

Statistische Auswertungen von Therapie-Ergebnissen

Einleitung

Um eine Behandlungstechnik zu rechtfertigen oder zu verbessern, wird eine Bilanz über bereits vorliegende Therapie-Ergebnisse benötigt. Solche Auswertungen werden begründet mit Vergleichen, Verbesserungen oder Nutzen von Therapieformen. Die Therapieergebnisse von definierten Patientengruppe können in Überlebenskurven anschaulich dargestellt werden, die aus objektiven und nachvollziehbaren Auswertemethoden resultieren. Die Methoden sind seit vielen Jahren standardisiert und basieren weitgehend auf Richtlinien und Empfehlungen für die Registrierung von Krebsfällen sowie für statistische Darstellungen.

Die Bestimmung von Überlebenszeiten

Als Grundlage für eine nachfolgende Auswertung müssen von jedem Patienten eines Kollektivs bestimmte und einheitliche Daten aufgelistet werden. Diese hängen teilweise von der Fragestellung ab. Das Sammeln der notwendigen Patientendaten ist vielfach eine langwierige Angelegenheit, wobei die angewandte Sorgfalt und Vollständigkeit die Aussagekraft und die Zuverlässigkeit der resultierenden Überlebenskurve bestimmt.

Die Grundlage der Analyse von Überlebenszeiten bildet das individuelle Zeitintervall zwischen einem definierten Beobachtungsbeginn - in der Regel das Diagnosedatum oder der Therapiebeginn - und einem Beobachtungsende. Meistens wird der Tod des Patienten als beobachtetes Ereignis definiert. In diesem Fall definiert der Tod eines Patienten eine echte Beobachtungszeit. Häufig lebt ein Teil der Patienten am Ende des Beobachtungszeitraumes der klinischen Studie noch. Mit der Wahl eines Ersatzendpunktes (z.B. letzte Kontrolle) erhält man eine sogenannte zensierte Beobachtungszeit. Aus den einzelnen Beobachtungszeiten wird das Resultat für das Kollektiv bestimmt. Die Darstellung von Überlebenszeiten in Form von arithmetischen Mittelwerten ist zwar ein beliebtes, aber unzweckmässiges Verfahren. Man erhält damit hohe mittlere Überlebenszeiten, da vor allem bei kleinen Patientengruppen schon ein Ausreisser nach oben den Mittelwert zu hohen Werten hin verschieben kann. Wenn schon ein Mittelwert angegeben wird, dann sollte es der Median-Wert sein.

Die in der Medizin am weitesten verbreitete statistische Methode ist die 5-Jahres-Überlebensrate. Diese ist nur unter der Voraussetzung anwendbar, dass bei allen Patienten in der Studie nach fünf Jahren bekannt ist, ob sie noch leben oder in der Zwischenzeit gestorben sind. Die Hauptbedeutung der Überlebenszeit-Analyse liegt jedoch in der statistischen Berücksichtigung von zensierten Daten. Die statistische Aussagekraft steigt wegen der grösseren Patientenzahl. Die graphische Darstellung von Überlebenszeiten erfolgt in der Regel anhand von Absterbekurven.

Das Kaplan-Meier-Verfahren

Das Kaplan-Meier-Verfahren basiert auf individuellen Überlebenszeiten und eignet sich auch bei einer kleineren Anzahl von Stichproben. Die Auswertung kann nach verschiedenen Kriterien vorgenommen werden. Anstelle des Todes kann als Ereignis das Auftreten von Rezidiven bzw. Metastasen oder der durch den Tumor bedingte Todesfall definiert werden. Die Überlebenskurve wird dann anstelle von Overall Survival entsprechend Disease-Free Survival, Event-Free Survival oder Adjusted Survival genannt. Die Überlebenszeiten mit solch definierten Ereignissen können nur mit der Sterbetafel-Methode oder der Methode nach Kaplan und Meier bestimmt werden, weil auch nach einer langen Beobachtungszeit nie alle Beobachtungen echt werden können.

Die echten und die zensierten Überlebenszeiten von allen Patienten in der beobachteten Gruppe werden der Grösse nach aufgelistet. Bei allen echten Zeitwerten wird die Überlebenswahrscheinlichkeit berechnet. Diese hängt von der Anzahl der Patienten ab, bei denen zum beobachteten Zeitpunkt das Risiko für das Ereignis noch besteht (Anzahl Patienten in der Risikogruppe bzw. "Patients at risk"). Die Multiplikation aller vorangegangenen Überlebenswahrscheinlichkeiten bis zum beobachteten Wert ergibt die kumulierte Überlebensrate für den betreffenden Zeitpunkt. Es wird also die Wahrscheinlichkeit berechnet, eine bestimmte Zeitspanne zu überleben.

Es hat sich gezeigt, dass als minimale Anzahl für die Auswertung ein Kollektiv von 15 Patienten vorliegen sollte. Die Darstellung der Überlebenskurve sollte abgebrochen werden, wenn sie nur noch durch 1 bis 2 Beobachtungszeiten gestützt wird. Die wichtigsten Eigenschaften der Auswertung nach Kaplan und Meier sind die folgenden:

- Berücksichtigung von Patienten mit kurzen Beobachtungszeiten
- Berücksichtigung aller detaillierter Beobachtungen von Patienten (echte und zensierte Beobachtungen)
- individuelle Zeitintervalle für die einzelnen Patienten
- zufriedenstellende Resultate auch bei kleinen Patientenkollektiven
- relativ aufwendige und komplexe Berechnung.

Relatives Überleben oder Standardisierung von Überlebenszeiten

Diese Auswertung haben wir speziell entwickelt, um vor allem bei der Untersuchung von Kollektiven mit älteren Patienten die Überlebensraten allein aufgrund der Todesursache einer untersuchten Krankheit zu bestimmen. In den meisten Studien wird die Todesursache nicht näher spezifiziert. Neben der durch die untersuchte Krankheit hervorgerufenen Todesursache kommen auch andere, sogenannte interkurrente Todesursachen in Frage. Zum Vergleich werden die Überlebensraten einer Vergleichsgruppe aus der Bevölkerung mit der gleichen Alters- und Geschlechtsverteilung bestimmt. Als Grundlage für diese Berechnungen dienen die amtlichen Sterbetafeln (statistische Jahrbücher der Schweiz). Zu jedem Zeitpunkt wird die beobachtete Überlebensrate (Observed Survival) durch die Überlebensrate der Vergleichsgruppe (Expected Survival) dividiert. Daraus resultiert die relative Überlebensrate.

Die Altersstandardisierung hat zum Beispiel bei der Auswertung von Daten aus einem Krebsregister eine grosse Bedeutung. Vor allem bei Kollektiven mit älteren Patienten (natürliches hohes Sterberisiko) wird die relative Überlebensrate wesentlich grösser als die beobachtete. Bei kleinen Kollektiven ist die Altersstandardisierung nicht sinnvoll. Mit zunehmender Beobachtungszeit weicht das relative Überleben immer stärker von der beobachteten Überlebenszeit ab. Sind alle in einem Kollektiv aufgetretenen Todesfälle nach der Sterbetafel statistisch zu erwarten, so bleibt der Wert der relativen Überlebensrate während der ganzen Beobachtungszeit auf 100 %.

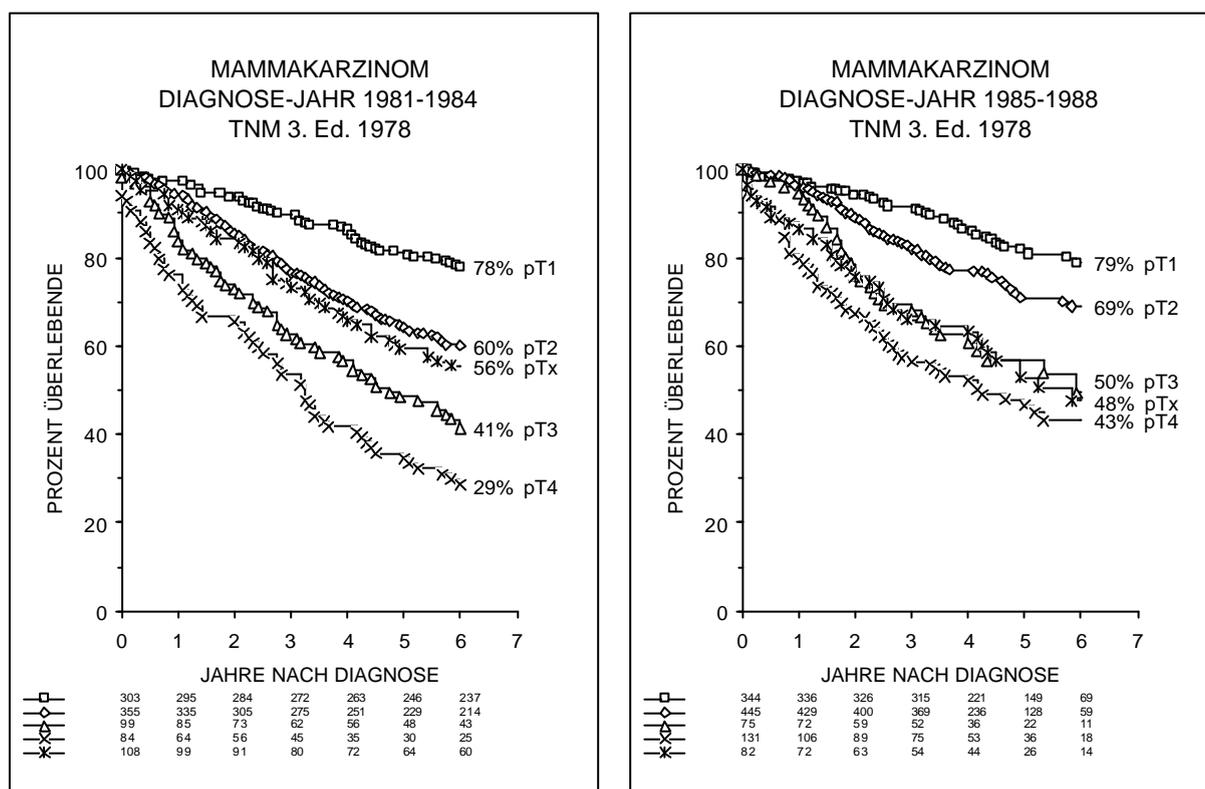


Abbildung 1: Beispiel für Überlebenskurven nach Kaplan und Meier. Vergleich von Mamma-Karzinomen in verschiedenen Stadien, diagnostiziert in den Jahren 1981 - 1984 (links) und 1985 - 1988 (rechts).

Vergleich von Überlebenskurven

Bei der Beurteilung des Therapieerfolges bei klinischen Studien spielt der Vergleich von Überlebenskurven eine entscheidende Rolle. Im einfachsten Fall werden die Ergebnisse von zwei verschiedenen Kollektiven, deren Patienten mit unterschiedlichen Therapieformen behandelt wurden, unter Zuhilfenahme von statistischen Tests miteinander verglichen. Grundsätzlich ist der gleichzeitige Vergleich mehrerer Kollektive bzw. mehrerer Therapieformen oder Stadien möglich. Überlebenszeiten folgen nicht einer Normalverteilung, so dass beim Vergleich nicht-parametrische Tests eine besondere Rolle spielen. Zum Vergleich von zwei Stichproben eignen sich: Wilcoxon-Test, Cox-Mantel-Test, Logrank-Test, Cox's F-Test und Mantel-Haenszel χ^2 -Test. Die genannten Tests sind in der Durchführung sehr aufwendig. Wir verwenden den Logrank-Test. Die Teststatistik besitzt in diesem Fall eine χ^2 -Verteilung mit einem Frei-

heitsgrad. Es wird die Wahrscheinlichkeit für die Hypothese bestimmt, dass in beiden Gruppen das gleiche Risiko besteht.

Während der Logrank-Test die Verläufe der Überlebensraten als Ganzes miteinander vergleicht, kann zu einem beliebigen Zeitpunkt der 95%-Vertrauensbereich einer Überlebensrate bestimmt werden. Dieser ist relativ einfach zu bestimmen und kann zu einer groben Beurteilung beim Vergleich von Überlebenszeiten dienen. Der 95%-Vertrauensbereich ist meistens grösser als allgemein erwartet wird und kann sich unter Umständen von 0 bis 100 % erstrecken. Wenn beim Vergleich von zwei Vertrauensbereichen der grössere den kleineren um nicht mehr als ein Viertel überdeckt, so kann auf einen signifikanten Unterschied geschlossen werden.

Die berechneten Überlebensraten beziehen sich auf eine bestimmte Gruppe von Patienten. Die Resultate sind nur dann von Bedeutung, wenn sie auf eine grössere Population verallgemeinert werden können, aus der die betrachtete Gruppe von Patienten stammt. Dies setzt genügend enge Grenzen der Stichprobenschwankung voraus. Beim umfangreichen Datenmaterial von Studien ist es oft von Interesse, die Abhängigkeit der Überlebenszeiten von verschiedenen Parametern - wie Krankheitsstadium, Behandlungsart, Alter und Geschlecht der Patienten usw. - darzustellen. Dafür bietet sich die Multivarianzanalyse an, eine nichtparametrische Regressionsanalyse nach Cox. Daraus resultiert die "Rangfolge" von prognostischen Faktoren, welche die Überlebenszeiten beeinflussen.

Erfolgreiche Zusammenarbeit mit der Radio-Onkologie

Im Jahre 1979 haben wir mit der Auswertung von Therapie-Ergebnissen begonnen, als uns der neu installierte Computer (Philips-Minicomputer) in der Radio-Onkologie jede Woche ab Freitag Nachmittag zur Verfügung gestellt wurde. Umfangreiche Programme in FORTRAN wurden erstellt für die Aufnahme der notwendigen Patientendaten, deren Auswertung und Darstellung. Mit dem Chefarzt der Radio-Onkologie, Prof. Dr. R. Hünig, entwickelte sich eine intensive und angeregte Zusammenarbeit bei der statistischen Auswertung von Therapieergebnissen von radio-onkologischen Patienten [3].

Das grösste Anwendungsgebiet für die Auswertung und Darstellung von Überlebenskurven war das Basler Protokoll für die brusterhaltende Therapie des Mammakarzinoms mit der Gruppe der Proff. R. Hünig, F.H. Harder, A.C. Almendral und J. Torhorst. Zahlreiche Zeitschriften-Publikationen und Buchbeiträge in den 80er Jahren waren ein Ergebnis davon [z. B. 2, 5, 6]. Ein weiteres grosses Gebiet für Auswertungen bildeten die Karzinom-Behandlungen im HNO-Bereich, womit sich Prof. R. Hünig, Prof. W. Wey und Dr. E. Walther beschäftigten [1, 4]. Dass damals die Radio-Onkologie Basel diesbezüglich das schweizerische Referenzzentrum bildete (wie Prof. R. Greiner in Bern im Jahre 2001 schrieb), war das Verdienst von Prof. Hünig und auf seine Initiative und seine Ideen zurückzuführen.

Weitere Anwendungen

Nachdem die Auswertung von Therapie-Ergebnissen in der Radio-Onkologie von Prof. Dr. R. Hünig, Dr. J. Kurz und Dr. E. Walther auf breiter Basis genutzt und vorgetragen wurde, gab es bald weitere Interessenten vor allem aus den Gebieten Gynäkologie, Hämatologie, Chirurgie, medizinische Onkologie, HNO, Kardiologie (schweizerische Studien über Herzklappenersatz), Basler Krebsregister usw. So durften wir beispielsweise mithelfen, praktizierende Ärzte zu Vorsorgeuntersuchungen zu motivieren, indem die verbesserten Überlebenschancen von Patienten dargestellt werden konnten.

Innerhalb von 15 Jahren wurden über 120 verschiedenen Patientenkollektive aufgebaut. Jährlich wurden zwischen 270 und 550 Überlebenskurven berechnet und dargestellt. Da die Anwendung auf dem Gebiet der statistischen Auswertung und der Überlebenskurven mit Hilfe des Computers von 1980 bis anfangs der 90er Jahre noch relativ unbekannt war, hatten wir auch zahlreiche Kunden anderer Spitäler der Schweiz und sogar aus den USA.

Im Jahre 1992 wurden bei uns die Programme auf EXCEL für PC umgeschrieben. Damit wurde die Auswertung von der Radio-Onkologie unabhängig. Auch der Datentransfer und die Auswertungen wurden vereinfacht.

Heute werden statistische Auswertungen hauptsächlich von Studienzentren durchgeführt. Es existieren auch kommerzielle Auswerteprogramme für PC. Die Nachfrage für Auswertungen von Patientenkollektiven ist deshalb stark zurückgegangen. Die Fragestellungen kommen heute noch von Interessenten bei Auswertungen für Dissertationen oder bei Vergleichen, die im wesentlichen innerhalb des KBS erfolgen.

Literatur

- [1] W. Wey, R. Hünig, W. Krauer, J. Roth, E. Walther: Präoperative Strahlentherapie von Pflasterzellkarzinomen in der Oto-Rhino-Laryngologie. In: Neue Aspekte der radiologischen Diagnostik und Therapie (Hrsg.: W. Bessler, W.A. Fuchs, J. Locher, J. Paunier), Huber-Verlag, Bern (1981), 134-142
- [2] R. Hünig, E. Walther, F. Harder, A.C. Almendral, J. Roth, J. Torhorst: The Basel lumpectomy protocol – 5 year experience with a prospective study for conservative management of breast cancer. In: Conservative Management of Breast Cancer (Eds.: J.R. Harris, S. Hellmann, W. Sillen). J.B. Lippincott Co., Philadelphia (1983), 23-33
- [3] J. Roth, R. Hünig: Life Table Analysis. In: Neue Aspekte radiologischer Diagnostik und Therapie. Jahrbuch der Schweiz. Gesellschaft für Radiologie und Nuklearmedizin. Verlag Hans Huber, Bern (1984), 338-349
- [4] J. Roth, R. Hünig: Zur "Life-Table"-Analyse nach Tumortherapie. ORL 9. Aktuelle Probleme der Otorhinolaryngologie. Verhandlungsberichte der Schweiz. Gesellschaft für Otorhinolaryngologie, Hals- und Gesichtschirurgie (Hrsg.: Kellerhals et al.), Verlag Hans Huber, Bern (1985), 103-109
- [5] R. Hünig, E. Walther, F. Harder, A.C. Almendral, J.P. Obrecht, J. Torhorst, J. Roth: Radio-Oncological Aspects. In: Breast Cancer, Conservative and Reconstructive Surgery (Ed.: H. Bohmert). Georg-Thieme Verlag, Stuttgart (1989), 95-102
- [6] E. Walther, A.C. Almendral, F. Harder, R. Hünig, J. Torhorst, J. Roth: Experience and Results. In: Breast Cancer, Conservative and Reconstructive Surgery (Ed.: H. Bohmert). Georg-Thieme Verlag, Stuttgart (1989), 102-104.