerade in der frühen Phase eines Ausbruchs, um ein rhetorisches Stilmittel handelte, das von Interessensgruppen verwendet wurde, um große Public Health Maßnahmen anzustoßen. Die Autoren verdeutlichten diese These am Beispiel der pandemischen Influenza von 2009, an der weniger Menschen als an der saisonalen Influenza im gleichen Zeitraum verstorben waren. Dennoch, so die Autoren, erfüllte die Bezeichnung als Pandemie ihren Zweck, nämlich den der großen medialen Aufmerksamkeit. Prinzipiell, so die Autoren, war eine Pandemie eine weltweit oder in zahlreichen Gebieten auftretende, staatliche Grenzen überwindende und einen Großteil der Bevölkerung betreffende Epidemie. Die WHO, so die Autoren, verstand unter einer Pandemie eine sich weltweit ausbreitende neue Krankheit. Singer und Kollegen folgerten, dass alle Pandemiemerkmale qualitativ (Singer et al., 2021) und damit in einem gewissen Maße anfällig für Willkür waren. Eine Pandemie ist also in dem hier verwendeten Kontext das

zeitgleiche, an verschiedenen, nichtbenachbarten Orten der Welt, Auftreten einer Virus-epidemie, üblicherweise der Influenza (Doshi, 2011; European Centre for Disease Prevention and Control, 2021; Kelly, 2011).

## II. Epidemische Kurve und epidemische Welle

Die Begriffe „epidemische Kurve“ (Tsang et al., 2020) und „epidemische Welle“ (Laxminarayan et al., 2021) wurden häufig synonym verwendet. Suchte man am 29.09.2021 nach *“epidemic wave“ AND definition* auf „PubMed“, einer durch das „National Center for Biotechnology Information“ und die „U.S. National Library of Medicine“ entwickelten und durch die „National Institutes of Health“ zur freien Verfügung gestellten Ressource zur Suche von biomedizinischer und gesundheitswissenschaftlicher Literatur (National Library of Medicine, N.D.), so erhielt man sieben Suchergebnisse. Eines davon war eine Arbeit von Zhang et al., die eine Definition der epidemischen Welle anbot. Die Autoren stellten fest, dass der Wellenbegriff eine in der Epidemiologie verwendete Metapher ohne präzise Definition war. Die WHO vertrat die Ansicht, dass eine Welle endete, wenn das Virus unter Kontrolle war und die Fallzahlen substantiell zurückgegangen waren; eine zweite Welle zeichnete sich durch einen neuen, anhaltenden Anstieg der Fallzahlen aus. Eine Welle war abzugrenzen von einem „Zacken“, also einem kurzzeitigen, hohen Anstieg und dann ähnlich schnell und stark absinkenden Fallzahlen. Eine Welle zeichnete sich also durch zwei Charakteristika aus:

1. einer Phase des Anstiegs und einer Phase des Absinkens der Fallzahlen und
  2. die Zeiträume des Anstiegs und des Absinkens waren „nachhaltiger“ und unterschieden sich von dem Bild des Zackens, der aber Teil einer Welle sein konnte (Zhang et al., 2021).
- Am 29.09.2021 erhielt man 26 Suchergebnisse für *“epidemic curve“ AND definition* auf PubMed. Keines dieser Ergebnisse versuchte den Begriff ähnlich grundsätzlich zu definieren wie der Beitrag von Zhang et al. (Zhang et al., 2021). Die CDC aus den USA definierten die epidemische Kurve als Darstellung des voranschreitenden Krankheitsausbruchs über die Zeit (Centers for Disease Control and Prevention, 2015). Jefferson und Heneghan aus Großbritannien verwendeten für die Beschreibung des gleichen Phänomens den Begriff der Welle (Jefferson and Heneghan, 2020). Torok meinte, dass die Begriffe „epidemische Kurve“ und „epidemische Welle“ Synonyme wären und mehrere In-



fektionswellen im Rahmen einer Epidemie auftreten könnten. Der zeitliche Abstand zwischen den Höhepunkten war im Idealbild die Inkubationszeit der Erkrankung, die epidemische Kurve in der Realität wich von dem Idealbild ab (Torok, N.D.).

Epidemische Welle und epidemische Kurve wurden synonym verwendet, die unterschiedliche Bezeichnung war kultureller Natur. Charakterisiert war das Phänomen der epidemischen Welle bzw. Kurve durch einen beständigen, über längere Zeit andauernden Zuwachs an Neuerkrankungen, einen Höhepunkt und eine anschließende, substantielle Abnahme. Stiegen die Neuerkrankungen erneut, so entstand das Bild einer mehrwelligen Epidemie. Jefferson und Heneghan vermuteten, dass mehrwellige Epidemien nur theoretisch vorkamen (Jefferson and Heneghan, 2020).

### III. Forschungsstand und historischer Hintergrund

Es gab mehrere Arbeiten, die aktuelle und historische Pandemien miteinander verglichen. Petersen et al. verglichen COV2 mit COV1 und einigen Influenzapandemien (Petersen et al.; Standl et al., 2021). Dieser historische, epidemiologische Ansatz lieferte eine Einschätzung des Schweregrads von Viruspanidemien auf der Grundlage epidemiologischer Daten aus früheren Pandemien und versuchte, das epidemiologische Wissen über Pandemien zu systematisieren.

In den letzten 130 Jahren wurden sieben Grippepanidemien beobachtet (Amato Gauci et al., 2010). Allerdings gab es wohl bislang keine öffentlich zugängliche Datenbank über Pandemien, die es ermöglichte, internationale vergleichende Analysen über die bekannten Pandemien auf breiter statistischer Basis durchzuführen. Die meisten Arbeiten im Bereich der historischen Epidemiologie verglichen entweder SG auf internationaler Ebene (Ansart et al., 2009) oder mehrere Pandemien in einem bzw. mehreren Ländern (Zürcher et al., 2016). Gemäß der für die vorliegende Arbeit durchgeführten Recherchen gab es bislang keinen veröffentlichten strukturierten quantitativen Vergleich der Epidemien der letzten 130 Jahre auf breiter internationaler Basis.

Zusammenfassend lässt sich feststellen, dass die Erforschung von Pandemien und die Vorhersage des künftigen Verlaufs einer Pandemie auf der Grundlage weniger historischer Daten eine häufig verwendete Methode ist. Es mangelte bisher jedoch an einer gemeinsamen Methodik und einer Datenbank, um eine solide, systematische, statistische

Analyse von Pandemien zu ermöglichen, um die Auswirkungen möglicher verzerrter historischer Bewertungen zu verringern, die bei Einzelbetrachtungen auftreten können und um die Validität von angenommenen Parametern für statistische Modelle zu verbessern.

### 1. Forschungsstand der historischen Epidemiologie

Suchte man am 13.05.2021 nach „*history viral pandemic*“ auf PubMed, konnte man anhand der Titel die häufigsten Themen feststellen: SG, COV2 und Human Immunodeficiency Virus. Ansart und Valleron nannten präzise Zahlen zu den Todesopfern von SG in Europa und den USA (Ansart and Valleron, 2008). Esparza analysierte 300 Jahre pandemischer Grippe und stellt die Verbindung zu COV2 her. COV2 war eine typische, über die Atemwege übertragene Krankheit und daher mit der Influenza vergleichbar. Die zehn Pandemien in dem Beitrag hatten folgende gemeinsame Merkmale: Ursprung i. d. R. in Asien, i. d. R. eine Dauer von mindestens zwei Jahren, mehrere Wellen und die Pandemien endeten, wenn mindestens 30% der Bevölkerung immun waren (Esparza, 2020).

In einer Arbeit von Wu und Cowling wurden mit Hilfe eines SIR-Modells (Susceptible, Infectious, Recovered) Schätzungen über den weiteren Verlauf der „Schweine-Grippe“ (SWG) vorgenommen. Zur Kalibrierung der Modelle wurden epidemiologische Parameter für die Basisreproduktionszahl ( $R_0$ ), also die durchschnittliche Zahl von neuinfizierten Personen pro bereits infizierter Person, und die mittlere Generationszeit auf der Grundlage der Pandemien von 1918, 1957 und 1968 geschätzt (Wu and Cowling, 2011).

Reina verglich die epidemiologischen Merkmale von SG, „Asiatischer Grippe“ (AG), „Hong-Kong Grippe“ (HKG) und „Zweiter Russische Grippe“ (RG2), die die erste H1-Pandemie nach SG war. Darüber hinaus wurden SWG und die H5-Grippe von 1997 berücksichtigt (Reina, 2009).

Um die Lehren aus HKG in Afrika zu analysieren, führten Ortiz et al. ein systematisches Review der Literatur durch, wobei AG und HKG berücksichtigt wurden. Analysiert wurden 14 Studien zu HKG. Die Autoren kamen zu dem Schluss, dass HKG grundlegende Auswirkungen auf die öffentliche Gesundheit in Afrika hatte, obwohl sie weltweit als mild beschrieben wurde. Es wurde empfohlen, SWG in Afrika nicht als mild zu bezeichnen, da nicht genügend Daten für diese Einschätzung vorlagen (Ortiz et al., 2012).

Taubenberger und Morens beschrieben epidemiologische Merkmale von SG, AG, HKG und SWG und erörterten, inwieweit sich diese Pandemien von den Beschreibungen anderer historischer Pandemien unterschieden. Man kam zu dem Schluss, dass es abwegig

war, dass die Vogelgrippeerreger H5 und H7 pandemisch würden, da die letzten fünf Pandemien durch H1-, H2- und H3-Viren verursacht wurden und somit keine Pandemie beim Menschen bekannt war, die durch ein hochpathogenes Vogelgrippevirus verursacht worden war (Taubenberger and Morens, 2010).

Morens und Taubenberger fassten in einer weiteren Arbeit den Stand des Wissens und dessen Unsicherheiten zusammen (Morens and Taubenberger, 2011). Als grundsätzliches Problem der historischen Epidemiologie von Influenzapandemien beschrieben die Autoren den Mangel an verlässlichen historischen Quellen, um Influenzapandemien nach modernen Kriterien zu identifizieren. Eine hinreichende Analyse von Influenzapandemien könne erst seit der Influenzapandemie von 1889 durchgeführt werden, da seither umfangreiche Informationen vorliegen. Bei der globalen Ausbreitung von Influenzapandemien stellte die Pandemie von 1968 einen Wendepunkt dar, da die globale Ausbreitung seither hauptsächlich über den Luftverkehr erfolgte – bis dahin waren die Hauptübertragungswege Handelsrouten und Schifffahrtslinien. Andererseits hatten sich die Risikogruppen, z. B. ältere Menschen und Menschen mit Vorerkrankungen, in den letzten Jahrhunderten nicht verändert (Morens and Taubenberger, 2011).

Barry analysierte die Pandemien von 1889, 1918, 1957 und 1968, um eine Vorhersage für SWG zu treffen. Ein Kriterium für eine Grippepandemie war hier, dass sie etwa 15% bis 40% der Bevölkerung infizierte, während die saisonale Grippe nur etwa 10% der Bevölkerung infizierte. Barry schloss daraus, dass SWG drei der vier Pandemien in Bezug auf die Intensität der Ausbrüche entsprach, die Ausbrüche aber örtlich verstreuter waren. Das Konzept einer möglichen zweiten Welle wurde aus den Pandemien von 1889, 1918 und 1957 abgeleitet (Barry, 2009).

Akin und Gözel analysierten die Dynamik von Pandemien, indem sie die Morbiditäts- und Mortalitätsmerkmale der Pandemien von 1918, 1958, 1968 und 2009 verglichen. Es zeigten sich Ähnlichkeiten mit COV2 in Bezug auf die Ursache der weltweiten Ausbreitung, und dass Menschen mit Vorerkrankungen i. d. R. stärker betroffen waren. Man hielt es für wichtig, dass epidemiologische Merkmale auf der Grundlage historischer Pandemien bewertet würden. Für die Beschreibung des Schweregrades wurde u. a. vorgeschlagen, den Zeitpunkt der Spitzenwerte zu messen (Akin and Gözel, 2020).

## 2. Forschungsstand zu epidemischen Wellen in viralen Pandemien

Suchte man am 13.05.2021 auf PubMed nach „*wave*“, jeweils in Kombination mit „*Russian Flu*“ (14 Suchergebnisse), „*Spanish Flu*“ (76 Suchergebnisse), „*Asian Flu*“ (25 Suchergebnisse), „*Hong Kong Flu*“ (53 Suchergebnisse) und *SARS* (947 Suchergebnisse, davon 353 aus dem Jahr 2020), um Literatur zu epidemischen Wellen in Influenza- und Coronapandemien zu identifizieren, stellte man fest, dass die meisten Artikel eine Pandemie untersuchten.

Xu und Li warnten vor einer zweiten COV2-Welle, wenn die Public Health Interventionen zu früh gelockert würden (Xu and Li, 2020). Neher et al. zeigten saisonale Muster auf, die das Auftreten von Folgewellen beeinflussen könnten. Die Autoren kamen zu dem Schluss, dass eine weitere Welle im Winter 2020/2021 auftreten könnte (Neher et al., 2020). Zur Vorhersage des künftigen Verlaufs einer Influenzapandemie modellierten Shi et al. virale Mutation und Saisonalität und folgerten, dass es für Entscheidungsträger unerlässlich sei, das Risiko einer Pandemie auf der Grundlage der pandemischen Merkmale im Frühstadium zu bewerten. Die Autoren verwiesen auf SG und SWG, um den Bedarf an solchen Modellen zu unterstreichen. Für ihre Analyse verwendeten sie ein SEIR-Modell (Susceptible, Exposed, Infected, Recovered) und verschiedene Saisonalitätsszenarien, die auf der Grundlage zahlreicher Kombinationen von Einflußvariablen zu verschiedenen epidemischen Kurven führten (Shi et al., 2010).

## 3. Bestehende Influenzadatenbanken

Das von der WHO bereitgestellte „FluNet“ lieferte ab 1995 virologische Informationen zur Influenza in zahlreichen Ländern (World Health Organization). Dies bedeutet, dass keine der großen Grippepandemien des 20. Jh. enthalten waren. „NCBI Virus“ bot Datensätze zu viralen Sequenzdaten (National Center for Biotechnology Information, 2021c) und als Unterabschnitt die „Influenza Virus Resource“ (National Center for Biotechnology Information, 2021a). Diese Datenbank enthielt Nukleotid- und Proteinsequenzen der bekannten Grippeviren und umfasste Daten ab 1998 (National Center for Biotechnology Information, 2021a; b). Die „Influenza Research Database“ (Influenza Research Database, 2021; Zhang et al., 2016) unterstützte insbesondere die virologische Erforschung von Erregern, die Infektionskrankheiten auslösen und verursachen (Influenza Research Database). „GISAID“ [Eigename] setzte sich für den schnellen

Austausch von Virusdaten bei Epi- und Pandemien ein. Dabei handelte es sich um genetische Sequenzen, klinische und mikrobiologisch-epidemiologische Daten (Shu and McCauley, 2017). „Our World in Data“ war eine Datenbank, die u. a. Zeitreihendaten zu verschiedenen Themen enthielt, die heruntergeladen und analysiert werden konnten. Man konnte z. B. Veränderungen auf epidemiologischer und gesundheitspolitischer Ebene messen (Our World in Data, 2021). Der Artikel über die globalen Auswirkungen von SG enthielt Daten zur Lebenserwartung in verschiedenen Ländern. Der abgedeckte Zeitraum reichte mindestens von 1918 bis 1920; einige, jährliche Daten reichten jedoch von 2019 bis 1751 zurück (Roser Max, 2020). Die Auswirkungen verschiedener Pandemien auf die Lebenserwartung konnten gemessen und mit anderen Ländern verglichen werden. Dies ermöglichte jedoch nur eine hochaggregierte, jährliche, thematisch stark begrenzte und indirekte Analyse. Eine Analyse der einzelnen Wellen zur Beobachtung bestimmter Muster, die genaue Vorhersagen über den Verlauf von Influenzapandemien ermöglichten, war nicht möglich.

Andere Ansätze zur Verfolgung und Vorhersage der saisonalen Grippe und von Grippeepidemien kombinierten Datenbanken und nutzen „Google Trends“ als zusätzliche Datenquelle. Diese Ansätze nutzten jedoch nur die Informationen eines oder weniger ausgewählter Länder. Darüber hinaus waren die abgedeckten Zeiträume eng und reichten z. B. von 2010 bis 2018 oder von 2003 bis 2015 (Ning et al., 2019; Zhang et al., 2019).

Die bestehenden Influenza-Datenbanken ermöglichten also keine historischen Vergleiche und erfüllten zwar virologische und mikrobiologische Zwecke, aber nur begrenzt epidemiologische Aspekte. Es gab also bisher wohl keine internationale Datenbank zu viralen Epidemien der letzten 130 Jahre, die epidemiologische, epochenübergreifende und globale, vergleichende Forschung ermöglichte.

#### 4. Physikalische Analogien in der Epidemiologie: Duty Cycle

Begriffe wie Welle und Schwingung ebenso wie die zugehörigen Formeln und Theorien wurden häufig im Zusammenhang mit Epidemien verwendet, um diese zu analysieren und zu beschreiben. Grundsätzliche Überlegungen zur erweiterten Verwendung von Diffusionswellen fanden sich etwa bei Mandelis (Mandelis, 2000). Für *“waves diffusion“ AND epidemiology* gab es am 10.10.2021 auf PubMed 102 Einträge. Nachweislichen Eingang in die Epidemiologie fand die Verwendung von Diffusionswellen 2003 (Lin et al., 2003), auch im Rahmen der aktuellen Coronapandemie wurden physikalische Analogien

verwendet (Cacciapaglia et al., 2021). Während manche Arbeiten eher allgemeiner, theoretischer und methodischer Art waren (Cacciapaglia et al., 2021; Huang and Wang, 2019; Pang and Xiao, 2019; Tsai and Weng, 2020; Yang et al., 2011), untersuchten andere Arbeiten konkrete Vorkommnisse (Lewis et al., 2006; Nelson et al., 2011; Steindorf and Maidana, 2019; Tosi and Campi, 2020). Eine dritte Gruppe untersuchte theoretische Maßnahmen in theoretischen Settings (Lupica et al., 2019).

Vereinfacht konnte man eine Epidemie als eine zyklische Sequenz von Neuinfektionen über die Zeit begreifen und darstellen. In der Schwingungslehre bezeichnete man solche Schwingungssequenzen auch als „periodic pulse train“ (Grami, 2016), die sich auch als „triangle wave“ darstellen ließen (Chaparro, 2015).

Der „Duty Cycle“ (DC) beschrieb den beobachtbaren zeitlichen Anteil eines Signals ( $\Delta t$ ) in Relation zu der Zeitspanne, die zwischen zwei aufeinanderfolgenden Zeitpunkten vergangen war. Dies war die Periode ( $T$ ), wobei die Punkte die gleiche Auslenkung und Auslenkungsrichtung hatten, also von gleicher Phase waren. Üblicherweise wurde  $T$  für direkt aufeinanderfolgende Zeitpunkte verwendet (Grami, 2016; Sun et al., 2013; Taylor et al., 2018). Dass die Verwendung des DC in der Epidemiologie u. a. für die Dynamikbeschreibung einer Epidemie sinnvoll war, wurde von Bin et al. gezeigt (Bin et al., 2021). Angewendet auf epidemische Wellen war  $\Delta t$  die Dauer einer epidemischen Welle,  $T$  war die Dauer zwischen zwei aufeinanderfolgenden Zeitpunkten gleicher Phase. Eine epidemische Welle konnte mittels dreier Zeitpunkte beschrieben (Anfang, Höhepunkt und Ende) und als Dreieck dargestellt werden. Ein mehrwelliges, epidemisches Geschehen war als eine zeitliche Abfolge von Dreiecken darstellbar. Der DC konnte also aus dem Quotienten der Dauer einer epidemischen Welle und der Dauer zwischen dem Höhepunkt dieser Welle und dem Höhepunkt der Folgewelle berechnet werden – da die Höhepunkte einfach identifiziert werden konnten. Der mathematische Ausdruck war:

$$DC = \frac{\Delta t}{T} \quad (1)$$

Abbildung 1 war die schematische Darstellung einer zweiwelligen Epidemie. Dargestellt waren die relevanten Zeitpunkte und die Auslenkung (Signalzustand). Die x-Achse zeigte die Zeitpunkte und die y-Achse den Signalzustand. Der Signalzustand  $0$  bedeutete, dass ein Signal noch nicht erkennbar war, oder sich nicht vom Grundrauschen unterschied. Der Signalzustand  $1$  bedeutete, dass das Signal seine maximale Auslenkung in dieser

Welle erreicht hatte. Die Signale  $< \mathbf{0}$  wurden in dieser Betrachtung ignoriert und als Zeitraum zwischen zwei Wellen verstanden. In der Darstellungsform absoluter Fallzahlen gab es schlicht keine sinnvollen negativen Werte – andere Darstellungsformen, wie etwa die prozentuale Abweichung eines Mittelwerts von z. B. täglichen Neuinfektionen könnten aber auch negative Werte sinnvoll abbilden. Ein Signal also, in dem hier verwendeten Kontext, das etwa in Neuerkrankungen pro Tag gemessen wurde, konnte die maximale Auslenkung in einer ersten Infektionswelle bei fünf Neuerkrankungen pro Tag und in der zweiten Welle bei zehn Neuerkrankungen pro Tag haben. Dennoch hatten beide Signale maximal den Signalzustand  $\mathbf{1}$ , da in der jeweiligen Welle die maximale Höhe bzw. Auslenkung erreicht war. Wollte man den Zeitpunkt des zweiten Wellengipfels bei bekanntem DC berechnen, so war  $\Delta t$  die Differenz von Zeitpunkt  $\mathbf{c}$  und  $\mathbf{a}$ ;  $\mathbf{T}$  war die Differenz von Zeitpunkt  $\mathbf{e}$  und  $\mathbf{b}$ . Es ergab sich also die Formel:

$$DC = \frac{(c - a)}{(e - b)} \quad (2)$$

Dieser Ausdruck ließ sich umformen und nach dem gesuchten Zeitpunkt  $\mathbf{e}$  auflösen:

$$e = \frac{(c - a)}{DC} + b \quad (3)$$

Rechenbeispiel:  $DC = 0,5$  und die Epidemie begann im Januar eines Jahres X. Man ordnete also dem Zeitpunkt  $\mathbf{a}$  den Zahlenwert des Monats Januar, also 1 zu. Das Ende der ersten Welle wurde im Juli des gleichen Jahres X beobachtet. Damit war der Zeitpunkt  $\mathbf{c}$  eine 7. Der Höhepunkt der ersten epidemischen Welle, also der Zeitpunkt  $\mathbf{b}$  lag im Mai und hatte damit den Zahlenwert 5. Setzt man die Zahlen ein, so erhält man:

$$e = \frac{(7 - 1)}{0,5} + 5 = 19 \quad (4)$$

Für  $\mathbf{e}$  erhielt man 19, man erwartete also den Gipfel der zweiten Welle im 19. Monat der Epidemie. Da ein Jahr 12 Monate hat, ist der Gipfel also im siebten Monat des Folgejahres von X, also im Juli des Jahres Y erwartet worden. Wenn Epidemien analog zur Schwingungslehre beschrieben werden könnten, so könnte man den DC historischer Epidemien messen. Man konnte erwarten, dass es für verschiedene Erreger typische DC gab, mit denen sich die Zeitpunkte zukünftiger epidemischer Wellen berechnen ließen.

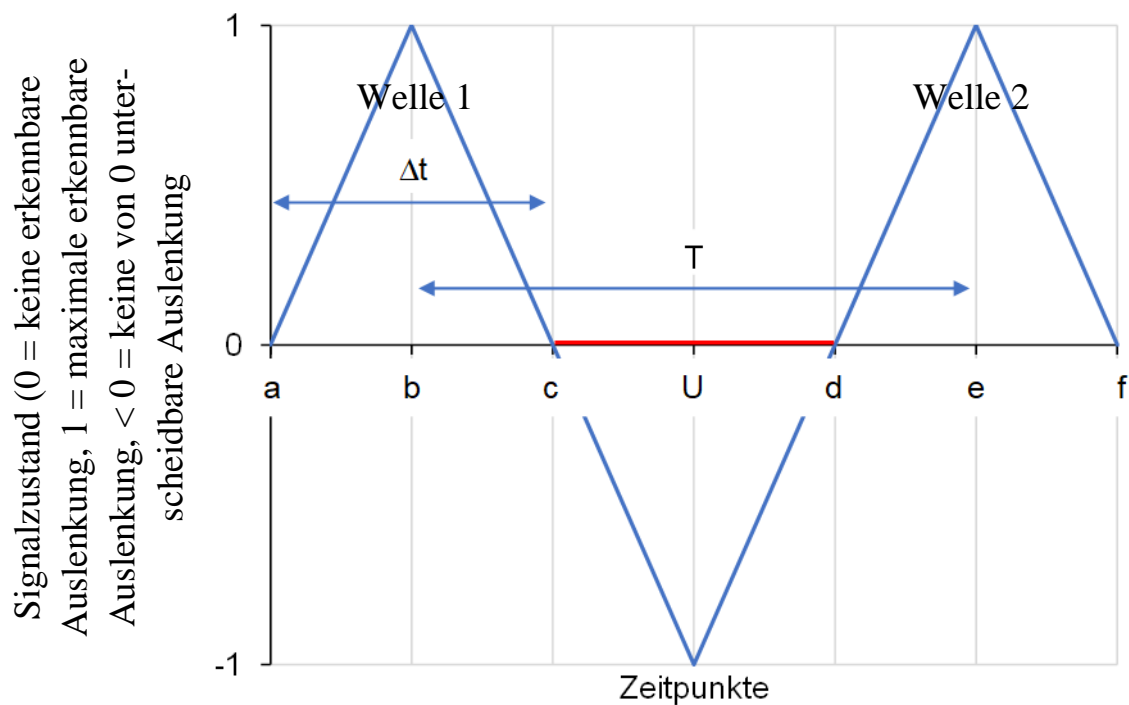


Abbildung 1: Schematische Darstellung einer zweiwelligen Epidemie

Eine epidemische Kurve kann als allgemeines Dreieck dargestellt werden. Eine Epidemie einer Infektionskrankheit ist die Abfolge epidemischer Wellen. Eine mehrwellige Epidemie ist darstellbar als Abfolge von Dreiecken. Dargestellt sind die relevanten Zeitpunkte und der Signalzustand. Die x-Achse zeigt die Zeitpunkte, die y-Achse die Auslenkung (Signalzustand) z.B. Neuerkrankungen pro Tag einer bestimmten Krankheit. Der Signalzustand **0** bedeutet, dass ein Signal noch nicht erkennbar ist. Der Signalzustand **1** bedeutet, dass das Signal hier maximal ist. Negative Signalzustände können bei der Verwendung absoluter Fallzahlen nicht sinnvoll dargestellt werden und werden deshalb behelfsweise durch die rote Linie, also einer in etwa horizontalen Weiterentwicklung der Fallzahlen ersetzt und als die Zeit zwischen zwei epidemischen Kurven verstanden. Damit entfällt auch der Zeitpunkt U an dem die negative Auslenkung maximal wäre.



## B. Fragestellungen

1. Häufigkeit von Inzidenzwellen: Wie häufig traten zweite, dritte, vierte und fünfte Wellen (Folgewellen) auf?
2. Höhe von Inzidenzwellen: Welche Höhen erreichten Folgewellen im Vergleich zu einer ersten Infektionswelle?
3. Vorhersage: Mit welchem Maß, welches sich aus Distanzmaßen von mindestens zweigipfligen epidemischen Kurven historisch ableiten ließ, war der Zeitpunkt des Gipfels der zweiten Welle in der aktuellen COV2 Pandemie (jedwede Länder mit Daten, Einschluss nur der Länder mit vollständigen Daten zu erster und zweiter Welle) am genauesten vorhersagbar?
4. Welche der untersuchten historischen Epidemien beschrieb die aktuelle COV2 Pandemie am besten?

## C. Material und Methoden

Es wurde mit dem Aufbau einer Datenbank begonnen, die die strukturierte, epidemiologische Erforschung von durch Viren ausgelösten und durch Tröpfcheninfektion übertragbaren, lokalen und globalen Epidemien ermöglichte. Die Informationen zu den Influenzapandemien wurden durch eine strukturierte Literaturrecherche gewonnen. Die Daten zu COV1 waren bereits zuvor gesammelt worden und wurden in die entstehende Datenbank aufgenommen. Die Daten zu den Ebolaepidemien wurden im Rahmen des weiteren Aufbaus der Datenbank ermittelt und flossen mit ihrem aktuellen Bearbeitungsstand in die vorliegende Promotionsarbeit ein. Der zu entwickelnde Inzidenzdatensatz (INZDS) umfasste die nachfolgenden Variablen: 1. Informationsquelle, 2. Virustyp, 3. Ort, 4. Startjahr der Epidemie, 5. zeitliche Auflösung, 6. zeitlicher Anfang, Höhepunkt und Ende jeder Welle sowie 7. Höhe jeder Welle, maximal bis einschließlich einer fünften Welle. Da die Definitionen von epidemischen Wellen mehrere Sichtweisen ermöglichten, wurden diese in den INZDS als einzelne Epidemiedatensätze (Perspektiven) aufgenommen.

I. Strukturierte Literaturrecherche zu Influenzapandemien mittels PubMed  
Dieser Abschnitt bezieht sich auf den 23.03.2021. Es gab 139.551 Einträge auf PubMed für den Suchbegriff „*Influenza*“ und 103.800 Einträge für „*Pandemic*“. Für die Kombination der Suchbegriffe „*Influenza AND Pandemic*“ gab es 18.696 Einträge. Durch die Verwendung von „MeSH-Terms“ wurde die Suche auf die beim Menschen auftretende

Influenza und die dazugehörigen Pandemien begrenzt. Die „Entry Terms“ gaben Aufschluss darüber, dass *Influenza* als Suchbegriff sowohl Formulierungen im Singular als auch im Plural abdeckte. Zudem waren auch Synonyme, wie etwa „Flu“ abgedeckt. Gleiches galt auch für den Suchbegriff *Pandemics*. Die MeSH-Term Suche ergab 4.383 Einträge, die auch nicht relevante Ergebnisse enthielt, die keinen Bezug zur Epidemiologie auf Public Health Ebene und Zeitreihenanalyse hatten. Durch eine erneute Ausdifferenzierung der MeSH-Term-Suche wurde schließlich die nachfolgende Sucheingabe erstellt, die 1.417 Einträge ergab:

*("Influenza, Human/analysis"[Mesh] OR "Influenza, Human/classification"[Mesh] OR "Influenza, Human/epidemiology"[Mesh] OR "Influenza, Human/history"[Mesh] OR "Influenza, Human/immunology"[Mesh] OR "Influenza, Human/mortality"[Mesh] OR "Influenza, Human/prevention and control"[Mesh] OR "Influenza, Human/statistics and numerical data"[Mesh] OR "Influenza, Human/virology"[Mesh]) AND ("Pandemics/classification"[Mesh] OR "Pandemics/history"[Mesh] OR "Pandemics/prevention and control"[Mesh] OR "Pandemics/statistics and numerical data"[Mesh]).*

Dieser strukturierte Ansatz wurde abgerundet durch eine Recherche der zentralen, kombinierten Begriffe: „*Influenza AND timeseries*“ (198 Einträge), „*Influenza AND time series*“ (1.020 Einträge) und „*Influenza AND wave*“ (1.023 Einträge). Da nicht alle Suchergebnisse relevant waren und es zahlreiche Dopplungen gab, die entfernt wurden, wurden zunächst die Titel gelesen und jene Artikel ausgeschlossen, die nicht die Influenza zum Thema hatten. Zudem wurden sämtliche Artikel ausgeschlossen, die nur eine spezifische Welle oder die saisonale Influenza betrachteten. Die übrigen Artikel wurden auf Darstellungen überprüft, die es erlaubten, epidemische Wellen auszuwerten. Erneut waren nicht alle Artikel relevant, da keine Darstellungen bzw. nicht verwertbare Abbildungen enthalten waren. Letzteres war der Fall, wenn 1. die Zeitachse zu kurz war, also nicht mehrere Wellen analysiert werden konnten und wenn 2. die Zeitachse oder die Datendarstellung zu ungenau für eine hinreichend präzise Extraktion der abgebildeten Werte war. Zeitachsen wurden als ausreichend präzise gewertet, wenn man einem beliebigen Datenpunkt mindestens einen bestimmten Kalendermonat in einem bestimmten Kalenderjahr zuordnen konnte. Manche Artikel enthielten zusätzliche Informationen, wie Wellen in bestimmten Altersgruppen, die nicht von jeder Infektionswelle betroffen waren. Solche De-

tailanalysen wurden nicht berücksichtigt, da es keine Erscheinungen der gesamten Population waren. Schlussendlich wurden alle Artikel ausgeschlossen, die nicht auf Englisch, Deutsch, Schwedisch oder Spanisch verfasst waren. Die nun übrigen Abbildungen der Zeitreihen wurden heruntergeladen und deren dargestellte Informationen, wenn nötig unter Zuhilfenahme eines Lineals, etwa bei Abwesenheit einer hinreichend präzisen Zeitachse, ausgemessen.

Insgesamt wurden mit diesem Vorgehen 201 Artikel identifiziert, die auswertbare Zeitreihendarstellungen enthalten. Zudem wurden drei weitere Arbeiten eingeschlossen, die bereits für einen älteren Beitrag (Standl et al., 2020) recherchiert worden waren. Aus diesen 204 Arbeiten wurden schließlich jene ausgenommen, die die Zeitreihen nicht in Inzidenz- sondern z. B. in Mortalitätsangaben darstellten, woraufhin schließlich 131 Artikel die Einschlusskriterien erfüllten. Hieraus konnten 730 Perspektiven zu vier Influenzapanemien generiert werden.

## II. INZDS Daten zu COV1, COV2 und Ebola

Die Daten zu COV1 wurden für einen älteren Beitrag (Standl et al., 2020) erhoben und in den INZDS übernommen. Die Daten zu den täglichen, neu bestätigten Erkrankungsfällen von COV2 pro Land, bereitgestellt von der Johns Hopkins Universität (Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE), 2021), wurden heruntergeladen und aufbereitet. Alle Zeitreihen begannen mit dem 22.01.2020 und endeten am 12.04.2021. Pro Tag war pro Land die bis dahin kumulative Anzahl an bestätigten Infektionen angegeben. Um epidemische Wellen pro Land über die Zeit darstellen zu können, wurden die täglichen Neuerkrankungen berechnet, indem von dem jeweiligen Tageswert der Wert des Vortages subtrahiert wurde. Anschließend wurde zur Glättung der Zeitreihen ein gleitender Mittelwert von sieben Tagen verwendet. Von den 194 Orten aus dem ursprünglichen Datensatz wurden zu 191 Orten Zeitreihenabbildungen erzeugt und visuell inspiziert; 190 Darstellungen flossen in den INZDS ein, die Datensammlungen von Schiffen und Ländern mit nicht plausiblen Darstellungen wurden zuvor entfernt und führten zu 356 Perspektiven. Dies stellte eine Neuerung dar, da nicht nur größere Ausbruchsgeschehen untersucht wurden, die eine beliebige Mindestanzahl von gemeldeten Fällen benötigten (N.N., 2021).

Die Daten zu den Ebolaepidemien entstammten mehreren Literaturrecherchen, die zwischen dem 15.12.2020 und dem 25.01.2021 mittels PubMed durchgeführt wurden. Untersucht wurden die Begriffe „*ebola AND timeseries*“ (13 Einträge, keiner verwendbar), „*ebola AND time AND series*“ (66 Einträge, fünf verwendbar), „*ebola AND wave*“ (22 Resultate, drei verwendbar) sowie „*ebola AND pandemic*“ (637 Einträge, fünf verwendbar). Nach der Entfernung von Dopplungen konnten aus 13 Arbeiten 36 Perspektiven generiert werden. Insgesamt enthielt der INZDS 1.141 Perspektiven.

### III. Interraterreliabilität bei epidemischen Wellen

Da die Definitionen epidemischer Wellen es nicht erlaubten, eine Perspektive als einzig richtige Variante zu identifizieren, ergab sich die Notwendigkeit der Überprüfung, inwieweit verschiedene Personen bei der Interpretation der gleichen Darstellungen übereinstimmten. Hierzu wurden die beiden Betreuer dieser Doktorarbeit und der Promovend als Rater bestimmt. Der Promovend erstellte ein Dokument mit 30 zufällig ausgewählten Darstellungen epidemischer Wellen aus der Fachliteratur und eigenen Grafiken zu COV2, wie oben beschrieben, die zu beurteilen waren. Die 30 Darstellungen stammten aus einem Dokument mit 413 Darstellungen epidemischer Kurven von ILI („Influenza-like Illness“) Epidemien, deren Perspektiven bis zum Stichtag (19.03.2021) noch nicht in den INZDS eingegangen waren. Die 30 Darstellungen wurden durch die Generierung von Zufallszahlen ermittelt. Für die Generierung der Zufallszahlen galt: Die Zahlen konnten einen Wert zwischen 1 und 413 haben, Dopplungen waren ausgeschlossen. Die Rater sollten die Darstellungen ohne Absprache mit anderen interpretieren und ihre Einschätzung schriftlich festhalten. Bei jeder Grafik gab es zwei wesentliche Aufgaben: 1. Die Anzahl der epidemischen Wellen zu identifizieren und festzuhalten sowie 2. den jeweils zeitlichen Beginn, Höhe- und Endpunkt jeder Welle kenntlich zu machen. Damit alle Rater ein ähnliches Verständnis von epidemischen Wellen hatten, sollte jeder zuvor den Passus zu epidemischen Wellen der vorliegenden Arbeit lesen.

Die Sichtweisen der drei Rater pro Grafik wurden in einer Tabelle gesammelt. Erfasst sind die gezählten Wellen pro Rater und Grafik, sowie die jeweils korrespondierenden, identifizierten Wellenzeitpunkte (Anfang, Höhepunkt und Ende). Es wurde:

1. Die von jedem Rater identifizierte Wellenanzahl pro Grafik als kategoriale Variable behandelt. Die Übereinstimmung der Rater wurde mittels Fleiss' Kappa ermittelt und das Ergebnis gemäß der Einordnung von Landis und Koch interpretiert (Fleiss, 1971).

2. Um zu ermitteln, wie stark die Sichtweisen der Rater übereinstimmen, wurde, wenn möglich pro Grafik und Rater, der DC der Welle Eins-Zwei ermittelt. Mittels linearer Regression wurde dann untersucht, wie sehr die DC der Rater übereinstimmen – der Zusammenhang der raterspezifischen DC-Werte, also der ermittelte DC bei einem Rater A und der ermittelte DC eines anderen Raters B bei der gleichen Abbildung, wurde mittels für den Stichprobenumfang adjustiertem  $R^2$  ( $R^2$  Adj.) ausgedrückt. Es galt also:

$$DC_A = a + b DC_B \quad (5)$$

#### IV. Aufbereitung und Auswertung INZDS

Der INZDS beinhaltete drei Arten von Dubletten:

1. Perspektiven des gleichen Ereignisses, etwa Inzidenz der Aufnahmen in einem Krankenhaus während SWG und dazu die Inzidenz der virologisch positiven Befunde,
2. Dubletten, die sich auf das gleiche Land bzw. die gleiche Zeit beziehen, etwa SG in Italien und SG in Rom, und
3. Dubletten des gleichen Ereignisses aus verschiedenen Quellen, etwa die Inzidenz von AG in Rom bei Autor A und B. Da die Dubletten nicht statistisch unabhängig waren, wurde beschlossen, dass pro Epidemie ein Land einmal vertreten sein durfte. Dies bedeutete, gab es für SG eine nationale Betrachtung für Italien und weitere lokale Epidemiedatensätze, so wurde nur ein, von einer dritten Person ohne Projektbeteiligung zufällig ausgewählter Epidemiedatensatz verwendet. Zudem schien es sinnvoll zu ermitteln, ob die Dynamiken auf lokaler Ebene ähnlich bzw. gleich den Dynamiken auf nationaler Ebene waren, weshalb sich ein Vergleich anbot.

Aufbereitung des INZDS: Aus dem INZDS wurden 41 Epidemiedatensätze entfernt, bei denen der Ort unbekannt war. Die übrigen Perspektiven wurden in drei Gruppen geteilt: Ausprägung **0** kennzeichnete Perspektiven auf substaatlicher, **1** die staatliche und **2** die suprastaatliche Ebene. Die Sowjetunion wurde, wie die USA, unter **1** gruppiert. England, Wales usw. wurden unter **0** eingruppiert, da sie Teile des Vereinigten Königreiches sind, welches unter **1** geführt wurde. Jetzt gab es 403 Perspektiven mit der Ausprägung **0**, 679 mit der Ausprägung **1** und 20 mit der Ausprägung **2**. Die Perspektiven mit Ausprägung **2** wurden entfernt und den Ausprägungen **0**, zuvor z. B. mittels Städtenamen identifiziert, wurden die zugehörigen Staatsnamen zugeordnet – die vier Perspektiven, die Epidemien auf Schiffen beschrieben, wurden entfernt. Nun wurden alle unvollständigen Perspektiven entfernt; sobald mindestens eine Ausprägung einer Welle ausgefüllt war, die anderen

aber nicht, erfolgte die Entfernung. Nun waren 388 Perspektiven mit der Ausprägung **0** zu SG, AG, RG2, SWG und COV1 übrig. Zudem verblieben 450 Perspektiven mit der Ausprägung **1** zu SG, AG, RG2, SWG, Ebolaepidemie von 2014 (EB14), COV1 und COV2. Bei Ebola hatte es jeweils Einzeldatensätze für die Jahre 1976, 1995 und 2018 gegeben, die entfernt wurden. Nun wurden vier Datensätze erzeugt (0erKurz, 0erLang, 1erKurz und 1erLang), jeweils zwei Ausführungen mit Perspektiven zur Ausprägung **0** und zwei für Perspektiven mit der Ausprägung **1**. Je ein Datensatz pro Ausprägung enthielt die Perspektiven mit weniger Wellen (z. B. 0erKurz) und der andere die Perspektiven mit mehr Wellen (z. B. 0erLang). Gab es mehrere Perspektiven pro Ort, die die gleiche Wellenanzahl hatten, so wurde eine Perspektive von einer dritten Person ohne Projektbeteiligung zufällig ausgewählt, die im Datensatz verblieb. Die beiden Datensätze zur Ausprägung **0** enthielten 28 Perspektiven: sieben zu SG, je eine zu AG und RG2, 17 zu SWG und zwei zu COV1. Die beiden Datensätze zur Ausprägung **1** enthielten 116 Perspektiven: zwei zu SG, drei zu AG, eine zu RG2, 31 zu SWG, drei zu EB14, vier zu COV1 und 72 zu COV2. Schließlich wurden unter der Annahme, dass die lokalen und nationalen Dynamiken ähnlich sind, die Datensätze zu den kurzen und den langen Perspektiven kombiniert, und es entstanden zwei weitere Datensätze (KombiKurz und KombiLang). Dabei entstanden erneut Dopplungen, die sogleich mittels zufälliger Auswahl durch eine dritte an dem Projekt unbeteiligte Person entfernt wurden. Diese beiden Datensätze enthielten jeweils 126 Perspektiven: sieben zu SG, drei zu AG, zwei zu RG2, 35 zu SWG, drei zu EB14, vier zu COV1 und 72 zu COV2.

Analyse und Vergleich der sechs Datensätze: Für die erste Analyse der Datensätze wurden in jedem Datensatz und jeweils pro Pandemie und Perspektive die DC gemessen und die Wellenhöhen bestimmt. Zudem wurde pro Datensatz und Epidemie die absolute Häufigkeit von Wellen gezählt. Berichtet wird in den Ergebnissen der Median (erstes Quartil [Q1]; drittes Quartil [Q3]) pro Epidemie und Welle. Bei Einzelwerten wurde der Einzelwert und bei zwei Werten der arithmetische Mittelwert berichtet. Das Höhenprofil wurde ermittelt, indem die jeweiligen Wellenhöhen einer Perspektive jeweils durch die Gipfelhöhe der ersten Welle geteilt wurden; berichtet wurden Median (Q1; Q3). Zudem wurde der Median der Mediane (Mdm) angegeben, stets ohne Berücksichtigung von COV2.

## V. Vorhersage des Zeitpunkts des zweiten Wellengipfels in COV2 auf Basis historischer Epidemien

Nachdem die DC bestimmt waren, galt es herauszufinden, ob es präzisere Berechnungsmethoden des zweiten Gipfels bei COV2 gab, die in ihrer Komplexität aber mit dem DC vergleichbar waren. Zur Evaluierung der Güte der Vorhersage wurde das Bestimmtheitsmaß  $R^2$  Adj. verwendet. Zudem galt es ein geeignetes Maß zu verwenden oder zu entwickeln, um die absolute Abweichung von Beobachtung und Vorhersage zu beschreiben.

### 1. Vorbereitung der Datensätze

Der erste Schritt bestand darin, in den Datensätzen (0erKurz, 0erLang, 1erKurz, 1erLang, KombiKurz und KombiLang) die aus Abbildung 1 ableitbaren Strecken zu messen. Für die weitere Berechnung wurden ausschließlich die Mediane verwendet, weshalb die korrespondierenden Ergebnisse nur die Mediane berichten. Zudem erschien es zunächst sinnvoll, die Zeitspannen in ihren absoluten und anteiligen Eigenschaften zu betrachten. Ersteres stellte sich als nicht zielführend heraus. Dies lag v. a. an der Datenheterogenität. Abfragen aus dem INZDS zeigten, dass die ersten beiden Wellen bei COV1 maximal 121 Tage dauerten, wohingegen es bei SG maximal 25 Monate waren. Betrachtete man aber etwa die MdM des Quotienten aus den Streckenlängen von **D** und **E**, also **D'** (genauer hierzu nachfolgend) in den Datensätzen, so fiel auf, dass diese zwischen 0,74 und 0,78 lagen. Dies bedeutete, dass der Anteil der Zeitstrecke von **a** bis **e** geteilt durch die Zeitstrecke von **a** bis **f**, verhältnismäßig konstant war. Der Ausdruck der relativen Strecken ermöglichte eine im Median wenig schwankende Messung, die in absoluten Zahlen verborgen blieb.

Tabelle 1 zeigt die 15 aus Abbildung 1 ableitbaren Strecken. Diese wurden in Paaren kombiniert und durch eine der vier Grundrechenarten (Addition, Subtraktion, Multiplikation und Division) verbunden. Diese Gleichungssysteme wurden nach dem Zeitpunkt **e**, also dem Höhepunkt der zweiten Welle, aufgelöst. Zudem wurde ermittelt, welche Gleichungssysteme die präziseste Schätzung ermöglichten und ob diese dem DC überlegen waren. Zentral war also die Annahme, dass aus den historischen Epidemien Werte abgeleitet werden konnten, die quasi als Konstante verwendet werden konnten, um die Gleichungssysteme sinnvoll nach **e** aufzulösen.

Da es um die Vorhersage des Zeitpunktes **e** ging, konnten alle Streckenkombinationen aus Tabelle 1 unberücksichtigt bleiben, die über den Zeitpunkt **e** hinausreichten. Im Ergebnisteil wird pro Datensatz, Epidemie und Streckenbezeichnung der jeweils ermittelte Median, die zugrundeliegende Anzahl an Beobachtungen und der MdM berichtet. Da die Strecken **E**, **I**, **L**, **N** und **O** also entfielen, blieben von den anfänglich 105 möglichen Streckenkombinationen zunächst 55 übrig. Entfernte man nun jene Kombinationen, die nicht nach dem Zeitpunkt **e** aufgelöst werden konnten, blieben insgesamt 30 Streckenkombinationen übrig. Dies ergab für alle vier Grundrechenarten 120 Gleichungssysteme zur Exploration. Nun wurden die Streckenanteile in den historischen Epidemien bestimmt, indem die Strecken durch die korrespondierende Strecke **E** geteilt wurden, um so den Anteil der jeweiligen Streckenbezeichnung an dem Zeitstrahl zu bestimmen. Die relativen Streckenbezeichnungen wurden durch die Verwendung eines ‘ hinter der jeweiligen Streckenbezeichnung kenntlich gemacht. Als relative Zahlenwerte für die **Erste** Streckenbezeichnung und die **Zweite** Streckenbezeichnung wurden die pro Datensatz ermittelten Werte des MdM verwendet.

Tabelle 1: Übersicht der aus Abbildung 1 ableitbaren Strecken mit Streckenbezeichnung

| Streckenbezeichnung | Anfangszeitpunkt | Endzeitpunkt |
|---------------------|------------------|--------------|
| A                   | a                | b            |
| B                   | a                | c            |
| C                   | a                | d            |
| D                   | a                | e            |
| E                   | a                | f            |
| F                   | b                | c            |
| G                   | b                | d            |
| H                   | b                | e            |
| I                   | b                | f            |
| J                   | c                | d            |
| K                   | c                | e            |
| L                   | c                | f            |
| M                   | d                | e            |
| N                   | d                | f            |
| O                   | e                | f            |

Die Tabelle zeigt die 15 aus Abbildung 1 ableitbaren Strecken. Aufgeführt sind pro Strecke jeweils der Anfangs- und Endzeitpunkt sowie die korrespondierende Streckenbezeichnung. Zur Ermittlung des Anteils einer Streckenbezeichnung an dem zweiwelligen Geschehen wurde jede Streckenbezeichnung, ausgenommen jener von COV2, durch die korrespondierende Streckenbezeichnung **E** geteilt.



## 2. Gleichungssysteme

Nun bestand jedes historische Gleichungssystem aus zwei bekannten Werten, die durch eine der vier Grundrechenarten verbunden waren und ein bekanntes, errechnetes Ergebnis hatten. In allen Gleichungssystemen enthielt die **Zweite** Streckenbezeichnung den gesuchten Wert **e**, weshalb diese Streckenbezeichnung fortan als Unbekannte behandelt wurde. Vereinzelt kam es vor, dass der Zeitpunkt **e** in beiden Strecken enthalten war. Da aber die **Erste** Streckenbezeichnung ein bekannter Wert war, ergaben sich ausschließlich Gleichungssysteme mit einer Unbekannten. Nun gab es vier Gleichungssysteme, eines pro Grundrechenart, die jeweils aus einer Konstanten **X** und dem bekannten Rechenergebnis der historischen Epidemien von **Erster** und **Zweiter** Streckenbezeichnung auf der linken Seite bestanden, sowie einer Konstanten **Y**, dem bekannten Messwert der historischen Epidemien der **Ersten** Streckenbezeichnung, und einem Ausdruck **V**, der die Unbekannte **e** enthielt, auf der rechten Seite. Die Gleichungen waren also:

$$X = Y + V \quad (6)$$

$$X = Y - V \quad (7)$$

$$X = Y * V \quad (8)$$

$$X = \frac{Y}{V} \quad (9)$$

**V** war die in ihre Einzelinformationen zerlegte **Zweite** Streckenbezeichnung und setzte sich zusammen aus der Division der jeweiligen Streckenbezeichnung durch die Streckenbezeichnung **E**. Da die Streckenbezeichnung **E** die Subtraktion der Zahlenwerte der Zeitpunkte **a** von **f** war und **f** der Zeitpunkt des Endes der zweiten Welle war, der unbekannt war, musste die Streckenbezeichnung **E**, deren relativer Wert **E'** stets 1,00 war, durch bekannte Strecken ausgedrückt werden, die sich aus den bislang beschriebenen Zeitpunkten zusammensetzten, die zeitlich vor **e**, also dem Höhepunkt der zweiten Welle lagen. Obwohl es auch andere Möglichkeiten gab, wurden hier die drei Strecken: **a** bis **b**, **a** bis **c** und **a** bis **d**, also die Streckenbezeichnungen **A**, **B** und **C** verwendet. Nun wurde der Faktor **Z** bestimmt, der angab, wie oft eine der drei genannten Strecken multipliziert werden musste, um 1,00 zu erhalten. Dies passierte, indem man jeweils 1,00 durch den Wert der Streckenbezeichnungen **A**, **B** und **C** teilte. Die drei **Z**-Faktoren mussten für jeden der sechs Datensätze (0erKurz, 0erLang, 1erKurz, 1erLang, KombiKurz und KombiLang) und für jede der Varianten ( $Z_A$ ,  $Z_B$ , und  $Z_C$ ) bestimmt werden. An diesem Punkt gab es

die vier nachfolgenden Gleichungen, die sich nach  $e$  als einzige Unbekannte auflösen ließen:

$$X = Y + \frac{(e - P_1)}{((P_3 - P_2) * Z_n)} \quad (10)$$

$$X = Y - \frac{(e - P_1)}{((P_3 - P_2) * Z_n)} \quad (11)$$

$$X = Y * \frac{(e - P_1)}{((P_3 - P_2) * Z_n)} \quad (12)$$

$$X = \frac{Y}{\left(\frac{(e - P_1)}{((P_3 - P_2) * Z_n)}\right)} \quad (13)$$

$P_1$  war der jeweilige bekannte Zeitpunkt der **Zweiten** Streckenbezeichnung,  $P_3$  war der jeweilige, zutreffende und bekannte Endzeitpunkt und  $P_2$  der jeweilige, korrespondierende und bekannte Anfangszeitpunkt der Streckenbezeichnungen **A**, **B** oder **C**.

Diese Gleichungssysteme, die ermittelten Z-Faktoren und die ermittelten historischen Schätzer wurden nun auf die kurzen und langen COV2-Perspektiven angewendet – dies wurde durch das jeweilige Suffix „2019Kurz“ oder „2019Lang“ kenntlich gemacht. „0erKurz2019Kurz“ bedeutete etwa, dass die historischen Werte aus dem Datensatz der wellenarmen Zählungsweise der regionalen Perspektiven (0erKurz) auf die wellenarme Zählweise (kurz) der COV2 Perspektiven angewendet wurden. Gerechnet wurden nun die 30 Formeln für jede der vier Transformationen, die wiederum für jede der drei Z-Faktoren ( $Z_A$ ,  $Z_B$ , und  $Z_C$ ) berechnet wurden, die wiederum für jede der 12 Versuchsreihen (0erKurz2019Kurz, 0erKurz2019Lang, 0erLang2019Kurz, 0erLang2019Lang, 1erKurz2019Kurz, 1erKurz2019Lang, 1erLang2019Kurz, 1erLang2019Lang, KombiKurz2019Kurz, KombiKurz2019Lang, KombiLang2019Kurz und KombiLang2019Lang) berechnet wurden, was insgesamt 4.320 Berechnungen (nachfolgend als Beobachtungen bezeichnet) ergab. Für die Ermittlung der Zahlenwerte der Zeitpunkte wurde der 01.01.1900 als Tag Eins gewählt. Jedes Datum ist also die Anzahl der Tage seit dem 31.12.1899.

Beispielrechnung Addition: Im Datensatz 0erKurz war  $Z_A = 7,14$ . In einer Streckenkombination war  $X = 0,92$  und  $Y = 0,14$ . Bekannt waren die kurzen COV2 Zeitpunkte für Afghanistan mit  $\mathbf{a} = 43.915$ ,  $\mathbf{b} = 43.981$ . Setzte man die Werte ein, so erhielt man

$e = 44.283$ , den 27.03.2021. Der tatsächliche Wert in diesem Datensatz, für den Höhepunkt der zweiten Welle in Afghanistan war 44.157, also der 22.11.2020. Der Betragswert der Differenz von Beobachtung zu Prädiktion betrug 126 Tage. Die nach  $e$  aufgelöste Formel lautete:

$$e = ((a - b)Y + (b - a)X)Z + a \quad (14)$$

Beispielrechnung Subtraktion: Im Datensatz 0erKurz war  $Z_B = 2,94$ . In einer Streckenkombination war  $X = -0,64$  und  $Y = 0,14$ . Bekannt waren die langen COV2 Zeitpunkte für Afghanistan mit  $a = 43.915$ ,  $c = 44.048$ . Setzte man die Werte ein, so erhielt man  $e = 44.220$ , den 23.01.2021. Der tatsächliche Wert in diesem Datensatz, für den Höhepunkt der zweiten Welle in Afghanistan war 44.157, also der 22.11.2020. Der Betragswert der Differenz von Beobachtung zu Prädiktion betrug 63 Tage. Die nach  $e$  aufgelöste Formel lautete:

$$e = ((c - a)Y + (a - c)X)Z + a \quad (15)$$

Beispielrechnung Multiplikation: Im Datensatz 0erLang galt für  $Z_C = 1,69$ . In einer Streckenkombination war  $X = 0,13$  und  $Y = 0,18$ . Bekannt waren die kurzen COV2 Zeitpunkte für Afghanistan mit  $a = 43.915$ ,  $d = 44.109$ . Setzte man die Werte ein, so erhielt man  $e = 44.152$ , den 16.11.2020. Der tatsächliche Wert in diesem Datensatz, für den Höhepunkt der zweiten Welle in Afghanistan war 44.157, der 22.11.2020. Der Betragswert der Differenz von Beobachtung zu Prädiktion betrug 5 Tage. Die nach  $e$  aufgelöste Formel lautete:

$$e = \frac{((d - a)XZ + aY)}{Y} \quad (16)$$

Beispielrechnung Division: Im Datensatz 0erLang galt für  $Z_C = 1,69$ . In einer Streckenkombination war  $X = 0,24$  und  $Y = 0,18$ . Bekannt waren die langen COV2 Zeitpunkte für Afghanistan mit  $a = 43.915$ ,  $d = 44.109$ . Setzte man die Werte ein, so erhielt man  $e = 44.161$ , den 25.11.2020. Der tatsächliche Wert in diesem Datensatz, für den Höhepunkt der zweiten Welle in Afghanistan war 44.157, also der 22.11.2020. Der Betragswert der Differenz von Beobachtung zu Prädiktion betrug also 4 Tage. Die nach  $e$  aufgelöste Formel lautete:

$$e = \frac{((d - a)YZ + aX)}{X} \quad (17)$$

Die zugehörigen Tabellen berichten pro Datensatz jeweils die untersuchten Streckenbezeichnungskombinationen und die dazugehörigen Ergebnisse für die verschiedenen mathematischen Operationskombinationen, sowie die zugehörigen Z-Werte.

### 3. Güte der Schätzungen

Als die 4.320 Beobachtungen berechnet waren, wurde ermittelt, welches Modell, also die Vereinigung von mathematischer Formel, langer oder kurzer Perspektive und die aus den historischen Epidemien abgeleiteten Werte den zweiten Wellengipfel am präzisesten vorhersagten. Hierzu wurden zwei Methoden verwendet: 1. die Güte der Zusammenhänge von beobachtetem und prognostiziertem zweiten Wellengipfel mittels  $R^2$  Adj. und 2. die in vollen Tagen errechnete, absolute Abweichung von Beobachtung und Vorhersage.

1. Pro Beobachtung, die durch die Versuchsreihe, die jeweilige Formel und die historischen Konstanten identifizierbar war, ergab sich ein Ergebnis für  $R^2$  Adj. Die Ergebnisse wurden in absteigender Reihenfolge sortiert, sodass das größte Ergebnis für  $R^2$  Adj., also die Beobachtung, die das meiste an Varianz erklärte, an oberster Stelle stand. Dadurch entstand eine Rangordnung, aus der sich bestimmen ließ, welche Beobachtung mit ihren jeweiligen Eigenschaften den zweiten Wellengipfel in COV2 am präzisesten vorhersagte. Zudem konnte so die Platzierung der DC in der Rangordnung bestimmt werden, und es war möglich, diesen Ansatz mit dem mathematisch explorativen Ansatz zu vergleichen.

2. Um die Abweichung von Beobachtung und Vorhersage in vollen Tagen zu beschreiben, wurde der Betragswert der Subtraktion von beobachtetem und vorhergesagtem Zeitpunkt ermittelt und kaufmännisch auf die nächste ganze Zahl gerundet. Durch die Betragswerte hatten die Werte das gleiche Vorzeichen, wodurch eine sinnvolle Deskription dieser absoluten Abweichung möglich war. Ermittelt wurden pro Beobachtung Median (Q1; Q3) und Interquartilsabstand (IQA). Die Ergebnisse wurden in eine geordnete Reihe gebracht, die absteigend sortiert war. Die Präzision wurde v. a. an zwei Variablen untersucht: 1. ein möglichst geringer Wert bei Q3, sodass 75% der Vorhersagen einer Beobachtung möglichst nahe an den tatsächlichen zweiten Wellenhöhepunkten lagen und 2. ein möglichst geringer Wert des IQA, was auf eine geringe Streuung der Berechnungsmethode hinwies. Um die Präzision der Kombinationen initial einzuordnen, wurden die medianen Betragswerte der absoluten Abweichung, getrennt nach Z-Faktor, visuell verglichen (siehe Abbildung 3).

## VI. Passende historische Epidemie

Die letzte Forschungsfrage war, welche der untersuchten Epidemien im Rahmen der hier ermittelten Charakteristika COV2 am ähnlichsten war und ob eine Taxonomie erkennbar war. In historischen Vergleichen wurden häufig Inzidenzen und Mortalitäten verglichen; dies erlaubte aber nur eine oberflächliche Interpretation der epidemiologischen Charakteristika. Die Frage war, ob sich Hinweise auf epidemische Gruppen fanden. Um diese Frage zu beantworten, wurden qualitative und quantitative Vergleiche angestellt. Da 0erKurz und 0erLang keine Werte zu COV2 enthielten, waren sie von der nachfolgenden Analyse ausgenommen. Auf qualitativer Ebene wurde für jeden Datensatz diskutiert, inwiefern sich die Messwerte von COV2 und den historischen Epidemien hinsichtlich der Wellenzahl, der DC und der Höhenprofile ähnelten bzw. sich unterschieden. Auch eine datensatzübergreifende Betrachtung wurde formuliert. Der quantitative Vergleich erfolgte dann in einem eigenständigen Kapitel für die Datensätze: 1erKurz, 1erLang, KombiKurz und KombiLang. Hier wurden pro Datensatz die zuvor ermittelten medianen Streckenanteile der historischen Epidemien mittels linearer Regression mit den medianen Streckenanteilen von COV2 verglichen. Das kaufmännisch auf zwei Nachkommastellen gerundete  $R^2$  Adj. gab Auskunft darüber, welche historischen Epidemien gut mit COV2 vergleichbar waren. Der zugehörige mathematische Ausdruck war:

$$\text{Streckenanteile}_{\text{Epidemie}} = a + b \text{ Streckenanteile}_{\text{COV2}} \quad (18)$$

Eine datensatzübergreifende Betrachtung ermöglichte sodann einen Vorschlag der Taxonomie. Der Vorschlag war hier nur für die jeweilige Abfolge der medianen Werte der Streckenbezeichnungen möglich, da man für eine Taxonomie des Höhenprofils auch die Inzidenzhöhe zu den einzelnen Zeitpunkten berücksichtigen sollte.

## D. Ergebnisse

Es entstand die vielleicht erste Datenbank, die eine Erforschung der epidemiologischen Charakteristika von durch Tröpfcheninfektion übertragbaren Virusepidemien und deren Inzidenzverläufe über die vergangenen 130 Jahre auf einer breiten statistischen Basis ermöglichte.

### I. Ergebnisse der Literaturrecherchen

Untersuchungen, die die zeitlichen Verläufe von Epidemien auf Public Health Ebene untersuchten, waren unter Verwendung des dargestellten Suchalgorithmus sehr selten. Von den 18.696 Suchergebnissen, die am 23.03.2021 auf PubMed zu dem Thema „*Influenza AND Pandemic*“ zur Verfügung standen, konnten lediglich 201 (1,1%) Artikel identifiziert werden, die verwertbare Zeitreihendarstellungen beinhalteten. Da sich die Arbeit auf die Untersuchung von Inzidenzverläufen beschränkte, waren es letztlich nur 131 (0,7%) Arbeiten, die berücksichtigt werden konnten. Die Literaturrecherche zu Ebola war anteilmäßig ergiebiger. So konnten etwa für „*Ebola AND Wave*“ drei der 22 (13,6%) Einträge verwendet werden. Allerdings gab es auch Recherchen, die ergebnislos blieben.

Ein anderer Aspekt, den die Literaturrecherchen verdeutlichten, ist, dass es ausgewählte Epidemien gab, die – im Vergleich untereinander – gut erforscht waren, zugleich gab es Epidemien, die wenig Niederschlag in der untersuchten Literatur hatten. Von den 131 Artikeln, die in den INZDS eingingen, konnten in 118 (90,1%) Arbeiten Darstellungen zu SWG gefunden werden. In zehn (7,6%) Artikeln fanden sich Darstellungen zu SG, zwei (1,5%) Artikel hatten Darstellungen zu AG und ein (0,8%) Artikel hatte Darstellungen zu RG2. Nimmt man also SWG und SG aus, so waren fünf der sieben möglichen Influenzapandemien der hier untersuchten epidemiologischen Literatur quasi unbekannt. Ein ähnliches Bild zeigte sich auch bei den Ebolaepidemien: Von 13 Arbeiten, die in den Datensatz einfließen, betrafen neun (69,2%) Artikel EB14; jeweils ein (7,7%) Artikel entfiel auf die Epidemien der Jahre 1976, 1995, 2017 und 2018.

### II. Interratervergleich

Die Übereinstimmung der Rater (Fleiss' Kappa) in Bezug auf die Wellenanzahl war 0,60 95%-Konfidenzintervall (0,45; 0,75) und damit knapp unter den definierten 0,61, die eine gute Übereinstimmung bedeuteten (Fleiss, 1971). Bei den jeweils berechneten DC der ersten beiden Wellen zeigte sich, wie aus Abbildung 2 ersichtlich, zwischen den Ratern

Stang und Kowall der schwächste Zusammenhang von  $R^2 = 0,44$ . Der Zusammenhang zwischen Standl und Stang war mit  $R^2 = 0,50$  etwas stärker und schließlich zwischen Standl und Kowall mit  $R^2 = 0,75$  am stärksten. Obwohl es unterschiedliche Auffassungen davon gab, wie die Stärke eines Zusammenhanges zu bewerten war (Baidoun, 2016), konnte man durchaus zu dem Schluss kommen, dass die ermittelten Zusammenhänge moderat bis substantiell waren (Fung, 2015).

### III. Häufigkeit von Inzidenzwellen, Duty Cycle und Wellenhöhenprofile

Anders als man es aus der aktuellen Forschung vermuten konnte, waren mehrwellige Epidemien nicht die Ausnahme (Hoen et al., 2015), sondern die Regel. Verallgemeinernd konnte man sagen, dass zweite Infektionswellen in 3/4 und dritte Infektionswellen in 1/3 der Fälle vorkamen. Man konnte sagen, dass sich ab der dritten Infektionswelle die Häufigkeit je Folgewelle halbierte.

#### 1. Häufigkeit von Folgewellen

Bei 0erKurz zeigte sich, dass jede Pandemie eine zweite Welle hatte und diese Zweitwellen in etwa 1/3 der Perspektiven auftraten. Dritte Wellen traten in zwei der drei untersuchten H1-Pandemien auf (Tabelle 2). Bei 1erKurz traten Zweitwellen bei drei von vier Influenzapandemien auf, bei EB14 in 1/3 der Perspektiven und bei den SARS-Epidemien ca. in 3/4 der Perspektiven (Tabelle 3). Bei KombiKurz wurden in allen Epidemien zweite Infektionswellen beobachtet. Vierte Infektionswellen traten in SWG und in COV2 auf. Fünfte Infektionswellen wurden nur in COV2 beobachtet (Tabelle 4). Über alle kurzen Perspektiven hinweg ließ sich also feststellen, dass epidemische Folgewellen keine Seltenheit waren und Zweitwellen in EB14 in knapp 1/3 der epidemischen Ereignisse vorkamen. Bei Influenzapandemien traten Zweitwellen etwa in der Hälfte der Fälle auf und in Coronapandemien mindestens in sechs von sieben epidemischen Ereignissen. Ein relevanter Unterschied zwischen den lokalen und nationalen Beobachtungen ließ sich hier – auch wegen der letztlich kleinen Anzahl an Beobachtungen – nicht feststellen. Es gab keinen Hinweis darauf, dass eine kombinierte Betrachtung von lokalen und nationalen epidemischen Ereignissen nicht zulässig gewesen wäre. Vereinfacht bedeutete das, dass es nicht relevant war, ob in einen Datensatz beispielsweise eine Perspektive für Stockholm oder für Schweden einging – mögliche Unterschiede hatten keine relevanten Auswirkungen.

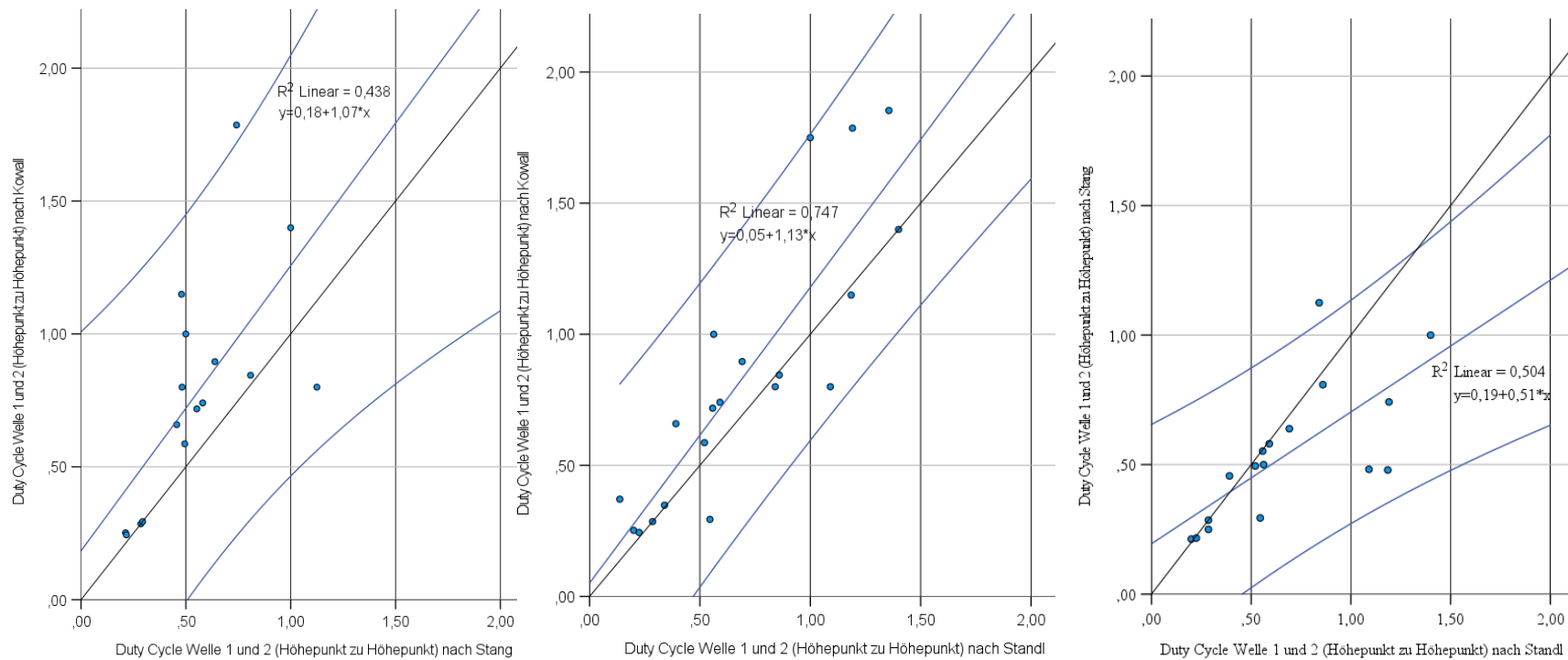


Abbildung 2: Güte der mittels  $R^2$  ausgedrückten Zusammenhänge der DC von Erst- und Zweitwelle bei den Ratern Kowall, Stang und Standl

Bei jedem Rater wurden die identifizierten Zeitpunkte der Erst- und Zweitwellen erfasst und daraus jeweils der DC dieser Wellen errechnet. Die ermittelten DC wurden dann den DC der anderen beiden Rater gegenübergestellt. Mittels linearer Regression wurde verglichen, wie gut die ermittelten DC der Rater übereinstimmen. Die Zusammenhänge zwischen Kowall und Stang (Darstellung links), sowie von Stang und Standl (Darstellung rechts) können mit  $R^2 = 0,44$  bzw.  $0,50$  als mäßig beschrieben werden. Der Zusammenhang zwischen Kowall und Standl (Darstellung Mitte) ist mit  $R^2 = 0,75$  substantiell. Die schwarze Gerade in jeder Darstellung ist die Winkeldiagonale, die blaue Gerade ist die jeweilige Regressionsgerade und die in jeder Darstellung vorkommenden blauen Kurven sind die Ränder der 95% Konfidenzintervalle.



Tabelle 2: Median (Q1; Q3) im Datensatz der regionalen, kurzen Perspektiven (0erKurz)

| Pandemie           |               | DC12              | DC23     | H2/H1             | H3/H1    |
|--------------------|---------------|-------------------|----------|-------------------|----------|
| SG                 |               | 0,58 (0,36; NA)   | 2,04     | 1,88 (0,17; NA)   | 0,13     |
|                    | Beobachtungen | 3 aus 7           | 1 aus 7  | 3 aus 7           | 1 aus 7  |
| AG                 |               | 0,69              |          | 0,6               |          |
|                    | Beobachtungen | 1 aus 1           |          | 1 aus 1           |          |
| RG2                |               | 0,39              |          | 0,58              |          |
|                    | Beobachtungen | 1 aus 1           |          | 1 aus 1           |          |
| SWG                |               | 0,64 (0,53; 1,34) | 1,20     | 0,50 (0,45; 1,02) | 0,50     |
|                    | Beobachtungen | 5 aus 17          | 1 aus 17 | 5 aus 17          | 1 aus 17 |
| Median der Mediane |               | 0,61              | 1,62     | 0,59              | 0,31     |

In dieser Tabelle sind die Mediane (Q1; Q3) dargestellt, sowie die Anzahl an Beobachtungen, auf denen die Angaben der Mediane basieren. Ausgangspunkt für die Angaben ist der Datensatz 0erKurz, also die regionalen, kurzen Perspektiven, die aus dem INZDS ausgewählt wurden. NA bedeutet, dass ein Wert nicht ermittelt werden konnte. Der Median der Mediane bezieht sich auf die medianen Angaben zu DC und Wellenhöhen und umfasst nicht die Anzahl an Beobachtungen. Die Wellenhöhenangaben ergeben sich aus der jeweiligen Wellenhöhe geteilt durch die Höhe der korrespondierenden ersten Welle. Nachfolgend ein Ablesebeispiel: In der Pandemie SG betrug der mediane DC12, also der DC der Wellen Eins und Zwei, der auf drei Beobachtungen basiert, 0,58. Der DC der Wellen Zwei und Drei basiert auf einer Perspektive, weshalb hier lediglich ein Einzelwert angegeben ist. H2/H1 gibt die mediane Höhe der Zweitwellen an. Der Lageparameter wurde ermittelt, indem die Höhe der Zweitwelle der drei Perspektiven durch die jeweilige Höhe der Erstwelle geteilt wurde. H3/H1 gibt dann die Höhe der Drittwelle(n) in Bezug auf die Erstwelle(n) an. Durch die Angabe der Anzahl an Beobachtungen lässt sich ablesen, dass in SG drei aus sieben Perspektiven eine Zweitwelle hatten und eine Perspektive aus sieben eine Drittwelle, was zugleich bedeutet, dass in einer aus drei Perspektiven mit einer Zweitwelle eine dritte Infektionswelle beobachtet wurde. Der Median der Mediane ist der mediane Wert der ermittelten Mediane einer Spalte.

Tabelle 3: Median (Q1; Q3) im Datensatz der nationalen, kurzen Perspektiven (1erKurz)

| Epidemie           | DC12              | DC23              | DC34              | DC45              | H2/H1             | H3/H1              | H4/H1             | H5/H1             |
|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|
| SG                 | 0,80              |                   |                   |                   | 4,99              |                    |                   |                   |
| Beob. <sup>1</sup> | 2 aus 2           |                   |                   |                   | 2 aus 2           |                    |                   |                   |
| AG                 | 0,83 (0,71; NA)   | 0,50 (0,27; NA)   |                   |                   | 0,88 (0,36; NA)   | 0,88 (0,86; NA)    |                   |                   |
| Beob. <sup>1</sup> | 3 aus 3           | 3 aus 3           |                   |                   | 3 aus 3           | 3 aus 3            |                   |                   |
| RG2                | 0,12              |                   |                   |                   | 12,50             |                    |                   |                   |
| Beob. <sup>1</sup> | 1 aus 1           |                   |                   |                   | 1 aus 1           |                    |                   |                   |
| COV1               | 0,44 (0,39; NA)   |                   |                   |                   | 1,50 (1,00; NA)   |                    |                   |                   |
| Beob. <sup>1</sup> | 3 aus 4           |                   |                   |                   | 3 aus 4           |                    |                   |                   |
| SWG                | 0,65 (0,51; 0,93) | 0,50 (0,38; 1,07) | 0,20              |                   | 1,53 (0,63; 4,53) | 6,00 (2,92; 11,72) | 1,67              |                   |
| Beob. <sup>1</sup> | 16 aus 31         | 5 aus 31          | 1 aus 31          |                   | 16 aus 31         | 5 aus 31           | 1 aus 31          |                   |
| EB14               | 1,83              |                   |                   |                   | 0,47;             |                    |                   |                   |
| Beob. <sup>1</sup> | 1 aus 3           |                   |                   |                   | 1 aus 3           |                    |                   |                   |
| COV2               | 0,51 (0,30; 0,91) | 0,94 (0,55; 1,20) | 0,64 (0,53; 0,94) | 0,87 (0,75; 1,34) | 1,48 (0,71; 3,13) | 2,02 (0,90; 7,94)  | 2,46 (1,21; 7,70) | 3,71 (1,52; 8,75) |
| Beob. <sup>1</sup> | 63 aus 72         | 29 aus 72         | 15 aus 72         | 6 aus 72          | 63 aus 72         | 29 aus 72          | 15 aus 72         | 6 aus 72          |
| MdM <sup>2</sup>   | 0,72              | 0,50              | 0,20              |                   | 1,51              | 3,44               | 1,67              |                   |

Beob.<sup>1</sup>=Beobachtungen. MdM<sup>2</sup>=Median der Mediane ohne COV2. In dieser Tabelle sind die Mediane (Q1; Q3) dargestellt, sowie die Anzahl an Beobachtungen, auf denen die Angaben der Mediane basieren. Ausgangspunkt für die Angaben ist der Datensatz 1erKurz, also die nationalen, kurzen Perspektiven, die aus dem INZDS ausgewählt wurden. NA bedeutet, dass ein Wert nicht ermittelt werden konnte. Der Median der Mediane bezieht sich auf die medianen Angaben zu DC und Wellenhöhen und umfasst nicht die Anzahl an Beobachtungen. Die Wellenhöhenangaben ergeben sich aus der jeweiligen Wellenhöhe geteilt durch die Höhe der korrespondierenden ersten Welle. Ablesbeispiel siehe Tabelle 2.

Tabelle 4: Median (Q1; Q3) im Datensatz der kombinierten regionalen und nationalen, kurzen Perspektiven (KombiKurz)

| Epidemie         | DC12               | DC23              | DC34              | DC45              | H2/H1             | H3/H1             | H4/H1             | H5/H1             |
|------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| SG               | 0,58 (0,36; NA)    | 2,04              |                   |                   | 1,88 (0,84; NA)   | 0,13              |                   |                   |
|                  | Beob. <sup>1</sup> | 3 aus 7           | 1 aus 7           |                   | 3 aus 7           | 1 aus 7           |                   |                   |
| AG               | 0,71 (0,69; NA)    | 0,64              |                   |                   | 0,60 (0,36; NA)   | 1,69              |                   |                   |
|                  | Beob. <sup>1</sup> | 3 aus 3           | 2 aus 3           |                   | 3 aus 3           | 2 aus 3           |                   |                   |
| RG2              | 0,25               |                   |                   |                   | 6,54              |                   |                   |                   |
|                  | Beob. <sup>1</sup> | 2 aus 2           |                   |                   | 2 aus 2           |                   |                   |                   |
| COV1             | 0,62               |                   |                   |                   | 1,59              |                   |                   |                   |
|                  | Beob. <sup>1</sup> | 2 aus 4           |                   |                   | 2 aus 4           |                   |                   |                   |
| SWG              | 0,74 (0,53; 1,04)  | 0,57 (0,41; 1,27) | 0,20              |                   | 1,33 (0,49; 2,51) | 2,27 (1,02; 6,60) | 0,10              |                   |
|                  | Beob. <sup>1</sup> | 17 aus 35         | 6 aus 35          | 1 aus 35          | 17 aus 35         | 6 aus 35          | 1 aus 35          |                   |
| EB14             | 1,83               |                   |                   |                   | 0,47              |                   |                   |                   |
|                  | Beob. <sup>1</sup> | 1 aus 3           |                   |                   | 1 aus 3           |                   |                   |                   |
| COV2             | 0,51 (0,30; 0,91)  | 0,94 (0,55; 1,20) | 0,64 (0,53; 0,94) | 0,87 (0,75; 1,34) | 1,48 (0,71; 3,13) | 1,91 (0,83; 4,18) | 1,49 (1,00; 2,92) | 1,20 (0,83; 2,59) |
|                  | Beob. <sup>1</sup> | 63 aus 72         | 29 aus 72         | 15 aus 72         | 6 aus 72          | 63 aus 72         | 29 aus 72         | 15 aus 72         |
| MdM <sup>2</sup> | 0,66               | 0,64              | 0,20              |                   | 1,46              | 1,69              | 0,10              |                   |

Beob.<sup>1</sup>=Beobachtungen. MdM<sup>2</sup>=Median der Mediane ohne COV2. In dieser Tabelle sind die Mediane (Q1; Q3) dargestellt, sowie die Anzahl an Beobachtungen, auf denen die Angaben der Mediane basieren. Ausgangspunkt für die Angaben ist der Datensatz KombiKurz, also die Kombination der regionalen und nationalen, kurzen Perspektiven, die aus dem INZDS ausgewählt wurden. NA bedeutet, dass ein Wert nicht ermittelt werden konnte. Der Median der Mediane bezieht sich auf die medianen Angaben zu DC und Wellenhöhen und umfasst nicht die Anzahl an Beobachtungen. Die Wellenhöhenangaben ergeben sich aus der jeweiligen Wellenhöhe geteilt durch die Höhe der korrespondierenden ersten Welle. Ablesebeispiel siehe Tabelle 2.

Bei 0erLang wurden in allen Epidemien Zweitwellen beobachtet. In den beiden großen H1-Pandemien SG und SWG waren die Häufigkeiten von Folgewellen etwa ähnlich (Tabelle 5). Das Bild in 1erLang unterschied sich z. T. deutlich von den langen lokalen Perspektiven. SG hatte hier jeweils in allen Fällen eine Zweit- und eine Drittwelle. In AG und SWG wurden vierte und fünfte Wellen beobachtet. Während in COV1 keine vierten und fünften Wellen beobachtet wurden, so war dies in COV2 in knapp 1/3 bzw. 1/7 der Perspektiven der Fall (Tabelle 6). Bei KombiLang zeigte sich, dass jede Epidemie mindestens eine zweite Welle hatte. Epidemien mit insgesamt drei Wellen waren SG und COV1. Die übrigen Epidemien zeigten fünf Wellen (Tabelle 7).

Insgesamt zeigte sich, dass sich von Welle zu Welle die Anzahl der Perspektiven halbierte und sich die Wellenhäufigkeiten von SWG und COV2 ähnelten. Ob nun die Wellenzählweise der kurzen oder langen Perspektiven sinnvoller bzw. richtig war, ließ sich an diesem Punkt nicht feststellen. Die Unterschiede aber der kurzen regionalen und nationalen Datensätze bzw. der Datensätze mit langen regionalen und nationalen Perspektiven schienen unerheblich zu sein, weshalb eine Kombination von regionalen und nationalen Perspektiven sinnvoll erschien, da so die Stichprobengrößen umfangreicher waren. Die Wellenhäufigkeiten zwischen kurzen und langen Perspektiven schienen relevant zu sein, es war aber fraglich, ob die Unterschiede sich auch bei umfangreicheren Stichproben zeigen würden. Da diese Frage hier nicht abschließend beantwortet werden konnte, empfiehlt sich für weitere Untersuchungen eine Trennung von kurzen und langen Perspektiven.

## 2. Duty Cycle in den verschiedenen Epidemien

In diesem Abschnitt beziehen sich alle Zahlenangaben auf die medianen Messwerte. In 0erKurz und KombiKurz passte RG2 nicht recht zu den anderen Pandemien (Tabelle 2 und 4). In 1erKurz fiel auf, dass EB14 die einzige Nicht-Pandemie war und als einzige Epidemie einen DC größer 1,00 hatte (Tabelle 3). In 0erLang zeigte sich bei der Wellenanzahl und den DC eine relevante Ähnlichkeit von SG und SWG (Tabelle 5). In 1erLang lag der DC von EB14 erneut bei 1,00. Die DC der Wellen Zwei und Drei lagen bis auf EB14 unterhalb von 1,00 (Tabelle 6). In KombiLang lag EB14 mit 1,07 erneut über der 1,00 (Tabelle 7). Besonders niedrige bzw. hohe DC basierten häufig auf Einzelwerten, weshalb damit gerechnet werden musste, dass es sich um Messfehler bzw. Ausreißer handelte. Eine Möglichkeit, die Datensätze miteinander zu vergleichen, war es, den MDM pro DC zu ermitteln und zu vergleichen. Da der Median von COV2 nicht in die Ermittlung

Tabelle 5: Median (Q1; Q3) im Datensatz der regionalen, langen Perspektiven (0erLang)

| Pandemie           | DC12              | DC23              | DC34              | DC45              | H2/H1             | H3/H1             | H4/H1             | H5/H1             |
|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| SG                 | 0,58 (0,36; 0,75) | 0,93 (0,81; 1,69) | 1,04 (0,78; NA)   | 0,89 (0,87; NA)   | 1,88 (0,43; 2,60) | 0,45 (0,20; 1,17) | 0,50 (0,33; NA)   | 0,28 (0,23; NA)   |
| Beob. <sup>1</sup> | 7 aus 7           | 6 aus 7           | 2 aus 7           | 2 aus 7           | 7 aus 7           | 6 aus 7           | 2 aus 7           | 2 aus 7           |
| AG                 | 1,20              | 0,46              |                   |                   | 0,16              | 0,60              |                   |                   |
| Beob. <sup>1</sup> | 1 aus 1           | 1 aus 1           |                   |                   | 1 aus 1           | 1 aus 1           |                   |                   |
| RG2                | 0,49              |                   |                   |                   | 1,13              |                   |                   |                   |
| Beob. <sup>1</sup> | 1 aus 1           |                   |                   |                   | 1 aus 1           |                   |                   |                   |
| COV1               | 0,86              | 0,14              |                   |                   | 2,31              | 0,53              |                   |                   |
| Beob. <sup>1</sup> | 2 aus 2           | 1 aus 2           |                   |                   | 2 aus 2           | 1 aus 2           |                   |                   |
| SWG                | 0,77 (0,50; 1,50) | 0,68 (0,62; 1,04) | 0,97 (0,44; 1,18) | 1,00 (0,40; 2,00) | 1,17 (0,29; 4,33) | 1,68 (0,17; 3,67) | 2,60 (0,43; 8,45) | 0,67 (0,33; 1,25) |
| Beob. <sup>1</sup> | 15 aus 17         | 10 aus 17         | 8 aus 17          | 5 aus 17          | 15 aus 17         | 10 aus 17         | 8 aus 17          | 5 aus 17          |
| MdM <sup>2</sup>   | 0,77              | 0,57              | 1,00              | 0,94              | 1,17              | 0,56              | 1,55              | 0,47              |

Beob.<sup>1</sup>=Beobachtungen. MdM<sup>2</sup>=Median der Mediane ohne COV2. In dieser Tabelle sind die Mediane (Q1; Q3) dargestellt, sowie die Anzahl an Beobachtungen, auf denen die Angaben der Mediane basieren. Ausgangspunkt für die Angaben ist der Datensatz 0erLang, also die regionalen, langen Perspektiven, die aus dem INZDS ausgewählt wurden. NA bedeutet, dass ein Wert nicht ermittelt werden konnte. Der Median der Mediane bezieht sich auf die medianen Angaben zu DC und Wellenhöhen und umfasst nicht die Anzahl an Beobachtungen. Die Wellenhöhenangaben ergeben sich aus der jeweiligen Wellenhöhe geteilt durch die Höhe der korrespondierenden ersten Welle. Ablesbeispiel siehe Tabelle 2.

Tabelle 6: Median (Q1; Q3) im Datensatz der nationalen, langen Perspektiven (1erLang)

| Epidemie           | DC12              | DC23              | DC34              | DC45              | H2/H1             | H3/H1             | H4/H1             | H5/H1             |
|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| SG                 | 1,38 (1,00; NA)   | 0,42 (0,28; NA)   |                   |                   | 1,72 (1,06; NA)   | 5,85 (2,00; NA)   |                   |                   |
| Beob. <sup>1</sup> | 2 aus 2           | 2 aus 2           |                   |                   | 2 aus 2           | 2 aus 2           |                   |                   |
| AG                 | 0,83 (0,71; NA)   | 0,50 (0,27; NA)   | 1,64 (1,14; NA)   | 0,86              | 0,88 (0,36; NA)   | 0,88 (0,86; NA)   | 0,15 (0,11; NA)   | 0,04              |
| Beob. <sup>1</sup> | 3 aus 3           | 3 aus 3           | 2 aus 3           | 1 aus 3           | 3 aus 3           | 3 aus 3           | 2 aus 3           | 1 aus 3           |
| RG2                | 0,12              |                   |                   |                   | 12,50             |                   |                   |                   |
| Beob. <sup>1</sup> | 1 aus 1           |                   |                   |                   | 1 aus 1           |                   |                   |                   |
| COV1               | 0,86 (0,39; NA)   | 0,19              |                   |                   | 1,50 (0,93; NA)   | 1,60              |                   |                   |
| Beob. <sup>1</sup> | 3 aus 4           | 1 aus 4           |                   |                   | 3 aus 4           | 1 aus 4           |                   |                   |
| SWG                | 0,80 (0,50; 1,25) | 0,69 (0,32; 0,91) | 0,75 (0,45; 1,11) | 0,80 (0,20; 1,22) | 1,60 (0,59; 5,00) | 2,39 (0,77; 4,92) | 1,03 (0,32; 1,86) | 0,28 (0,09; 0,56) |
| Beob. <sup>1</sup> | 31 aus 31         | 22 aus 31         | 13 aus 31         | 4 aus 31          | 31 aus 31         | 22 aus 31         | 13 aus 31         | 4 aus 31          |
| EB14               | 1,00 (1,00; NA)   | 1,13 (0,75; NA)   | 0,82 (0,44; NA)   | 1,07              | 0,83 (0,30; NA)   | 0,72 (0,44; NA)   | 1,58 (0,56; NA)   | 7,00              |
| Beob. <sup>1</sup> | 3 aus 3           | 2 aus 3           | 2 aus 3           | 1 aus 3           | 3 aus 3           | 2 aus 3           | 2 aus 3           | 1 aus 3           |
| COV2               | 0,51 (0,31; 0,85) | 0,94 (0,55; 1,24) | 0,64 (0,44; 0,95) | 0,92 (0,75; 1,24) | 1,52 (0,61; 3,57) | 1,99 (0,88; 6,74) | 2,06 (0,76; 7,36) | 1,98 (1,23; 8,69) |
| Beob. <sup>1</sup> | 65 aus 72         | 38 aus 72         | 22 aus 72         | 10 aus 72         | 65 aus 72         | 38 aus 72         | 22 aus 72         | 10 aus 72         |
| MdM <sup>2</sup>   | 0,84              | 0,50              | 0,82              | 0,86              | 1,55              | 1,60              | 1,03              | 0,28              |

Beob.<sup>1</sup>=Beobachtungen. MdM<sup>2</sup>=Median der Mediane ohne COV2. In dieser Tabelle sind die Mediane (Q1; Q3) dargestellt, sowie die Anzahl an Beobachtungen, auf denen die Angaben der Mediane basieren. Ausgangspunkt für die Angaben ist der Datensatz 1erLang, also die nationalen, langen Perspektiven, die aus dem INZDS ausgewählt wurden. NA bedeutet, dass ein Wert nicht ermittelt werden konnte. Der Median der Mediane bezieht sich auf die medianen Angaben zu DC und Wellenhöhen und umfasst nicht die Anzahl an Beobachtungen. Die Wellenhöhenangaben ergeben sich aus der jeweiligen Wellenhöhe geteilt durch die Höhe der korrespondierenden ersten Welle. Ablesbeispiel siehe Tabelle 2.

Tabelle 7: Median (Q1; Q3) im Datensatz der kombinierten regionalen und nationalen, langen Perspektiven (KombiLang)

| Epidemie           | DC12              | DC23              | DC34              | DC45              | H2/H1             | H3/H1             | H4/H1             | H5/H1             |
|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| SG                 | 0,75 (0,36; 1,50) | 0,76 (0,49; 2,37) |                   |                   | 1,06 (0,43; 2,37) | 1,31 (0,35; 5,41) |                   |                   |
| Beob. <sup>1</sup> | 7 aus 7           | 6 aus 7           |                   |                   | 7 aus 7           | 6 aus 7           |                   |                   |
| AG                 | 0,83 (0,71; NA)   | 0,50 (0,46; NA)   | 1,64 (1,14; NA)   | 0,86              | 0,36 (0,16; NA)   | 2,37 (1,00; NA)   | 0,18 (0,12; NA)   | 0,35              |
| Beob. <sup>1</sup> | 3 aus 3           | 3 aus 3           | 2 aus 3           | 1 aus 3           | 3 aus 3           | 3 aus 3           | 2 aus 3           | 1 aus 3           |
| RG2                | 0,31              |                   |                   |                   | 6,81              |                   |                   |                   |
| Beob. <sup>1</sup> | 2 aus 2           |                   |                   |                   | 2 aus 2           |                   |                   |                   |
| COV1               | 0,86 (0,39; NA)   | 0,19              |                   |                   | 1,50 (0,93; NA)   | 1,72              |                   |                   |
| Beob. <sup>1</sup> | 3 aus 4           | 1 aus 4           |                   |                   | 3 aus 4           | 1 aus 4           |                   |                   |
| SWG                | 0,86 (0,50; 1,30) | 0,70 (0,33; 1,00) | 0,81 (0,45; 1,16) | 0,85 (0,23; 1,22) | 1,40 (0,57; 5,00) | 1,94 (0,95; 4,00) | 0,93 (0,12; 1,45) | 0,24 (0,12; 1,08) |
| Beob. <sup>1</sup> | 34 aus 35         | 23 aus 35         | 14 aus 35         | 4 aus 35          | 34 aus 35         | 23 aus 35         | 14 aus 35         | 4 aus 35          |
| EB14               | 1,00 (1,00; NA)   | 1,12              | 0,82              | 1,07              | 0,83 (0,30; NA)   | 0,68 (0,53; NA)   | 1,92 (1,25; NA)   | 2,69              |
| Beob. <sup>1</sup> | 3 aus 3           | 2 aus 3           | 2 aus 3           | 1 aus 3           | 3 aus 3           | 2 aus 3           | 2 aus 3           | 1 aus 3           |
| COV2               | 0,51 (0,31; 0,85) | 0,94 (0,55; 1,24) | 0,64 (0,44; 0,95) | 0,92 (0,75; 1,24) | 1,52 (0,61; 3,57) | 1,33 (0,78; 3,26) | 1,66 (0,99; 3,14) | 1,20 (0,41; 1,66) |
| Beob. <sup>1</sup> | 65 aus 72         | 38 aus 72         | 22 aus 72         | 10 aus 72         | 65 aus 72         | 38 aus 72         | 22 aus 72         | 10 aus 72         |
| MdM <sup>2</sup>   | 0,84              | 0,70              | 0,82              | 0,86              | 1,23              | 1,72              | 0,93              | 0,35              |

Beob.<sup>1</sup>=Beobachtungen. MdM<sup>2</sup>=Median der Mediane ohne COV2. In dieser Tabelle sind die Mediane (Q1; Q3) dargestellt, sowie die Anzahl an Beobachtungen, auf denen die Angaben der Mediane basieren. Ausgangspunkt für die Angaben ist der Datensatz KombiLang, also die Kombination von regionalen und nationalen, langen Perspektiven, die aus dem INZDS ausgewählt wurden. NA bedeutet, dass ein Wert nicht ermittelt werden konnte. Der Median der Mediane bezieht sich auf die medianen Angaben zu DC und Wellenhöhen und umfasst nicht die Anzahl an Beobachtungen. Die Wellenhöhenangaben ergeben sich aus der jeweiligen Wellenhöhe geteilt durch die Höhe der korrespondierenden ersten Welle. Ablesebeispiel siehe Tabelle 2.

eingegangen war, konnte überprüft werden, wie nahe die historischen Mediane an die ermittelten Werte in COV2 herankamen. Die historische Ableitung und die Beobachtung in COV2 legten nahe, dass in der Praxis DC-Schätzwerte von ca. 0,50 bis etwa 1,00 verwendet werden könnten.

### 3. Wellenhöhen

Aus dem Datensatz 0erKurz ging ein gemischtes Bild hervor (Tabelle 2). Der Datensatz 1erKurz legte nahe, dass zweite Wellen in Pandemien höher waren als erste Wellen. Dritte Wellen schienen in Pandemien ebenfalls höher als erste Wellen zu sein (Tabelle 3). Der Datensatz KombiKurz zeigte erneut, dass zweite Wellen höher waren als erste Wellen. Auch dritte Wellen waren i. d. R. höher als erste Wellen (Tabelle 4).

Die Auswertung von 0erLang ergab, dass Zweitwellen höher waren als Erstwellen. Fünfte Wellen, so zeigte die Auswertung, waren i. d. R. niedriger als Erstwellen (Tabelle 5). Auch im Datensatz 1erLang waren Zweitwellen i. d. R. höher als Erstwellen, wobei EB14 hier die Ausnahme war; gleiches galt auch für Drittwellen. Fünfte Wellen waren in EB14 und COV2 deutlich höher als Erstwellen (Tabelle 6). Im Datensatz KombiLang waren ebenfalls die zweiten Wellen in Pandemien höher als die ersten Wellen (Tabelle 7).

Verglich man den MdM, so schienen kurze und lange Perspektiven überwiegend darin übereinzustimmen, dass zweite und dritte Wellen i. d. R. höher waren als erste Wellen, wohingegen vierte Wellen niedriger waren. Dieses Bild konnte man auch in COV2 erwarten. Allerdings waren die Folgewellen in COV2 tatsächlich stets höher als Erstwellen, weshalb sich die Erwartungen nur teilweise bestätigten. Für fünfte Wellen gab es in den kurzen Perspektiven keinen erwartbaren, historischen Wert, weshalb hier keine Überprüfung durchgeführt werden konnte. In den langen Perspektiven aber sollten fünfte Wellen stets niedriger sein als Erstwellen – diese Erwartung erfüllte COV2 nicht.

### IV. Vorhersage des zweiten Wellengipfels: Anteile der Streckenbezeichnungen, Güte und Präzision der Schätzung

Ausgangspunkt für die Vorhersage der Zeitpunkte der Zweitwellengipfel in COV2 war die Analyse der historischen Epidemien. Bei den Streckenbezeichnungen zeigten sich verschiedene Auffälligkeiten: So streute  $\mathbf{D}'$  über die verschiedenen Epidemien und Datensätze hinweg nur wenig. Es konnten hier keine Unterschiede zwischen kurzen, regio-



nen und nationalen Perspektiven festgestellt werden. Auch die Ähnlichkeit zu den langen Perspektiven war gegeben, weshalb eine Trennung in kurze und lange Perspektiven bzw. in regionale und nationale Perspektiven, zumindest bei der Verwendung des DC, unerheblich schien (Tabelle 8).

Während  $Z_A$  bei den kurzen Perspektiven durchaus stark streute, so war dies bei den langen Perspektiven nicht der Fall.  $Z_B$  streute bei den kurzen Perspektiven weniger und bei den langen Perspektiven noch weniger.  $Z_C$  streute in den langen Perspektiven am wenigsten. Während in den Datensätzen mit den langen Perspektiven die jeweiligen Z-Faktoren sehr ähnlich zu sein schienen, war dies in den kurzen Perspektiven nur in begrenztem Maße der Fall. Insgesamt deutete die abnehmende Streuung der Z-Faktoren darauf hin, dass längere Streckenanteile über die Datensätze hinweg einander ähnlicher wurden.

Insgesamt wurden 4.320 Beobachtungen berechnet, davon waren 4.272 auswertbar; 48 Beobachtungen konnten nicht ausgewertet werden, da für X im Rahmen der Messgenauigkeit kein sinnvoller Wert ermittelt werden konnte, diese also als „Not a Number“ (NaN) ausgegeben wurden ( $n = 12$ ) oder die Werte für X und Y gleich 0,00 und somit Median, IQA und  $R^2$  nicht bestimmbar waren. Alle diese nicht auswertbaren Varianten waren den Versuchsreihen 1erLang2019Kurz und 1erLang2019Lang, jeweils mit  $n = 24$  zuzuordnen.

Tabelle 8: Median der Mediane (MdM) der Anteile der Streckenbezeichnungen nach Datensatz

| Datensatz | Beob. | A'   | B'   | C'   | D'   | F'   | G'   | H'   | J'   | K'   | M'   |
|-----------|-------|------|------|------|------|------|------|------|------|------|------|
| 0erKurz   | 10    | 0,14 | 0,34 | 0,49 | 0,78 | 0,18 | 0,36 | 0,65 | 0,17 | 0,34 | 0,26 |
| 1erKurz   | 89    | 0,20 | 0,40 | 0,65 | 0,78 | 0,21 | 0,38 | 0,55 | 0,14 | 0,35 | 0,14 |
| KombiKurz | 91    | 0,11 | 0,41 | 0,55 | 0,78 | 0,18 | 0,37 | 0,62 | 0,11 | 0,33 | 0,17 |
| 0erLang   | 26    | 0,18 | 0,45 | 0,59 | 0,74 | 0,25 | 0,36 | 0,64 | 0,15 | 0,33 | 0,18 |
| 1erLang   | 108   | 0,19 | 0,44 | 0,58 | 0,78 | 0,22 | 0,34 | 0,51 | 0,00 | 0,26 | 0,19 |
| KombiLang | 117   | 0,18 | 0,41 | 0,56 | 0,76 | 0,23 | 0,40 | 0,54 | 0,07 | 0,27 | 0,19 |

MdM=Median der Mediane ohne COV2. Beob.=Anzahl der Beobachtungen, aus denen der MdM ermittelt wurde. Die Tabelle zeigt den Anteil der medianen Streckenbezeichnung, die aus den korrespondierenden medianen Werten der einzelnen Epidemien pro Datensatz ermittelt wurden. Die Streckenbezeichnungen wurden als Anteil der jeweiligen Streckenbezeichnung E ausgedrückt (Tabelle 1).

### 1. $R^2$ Adj. als Gütekriterium der Schätzung

Bei  $R^2$  als Gütekriterium der Schätzung zeigte sich bei 18 Varianten ein Wert von 0,80. Diese starken Zusammenhänge von Vorhersage und Beobachtung der Zweitwellengipfel

traten ausschließlich in „1erKurz2019Kurz“ auf und umfassten alle Grundrechenarten, aber nur Konstanten aus  $Z_C$ . Die absolute Abweichung von vorhergesagten und beobachteten Wellengipfeln lag zwischen 22 (13; 40) und 25 (11; 38) Tagen. Das bedeutete, dass der zweite Wellengipfel in 25% der Fälle mit einer Präzision von bis zu 13 Tagen, in 50% der Fälle mit bis zu 25 Tagen und in 75% der Fälle mit bis zu 40 Tagen genau vorhergesagt wurden. Auch die zugehörigen IQA waren schmal und reichten von 27 bis 29 Tagen (Tabelle 9).

Insgesamt hatten 869 der 4.272 Beobachtungen ein  $R^2$  von  $\geq 0,75$ . 1.046 Beobachtungen hatten ein  $R^2$  von  $\geq 0,70$ , 1.799 von  $\geq 0,60$ , 2.080 von  $\geq 0,50$ . In 458 Beobachtungen war  $R^2 < 0,10$ . In  $Z_A$  war der DC mit dem größten  $R^2$  in der Versuchsreihe „1erLang2019Kurz“ mit  $R^2$  von 0,13 einer medianen Tagesabweichung von 80 (45; 149) und auf Rang 3.551. In  $Z_B$  war der DC mit dem größten  $R^2$  (0,29) in der Versuchsreihe „0erLang2019Kurz“ und hatte eine mediane Tagesabweichung von 61 (23; 121). In  $Z_C$  war der DC mit dem stärksten  $R^2$  (0,65) in der gleichen Versuchsreihe und hatte eine mediane Abweichung in Tagen von 26 (12; 55). Bereits hier ließen sich zwei Aspekte festhalten:

1. Die Berechnung eines linearen Zusammenhangs von beobachteten und berechneten Zweitwellengipfeln in COV2 funktionierte besonders gut in den kurzen Coronaperspektiven, allerdings waren die langen Coronaperspektiven nur in geringem Maße unterlegen.
2. Das größte  $R^2$  deutete nicht unbedingt auf die geringste Abweichung in Tagen hin.

## 2. Die Absolute Abweichung als Gütekriterium der Schätzung

Aus Abbildung 3 geht hervor, dass die Präzision der Vorhersage von  $Z_A$  nach  $Z_C$  zunahm. Dies war nicht überraschend, da von  $Z_A$  nach  $Z_C$  anteilig mehr Zeit berücksichtigt wurde und somit der Bereich, in dem sich der wahre Wert des Zweitwellengipfels befand, ständig kleiner wurde. Die Präzision nahm also erwartbar zu, da es schlicht weniger Möglichkeiten für falsche Werte gab.

Die Verfahren mit den laufenden Nummern 16, 20, 26 und 27 fielen durch ihre, im Vergleich zu den anderen Verfahren, starke Streuung bei der Vorhersage auf. Gemeinsam hatten diese Verfahren mit mangelnder Präzision v. a., dass die Teilstrecke **D** verwendet wurde. Zudem fiel auf, dass entgegen des sonst beobachteten Trends die Präzision der Vorhersage von  $Z_A$  nach  $Z_C$  abnahm.

Tabelle 9: Die präzisesten Schätzungen für die Zweitwellengipfel mit  $R^2$  Adj. als Gütekriterium

| Versuchsreihe   | Variante | Formel                 | X     | Y    | Z    | Absolute Abweichung |    |    |     |         | Rang |
|-----------------|----------|------------------------|-------|------|------|---------------------|----|----|-----|---------|------|
|                 |          |                        |       |      |      | Median              | Q1 | Q3 | IQA | R2 Adj. |      |
| 1erKurz2019Kurz | * $Z_C$  | $e=((d-a)XZ+aY)/Y$     | 0,98  | 0,20 | 1,54 | 22                  | 13 | 40 | 28  | 0,80    | 1    |
| 1erKurz2019Kurz | / $Z_C$  | $e=((d-a)YZ+aX)/X$     | 1,18  | 0,40 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 2    |
| 1erKurz2019Kurz | * $Z_C$  | $e=((d-a)XZ+aY)/Y$     | 1,43  | 0,65 | 1,54 | 25                  | 11 | 38 | 27  | 0,80    | 3    |
| 1erKurz2019Kurz | + $Z_C$  | $e=((a-d)Y+(d-a)X)Z+a$ | 0,99  | 0,21 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 4    |
| 1erKurz2019Kurz | + $Z_C$  | $e=((a-d)Y+(d-a)X)Z+a$ | 1,16  | 0,38 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 5    |
| 1erKurz2019Kurz | + $Z_C$  | $e=((a-d)Y+(d-a)X)Z+a$ | 0,92  | 0,14 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 6    |
| 1erKurz2019Kurz | + $Z_C$  | $e=((a-d)Y+(d-a)X)Z+a$ | -0,58 | 0,20 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 7    |
| 1erKurz2019Kurz | + $Z_C$  | $e=((a-d)Y+(d-a)X)Z+a$ | -0,38 | 0,40 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 8    |
| 1erKurz2019Kurz | + $Z_C$  | $e=((a-d)Y+(d-a)X)Z+a$ | -0,13 | 0,65 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 9    |
| 1erKurz2019Kurz | - $Z_C$  | $e=((d-a)Y+(a-d)X)Z+a$ | 0,31  | 0,40 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 10   |
| 1erKurz2019Kurz | - $Z_C$  | $e=((d-a)Y+(a-d)X)Z+a$ | 0,51  | 0,65 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 11   |
| 1erKurz2019Kurz | - $Z_C$  | $e=((d-a)Y+(a-d)X)Z+a$ | 0,09  | 0,65 | 1,54 | 24                  | 11 | 38 | 27  | 0,80    | 12   |
| 1erKurz2019Kurz | * $Z_C$  | $e=((d-a)XZ+dY)/Y$     | 0,16  | 0,21 | 1,54 | 23                  | 11 | 38 | 27  | 0,80    | 13   |
| 1erKurz2019Kurz | / $Z_C$  | $e=((d-a)YZ+aX)/X$     | 0,11  | 0,14 | 1,54 | 24                  | 11 | 39 | 28  | 0,80    | 14   |
| 1erKurz2019Kurz | / $Z_C$  | $e=((d-a)YZ+aX)/X$     | 0,05  | 0,38 | 1,54 | 24                  | 11 | 39 | 28  | 0,80    | 15   |
| 1erKurz2019Kurz | * $Z_C$  | $e=((d-a)XZ+aY)/Y$     | 0,26  | 0,20 | 1,54 | 24                  | 11 | 39 | 28  | 0,80    | 16   |
| 1erKurz2019Kurz | * $Z_C$  | $e=((d-a)XZ+aY)/Y$     | 0,51  | 0,40 | 1,54 | 24                  | 11 | 39 | 29  | 0,80    | 17   |
| 1erKurz2019Kurz | * $Z_C$  | $e=((d-a)XZ+dY)/Y$     | 0,83  | 0,65 | 1,54 | 24                  | 11 | 40 | 29  | 0,80    | 18   |

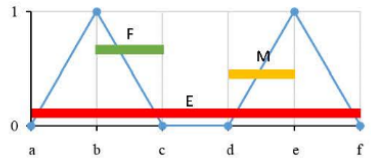
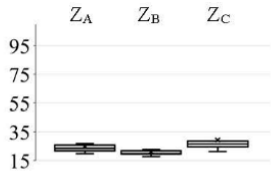
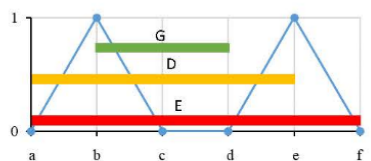
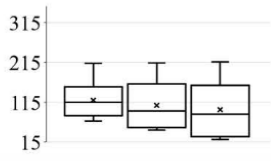
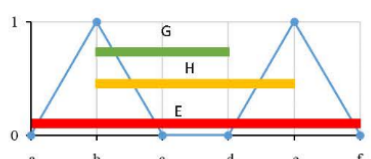
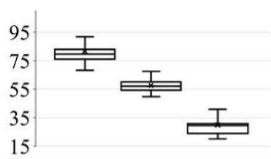
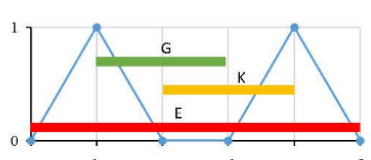
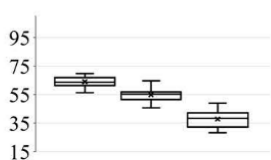
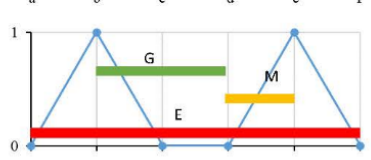
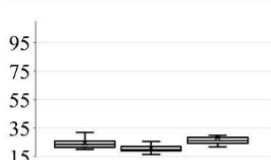
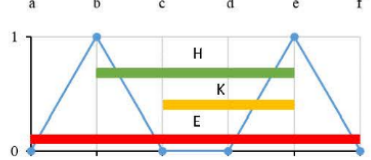
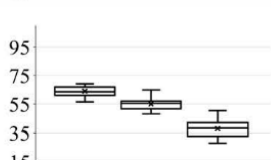
Die Tabelle zeigt, in absteigender Reihenfolge angeordnet, die präzisesten Schätzungen der Zweitwellengipfel unter der Verwendung von  $R^2$  Adj. als Gütekriterium der Schätzung. Aufgeführt sind die Ergebnisse für  $R^2$  Adj., die bei der Rundung auf zwei Nachkommastellen den maximalen beobachteten Wert von 0,80 erreichten. Angegeben ist zudem die Versuchsreihe, die Variante, die Formel, sowie die Schätzparameter für X, Y und Z unter denen das  $R^2$  Adj. beobachtet wurde. Zudem ist die jeweilige Abweichung in Tagen angegeben. Die Absolute Abweichung ist der Betragswert der Differenz von Beobachtung zu Prädiktion in Tagen.  $R^2$  Adj. ist die aufgeklärte Varianz von dem Zahlenwert der beobachteten Tage, wobei der erste Tag der 01.01.1900 ist, und dem Zahlenwert der jeweils vorhergesagten Tage der Zweitwellengipfel.

| Laufende Nummer | Abbildung | Involvierte Ausdrücke  | Transformationen   | Betragswerte der medianen Abweichung in Tagen (Vorhersage zu Beobachtung) der Varianten $Z_{A,B,C}$ |
|-----------------|-----------|--|--|---|
| 1               |           | $A' = (b-a)/E$<br>$D' = (e-a)/E$<br>$E' = (f-a)/1 =$<br>1. $(b-a) * (Z_A)$<br>2. $(e-a) * (Z_B)$<br>3. $(d-a) * (Z_C)$ | Addition: $X = A' + D' = A' + ((e-a)/E') = A' + ((e-a)/((f-a)/1)) = A' + ((e-a)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = A' - D' = A' - ((e-a)/E') = A' - ((e-a)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = A' * D' = A' + ((e-a)/E') = A' * ((e-a)/b,c,d-a) * Z_{ABC}$<br>Division: $X = A' / D' = A' / ((e-a)/E') = A' / ((e-a)/b,c,d-a) * Z_{ABC}$ |   |
| 2               |           | $A' = (b-a)/E$<br>$H' = (e-b)/E$<br>$E' = (f-a)/1 =$<br>1. $(b-a) * (Z_A)$<br>2. $(e-a) * (Z_B)$<br>3. $(d-a) * (Z_C)$ | Addition: $X = A' + H' = \dots = A' + ((e-b)/((f-a)/1)) * Z_{ABC}$<br>Subtraktion: $X = A' - H' = \dots = A' - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = A' * H' = \dots = A' * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = A' / H' = \dots = A' / ((e-b)/b,c,d-a) * Z_{ABC}$  |   |
| 3               |           | $A' = (b-a)/E$<br>$K' = (e-c)/E$<br>$E' = (f-a)/1 =$<br>1. $(b-a) * (Z_A)$<br>2. $(e-a) * (Z_B)$<br>3. $(d-a) * (Z_C)$ | Addition: $X = A' + K' = \dots = A' + ((e-c)/((f-a)/1)) * Z_{ABC}$<br>Subtraktion: $X = A' - K' = \dots = A' - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = A' * K' = \dots = A' * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = A' / K' = \dots = A' / ((e-c)/b,c,d-a) * Z_{ABC}$  |   |
| 4               |           | $A' = (b-a)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$<br>1. $(b-a) * (Z_A)$<br>2. $(e-a) * (Z_B)$<br>3. $(d-a) * (Z_C)$ | Addition: $X = A' + M' = \dots = A' + ((e-d)/((f-a)/1)) * Z_{ABC}$<br>Subtraktion: $X = A' - M' = \dots = A' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = A' * M' = \dots = A' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = A' / M' = \dots = A' / ((e-d)/b,c,d-a) * Z_{ABC}$  |   |
| 5               |           | $B' = (c-a)/E$<br>$D' = (e-a)/E$<br>$E' = (f-a)/1 =$<br>1. $(b-a) * (Z_A)$<br>2. $(e-a) * (Z_B)$<br>3. $(d-a) * (Z_C)$ | Addition: $X = B' + D' = \dots = B' + ((e-a)/((f-a)/1)) * Z_{ABC}$<br>Subtraktion: $X = B' - D' = \dots = B' - ((e-a)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = B' * D' = \dots = B' * ((e-a)/b,c,d-a) * Z_{ABC}$<br>Division: $X = B' / D' = \dots = B' / ((e-a)/b,c,d-a) * Z_{ABC}$  |   |
| 6               |           | $B' = (c-a)/E$<br>$H' = (e-b)/E$<br>$E' = (f-a)/1 =$<br>1. $(b-a) * (Z_A)$<br>2. $(e-a) * (Z_B)$<br>3. $(d-a) * (Z_C)$ | Addition: $X = B' + H' = \dots = B' + ((e-b)/((f-a)/1)) * Z_{ABC}$<br>Subtraktion: $X = B' - H' = \dots = B' - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = B' * H' = \dots = B' * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = B' / H' = \dots = B' / ((e-b)/b,c,d-a) * Z_{ABC}$  |   |

| Laufende Nummer | Abbildung | Involvierte Ausdrücke   | Transformationen  | Betragswerte der medianen Abweichung in Tagen (Vorhersage zu Beobachtung) der Varianten $Z_{A,B,C}$ |
|-----------------|-----------|---|---|---|
| 7               |           | $B' = (c-a)/E$<br>$K' = (e-c)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = B' + K' = \dots = B' + ((e-c)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = B' - K' = \dots = B' - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = B' * K' = \dots = B' * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = B' / K' = \dots = B' / ((e-c)/b,c,d-a) * Z_{ABC}$ |   |
| 8               |           | $B' = (c-a)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = B' + M' = \dots = B' + ((e-d)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = B' - M' = \dots = B' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = B' * M' = \dots = B' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = B' / M' = \dots = B' / ((e-d)/b,c,d-a) * Z_{ABC}$ |   |
| 9               |           | $C' = (d-a)/E$<br>$D' = (e-a)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = C' + D' = \dots = C' + ((e-a)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = C' - D' = \dots = C' - ((e-a)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = C' * D' = \dots = C' * ((e-a)/b,c,d-a) * Z_{ABC}$<br>Division: $X = C' / D' = \dots = C' / ((e-a)/b,c,d-a) * Z_{ABC}$ |   |
| 10              |           | $C' = (d-a)/E$<br>$H' = (e-b)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = C' + H' = \dots = A' + ((e-b)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = C' - H' = \dots = A' - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = C' * H' = \dots = A' * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = C' / H' = \dots = A' / ((e-b)/b,c,d-a) * Z_{ABC}$ |   |
| 11              |           | $C' = (d-a)/E$<br>$K' = (e-c)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = C' + K' = \dots = A' + ((e-c)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = C' - K' = \dots = A' - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = C' * K' = \dots = A' * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = C' / K' = \dots = C' / ((e-c)/b,c,d-a) * Z_{ABC}$ |   |
| 12              |           | $C' = (d-a)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = C' + M' = \dots = C' + ((e-d)/(b,c,d-a)) * Z_{ABC}$<br>Subtraktion: $X = C' - M' = \dots = C' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = C' * M' = \dots = C' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = C' / M' = \dots = C' / ((e-d)/b,c,d-a) * Z_{ABC}$ |   |

| Laufende Nummer | Abbildung | Involvierte Ausdrücke   | Transformationen  | Betragswerte der medianen Abweichung in Tagen (Vorhersage zu Beobachtung) der Varianten $Z_{A,B,C}$ |
|-----------------|-----------|---|---|---|
| 13              |           | $D' = (e-a)/E$<br>$H' = (e-b)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = D' + H' = \dots = D' + ((e-b)/((b,c,d-a)) * Z_{ABC})$<br>Subtraktion: $X = D' - H' = \dots = D' - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = D' * H' = \dots = D' * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = D' / H' = \dots = D' / ((e-b)/b,c,d-a) * Z_{ABC}$ |   |
| 14              |           | $D' = (e-a)/E$<br>$K' = (e-c)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = D' + K' = \dots = K' + ((e-c)/((b,c,d-a)) * Z_{ABC})$<br>Subtraktion: $X = D' - K' = \dots = K' - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = D' * K' = \dots = D' * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = D' / K' = \dots = D' / ((e-c)/b,c,d-a) * Z_{ABC}$ |   |
| 15              |           | $D' = (e-a)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = D' + M' = \dots = D' + ((e-d)/((b,c,d-a)) * Z_{ABC})$<br>Subtraktion: $X = D' - M' = \dots = D' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = D' * M' = \dots = D' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = D' / M' = \dots = D' / ((e-d)/b,c,d-a) * Z_{ABC}$ |   |
| 16              |           | $F' = (c-b)/E$<br>$D' = (e-a)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = F' + D' = \dots = F' + ((e-a)/((b,c,d-a)) * Z_{ABC})$<br>Subtraktion: $X = F' - D' = \dots = F' - ((e-a)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = F' * D' = \dots = F' * ((e-a)/b,c,d-a) * Z_{ABC}$<br>Division: $X = F' / D' = \dots = F' / ((e-a)/b,c,d-a) * Z_{ABC}$ |   |
| 17              |           | $F' = (c-b)/E$<br>$H' = (e-b)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = F' + H' = \dots = F' + ((e-b)/((b,c,d-a)) * Z_{ABC})$<br>Subtraktion: $X = F' - H' = \dots = F' - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = F' * H' = \dots = F' * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = F' / H' = \dots = F' / ((e-b)/b,c,d-a) * Z_{ABC}$ |   |
| 18              |           | $F' = (c-b)/E$<br>$K' = (e-c)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = F' + K' = \dots = F' + ((e-c)/((b,c,d-a)) * Z_{ABC})$<br>Subtraktion: $X = F' - K' = \dots = F' - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = F' * K' = \dots = F' * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = F' / K' = \dots = F' / ((e-c)/b,c,d-a) * Z_{ABC}$ |   |



| Laufende Nummer | Abbildung   | Involvierte Ausdrücke  | Transformationen   | Betragswerte der medianen Abweichung in Tagen (Vorhersage zu Beobachtung) der Varianten $Z_{A,B,C}$ |
|-----------------|---|--|--|---|
| 19              |    | $F^* = (c-b)/E$<br>$M^* = (e-d)/E$<br>$E^* = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = F^* + M^* = \dots = F^* + ((e-d)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = F^* - M^* = \dots = F^* - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = F^* * M^* = \dots = F^* * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = F^* / M^* = \dots = F^* / ((e-d)/b,c,d-a) * Z_{ABC}$ |                  |
| 20              |    | $G^* = (d-b)/E$<br>$D^* = (e-a)/E$<br>$E^* = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = G^* + D^* = \dots = G^* + ((e-a)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = G^* - D^* = \dots = G^* - ((e-a)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = G^* * D^* = \dots = G^* * ((e-a)/b,c,d-a) * Z_{ABC}$<br>Division: $X = G^* / D^* = \dots = G^* / ((e-a)/b,c,d-a) * Z_{ABC}$ |                  |
| 21              |    | $G^* = (d-b)/E$<br>$H^* = (e-b)/E$<br>$E^* = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = G^* + H^* = \dots = G^* + ((e-b)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = G^* - H^* = \dots = G^* - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = G^* * H^* = \dots = G^* * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = G^* / H^* = \dots = G^* / ((e-b)/b,c,d-a) * Z_{ABC}$ |                  |
| 22              |    | $G^* = (d-b)/E$<br>$K^* = (e-c)/E$<br>$E^* = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = G^* + K^* = \dots = G^* + ((e-c)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = G^* - K^* = \dots = G^* - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = G^* * K^* = \dots = G^* * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = G^* / K^* = \dots = G^* / ((e-c)/b,c,d-a) * Z_{ABC}$ |                  |
| 23              |   | $G^* = (d-b)/E$<br>$M^* = (e-d)/E$<br>$E^* = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = G^* + M^* = \dots = G^* + ((e-d)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = G^* - M^* = \dots = G^* - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = G^* * M^* = \dots = G^* * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = G^* / M^* = \dots = G^* / ((e-d)/b,c,d-a) * Z_{ABC}$ |                 |
| 24              |  | $H^* = (b-e)/E$<br>$K^* = (e-c)/E$<br>$E^* = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = H^* + K^* = \dots = H^* + ((e-c)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = H^* - K^* = \dots = H^* - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = H^* * K^* = \dots = H^* * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = H^* / K^* = \dots = H^* / ((e-c)/b,c,d-a) * Z_{ABC}$ |                |

| Laufende Nummer | Abbildung | Involvierte Ausdrücke   | Transformationen   | Betragswerte der medianen Abweichung in Tagen (Vorhersage zu Beobachtung) der Varianten $Z_{A,B,C}$ |
|-----------------|-----------|---|--|---|
| 25              |           | $H' = (b-e)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = H' + M' = \dots = H' + ((e-d)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = H' - M' = \dots = H' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = H' * M' = \dots = H' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = H' / M' = \dots = H' / ((e-d)/b,c,d-a) * Z_{ABC}$ |   |
| 26              |           | $J' = (d-c)/E$<br>$D' = (e-a)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = J' + D' = \dots = J' + ((e-a)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = J' - D' = \dots = J' - ((e-a)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = J' * D' = \dots = J' * ((e-a)/b,c,d-a) * Z_{ABC}$<br>Division: $X = J' / D' = \dots = J' / ((e-a)/b,c,d-a) * Z_{ABC}$ |   |
| 27              |           | $J' = (d-c)/E$<br>$H' = (e-b)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = J' + H' = \dots = J' + ((e-b)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = J' - H' = \dots = J' - ((e-b)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = J' * H' = \dots = J' * ((e-b)/b,c,d-a) * Z_{ABC}$<br>Division: $X = J' / H' = \dots = J' / ((e-b)/b,c,d-a) * Z_{ABC}$ |   |
| 28              |           | $J' = (d-c)/E$<br>$K' = (e-c)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = J' + K' = \dots = J' + ((e-c)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = J' - K' = \dots = J' - ((e-c)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = J' * K' = \dots = J' * ((e-c)/b,c,d-a) * Z_{ABC}$<br>Division: $X = J' / K' = \dots = J' / ((e-c)/b,c,d-a) * Z_{ABC}$ |   |
| 29              |           | $J' = (d-c)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = J' + M' = \dots = J' + ((e-d)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = J' - M' = \dots = J' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = J' * M' = \dots = J' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = J' / M' = \dots = J' / ((e-d)/b,c,d-a) * Z_{ABC}$ |   |
| 30              |           | $K' = (e-c)/E$<br>$M' = (e-d)/E$<br>$E' = (f-a)/1 =$ <ol style="list-style-type: none"> <li><math>(b-a) * (Z_A)</math></li> <li><math>(c-a) * (Z_B)</math></li> <li><math>(d-a) * (Z_C)</math></li> </ol> | Addition: $X = K' + M' = \dots = K' + ((e-d)/((b,c,d-a) * Z_{ABC})$<br>Subtraktion: $X = K' - M' = \dots = K' - ((e-d)/b,c,d-a) * Z_{ABC}$<br>Multiplikation: $X = K' * M' = \dots = K' * ((e-d)/b,c,d-a) * Z_{ABC}$<br>Division: $X = K' / M' = \dots = K' / ((e-d)/b,c,d-a) * Z_{ABC}$ |   |



Abbildung 3: Abweichung der Vorhersage der medianen absoluten Abweichung in Tagen von dem beobachteten Zweitwellengipfel pro Streckenkombination

Dargestellt sind die 30 Streckenkombinationen, die sich für die Vorhersage des Zweitwellengipfels sinnvoll aus Abbildung 1 ableiten lassen. Die dargestellten Boxplots stellen die Betragswerte der medianen Abweichung in Tagen von Vorhersage und Beobachtung pro Z Faktor dar. Die Spalte mit den Transformationen zeigt, welche Formeln in den Boxplots enthalten sind.

Die Verfahren mit den laufenden Nummern 1, 2, 3, 5, 6, 7, 9, 10, 11, 13, 14, 17, 18, 21, 22, 24 und 28 ähnelten sich dahingehend, dass die Präzision der Vorhersage von  $Z_A$  nach  $Z_C$  zunahm. Zu diesen Verfahren gehörte auch die laufende Nummer 6, die den DC beinhaltete. Bereits an dieser Stelle war klar, dass der DC nicht das präziseste Verfahren war. Obwohl der  $D_C$  in  $Z_C$  mit den präziseren Verfahren vergleichbar war, gab es Verfahren, die von  $Z_A$  bis  $Z_C$  eine recht konstante Vorhersagequalität hatten und in etwa der Präzision des DC in  $Z_C$  entsprachen.

Diese Verfahren hatten die laufenden Nummer 4, 8, 12, 15, 19, 23, 25, 29 und 30. Gemeinsam war diesen Verfahren, dass sie alle die Teilstrecke M verwendeten und in  $Z_B$  die medianen Vorhersagen nicht nur am präzisesten waren, sondern auch die Streuung am geringsten war. Zusammenfassend boten sich diese Verfahren also für die Verwendung in der epidemiologischen Praxis an, da die medianen Ergebnisse die präzisesten waren, die Ergebnisse im geringsten Umfang streuten und auch die Präzision der Vorhersage in  $Z_A$ , also frühzeitig, anderen Verfahren überlegen war.

Bezog man auch noch die einzelnen Transformationen in die Bewertung der Präzision mit ein, so konnte man die mediane und die Abweichung in Q3 als Gütekriterium ansetzen. Interessanterweise bedeutete ein niedriger Median nicht unbedingt einen niedrigen Wert in Q3. So war die präziseste mediane Schätzung in der Versuchsreihe „1er-Kurz2019Lang“ in der Variante  $Z_B$  und einem  $R^2$  von 0,76 bei 17 (8; 43) Tagen. Das in Q3 präziseste Verfahren war aus der Versuchsreihe „KombiLang2019Lang“ in der Variante  $Z_C$ , hatte ein  $R^2$  von 0,78 und eine Abweichung von 25 (12; 35) Tagen.

Der in Q3 präziseste DC für die Variante  $Z_A$  war aus der Versuchsreihe „1er-Lang2019Lang“, hatte ein  $R^2$  von 0,08 und eine Abweichung von 77 (43; 131) Tagen. Der in Q3 präziseste DC für die Variante  $Z_B$  war aus der Versuchsreihe „0er-Lang2019Kurz“, hatte ein  $R^2$  von 0,27 und eine Abweichung von 57 (27; 105) Tagen.

Der in Q3 präziseste DC für die Variante  $Z_C$  war aus der Versuchsreihe „KombiLang2019Lang“, hatte ein  $R^2$  von 0,64 und eine Abweichung von 20 (7; 48) Tagen (Tabelle 10).

Zog man das IQA als Kriterium der Schätzgüte heran, so zeigte sich, dass es 446 Beobachtungen gab, bei denen das  $IQA \leq 30$  Tage war. Die besten 83 Versuchsreihen hatten dabei ein  $R^2$  von 0,77. Alle 83 waren aus der Variante  $Z_C$ , und die besten zehn aus der Versuchsreihe „1erLang2019Lang“.

Auch bei den DC gab es schmale und breite IQA, wobei die engsten IQA in der  $Z_C$ -Variante beobachtet wurden. Jeweils pro Variante waren die präzisesten Vertreter:

1. für  $Z_C$  aus der Versuchsreihe „KombiKurz2019Lang“ mit  $R^2$  von 0,65, einer Abweichung von 27 (15;51) und einem IQA von 36 Tagen,
2. für  $Z_B$  aus der Versuchsreihe „0erLang2019Kurz“ mit einem  $R^2$  von 0,27, einer Abweichung von 57 (27; 105) und einem IQA von 78 Tagen und
3. für  $Z_A$  aus der Versuchsreihe „1erLang2019Lang“ mit einem  $R^2$  von 0,08, einer Abweichung von 77 (43; 131) und einem IQA von 88 Tagen.

Bislang zeigte sich, dass man mit den aus den historischen Perspektiven abgeleiteten Schätzern gute Vorhersagen in den verschiedenen Perspektivenformen (kurz, lang, kombiniert) berechnen konnte.

## V. Passende historische Epidemie anhand quantitativer Merkmale

Tabelle 11 fasst die Ergebnisse der Regressionsanalysen zusammen und zeigt das jeweilige  $R^2$  für den Zusammenhang der COV2-Werte des jeweiligen Datensatzes mit den Werten der jeweiligen Epidemie. Es zeigte sich, dass EB14 nicht zu COV2 passte. Auch RG2 passte in den „1er-Datensätzen“ nicht zu COV2. In den kombinierten Datensätzen ließ sich eine Ähnlichkeit feststellen. Die Ähnlichkeit von COV2 und COV1 war in den kurzen Datensätzen vorhanden, allerdings zeigte sich der Zusammenhang in den langen Datensätzen nicht. SG hatte in den verschiedenen Datensätzen, bis auf 1erLang, stets starke Zusammenhänge mit COV2. Über alle Datensätze hinweg hatte SWG stets starke Zusammenhänge mit COV2. Insgesamt war der Schluss möglich, dass SWG COV2 besonders ähnlich war. Auffällig war auch die Ähnlichkeit von SWG zu SG, was vermuten ließ, dass hier ein pandemischer Mechanismus beobachtbar war; diese Vermutung konnte wegen des deutlichen Unterschieds zu EB14 gestützt werden.

Tabelle 10: Duty Cycle geordnet nach Abweichung in Tagen in Q3

| Versuchsreihe     | Variante        | Formel             | X    | Y    | Z    | Absolute Abweichung |    |     |     | R <sup>2</sup> Adj. |
|-------------------|-----------------|--------------------|------|------|------|---------------------|----|-----|-----|---------------------|
|                   |                 |                    |      |      |      | Median              | Q1 | Q3  | IQA |                     |
| KombiLang2019Lang | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,76 | 0,41 | 1,79 | 20                  | 7  | 48  | 42  | 0,64                |
| KombiKurz2019Lang | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,66 | 0,41 | 1,82 | 27                  | 15 | 51  | 36  | 0,65                |
| KombiKurz2019Kurz | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,66 | 0,41 | 1,82 | 31                  | 15 | 52  | 37  | 0,65                |
| 0erLang2019Lang   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,70 | 0,45 | 1,69 | 23                  | 11 | 54  | 43  | 0,65                |
| 0erLang2019Kurz   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,70 | 0,45 | 1,69 | 26                  | 12 | 55  | 43  | 0,65                |
| KombiLang2019Kurz | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,76 | 0,41 | 1,79 | 21                  | 7  | 55  | 48  | -0,01               |
| 1erLang2019Lang   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,86 | 0,44 | 1,72 | 28                  | 10 | 58  | 48  | 0,63                |
| 1erLang2019Kurz   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,86 | 0,44 | 1,72 | 29                  | 11 | 58  | 48  | 0,64                |
| 1erKurz2019Lang   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,73 | 0,40 | 1,54 | 31                  | 14 | 60  | 46  | 0,62                |
| 1erKurz2019Kurz   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,73 | 0,40 | 1,54 | 31                  | 14 | 61  | 48  | 0,63                |
| 0erKurz2019Lang   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,52 | 0,34 | 2,04 | 41                  | 16 | 76  | 61  | 0,64                |
| 0erKurz2019Kurz   | /Z <sub>C</sub> | $e=((d-a)YZ+bX)/X$ | 0,52 | 0,34 | 2,04 | 52                  | 20 | 81  | 60  | 0,64                |
| 0erLang2019Kurz   | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,70 | 0,45 | 2,22 | 57                  | 27 | 105 | 78  | 0,27                |
| KombiKurz2019Lang | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,66 | 0,41 | 2,44 | 54                  | 26 | 106 | 80  | 0,18                |
| 1erKurz2019Kurz   | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,73 | 0,40 | 2,50 | 60                  | 29 | 108 | 79  | 0,28                |
| 0erLang2019Lang   | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,70 | 0,45 | 2,22 | 55                  | 25 | 108 | 83  | 0,19                |
| KombiKurz2019Kurz | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,66 | 0,41 | 2,44 | 59                  | 28 | 109 | 81  | 0,27                |
| 1erKurz2019Lang   | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,73 | 0,40 | 2,50 | 53                  | 22 | 110 | 88  | 0,19                |
| KombiLang2019Kurz | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,76 | 0,41 | 2,44 | 57                  | 28 | 110 | 82  | 0,28                |
| KombiLang2019Lang | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,76 | 0,41 | 2,44 | 55                  | 19 | 112 | 92  | 0,19                |
| 0erKurz2019Lang   | /Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,52 | 0,34 | 2,94 | 52                  | 25 | 115 | 90  | 0,16                |

| Versuchsreihe     | Variante         | Formel             | X    | Y    | Z    | Absolute Abweichung |    |     |     | R <sup>2</sup> Adj. |
|-------------------|------------------|--------------------|------|------|------|---------------------|----|-----|-----|---------------------|
|                   |                  |                    |      |      |      | Median              | Q1 | Q3  | IQA |                     |
| 1erLang2019Lang   | / Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,86 | 0,44 | 2,27 | 60                  | 22 | 118 | 96  | 0,20                |
| 1erLang2019Kurz   | / Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,86 | 0,44 | 2,27 | 61                  | 23 | 121 | 99  | 0,29                |
| 0erKurz2019Kurz   | / Z <sub>B</sub> | $e=((c-a)YZ+bX)/X$ | 0,52 | 0,34 | 2,94 | 67                  | 35 | 124 | 90  | 0,25                |
| 1erLang2019Lang   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,86 | 0,44 | 5,26 | 77                  | 43 | 131 | 88  | 0,08                |
| KombiLang2019Lang | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,76 | 0,41 | 5,56 | 76                  | 38 | 140 | 102 | 0,07                |
| 1erKurz2019Lang   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,73 | 0,40 | 5,00 | 76                  | 39 | 141 | 103 | 0,07                |
| 1erLang2019Kurz   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,86 | 0,44 | 5,26 | 80                  | 45 | 149 | 105 | 0,13                |
| 0erLang2019Lang   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,70 | 0,45 | 5,56 | 80                  | 37 | 152 | 116 | 0,06                |
| 1erKurz2019Kurz   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,73 | 0,40 | 5,00 | 82                  | 40 | 159 | 120 | 0,12                |
| KombiLang2019Kurz | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,76 | 0,41 | 5,56 | 82                  | 39 | 160 | 120 | 0,12                |
| 0erKurz2019Lang   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,52 | 0,34 | 7,14 | 70                  | 24 | 177 | 152 | 0,04                |
| 0erLang2019Kurz   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,70 | 0,45 | 5,56 | 83                  | 43 | 182 | 139 | 0,10                |
| KombiKurz2019Lang | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,66 | 0,41 | 9,09 | 73                  | 28 | 202 | 175 | 0,03                |
| 0erKurz2019Kurz   | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,52 | 0,34 | 7,14 | 91                  | 47 | 211 | 164 | 0,08                |
| KombiKurz2019Kurz | / Z <sub>A</sub> | $e=((b-a)YZ+bX)/X$ | 0,66 | 0,41 | 9,09 | 111                 | 29 | 236 | 207 | 0,07                |

Die Tabelle zeigt, in absteigender Reihenfolge angeordnet, die präzisesten Schätzungen der Zweitwellengipfel unter der Verwendung der absoluten Abweichung (Betragswert der Differenz von Beobachtung zu Prädiktion in Tagen) in Q3 als Gütekriterium der Schätzung. Angegeben ist die Versuchsreihe, die Variante, die Formel, sowie die Schätzparameter für X, Y und Z unter denen die Absolute Abweichung beobachtet wurde. R<sup>2</sup> Adj. ist die aufgeklärte Varianz von dem Zahlenwert der beobachteten Tage, wobei der erste Tag der 01.01.1900 ist, und dem Zahlenwert der jeweils vorhergesagten Tage der Zweitwellengipfel.

Tabelle 11: Ergebnisse der Regressionsanalysen für den Zusammenhang der COV2-Werte des jeweiligen Datensatzes mit den anderen Epidemien des jeweiligen Datensatzes

| Datensatz | Epidemie | R <sup>2</sup> Adj. |
|-----------|----------|---------------------|
| 1erKurz   | SG       | 0,92                |
|           | AG       | 0,67                |
|           | RF2      | 0,56                |
|           | COV1     | 0,95                |
|           | SWG      | 0,94                |
|           | EB14     | 0,38                |
| 1erLang   | SG       | 0,35                |
|           | AG       | 0,68                |
|           | RF2      | 0,58                |
|           | COV1     | 0,67                |
|           | SWG      | 0,84                |
|           | EB14     | 0,65                |
| KombiKurz | SG       | 0,92                |
|           | AG       | 0,70                |
|           | RF2      | 0,73                |
|           | COV1     | 0,87                |
|           | SWG      | 0,92                |
|           | EB14     | 0,38                |
| KombiLang | SG       | 0,95                |
|           | AG       | 0,66                |
|           | RF2      | 0,86                |
|           | COV1     | 0,67                |
|           | SWG      | 0,87                |
|           | EB14     | 0,65                |

Die Tabelle fasst die Ergebnisse der Regressionsanalysen zusammen und zeigt das jeweilige R<sup>2</sup> Adj. für den Zusammenhang der COV2-Werte des jeweiligen Datensatzes mit den anderen Epidemien des jeweiligen Datensatzes.

## E. Diskussion

Die zentralen Beiträge der vorliegenden Arbeit konnten vor den Hintergründen der Empfehlungen des Fineberg Komitees gesehen werden. Dieses empfahl der WHO im Jahr 2011, adäquate Maße zu entwickeln und anzuwenden, die es erlauben, die Schwere eines epidemischen Geschehens im Sinne des Verbreitungsgrades einzuordnen (European Centre for Disease Prevention and Control, 2011). Diese Maße sollten auch dabei helfen, willkürliche Pandemieausrufungen zu verhindern (Singer et al., 2021). Ein geeignetes Merkmal hierzu könnte der DC sein, da die Ergebnisse der vorliegenden Arbeit darauf hindeuten, dass Pandemien eher einen medianen DC unter 1,00 hatten, wohingegen der mediane DC in EB14 i. d. R. über diesem Wert lag. Sollte sich in weiteren Untersuchungen dieser Befund bestätigen, so gäbe es ein quantitatives Merkmal zur Unterscheidung von Epi- und Pandemien.

### I. Zusammenfassung der Ergebnisse

Die Literaturrecherche ergab, dass von den 18.696 Literatureinträgen, die am 23.03.2021 auf PubMed zu Influenzapandemien zur Verfügung standen, letztlich nur 131 (0,7%) für die vorliegende Arbeit verwendet werden konnten. Von diesen 131 Einträgen behandelten drei Artikel andere Pandemien als SG und SWG. Man konnte also zu dem Schluss kommen, dass unter Verwendung des dargestellten Suchalgorithmus auf PubMed fünf der sieben Influenzapandemien seit 1889 der epidemiologischen Forschung fast gänzlich entgingen. Auch bei Ebola zeigte sich eine Schieflage: Von den 13 Artikeln, die in den INZDS einfließen, behandelten neun EB14 und jeweils ein Beitrag die Ebolaepidemien von 1976, 1995, 2017 und 2018. Es entstand der Eindruck, dass das epidemiologische Verständnis der Charakteristika von Epidemien und Pandemien auf wenigen, ausgewählten Epidemien fußte.

Der Interratervergleich ergab, dass verschiedene Epidemiologen ein recht ähnliches Verständnis von dem Aussehen und der Häufigkeit epidemischer Wellen hatten. Auch der abgeleitete DC war bei den besagten Personen recht ähnlich. Da man sich im Rahmen der Arbeit allerdings nur auf die DC der Wellen Eins und Zwei fokussiert hatte und die Bewertung von drei fachlich versierten Ratern zwar ein erster Hinweis ist, sollte dieser Interratervergleich mit mehr Abbildungen, DC und mehr Ratern wiederholt werden.

Die Analyse der Häufigkeit von Folgewellen ergab, dass diese nicht nur eine Theorie waren (Jefferson and Heneghan, 2020), sondern dass das Auftreten von Folgewellen die

Regel war. Verallgemeinernd zeigte die vorliegende Arbeit, dass zweite Infektionswellen in 3/4 und dritte Infektionswellen in 1/3 der untersuchten Perspektiven auftraten. In Influenzapandemien traten zweite Wellen etwa in der Hälfte der Fälle auf. Man konnte vermuten, dass die überdurchschnittliche Anzahl an Folgewellen in COV2 ein Artefakt war, die wellenphysikalische Analogie schien davon aber unberührt. Dies zeigte sich auch daran, dass es keinen nennenswerten Unterschied in den regionalen und nationalen Datensätzen gab und es somit keinen Hinweis darauf gab, dass eine kombinierte Betrachtung regionaler und nationaler Perspektiven ohne Doppelungen des Ortes nicht möglich war, was es in Zukunft deutlich einfacher macht, eine umfangreichere Stichprobe zu generieren. Ob die Unterschiede zwischen den kurzen und langen Perspektiven bei umfangreicheren Stichproben Bestand haben, müssen künftige Untersuchungen zeigen. Auf Basis der hier erarbeiteten Ergebnisse waren die Unterschiede in den Wellenhäufigkeiten relevant, weshalb man auch in zukünftigen Untersuchungen zwischen kurzen und langen Perspektiven unterscheiden sollte. Da die Wellenhäufigkeiten von SWG und COV2 im Verhältnis zu den anderen Epidemien, die COV2 im Rahmen dieser Arbeit gegenübergestellt wurden, am ähnlichsten waren, konnte man COV2 – in Bezug auf die Wellenhäufigkeit – als vom Typus SWG bezeichnen.

Bei den DC zeigte sich, dass die untersuchten Pandemien i. d. R. mediane DC von unter 1,00 hatten, während in EB14, der einzigen hier untersuchten Nicht-Pandemie, der mediane DC häufig 1,00 und höher war. Sollten weitere Untersuchungen bestätigen, dass Epidemien zu medianen DC von 1,00 und höher neigten, wohingegen Pandemien eher darunterblieben, so hätte man eine „harte“ Unterscheidung von Epi- und Pandemien im Sinne des Fineberg Komitees.

Bei den Wellenhöhen zeigte sich, dass kurze und lange Perspektiven darin übereinstimmen, dass Zweit- und Drittwellen höher waren als Erstwellen und Viertwellen niedriger waren als diese. In COV2 wurden die Wellenhöhen von Welle zu Welle höher, weshalb COV2 nicht den historischen Vorgängern entsprach. Ob dies tatsächlich eine Eigenart von COV2 war, oder ob sich hier Verzerrungen durch die Beobachtungsfähigkeiten ergaben, etwa wegen der virusspezifischen Antigentests, die bei den meisten historischen Epidemien nicht zur Verfügung standen, konnte hier nicht geklärt werden.

In dem Kapitel zur Vorhersage des Zweitwellengipfels zeigte sich, dass  $D'$  über alle untersuchten Datensätze hinweg wenig schwankte und die beiden großen H1-Pandemien

des 20. Jh. ähnliche Dynamiken zeigten. Ein vergleichsweise großes  $R^2$  Adj. erhielt man im Vergleich zu den anderen untersuchten Datensätzen bei der Bestimmung des linearen Zusammenhangs von berechneten und beobachteten Zweitwellengipfeln in den kurzen Coronaperspektiven. Die mathematische Exploration ergab z. T. präzisere Verfahren als die Methode des DC, denen aber eine physikalische Analogie, wie die der Schwingungseigenschaft, fehlte. Unzweifelhaft wäre eine Schätzmethode mit einer besonders früh prognostizierbaren und besonders geringen Abweichung in Tagen von größter klinischer bzw. Public Health Relevanz, da man sich so aus zeitlicher Sicht bestmöglich auf eine Infektionswelle vorbereiten könnte.

Der quantitative Vergleich von COV2 mit den historischen Epidemien ergab, dass EB14 nicht zu COV2 passte. Die Ähnlichkeit von COV2 und COV1, die besonders in den Datensätzen mit den kurzen Perspektiven erkennbar war, ließe sich kausal erklären, da es sich um die gleiche Virusart handelte. Über alle Datensätze hinweg war es SWG, die die größte Ähnlichkeit mit COV2 aufwies.

In der Summe also bestätigte der quantitative Vergleich die Ergebnisse der qualitativen Vergleiche. Auch die Ähnlichkeit zu SG war hoch, weshalb man spekulieren könnte, ob sich hier direkt oder indirekt ein pandemischer Mechanismus beobachten ließe. Diese Vermutung konnte dadurch gestützt werden, dass es deutliche Unterschiede zu EB14, der einzigen hier untersuchten Nicht-Pandemie, gab.

## II. Vergleich zu anderen Untersuchungen

Da es bislang wohl keine Arbeit gab, die die gleiche oder eine ähnliche Methodik verwendet hatte, war ein Vergleich zu anderen Untersuchungen, die die hier generierten Ergebnisse einordneten, nur bedingt möglich. Ausgehend von der dargestellten Idee einer Analogie eines physikalisch-epidemiologischen Verständnisses der epidemischen Zeitreihen ließen sich in der Forschung zwei große Strömungen erkennen: 1. epidemische Ereignisse in Abhängigkeit von Zeit und Ort zu analysieren und 2. signaltheoretische Ansätze, die sich mit grundlegenden Eigenschaften von epidemischen Zyklen beschäftigten. Hoen und Kollegen analysierten die Kontakte zwischen Benutzern eines öffentlichen Netzwerks für kabelloses Internet und schätzten damit die theoretische, epidemische Dynamik, die ein solches Netzwerk in einer Epidemie aufweisen würde. Verschiedene epidemische Kurven wurden simuliert und mehrwellige Epidemien automatisch identifiziert. Bei mittlerem  $R_0$  traten in 44,6% der Simulationen mehrwellige Epidemien auf und man



schloss daraus, dass mehrwellige Epidemien selten waren (Hoen et al., 2015). Im Gegensatz zu diesen Simulationen wurden in der vorliegenden Arbeit die empirischen Häufigkeiten der Folgewellen gezählt, und es konnte gezeigt werden, dass die Eigenschaft der Mehrwelligkeit in knapp 71% der kurzen und ca. 96% der langen Perspektiven beobachtet wurde.

Castilla und Kollegen untersuchten zwei Wellen von SWG 2009 in Navarra, Spanien, und stellten fest, dass die zweite Welle, bezogen auf die Inzidenz, sieben Mal höher war als die erste Welle (Castilla et al., 2011). Auch die vorliegende Arbeit zeigt erstens, dass Zweitwellen generell höher waren als Erstwellen, und zweitens, dass im Q3 der Zweitwellen von SWG z. T. Höhen gemessen wurden, die das vier- bzw. fünffache an Höhe der Erstwelle erreichten.

Cliff und Kollegen stellten fest, dass sich Influenzapandemien schneller verbreiteten als die saisonale Influenza und früher in der Grippesaison auftraten. Die Autoren kamen zu dem Schluss, dass die Bestimmung der Wellengeschwindigkeit auch in regionalen und lokalen Daten möglich war (Cliff et al., 2007). Auch die vorliegende Promotionsarbeit kam zu dem Schluss, dass die Unterschiede von nationalen und regionalen Dynamiken bei den hier untersuchten Epidemien für die Charakterisierung von epidemischen Wellen nicht relevant waren.

Akin und Gözel beschrieben die Ähnlichkeiten von SG, AG, HKG, SWG und COV2 anhand der üblichen und etablierten Dynamikbeschreibungen: Morbidität, Mortalität, Übertragbarkeit, Wachstumsrate und Dauer (Akin and Gözel, 2020). Diesem Arsenal wurde durch die vorliegende Arbeit mit dem DC eine weitere Beschreibung mit hinreichender prognostischer und Public Health Relevanz hinzugefügt.

Bei Schwartz und Kollegen stimmten die Beobachtung und die Modelle nur begrenzt überein, was die Autoren auch auf eine Untererfassung der Fallzahlen zu Beginn von SWG zurückführten (Schwartz et al., 2015). Mit Hilfe der in der vorliegenden Arbeit vorgestellten Methoden wäre es möglich, die Zeitpunkte einer vermuteten Vorwelle zu berechnen, indem man in den dargestellten Formeln nicht nur die Höhepunkte als Bezugspunkte nimmt, sondern auch die Wellenanfangs- und -endzeitpunkte. In der Folge könnte man überprüfen, ob es in den Datensätzen ein passendes Signal gibt.

Brockmann und Helbing nahmen konstante Werte für epidemische Parameter an und kamen zu dem Schluss, dass so hinreichend präzise Ergebnisse etwa für die Berechnung der

Ankunftszeit einer Epidemie in einem neuen Gebiet erreicht wurden (Brockmann and Helbing, 2013). Wie stark die Ergebnisse streuten und was die Autoren unter hinreichender Präzision verstanden, ließ sich nur schwierig ermitteln, jedoch fand sich zu Abbildung 4D der Publikation der Hinweis, dass die zeitliche Präzision um den Prävalenzhöhepunkt etwa bei zwei Monaten lag. Die in dieser Arbeit erzielte präziseste Schätzung der absoluten Abweichung in Tagen in Q3 betrug 35 Tage und war damit deutlich präziser.

Mummert und Kollegen nahmen an, dass es mehrere Mechanismen gab, die das Mehrwellenbild verursachten bzw. beeinflussen konnten. Die Autoren verwendeten in ihrer Arbeit fünf mathematische Modelle, wobei vier dieser Modelle zwei Wellen von SWG quantitativ „reproduzieren“ konnten (Mummert et al., 2013). Eine Definition, wie in der vorliegenden Promotionsarbeit, was eine Infektionswelle war, boten die Autoren nicht; auch eine Thematisierung der unterschiedlichen Sichtweisen und deren Einflüsse auf die Analyse von epidemischen Wellen fehlten dem Beitrag der Kollegen. Insgesamt begrenzte sich der Beitrag von Mummert et al. auf zwei Wellen von SWG in den USA und verglich die generierten Ergebnisse nicht mit anderen Pandemien oder Regionen.

Chowell und Kollegen beschrieben, dass aktuelle Modelle häufig nur epidemische Ereignisse mit einem Gipfel abdecken konnten, epidemische Kurven in Wirklichkeit aber häufig komplexere Muster zeigten. Für ihre Vorhersagen nahmen die Kollegen an, dass die Inzidenzverläufe einer Epidemie den Regeln von anhaltenden oder gedämpften Schwingungen unterlagen. Die Autoren schlossen, dass die unterstellten Schwingungseigenschaften von Epidemien ein essentieller Bestandteil deren Natur waren. Die Autoren regten schließlich dazu an, diese Schwingungseigenschaften genauer zu erforschen, da sie vermuteten, dass jede Epidemie eigenen Schwingungscharakteristika unterliegt und damit präzise Vorhersagen möglich wären (Chowell et al., 2019). Die vorliegende Promotionsarbeit hat genau das getan und sogar weitere mögliche Verfahren von ähnlicher Komplexität wie dem DC untersucht.

Um zeitliche Bezüge, Trends und Zyklen festzustellen, wurden in epidemiologischen Zeitreihenuntersuchungen üblicherweise Spektralanalysen durchgeführt. Verwendet wurde im nachfolgenden Beitrag ein Modell mit harmonischen Zyklen, um den Trend und das saisonale Muster in den wöchentlichen Daten zu erfassen. Es konnte gezeigt werden, dass diese Influenzaausbrüche stark kohärent mit den annähernd jährlichen meteorologischen Zyklen waren (Yan et al., 2019). Auch die vorliegende Arbeit kommt zu dem

Schluss, dass die untersuchten Infektionskrankheiten zyklischen Mustern unterlagen, die hier aber dafür verwendet wurden, um auch den nächsten Wellenhöhepunkt vorauszusagen.

Cacciapaglia und Kollegen zeigten am Beispiel von COV2 und SG, dass Pandemien sich in Folgewellen über die Welt verbreiteten und schlugen dabei ein Wellenbild vor, das auf dem Rahmenwerk der „Epidemic Renormalization Group“ basierte. Gemäß ihren Ergebnissen gingen die Autoren davon aus, dass sich eine Folgewelle im Zeitraum zwischen zwei Wellen, also während einer Phase des linearen Fallwachstums, verzögern oder gar verhindern ließ. Bislang, so die Autoren, war der Ursprung der Wellenausbreitung von Epidemien noch nicht verstanden. Die etablierten Algorithmen konnten aber nicht das zeitliche Eintreten der Zweitwellen im Herbst 2020 vorhersagen (Cacciapaglia et al., 2021). Das Problem war, dass es sich um Modelle und nicht um tatsächliche Wellen handelte. Bei den Kollegen blieb auch offen, wie viele Ereignisse von SG und COV2 berücksichtigt wurden und nach welchen Kriterien die Länder, die in die Untersuchung eingingen, ausgesucht wurden. Zudem fehlte ein Vergleich von SG und COV2. Die Vorhersage, dass die dritte Welle in Japan Anfang Dezember 2021 ihren Höhepunkt erreichen würde, war nicht vollkommen falsch. Vergleich man die 40 Tage Präzision mit den in  $Z_C$ -Varianten erreichten Präzisionen der vorliegenden Promotionsarbeit, so stellte man fest, dass es von 1.440 möglichen Beobachtungen 368 gab, die die gleiche oder eine schlechtere mediane Abweichung hatten – in Q3 sind dies 1.181 Beobachtungen.

Nicht neu an der vorliegenden Promotionsarbeit war die Verwendung physikalischer Schwingungseigenschaften, um Zyklen in Epidemien und epidemischen Kurven zu beschreiben. Wichtige Neuerungen dieser Arbeit waren allerdings die Verwendung und Beschreibung der Zeitanteile, die Charakterisierung der Schwingungen mittels DC und die Beschreibung von Beobachtung und Vorhersage mittels der Absoluten Abweichung, die von hoher klinischer und Public Health Relevanz ist. Auch die Schaffung der zugehörigen Datenbank, die die unterschiedlichen Regionen, Länder und Epidemien abdeckt, war allem Anschein nach neu.

### III. Stärken und Limitationen

Limitationen der vorliegenden Promotionsarbeit waren, dass ausschließlich die Inzidenzkurven von Epidemien untersucht wurden. Ob aber die generierten Ergebnisse auch in den dazugehörigen Mortalitätskurven galten bzw., ob und inwieweit Unterschiede in den

Charakteristika von Inzidenz- und Mortalitätskurven bestanden, wurde zukünftigen Arbeiten überlassen. Das Höhenprofil der Epidemien hätte präziser erfasst werden können und damit auch eine mögliche Präzisierung der Taxonomie, wenn man auch die Wellenhöhen zu den jeweiligen Anfangs- und Endzeitpunkten der Wellen gemessen und ausgewertet hätte. Das Interraterverfahren sollte in zukünftigen Arbeiten mehr Rater und mehr Abbildungen umfassen. Die mittels Regression abgeleitete Taxonomie beruhte nur auf den Streckenbezeichnungen, sollte aber mehrere Charakteristika mit gleicher Analyse-methode erfassen – hierzu würden sich die Höhenprofile anbieten. Trotz des großen Aufwandes bei der Datengenerierung müssten mehr Daten für alle Epidemien außer SG, SWG und COV2 generiert werden. Erste Russische Grippe, HKG und die namenlose Pandemie um 1900 konnten, mangels einschlägiger Literatur bei Verwendung des dargestellten Suchalgorithmus auf PubMed, nicht in den INZDS eingehen. Bei den übrigen Epidemien, außer SG, SWG und COV2, war die Datenlage auf PubMed bei Verwendung des dargestellten Suchalgorithmus durchaus schlecht. Diese Lücken sollten mit höherer Priorität gefüllt werden. Auch weitere Daten zu SG, SWG und COV2, bei letzterer gerade auf der regionalen Ebene, sollten weiterhin aufgenommen werden, jedoch sollten auch andere Infektionskrankheiten wie Dengue, Malaria usw. nicht unberücksichtigt bleiben. Auch bei der mathematischen Exploration sollten in Zukunft weitere Transformationen, wie etwa der Logarithmus, berücksichtigt werden.

Die zentralen Limitationen sind also: EB14 ist die einzige untersuchte Nicht-Pandemie, die Arbeit untersucht ein begrenztes Erregerspektrum, es sind keine Sensitivitätsanalysen bzgl. der kurzen und langen Perspektiven eingegangen und mögliche Messungenauigkeiten bei der Datenerhebung wurden zwar über die Erfassung in Perspektiven umgangen, aber ein Plausibilitätsmechanismus wurde nicht überlegt.

Eine Stärke der vorliegenden Arbeit ist der Datensatz, der eine epidemieübergreifende, internationale, regionale und vergleichende Analyse ermöglichte. Auch der Umfang an Daten und Datenquellen war deutlich größer als in vergleichbaren Untersuchungen. Der Umstand, dass pro ausgewerteter Darstellung jeweils alle Perspektiven erfasst wurden, war eine zusätzliche Stärke. Auch der Interrateransatz zu dem Verständnis epidemischer Wellen war wohl bislang noch nicht publiziert worden. Auch wurde wohl bislang nicht die absolute Abweichung als Maß der Präzision einer Vorhersage publiziert, obwohl diese im Vergleich zu den üblichen biometrischen Maßen von größerer klinischer und Public

Health Relevanz ist. Die mathematische Exploration ermöglichte die rationale Einordnung der physikalischen Analogie des DC.

#### IV. Implikationen

Die wichtigste Implikation war, dass sich durch die Verwendung einer physikalischen Analogie und die Vermessung historischer Epidemien Maße ableiten ließen, die es auch in aktuellen Seuchen ermöglichten, einen zeitlichen Korridor zu berechnen, in dem im Falle einer weiteren Infektionswelle der nächste Wellenhöhepunkt lag. Dabei war die verwendete Methodik derart einfach, dass sie im Grunde überall, mittels Block und Bleistift, angewendet werden kann; es waren also keine komplizierten mathematischen Modelle notwendig, um ähnlich präzise Vorhersagen zu treffen. Zudem zeigte sich, dass die Verwendung von Zeitanteilen Harmonien erkennen ließen – wie etwa bei  $\mathbf{D}'$  – die den bislang üblichen Betrachtungen der absoluten Zeitachsen verborgen geblieben sind. Die dargestellten Unterschiede von Epidemie und Pandemie könnten, wenn weitere Untersuchungen die Vermutungen bestätigen, zu einem messbaren, rational beschreibbaren Maß zur Unterscheidung von Epidemie und Pandemie führen, wie es das Fineberg Komitee 2011 gefordert hatte. Es galt zu beachten: Je größer der DC, desto länger war die epidemische Welle im Zähler im Verhältnis zu der Dauer von Wellenhöhepunkt zu Wellenhöhepunkt im Nenner. Ein DC von 1,00 zeigte an, dass aus dem Ende einer epidemischen Welle der Anfang einer neuen Welle hervorging. Größere DC bedeuteten, dass die letzte Welle noch nicht vollständig abgeklungen war, eine neue aber bereits begonnen hatte. DC kleiner 1,00 konnten als Latenzzeit zwischen dem Ende einer epidemischen Welle und dem Beginn einer neuen Welle verstanden werden. Breitete sich dann die Epidemie in eine weitere Population aus, so kam es vielleicht zu einer Überlagerung der Wellenbilder. Die skizzierte Taxonomie ermöglichte es, wenn sie beständig gepflegt und verbessert wird, Epidemien in Gruppen einzuteilen, was den zukünftigen Verlauf für Kliniker und die Gesundheitspolitik abschätzbarer macht, da die historischen Vorgänger häufig ausreichend beschrieben sind. Durch die Verbindung von Epidemietypen, der Abschätzbarkeit des nächsten Wellengipfels – wenn es eine Folgewelle gibt – und der möglichen Unterscheidbarkeit zwischen einem epidemischen oder pandemischen Verlauf ließen sich adäquatere und präzisere Gegenmaßnahmen ergreifen. Der Umstand, dass fünf von sieben Influenzapandemien der epidemiologischen Forschung bei der Suche über PubMed mittels des

dargestellten Suchalgorithmus entgehen, zeigt den großen Forschungsbedarf im Bereich der historischen Epidemiologie.

#### V. Unbeantwortete Fragen und zukünftiger Forschungsbedarf

Scarpino und Petri zeigten anhand der Modellierung von 20.000 Simulationen verschiedener Infektionskrankheiten, dass eine Vorhersage, die über vier Monate hinausgeht, in den meisten Fällen nur noch wenig präzise ist (Scarpino and Petri, 2019). Es ist die Frage, ob sich diese Vermutungen auch in Echtweltdaten zeigen. Der INZDS bietet hierfür eine gute Ausgangsgrundlage. Man sollte bei zukünftigen Forschungen untersuchen, ob sich die Unterschiede von EB14 und den untersuchten Pandemien bei der Untersuchung von weiteren Epidemien bestätigen. Wäre dies der Fall, so wüsste man sehr früh in einer Epidemie, ob die Gefahr einer Pandemie besteht, und könnte adäquate Maßnahmen ergreifen. Um dies zu ermöglichen, muss ein deutliches Mehr an historischer Epidemiologie beforscht werden, da nur in geringem Ausmaß die nötigen Informationen vorlagen. Auch die Untersuchung der Mortalitätskurven und deren Vergleich mit den Inzidenzkurven ist wichtig, wobei darauf geachtet werden sollte, die Wellenhöhen nicht nur zu den Zeitpunkten der maximalen Wellenauslenkung zu erfassen.

## F. Zusammenfassung

Hintergrund: Es wurden vier Fragen formuliert: 1. wie häufig traten weitere Infektionswellen in vergleichbaren Pandemien auf, 2. gipfelten weitere Infektionswellen höher als Erstwellen, 3. kann man das zeitliche Auftreten von weiteren Infektionswellen vorhersagen und 4. gibt es bestimmte Typen von Pandemien?

Methoden: Es wurden Inzidenzinformationen der Influenzapandemien der letzten 130 Jahre, der bekannten Coronapandemien und der Ebolaepidemien ausgewertet, die Häufigkeiten von epidemischen Wellen gezählt und deren Höhenprofil ermittelt. Mittels Interratervergleich wurde untersucht, inwieweit sich die Auffassungen von Infektionswellen ähnelten. Durch die analoge Verwendung des Tastgrades wurden zeitliche Vorhersagen für Zweitwellengipfel in der 2019 Coronavirus Erkrankung (COVID-19) Pandemie berechnet, wobei die Berechnungsparameter aus den Epidemien der Vergangenheit abgeleitet wurden. Mittels mathematischer Exploration wurde überprüft, ob die verwendete physikalische Analogie die präzisesten Resultate lieferte. Schließlich wurden die ermittelten Charakteristika der vergangenen Epidemien der aktuellen Coronapandemie gegenübergestellt, um mögliche Ähnlichkeiten festzustellen.

Ergebnisse: Folgewellen in Epidemien waren häufiger als bislang gedacht und die COVID-19 Pandemie entsprach im Wellenhöheprofil nicht den bekannten Mustern. Die Interpretation von epidemischen Wellen bei verschiedenen Ratern war ähnlich. Die analoge Verwendung des Tastgrades ermöglichte eine hinreichend präzise Schätzung der Zeitpunkte der Wellenhöhepunkte in Zweitwellen. Die erreichte Präzision war denen anderer Untersuchungen und Methoden gleichrangig und in manchen Fällen überlegen. Gemäß den erarbeiteten Charakteristika entsprach die COVID-19 Pandemie am ehesten der Schweingrippe von 2009. Möglicherweise unterschieden sich Epidemien von Pandemien darin, dass die Werte der Tastgrade in Pandemien in der Regel unterhalb von 1,00 – in Epidemien wohl über diesem Wert lagen.

Fazit: Folgewellen in Epidemien sind mehr die Regel als die Ausnahme, und mögliche Wellenfolgegipfel lassen sich mit hinreichender Präzision mittels aus historischen Epidemien abgeleiteter Schätzer berechnen.

## G. Summary

Background: Four questions were formulated: 1. how often did subsequent waves of infection occur in comparable pandemics, 2. did subsequent waves of infection peak higher than initial waves, 3. can the timing of subsequent waves of infection be predicted, and 4. are there specific types of pandemics?

Methods: Incidence information of influenza pandemics of the last 130 years, known corona pandemics and Ebola epidemics were analyzed, the frequencies of epidemic waves were counted and their height profiles were calculated. Interrater comparison was used to examine the extent to which the perceptions of infectious waves are similar. By analogous use of duty cycle, temporal predictions for second-wave peaks in the 2019 Coronavirus disease (COVID-19) pandemic were calculated, with calculation parameters derived from past epidemics. Mathematic exploration was used to verify that the physical analogy used yielded the most accurate results. Finally, the determined characteristics of past epidemics were contrasted with the current corona pandemic to determine possible similarities.

Results: Subsequent waves in epidemics were more frequent than previously thought, and the COVID-19 pandemic did not conform to known patterns in the wave height profile. Interpretation of epidemic waves in different raters was similar. The analogous use of duty cycle provided a sufficiently precise estimate of the timing of wave height peaks in second waves. The precision achieved was equal to, and in some cases superior to, those of other studies and methods. According to the characteristics identified, the COVID-19 pandemic corresponded most closely to the 2009 swine flu. It is possible that epidemics differed from pandemics in the sense that the values of duty cycles in pandemics were usually below 1.00 - in epidemics they were probably above this value.

Conclusion: Subsequent waves in epidemics are more common than an exception, the timing of possible wave peaks can be calculated with sufficient precision by means of estimators derived from historical epidemics.



## H. Literaturverzeichnis

1. Akin, L., and Gözel, M.G. (2020). Understanding dynamics of pandemics. *Turk J Med Sci* 50, 515-519. 10.3906/sag-2004-133.
2. Amato Gauci, A., Zucs, P., Snacken, R., Ciancio, B., Lopez, V., Broberg, E., Penttinen, P., Plata, F., and Nicoll, A. (2010). The 2009 A(H1N1) pandemic in Europe. ECDC. [https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/101108\\_SPR\\_pandemic\\_experience.pdf](https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/101108_SPR_pandemic_experience.pdf).
3. Ansart, S., Pelat, C., Boelle, P.-Y., Carrat, F., Flahault, A., and Valleron, A.-J. (2009). Mortality burden of the 1918–1919 influenza pandemic in Europe. *3*, 99-106. 10.1111/j.1750-2659.2009.00080.x.
4. Ansart, S., and Valleron, A.J. (2008). [The history of viral pandemics]. *Rev Mal Respir* 25, 490-491. 10.1016/s0761-8425(08)71589-4.
5. Baidoun, N. (2016). What is the acceptable r-squared value? [https://www.researchgate.net/post/what\\_is\\_the\\_acceptable\\_r-squared\\_value](https://www.researchgate.net/post/what_is_the_acceptable_r-squared_value).
6. Barry, J.M. (2009). Observations on past influenza pandemics. *Disaster Med Public Health Prep* 3 *Suppl* 2, S95-99. 10.1097/DMP.0b013e3181be82c5.
7. Bin, M., Cheung, P.Y.K., Crisostomi, E., Ferraro, P., Lhachemi, H., Murray-Smith, R., Myant, C., Parisini, T., Shorten, R., Stein, S., and Stone, L. (2021). Post-lockdown abatement of COVID-19 by fast periodic switching. *PLOS Computational Biology* 17, e1008604. 10.1371/journal.pcbi.1008604.
8. Brockmann, D., and Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science* 342, 1337-1342. 10.1126/science.1245200.
9. Cacciapaglia, G., Cot, C., and Sannino, F. (2021). Multiwave pandemic dynamics explained: how to tame the next wave of infectious diseases. *Sci Rep* 11, 6638. 10.1038/s41598-021-85875-2.

10. Castilla, J., Guevara, M., García Cenoz, M., Reina, G., Martínez Artola, V., Zamora, M.J., Irisarri, F., Fernández Alonso, M., and Salcedo, E. (2011). [Differences between summer and fall waves of influenza in the pandemic of influenza (H1N1) 2009 in Navarre, Spain]. *Rev Esp Salud Publica* 85, 47-56. 10.1590/s1135-57272011000100006.
11. Center for Infectious Disease Research and Policy (2008). WHO's draft pandemic flu guidance revises phases. <https://www.cidrap.umn.edu/news-perspective/2008/10/whos-draft-pandemic-flu-guidance-revises-phases>.
12. Centers for Disease Control and Prevention (2009). CDC Press Conference on Investigation of Human Cases of Novel Influenza A H1N1. <https://www.cdc.gov/media/transcripts/2009/t090611.htm>.
13. Centers for Disease Control and Prevention (2015). Interpretation of Epidemic Curves. <https://www.cdc.gov/foodsafety/outbreaks/investigating-outbreaks/epi-curves.html>.
14. Chaparro, L. (2015). Chapter 1 - Continuous-time Signals. In *Signals and Systems Using MATLAB (Second Edition)*, L. Chaparro, ed. (Academic Press), pp. 65-122. <https://doi.org/10.1016/B978-0-12-394812-0.00001-2>.
15. Chowell, G., Tariq, A., and Hyman, J.M. (2019). A novel sub-epidemic modeling framework for short-term forecasting epidemic waves. *BMC Medicine* 17, 164. 10.1186/s12916-019-1406-6.
16. Cliff, A.D., Haggett, P., and Smallman-Raynor, M. (2007). An exploratory method for estimating the changing speed of epidemic waves from historical data. *International Journal of Epidemiology* 37, 106-112. 10.1093/ije/dym240.
17. Doshi, P. (2011). The elusive definition of pandemic influenza. *Bulletin of the World Health Organization* 89, 532-538. 10.2471/BLT.11.086173.

18. Esparza, J. (2020). Lessons From History: What Can We Learn From 300 Years of Pandemic Flu That Could Inform the Response to COVID-19? *Am J Public Health* *110*, 1160-1161. 10.2105/ajph.2020.305761.
19. European Centre for Disease Prevention and Control (2011). Issues around the definitions and severity of pandemic influenza. <https://www.ecdc.europa.eu/en/news-events/issues-around-definitions-and-severity-pandemic-influenza>.
20. European Centre for Disease Prevention and Control (2021). Questions and answers on influenza pandemics. <https://www.ecdc.europa.eu/en/pandemic-influenza/facts/questions-and-answers>.
21. Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin* *76*, 378-382. 10.1037/h0031619.
22. Fung, H.P. (2015). Re: What is a good r square value in regression analysis? [https://www.researchgate.net/post/what\\_is\\_a\\_good\\_r\\_square\\_value\\_in\\_regression\\_analysis/5527fb76f079eddf718b45bc/citation/download](https://www.researchgate.net/post/what_is_a_good_r_square_value_in_regression_analysis/5527fb76f079eddf718b45bc/citation/download).
23. Grami, A. (2016). Chapter 3 - Signals, Systems, and Spectral Analysis. In *Introduction to Digital Communications*, A. Grami, ed. (Academic Press), pp. 41-150. <https://doi.org/10.1016/B978-0-12-407682-2.00003-X>.
24. Hoen, A.G., Hladish, T.J., Eggo, R.M., Lenczner, M., Brownstein, J.S., and Meyers, L.A. (2015). Epidemic Wave Dynamics Attributable to Urban Community Structure: A Theoretical Characterization of Disease Transmission in a Large Network. *J Med Internet Res* *17*, e169. 10.2196/jmir.3720.
25. Huang, X.M., and Wang, X.S. (2019). Traveling waves of diffusive disease models with time delay and degeneracy. *Math Biosci Eng* *16*, 2391-2410. 10.3934/mbe.2019120.
26. Influenza Research Database (2021). Influenza Research Database. <https://www.fludb.org/brc/home.spg?decorator=influenza>.

27. Influenza Research Database (N.D.). [www.fludb.org](http://www.fludb.org). [www.viprbrc.org](http://www.viprbrc.org). What they are & what they do. [https://www.fludb.org/brcDocs/documents/IRD\\_brochure.pdf](https://www.fludb.org/brcDocs/documents/IRD_brochure.pdf).
28. Jefferson, T., and Heneghan, C. (2020). Covid 19 – Epidemic ‘Waves’. <https://www.cebm.net/covid-19/covid-19-epidemic-waves/>.
29. Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE) (2021). `time_series_covid19_confirmed_global_narrow.csv`. <https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases>.
30. Johnson, N.P.A.S., and Mueller, J. (2002). Updating the Accounts: Global Mortality of the 1918-1920 "Spanish" Influenza Pandemic. *Bulletin of the History of Medicine* 76, 105-115.
31. Kelly, H. (2011). The classical definition of a pandemic is not elusive. *Bulletin of the World Health Organization* 89, 540-541. 10.2471/BLT.11.088815.
32. Laxminarayan, R., B, C.M., G, V.T., Arjun Kumar, K.V., Wahl, B., and Lewnard, J.A. (2021). SARS-CoV-2 infection and mortality during the first epidemic wave in Madurai, south India: a prospective, active surveillance study. *The Lancet Infectious Diseases*. 10.1016/S1473-3099(21)00393-5.
33. Lewis, M., Renclawowicz, J., and van den Driessche, P. (2006). Traveling waves and spread rates for a West Nile virus model. *Bull Math Biol* 68, 3-23. 10.1007/s11538-005-9018-z.
34. Lin, J., Andreasen, V., Casagrandi, R., and Levin, S.A. (2003). Traveling waves in a model of influenza A drift. *J Theor Biol* 222, 437-445. 10.1016/s0022-5193(03)00056-0.
35. Lupica, A., Manfredi, P., Volpert, V., Palumbo, A., and d'Onofrio, A. (2019). Spatio-temporal games of voluntary vaccination in the absence of the infection: the interplay of local versus non-local information about vaccine adverse events. *Math Biosci Eng* 17, 1090-1131. 10.3934/mbe.2020058.

36. Mandelis, A. (2000). Diffusion Waves and their Uses. *Physics Today* 8. <https://doi.org/10.1063/1.1310118>.
37. Morens, D.M., and Taubenberger, J.K. (2011). Pandemic influenza: certain uncertainties. *Rev Med Virol* 21, 262-284. 10.1002/rmv.689.
38. Mummert, A., Weiss, H., Long, L.-P., Amigó, J.M., and Wan, X.-F. (2013). A Perspective on Multiple Waves of Influenza Pandemics. *PLOS ONE* 8, e60343. 10.1371/journal.pone.0060343.
39. N.N. (2021). Ländervergleich des Fallverlaufs Covid-19. <https://2222029.github.io/>.
40. National Center for Biotechnology Information (2021a). Influenza virus database. <https://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database>.
41. National Center for Biotechnology Information (2021b). Influenza Virus Resource help center. How to use. <https://www.ncbi.nlm.nih.gov/genome/viruses/variation/help/flu-help-center/how-to-use/#search-for-keyword>.
42. National Center for Biotechnology Information (2021c). NCBI Virus. [https://www.who.int/influenza/gisrs\\_laboratory/flunet/en/](https://www.who.int/influenza/gisrs_laboratory/flunet/en/).
43. National Library of Medicine (N.D.). PubMed Overview. <https://pubmed.ncbi.nlm.nih.gov/about/>.
44. Neher, R.A., Dyrdak, R., Druelle, V., Hodcroft, E.B., and Albert, J. (2020). Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *Swiss Med Wkly* 150, w20224. 10.4414/smw.2020.20224.
45. Nelson, M.I., Tan, Y., Ghedin, E., Wentworth, D.E., St George, K., Edelman, L., Beck, E.T., Fan, J., Lam, T.T., Kumar, S., Spiro, D.J., Simonsen, L., Viboud, C., Holmes, E.C., Henrickson, K.J., Musser, J.M. (2011). Phylogeography of the spring and fall waves of the H1N1/09 pandemic influenza virus in the United States. *J Virol* 85, 828-834. 10.1128/jvi.01762-10.

46. Ning, S., Yang, S., and Kou, S.C. (2019). Accurate regional influenza epidemics tracking using Internet search data. *Scientific Reports* 9, 5238. 10.1038/s41598-019-41559-6.
47. Ortiz, J.R., Lafond, K.E., Wong, T.A., and Uyeki, T.M. (2012). Pandemic influenza in Africa, lessons learned from 1968: a systematic review of the literature. *Influenza Other Respir Viruses* 6, 11-24. 10.1111/j.1750-2659.2011.00257.x.
48. Our World in Data (2021). About. <https://ourworldindata.org/about>.
49. Pang, D.F., and Xiao, Y.N. (2019). The SIS model with diffusion of virus in the environment. *Math Biosci Eng* 16, 2852-2874. 10.3934/mbe.2019141.
50. Petersen, E., Koopmans, M., Go, U., Hamer, D.H., Petrosillo, N., Castelli, F., Storgaard, M., Al Khalili, S., and Simonsen, L. (2020). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *The Lancet Infectious Diseases*. 10.1016/S1473-3099(20)30484-9.
51. Potter, C.W., and Jennings, R. (2011). A definition for influenza pandemics based on historical records. *J Infect* 63, 252-259. 10.1016/j.jinf.2011.04.013.
52. Reina, J. (2009). [The influenza pandemics: learned lessons]. *Rev Esp Quimioter* 22 *Suppl 1*, 2-4.
53. Roser Max (2020). The Spanish flu (1918-20): The global impact of the largest influenza pandemic in history. <https://ourworldindata.org/spanish-flu-largest-influenza-pandemic-in-history>.
54. Scarpino, S.V., and Petri, G. (2019). On the predictability of infectious disease outbreaks. *Nature Communications* 10, 898. 10.1038/s41467-019-08616-0.
55. Schwartz, E.J., Choi, B., and Rempala, G.A. (2015). Estimating epidemic parameters: Application to H1N1 pandemic data. *Math Biosci* 270, 198-203. 10.1016/j.mbs.2015.03.007.

56. Shi, P., Keskinocak, P., Swann, J.L., and Lee, B.Y. (2010). Modelling seasonality and viral mutation to predict the course of an influenza pandemic. *Epidemiol Infect* 138, 1472-1481. 10.1017/s0950268810000300.
57. Shu, Y., and McCauley, J. (2017). GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill* 22. 10.2807/1560-7917.Es.2017.22.13.30494.
58. Singer, B.J., Thompson, R.N., and Bonsall, M.B. (2021). The effect of the definition of ‘pandemic’ on quantitative assessments of infectious disease outbreak risk. *Scientific Reports* 11, 2547. 10.1038/s41598-021-81814-3.
59. Standl, F., Jöckel, K.-H., Kowall, B., Schmidt, B., and Stang, A. (2020). Subsequent waves of viral pandemics, a hint for the future course of the SARS-CoV-2 pandemic. *medRxiv*, 2020.2007.2010.20150698. 10.1101/2020.07.10.20150698.
60. Standl, F., Jöckel, K.H., Brune, B., Schmidt, B., and Stang, A. (2021). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *Lancet Infect Dis* 21, e77. 10.1016/s1473-3099(20)30648-4.
61. Steindorf, V., and Maidana, N.A. (2019). Modeling the Spatial Spread of Chagas Disease. *Bull Math Biol* 81, 1687-1730. 10.1007/s11538-019-00581-5.
62. Sun, F.-B., Lock, S., and Wang, W. (2013). Duty Cycle Definition And Application In Reliability Engineering — A Unified Review, Classification, And Clarification. *International Journal of Reliability, Quality and Safety Engineering* 20, 1350017. 10.1142/s0218539313500174.
63. Taubenberger, J.K., and Morens, D.M. (2010). Influenza: the once and future pandemic. *Public Health Rep* 125 Suppl 3, 16-26.
64. Taylor, M.J., Fornusek, C., and Ruys, A.J. (2018). The duty cycle in Functional Electrical Stimulation research. Part II: Duty cycle multiplicity and domain reporting. *Eur J Transl Myol* 28, 7733-7733. 10.4081/ejtm.2018.7733.

65. Torok, M. (N.D.). Epidemic Curves Ahead. *Focus on Field Epidemiology I*.
66. Tosi, D., and Campi, A. (2020). How Data Analytics and Big Data Can Help Scientists in Managing COVID-19 Diffusion: Modeling Study to Predict the COVID-19 Diffusion in Italy and the Lombardy Region. *J Med Internet Res* 22, e21081. 10.2196/21081.
67. Tsai, J.C., and Weng, Y.Y. (2020). Propagation direction of traveling waves for a class of bistable epidemic models. *J Math Biol* 81, 1465-1493. 10.1007/s00285-020-01546-2.
68. Tsang, T.K., Wu, P., Lin, Y., Lau, E.H.Y., Leung, G.M., and Cowling, B.J. (2020). Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *Lancet Public Health* 5, e289-e296. 10.1016/s2468-2667(20)30089-x.
69. World Health Organization (2003). Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003. <https://www.who.int/publications/m/item/summary-of-probable-sars-cases-with-onset-of-illness-from-1-november-2002-to-31-july-2003>.
70. World Health Organization (2020). WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>.
71. World Health Organization (2021). FluNet. [https://www.who.int/influenza/gisrs\\_laboratory/flunet/en/](https://www.who.int/influenza/gisrs_laboratory/flunet/en/).
72. World Health Organization (WHO) (2009). Pandemic Influenza Preparedness and Response: A WHO Guidance Document. In.
73. Wu, J.T., and Cowling, B.J. (2011). The use of mathematical models to inform influenza pandemic preparedness and response. *Exp Biol Med (Maywood)* 236, 955-961. 10.1258/ebm.2010.010271.



74. Xu, S., and Li, Y. (2020). Beware of the second wave of COVID-19. *Lancet* 395, 1321-1322. 10.1016/s0140-6736(20)30845-x.
75. Yan, Q., Tang, S., Jin, Z., and Xiao, Y. (2019). Identifying Risk Factors Of A(H7N9) Outbreak by Wavelet Analysis and Generalized Estimating Equation. *Int J Environ Res Public Health* 16, 1311. 10.3390/ijerph16081311.
76. Yang, J., Liang, S., and Zhang, Y. (2011). Travelling waves of a delayed SIR epidemic model with nonlinear incidence rate and spatial diffusion. *PLoS One* 6, e21128. 10.1371/journal.pone.0021128.
77. Zhang, S.X., Arroyo Marioli, F., Gao, R., and Wang, S. (2021). A Second Wave? What Do People Mean by COVID Waves? - A Working Definition of Epidemic Waves. *Risk Manag Healthc Policy* 14, 3775-3782. 10.2147/rmhp.S326051.
78. Zhang, Y., Aevermann, B.D., Anderson, T.K., Burke, D.F., Dauphin, G., Gu, Z., He, S., Kumar, S., Larsen, C.N., Lee, A.J., Li, X., Macken, C., Mahaffey, C., Pickett, B.E., Reardon, B., Smith, T., Stewart, L., Suloway, C., Sun, G., Tong, L., Vincent, A.L., Walters, B., Zaremba, S., Zhao, H., Zhou, L., Zmasek, C., Klem, E.B., Scheuermann, R.H. (2016). Influenza Research Database: An integrated bioinformatics resource for influenza virus research. *Nucleic Acids Research* 45, D466-D474. 10.1093/nar/gkw857.
79. Zhang, Y., Yakob, L., Bonsall, M.B., and Hu, W. (2019). Predicting seasonal influenza epidemics using cross-hemisphere influenza surveillance data and local internet query data. *Scientific Reports* 9, 3262. 10.1038/s41598-019-39871-2.
80. Zürcher, K., Zwahlen, M., Ballif, M., Rieder, H.L., Egger, M., and Fenner, L. (2016). Influenza Pandemics and Tuberculosis Mortality in 1889 and 1918: Analysis of Historical Data from Switzerland. *PLoS One* 11, e0162575. 10.1371/journal.pone.0162575.

## I. Anhang

## I. Tabellenverzeichnis

| Tabelle   | Seite |
|---|-------|
| Tabelle 1: Übersicht der aus Abbildung 1 ableitbaren Strecken mit Streckenbezeichnung   | 24    |
| Tabelle 2: Median (Q1; Q3) im Datensatz der regionalen, kurzen Perspektiven (0erKurz)   | 33    |
| Tabelle 3: Median (Q1; Q3) im Datensatz der nationalen, kurzen Perspektiven (1erKurz)   | 34    |
| Tabelle 4: Median (Q1; Q3) im Datensatz der kombinierten regionalen und nationalen, kurzen Perspektiven (KombiKurz)                                     | 35    |
| Tabelle 5: Median (Q1; Q3) im Datensatz der regionalen, langen Perspektiven (0erLang)   | 37    |
| Tabelle 6: Median (Q1; Q3) im Datensatz der nationalen, langen Perspektiven (1erLang)   | 38    |
| Tabelle 7: Median (Q1; Q3) im Datensatz der kombinierten regionalen und nationalen, langen Perspektiven (KombiLang)                                     | 39    |
| Tabelle 8: Median der Mediane (Mdm) der Anteile der Streckenbezeichnungen nach Datensatz  | 41    |
| Tabelle 9: Die präzisesten Schätzungen für die Zweitwellengipfel mit $R^2$ Adj. als Gütekriterium   | 43    |
| Tabelle 10: Duty Cycle geordnet nach Abweichung in Tagen in Q3  | 51    |
| Tabelle 11: Ergebnisse der Regressionsanalysen für den Zusammenhang der COV2-Werte des jeweiligen Datensatzes mit den anderen Epidemien des Datensatzes | 53    |

## II. Abbildungsverzeichnis

| Abbildung   | Seite |
|---|-------|
| Abbildung 1: Schematische Darstellung einer zweiwelligen Epidemie   | 16    |
| Abbildung 2: Güte der, mittels $R^2$ ausgedrückten, Zusammenhänge der DC von Erst- und Zweitwelle bei den Ratern Kowall, Stang und Standl | 32    |

| Abbildung  | Seite |
|--|-------|
| Abbildung 3: Abweichung der Vorhersage der medianen absoluten Abweichung in Tagen von dem beobachteten Zweitwellengipfel pro Streckenkombination | 44    |

### III. Abkürzungsverzeichnis

| Abkürzung  | Bedeutung   |
|------------|---|
| AG         | Asiatischen Grippe  |
| CDC        | Centers for Disease Control and Prevention  |
| COV1       | Pandemie des Severe Acute Respiratory Syndrom Coronavirus 1   |
| COV2       | Pandemie des Severe Acute Respiratory Syndrom Coronavirus 2   |
| COVID-19   | Coronavirus Krankheit 19  |
| DC         | Duty Cycle  |
| EB14       | Ebolaepidemie 2014  |
| eRG        | epidemic Renormalisation Group  |
| HKG        | Hong-Kong Grippe  |
| INZDS      | Promotionsdatensatz der Inzidenzverläufe  |
| IQA        | Interquartilabstand   |
| MdM        | Median der Mediane  |
| n          | Anzahl  |
| N.D.       | Kein Veröffentlichungsdatum   |
| N.N.       | Nomen nominandum  |
| N.T.       | Kein Titel  |
| NA         | Not applicable  |
| NAN        | Not a Number  |
| Q1         | erstes Quartil  |
| Q3         | drittes Quartil   |
| $R_0$      | Basisreproduktionszahl: durchschnittliche Zahl von neuinfizierten Personen pro bereits infizierter Person |
| $R^2$ Adj. | Für den jeweiligen Stichprobenumfang adjustiertes $R^2$   |
| RG2        | zweiten Russischen Grippe   |
| SARS-CoV-2 | Severe Acute Respiratory Syndrom Coronavirus 2  |
| SG         | Spanischen Grippe   |
| SIR        | Susceptible, Infectious, Recovered  |

---

| Abkürzung | Bedeutung                   |
|-----------|-----------------------------|
| SWG       | Schweinegrippe              |
| T         | Periode                     |
| WHO       | Weltgesundheitsorganisation |

---

## J. Danksagung

Zunächst möchte ich mich bei meinen Betreuern, Professor Stang und Professor Kowall für die hervorragende Unterstützung und Anleitung während des gesamten Prozesses bedanken. Ohne ihre Geduld, Expertise und wertvollen Ratschläge hätte es diese Doktorarbeit nicht gegeben. Ich habe die gemeinsame Arbeit sowie die vielen Diskussionen in der Gruppe und im Einzelgespräch stets geschätzt und werde das in guter Erinnerung behalten. Zudem gilt mein Dank auch der medizinischen Fakultät der Universität Duisburg-Essen, da ich die Chance nutzen durfte mich als Wissenschaftler mit meiner Forschung weiterzuentwickeln.

Ich danke auch meinen Eltern, Geschwistern und meiner Familie. Ihre bedingungslose Unterstützung und Ermutigung haben mich stets motiviert, meine Ziele zu verfolgen. Besonders möchte ich mich bei meiner Frau Mara bedanken, denn letztlich war es ihre Geduld, Liebe und Verständnis die mir zusammen mit der Geburt meines Sohnes Linus die Kraft gegeben haben die Doktorarbeit zu vollenden. Auch bei meinen Freunden, Kollegen möchte ich mich bedanken, da auch sie stets ein verlässlicher Rückhalt und ein Quell guter Ratschläge waren.

## K. Lebenslauf

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.