



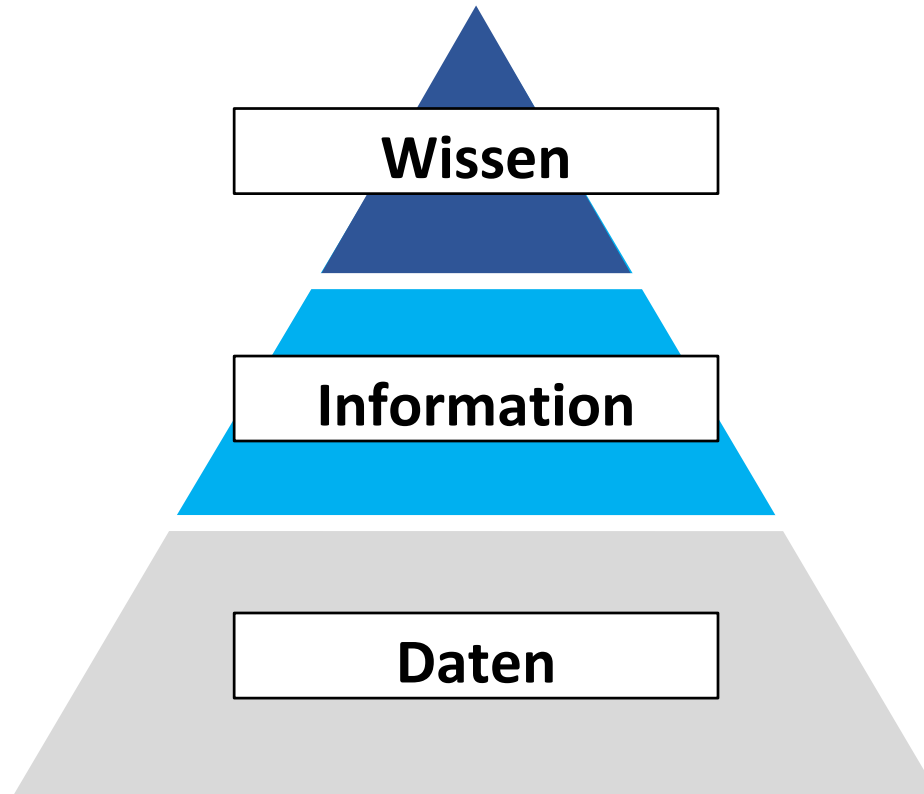
Zur Bedeutung einer guten Dateninfrastruktur in Pandemiezeiten

Dr. Amke Caliebe

Institut für Medizinische Informatik und Statistik,
Christian-Albrechts-Universität zu Kiel
Ethik-Kommission der Medizinischen Fakultät der
Christian-Albrechts-Universität zu Kiel

21. SOMMERTAGUNG AKEK 17. JUNI 2021

Wissenspyramide



Aamodt A, Nygard M (1995) Different roles and Mutual Dependencies of Data, Information and Knowledge: An AI Perspective on their Integration. In: Data & Knowledge Engineering, Vol. 16, Nr. 3, Elsevier Science, Amsterdam

**Zu wenig
aussagekräftige
Daten**

Pandemie als Datenkrise



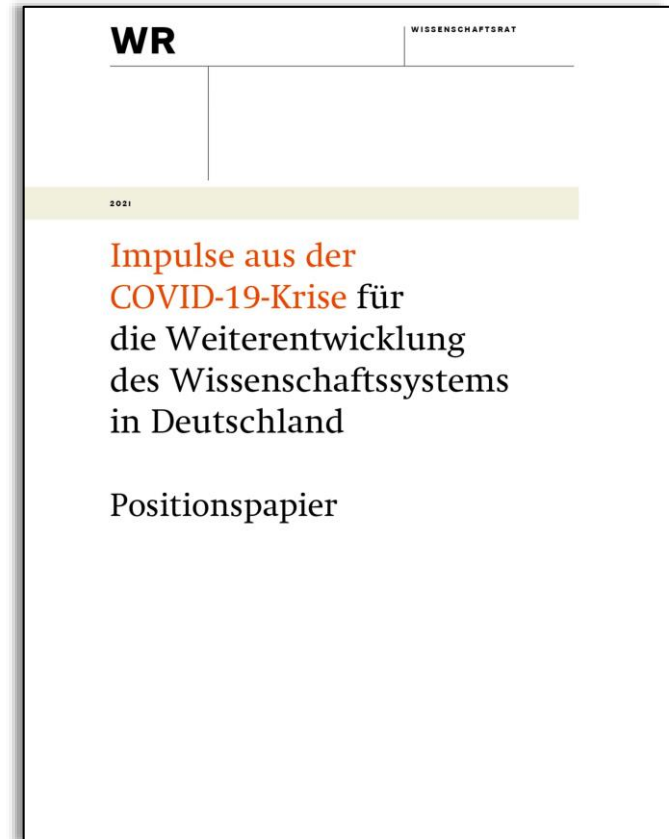
- Die Corona-Pandemie ist auch eine **Missing-Data-Krise**.
- Viele **schwere gesundheitspolitische Entscheidungen** müssen auf der Grundlage unzulänglicher oder fehlender Daten getroffen werden.
- Die **Wirksamkeit ergriffener Maßnahmen** kann häufig nicht überprüft werden, weil notwendige informationstechnische Voraussetzungen fehlen.

Caliebe A, Antes G, Leverkus F, Krawczak M (2021) Dtsch Ärztebl 118: A501.

Pandemie als Datenkrise

Aussagekräftige Daten fehlen
auf allen Ebenen!

- Epidemiologie
- Klinische Forschung
- Versorgung



https://www.wissenschaftsrat.de/download/2021/8834-21.pdf?__blob=publicationFile&v=15

**Zwei
Wissenschaftstheorien
– ein Ziel**

Modelle vs. Evidenz

BOSTON REVIEW
A POLITICAL AND LITERARY FORUM

JOIN SHOP DONATE Q = MENU

f t Join our email newsletter >

May 05, 2020

Topics
COVID-19
HEALTH
PHILOSOPHY
SCIENCE AND TECHNOLOGY

Share
f t e +

LOWER AND DELAY THE EPIDEMIC PEAK

of cases

uncontrolled transmission

control measures* slow the spread of disease and reduce burden on hospitals

with controls

healthcare system capacity (ICUs, ER visits, etc.)

time since first case

SCIENCE & NATURE

Models v. Evidence

COVID-19 has revealed a contest between two competing philosophies of scientific knowledge. To manage the crisis, we must draw on both.

JONATHAN FULLER

Boston Review is nonprofit and reader funded.

Contributions from readers enable us to provide a public space, free and open, for the discussion of ideas. Join this effort—**become a supporting reader** today.

<http://bostonreview.net/science-nature/jonathan-fuller-models-v-evidence>

Modelle



„end, it's Fleißarbeit“, sagt die Göttinger Max-Planck-Forscherin. © Julia Steiniger

*Viola Priesemann, Physikerin
MPI Göttingen*

- **Theoretisch**

„simplified representations of reality“

Neil Ferguson, Epidemiologe, Oxford

- **Pragmatisch**

„We know enough to act; indeed there is an imperative to act strongly and swiftly.“

Marc Lipsitch, Epidemiologe, Harvard

- Hängt von vielen Parametern und Annahmen ab

Evidenz



*Gerd Antes, Medizinstatistiker
Universität Freiburg*

- **Empirisch**
Daten basiert
hochqualitative Daten
Randomisierung, Repräsentativität
- **Kritisch**
„Kein noch so ausgefeiltes
statistisches oder theoretisches
Modell kann eine unzureichende
Datenlage kompensieren“

*Caliebe A, Antes G, Leverkus F,
Krawczak M (2021)
Dtsch Ärztebl 118: A501*

Ein Ziel



*Jonathan Fuller, Philosoph
University of Pittsburgh*

- „Epidemiology should be of **two minds**. It must **combine theory with evidence** and make use of diverse data while demanding data of increasingly higher quality.”
- „Models without evidence are **blind**, while evidence without models is **inert**.”
- „diverse range of experts”

**Mehr gut
geplante Studien**

Beobachtungsstudien

Confounder auf allen Ebenen!

Ein **Außerachtlassen von Störgrößen** „führt potenziell nicht nur zu Fehldeutungen, es **kann auch die Aussagekraft von Daten ruinieren** und den Blick auf zielgerichtete, effektive Maßnahmen gegen die Pandemie verstellen.“

Caliebe A, Antes G, Leverkus F, Krawczak M (2021) Dtsch Ärztebl 118: A501.



Dateninfrastruktur
Ein Baustein gegen COVID-19

Zur Bekämpfung der SARS-CoV-2-Pandemie fehlte es nicht nur an brauchbaren „guten“ Daten. Erforderlich für die künftige Prävention ist auch die Sicherstellung von interdisziplinärem Input sowie ein Austausch wissenschaftlicher Expertise und praktischer Erfahrung.

Amke Caliebe, Gerd Antes, Friedhelm Leverkus, Michael Krawczak

Fast alle für die Bewältigung einer Pandemie entscheidenden Fragen lassen sich nur mit geeigneten Daten beantworten. Persönliche Erfahrung, Wissen und Kompetenz sind in einer Notlage wie der aktuellen Corona-Krise zwar unersetzlich für die wissenschaftsgeleitete Unterstützung von Politik und Verwaltung, doch selbst ein Rat von Experten kann nur sinnvoll und hilfreich sein, wenn er sich auf aussagekräftige Daten stützt.

Die Qualität und Quantität der Daten haben dabei mindestens den gleichen Stellenwert wie die Auswahl der auf sie angewandten Methoden: Kein noch so ausgefeiltes statistisches oder theoretisches Modell kann eine unzureichende Datengrundlage kompensieren.

Bei der bisherigen Akquisen von Daten zu COVID-19 und SARS-CoV-2 gab es viele Versträmnisse, die aber nicht nur als Fehler begriffen werden sollten, sondern primär als Richtschnur für ihre künftige Vermeidung – auch abseits aktueller Bedrohungsszenarien.

Zu wenig aussagekräftige Daten
Die Corona-Krise ist auch eine „Missling-Dauer-Krise. Sie hat schmerzlich gezeigt, dass Deutschland über zu wenig Infrastruktur verfügt, um während einer Pandemie wichtige Daten zum Test- und Infektionsgeschehen und zur Wirksamkeit von Eindämmungs- und Therapiemaßnahmen zu akquirieren und zeitnah verfügbar zu machen. Dabei geht es nicht so sehr um Nebelthemen wie Infektionswege, klinische Diagnosen und

Todesmeldungen, sondern um die zahlreichen möglichen Störgrößen der Interpretation dieser Daten: geografische Herkunft der erkrankten beziehungsweise geisteten Personen, sozialer Hintergrund, Alter, Mobilität, Testgrund, lokale Prävalenz von Infektionen und symptomatischen Erkrankungen, lokale Testintensität und so weiter. Ein Außerachtlassen dieser Variablen führt potenziell nicht nur zu Fehldeutungen, es kann auch die Aussagekraft von Daten mindern und den Blick auf zielgerichtete, effektive Maßnahmen gegen die Pandemie verstellen.

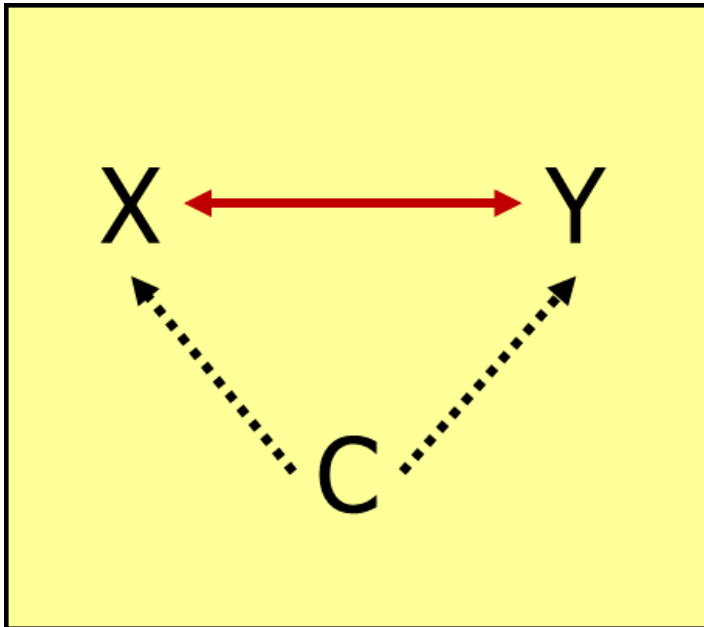
Zudem waren die meisten der derzeit verfügbaren Daten ursprünglich für andere Zwecke als das Pandemiemanagement gedacht. Sie können daher dem qualitativen Ansprechen hinsichtlich Vollständigkeit

Institut für Medizinische Informatik und Statistik, Universitätsklinikum Carl-Neuberg-Campus Göttingen, Prof. Dr. rer. med. Caliebe
Medizinische Fakultät, Albert-Ludwigs-Universität Freiburg, Institut für Diagnostik und Auswertungsrechnung in der Medizin (Senior Fellow) Prof. Dr. rer. med. Antes
Pflanzengarten, Göttingen, Universitätsklinikum Carl-Neuberg-Campus Göttingen, Prof. Dr. rer. med. Krawczak

Deutsche Ärztezeitung | 118 | Heft 10 | 12. März 2021

A 501

Confounder



Zahlreiche möglichen Störgrößen

- lokale Prävalenz von Infektionen
- lokale Prävalenz von symptomatischen Erkrankungen
- lokale Testintensität (PCR, Schnell)
- Testgrund
- Vergleichsgruppe
- zeitliche Dynamik
- Alter
- Morbidität
-

Studiendesign und Koordination

Medizin

COVID-19: Weitere kleinere Studie findet Hinweise auf Wirkung von Hydroxy-Chloroquin



Dtsch Ärzteblatt, April 2020

- Ad-hoc-Behandlungsstudien
- ohne Kontrollarm bzw standardisierte Kontrollbehandlung
- zu geringe Fallzahl
- mangelhaftes Studiendesign
- unkoordinierte Planung
- konkurrierende Rekrutierungsaktivitäten

„**widersprüchliche Ergebnisse**, die die **Verunsicherung der Bevölkerung** eher vermehren als verringern.“

Caliebe A, Antes G, Leverkus F, Krawczak M (2021) Dtsch Ärztebl 118: A501.

High risk of bias

Table 1 Basic characteristics of COVID-19-related, peer-reviewed original articles

Type of original article studies involving patients	Number of studies (%)	Number of patients Median (IQR)	Number of studies at risk of bias (%)		
			Low	Intermediate	High
Case-control	68 (9.5)	108 (62–211)	11 (16.2)	25 (36.7)	32 (47.1)
Cohort	50 (7.0)	110 (54–327)	7 (14.0)	20 (40.0)	23 (46.0)
Cross-sectional	306 (42.9)	217 (80–730)	10 (3.3)	43 (14.0)	253 (82.7)
Case series	129 (18.1)	18 (9–53)	9 (6.9)	26 (20.2)	94 (72.9)
Diagnostic	37 (5.2)	84 (49–215)	0 (0)	0 (0)	37 (100.0)
Prognostic	8 (1.1)	143 (66–217)	3 (37.5)	1 (12.5)	4 (50.0)
Simulation	185 (25.9)	1428 (14–40,696)	16 (8.6)	47 (25.4)	122 (66.0)
Non-randomized trial	8 (1.1)	35 (29–58)	1 (12.5)	1 (12.5)	6 (75.0)
Randomized controlled trial	4 (0.6)	56 (29–111)	0 (0)	2 (50.0)	2 (50.0)

Stand 01.05.2020

RCTs

The Landscape of Emerging Randomized Clinical Trial Evidence for COVID-19 Disease Stages: A Systematic Review of Global Trial Registries

The article was published in the following Dove Press journal:
Infection and Drug Resistance

Alison Dillman¹
Michael J Zoratti²
Jay JH Park³
Grace Hsu⁴
Louis Dron²
Gerald Smith²
Ofir Harari⁵
Craig R Rayner⁴
Noor-E Zannat²
Alind Gupta²
Eric Mackay⁵
Paul Arora⁴
Zelyn Lee⁶
Edward J Mills²

¹School of Public Health, Faculty of Medicine, Imperial College London, London, England; ²Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Canada; ³Department of Experimental Medicine, University of British Columbia, Vancouver, British Columbia, Canada; ⁴Monash Institute of Pharmaceutical Sciences, Monash University, Melbourne, Australia; ⁵Department of Statistical Sciences, University of Toronto, Toronto, Canada; ⁶Dalla Lana School of Public Health, University of Toronto, Toronto, Canada; ⁷Department of Physiology & Department of Neuroscience, University of Toronto, Toronto, Canada

Correspondence: Jay JH Park
Tel +1 404 294-3823
Email jayhpark1@alummi.ubc.ca

Purpose: A multitude of randomized controlled trials (RCTs) have emerged in response to the novel coronavirus disease (COVID-19) pandemic. Understanding the distribution of trials among various settings is important to guide future research priorities and efforts. The purpose of this review was to describe the emerging evidence base of COVID-19 RCTs by stages of disease progression, from pre-exposure to hospitalization.

Methods: We collated trial data across international registries: ClinicalTrials.gov; International Standard Randomised Controlled Trial Number Registry; Chinese Clinical Trial Registry; Clinical Research Information Service; EU Clinical Trials Register; Iranian Registry of Clinical Trials; Japan Primary Registries Network; German Clinical Trials Register (up to 7 October 2020). Active COVID-19 RCTs in international registries were eligible for inclusion. We extracted trial status, intervention(s), control, sample size, and clinical context to generate descriptive frequencies, network diagram illustrations, and statistical analyses including odds ratios and the Mann-Whitney U-test.

Results: Our search identified 11503 clinical trials registered for COVID-19 and identified 2388 RCTs. After excluding 45 suspended RCTs and 480 trials with unclear or unreported disease stages, 1863 active RCTs were included and categorized into four broad disease stages: pre-exposure (n=107); post-exposure (n=208); outpatient treatment (n=266); hospitalization, including the intensive care unit (n=1376). Across all disease stages, most trials had two arms (n=1500/1863, 80.52%), most often included (hydroxy)chloroquine (n=271/1863, 14.55%) and were US-based (n=408/1863, 21.90%). US-based trials had lower odds of including (hydroxy)chloroquine than trials in other countries (OR: 0.63, 95% CI: 0.45–0.90) and similar odds of having two arms compared to other geographic regions (OR: 1.05, 95% CI: 0.80–1.38).

Conclusion: There is a marked difference in the number of trials across settings, with limited studies on non-hospitalized persons. Focus on pre- and post-exposure, and outpatients, is worthwhile as a means of reducing infections and lessening the health, social, and economic burden of COVID-19.

Keywords: COVID-19, novel coronavirus 2019, systematic review, randomized controlled trials

Introduction

The COVID-19 pandemic, caused by the novel severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2), has mobilized the global health research community in an effort to rapidly identify preventative strategies and effective therapies.¹ Indeed, we are witnessing an international research community aligned with

- Stand 01.10.2020
1863 aktive RCTs
13% geschlossen

- 15 % Chloroquin

- sample size
Median 100 (IQR: 50, 308)

Digitalisierung und Datenschutz

Sekundärdatennutzung in der Forschung

Voraussetzungen (EU-DSGVO, Artikel 6)

Forschung mit medizinischen Versorgungsdaten ist verboten, außer

1. der Patienten/die Patientin hat informiert eingewilligt, oder
2. die Forschung steht in hinreichendem öffentlichen Interesse und wird durch ein Gesetz erlaubt.

Probleme

- Feststellung eines „öffentlichen Interesses“
- mangelnde Repräsentativität
- schwierige Verankerung in der klinischen Routine
- begrenztes inhaltliches Verständnis durch Patient*innen

angemessene Nutzen-Risiko-Abwägung

Datenspende: ein „dritter Weg“?

Wissenschaftliches Gutachten

„Datenspende“ – Bedarf für die Forschung,
ethische Bewertung, rechtliche,
informationstechnologische und
organisatorische Rahmenbedingungen

Erstellt für das Bundesministerium für Gesundheit



Erarbeitet und vorgelegt von der Gutachtergruppe

Prof. Dr. med. Dr. phil. Daniel Strech, Berlin;
Prof. Dr. Sebastian Graf von Kielmansegg, Kiel;
PD Dr. Sven Zenker, Bonn;
Prof. Dr. Michael Krawczak, Kiel;
Sebastian C. Semler, Berlin
Koordination: TMF e. V.

Version 1.1
Berlin, 30.03.2020

TMF – Technologie- und Methodenplattform
für die vernetzte medizinische Forschung e.V.



1. Datenspende zeitlich und räumlich **von medizinischer Behandlung entkoppeln**
2. Einwilligung durch **einfach auszuübendes Widerspruchsrecht** ersetzen („Opt-out“)
3. Entscheidung über durch **unabhängige, rechtlich legitimierte Use-and-Access-Committees**
4. bundeseinheitliche **infrastrukturelle Voraussetzungen schaffen** (z.B. Identitäts- und Statusmanagement, Treuhandstellen)

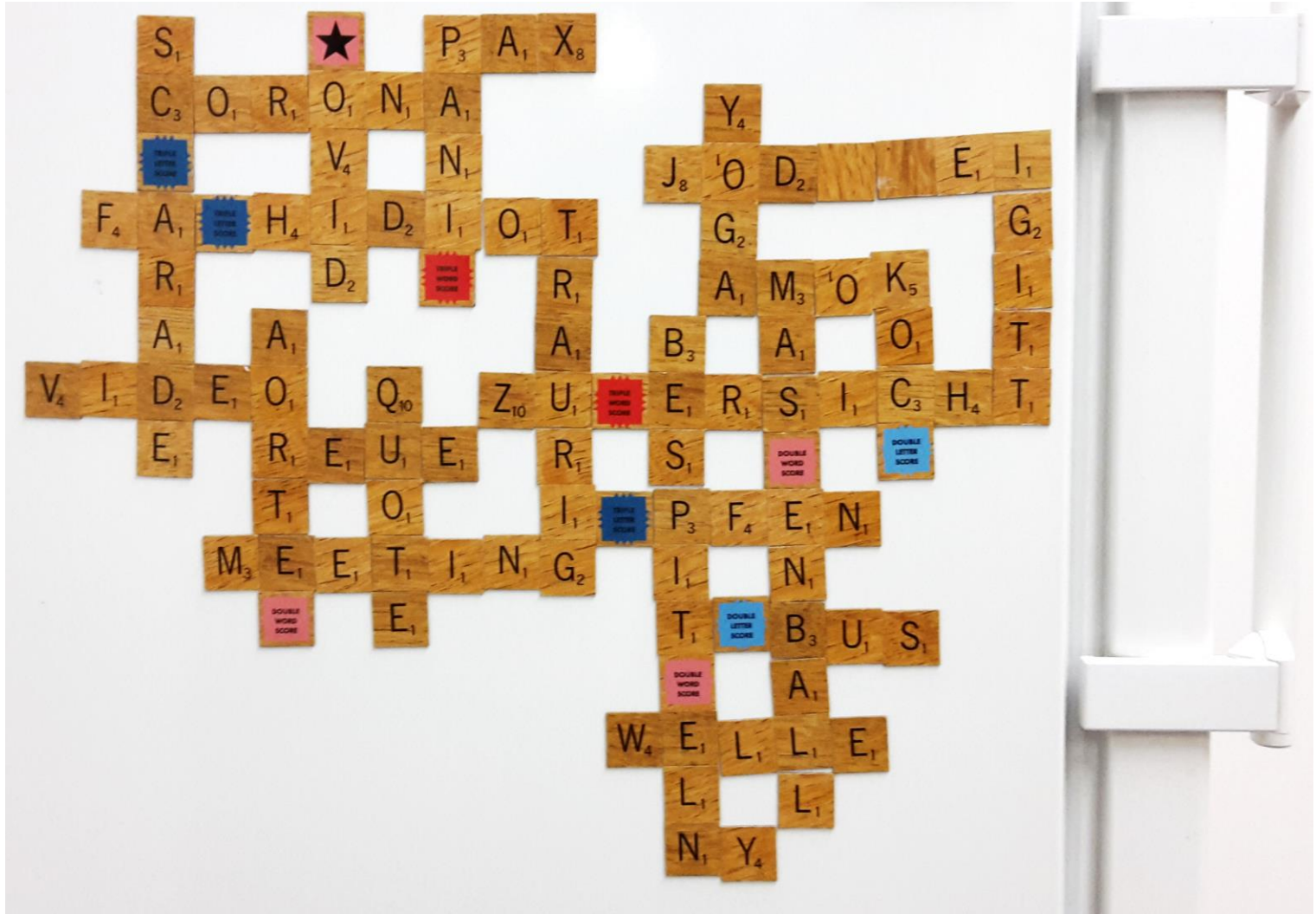
Fazit

Lernen für die nächste ~~Pandemie~~ Gesundheitskrise

- kein erfolgreiches Krisenmanagement ohne **Daten von hinreichender Qualität und Quantität**
- Daten unterstützen **Versachlichung der politischen Diskussion** und **Akzeptanz in der Bevölkerung**
- Vorantreiben der Digitalisierung in **Gesundheitswesen, öffentlicher Verwaltung und privatem Umfeld** erforderlich
- Notwendigkeit des **interdisziplinären Austauschs** wissenschaftlicher Expertise und praktischer Erfahrung in vielen Bereichen

Herausforderungen für Ethikkommissionen

- einarmige Studien
 - sehr kleine Studien
 - mangelnde Erfassung von Confoundern
 - Zeitdruck vs gut geplante Studien
-
- Eine verlässliche Datenbasis dient der **evidenz-basierten** Diskussion über politische Maßnahmen, Behandlungsmöglichkeiten, Impfstoffe
 - **Wir in der Ethikkommission** können mit dazu beitragen, dass eine solche verlässliche Datenbasis erzeugt wird.



Danke für Ihr Interesse