

Statistik in Pandemiezeiten

Prof. Dr. Thomas K. Bauer

Vortrag Wissenschaftsforum Ruhr

- Seit 2012: Hinterfragt jeden Monat jüngst publizierte Zahlen und deren Interpretation, um einen Beitrag zu einer verbesserten „Data-Literacy“ zu leisten.
- **Team:** Thomas Bauer (RWI), Gerd Gigerenzer (MPI – Harding-Zentrum für Risikokompetenz), Walter Krämer (TU Dortmund) und Katharina Schüller (STAT-UP)
- **Seit März 2020:** Aufklärung über zentrale Statistiken der Covid-19-Pandemie. Situation radikaler Unsicherheit!!



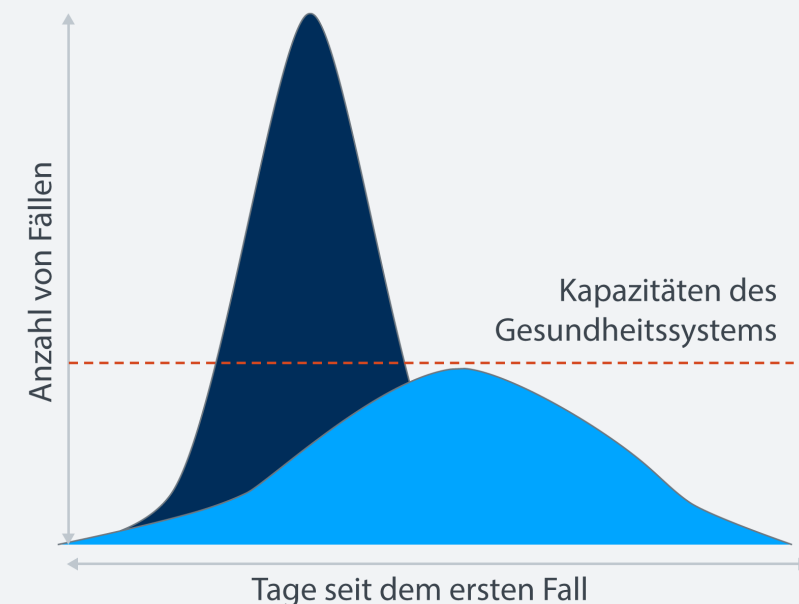
Die Reproduktionszahl

- **Reproduktionszahl R:** gibt an, wie viele andere Personen eine infizierte Person ansteckt. Ist $R = 2$, steckt jede infizierte Person im Durchschnitt zwei weitere Personen an. **Ziel:** $R < 1$.
- Aufgrund des exponentiellen Wachstums der Pandemie wurde dieser Kennzahl eine zentrale Rolle beigemessen.
- **Annahme:** Wachstumsrate von 23%: Bei 6.000 Infizierten (15. März 2020), wären innerhalb von 14 Tagen knapp 109.000, nach 30 Tagen nahezu 3 Millionen Personen infiziert.

COVID-19: Exponentielles Wachstum verhindern

Abstand halten – das Virus ausbremsen

- Pandemieverlauf ohne Schutzmaßnahmen
- Pandemieverlauf mit Schutzmaßnahmen



Die Reproduktionszahl

- **Schätzung von R ist nicht trivial:** Quotient aus Neuinfektionen des aktuellen und der vorangegangenen drei Tage und der Summe der Neuinfektionen der vier Tage zuvor.
- **Annahmen:**
 - **Generationszeit:** durchschnittliche Zeitspanne zwischen der Infektion einer Person und der Infektion der von ihr angesteckten Personen (4 Tage).
 - Es werden nicht die gemeldeten Neuinfektionen, sondern die daraus **geschätzten Neuerkrankungen** verwendet (Nowcasting) → **Schätzfehler**
 - **95% - Konfidenzintervall am 9. April 2020:** 0,8 – 1,1
 - **Dunkelziffer bleibt konstant:** Verändert sich mit der Anzahl der Tests (mehr Tests → geringere Dunkelziffer → R steigt) und dem Verhalten der Bevölkerung.
- **Forderung:** Repräsentative Panelstichprobe (wurde leider nicht realisiert).

Die 7-Tage Inzidenz

- **7-Tage-Inzidenz:** Innerhalb der vergangenen sieben Tagen registrierten Neuinfektionen je 100.000 Einwohner. Hat eine Stadt mit 250.000 Einwohner in den letzten sieben Tagen 50 Neuinfektionen verzeichnet, so beträgt die Sieben-Tages-Inzidenz $50 * (100.000 / 250.000) = 20$.
- Steht im Fokus der derzeitigen Politik.
- Siehe Dash-Board des RKI.

Probleme:

- **Anzahl der Tests** (je mehr, desto mehr positive Ergebnisse):
Ende April und Anfang September 2020 etwa 8.300 Neuinfektionen, aber im September mehr als dreimal so viele Test.
→ **Positiv-Test-Rate:** April 2,5%, September 0,8%, Oktober 3,6%
- **Andere Personengruppen werden getestet:** April nur Personen mit Symptomen, später vermehrt Massentests (Dunkelziffer wird besser erfasst).

Probleme:

- **Neuinfektionen im Verhältnis zur Anzahl der Verstorbenen setzen:**
Steigende Zahl von Neuinfektionen, aber Verhältnis der Verstorbenen zu den zwei Wochen zuvor Infizierten sinkt (für alle Altersgruppen) → Verbesserte Behandlung (vorschnelle Beatmung).
- **Schlussfolgerung:** Nicht auf eine einzige Zahl blicken!
Rechtfertigung 7-Tage-Inzidenz: Nachverfolgung der Kontaktpersonen durch Gesundheitsämter. Mindestens ebenso bedeutsam: Veränderung der Positiv-Test-Raten und Todesraten, Anteil Corona-Patienten in Intensivstationen

- **Coronamortalität:**

- Statista: 3,0%
- Deutsches Ärzteblatt: 1,4%
- Stanford Universität: < 0,5%

- **Probleme:**

- Zähler und Nenner **nicht einfach zu messen** (Zähler: man müsste die an Corona und nicht die mit Corona verstorbenen Menschen zählen).
- **Zähler ist eine „Flussgröße“, der Nenner eine „Bestandsgröße“:** Die innerhalb eines bestimmten Zeitraums Verstorbenen geteilt durch eine bestimmte Anzahl Menschen an einem bestimmten Tag.

Aber welcher Zeitraum, welche Menschen und an welchem Tag? Alle Menschen eines Landes insgesamt, die mit dem Coronavirus Infizierten oder die an Corona tatsächlich auch Erkrankten?

- **Mortalität vs. Letalität:**
 - a) Mortalität:** Todesfälle in Bezug zur Population
 - b) Letalität:** Todesfälle in Bezug zur Anzahl der Erkrankten
 - c) Fall-Verstorbenen-Anteil (Case Fatality-Rate, CFR):** Verstorbene in Bezug zu den bestätigten Fällen
 - d) CFR:** nur tatsächlich Erkrankte oder auch bestätigte Infektionen ohne Symptome?
 - e) Infizierten-Verstorbenen-Anteil (Infection Fatality Rate):** berücksichtigt die Dunkelziffer der klinisch relevanten, aber nicht diagnostizierten Fälle.

- **Berücksichtigte Zeitspanne:** Verstorbene an einem Tag, einer Woche, einem Monat, oder dem kompletten Zeitraum seit Beginn der Pandemie?

Statistisches Bundesamt: Zahl der Verstorbenen seit Beginn der Pandemie geteilt durch die Bevölkerung an einem bestimmten Tag → theoretisch Mortalitätsraten von über 100% möglich.

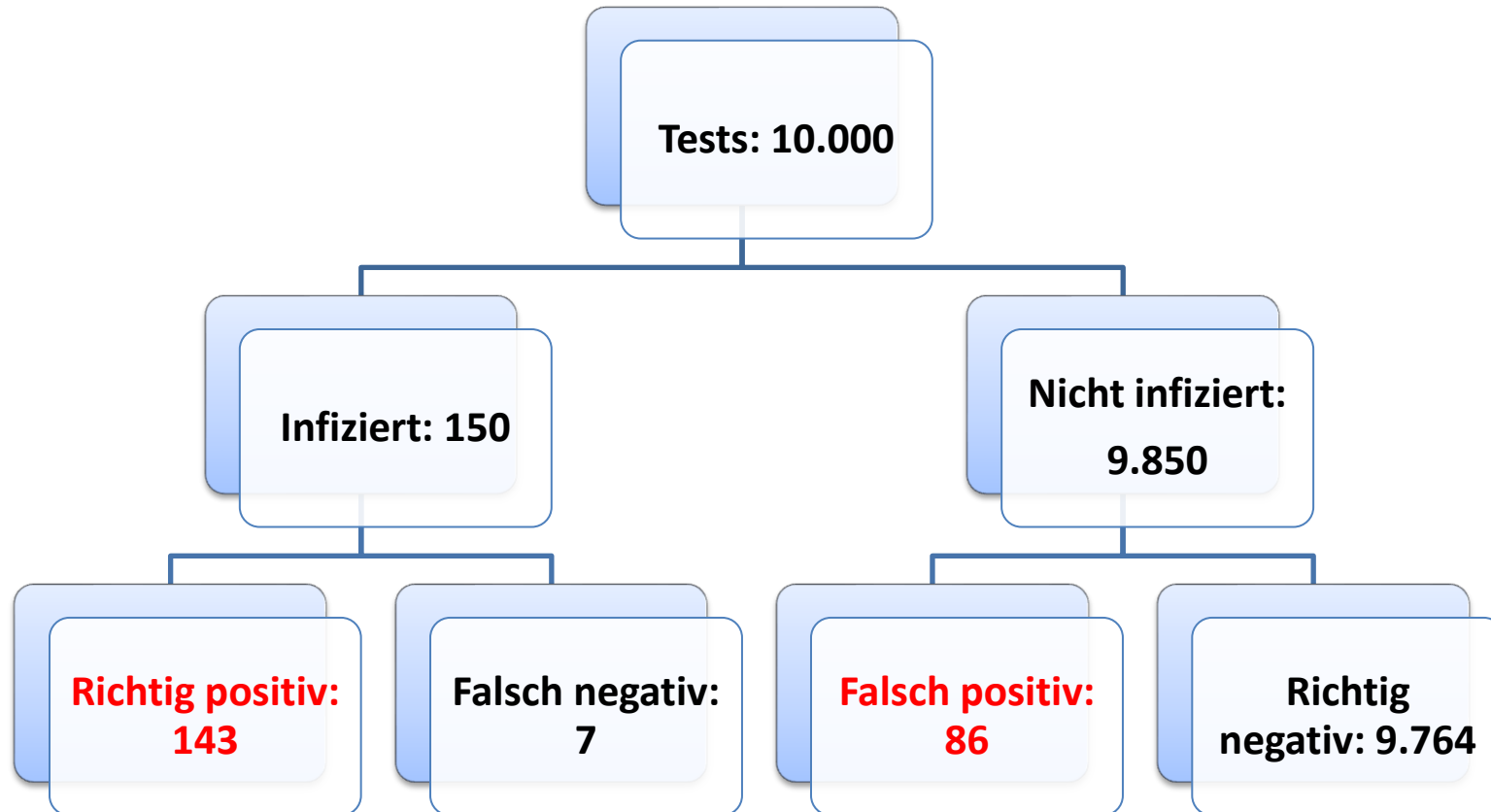
- Selbst bei einheitlicher Definition sind internationale Vergleiche aufgrund unterschiedlicher Altersstruktur der Bevölkerung (Amerikaner sind im Schnitt 6 Jahre jünger) nicht möglich → **standardisierte Altersstruktur.**

- **Jeder Test kann zwei Fehler machen:** Er kann eine Infektion übersehen (falsch-negativ) oder eine nicht-infizierte Person als infiziert (falsch-positiv) klassifizieren.
- Tests liefern üblicherweise zwei Angaben:
 - **Sensitivität:** Wahrscheinlich, dass ein Infizierter richtig positiv getestet wird.
 - **Spezifität:** Wahrscheinlichkeit, dass ein Nicht-Infizierter als solcher erkannt wird.
 - **Falsch-Positiv-Rate:** Gegenwahrscheinlichkeit zur Spezifität.
- Angaben des Schnelltests auf meinem Schreibtisch:
 - **Sensitivität:** 95,3%
 - **Spezifität:** 99,13%

Frage: Mit welcher Wahrscheinlichkeit bin ich bei einem positiven Testergebnis infiziert?

Einordnung von Schnelltest

Die Antwort auf diese Frage hängt von der Inzidenzrate (Vortestwahrscheinlichkeit) ab.
Annahme: 150 von 10.000 Getesteten sind tatsächlich infiziert.



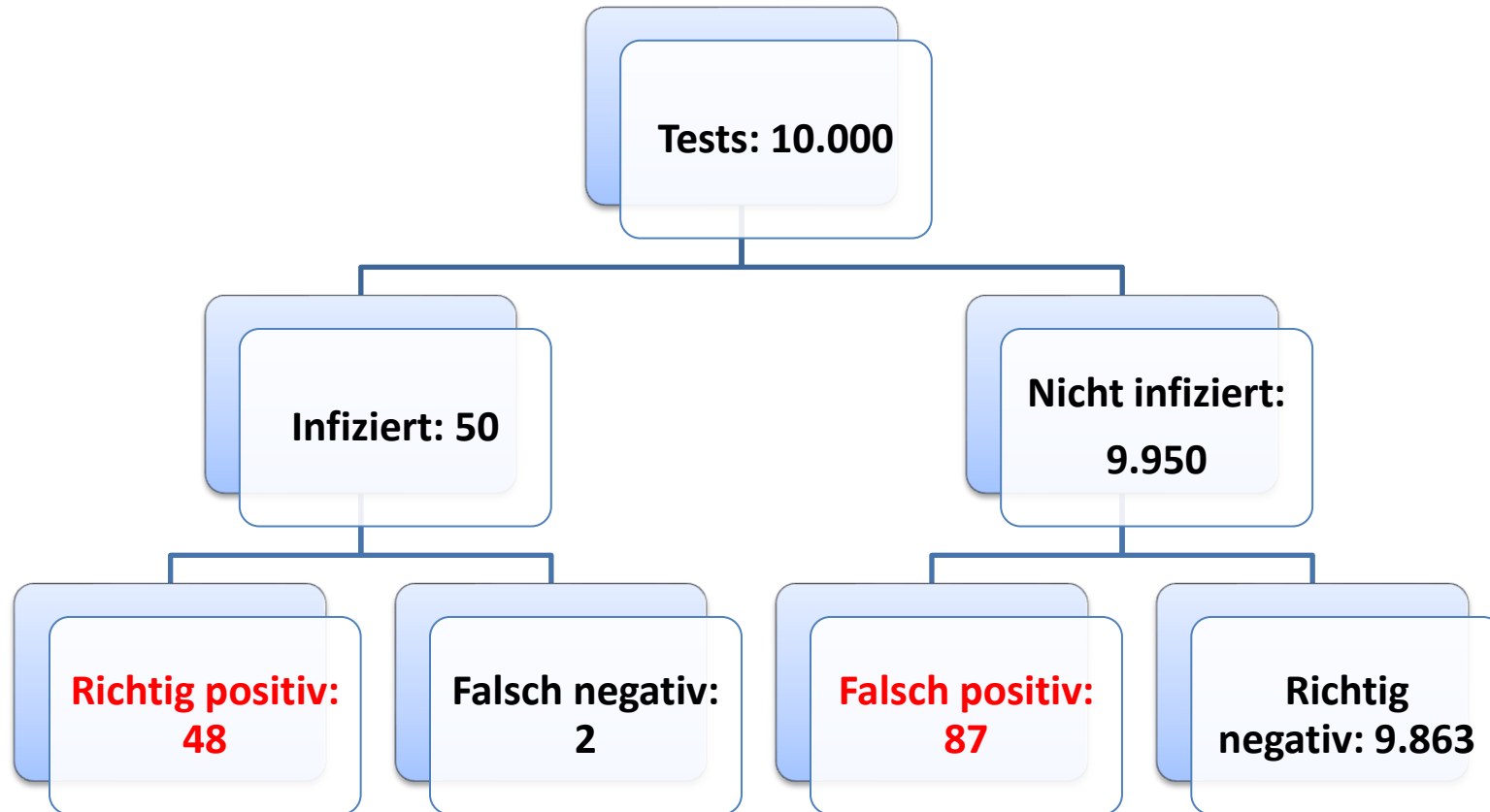
Ein positives Testergebnis bedeutet, das ich mit Wahrscheinlichkeit

$$P = 100 \times \frac{143}{143 + 86} = \mathbf{62,44\%}$$

infiziert bin.

Einordnung von Schnelltest

Die Antwort auf diese Frage hängt von der Inzidenzrate (Vortestwahrscheinlichkeit) ab.
Annahme: 50 von 10.000 Getesteten sind tatsächlich infiziert.



Ein positives Testergebnis bedeutet, das ich mit Wahrscheinlichkeit

$$P = 100 \times \frac{48}{48 + 87} = 35,56\%$$

infiziert bin.

Was bedeutet „Ein Impfstoff ist zu 90 Prozent wirksam?“

- **Falsch:** „9 von 10 Menschen können durch die Impfung vor einer Infektion geschützt werden.“
Das würde bedeuten: wenn man alle 83 Millionen Deutschen impft, dann sind davon 90 Prozent geschützt; nur die restlichen 8,3 Millionen können sich anstecken. Das wären aber immer noch weit mehr Erkrankte als es bisher der Fall ist.
- **Richtig:** Relative Risikoreduktion. Studienprotokoll von Pfizer:

$$\text{Wirksamkeit} = 100 \times \left(1 - \frac{\text{Anteil der Covid-19-Fälle in der Impfgruppe}}{\text{Anteil der Covid-19-Fälle in der Kontrollgruppe}} \right)$$

Teilnehmer: 43.000, ~ 50% in der Kontrollgruppe. Impfgruppe: 8 Infizierte; Placebogruppe: 86 Infizierte.

Angela Merkel zu Mutation B.1.1.7 des Corona-Virus: „Wenn wir es nicht schaffen, dieses britische Virus abzuhalten, dann haben wir bis Ostern eine zehnfache Inzidenz“

War das starker Tobak?

- **Annahmen:**
 - B.1.1.7 ist ungefähr 35-70% ansteckender als die bisher bekannte Form.
 - Anteil der mit der neuen Virusmutation infizierten Bevölkerung: 0,1 – 1 %
 - Exponentielles Wachstum, R-Wert von 1,1, 7-Tages-Inzidenz von 164

→ Bis Ostern:

Inzidenzwert zwischen 500 und 3.500 Infektionen je 100.000 Einwohner.
Mittelwert entspricht der Vorhersage von Frau Merkel.

Probleme der Prognose:

- **Mangelhafte Qualität** der zugrundeliegenden Daten (geschätzte Ansteckungsgefahr, Anteil der mit der Mutation infizierten Personen) → hohe Prognoseunsicherheit
- **Zentrale Annahme von Prognosen:** Die Zukunft verläuft wie die Vergangenheit

Darin liegt gerade **eine der zentralen Aufgaben einer Prognose:** Man schätzt die zukünftige Entwicklung, wenn alles so bleibt, wie es ist, um auf Basis dieser Prognose Handlungsnotwendigkeiten und -optionen diskutieren zu können.

Damit ist aber jedwede Prognose bereits mit ihrer Veröffentlichung notwendigerweise falsch!

(Verhaltensänderungen aufgrund der Prognose → Zukunft unterscheidet sich von Vergangenheit)

Vergangene Prognosen ex-post zu kritisieren ist unfair – die Prognose hat mehr oder weniger selbst dazu beigetragen, dass sie nicht stimmt.