



# Zufall oder Signal? Statistische Bewertung von zeitlichen Zusammenhängen

André Michael Toschke, Rüdiger von Kries

Abteilung Epidemiologie, Institut für Soziale Pädiatrie und Jugendmedizin  
Ludwig-Maximilians-Universität München

## Zusammenfassung

Bei zeitlichem Zusammentreffen zweier Ereignisse – wie z. B. einer Impfung und einer fraglichen Komplikation – muß beurteilt werden, ob eine zufällige Koinzidenz oder ein statistisch signifikanter Zusammenhang vorliegt. Ein statistisch signifikanter Zusammenhang besagt jedoch nicht, daß auch ein kausaler Zusammenhang bestehen muß. Statistische Signifikanz besagt ausschließlich, daß die beobachtete Assoziation nur mit geringer Wahrscheinlichkeit rein durch den Zufall bedingt ist.

Da mögliche Komplikationen nach Impfungen nur sehr selten auftreten können, sind die Fallzahlen aus Zulassungs-

studien meist zu gering, um sehr seltene Komplikationen zu erkennen. Mögliche statistische Zusammenhänge können dann oft nur in Fall-Kontrollstudien oder durch den Vergleich mit der Basisinzidenz der möglichen Komplikation anhand von standardisierten Mortalitätsraten (SMR) erkannt werden. Ein Beispiel zur Berechnung und statistischen Bewertung einer SMR wird gegeben.

Die Wahrscheinlichkeit einer möglichen Kausalität bei einem statistisch signifikanten Zusammenhang muß anhand weiterer Kriterien zusätzlich beurteilt werden. Für Beobachtungsstudien sind die Kausalitätskriterien nach Bradford Hill hilfreich.

## Einleitung

Zeitliche Zusammenhänge zwischen Erkrankungen, die eventuell als Folgeerkrankungen gewertet werden könnten und medizinischen Maßnahmen bzw. Eingriffen wie zum Beispiel einer Impfung oder Operation müssen immer unter Berücksichtigung der jeweiligen Basisinzidenz dieser Erkrankungen beurteilt werden. Allein die korrekte zeitliche Sequenz (zunächst die Impfung, dann das Auftreten einer Erkrankung) ist noch lange kein Beweis für einen kausalen Zusammenhang und Klassifizierung der Erkrankung als Folgeerkrankung.

## Kausalität aus randomisierten Studien

Der Nachweis einer Kausalität zwischen einer medizinischen Maßnahme und späterer Komplikation läßt sich theoretisch mit dem Gold Standard, der ran-

domisierten Studie, führen. Allerdings kommen bei seltenen Nebenwirkungen hier zwei Schwierigkeiten zusammen: Zur Erkennung sehr seltener Nebenwirkungen müßten in entsprechenden Studien enorm hohe Fallzahlen angestrebt werden. Dies ist weder logistisch möglich noch finanzierbar. Darüber hinaus ist eine randomisierte Studie mit dem Ziel der Detektion möglicher schädlicher Wirkungen ethisch nicht zu vertreten. So würde z. B. eine randomisierte Studie zu Expositionen wie dem Rauchen oder Passiv-Rauchen vernünftigerweise aus ethischen Gründen nie realisiert werden.

tigerweise aus ethischen Gründen nie realisiert werden.

**Randomisierte Studien sind bei sehr seltenen Ereignissen nicht realisierbar.**

Daß die Forderung nach randomisierten Studien für den Nachweis einer Kausalität auch bei anderen Expositionen nicht immer aufrechtzuerhalten ist, wurde einprägsam in einem ironischen Artikel in der Weihnachtsausgabe des British Medical Journal von 2003 auf den Punkt gebracht (1). Dort wurde eine Meta-Analyse zum Nutzen von Fallschirmbenutzung bei Sprüngen aus mehr als 100 m Höhe beschrieben, zu der keine randomisierten Studien gefunden werden konnten. Die beiden Autoren sind zu dem Schluß gekommen, daß daher der Nutzen von Fallschirmen in Frage zu stellen sei und daß randomisierte Studien, in denen manche Probanden mit



*„Ob mögliche seltene Komplikationen im zeitlichen Zusammenhang mit einer Impfung Zufall sind oder ein Signal darstellen, ist bei ausreichend großer Fallzahl und entsprechender Teststärke (Power) durchaus beurteilbar.“*



Zufällige Koinzidenz oder statistischer Zusammenhang – für diese Entscheidung reicht eine Einzelbeobachtung nicht.

und andere ohne Fallschirm aus dem Flugzeug springen, zu diesem Thema dringend durchgeführt werden sollten, um den Grad der Evidenz für den Nutzen von Fallschirmen zu belegen (1). Bei möglichen Impfnebenwirkungen sind solche unmittelbar evidenten biologischen Zusammenhänge jedoch nie gegeben – gäbe es diese, würde keine Zulassungsbehörde solche Impfstoffe zulassen, es sei denn, es handelte sich um relativ banale Nebenwirkungen, die durch den erreichten Impfschutz weit aufgewogen würden, wie z. B. Impfmüdigkeit nach der Masernimpfung.

Wenn aus unterschiedlichen Gründen in randomisierten Studien der Nachweis von Kausalzusammenhängen von Impfung und Nebenwirkung nicht erbracht werden kann, bedeutet das dann gleichzeitig, daß keine Evidenz für einen möglichen Schaden einer Maßnahme/Exposition gewonnen werden kann? Diese Frage ist klar mit Nein zu beantworten.

### Evidenz aus Beobachtungsstudien

Wenn randomisierte Studien nicht möglich sind, kann Evidenz auch aus

Beobachtungsstudien gewonnen werden. Allerdings bleibt bei seltenen Ereignissen in prospektiven Beobachtungsstudien (z. B. Kohortenstudien) das Fallzahlproblem und die damit verbundene Frage der Realisierbarkeit bestehen. Daher bleibt oft nur die Möglichkeit einer Fall-Kontroll-Studie oder der Vergleich mit der eigentlichen Basisinzidenz des jeweiligen Ereignisses übrig, um zu entscheiden, ob es sich um ein Signal (statistisch signifikanter Zusammenhang) handelt oder ob man nur weißes Rauschen (Zufall) beobachtet.

**Ein Signal bedeutet einen statistisch signifikanten Zusammenhang. Der p-Wert macht eine Aussage darüber, mit welcher Wahrscheinlichkeit das Ereignis rein zufällig beobachtet wurde. Weißes Rauschen beschreibt den Zufall (kein Zusammenhang).**

Der „magische“ p-Wert, der für solche Studien berichtet wird, macht nur eine Aussage darüber, mit welcher Wahrscheinlichkeit ein solches Ereignis rein zufällig in zeitlichem Zusammenhang mit der Impfung aufgetreten ist:  $p < 0,05$  heißt, mit einer Wahrscheinlichkeit von knapp unter 1/20 könnte das Ereignis durchaus auch zufällig in zeitlichem Zusammenhang mit der Impfung aufgetreten sein. Bei niedrigeren p-Werten ist diese Wahrscheinlichkeit entsprechend geringer.

Neben der Beurteilung eines statistischen Zusammenhangs ist die rationale Beurteilung einer möglichen Kausalität genauso wichtig, da das Vorliegen eines statistischen Zusammenhangs noch lange nicht bedeutet, daß dieser auch kausal ist (2). Zur Beurteilung der Kausalität führte Sir Austin Bradford Hill bereits im Jahre 1965 Kriterien ein (3). Je mehr dieser 9 Kriterien (Tabelle 1) vorhanden sind, desto wahrscheinlicher wird ein kausaler Zusammenhang. Ein wesentlicher Bestandteil ist die zeitliche Sequenz.

### Zeitliche Sequenz

Eine korrekte zeitliche Abfolge ist gegeben durch die vorherige Exposition (Impfung) und ein nachfolgendes Ereignis. Diese reicht aber allein für eine mögliche Kausalität nicht aus. Vielmehr stellt eine korrekte zeitliche Sequenz eine notwendige, aber nicht hinreichende Bedingung dar. Ohne notwendige Bedingung kann eine mögliche kausale Folge nicht eintreten, wobei eine hinreichende Bedingung sicher die kausale Folge nach sich zieht. Bei den meisten medizinischen Zusammenhängen gibt es jedoch keine „hinreichenden Bedingungen“.

**Die korrekte zeitliche Abfolge zwischen Exposition und Erkrankung ist eine notwendige, aber nicht hinreichende Bedingung für einen kausalen Zusammenhang.**

### Koinzidenz

Auch wenn die zeitliche Sequenz stimmt, kann es sich noch immer um eine reine Koinzidenz handeln. Eine Koinzidenz von Exposition und Erkrankungen – bzw. Impfung und Nebenwirkung – liegt vor, wenn diese trotz ihrer Unabhängigkeit im selben Zeitraum auftreten. Die Entscheidung, ob es sich lediglich um eine zufällige Koinzidenz oder einen statistischen Zusammenhang handelt, kann bei einer Einzelbeobachtung nicht getroffen werden. Wird z. B. bei einem Kind innerhalb von 5 Tagen nach einer Impfung eine ITP diagnostiziert, besagt dies zunächst gar nichts. Tritt dies jedoch bei mehreren Kindern auf, könnte eventuell ein statistisch signifikanter Zusammenhang zugrunde liegen.

Um eine solche Entscheidung zu treffen und einen möglichen Zusammenhang auf Signifikanz zu prüfen, bedarf es genügender Teststärke (Power), die sich unter anderem durch eine ausreichend hohe Fallzahl erzielen läßt. Hohe Fallzahlen bei seltenen möglichen Komplikationen nach Impfungen (0,1 bis 0,01 %) können nur durch Beobach-



tung einer größeren Population oder eines genügend langen Zeitraums erreicht werden.

### Statistischer Zusammenhang: Beziehung erwartete versus beobachtete Fälle – SMR

Bei ausreichend großen Datenmengen, zum Beispiel aus nationalen Quellen, kann man die im zeitlichen Rahmen nach Impfung beobachteten Fälle in Beziehung zu den erwarteten Fällen der Erkrankung stellen. Übersteigt die Zahl der beobachteten Fälle die erwarteten Fälle signifikant, geht man von einem statistischen Zusammenhang (Signal) aus.

Die erwarteten Fälle ergeben sich aus der allgemeinen Basisinzidenz und werden für das jeweilig interessierende Intervall (z. B. bis 4 Wochen nach Impfung) berechnet. Dabei geht man zunächst von einer Unabhängigkeit (Koinzidenz) der Ereignisse als Nullhypothese (H<sub>0</sub>) aus und prüft, ob diese signifikant verworfen werden kann. Dieses läßt sich zum Beispiel mit standardisierten Mortalitätsraten (SMR) durchführen. Eine SMR ist gegeben durch

$$SMR = \frac{N_{obs}}{N_{exp}}$$

wobei N<sub>exp</sub> die erwarteten Fälle sind und N<sub>obs</sub> die beobachteten Fälle im interessierenden Zeitintervall.

Die wesentliche Herausforderung bei der Bestimmung der SMR liegt in der Bestimmung des Nenners (erwartete Fälle = N<sub>exp</sub>). Für die Bestimmung der erwarteten Fälle muß zunächst die Population unter Risiko bestimmt werden.

### Population unter Risiko/erwartete Fälle

Die Population unter Risiko ist gegeben durch alle Personen oder Kinder in der entsprechenden Altersgruppe, die im Studienzeitraum eine Impfdosis er-

Tabelle 1: Kausalitätskriterien nach Sir Austin Bradford Hill (3).

Stärke des Effektes	Ist das Risiko so hoch, daß andere Faktoren einfach ausgeschlossen werden können?
Konsistenz	Konnten die Ergebnisse von anderen Untersuchern unter anderen Bedingungen repliziert werden?
Spezifität	Ist die Exposition mit einer spezifischen Erkrankung assoziiert?
zeitliche Sequenz	Abfolge – 1. Exposition, 2. Erkrankung
Dosis-Wirkung	je höher die Expositions-dosis, desto höher die Erkrankungswahrscheinlichkeit
Plausibilität	Gibt es einen biologischen Erklärungsmechanismus?
Kohärenz	Steht der angenommene Zusammenhang unter Umständen völlig in Widerspruch zum aktuellen Wissensstand zu Ätiologie und Pathogenese der Erkrankung?
experimentelle Evidenz	Erzielt eine Intervention Ergebnisse, die mit der Assoziation konsistent sind (z. B. Rückenschlafkampagne und Reduktion der Inzidenz von plötzlichem Kindstod)?
Analogie	Gibt es ähnliche Effekte (z. B. Exposition in der Schwangerschaft und spätere Erkrankung)?

halten haben. Für diese kann dann mit der bekannten Basisinzidenz des Zielerignisses die Zahl der erwarteten Fälle für den gesamten Studienzeitraum berechnet werden. Geht man nun davon aus, daß das Zielereignis gleichverteilt auftritt, das heißt, daß genauso viele Fälle im Januar wie im Februar oder jedem anderen Monat auftreten, muß man nur noch die erwarteten Fälle für den gesamten Zeitraum mit dem Anteil des interessierenden Intervalls am Gesamtzeitraum gewichten. Wenn man z. B. davon ausgeht, daß 4 Wochen nach Impfung ein Ereignis eintreten kann und danach nicht mehr, und der Studienzeitraum 2½ Jahre (130 Wochen) beträgt, entsprechen die erwarteten Fälle nach Impfung 4 Wochen/130 Wochen ≈ 3 % der Fälle vom gesamten Zeitraum. Falls das Risiko, am Zielereignis zu erkranken, nicht gleichverteilt ist, muß entsprechend adjustiert werden.

### Interpretation der SMR

Bei einer SMR von 1 werden genauso viele Fälle beobachtet wie erwartet, während eine SMR < 1 einen protektiven Effekt signalisiert und eine SMR > 1

ein Risiko. Eine SMR von 2 würde eine Verdopplung der beobachteten Fälle anzeigen.

**Eine signifikante SMR < 1 bedeutet einen schützenden Zusammenhang; eine signifikante SMR > 1 kennzeichnet einen potentiellen Risikofaktor.**

### Beispiel

Es handelt sich um ein hypothetisches Beispiel: In einem Land mit 80 Millionen Einwohnern wurden 7 Fälle von Typ-1-Diabetes innerhalb von 4 Wochen nach Masern-Mumps-Röteln- (MMR-) Impfung von 1998 bis 2003 beobachtet. Daher haben wir den Verdacht, daß Typ-1-Diabetes innerhalb von 4 Wochen nach der MMR-Impfung auftreten kann. Wir setzen einen Beobachtungszeitraum von 5 Jahren voraus. Um die Berechnungen nicht zu kompliziert zu machen, gehen wir von konstant großen Geburtsjahrgängen von 700.000 Kindern aus. Da wir 5 Jahre beobachten, handelt es sich um 5 x 700.000 = 3,5 Millionen Kinder. Von diesen 3,5 Millionen Kindern werden im Durchschnitt aber nur 80 % mit MMR geimpft, so daß die Population unter Risiko 80 %

x 3.500.000 = 2.800.000 Kinder ist. Die Basisinzidenz von Typ-1-Diabetes wird mit  $10,6 \times 10^{-5}$  für Kinder unter 5 Jahren im Jahr 2000 für Deutschland angegeben (ESPED-Jahresbericht 2000). Also würden wir im gesamten Beobachtungszeitraum genau  $10,6 \times 10^{-5} \times 2.800.000 = 298$  Fälle erwarten. Innerhalb eines Zeitraums von 4 Wochen, der  $1/60$  des gesamten Studienzeitraumes beträgt, würden wir dementsprechend  $298/60 = 5$  Fälle erwarten. Nun haben wir in diesem hypothetischen Beispiel in den 5 Jahren 7 Fälle beobachtet, so daß die SMR nun  $7/5 = 1,4$  beträgt und etwas über 1 liegt.

#### Konfidenzintervall für die SMR

Die wesentliche Frage ist nun aber, ob dies reiner Zufall ist oder nicht. Dies wird anhand der Signifikanz der SMR beurteilt. Bei der Beurteilung etwaiger signifikanter Zusammenhänge wird neben dem p-Wert auch das 95 %-Konfidenzintervall verwendet. Das 95 %-Konfidenzintervall entspricht einem p-Wert  $< 0,05$ . Ist die untere Grenze des 95 %-Konfidenzintervalls größer als 1, kann von einem, mit einem p-Wert  $< 0,05$  signifikanten Zusammenhang ausgegangen werden. Eine SMR von 2,8 mit einem Konfidenzintervall von 1,2 bis 4,8 würde bedeuten, daß die wahre SMR mit 95 % Wahrscheinlichkeit mindestens bei 1,2 und maximal

#### Wesentliches für die Praxis...

- Randomisierte Studien und Kohortenstudien sind zur Erfassung zwar möglicher, aber seltener Impfnebenwirkungen auf Grund nötiger hoher Fallzahlen oft nicht realisierbar.
- Neben Fall-Kontrollstudien bietet der Vergleich mit der Basisinzidenz des Ereignisses in Form von standardisierten Mortalitätsraten (SMR) eine Möglichkeit der statistischen Beurteilung.
- Konfidenzintervalle einer SMR sind Indikatoren einer möglichen statistischen Signifikanz.
- Neben der statistischen Signifikanz müssen weitere Kriterien zur Beurteilung einer möglichen Kausalität analysiert werden.

bei 4,8 liegt. Die entsprechenden Konfidenzintervalle sind nachzuschlagen, zu berechnen oder im Internet zu recherchieren. Eine nützliche Internetquelle (<http://www.sph.emory.edu/~cdckms/exact-midP-SMR.html>) erlaubt es, die Anzahl der beobachteten und der erwarteten Fälle einzugeben, woraufhin eine SMR mit dazugehörigem exaktem Konfidenzintervall ausgegeben wird. Das 95 %-Konfidenzintervall für unseren hypothetischen Fall (SMR=1,4) ist mit 0,6-2,8 gegeben. Es handelte sich also nicht um einen statistisch signifikanten Zusammenhang des Auftretens von Diabetes nach der Impfung. Vielmehr muß aufgrund der hypothetischen Datenlage von einem zufälligen Zusammenhang, einer Koinzidenz, ausgegangen werden. Die Beispielrechnung ist in Tabelle 2 zusammengefaßt.

#### Fazit

Ob mögliche seltene Komplikationen im zeitlichen Zusammenhang mit einer Impfung Zufall sind oder ein Signal, ist bei ausreichend großer Fallzahl und entsprechender Teststärke (Power) durchaus beurteilbar. Dies ist in Fall-Kontroll-Studien oder Vergleichen mit der Basisinzidenz durch Bestimmung der SMR möglich. Dabei wird geprüft, ob z. B. die Zahl der Fälle nach Impfung die Basisinzidenz signifikant überschreitet oder ob es sich um eine reine Koinzidenz zweier unabhängiger Ereignisse handelt. Ein statistisch signifikanter Zusammenhang ist nicht mit Kausalität gleichzusetzen. Die Wahrscheinlichkeit einer möglichen Kausalität für einen signifikanten Zusammenhang muß zusätzlich beurteilt werden. Für solche Beobachtungsstudien sind die Kausalitätskriterien nach Bradford Hill hilfreich.

Tab. 2: Hypothetische Beispielrechnung zur Bestimmung einer SMR bei Typ-1-Diabetes nach MMR Impfung.

Beobachtungszeitraum:	5 Jahre
Geburtsjahrgangstärke:	700.000 Kinder
Kinder unter Beobachtung im gesamten Zeitraum:	$5 \times 700.000 = 3.500.000$
MMR-Impfrate:	80 %
Kinder unter Risiko:	$80 \% \times 3.500.000 = 2.800.000$
Basisinzidenz Typ-1-Diabetes (<5 Jahre):	$10,6 \times 10^{-5}$
erwartete Fälle im gesamten Zeitraum:	$10,6 \times 10^{-5} \times 2.800.000 = 298$
erwartete Fälle im 4- Wochen-Intervall:	$1/60 \times 298 = 5$
Beobachtete Fälle:	7
SMR:	$7/5 = 1,4$ (95%KI 0,6-2,8)

#### Literatur

1. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *Bmj*. 2003;327(7429):1459-1461.
2. Phillips CV, Goodman KJ. The missed lessons of Sir Austin Bradford Hill. *BioMed Central*.
3. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med*. 1965;58:295-300.

#### Korrespondenzadresse:

Dr. André Michael Toschke, MPH, MSc  
 Institut f. Soziale Pädiatrie u. Jugend-  
 medizin, Abt. Epidemiologie  
 Heiglhofstr. 63, 81377 München  
 Tel./Fax: 089/71009-307/315  
 E-Mail: [toschke@biostats.info](mailto:toschke@biostats.info)