

(14) Die Magie der Statistik in Aktion: Tamiflu – „der größte Raub der Geschichte“ – ziemlich unbrauchbar

Ich hatte in meinem letzten Beitrag zur Methodenlehre den Zusammenhang zwischen Effektgröße, Stichprobengröße und Signifikanz durchdekliniert. Wer sich das Wichtigste gemerkt hat weiß: Auch jeder noch so kleine Effekt kann signifikant gemacht werden, wenn er vorhanden ist, vorausgesetzt man hat die entsprechenden Mittel. Was sich alle Leser immer wieder fragen sollten: Ist der Effekt sein Geld wert? Ist die Größe des Effekts für praktische Zwecke groß genug? Darüber hinaus stellt sich bei kleinen Effekten natürlich immer die Frage: Sind denn auch wirklich alle Studien bekannt und berücksichtigt? Denn klarerweise macht bei kleinen Effekten, wenn man sie nun in Meta-Analysen oder Reviews zusammenfasst, das Fehlen einiger negativen oder weniger stark positiver Studien gleich einen großen Unterschied aus.

Wie dies alles zu Buche schlägt kann man an einem aktuellen Beispiel besichtigen: dem Cochrane- Review über die Wirksamkeit von Neuraminidase-Hemmern zur Behandlung von Grippe und zur Vorbeugung gegen Grippe-Komplikationen, der von Jefferson und Kollegen soeben in neuer Version vorgelegt wurde [1]. Wir erinnern uns:

Im Jahr 2009 bricht eine neue Welle der Schweine-Grippe aus, die vom Virus A/H1N1 verursacht wird. Ich kann mich noch gut erinnern und war damals in England. Die Zeitungen berichteten fast hysterisch.

Es lohnt sich, heute, aus fünfjähriger Distanz, den Originalton noch einmal Revue passieren zu lassen.

Man liest, dass höchstwahrscheinlich eine Pandemie ausbrechen werde. In anderen Publikationen warnt die WHO hoch-offiziös vor einer Pandemie – und bekniert Regierungen, ausreichend Tamiflu einzulagern. Dieses Medikament wird in der Frage-Antwort-Publikation der englischen Tageszeitung „The Independent“, die ich damals regelmäßig verfolgte, der Impfung sogar vorgezogen. Denn die sei schwierig, nicht rechtzeitig für alle einführbar. Daher sei die logische Alternative: Tamiflu. Dieser Neuraminidasehemmer der Firma Roche, genauso wie das Konkurrenzprodukt Relenza der Firma GlaxoSmithKline (GSK), hemme angeblich die Weiterverbreitung der Viruspartikel, heißt es da. Dadurch könne es nicht zur Ansteckung kommen, außerdem könne man es gleich auch prophylaktisch geben. Wird es eine Pandemie geben? Ja, höchstwahrscheinlich. Aber, alles kein Problem: wir haben ja Tamiflu und Relenza eingelagert. Zwar leider nicht genug, aber immerhin. Damit kriegen wir die schlimmsten Probleme in den Griff. Wunderbar, denkt man. Oder soll man nicht doch lieber einfach Vitamin C nehmen? Ist doch billiger, oder? Die kluge Antwort aus dem Off: *„Alternative health practitioners of homeopathy, herbal remedies and nutritional medicine are recommending measures to protect against the flu. They are about as useful as a water pistol against a forest fire.“* Klar. Nur die Pharmakologie, „evidence based“, mit bekannten, im Experiment erhärteten Wirkprinzipien kann hier helfen.

(14) Die Magie der Statistik in Aktion: Tamiflu – „der größte Raub der Geschichte“ – ziemlich unbrauchbar

Mit der Schützenhilfe derartiger Publicity ausgerüstet, unter dem Druck der WHO und der Öffentlichkeit, machten sich unsere Regierungen daran, geschätzte 7.6 Milliarden CHF in die Vorratshaltung von Tamiflu zu investieren. Wieviel Geld bei GSK für Relenza über die Theke gegangen sind ist meines Wissens nicht bekannt. Um die Sache für die ungeduldigen Leser abzukürzen: Das Geld war zum Fenster hinaus geschmissen. Erstens wurde aus der A H1/N1 Grippe keine Epidemie, jedenfalls keine, die schlimmer gewesen wäre als andere Grippe-Wellen. Zweitens wirkt Tamiflu nur sehr begrenzt gegen Grippe, wenn man sie einmal hat und praktisch gar nicht zur Vorbeugung. Und drittens dürfte die Wirkung nicht über den behaupteten Mechanismus laufen.

Das wird durch den Review von Jefferson und Kollegen klargestellt [1]. Peter Gøtzsche, der Leiter des nordischen Cochrane Centers in Kopenhagen, sagt in diesem Zusammenhang sogar: *„Roche has committed what looks to me like the biggest theft in history, but no one has yet dragged the company to court“* [2, p. 28].

Starke Worte. Ich bin zwar nicht der Meinung, dass Gøtzsche Recht hat; vermutlich war der größte Raub der neueren Geschichte die Enteignung der Klöster in England durch Heinrich VIII., oder vielleicht der Raub der Goldschätze der indigenen Völker Amerikas, aber lassen wir diese Detailgeplänkel. Tatsache ist: in einem Spiegelgefecht scheinen uns Roche und GSK an der Nase herumgeführt zu haben, indem sie zunächst nur diejenigen Studien publizierten, die die These von der Wirksamkeit untermauerten. Schon die erste Analyse der Autorengruppe aus dem Jahr 2009 meldete Zweifel an [3]. Die Effekte waren signifikant. Symptomreduktion um etwa einen Tag bei experimentell im Labor induzierten Infektionen. Das war zwar nicht viel, aber immerhin. Schon damals war sichtbar, dass die Behauptung, man könne mit Neuraminidasehemmern Grippe ganz verhindern, falsch war. Dann tauchte plötzlich das Argument in der Debatte auf: man könne aber Komplikationen, Krankenhausaufenthalte, Lungenentzündungen und dergleichen verhindern. Das ist ja auch schon was (obwohl man mit 7 oder so Milliarden Franken ganze Armeen mit Lungenentzündung im Krankenhaus behandeln könnte).

Aber nun zerrinnt der ganze Goldanstrich in der Hand der Cochrane-Reviewer zu billigem Abklatsch. Der neue Review hat nun alle greifbaren Studien eingeschlossen, auch diejenigen, die Roche lange zurückgehalten hat und die lange Zeit nicht publiziert waren, auch solche die über die Zulassungsbehörden nicht erhältlich waren. Das ist Peter Doshi und einigen Leuten der Autorengruppe zu danken, die unermüdlich über das British Medical Journal und die Öffentlichkeit Druck ausgeübt haben, so dass Roche am Schluß alle Dokumente zur Verfügung stellte. Die Autoren des Reviews mussten sich dann, wie sie selber sagen, durch Tausende und Abertausende Seiten schlecht aufgearbeiteter Berichte wühlen. Im spröden Text der offiziellen Cochrane-Publikation geht es unter, aber wer Ohren hat zu hören, der hört sehr genau: Die Berichte waren teilweise schlecht, die Studien waren

(14) Die Magie der Statistik in Aktion: Tamiflu – „der größte Raub der Geschichte“ – ziemlich unbrauchbar

voller Fehler, Protokolle waren nachträglich verändert, Zielkriterien während laufender Studien verändert, Auswertungen angepasst worden, Ergebnisvariablen umdefiniert worden. Das ganze Pandämonium methodischer Fehler begegnet einem auf den ersten 20 Seiten der Beschreibung der eingeschlossenen Studien.

Das ging soweit, dass sogar eine ganze Reihe der freigegebenen Studien gar nicht verwendbar waren, weil die Datenauswertung zu fragmentarisch war. Und am Ende zeigt sich ein geradezu peinliches, wenn nicht tragisches Ergebnis: Zwar haben, über alle Studien gesehen, Neuraminidasehemmer, Tamiflu und Relenza, einen statistisch signifikanten Effekt. Aber was heißt das konkret? Das Ergebnis kommt zustande durch die Zusammenfassung von Daten aus 46 Studien, davon 20 mit Oseltamivir (Tamiflu; 9623 Patienten) und 26 mit Zanamivir (Relenza; 14.628 Patienten). Die Verkürzung der Zeit mit Symptomen betrug demnach 16,8 Stunden. Also: statt 7 Tagen ist man nur 6,3 Tage krank. Bei Zanamivir war es so ähnlich: Hier wurde die Zeit mit Symptomen um 0,6 Tage verkürzt, machte eine Reduktion der mittleren Symptombdauer von 6,6 auf 6 Tage. Krankenausesweisungen und Prophylaxe, mit denen so geworben wurde? Kein Effekt. Komplikationen? Keine Effekte. Lungenentzündung? Nur insofern, als man selbstberichtete Symptome auswertete, nicht, wenn man radiologisch, also hart untersuchte Lungenentzündung wertete. Dabei war der Effekt klein: 100 Patienten muss man behandeln, um eine Lungenentzündung zu verhüten. Nur Zanamivir reduziert das Risiko einer Bronchitis. 56 Menschen müssen behandelt werden, bis einer diesen Vorteil erleben kann. Oseltamivir zeigt diesen Vorteil nicht, und keine der beiden Präparate verbessert die Gefahr an Otitis media oder Sinusitis zu erkranken, weder bei Kindern noch bei Erwachsenen.

Andersherum die Nebenwirkungen: Tamiflu erhöht die Gefahr von Übelkeit (einer von 28 Behandelten erlebt das) und von Erbrechen (einer von 22 Behandelten). Prophylaxe-Effekte sind vorhanden aber klein. Die Autoren folgern:

„Oseltamivir and zanamivir have small, non-specific effects on reducing the time to alleviation of influenza symptoms in adults, but not in asthmatic children. Using either drug as prophylaxis reduces the risk of developing symptomatic influenza. Treatment trials with oseltamivir or zanamivir do not settle the question of whether the complications of influenza (such as pneumonia) are reduced, because of a lack of diagnostic definitions. The use of oseltamivir increases the risk of adverse effects, such as nausea, vomiting, psychiatric effects and renal events in adults and vomiting in children. The lower bioavailability may explain the lower toxicity of zanamivir compared to oseltamivir. The balance between benefits and harms should be considered when making decisions about use of both NIs for either the prophylaxis or treatment of influenza. The influenza virus-specific mechanism of action proposed by the producers does not fit the clinical evidence.“ [1, p.3]

Auf gut deutsch: Es funktioniert zwar so lala, aber auf jeden Fall nicht so, wie man aufgrund

(14) Die Magie der Statistik in Aktion: Tamiflu – „der größte Raub der Geschichte“ – ziemlich unbrauchbar

der Theorie denkt – und schon gar nicht so stark, wie behauptet. Und wichtige Fragen, etwa ob Lungenentzündung als Komplikation verhindert werden kann, sind immer noch offen. Angesichts der Gefahr, etwa auch der möglichen Probleme psychiatrischer Erkrankungen oder Nierenerkrankungen, und der geringen Effektgrößen, sollte man sich gut überlegen, ob öffentliche Geldgeber diese Substanzen finanzieren sollten. Wie war das gleich noch mit Vitamin C und den Empfehlungen der komplementärmedizinischen Praktiker?

Wasserpistolen gegen Waldbrände? Das scheint ja auf Tamiflu und Co. in der Tat zuzutreffen und die Autoren des Reviews denken denn auch laut darüber nach, ob nicht alte Entzündungshemmer wie das gute alte Aspirin nicht am Ende besser wären.

Die hier analysierten Studien sind übrigens nur ein Teil der vorhandenen. Weil bei vielen Studien die Daten nicht ausreichend detailliert berichtet wurden, konnten sie gar nicht ausgewertet werden. Insgesamt lagen 121 Studien zu Oseltamivir vor, von denen 83 von Roche zur Verfügung gestellt wurden, so dass die 20 Oseltamivir-Studien nur einen Bruchteil der Daten darstellen. Wieviel all das wohl gekostet hat? Darüber geben die Berichte keine Auskunft. Wenn wir pro Studie im Durchschnitt 2 Millionen Euro rechnen, was vermutlich zu billig ist, dann landen wir nur für Tamiflu bei mehr als 240 Millionen Euro, schätzungsweise mehr. Aber immer noch weit unter den mehr als 7 Milliarden Franken, die Roche mit Tamiflu eingenommen hat.

Was lernen wir aus diesem Debakel, denn anders kann man es nicht bezeichnen?

1. Effektgrößen sind mit die wichtigsten Größen, die man sich anschauen muss, nicht die Signifikanz.
2. Wenn es um Arzneimittelstudien und andere Produkte geht, bei denen Geld oder andere Interessen im Spiel sind, kann man davon ausgehen, dass wichtige Informationen von Interessengruppen zurückgehalten werden. Es lohnt sich, skeptisch zu sein und konservativ bis zum Beweis des Gegenteils.
(2a: Ich habe es mir angewöhnt, vor allem dann, wenn in den Mainstream-Medien mit großem Hallo und viel Hype die neue Lösung, oder die Ankündigung eines neuen Problems bejubelt oder verkündet wird, zunächst mal probatorisch vom Gegenteil auszugehen, bis ich mich durch Daten von der Richtigkeit der Behauptung überzeugen konnte.)
3. Jeder Effekt muss in Relation zu den Kosten gesehen werden. Zu den Kosten gehören nicht nur die monetären, sondern auch Nebenwirkungen und andere Probleme. Man beachte, dass im hier vorliegenden Beispiel die Effektgrößen der Nebenwirkungen wesentlich größer sind als die Effektgrößen der Vorteile.
4. Mit ausreichend viel Geld kann man jeden Effekt „wissenschaftlich“ belegen. Diese wissenschaftliche Prostitution wird erst aufhören, wenn wir von der Signifikanz weg und auf die Effektgrößen schauen.

(14) Die Magie der Statistik in Aktion: Tamiflu – „der größte Raub der Geschichte“ – ziemlich unbrauchbar

Wissenschaftlich scheinen wir mit diesem Beispiel an die Grenze des bestehenden Systems der Zulassungsforschung zu stoßen. Die Autoren des Reviews sagen immer wieder, dass das derzeit gültige System der wissenschaftlichen Bewertung unbrauchbar ist. Es erlaubt es, dass man nur diejenigen Daten publiziert, die einem in den Kram passen und den Rest versteckt. Auch Studienregister, die es ja mittlerweile gibt, helfen da nur begrenzt. Denn die Daten sind Eigentum der Firma, die sie erzeugt hat. Und der ethische Standard, der einen Wissenschaftler zur Publikation seiner Daten verpflichtet, ist nicht einklagbar. Allenfalls öffentlicher Druck, wie hier, kann dazu führen, dass eine Firma unpublizierte Daten herausgibt. Erst wenn es ein selbstverständlicher Standard werden würde, dass alle Daten öffentlich, über Plattformen, zur Verfügung gestellt werden, würden sich Firmen oder Einzelne, die dies nicht tun, unglaublich machen. Aber wer könnte Hüter und Pfleger der Daten sein? Wer würde die Plattformen finanzieren? Wer den Standard einklagen? Wenn man ausführlich über alle möglichen Lösungen nachdenkt, dann kommt man, scheint mir, rasch immer wieder an das Paradox, dass man ethisch-moralisches Verhalten nicht durch Regeln erzwingen kann.

Literatur

1. Jefferson, T., Jones, M. A., Doshi, P., Del Mar, C. B., Hama, R., Thompson, M. J., et al. (2014). Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews, CD008965(April).
2. Gøtzsche, P. C. (2013). Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Health Care. London: Radcliff.
3. Jefferson, T., Jones, M., Doshi, P., & Del Mar, C. (2009). Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis. British Medical Journal, 339, b5106.

