

## Mathematische Methoden für die Interpretation von Risiken

*Klaus FELSENSTEIN*

### Zusammenfassung

Es wurden mathematische Verfahren entwickelt, mit denen Ergebnisse von klinischen Studien in allgemein vergleichbare Größen umgerechnet werden können. Der Ausgangspunkt besteht aus unterschiedlichen Risiken bei behandelten und nicht behandelten Patienten für ein klinisches „Event“. Die Auswirkungen unterschiedlicher Risiken auf verschiedene Kenngrößen werden berechnet. Eine solche Kenngröße ist die erwartete Zeit bis zu einem Event oder bis zum Tod. Eine wesentliche Verbesserung der Plausibilität der Berechnungen wird durch die Verknüpfung mit der allgemeinen Sterbetafel erreicht. Das allgemeine Risiko wird auf verschiedene spezielle Risiken im Gesamtkollektiv aufgeteilt und an das Patientenalter angepasst. Dadurch ist es möglich, die Auswirkungen sich ändernder Risiken gezielt für spezielle Behandlungssituationen darzustellen.

**Schlüsselwörter:** relatives Risiko, erwartete Lebensdauer, Sterbetafel, Effizienz der Therapie, stochastischer Prozess

### Abstract

We developed mathematical procedures for the calculation of several variables based on the results of a clinical trial. Usually the effects of a therapy are proved by relative reduction of risk. The risk of death or a well defined event for a group of patients receiving therapy is compared to the risk for a group of patients receiving placebo. We calculate the expected lifetime or the time until an event occurs. The estimation of that time is tied to official life tables to be close to reality. The combination of risks concerning the disease and independent risks leads to more precise and realistic results. Frequently it is necessary to split the risk into several parts representing different causes and adjust the risk to the age of the patients. Then we are able to assess the benefit of the therapy more precisely in each individual situation.

**Keywords:** Expected Lifetime, Hazard Rates, Life Tables, Efficiency of Therapy, Stochastic Process

*Anschrift des Autors:* Univ.-Prof. Dr. Klaus FELSENSTEIN  
Institut für wissenschaftliche Statistik, Technische Universität Wien  
Wiedner Hauptstraße 8-10, A-1040 Wien  
*Klaus.Felsenstein@tuwien.ac.at*

## 1 Einleitung und Problemstellung

Das Ergebnis von klinischen Studien besteht zumeist aus einer Darstellung der Wirksamkeit (oder Ineffizienz) einer Therapie in Form von Risikoreduktionen. In einem Vergleich wird eine Gruppe behandelter Patienten einer Kontrollgruppe gegenübergestellt. Die absolute Anzahl oder der relative Anteil der wirkungsvoll behandelten Patienten dient als Evidenz für die generelle Wirksamkeit der Therapie. Die Unterschiede in den beiden Gruppen werden dann mit standardisierten statistischen Verfahren analysiert. Mit solchen Verfahren (wie etwa t-Test, exakter Test, Varianzanalyse, etc.) lassen sich signifikante Unterschiede nachweisen. Die Voraussetzungen für eine prinzipielle Aussagekraft der statistischen Verfahren, wie vergleichbare Gruppengröße und Zusammensetzung der Gruppen, vergleichbare Altersstruktur und Allgemeinzustand der Patienten und vor allem hinreichende Anzahl an Patienten und Dauer der Studie, ist bei den meisten Untersuchungen gegeben und wird auch überprüft.

Auch wenn also die Vorbedingungen für die Relevanz einer statistischen Aussage erfüllt sind, bleibt das Problem der Interpretation dieser Aussage. Selbst wenn in der Studie die Patienten randomisiert den beiden Gruppen zugewiesen wurden, können nicht alle unerwünschten Einflussfaktoren ausgeschaltet werden. Die statistische Aussage bleibt auf eine Versuchsanordnung beschränkt, von der gehofft wird, dass sie möglichst allgemein vorliegt. Es besteht die Notwendigkeit, die aus der Studie gewonnene Erkenntnis in verschiedener Art zu interpretieren. Die prinzipielle Frage lautet: „Welche allgemeinen Auswirkungen hat der Unterschied in den beiden Gruppen?“. Mit mathematischen Modellen lassen sich für verschiedene Kenngrößen die Unterschiede berechnen. So lässt sich beispielsweise die Frage „Welcher Unterschied ergibt sich

für die Lebenserwartung in der Gruppe behandelter Patienten zur Kontrollgruppe aus der Risikoreduktion?“ mit statistischen Berechnungen beantworten. Die Resultate der Studie können dann auf ihre Auswirkungen in der allgemeinen Anwendung besser diskutiert werden. Zudem können Widersprüche verschiedener Studien leichter erkannt werden, wenn die spezifischen Aussagen zu einer Erkrankung auf unterschiedliche Auswirkungen führen. Stark abweichende Berechnungen der Lebenserwartungen geben beispielsweise einen Hinweis darauf, dass die verschiedenen Studien nicht direkt vergleichbar sind oder in einer Studie ein methodischer Fehler vorliegt. Prinzipiell lassen sich beliebige statistische Kenngrößen als Funktion der Studienergebnisse, etwa der Risiken, darstellen. Es liegt an der angewandten Sorgfalt bei der Auswahl des zugrunde liegenden Modells, ob die Resultate der Berechnungen der Realität entsprechen. Dies lässt sich am ehesten dann bejahen, wenn sich aus dem Ergebnis der Studie möglichst viele Kenngrößen errechnen lassen, die in der medizinischen Praxis Bestätigung finden. Der Wert einer Studie steigt enorm, wenn die Auswirkungen in vielfältiger Hinsicht darstellbar sind und möglichst wenig Widerspruch zu anderen Ergebnissen liefern.

## 2 Interpretation des Risikos

Wenn die Ergebnisse einer klinischen Studie als relative Risiken für Erkrankung oder Tod bzw. in Form von absoluten Zahlen oder Prozenten für Patienten unter Behandlung und Patienten in der Kontrollgruppe vorliegen, ist zu berücksichtigen, dass sich die Risiken meist nur auf einen relativ kurzen Zeitraum, in dem die Patienten beobachtet wurden, beziehen. Wenn es auch verführerisch ist, die Risiken linear auf einen beliebigen Zeitraum umzurechnen, so führt dies

zu irreführenden Angaben. Risiken werden als Faktoren verwendet und können daher nicht linear verknüpft werden (das ist nur für die jeweiligen Logarithmen möglich).

Angenommen das Risiko pro Jahr zu erkranken, beträgt in der Kontrollgruppe 5% und in der Behandlungsgruppe nur 4%; die Behandlung reduziert also das Risiko um 1%. Das Risiko für eine Erkrankung innerhalb von 20 Jahren ist aber in der Kontrollgruppe nicht um 20% höher als in der Behandlungsgruppe. Der Unterschied zwischen beiden Gruppen beträgt 6,8%, vorausgesetzt, das Risiko verändert sich über die Jahre nicht.

Aus den angegebenen Risiken bestimmt man nun die Verteilung der Zeit  $T$  bis zum Eintritt des Events. Das Event ist gemäß der Untersuchung exakt zu definieren. Es kann den Todeszeitpunkt bedeuten, den Eintritt eines medizinischen Vorfalls, wie Infarkt oder Knochenbruch etc., oder auch die erstmalige Diagnose einer Erkrankung. Wenn die Risiken für verschiedene Altersgruppen von Patienten bekannt sind, ergibt sich die *Überlebensfunktion* als Produkt. Die Überlebensfunktion gibt für jeden Zeitpunkt  $t$  die Wahrscheinlichkeit an, dass bis zu diesem Zeitpunkt kein Event eingetreten ist,  $P[T > t]$ . Eine präzise Überlebensfunktion erhält man, wenn das Risiko für jedes Alter in der Studie angegeben wurde. Das ist kaum der Fall. In der Regel gibt es Angaben über eine Periode von 5 bis 10 Jahren um den Median des Alters der Patienten. Wenn es keine zusätzlichen Informationen über die Veränderung des Risikos bei Patienten, deren Alter nicht von der Studie abgedeckt wird, gibt, kann nur ein über die Zeit konstantes Risiko angenommen werden, was manchmal nicht der Realität entspricht.

Bei langfristigen Studien wird das Resultat häufig in Form einer *Kaplan-Meier-Kurve* abgebildet, da die an der Untersuchung teilnehmenden Patienten nur einen Teil der Dauer der Studie beobachtet werden. Dann ist es möglich, das Risiko für jedes Alter

abzulesen bzw. für fehlende Risikowerte eine brauchbare Interpolation durchzuführen. Jedenfalls können unterschiedliche Parameter aus der Verteilung der Zeit bis zum Event berechnet werden. Die häufigste verwendete Kenngröße ist der Erwartungswert. Als Hilfe für die Interpretation des Risikos eignen sich oft andere Verteilungsparameter besser. Der Median (Wert, wobei die Hälfte der Patienten eine kürzere und die Hälfte eine längere Zeit bis zum Event hat) oder der Modus (Dauer bis zum Event, den die größte Untergruppe des Kollektivs hat) bieten oft einen verständlicheren Ansatzpunkt für die Erklärung der Auswirkungen für den einzelnen Patienten.

Der Durchschnitts- oder Erwartungswert dient als markanter Wert für das gesamte Merkmal und beschreibt die Allgemeinsituation. Als aussagekräftige Prognosezahl ist der Erwartungswert im individuellen Einzelfall kaum aussagekräftig. „Kein Individuum fühlt sich durch einen Durchschnittswert (insbesondere bei der Lebenserwartung) charakterisiert.“ Im Allgemeinen wird hier etwa der Modus als individuelle Kennzahl besser akzeptiert. Die Verteilung der Zeit bis zum Event ist im Allgemeinen nicht symmetrisch, was bewirkt, dass sich Erwartungswert, Median und Modus unterscheiden. Erst wenn die Gruppe, zu der der Patient gehört, bezüglich anderer Risiken, Alter, Geschlecht, Vorerkrankungen etc. genauer eingegrenzt ist, kommt dem Durchschnittswert mehr individuelle Aussagekraft zu (Siehe Abschnitt 5 und 6).

### 3 Vergleich mit Sterbetafeln

Bei Erkrankungen mit kleinerem Risiko ergibt sich eine hohe Durchschnittsdauer bis zum Event. Hat beispielsweise ein 50jähriger Patient ein (gleich bleibendes) Risiko pro Jahr von 1% für ein Event, dann erhält man eine 99jährige Durchschnittsdauer bis zum Eintritt der Erkrankung. Der

Patient müsste 149 Jahre werden, um dieses Event zu erleben. Bei älteren Patienten übersteigt auch die Dauer bis zum Event unter wesentlich höherem Risiko noch die allgemeine Lebenserwartung des Patienten. Es ist daher notwendig, die Dauer bis zum Event an die allgemeine Sterbewahrscheinlichkeit zu knüpfen, um realistische Durchschnittszeiten zu erhalten.

Als Referenz für die Berechnung der Lebenserwartung dient die allgemeine Sterbetafel. Für die österreichische Sterbetafel<sup>1</sup> kann gezeigt werden, dass die allgemeine Lebensdauerverteilung einer stetigen Gumbel-Verteilung entspricht. Bei der Berechnung der Durchschnittszeiten wurde trotzdem auf die Modellierung mit einer stetigen Lebensdauerverteilung<sup>2</sup> verzichtet. Jede auch noch so sorgfältig angepasste stetige Lebensdauerverteilung impliziert Gesetzmäßigkeiten, die sich auf die Berechnungen des Erwartungswertes auswirken. Im Allgemeinen überschätzen solche parametrischen Modelle die erwartete Lebensdauer. Um ein möglichst realistisches Modell zu erstellen, wurden für alle stochastischen Größen (Lebensdauern, ereignisfreie Zeiten, etc.) diskrete Wahrscheinlichkeitsverteilungen mit Punktwahrscheinlichkeiten verwendet. Die Sterberisiken veränderten sich infolge einer steigenden Lebenserwartung innerhalb der letzten 10 Jahre merklich. Es ist bei der Verwendung einer Sterbetafel darauf zu achten, dass die Tabelle für die Patienten der Studie gültig ist. Auch die Zusammensetzung (Anteil von Männern und Frauen) soll der entsprechend gewichteten Summe der Sterberisiken für Männer und Frauen entsprechen.

Die Lebensdauerverteilung aus der Sterbetafel wird der Verteilung der Dauer bis zum Event, die sich aus der Studie ergibt, überlagert. Dabei wird das Event  $E$  aus der Studie durch ein Event, das entweder  $E$  oder „Tod des Patienten“ bedeutet, ersetzt. Die Zeit bis zum Event  $T$  wird entsprechend durch das Minimum von  $T$  und der Lebens-

dauer  $L$  mit der aus der Sterbetafel berechneten Verteilung ersetzt.

Die Verteilung des Minimums von  $T$  und der Lebensdauer  $L$  errechnet sich zunächst aus dem Produkt der Überlebensfunktionen beider Größen. Das impliziert, dass die beiden Zeiten als unabhängig verstanden werden. Für seltene Krankheiten, die keinen wesentlichen Anteil am allgemeinen Sterberisiko haben, ist diese Annahme sicher gerechtfertigt. Auch wenn die Krankheit nicht zum Tode führt bzw. das Event keinen direkten Einfluss auf die weitere Lebenserwartung besitzt, ist die Annahme der Unabhängigkeit bzw. einer vernachlässigbaren Korrelation vertretbar. Auch wenn die untersuchte Patientengruppe aus vorwiegend recht jungen Patienten mit an sich geringem Sterberisiko besteht, sind die Berechnungen des Minimums unter unabhängigen Variablen zu rechtfertigen.

Trifft die Unabhängigkeitshypothese nicht zu, etwa bei Krankheiten, die einen erheblichen Anteil am allgemeinen Sterberisiko haben, sind weitere Unterteilungen und Spezifikationen bei der Berechnung von Risiko und Durchschnittszeit notwendig (siehe Abschnitt 5 und 6).

Die Überlagerung der Verteilung der Dauer bis zum Event mit einer Tabelle von Eintrittswahrscheinlichkeiten für das Kollektiv ist nicht auf Sterbewahrscheinlichkeiten beschränkt. Es kann in manchen Fällen sinnvoll sein, das untersuchte Event mit allgemeinen Unfallereignissen zu verknüpfen. Dann gibt die Unfallstatistik das allgemeine Risiko und daher die Dauer bis zu einem Unfallereignis an, und es wird die Dauer bis zum Event wie bei der Lebensdauerverteilung mit dem Minimum der entsprechenden Zeiten berechnet. Die enge Anbindung an gesicherte epidemiologische Daten (etwa Sterbetafel oder allgemeine Unfallstatistik) ermöglicht es, auch Resultate aus kleineren Studien, die für sich genommen weniger allgemeingültige Aussagekraft haben, zu universellen Aussagen zu führen.

## 4 Berechnung der Lebensdauer

Die längerfristigen Effekte der Behandlung sollen für die Patienten aus den eher kurzfristigen Untersuchungen in der klinischen Studie mit stochastischen Modellen prognostiziert werden. Die Auswirkungen lassen sich durch verschiedene Kenngrößen messen. Wichtige Kenngrößen sind etwa die Lebenserwartung des Patienten, die krankheitsfreie Zeit (Dauer bis zum Krankheitsvorfall oder Event-freie Zeit), der Median oder Modus dieser stochastischen Größen oder die Wahrscheinlichkeit, an einer bestimmten Krankheit überhaupt jemals zu erkranken. Durch diese Größen werden die relativen Risiken aus einer Studie in einen für Arzt und Patienten greifbaren und vergleichbaren Ausdruck übersetzt.

Mit der Überlebensfunktion, als Kombination der Funktion, die sich aus der Studie ergibt, und der allgemeinen Sterbetafel erhält man die restliche durchschnittliche Event-freie Zeit als Summe aller (bedingten) Überlebenswahrscheinlichkeiten beginnend beim aktuellen Alter des Patienten. Die bedingte Überlebenswahrscheinlichkeit entspricht dem Quotienten aus der allgemeinen Überlebenswahrscheinlichkeit und der Wahrscheinlichkeit, das aktuelle Alter überhaupt zu erreichen.

Der Median der Event-freien Zeit  $med(T)$  ist dann erreicht, wenn die bedingte Überlebenswahrscheinlichkeit erstmals 50% unterschreitet. Der Modus  $mod(T)$  wird durch die maximale Veränderung der bedingten Überlebenswahrscheinlichkeit berechnet. Das Jahr, in dem die Differenz zwischen der Überlebenswahrscheinlichkeit am Beginn des Jahres und am Ende des Jahres am größten ist, entspricht dem Modus.

Der Erwartungswert (oder der Modus bzw. Median) der Event-freien Zeit dient auch als Kriterium, ob eine spezielle Therapie effizient ist oder nicht. Das Resultat einer Studie zeige wieder die relative Reduktion der Ri-

siken in der Behandlungsgruppe gegenüber der Kontrollgruppe. Dadurch erhält man verschiedene Überlebenswahrscheinlichkeiten und damit auch verschiedene Verteilungen der Event-freien Zeit für die beiden Gruppen. Es bezeichne etwa  $T_C$  die Event-freie Zeit in der Kontrollgruppe und  $T_B$  jene in der Behandlungsgruppe. Beide stochastischen Größen werden wieder mit der allgemeinen Sterbetafel verknüpft. Die Differenz der beiden Erwartungswerte (Dauer in der Kontrollgruppe und der Behandlungsgruppe) lässt sich zunächst als Gewinn der Therapie ansehen.

Eine ehrlichere Bewertung der Qualität der Therapie erhält man, wenn man den Unterschied in der Zeit bis zum Event nur unter jenen Patienten misst, die tatsächlich ein Event erlebt haben. Ist die verbleibende Lebenserwartung eines Patienten nicht groß, dann wird die Therapie ihren etwaigen Nutzen durch eine Verlängerung der Event-freien Zeit kaum beweisen können. Mathematisch formuliert man das so, dass die Zeit in der Kontrollgruppe  $T_C$  sich als Minimum der Zeit in der Behandlungsgruppe und einer weiteren Zeit  $T_N$  ergibt,  $T_C = \min(T_B; T_N)$ . Diese zusätzliche Variable  $T_N$  beschreibt ein durch 'Nichtbehandlung' entstandenes Risiko, was zu einer eventuell kürzeren Event-freien Zeit bei nichtbehandelten Patienten führt. Die Verteilung dieser Variablen  $T_N$  bestimmt die in der Studie angegebene relative Reduktion der Risiken.

Im konkreten Fall eines Patienten hat die Therapie nur dann einen Nutzen gebracht, wenn  $L > T_B > T_C$ , wobei  $L$  die Lebenszeit des Patienten, verteilt nach der allgemeinen Sterbetafel, bezeichnet. Die Wahrscheinlichkeit für dieses Ereignis (Wirksamkeit = WI)  $WI = \{L > T_B > T_C\}$ , also wenn der Patient erst nach dem Event verstirbt und durch die Behandlung der Zeitpunkt des Events hinausgeschoben wurde, gibt den Anteil an, wie oft die Therapie wirklich nützlich war.

Die durchschnittliche Verlängerung der Zeit bis zum Event durch die Therapie wird

demfolgend als bedingter Erwartungswert berechnet. Die Bedingung ist das Ereignis  $WI$ , die Wirksamkeit der Therapie. Diese Aufteilung der Patienten in solche, die von der Therapie profitieren und in solche, die den Nutzen nicht erleben, erleichtert insbesondere dann die Interpretation des individuellen Risikos, wenn die Therapie tatsächlich nur bei einem Teil der Gesamtheit effektiv ist und beim Rest keinerlei Wirkung zeigt.

Konkrete Berechnungen haben beispielsweise bei der UKPDS 33-Studie<sup>3</sup> hohe Übereinstimmungen der prognostizierten und der in der Studie auch angegebenen Zeiten bis zum Event ergeben.

Ist nach Beginn der Therapie zumindest ansatzweise zu erkennen, dass die Therapie wirkt, dann lässt sich auch eine genauere Prognose für die zu erwartende Verlängerung der Event-freien Zeit angeben.

## 5 Kombination verschiedener Risiken

Die obige Zerlegung des Kollektivs in Gruppen mit stark unterschiedlichem Risiko ist ein Weg, die individuelle Interpretation eines abstrakten Risikos für eine Erkrankung zu verbessern. Allgemein wird für jeden Patienten eine große Zahl an individuellen Risiken zu berücksichtigen sein. Bei gleichzeitiger Behandlung mit mehreren Therapien unterschiedlicher Wirksamkeit müssen die Risiken kombiniert werden.

Wieder ist für die Interpretation von Risiken prinzipiell zu beachten, dass die Risikofaktoren multiplikativ verknüpft werden müssen. Der arithmetische Mittelwert besitzt keine Aussagekraft über das durchschnittliche Risiko. Angenommen das Risiko für eine Erkrankung beträgt in einer Zeiteinheit 1% und das Risiko für eine andere Erkrankung, die nicht im Zusammenhang mit der ersten Erkrankung steht, beträgt 10%: dann ist das Gesamtrisiko, also dass mindestens eine der beiden Erkrankungen auf-

tritt, nicht 11%, sondern nur 10,9%. Bei mehr als zwei Krankheiten, bzw. bei der Berücksichtigung von etlichen Risikofaktoren, wie es realistischer Weise notwendig ist, verändert sich das Resultat der Berechnung des Gesamtrisikos noch wesentlich deutlicher.

Auch für die Berechnung von Risikofaktoren in verschiedenen Patientengruppen ist der arithmetische Mittelwert der Risiken nicht interpretierbar. Eine Patientengruppe bestehe etwa aus 10 Personen, von denen 5 ein Risiko für eine Erkrankung von 80% und die restlichen 5 Personen ein Risiko von 10% haben. Das durchschnittliche Risiko in dieser Gruppe beträgt dann nicht 45%, sondern 57,6%. Allgemein werden hier geometrische Mittelwerte anstelle von arithmetischen Mittelwerten für Faktoren berechnet. Oft wird sich die Gesamtbevölkerung sehr heterogen aus verschiedenen Gruppen mit sehr unterschiedlichem Risiko zusammensetzen. Durch die Festlegung differenzierter Patientengruppen mit unterschiedlichen Risiken steigt die Qualität von Aussagen über das Kollektiv. Die Risikofaktoren ändern sich oft gravierend mit dem Alter der Patienten. Die kontinuierliche Untersuchung des Risikos über alle Altersgruppen verstärkt die Aussagekraft einer Untersuchung wesentlich.

Bei vielen Behandlungen ist es notwendig, die Nebenwirkungen als eigene Risikofaktoren zu betrachten. Die verringerten Risiken für die Primärerkrankung werden durch ein infolge der Behandlung höheres Risiko für andere Events teilweise wieder aufgehoben. Die Vorgangsweise für die Verknüpfung des Erkrankungsrisikos und des Risikos für Nebenwirkungen entspricht genau jener bei der Verbindung der Verteilung der Event-freien Zeit mit der Verteilung der Lebensdauer aus der Sterbetafel. Jedem einzelnen Risikofaktor  $i$  entspricht eine zugehörige Zeit  $T_{i,j}$  bis zum Event. Die für den Patienten maßgebliche Variable ist dann das Minimum aller dieser Zeiten, also die Zeit bis zum ersten aller dieser Events. Solange die einzel-

nen Faktoren als voneinander unabhängig zu sehen sind, ist diese Verteilung des Minimums einfach durch das Produkt der Überlebensfunktionen zu berechnen. Ist die Unabhängigkeitshypothese nicht aufrecht zu halten, dann muss bei der Verknüpfung der Verteilungen die Korrelation berücksichtigt werden. Für das Erkrankungsrisiko und das Risiko für Nebenwirkungen besteht meist eine direkte Abhängigkeit. Die Korrelation zwischen diesen beiden Risiken ist, zumindest was numerische Angaben betrifft, nur selten in der Praxis untersucht. Oft beschränkt sich die Angabe über die Nebenwirkung nur auf einen Prozentsatz von Behandlungen, bei denen Nebenwirkungen aufgetreten sind.

Um die Verbindung zwischen abhängigen Risikofaktoren im Modell auszudrücken, beschreitet man bei der Analyse von Lebensdauern den Weg, ein Regressionsmodell zwischen den Überlebensfunktionen der verschiedenen Risikofaktoren zu verwenden. In einer Gleichung wird der Zusammenhang zwischen den Überlebenswahrscheinlichkeiten ausgedrückt. Die Koeffizienten in dieser Gleichung werden aus den Daten der Studie berechnet. So werden für die Umrechnung von Kaplan-Meier-Kurven in Lebenserwartungen oder Eintrittswahrscheinlichkeiten Regressionsmodelle verwendet. Aus der Zuverlässigkeitsanalyse erhält man Vorschläge, wie die Gleichungen zwischen den Überlebenswahrscheinlichkeiten zweckmäßigerweise gewählt werden sollen. In COX<sup>4</sup> und ANDERSEN et al.<sup>5</sup> wird die Gestalt solcher Regressionsmodelle analysiert. Wenn zu einem speziellen Risiko mehrere Untersuchungen bereits durchgeführt wurden, lassen sich die Informationen über die Koeffizienten im Regressionsmodell übertragen. Anstelle von Kleinst-Quadrat-Schätzern, die eine möglichst gute geometrische Anpassung der beiden Überlebensfunktionen ergeben, werden sogenannte Bayes-Schätzer eingesetzt, die zusätzlich Vorinformationen (etwa aus anderen Studien) einbringen. Die Verfahren

lassen sich auch auf Resultate von Studien anwenden, bei denen die Patienten in Klassen eingeteilt sind.<sup>6</sup>

Herzerkrankungen stellen ein Beispiel dar, wo man für Patienten höheren Alters annehmen kann, dass die Überlebensfunktion (also für jeden festen Zeitpunkt die Wahrscheinlichkeit, dass bis zu diesem Zeitpunkt kein Event eintritt) eine ähnliche Gestalt hat wie die allgemeine Überlebensfunktion aus der Sterbetafel. Im einfachsten Fall kann man annehmen, dass für die Eventfreie Zeit  $T$  und die allgemeine Lebensdauer  $L$  der Zusammenhang  $P[T > t] = aP[L > t]$  für einen Koeffizienten  $0 < a < 1$  gilt. Der Koeffizient  $a$  wird aus den Daten berechnet. Das wäre das einfachste Regressionsmodell, in der Praxis wird man wesentlich komplexere Regressionsmodelle verwenden. Die Wahrscheinlichkeiten werden bedingt unter dem Alter angesetzt, die Koeffizienten hängen von der Zeit und anderen Parametern ab und andere Modifikationen werden vorgenommen. Auch für die Umrechnung von Kaplan-Meier Kurven in Lebenserwartungen oder Eintrittswahrscheinlichkeiten werden allgemeine Regressionsmodelle verwendet. Auch diese Regressionsmodelle beinhalten die Möglichkeit, a-priori Information über die Kerngrößen wie Lebenserwartung oder Eintrittswahrscheinlichkeit einfließen zu lassen.

## 6 Risikozerlegung

Die Überlagerung einer durch die Krankheit verursachten Lebensdauer bis zum Event mit der allgemeinen Lebensdauerverteilung, wie sie im Abschnitt 3 erklärt wurde, erfolgt unter der Unabhängigkeitshypothese der Risiken durch Multiplikation der jeweiligen Überlebensfunktionen. Dies ist gerechtfertigt, wenn das Erkrankungsrisiko im Vergleich zum allgemeinen Sterberisiko klein ist und daher die spezielle Krankheit

keinen hohen Anteil an der allgemeinen Sterblichkeitsrate besitzt. Ist eine Krankheit hingegen weit verbreitet bzw. trägt das Risiko dieser Krankheit für Patienten dieses Alters einen wesentlichen Anteil an der Sterbewahrscheinlichkeit, dann besteht eine hohe Wechselwirkung zwischen dem allgemeinen Risiko und dem speziellen Krankheitsrisiko. Die Abhängigkeit der beiden Risiken muss ins Modell eingehen, da sonst Verstärkungseffekte für einzelne Risikofaktoren eintreten.

Die Zerlegung des allgemeinen Risikos in verschiedene einzelne Risikofaktoren bedarf der Kenntnis bzw. einer statistischen Schätzung des Anteils erkrankter Personen in der Gesamtbevölkerung. Das Gesamtrisiko, ausgedrückt durch die Überlebensfunktion  $P[L > t]$ , wird in die Überlebensfunktion erkrankter Personen  $P[T_K > t]$  und die Überlebensfunktion nicht erkrankter Personen  $P[T_N > t]$  mit dem Gewicht  $0 < A < 1$  gemäß  $P[L > t] = AP[T_K > t] + (1 - A)P[T_N > t]$  aufgeteilt. Dabei ist  $A$  der Prozentanteil der Bevölkerung, die an dieser Krankheit leidet; über  $P[T_K > t]$  bzw. das Krankheitsrisiko erlangt man aus der Studie Informationen, die Überlebensfunktion  $P[T_K > t]$  kann beispielsweise durch eine Kaplan-Meier-Kurve bestimmt werden. Aus der obigen Gleichung lässt sich dann auch der Überlebensfunktion des nicht erkrankten Kollektivs berechnen.

Ein sehr anschauliches Beispiel für eine solche Zerlegung des allgemeinen Risikos wird in BONELLI et. al (2004)<sup>7</sup> demonstriert. Dort wird das allgemeine Risiko in jenes für Raucher und Nichtraucher zerlegt. In der amtlichen Sterbetabelle sind natürlich die durch das Rauchen verursachten Sterbefälle enthalten.

Die Erhöhung des Sterberisikos durch das Rauchen ist in vielen Statistiken epidemiologisch nachgewiesen, daher kann die Überlebensfunktion  $P[T_K > t]$  für Raucher genau bestimmt werden. Der Anteil  $A$ , den Raucher bei Männern und Frauen in der Bevölkerung stellen, ist ebenfalls bereits genau

ermittelt. Durch die obige Gleichung wird dann die Überlebensfunktion  $P[T_N > t]$  der Nichtraucher berechnet. Mit den beiden Verteilungen lässt sich ein Vergleich der Lebenserwartung für Raucher und Nichtraucher ziehen. Die Resultate für die verminderte Lebenserwartung von Rauchern in BONELLI et. al (2004) decken sich mit Ergebnissen aus anderen getrennten Untersuchungen für Raucher und Nichtraucher. Durch die Bestimmung der Überlebensfunktion  $P[T_K > t]$  für Raucher kann man die Anpassung eines beliebigen Risikos einer Erkrankung an die Sterbetafel verbessern. Wenn man weiß, dass der Patient Raucher ist, wird der krankheitsspezifischen Überlebensfunktion nicht die allgemeine Überlebensfunktion, sondern jene für Raucher überlagert. Dadurch wird die Effizienz einer Therapie und die erwartete Verlängerung der Event-freien Zeit durch eine Therapie neu berechnet und in vielen Fällen merklich geringer werden.

Wendet man das Prinzip der Risikozerlegung auf eine beliebige Krankheit an, wird der Anteil von Erkrankten im Kollektiv für jedes Alter meistens zu schätzen sein. Auch die Zuteilung eines Patienten zu einer Gruppe lässt sich selten so einfach wie bei Rauchern und Nichtrauchern durchführen. Dieser Zerlegungsalgorithmus des Risikos ermöglicht wie bei der Lebensdauer von Rauchern und Nichtrauchern getrennte Prognosen für die unterschiedlichen Patientengruppen. Diese Zerlegung ist im mathematischen Modell auch dann möglich (und dann erst recht aufschlussreich), wenn für einen einzelnen Patienten keine definitive Zuordnung zu einer Risikogruppe getroffen werden kann.

## 7 Bewertung des Nutzens

Für die allgemeine Bewertung einer Therapie werden auch Kenngrößen berechnet, die die ökonomische Zweckmäßigkeit betreffen. Die erwartete Event-freie Zeit un-



ter der Therapie wird größer als die erwartete Event-freie Zeit für nicht behandelte Patienten sein. Die Differenz der beiden Zeiten sollte aber in einem rationellen Verhältnis zum Aufwand stehen.

Die Effektivität einer Therapie lässt sich durch die entstandenen Kosten pro erwarteter Zeitdifferenz messen. Wenn die Kosten eine Funktion von der Dauer der Behandlung sind, ist der Erwartungswert mit der Verteilung der Zeit bis zum Event bzw. Tod zu berechnen. Das Verhältnis zwischen erwarteten Kosten und zu erwartenden Lebensverlängerung gibt die Kosten pro gewonnenem Jahr an.

Für die Bewertung der Effizienz wird in vielen Publikationen auch die Anzahl der benötigten Behandlungsjahre bis zu dem Zeitpunkt, an dem eine merkliche Verlängerung der Restlebensdauer erwartet werden kann, angegeben. Man kann etwa die Anzahl der Jahre, bis ein behandelter Patient eine um ein Jahr höhere Lebenserwartung hat als eine Person aus der Kontrollgruppe, als Bewertung für die Effizienz betrachten. Dabei ist zu berücksichtigen, dass dies innerhalb der restlichen Lebenszeit geschehen muss. Es wird der Erwartungswert für die Lebensdauer eines therapierten bzw. für einen Patienten aus der Kontrollgruppe unter der Bedingung, dass beide diese Zeitspanne überlebt haben, berechnet. Die Anzahl benötigter Jahre Behandlung für den Gewinn eines Jahres (Number Needed to Treat, NNT) ist dann die kleinste Anzahl an Jahren, für die der Unterschied beider Erwartungswerte mindestens 1 beträgt.

Ein mathematisches Konzept zur Entwicklung einer optimalen Strategie bilden sogenannte Nutzenfunktionen.<sup>8</sup> Darunter versteht man eine Funktion  $U$ , die allen Ergebnissen eines Experiments und den Parametern einen Wert zuordnet, der eine numerische Bewertung des Nutzens darstellt. Es sei etwa  $T$  die Lebensverlängerung,  $C$  die Kosten der Behandlung,  $D$  die Dauer der Behandlung,

$NW$  ein Grad für die Nebenwirkungen, und  $B$  die gewählte Behandlungsmethode, dann wird  $B$  durch  $U(T,C,D,NW)$  bewertet. Die Auswahl der optimalen Therapie erfolgt dann durch den Erwartungswert des Nutzens  $EU$ .

Durch ein adaptives Verfahren kann über Vergleiche eine Nutzenfunktion bestimmt werden, die eine eindeutige Präferenzordnung einhält. Mit der Nutzenfunktion lassen sich dann je zwei Situationen mit unterschiedlichen Variablen  $(T,C,D,NW)$  vergleichen. Die optimale Strategie besteht in der Auswahl einer Behandlung, sodass keine andere Behandlung einen höheren erwarteten Nutzen besitzt. Die mathematische Optimierungsaufgabe bei der Suche nach optimalen Strategien ist weitgehend gelöst und stellt kein Hindernis für den Einsatz von Nutzenfunktionen in der Praxis dar. Das Problem besteht eher darin, dass es kaum gelingt, allgemein akzeptierte Nutzenfunktionen angeben zu können. Das entscheidungstheoretische Konzept wird nur dort erfolgreich einsetzbar sein, wo es übereinstimmende Einschätzungen von Behandlungsvorteilen und Nebenwirkungen gibt.

#### Referenzen:

- 1 ÖSTERREICHISCHES STATISTISCHES ZENTRALAMT, *Sterbetafel 1990/92 für Österreich*, Österreichisches Statistisches Zentralamt, Wien 1993
- 2 HARTUNG J., *Statistik*, Oldenbourg Verlag, München 1993
- 3 UK PROSPECTIVE DIABETES STUDY (UKPDS) GROUP, *Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33)*, Lancet (1998); 352: 837-853
- 4 COX D. R., *Regression models and life-tables*, J Royal Stat Soc B (1972); 34: 187-220
- 5 ANDERSEN P., GILL R., *Cox regression model for counting processes: a large sample study*, Ann Statist (1982); 10: 1100-1120
- 6 FELSENSTEIN K., PÖTZELBERGER K., *Asymptotically optimal transformations for grouped data*, Computing Science and Statistics (1997); 28: 602-607
- 7 BONELLI J., FELSENSTEIN K., *Jeder Raucher schädigt sich selbst*, Imago Hominis (2004) 11: 51-54
- 8 ROBERT C., *The Bayesian choice: a decision-theoretic motivation*, Springer Verlag, New York 2002