

**Prof. Dr. Werner Bergholz**

**Deutscher Bundestag**  
Ausschuss f. Gesundheit

Ausschussdrucksache  
**19(14)233(4)**  
zur 6Anh am 28.10.2020 -  
Covid-19 Teststrategie  
30.10.2020

## **Schriftliche Stellungnahme**

**Öffentliche Anhörung des**

**Ausschusses für**

**Gesundheit**

**des Deutschen Bundestages**

**zum Antrag der Abgeordneten Dr. Andrew Ullmann, Michael Theurer, Renata Alt, Christine Aschenberg-Dugnus, Nicole Bauer, Jens Beeck, Dr. Jens Brandenburg (Rhein-Neckar), Mario Brandenburg (Südpfalz), Dr. Marco Buschmann, Carl-Julius Cronenberg, Britta Katharina Dassler, Dr. Marcus Faber, Daniel Föst, Otto Fricke, Thomas Hacker, Reginald Hanke, Peter Heidt, Markus Herbrand, Torsten Herbst, Katja Hessel, Manuel Höferlin, Reinhard Houben, Ulla Ihnen, Olaf in der Beek, Dr. Christian Jung, Karsten Klein, Carina Konrad, Konstantin Kuhle, Till Mansmann, Roman Müller-Böhm, Dr. Martin Neumann, Matthias Nölke, Hagen Reinhold, Dr. Wieland Schinnenburg, Matthias Seestern-Pauly, Frank Sitta, Judith Skudelny, Dr. Hermann Otto Solms, Benjamin Strasser, Katja Suding, Stephan Thomae, Dr. Florian Toncar, Gerald Ullrich, Johannes Vogel (Olpe), Sandra Weeser, Nicole Westig, Katharina Willkomm und der Fraktion der FDP**

**Praxistaugliche und intelligente COVID-19-Teststrategie**

## Inhaltsübersicht

1. **Zusammenfassung** 3
2. **Stärken des FDP Konzepts**
3. **Mängel des momentanen Testkonzepts , Fünf Punkte Plan zur Beseitigung**
4. **Schlussfolgerungen und Handlungsempfehlungen**

## 1. Zusammenfassung

Die Stellungnahme ergänzt das FDP - Konzept für die Ausgestaltung einer praxistauglichen und intelligenten Teststrategie für Covid -19 um **fünf essentielle Anforderungen**, damit gravierende Mängel der bisherigen Teststrategie beseitigt werden. Sonst wird das Ziel einer effektiven und effizienten Teststrategie zur Ableitung angemessener Maßnahmen verfehlt:

- 1) Umgehende **Standardisierung des PCR – Test Verfahrens**, sonst gibt es keine zeitliche und örtliche **Vergleichbarkeit der Daten**.
- 2) Laufendes **Monitoring der Spezifität und Sensitivität** des Tests für jedes beteiligte Labor und **Berücksichtigung der falsch positiven Rate** bei der statistischen Analyse auf kommunaler, Landes- und Bundesebene, sonst entstehen **irreführende Trends**.
- 3) Die **Kennzahl „Fälle in den letzten 7 Tagen pro 100.000 Einwohner“** ist **keine valide Messgröße**, sondern eine Zahl, die direkt von der Anzahl der Testungen abhängt. Abhilfe durch Berücksichtigung der **Anzahl Tests in 7 Tagen pro 100.000 Einwohner auf kommunaler Ebene**.
- 4) Die laufende **Bewertung des Risikos** muss auf der Basis der Kennzahlen „Fälle“ (= **positive Tests**), **Hospitalisierung** und **Todesfälle** nach einer nachvollziehbaren, transparenten und **standardisierten Systematik** aufbauen. Eine solche Systematik ist zurzeit nicht erkennbar, die momentan festzustellende „eindimensionale“ Fixierung auf die Kennzahl Fälle bei der Risikobewertung führt mit Sicherheit zu **Fehlsteuerungen**.
- 5) Die **wissenschaftliche Aufarbeitung** des Infektionsgeschehens ist nur in Ansätzen erkennbar, dies betrifft ausdrücklich auch die nach wie vor **ausstehende wissenschaftlich fundierte Validierung des PCR Tests**. Ohne eine laufende Bewertung aller relevanten Fakten zum Aufbau eines **fundierten evidenzbasierten Modells** bleibt die Steuerung von Gegenmaßnahmen im besten Fall Stückwerk, im schlimmsten Fall resultiert eine **gravierende Fehlsteuerung mit vermeidbaren Kollateralschäden**.

Ohne konsequente und schnelle Umsetzung der Maßnahmen 1) – 5) gibt es **keine wirklich aussagefähigen und belastbaren Zahlen zum Infektionsgeschehen** und es sind **keine**

**angemessenen staatlichen Maßnahmen möglich.** Die Umsetzung von 4) und 5) erfordert zwingend den Einsatz von „**Interdisziplinären Teams**“, in denen alle relevanten Kompetenzen vertreten sind.

**Im momentanen Zustand entspricht die Teststrategie in keiner Weise den Qualitätsanforderungen der Technik oder dem Stand der Wissenschaft.**

Die 5 Anforderungen sind von 4 relevanten internationalen Normdokumenten abgeleitet:

- ISO 31000: „Risk Management“ (<https://www.iso.org/iso-31000-risk-management.html>)
- ISO 9001-2015: “Quality management systems — Requirements” (<https://www.iso.org/standard/62085.html>)
- ISO 15189:2012: “Medical laboratories — Requirements for quality and competence” (<https://www.iso.org/standard/56115.html>)
- JCGM 100:2008 “Evaluation of measurement - — Guide to the expression of uncertainty in measurement” (<https://www.iso.org/sites/JCGM/GUM/JCGM100/C045315e.html/C045315e.html?csnumber=50461>)

## 2. Stärken des FDP Konzepts

Das von der FDP Fraktion vorgeschlagene Konzept ist definitiv eine Verbesserung. Die m.E. drei wichtigsten Elemente des Konzepts sind:

1. Vermeidung von anlasslosem Testen und vordringlich sollten Patienten mit einschlägigen Symptomen getestet werden
2. Umgehende Testung von Kontaktpersonen (*Ergänzung von mir: sofern klinische Symptome bei der positiv getesteten „Primär“-Person vorliegen*)
3. Stärkung Verantwortung der Gesundheitsämter

Es gibt aber darüber hinaus selbst unter Berücksichtigung der Vorschläge der FDP Fraktion **so gravierende Mängel** im momentanen Testkonzept, dass das Konzept bei weitem nicht technische und wissenschaftliche Standards erfüllt. Dies wird im folgenden Abschnitt 3 dargelegt.

## 3. Mängel des momentanen Testkonzepts , Fünf Punkte Plan zur Beseitigung

Vorbemerkung:

Das Testkonzept wird in seiner Gesamtheit, von der Probennahme bis zur Bewertung der Ergebnisse und der aus den Ergebnissen zu folgernden Maßnahmen betrachtet.

### 3.1 Standardisierung des PCR – Tests

Es bedarf keiner weiteren Erläuterung, dass ein Monitoring des Infektionsgeschehens durch positive Testergebnisse eine **Vergleichbarkeit**

- von Labor zu Labor
- zeitlich (für die Interpretation von Trends in Zeitreihen)
- örtlich (Vergleich von Betroffenheit in Landkreisen und Stadtkreisen)

als **unverzichtbare Anforderung** hat.

Nach der Norm ISO 9001 für Qualitätsmanagementsysteme allgemein und der Norm ISO 15189 für medizinische Labore ist eine „conditio sine qua non“, um dies zu erreichen, dass der Prozess von der Anforderung der Probenahme bis zur Mitteilung der ausgewerteten Ergebnisse nach einem standardisierten **immer in der gleichen Art ablaufenden Prozess mit standardisierten Hilfsstoffen** zu erfolgen hat.

Die folgenden Anforderungen der Normen sind beim PCR Test für SARS Cov-2 **nicht** erfüllt:

- Eine Probennahme in nicht dafür vorgesehenen Räumen (Flughäfen, Bahnhöfen) oder gar im Freien an Autobahnen stellt eine signifikante **Abweichung von einer definierten Testumgebung** statt. Ein geeigneter Raum ist Voraussetzung dafür, dass kein Eintrag von Fremdmaterial, das zu falsch positiven Ergebnissen führen kann, verhindert bzw. reduziert wird.
- Jedes Labor definiert zurzeit seinen eigenen  $c_t$  Wert ( $c_t$  = cycle threshold, bei dem entweder ein Signal vorliegt (= Test positiv) oder nicht (= Test negativ)). Der Einfluss dieses Parameters auf den Schwellwert, ab dem ein Testergebnis als positiv gewertet wird, ist signifikant, nach einem Bericht der New York Times /Ref. 1/ um bis zu einem Faktor 3.
- Die Reagenzien (z.B. die sogenannten Primer) für den Test werden entweder von einer der diversen Hersteller bezogen oder bei entsprechendem Know How und Infrastruktur selbst hergestellt, es liegen also keine standardisierten Hilfsmaterialien vor. Eine wissenschaftliche Veröffentlichung, die den Einfluss der Primer untersucht, /Ref. 2/ kommt zum Ergebnis, dass je nach Lieferanten die Sensitivität um einen Faktor 10 oder mehr variieren kann, was zu unterschiedlichen positiven und negativen Raten führt, für die selbe zu untersuchende Probe!
- Der eigentliche Analyseprozess (welche Teile des Genoms werden zur Identifizierung genutzt) wird verschieden gehandhabt, und die Umweltbedingungen dafür (Temperatur, relative Luftfeuchte, Reinraum oder Nichtreinraum) sind ebenfalls nicht festgelegt, auch hier liegt mit Sicherheit ein Einfluss auf das Testergebnis vor.
- Die Auswertung der Rohdaten, das Umgehen mit Anzeichen von Abweichungen vom normalen Analyseprozess und Inhalt und Format des Berichts an die Auftraggeber ist nicht festgelegt. Das heißt, es werden teilweise unvollständige Daten übermittelt.

**Handlungsbedarf 1: Standardisierung des PCR Tests**

### 3.2 Kein Monitoring der falsch positiv und falsch negativ Raten

Jede Messtechnik ist mit einer Messunsicherheit behaftet. Die Messunsicherheit führt zu einer Abweichung des Messwerts von dem wahren Wert. Eine Messung ohne die Angabe der Messunsicherheit ist eigentlich wertlos, es sei denn es gibt wenigstens eine grobe Abschätzung des Messfehlers. **In der Wissenschaft ist ein Messwert ohne Angabe der Unsicherheit nicht akzeptabel.**

Im Fall der PCR Messung ist das Messergebnis eine binäre Aussage: positiv oder negativ. Der Messfehler wird in diesem speziellen Fall durch die Sensitivität und Spezifität determiniert, und wird **positiver bzw. negativer Vorhersagewert** genannt, also mit welcher Wahrscheinlichkeit der Test richtig ist.

Das Problem bei einer geringen Inzidenz im unteren einstelligen Prozentbereich ergeben sich überraschend schlechte positive Vorhersagewerte, folgendes Zahlenbeispiel:

Bei 10 000 Tests, 1% wirklich Infizierte (also echt Positive), einer Spezifität von 1,4% und 98% Sensitivität (Werte aus dem Ringversuch vom April) ergibt der Test 98 echt Positive und zusätzlich 140 falsch Positive. Das heißt der positive Vorhersagewert (PPV = positive predictive value) beträgt  $98/(98+140) \times 100\% = 41\%$ . Mit anderen Worten, es gibt mehr falsch Positive als echt Positive, also ein **Messfehler grösser als ein Faktor 2.**

**Es ist schlicht unverständlich, dass dies bei der Aufbereitung der Rohdaten unberücksichtigt bleibt!**

**Handlungsbedarf 2:** Die Labore müssen eine angemessene Zahl Kontrollproben verarbeiten, damit die **Sensitivität und Spezifität** regelmäßig bestimmt werden und damit aus den Rohdaten die wirklich positiven Befunde errechnet werden können.

### 3.3 keine zeitliche und örtliche Vergleichbarkeit der Anzahl positiver Tests pro 7 Tage und pro 100 000 Einwohner

Ein dramatischer Mangel an der bisherigen Strategie ist die Definition der Kennzahl „**Positiver Tests pro 7 Tage pro 100 000 Einwohner**“. Der Mangel besteht darin, dass diese Zahl **keine sinnvoll definierte, also keine valide Messgröße** ist, denn:

Eine elementare Regel bei Messung auf der Basis einer Stichprobe ist, dass der festgestellte Zahlenwert **nicht von der Größe der Stichprobe abhängen darf**. Sondern man muss entweder den Prozentsatz positiver Befunde errechnen, und das ist dann eine valide Messgröße, oder alternativ kann man das Ergebnis auf eine bestimmte Stichprobengröße (z.B. auf 1000 Test pro 100 000 Einwohner) normieren.

Praktisches Beispiel: Angenommen, in Woche 1 werden in einem Landkreis mit 100 000 Einwohnern 2500 Tests durchgeführt und man erhält 25 positive Tests, in Woche 2 werden 5000 Tests durchgeführt,

und man erhält 50 positive Tests. Das Infektionsgeschehen ist identisch (nämlich 1% positiver Anteil), aber die Kennzahl ist doppelt so hoch, **also falsch und irreführend**.

Dass die Zahlen in einer Zeitreihe oder bei einem Vergleich von Ort zu Ort wirklich vergleichbar sein müssen, ist unmittelbar einleuchtend und ein absolute MUSS – Bedingung. Diese Anforderung ist aber durch die unterschiedlichen Testraten in den Land- und Stadtkreisen auf gröbste Weise verletzt und damit ein unhaltbarer Zustand, der umgehend abgestellt werden muss! Noch schlimmer: die Zahl „Fälle“ in den letzten 7 Tagen pro 100.000 Einwohner kann durch willkürliche Erhöhung oder Erniedrigung der lokalen Testzahlen nach Belieben manipuliert werden!

**Handlungsbedarf 3 (dringendst):** Kennzahl „Fälle pro 7 Tage und 100 000 Einwohner“ durch eine messtechnisch und mathematisch valide Kennzahl ersetzen!

### 3.4 keine Systematik und Transparenz bei der Risikobewertung

In dem Dokument des RKI zu Covid-19 vom 7. Juli /Ref. 3/ zu den Laborkapazitäten wird auch die Kommunikation des Infektionsgeschehens angesprochen:

**„Information der Bevölkerung: Die Bevölkerung soll über die Möglichkeiten und Grenzen von Tests geeignet informiert werden (s. KBV und ggf. BZgA, sowie die Informationsangebote der Länder). Eine effiziente Teststrategie schließt die Information der Bürger über die Situationen, in denen eine Testung tatsächlich sinnvoll ist, ein.“**

Dem ist eigentlich nichts mehr hinzuzufügen, außer dem ausdrücklichen Hinweis, dass auf Land- und Stadtkreisebene keine Informationen zur Zahl der Tests pro 7 Tage bzw. pro Tag öffentlich zugänglich ist, somit keine Transparenz besteht, ob ein Anstieg von positiven Tests auf lokaler Ebene einen realen Anstieg des Infektionsgeschehens wiedergibt oder ausschließlich durch die Erhöhung der Testzahlen verursacht wird. Die Zahlen müssen auf lokaler Ebene vorliegen, da den Gesundheitsämtern nicht nur die Zahl der positiven Tests, sondern auch die Zahl der negativen Tests gemeldet werden müssen, zur Weitergabe an das RKI.

**Handlungsbedarf 4 (dringend):** Anzahl der Tests für alle Land- und Stadtkreise wird lokal bekannt gegeben und auf der RKI Webseite tagesfein veröffentlicht. Nur so kann das RKI ein Risiko-Management implementieren, das die eigenen Vorgaben erfüllt.

Die Risikobewertung soll nach den Vorgaben des offiziellen Pandemieplans des RKI /Ref. 4/ nach 3 Kriterien (mit Unterkriterien, die hier nicht gelistet werden) erfolgen:

- 1) „Epidemisches Potenzial in der Bevölkerung“
- 2) „Epidemiologisches (Schwere)-Profil“

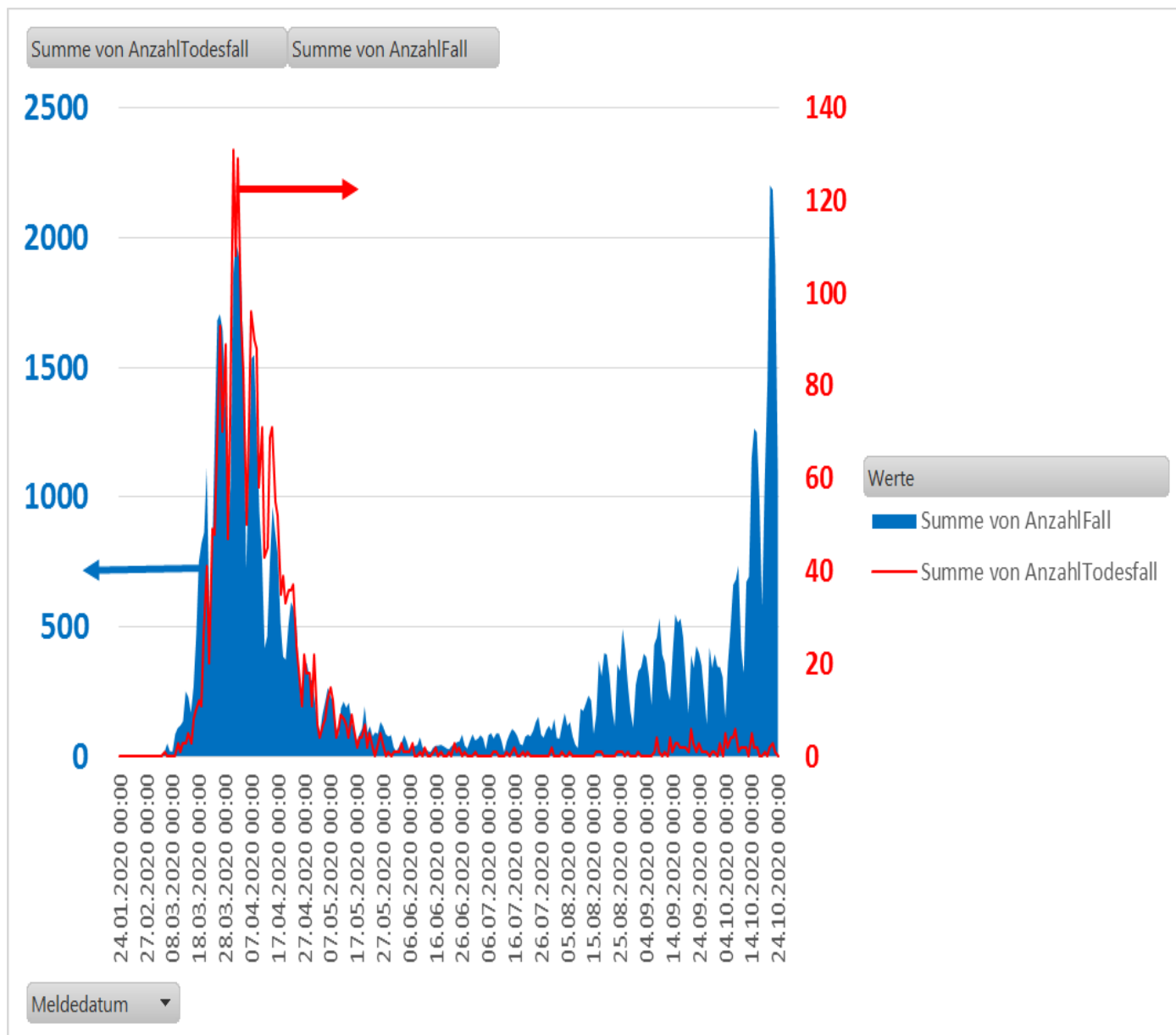
### 3) „Ressourcenbelastung“

Aus diesen 3 Hauptkriterien wird in den täglichen Situationsberichten des RKI eine Gesamtrisikobewertung entwickelt. Es ist aber nicht nachvollziehbar, dass einerseits von Mai bis September die Kennzahlen bezüglich der 3 Hauptkriterien immer kleiner wurden, aber die Risikobewertung konstant „hoch“ blieb.

Hier ist aus Gründen der Effektivität, der Glaubwürdigkeit und der Transparenz unbedingt eine **standardisierte Vorgehensweise für eine quantitative Risikobewertung** einzuführen und anzuwenden. Eine sehr einfache Möglichkeit wäre, ein im Qualitätsmanagement allgemein übliches Werkzeug, die FMEA, anzupassen /Ref. 5/. Die **FMEA (Fehlermöglichkeits- und Einflussanalyse)** verwendet ebenfalls 3 Kenngrößen zur Risikobewertung nach standardisierten Wertetabellen, und hat als Ergebnis eine **Risikoprioritätszahl**. Diese Zahl kann dann als Zeitreihe in einer Graphik die Entwicklung der Risikobewertung griffig und transparent darstellen. Eine Anpassung der FMEA für QM und RM auf die Bewertung des Covid-19 Risikos wäre in kurzer Zeit zu leisten.

Es versteht sich von selbst, dass in der Risikobewertung die Anzahl positiver Tests als „Frühindikator“, der für sich **keine Infektion nachweist**, überbewertet werden darf. Vielmehr ist **mit größerem Gewicht als für diesen Frühindikator** die tatsächliche Hospitalisierungsrate und die Sterberate (nach Abzug von Unfallopfern und andere Sterbeursachen, die keinen Zusammenhang mit einer Infektion haben) in die Risikobewertung einzubeziehen. Denn nur diese beiden Parameter bilden die Gefährlichkeit des Infektionsgeschehens **direkt** ab.

Ein aktueller Vergleich der Anzahl positiver Tests pro Tag und der Verstorbenen an oder mit Covid-19 zeigt sehr anschaulich, warum dies so ist. Die Fallzahlen steigen dramatisch, die Anzahl der Verstorbenen zeigt ein komplett anderes Bild.



**Abbildung 1: Vergleich der täglichen positiven PCR Tests in Bayern (die Graphik für Deutschland sieht ähnlich aus) mit der Anzahl pro Tag an oder mit Covid-19 Verstorbenen, Datenbasis RKI Datenbank, Download am 25.10.2020. Es ist sehr wahrscheinlich, dass von Mai bis August überwiegend falsch positive Befunde vorlagen.**

Darüber hinaus wird empfohlen, den Frühindikator „Anzahl positiver Tests“ solange aus der Risikobewertung auszuschließen, solange die Handlungsempfehlungen 1 – 4 nicht umgesetzt sind, da beim momentanen Zustand die Fallzahlen eine stark verringerte Aussagekraft haben.

**Handlungsbedarf 5:** Risikobewertung standardisieren und den Indikatoren Hospitalisierungsrate und Sterberate mehr Gewicht geben.



### 3.5 Laufende wissenschaftliche Analyse des Gesamtgeschehens und kontinuierliche Verfeinerung eines fundierten evidenzbasierten Modells

Bei jeder komplexen Risikolage ist lt. ISO 31000 eine regelmäßig aktualisierte Bewertung des Risikogeschehens einschließlich einer wissenschaftlichen Durchdringung und einer Modellvorstellung von den zugrundeliegenden Mechanismen gefordert. Ein speziell für komplexe Probleme entwickeltes Werkzeug aus dem Qualitätsmanagement, die 8D Analyse ist sehr gut einsetzbar /Ref.5/. Es gilt auch bei 8D das wissenschaftliche Grundprinzip, das ein einziger Widerspruch zwischen Beobachtungen und Modell das jeweils favorisierte Modell zu Fall bringt und ein verbessertes Modell entwickelt werden muss.

Konkret bedeutet das, dass **sämtliche** zeitlichen und räumlichen Phänomene erklärbar sein müssen. Als Beispiel für auffällige Beobachtungen seien genannt die ungewöhnliche Häufung positiver Tests im Landkreis Berchtesgaden und als weiteres Beispiel die ungewöhnlich hohen aktuellen Inzidenzen im Emsland, also einem relativ dünn besiedeltem Gebiet – im Gegensatz zum Landkreis Wittmund, in dem die 7 Tage Inzidenz am 21.10. Null war. Eine Spezialität der 8D Methode ist, dass „Problem nicht vorhanden“ wie im Landkreis Wittmund gleichwertig mit einem örtlich besonders ausgeprägten Problem zu berücksichtigen und erklären ist.

Ein wichtige „offene Flanke“ bei der Charakterisierung des Infektionsgeschehens ist die Tatsache, dass nach Aussagen von Prof. Tanner aus der Schweiz es immer noch keine Validierung des PCR Tests gibt, obwohl seit dem großflächigen Einsatz schon über 6 Monate vergangen sind.

**Handlungsbedarf 6:** Systematisch wissenschaftliche Aufarbeitung aller Daten und Fakten in einem interdisziplinären Team

## 4. Schlussfolgerungen und Handlungsempfehlungen

Selbst unter Berücksichtigung der Vorschläge der FDP Fraktion weist das momentane Testkonzept so gravierende Mängel auf, dass es in keiner Weise den Qualitätsanforderungen der Technik oder dem Stand der Wissenschaft entspricht:

- Weder sind die PCR Tests standardisiert noch gibt es belastbare Daten zur Sensitivität und Spezifität, aus denen wiederum belastbare Daten zur falsch positiv Rate abgeleitet werden können, um diese von den Rohdaten für eine wissenschaftlich fundierte statistische Analyse abziehen zu können.
- Die zentralen Kennzahlen „Fälle in 7 Tagen pro 100 000 Einwohner“ und der R-Wert (so wie er momentan ermittelt wird) sind **keine** validen Messgrößen.

- Die Tunnelblick – artige Fokussierung der Risikobewertung auf der Basis der PCR Testzahlen ist nicht nachvollziehbar, eine angemessene Berücksichtigung der viel relevanteren Messgrößen Hospitalisierte und Verstorbene ist nicht erkennbar.
- Es ist unverständlich, dass es offenbar keine systematische Weiterentwicklung der Basis für die PCR Test gibt in Richtung Validierung (steht immer noch aus) und Erfassung der denkbaren Kreuzreaktionen, etwa mit den in Deutschland endemischen Coronaviren. Auch ist eine regelmäßige Modellierung des Infektionsgeschehens nicht erkennbar.

Als Konsequenz drängt sich auf:

**Aussetzung der Bewertung der PCR Testergebnisse, bis die genannten Schwächen beseitigt sind und stattdessen bis dahin Beurteilung des Infektionsgeschehens ausschließlich auf der Basis der Zahl der Erkrankten und Verstorbenen.**

## Referenzen

Ref. 1: <https://www.nytimes.com/2020/08/29/health/coronavirus-testing.html>

Ref. 2: C.B.F. Vogels, „Analytical sensitivity and efficiency comparisons of SARS-COV-2 qRT-PCR primer-probe sets“, <https://www.medrxiv.org/content/10.1101/2020.03.30.20048108v3>

Ref. 3: AG LABORKAPAZITÄT BEIM RKI (7.7.2020), „Bericht zur Optimierung der Laborkapazitäten zum direkten und indirekten Nachweis von SARS-CoV-2 im Rahmen der Steuerung von Maßnahmen“

Ref. 4: Nationaler Pandemieplan Teil II

Ref. 5: J. Wittmann & W. Bergholz, „Introduction to Quality Management in the Semiconductor Industry – Vol. 1“ ISBN-10 1535046341

## Anhang

**Zahlenbeispiel für falsch positiv und falsch negativ Rate ausführlicher dargestellt.**

**Sensitivität** = wie empfindlich ist der Test, d.h. mit welcher Wahrscheinlichkeit weist er eine **positive** Probe wirklich als positiv nach. Nach dem im April durchgeführten Ringversuch zum Covid-19 PCR Test beträgt die Sensitivität des Covid-19 PCR Tests im Durchschnitt 98%.

**Spezifität** = selektiv ist der Test, d.h. mit welcher Wahrscheinlichkeit weist er ein **negative** Probe wirklich als negativ nach. Auf der Basis des April Ringversuchs liegt die Spezifität bei 98,6%, bei Anwesenheit von anderen Coronaviren (die es in Deutschland gibt) bei 98,0%.

Der eigentliche Messfehler besteht aus den falsch negativen Befunden, nämlich  $100\% - 98\% = 2\%$  und den falsch positiven Befunden, bei Abwesenheit von anderen Coronaviren  $100\% - 98,6\% = 1,4\%$ .

Das klingt wenig dramatisch, die Auswirkung auf die Aussage des Tests kann aber je nach Anteil der **wahren** positiven Befunde erheblich sein.

Um das zu verdeutlichen, folgendes konkretes Beispiel mit 10.000 Tests und der Annahme, dass der **wahre Anteil** positiver Testpersonen 1% beträgt, also der wahre Anteil der negativen Probanden 99%. Es gibt also 9900 wahre negative Probanden und 100 wahre positive Probanden.

In diesem Fall wirken sich die Sensitivität und die Spezifität sehr unterschiedlich auf die Korrektheit der Aussage des Tests aus:

Die falsch negativen Befunde sind 2% von den 100 wirklich vorhandenen positiven Probanden, also 2 zusätzliche falsche negative Tests, die in Wirklichkeit positiv hätten sein sollen, **es werden also 98 statt 100 positive Tests nachgewiesen**.

Die falsch positiven Befunde sind 1,4% von den wirklich vorhanden 9900 negativen Probanden, also gerundet **139 zusätzlich falsche positive Tests**, es bleiben also  $9900 - 139 = 9761$  negative Testergebnisse.

Das bedeutet: Es werden statt 9900 negative Test 9761 + 2 = 9763 negative Ergebnisse gemessen, d.h. der **Vorhersagewert** für ein negatives Testergebnis ist mit  $9763/9900 \times 100\% = 98,6\%$  Wahrscheinlichkeit richtig (= Negative Predictive Value NPV).

Es werden statt 100 positive Tests 98 + 139 = 237 positive Testergebnisse gezählt, d.h. der Vorhersagewert für ein positives Testergebnis ist nur mit  $100/237 \times 100\% = 42\%$  richtig (= Positive Predictive Value PPV).

**Fazit: bei einer Inzidenz von 1 % infizierten Personen und der Spezifität von 98,6% gibt es mehr falsch positiv Getestete als wirklich positive Personen!**

Somit ist klar:

Um bei niedriger Inzidenz das wahre Infektionsgeschehen abzubilden, **muss in einem Labor, das für diesen Test die Anforderungen von ISO 9001 oder der Norm ISO 15189 erfüllen soll, die Sensitivität und die Spezifität – ähnlich wie in dem Ringversuch – mit Kontrollproben gemonitort werden und dem Auftraggeber mitgeteilt werden**, so dass dieser für das Monitoring des Infektionsgeschehens die falsch positiv Rate abziehen kann.

Geschieht dies nicht, ergibt sich ein völlig falsches Bild vom Infektionsgeschehen, besonders wenn die Anzahl der Tests pro Tag oder Woche sich wesentlich ändert.

**Genau dies ist zurzeit der Fall, es wird der systematische Fehler der falsch positiv Rate in den veröffentlichten Zahlen nicht berücksichtigt!**

Schwanewede, den 26. Oktober 2020

gez. Prof. Dr. Werner Bergholz