

Betreff: "Unschärfe" bei der Berechnung der Impfwirksamkeit  
Datum: Sat, 25 Dec 2021 15:47:32 +0100  
Von: carloxx <carloxx@web.de>  
An: simon.hegeli@hfp.tum.de

Hallo Prof. Hegelich  
zu meiner Person:

Ich war angestellter Lehrer an einer beruflichen Schule in Kirchheim/Teck und bin seit diesem Schuljahr in Rente.

Da ich durch meinen Bekanntenkreis immer wieder mit der Corona-Problematik konfrontiert werde, versuche ich mich gerade mit diesem Thema zu beschäftigen.

Im Internet habe ich den hochinteressanten Artikel von Ihnen gefunden, auf den ich mich im Folgenden beziehen werde:

<https://politicaldatascience.blogspot.com/2021/12/deep-dive-impfeffektivitat-eine.html>

Dazu habe ich folgende Fragen und bitte Sie, mir diese zu beantworten:

1) Farrington-Formel:

Bezeichnungen:

PCV :=  $P(G | E)$  = Verhältnis Anzahl geimpfter Kranker unter den kranken Personen = Maß für Impfdurchbruch

PPV :=  $P(G) = G/N$  = Verhältnis Anzahl Geimpfter bzgl. Gesamtzahl = Maß für Impfquote

$N := U + G$

$VE = \text{Impfeffektivität} = 1 - [PCV/(1-PCV)] * [(1-PPV)/PPV]$

Mir ist leider nicht klar welche intuitive Vorstellung der „Impfeffektivität“ durch diese Farrington-Formel „abgebildet“ werden soll.

Insbesondere ist mir nicht klar, ob sie evtl. kontraintuitiv ist.

Allerdings hilft mir dieser Zusammenhang weiter (den ich im Folgenden verwende):

$v := v(PCV, PPV) := v(VE) := P(E|U) / P(E|G) =$

= Anteil der Erkrankten unter den Ungeimpften / Anteil der Erkrankten unter den Geimpften

=  $1 / (1-VE)$

wobei:

Wertebereich von VE:  $VE \leq 1$

Wertebereich von  $1/(1-VE)$ :  $1/(1-VE) \geq 0$

Neben Ihrer 3D-Funktionsplott-Grafik wäre eine weitere Grafik interessant, die die Werte  $1 / (1-VE)$  angibt. (siehe dazu mein angefügtes Excel-Diagramm).

Frage:

Was meinen Sie dazu?

2) Zitat:

„Sobald aber entweder der Anteil der Geimpften an den Erkrankten sehr niedrig ist ODER die Impfquote sehr hoch ist, kann die Funktion nur Werte produzieren, die nahe 1 sind (also für eine sehr hohe Impfeffektivität sprechen. In anderen Worten:

Die Methode wird bei einer hohen Impfquote fast immer eine hohe Effektivität ermitteln.

Diese Probleme mit der Farrington-Methode wurden von Cohen et al. 2012 aufgezeigt.“

Ich versuche den Inhalt des Zitates mit meinen Worten wiederzugeben:

Die Werte PCV und PPV werden in zeitlichen Abständen (z.B. wöchentlich) durch die

empirisch festgestellten Daten berechnet :  
Anzahl Geimpfter , Anzahl Erkrankter , Gesamtanzahl.

Daraus berechnet man dann:

PCV = Anzahl geimpfter Kranker/ Anzahl Erkrankter (\*)

PPV = Anzahl geimpfter Kranker/ Gesamtanzahl (\*\*)

Wenn diese empirischen Daten ( Anzahl Geimpfter , Anzahl Erkrankter , Gesamtanzahl ) fehlerbehaftet („unscharf“) sind (innerhalb eines bestimmten unbekanntes Intervalls liegen), dann sind PCV und PPV ebenfalls fehlerbehaftet („unscharf“), liegen also ebenfalls jeweils in einem bestimmten, unbekanntes Intervall.

Deshalb unterliegt  $v(\text{PCV}, \text{PPV})$  auch einer „Unschärfe“ und befindet sich innerhalb eines bestimmten, unbekanntes Intervalls.

Interessant ist nun, wie weit die wirklichen Werte von PCV bzw. PPV von den durch die o.g. ermittelten Berechnungen (\*) bzw. (\*\*)

entfernt sind und wie weit sich dadurch der wirkliche Wert  $v(\text{PCV}, \text{PPV})$  vom fehlerhaften unterscheidet.

Anschaulich:

Wenn man PCV und PPV auf den Koordinatenachsen der Ebene aufträgt und  $v$  auf der z-Achse, dann sind die Gebiete interessant, wo  $z$  massiv ansteigt bzw. abfällt, was bedeutet, dass kleine Fehler von PCV und PPV große Auswirkungen auf  $v$  haben.

Das ist für „kleine“ PCV und „große“ PPV der Fall:

Wird statt eines kleinen PCV-Wertes der wirkliche, noch kleinere zur Berechnung von  $v$  verwendet, so steigt  $v$  stark an.

Wird statt eines großen PPV-Wertes der wirkliche, noch größere zur Berechnung von  $v$  verwendet, so steigt  $v$  stark an.

Ebenso ist z.B. Folgendes bemerkenswert: (Parameterreihenfolge PCV,PPV)

$v(0,75 ; 0,8) = 1,33 \implies$  Ungeimpfte haben größeres Risiko als Geimpfte

$v(0,85 ; 0,8) = 0,71 \implies$  Ungeimpfte haben kleineres Risiko als Geimpfte

Frage:

Habe ich damit den Inhalt des o.g. Zitates richtig wiedergegeben?

3)

Zitat:

„Bei den Endpunkten „Hospitalisierung“ und „Tod“ sehen wir Sprünge von 30% und 20%. Offenbar sind die Fallzahlen nicht ausreichend, um eine solide Schätzung vorzunehmen.“

Frage:

Warum folgern Sie von Sprüngen von 30% und 20% auf nicht ausreichende Fallzahlen?

Sind das „Beweise“ oder nur Indizien?

4)

Die Unterteilung: Geimpfte, Ungeimpfte muss erweitert werden.

4.1)

Erweiterung in: Geimpfte, Ungeimpfte, Unbekannte

Da Unbekannte entweder Geimpfte oder Ungeimpfte sind, werden auch die Werte der Impfdurchbrüche  $\text{PCV} = P(G | E)$  bzw. die Impfquote  $\text{PPV} = P(G) = G/N$  verändert.

Damit wird auch  $v$  verändert.

Wenn man die Verteilung der Unbekannten auf die Geimpften, Ungeimpften nicht kennt, könnte man worst case / best case Szenarien berechnen:

a) Für PCV

Alle Unbekannte sind erkrankte Geimpfte versus alle Unbekannte sind erkrankte Ungeimpfte

Frage:

Wurde das schon berechnet, evtl. auch vom RKI?

b) Für PPV

Zitat:

„Das RKI bezieht sich nur auf die Fälle, bei denen der Impfstatus bekannter Maßen geimpft oder ungeimpft ist und zwar sowohl bei den Impfdurchbrüchen (also PCV in der Farrington-Formel) als auch bei der Impfquote (PPV).“

Mir ist nicht klar, warum es bei der Impfquote Unbekannte gibt.

Frage:

Werden nicht nur genau diejenigen geimpft, die ungeimpft sind ?

4.2)

Erweiterung in: Geimpfte, Ungeimpfte, Unbekannte, unvollständig Geimpfte

Die Unbekannten könnte man wieder (wie oben) als variable Größen auf die Geimpften und Ungeimpften verteilen. Allerdings ist das nicht mit den unvollständig Geimpften möglich.

Die unvollständig Geimpften könnten theoretisch PCV, also die Impfdurchbrüche beeinflussen (indem sie z.B. nur die Geimpften anstecken) und damit  $v$  beeinflussen.

Man hat also evtl. ein anderes Szenario und einen anderen Wert von  $v$ , als wenn es keine unvollständig Geimpften geben würde.

Frage:

Haben Sie das mit dem folgenden Zitat gemeint?

„Als Statistiker würde man sich fragen, ob die Gruppe der Ungeimpften zusammen mit den vollständig Geimpften ein repräsentatives Sample der Gesamtbevölkerung ist.“

5) Berechnung der Impfquote:

<https://politicaldatascience.blogspot.com/2021/12/deep-dive-impfeffektivitat-eine.html>

Zitat:

„Die Impfquote ist der Anteil aller bisher Geimpften in der Gesamtbevölkerung.

Da in Deutschland kein einheitliches umfassendes System zur Erhebung von Impfdaten besteht, werden Teilstichproben und Querschnittsuntersuchungen herangezogen, um eine Einschätzung der Impfsituation zu ermöglichen.“

Frage:

Hält sich das RKI an diesen Vorschlag und wie setzt das RKI diesen konkret um?

6)

<https://politicaldatascience.blogspot.com/2021/12/deep-dive-impfeffektivitat-eine.html>

Zitat:

„Das RKI verwendet für die Berechnung der “Impfquote” (eigentlich Verhältnis Geimpft zu Geimpft + Ungeimpft, PPV in der Farrington-Formel) den Durchschnitt der letzten vier Wochen. Auch das klingt zunächst vernünftig.

Aber: Die Impfquote erreicht eine Sättigung. D. h. in der Regel werden jede Woche weniger Impfungen hinzu kommen, aber die Zahl der Impfungen wird weiter steigen.“

Frage:

Können Sie mir bitte die Quelle nennen, wie das RKI die Impfquote berechnet?

Frage:

Ich verstehe den Sinn dieser Berechnung (des RKI) nicht:

Wenn man die aktuelle Impfquote errechnen will, muss man doch alle gemeldeten Verimpfungen pro Zeitabschnitt (z.B. pro Woche) bis zum aktuellen Datum aufaddieren.

Warum rechnet das RKI anders (Durchschnitt der letzten 4 Wochen)?

7)

Frage:

Wissen Sie , was aus den Forderungen nach tagesaktuellen und maschinenlesbare Daten geworden ist?

<https://netzpolitik.org/2020/datenjournalistinnen-fordern-offene-corona-daten/>

8)

In der folgenden Quelle wird ein Zitat des RKI angegeben:

<https://www.regensburg-digital.de/rki-und-impfdurchbrueche-zwischen-dunkelfeld-und-grober-verzerrung/08112021/>

Zitat:

„Seit dem 30.09.2021 enthält der Wochenbericht des RKI zur Beschreibung und Auswertung der Impfdurchbrüche

nur noch die symptomatischen COVID-19-Fälle, für die ausreichend Angaben zur Beurteilung des Impfstatus vorlagen.“

Frage:

Dadurch wird die Berechnung von  $v$  nur noch für die symptomatischen COVID-19-Fälle möglich.

Welchen Sinn soll das bezwecken ?

9)

Warum wird vom RKI keine Kohortenstudie durchgeführt, wie dies z.B. von dem Team um Prof. Matthias Schrappe empfohlen wird?

[https://www.schrappe.com/ms2/index\\_htm\\_files/thesenpapier\\_adhoc4\\_211128\\_endfass.pdf](https://www.schrappe.com/ms2/index_htm_files/thesenpapier_adhoc4_211128_endfass.pdf)

mfg

Karl-Heinz Barner