

Confounding – und wie man damit umgeht

Ätiologische Studien untersuchen das Verhältnis zwischen einer Exposition (Risikofaktor) und dem Risiko einer Erkrankung (Outcome). Ein Confounder ist ein dritter Faktor, der zwar das Risiko beeinflusst, aber nicht Teil der Kausalkette zwischen Exposition und Risiko ist.

Roland Müller-Waldeck

■ Ein Confounder hat drei Eigenschaften: Er beeinflusst das Risiko für die Krankheit, ist also ein Risikofaktor. Er ist zwischen Expositions- und Kontrollgruppe ungleich verteilt, sonst würde sich seine Wirkung aufheben. Weil der Confounder kein Effekt der Exposition ist, ist er nicht Teil der Kausalkette der Erkrankung (siehe Abb.). Mögliche Confounder lassen sich nur mit viel Aufwand und beträchtlichem Wissen um die pathophysiologischen und möglichen kausalen Zusammenhänge identifizieren. Grundsätzlich kann Confounding mit dem Studiendesign oder mit statistischen Methoden begegnet werden:

Confounding mit dem Studiendesign verhindern

Randomisierung: Durch die zufällige Verteilung der Teilnehmer wird auch der Einfluss von Confoundern zufällig zwischen den Studiengruppen verteilt. Auch wenn es unwahrscheinlich ist, dass alle potenziellen Confounder gleichmäßig zwischen den Studiengruppen verteilt sind, reduziert sich das Risiko so weit wie möglich. Je kleiner allerdings die Studienpopulation ist, desto weniger lässt sich das Risiko reduzieren, denn bei beispielsweise 300.000 Studienteilnehmern werden 2 Studiengruppen homogener sein als bei 30 Teilnehmern.

Ausschluss: Wenn Rauchen ein möglicher Confounder ist, können Raucher aus einer Studie ausgeschlossen werden. Das allerdings erschwert die Rekrutierung und kann wegen geringerer Teilnehmerzahl zu verringerter Power führen und beeinträchtigt die externe Validität, weil die Ergebnisse nicht mehr auf Raucher übertragen werden können.

Matching: In diesem Fall wird für jeden Raucher in der Expositionsgruppe auch ein Raucher in die Kontrollgruppe aufgenommen. So heben sich die Effekte des Confounders statistisch auf. Matching hat zwei Nachteile: Sollen mehrere Confounder ausgeglichen werden, kann die Rekrutierung sehr problematisch werden und der Effekt der gematchten Größe(n) kann nicht untersucht werden. Diese Methode wird häufig in Fall-Kontrollstudien genutzt, selten in Kohortenstudien.

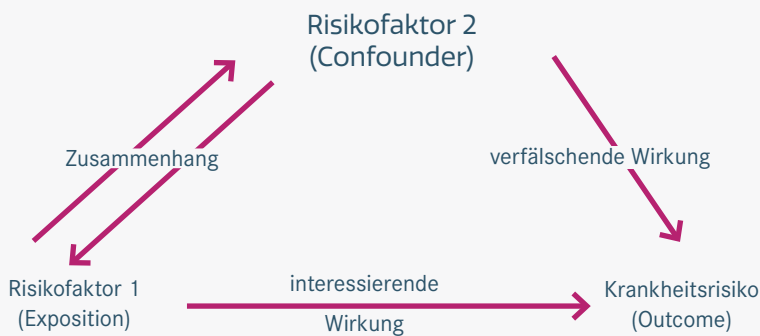
Confounding statistisch korrigieren

Stratifikation: Nach der Datenerhebung kann Confounding mit Hilfe von Stratifikation untersucht und korrigiert werden. Dafür wird die Studienpopulation in mehrere Subgruppen (Strata) geteilt, innerhalb derer der Confounder nicht oder nur wenig variiert. Anschließend wird das relative Risiko für das

Outcome in jedem Stratum und das relative Risiko für die gesamte Studienpopulation berechnet. Wenn das relative Risiko in den verschiedenen Strata von dem der gesamten Studienpopulation um mehr als 10% abweicht, liegt Confounding vor. Wenn die Risiken auch zwischen den Strata variieren, spielt eine dritte Variable (neben Outcome und Stratifikationskriterium) eine Rolle, die innerhalb der Strata unregelmäßig verteilt ist. Diese Effektmodifikation muss besonders ausgewertet werden.

Mantel-Haenszel-Schätzer: Dabei handelt es sich um einen gewichteten Mittelwert aus den Risiken der einzelnen Strata. Das Gewicht eines einzelnen Werts hängt von der Größe des Stratums ab. Die Methode erfordert einen stetigen Confounder, der in eine begrenzte Anzahl von Strata geteilt wird. Besonders bei wenigen Strata kann Restconfounding in den einzelnen Strata auftreten. Um dem zu begegnen, kann die Anzahl der Strata erhöht werden (dabei muss immer die Größe der Strata ausreichend bleiben) oder es kann eine multivariate Analyse durchgeführt werden, die den Confounder als stetige Variable behandelt.

Diese Methode hat zwei wesentliche Einschränkungen: Bei mehr als einem Confounder steigt der Aufwand und die Stu-



Definition: In einer ätiologischen Studie wird die Gesetzmäßigkeit untersucht, mit der ein Risikofaktor zu einer Erkrankung führt, indem Exposition und Erkrankungsrisiko gemessen und zueinander ins Verhältnis gesetzt werden. Hängt mit dem untersuchten ersten Risikofaktor ein zweiter Risikofaktor zusammen, ohne dass der zweite Risikofaktor ein Effekt des ersten ist, ist der zweite Risikofaktor ein Confounder.

Beispiel: In frühen Studien über orale Kontrazeptiva war das Risiko für Herzinfarkt erhöht. Später zeigte sich, dass Frauen, die orale Kontrazeptiva einnahmen, häufiger rauchten als die der Kontrollgruppe und daher ein erhöhtes Risiko für Herzinfarkt hatten: Rauchen war in dem Fall ein Confounder.

dienpopulation muss relativ groß sein. Bestehen einzelne Strata aus zu wenigen Fällen und Kontrollen, muss ein Regressionsmodell eingesetzt werden.

Standardisierung ist eine zweite Methode, Confounding nach der Datenerhebung auszugleichen.

Beispiel: In einer Studie erleiden in der Expositionsgruppe (n=80.000) 850 Teilnehmer den Endpunkt Tod, in der Kontrollgruppe (n=80.000) sterben 630. Das nichtadjustierte (crude) relative Risiko (RR) für den Endpunkt beträgt 1,35 – die Exposition scheint das Risiko um 35% zu erhöhen. Stratifiziert man beide Gruppen nach dem Alter, wird erkennbar, dass sich Expositions- und Kontrollgruppe in ihrer Altersstruktur unterscheiden. Diese unterschiedliche Altersstruktur beeinflusst das Risiko für den Endpunkt in Kontroll- und Expositionsgruppe unterschiedlich stark, das ist der verzerrende Effekt des Alters auf die Mortalität. Wäre also die Altersverteilung in beiden Gruppen die gleiche, wäre das Problem gelöst.

Um auf das Alter zu adjustieren, führt man ein „virtuelles Matching“ durch: Man addiert die Anzahl der Teilnehmer in einem Stratum (Exponierte + Nicht-exponierte), die Summe setzt man als Teilnehmerzahl in den beiden Gruppen dieses Stratums ein: Damit steht jedem 60-Jährigen in der Kontrollgruppe ein 60-Jähriger in der Expositionsgruppe gegenüber. Multipliziert man diese fiktiven, adjustierten Teilnehmerzahlen mit der gemessenen Mortalitätsrate, erhält man für jedes Stratum eine Mortalitätsrate, die für das Alter adjustiert ist. Aus ihr lässt sich ein adjustiertes RR errechnen. Das adjustierte RR im Beispiel beträgt 1,07. Der Unterschied von 28% zum nichtadjustierten RR zeigt, dass das Alter das Ergebnis stark verzerrt.

Überadjustierung von Confounding

Von Überadjustierung (overadjustment bias) spricht man, wenn auf eine Variable adjustiert wird, die Teil der Kausalkette zwischen Exposition und Studien-

endpunkt und damit kein Confounder ist. In diesem Fall wird die Wirkung der Variablen auf den Endpunkt „heraus-adjustiert“, also das Endergebnis verzerrt. Weil der Computer relativ einfach auf viele Confounder adjustieren kann und es sehr arbeitsaufwändig ist, pathophysiologische und kausale Zusammenhänge sorgfältig auf mögliche Confounder zu untersuchen, ist Überadjustierung nicht selten. Auf der anderen Seite kann es sinnvoll sein, auf einen Teil der Kausalkette zu adjustieren, um zu untersuchen, wie stark diese Komponente den Studienendpunkt beeinflusst. Ob Überadjustierung vorliegt, hängt also von der Fragestellung ab.

P-Werte können keinen Aufschluss darüber geben, ob eine Variable ein Confounder ist oder nicht: Man spricht von Confounding, wenn die adjustierten Ergebnisse von den Rohdaten um mehr als 10% abweichen, es geht also um die Größe des Effekts. Diese Information fehlt dem p-Wert, sie steckt jedoch in dem Konfidenzintervall. ■

Literatur: 1. Grimes DA et al. Bias and causal associations in observational research. *Lancet* 2001;359:248–252. 2. Jager KJ et al. Confounding: What is it and how to deal with it. *Kidney International* (2008) 73:256–260. 3. Tripepi G et al. Stratification for Confounding – Part 1: The Mantel-Haenszel Formula. *Nephron Clin Pract* (2010) 116:c317–c321. 4. Tripepi G et al. Stratification for Confounding – Part 2: Direct and Indirect Standardization. *Nephron Clin Pract* (2010) 116:c322–c325. 5. Fletcher RH, Fletcher SW. *Klinische Epidemiologie. Grundlagen und Anwendung.* Ullstein Medical, 1999. 6. Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health.* 2001;22:189–212.