

# Effizientes Screening zur Ermittlung des Anteils Corona-Infizierter in einer Population

Wolfgang Tschirk, wolfgang.tschirk@mathecampus.at, 27.03.2020

Zur zuverlässigen Ermittlung des Anteils jener Menschen in einer Population, die mit dem Coronavirus (SARS-CoV-2) infiziert sind, ist eine große Anzahl von Tests nötig. Wären beispielsweise 0,3 % der Population infiziert (das entspricht 26700 Personen in Österreich) und wollte man den Anteil mit einer Schwankungsbreite von 0,1 % ermitteln, bräuchte man rund 11500 Abstriche und, sofern man jeden Abstrich als eigene Probe testet, 11500 Tests. Führt man jedoch die Abstriche von beispielsweise 64 Personen in eine Probe zusammen und testet nur diese, so braucht man bei etwas mehr Abstrichen nur 198 Tests, ohne dass die Schwankungsbreite höher und damit das Screening ungenauer würde.

## 1 Problem und herkömmliche Lösung

Die Kenntnis und Verfolgung des Anteils jener Menschen in einer Population, die mit dem Coronavirus (SARS-CoV-2) infiziert sind, wäre hilfreich beim Evaluieren bestehender und Planen neuer Methoden zur Eindämmung der Pandemie.

Wir betrachten das Problem, den Anteil  $p$  von Corona-Infizierten in einer Population auf Basis einer Stichprobe zu schätzen. Die herkömmliche Lösung wäre, von ausreichend vielen Probanden einen Abstrich zu nehmen und diesen auf Vorhandensein des Virus zu testen. Unter der Annahme einer repräsentativen Stichprobe (wie jener bei Wahlumfragen, wo die Stimmenanteile der wahlwerbenden Parteien geschätzt werden) gilt für die Anzahl  $n$  der Probanden

$$n \geq \left( \frac{1,96}{\Delta p} \right)^2 p(1-p), \quad (1)$$

wobei  $p$  der zu schätzende Anteil ist und  $\Delta p$  die zulässige Schwankungsbreite, das ist die halbe Breite des 95 %-Konfidenzintervalls [1]. Mit  $p = 0,3 \%$  und  $\Delta p = 0,1 \%$  ergibt sich aus (1):  $n \geq 11491$ . Wollte man also einen Anteil von 0,3 % (entsprechend 26700 Infizierten in Österreich) mit einer Schwankungsbreite von 0,1 % ermitteln, bräuchte man 11491 Abstriche und ebensoviele Tests. Vor allem die hohe Anzahl der Tests lässt ein solches Screening zurzeit unmöglich erscheinen.

**Tabelle 1:** Minimal nötige Anzahl  $n$  von Abstrichen und Tests zur Ermittlung des Anteils  $p$  Corona-Infizierter mit der maximal zulässigen Schwankungsbreite  $\Delta p$ , für plausible Werte von  $p$  und  $\Delta p$  (herkömmliche Methode)

Anteil Infizierter	$p$	0,10 %	0,30 %	1,00 %			
Schwankungsbreite	$\Delta p$	0,01 %	0,03 %	0,03 %	0,10 %	0,10 %	0,30 %
Anzahl Abstriche und Tests	$n$	383 776	42 642	127 670	11 491	38 032	4 226

## 2 Lösung durch Pooling der Abstriche

Dina Berenbaum vom Technion Israel Institute of Technology hat vorgeschlagen, Abstriche jeweils einer Gruppe von Personen zu einer Probe zusammenzuführen (Pooling) und diese Probe zu testen [3, 4]. Das Resultat wäre wie folgt zu interpretieren: Ist der Test negativ, dann ist niemand in der Gruppe infiziert; ist er positiv, dann ist zumindest eine Person infiziert, und durch nachfolgende Einzeltests an den Abstrichen ließe sich herausfinden, welche Person(en) dies ist(sind). Aspekte dieser Idee wurden von Rudolf Hanel und Stefan Thurner, Complexity Science Hub Vienna, untersucht [2].

Eine besonders hohe Einsparung an Tests erlaubt das Pooling im vorliegenden Problem; denn solange nur der Anteil der Infizierten interessiert, können die im Positivfall folgenden Einzeltests unterbleiben (wiewohl man sie zur Identifikation der infizierten Personen natürlich auch durchführen kann). Wir berechnen nun, wie viele Abstriche und Tests man mit Pooling braucht, um den Anteil Infizierter mit gegebener Genauigkeit zu ermitteln.

Denken wir uns die Abstriche von jeweils  $g$  Personen (diese bezeichnen wir als Gruppe) in eine Probe zusammengeführt. Nach Angabe der israelischen Forscher sind Gruppen bis zu  $g = 64$  möglich: enthält auch nur einer von 64 Abstrichen das Virus, spricht der von ihnen verwendete Test an. Die Wahrscheinlichkeit dafür, dass eine Probe mindestens *einen* Abstrich einer infizierten Person enthält, und damit der ungefähre Anteil  $q$  infizierter Proben ist

$$q = 1 - (1 - p)^g . \quad (2)$$

Daraus lässt sich  $p$  berechnen:

$$p = 1 - (1 - q)^{\frac{1}{g}} . \quad (3)$$

Man kann also, indem man den Anteil  $q$  infizierter *Proben* bestimmt, den Anteil  $p$  infizierter *Personen* berechnen. Der Vorteil des Poolings liegt darin, dass der Anteil infizierter Proben mit weit weniger Tests gefunden werden kann, als man nach (1) brauchen würde. Wir berechnen nun, wie viele Tests zum Erreichen einer geforderten Genauigkeit nötig sind. Dazu leiten wir (2) ab:

$$\frac{dq}{dp} = g(1 - p)^{g-1} . \quad (4)$$

Für kleine Differenzen  $\Delta q$  und  $\Delta p$  ergibt sich daraus

$$\Delta q = g(1 - p)^{g-1} \Delta p . \quad (5)$$

Aus einer maximal zulässigen Schwankungsbreite  $\Delta p$  von  $p$  folgt damit eine maximal zulässige Schwankungsbreite  $\Delta q$  von  $q$ ; denn wenn  $q$  um nicht mehr als  $\Delta q$  vom geschätzten Wert abweicht, dann weicht  $p$  um nicht mehr als  $\Delta p$  von dem aus (3) berechneten Wert ab.

Um nun den Anteil  $q$  infizierter Proben mit der zulässigen Schwankungsbreite  $\Delta q$  zu ermitteln, braucht man eine bestimmte Anzahl  $t$  von Proben und damit von Tests, und zwar analog zu (1)

$$t \geq \left( \frac{1,96}{\Delta q} \right)^2 q(1 - q) . \quad (6)$$

Hier setzen wir (2) und (5) ein und erhalten

$$t \geq \left( \frac{1,96}{\Delta p} \right)^2 \frac{(1 - p)^{2-g} - (1 - p)^2}{g^2} . \quad (7)$$

Damit ist die Mindestanzahl der nötigen Tests bestimmt. Die Mindestanzahl  $n$  der nötigen Abstriche ergibt sich durch Multiplikation mit der Gruppengröße:

$$n = gt . \quad (8)$$

**Tabelle 2:** Minimal nötige Anzahl  $t$  von Tests und  $n$  von Abstrichen zur Ermittlung des Anteils  $p$  Corona-Infizierter mit der maximal zulässigen Schwankungsbreite  $\Delta p$ , für plausible Werte von  $p$  und  $\Delta p$  (Pooling mit einer Gruppengröße von 64)

Anteil Infizierter	$p$	0,10 %		0,30 %		1,00 %	
		$\Delta p$	0,01 %	0,03 %	0,03 %	0,10 %	0,10 %
Anzahl Tests	$t$	6 190	688	2 197	198	830	93
Anzahl Abstriche	$n$	396 160	44 032	140 608	12 672	53 120	5 952

Wie der Vergleich der Tabellen 1 und 2 zeigt, lässt sich durch Pooling mit etwas mehr Abstrichen als im herkömmlichen Fall die Anzahl der Tests beträchtlich verkleinern (bei einer Gruppengröße von 64 um beinahe den Faktor 64), ohne dass die Genauigkeit leiden würde. Das gilt für heute angenommene Werte des Anteils infizierter Personen und der nötigen Genauigkeit beim Ermitteln dieses Anteils.

### 3 Zusammenfassung

Zur zuverlässigen Ermittlung des Anteils jener Menschen in einer Population, die mit dem Coronavirus (SARS-CoV-2) infiziert sind, ist beim herkömmlichen Vorgehen, wo jeder Abstrich als eigene Probe getestet wird, eine so große Anzahl von Tests nötig, dass ein flächendeckendes Screening zurzeit unmöglich scheint. Dina Berenbaum vom Technion Israel Institute of Technology hat vorgeschlagen, Abstriche jeweils einer Gruppe von Personen zu einer Probe zusammenzuführen (Pooling) und diese Probe zu testen. In Anwendung der Idee auf das vorliegende Problem lässt sich aus dem Anteil infizierter *Proben* der Anteil infizierter *Personen* berechnen. Um die Genauigkeit der Ermittlung zu bewahren, sind etwas mehr Abstriche erforderlich als beim herkömmlichen Screening, die Anzahl der nötigen Tests wird aber bei einer – laut Technion möglichen – Gruppengröße von 64 um beinahe den Faktor 64 kleiner. Das gilt für heute angenommene Werte des Anteils infizierter Personen und der nötigen Genauigkeit beim Ermitteln dieses Anteils.

### Literatur

- [1] Bortz, J.: Statistik für Human- und Sozialwissenschaftler. Springer, Heidelberg (2005)
- [2] Hanel, R.; Thurner, S.: Boosting test-efficiency by pooled testing strategies for SARS-CoV-2. <https://www.csh.ac.at/wp-content/uploads/2020/03/PoolTestingv5.pdf> (24.03.2020)
- [3] Jeffay, N.: To ease global virus test bottleneck, Israeli scientists suggest pooling samples. The Times of Israel, <https://www.timesofisrael.com/to-ease-global-virus-test-bottleneck-israeli-scientists-suggest-pooling-samples/> (24.03.2020)
- [4] Technion Israel Institute of Technology: Pooling Method for Accelerated Testing of COVID-19. <https://www.technion.ac.il/en/2020/03/pooling-method-for-accelerated-testing-of-covid-19/> (24.03.2020)