

Metodologia modelu proporcjonalnego ryzyka Cox'a

Bernard Asselain¹, Richard F. Mould²

Model proporcjonalnego ryzyka Cox'a jest jednym z najszerzej stosowanych modeli w onkologicznych publikacjach naukowych, ale także jedną z najmniej rozumianych metod statystycznych. Wynika to z łatwego dostępu do pakietów statystycznych zawierających programy do analizy przeżyć, modeli regresji i analiz wielowariantowych, ale prawie nigdy nie zawierających dobrego opisu podstawowych zasad działania modelu Cox'a. Dostarczają one wyłącznie instrukcje, jak wprowadzić dane i uruchomić odpowiednie procedury w celu uzyskania wyniku. Poniższy artykuł zawiera pełny opis metodologii modelu proporcjonalnego ryzyka Cox'a, w tym wyjaśnienie następujących pojęć: funkcja ryzyka, funkcja skumulowanego rozkładu i funkcja przeżycia.

Methodology of the Cox proportional hazards model

The Cox proportional hazards model is one of the most well used models in the field in oncology papers: but it is also one of the least well understood statistical methodologies. This is due to although being a segment of many software packages which include survival analyses, regression models and multivariate analysis programmes, such software packages almost never give a good description of the underlying principles of this Cox model. Instead, they only give instructions on how to input data and provide standard wording to give the results. This paper describes in full the methodology of the Cox proportional hazards model. This includes an explanation of the following terms: hazard function, cumulative distribution function and survivor function.

Słowa kluczowe: model proporcjonalnego ryzyka Cox'a, modele regresji, analiza wielowymiarowa, funkcja ryzyka, funkcja skumulowanego rozkładu, funkcja przeżycia, metoda największej wiarygodności, test Wald'a, ryzyko względne

Key words: Cox proportional hazards model, regression modelling, multivariate analysis, hazard function, cumulative distribution function, survivor function, maximum likelihood, Wald test, relative risk

Wprowadzenie

Prosty model regresji zakłada, że wynik (np. przeżycie lub odpowiedź na leczenie onkologiczne) może być wyjaśniony zasadniczo w oparciu o jedną zmienną objaśniającą lub predykcyjną. Podstawą tego założenia jest stwierdzenie, że wiele innych czynników, mogących mieć wpływ na wynik, ma osobno niewielkie znaczenie, a ich sumaryczny wpływ jest równy zero. To założenie jest bardzo rzadko spełnione dla czynników prognostycznych, mogących mieć wpływ na przeżycia onkologiczne i z tego powodu konieczne jest zastosowanie złożonych modeli regresji.

Cox [1] opracował tego typu model dla tabeli przeżyć (zobacz także pracę Breslow'a [2]), który umożliwia

analizę wpływu czynników prognostycznych na przeżycie. Autor zilustrował zastosowanie modelu dla przypadku leukemii, ale model może być stosowany do obliczania przeżyć w odniesieniu do wszystkich innych chorób, jak w przypadku przeżyć w chorobach nowotworowych lub kardiologicznych po transplantacji serca lub zawałach serca [3].

W wyniku analizy wielowymiarowej otrzymujemy zbiór czynników z określoną miarą wpływu na wynik dla każdego czynnika osobno, przy założeniu stałości pozostałych.

Model Cox'a jest bardzo często stosowany i jest oferowany przez większość komercyjnych pakietów statystycznych, chociaż na ogół bez szczegółowego wyjaśnienia jego funkcjonowania.

Pomimo, że model ten jest najbardziej rozpowszechniony w onkologii, istnieją w tej dziedzinie inne modele wielowymiarowe, jak np.: model log-logitowy [4, 5], czy liniowy model logitowy [6]. W dalszej części zostały one zastosowane do analizy *case-control* [7].

¹ Unité de Biostatistique
Institut Curie
Paris
France

² Cartmel
Cumbria
United Kingdom

Funkcja ryzyka (hazardu), skumulowana funkcja rozkładu i funkcja przeżycia

Funkcja ryzyka nazywana jest również funkcją hazardu, intensywnością umieralności (*force of mortality*), umieralnością chwilową (*instantaneous death rate*) lub chwilową częstością niepowodzeń (awarii) (*failure rate*). Ostatniego określenia używa się w teorii odnowy [8], w której analizuje się awaryjność elementów przemysłowych. Wyrażona językiem prawdopodobieństwa, funkcja ryzyka $\lambda(t)$ jest stosunkiem funkcji gęstości prawdopodobieństwa $f(t)$ do funkcji przeżycia $(1 - F(t))$, gdzie $F(t)$ jest funkcją rozkładu.

$$\lambda(t) = f(t)/[1 - F(t)]$$

Funkcja rozkładu lub skumulowana funkcja rozkładu, $F(t)$, rośnie wraz z czasem t i przyjmuje wartość 1 dla maksymalnej wartości t . Funkcja przeżycia $(1 - F(t))$ wyraża prawdopodobieństwo, że zmienna czasu przyjmie wartość mniejszą lub równą t . Funkcja gęstości prawdopodobieństwa $f(t)$ jest funkcją, której całka po czasie t w przedziale $[t_1, t_2]$ jest równa prawdopodobieństwu, że zmienna t przyjmie wartość z tego przedziału. Ryciny 1-3 schematycznie ilustrują funkcje $f(t)$, $F(t)$ i $1 - F(t)$, dla $f(t)$ będącej funkcją wykładniczą:

$$f(t) = \lambda \exp[-\lambda t]$$

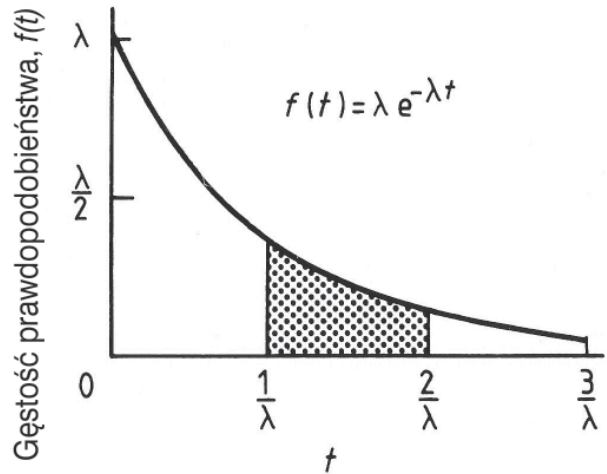
Widać, że dla tego konkretnego rozkładu funkcja ryzyka jest stała i równa λ . Model wykładniczy, zastosowany do analizy przeżyć w chorobach nowotworowych, oznacza stałe ryzyko zgonu, niezależne od czasu, jaki upłynął od zakończenia leczenia.

Hastings i Peacock podali przykłady rozkładów gęstości funkcji przeżycia i ryzyka dla wielu rodzin funkcji [9], a graficzne metody weryfikacji funkcji ryzyka zostały przedyskutowane przez Nelson'a [10]. Natomiast rozkłady funkcji przeżycia i metody estymacji parametrów rozkładu zostały opisane przez Gross'a i Clark'a [11] oraz Kalbfleisch'a i Prentice'a [4].

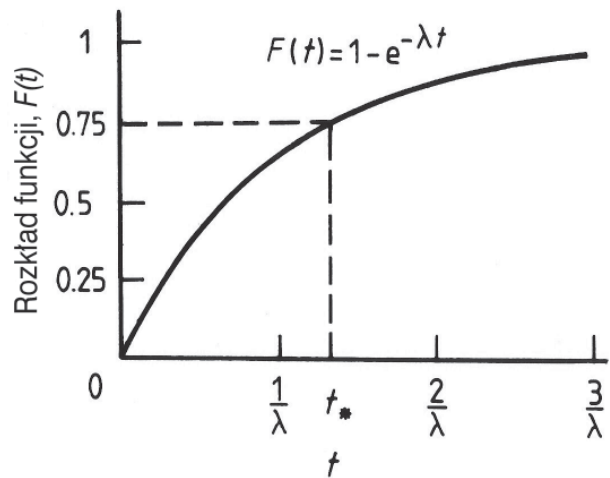
Założenia modelu proporcjonalnego ryzyka Cox'a

Model Cox'a zakłada, że względne ryzyko zgonu w dwóch podgrupach jest stałe w czasie (założenie proporcjonalności ryzyk) oraz, że ryzyko względne związane z jednym parametrem jest stałe i niezależne od wartości pozostałych parametrów (poza przypadkiem wprowadzenia do modelu efektu interakcji). To ostatnie założenie jest wspólne dla wszystkich modeli regresji, nie tylko dla modelu Cox'a. Dodatkowo model zakłada, że wpływ zmiennych niezależnych na funkcję ryzyka ma charakter log-linearny. Cox [1, 12] zaproponował model proporcjonalnego ryzyka w postaci:

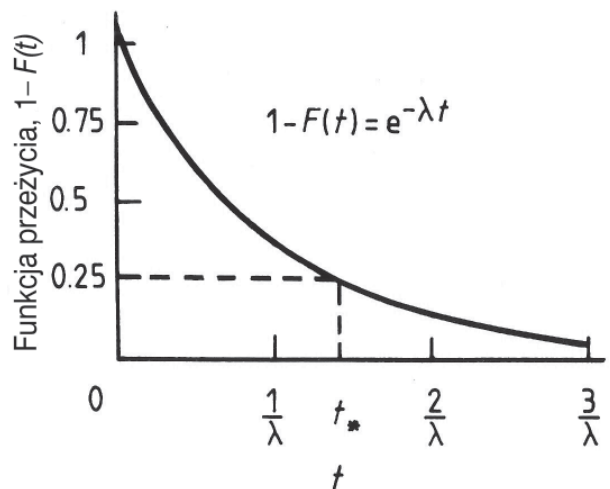
$$\lambda(t, Z) = \lambda_0(t) \exp[Z\beta]$$



Ryc. 1. Jako że funkcja ryzyka $\lambda(t)$ jest stosunkiem funkcji gęstości prawdopodobieństwa $f(t)$ do funkcji przeżycia $(1 - F(t))$, to dla rozkładu wykładniczego $\lambda(t) = \lambda$. Prawdopodobieństwo, że t należy do przedziału $(1/\lambda, 2/\lambda)$, jest równe polu pod wykresem krzywej w zaznaczonych granicach



Ryc. 2. Funkcja gęstości prawdopodobieństwa $f(t)$ jest pierwszą pochodną funkcji rozkładu prawdopodobieństwa $F(t)$.
 $f(t) = d/dt[F(t)] \Rightarrow d/dt\{1 - \exp[-\lambda t]\} = \lambda \exp[-\lambda t]$
 Jak pokazano na wykresie, prawdopodobieństwo, że t znajduje się w przedziale $(0, t_*)$, jest równe 0,75



Ryc. 3. Prawdopodobieństwo, że wartość t przekroczy t_* , jest równe 0,25. Jeśli funkcję przeżycia wyrazimy w procentach, a czas t w latach, wówczas t_* -letni odsetek przeżyć będzie równy 25%

gdzie $\lambda(t, Z)$ oznacza funkcję ryzyka dla czasu t i wektora zmiennej niezależnej Z , to jest:

$$\lambda(t, Z) = \lambda_0(t) \exp[\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p]$$

gdzie β jest wektorem nieznanych stałych parametrów (wag), a $\lambda_0(t)$ jest nieznaną funkcją ryzyka dla Z równego zero (tzw. ryzyko bazowe).

Sformułowany model ryzyka nie jest wyrażony w kategoriach rozkładu, jak np. rozkład Weibull'a, ale jest określony przez rozkład analizowanych danych.

Dla danego Z , $\exp[\beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p]$ jest wartością stałą, a więc funkcja ryzyka $\lambda(t, Z)$ jest funkcją bazową $\lambda_0(t)$, pomnożoną przez stałą. W konsekwencji Cox wprowadził określenie: model proporcjonalnego ryzyka. Trzeba jednak pamiętać, że ograniczeniem modelu Cox'a jest konieczność dysponowania wartościami Z dla każdego pacjenta. Istnieją wprawdzie metody szacowania wartości brakujących danych, ale wymagają one spełnienia dodatkowych założeń, które na ogół trudno jest uzasadnić [1].

Funkcja ryzyka dla trzech czynników prognostycznych

Dla zilustrowania działania modelu założymy, że mamy trzy czynniki prognostyczne $Z_1 =$ wiek pacjenta – 50 (w latach), $Z_2 =$ płeć i $Z_3 =$ zaawansowanie choroby. Z_2 może przyjmować wartości 1 dla mężczyzn lub 0 dla kobiet, a Z_3 jest zmienną porządkową o wartościach całkowitych nieujemnych. Funkcja $\lambda_0(t)$ przedstawia ryzyko dla referencyjnego pacjenta w wieku 50 lat ($Z_1 = 0$), płci żeńskiej ($Z_2 = 0$) i w najniższym stopniu zaawansowania o kodzie „0” ($Z_3 = 0$).

Funkcja $\lambda_0(t)$ przedstawia *intensywność umieralności (force of mortality)* dla wszystkich pacjentów, dla których zmienna zależna Z jest równa zero. Gdyby pacjent miał 55 lat zamiast 50, był mężczyzną zamiast kobietą i był w stopniu zaawansowania odpowiadającym kodowi 1 zamiast 0; oraz gdyby oszacowane wagi β dla danych czynników prognostycznych wyniosły odpowiednio $\beta_1 = 0,1$; $\beta_2 = 0,3$ i $\beta_3 = 0,3$, to funkcja ryzyka byłaby równa:

$$\lambda_0(t) \exp[0,1 \times 5 + 0,3 \times 1 + 0,3 \times 1] = \lambda_0(t) \exp[1,1] = 3,0 \times \lambda_0(t)$$

To znaczy, że odpowiednie ryzyko byłoby 3,0 razy większe niż dla referencyjnego pacjenta i ten stosunek ryzyk byłby stały w czasie.

Złagodzenie założenia o proporcjonalności ryzyka

Przykład leukemii, który wykorzystał Cox [1] jako ilustrację modelu, podobnie jak przykład nieoperacyjnego raka piersi w III stopniu zaawansowania [13], nie są szczególnie dobre, ale założenie o proporcjonalności ryzyka jest do zaakceptowania. Jednakże w przypadku danych

onkologicznych ze znaczącą liczbą długoletnich przeżyć, założenie stałości ryzyka względnego w czasie może już nie być rozsądne. Taka sytuację zaobserwowali Gore & Pocock [14] dla chorych na raka piersi, których wyniki odległe opublikował Langlands i inni [15]. Stwierdzili oni, że dla tych danych założenie proporcjonalności ryzyka nie było spełnione. Ponadto w komentarzu Gore zauważył, że stopień zaawansowania, który ma początkowo znaczenie, traci je w czasie i po 10 latach obserwacji roczna śmiertelność jest już od niego niezależna.

Problem niespełnienia założenia proporcjonalności można próbować rozwiązać na różne sposoby. Można, na przykład, uwzględnić w modelu czynniki prognostyczne zależne od czasu, wtedy:

$$\lambda(t, Z(t)) = \lambda_0(t) \exp[Z(t)\beta]$$

Iloczyn $\exp[Z(t)\beta]$ będąc funkcją czasu, nie jest już stały i współczynnik proporcjonalności także się zmienia. To prowadzi do mniej eleganckiego, ale być może bardziej realistycznego spojrzenia na historię choroby, w której wpływ czynników prognostycznych zmienia się wraz z czasem obserwacji.

Przykład: nawrót lokoregionalny w raku piersi

W tej części przedstawiono, zaczerpnięty z literatury [16], typowy przykład wyników wielowymiarowej analizy z zastosowaniem modelu Cox'a. Dotyczy lokoregionalnego nawrotu w raku piersi i jest to typowa publikacja, w której jedyną referencją jest oryginalny artykuł Cox'a z 1972 r. [1]. Inną referencją wydaje się być informacja o pakiecie statystycznym, którym się posłużono. Naukowe publikacje onkologiczne rzadko, o ile w ogóle, zawierają opis metodologii modelu Cox'a, a zainteresowanego czytelnika odsyłają do raczej niezbyt przystępnych publikacji statystycznych. Niniejszy artykuł dla *Nowotworów* został napisany w celu usunięcia tajemniczości otaczającej model Cox'a.

W tym przykładzie [16], jak we wszystkich podobnych analizach, przeprowadzono najpierw analizę jednowymiarową, a następnie czynniki, które okazały się statystycznie istotne na wybranym poziomie istotności (który niekoniecznie musi być równy 0,05), zostały włączone do analizy wielowymiarowej. Jednakże zdarza się nagminnie, że kiedy nawet dany czynnik w analizie jednowymiarowej jest nieistotny, jest on włączany do analizy wielowymiarowej.

W literaturze można znaleźć wiele przykładów zastosowań modelu Cox'a, ale analizy czynników prognostycznych dla nawrotów lokoregionalnych w raku piersi cieszą się największym zainteresowaniem. Zidentyfikowanie podgrupy pacjentów z korzystną prognozą jest bardzo ważne, ze względu na wysoką zachorowalność, a także ze względu na złą prognozę pacjentów po nawrocie lokoregionalnym po mastektomii.

Tabela I z publikacji Willner'a i innych [16] wyraźnie ukazuje, że niektóre czynniki prognostyczne, które były istotne w analizie jednowymiarowej, nie są istotne w ana-

Tab. I. Wybrane wyniki z publikacji Willner'a i in. [16], przedstawiające jednowymiarową i wielowymiarową analizę czynników prognostycznych, wpływających na przeżycie po nawrocie. NI = nieistotny, pT – stopień pierwotnego zaawansowania T, PD = pierwotna diagnoza, DN = diagnoza nawrotu, * oznacza wykluczenie z analizy wielowymiarowej z powodu dużej liczby braków danych

Czynniki prognostyczne	Analiza jedno-wymiarowa	Analiza wielo-wymiarowa
pT (T1,2/3,4)	P < 0,001	P < 0,01
Zróżnicowanie G (G1,2/3,4)	P < 0,01	*
Zajęcie naczyń krwionośnych	P < 0,01	NI
Martwica guza	P < 0,001	P < 0,01
Wiek PD (< / >50 lat)	NI	NI
Status węzłów pachowych PD	P < 0,001	P < 0,05
Chemioterapia po mastektomii	P < 0,01	NI
Lokalizacja nawrotu	P < 0,01	P < 0,001
Klatka piersiowa		
Węzły pachowe		
Nadobojczyk		
Wielokrotna		
Czas do nawrotu (< / >12 miesięcy)	P < 0,001	P < 0,01
Wiek DN (< / >50 lat)	NI	P < 0,05
Czas kontroli lokalizacji nawrotu	P < 0,001	P < 0,05

lizie wielowymiarowej. Taka sytuacja jest typowa i jasno wskazuje, że analiza jednowymiarowa nie jest adekwatna do badania wpływu grupy potencjalnych czynników.

Konkluzja Willner'a i innych [16] była następująca: „Najlepiej rokują pacjenci w wieku > 50 lat z pojedynczym nawrotem w klatce piersiowej lub węzle pachowym, z czasem wolnym od nawrotu nie mniejszym niż 12 miesięcy, z pierwotnym rakiem w stanie zaawansowania pT1-2pN0, bez martwicy guza i z miejscowo kontrolowanym nawrotem“.

Jest również możliwe, że czynnik nieistotny w analizie jednowymiarowej może być istotny w analizie wielowymiarowej, jak w przypadku wieku pacjenta DN (< / > 50 lat) w Tabeli I. Chociaż w większości badań czynniki nieistotne w analizie jednowymiarowej byłyby odrzucone. Taka sytuacja nie jest korzystna i można by zaproponować podwyższenie progu włączenia do analizy wielowariantowej do wartości 0,20. Jeśli istotność danego czynnika, która była wyraźna w analizie jednowymiarowej, znika w analizie wielowariantowej, oznacza to, że (a) powinien on być uważany za nieistotny i (b) jest on naj-

prawdopodobniej czynnikiem zastępczym innej zmiennej, to znaczy silnie z nią skorelowanym.

Dwie główne prace dotyczące nawrotu lokoregionalnego w raku piersi, opublikowane przed analizą Willner'a i wsp. [16] z 1997 r., na danych pochodzących z Uniwersytetu w Würzburgu, to praca Halverson'a i wsp. [17] z 1992 r. z Instytutu Radiologii Mallinckrodt oraz St. Louis'a i Schwaibold'a i wsp. [18] z 1991 r., z Centrum Onkologii Fox Chase w Filadelfii.

Wszystkie są przykładem analizy wielowymiarowej, ale praca Willner'a jest pracą przeglądową najbardziej wyczerpującą. We wszystkich trzech publikacjach [16-18], z powodu braku danych, wyłączono z analizy wielowymiarowej grupę czynników, pomimo ich istotności w analizie jednowymiarowej. Jest warte szczególnego podkreślenia, że analiza wielowariantowa wymaga odpowiednio wysokiej jakości danych.

Następujący cytat pracy z Willner'a i wsp. [16] daje dobrą perspektywę dla uzasadnienia stosowania analizy wielowymiarowej w onkologii: „Złożony charakter wzajemnych zależności czynników prognostycznych podkreśla znaczenie wielowymiarowej analizy przeżycia dla wszystkich dostępnych czynników prognostycznych i metod leczenia.“

Należy jednak podkreślić, że analizie wielowymiarowej, z powodu większych braków danych, powinno się z dużą ostrożnością przyznawać pierwszeństwo przed analizą jednowymiarową. Aczkolwiek, różnice pomiędzy obu analizami dają pewne informacje o korelacjach pomiędzy różnymi zmiennymi, czasem nieoczekiwanych korelacjach.

Przykład modelu Cox'a dla trzech zmiennych

Praca Christensen'a [19], zamieszczona w czasopiśmie *Hepatology*, zawiera wciąż (20 lat po jej publikacji) jeden z najpełniejszych opisów metodologii modelu Cox'a i dlatego tę pracę przedstawiono tu jako przykład praktyczny, wraz z zawartą w niej dyskusją.

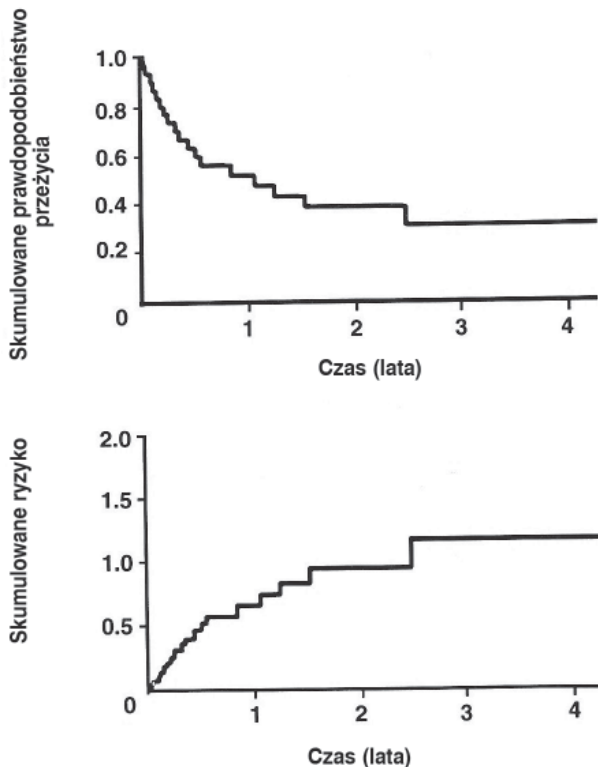
Grupa pacjentów Christensen'a została skonstruowana teoretycznie. Stanowi ją w sumie 30 przypadków z jednym kryterium oceny (*endpoint*) oraz trzema zmiennymi niezależnymi (Tab. II).

Związek pomiędzy prawdopodobieństwem przeżycia a ryzykiem zgonu

Związek ten był już w pewnym zakresie dyskutowany wcześniej, a dalsza część znajduje się na Rycinie 4.

Tabela II. Hipotetyczne dane użyte przez Christensen'a [19] dla zilustrowania metodologii modelu Cox'a. Wartości kryterium oceny (*endpoint*): 1 – zgon, 0 – żyje lub stracony z obserwacji (informacja cenzurowana)

Zmienne niezależne (czynniki prognostyczne)	Nazwa/Kodowanie
Białko w surowicy (mg/l) mierzone w skali ciągłej (21-36)	ALB
Bilirubina mierzona (m mole/litr) w skali ciągłej (19-332) i włączona do modelu Cox'a jako \log_{10} bilirubina	LGB
Alkoholizm	ALC
	Obecny (1), Nieobecny (0)



Ryc. 4. Jeżeli prawdopodobieństwo przeżycia w funkcji czasu t oznaczymy przez $S(t)$, a funkcję skumulowanego ryzyka przez $\Lambda(t)$, wtedy zależność pomiędzy obiema funkcjami można zapisać jako: $\Lambda(t) = -\log_e S(t)$ lub równoważnie: $S(t) = e^{-\Lambda(t)}$. Funkcja przeżycia $[1 - F(t)]$ na Rycinie 3. jest tą samą funkcją co $S(t)$ na Rycinie 4. Jednym z problemów w rozumieniu modelu Cox'a jest wielość symboli, które nie zawsze są spójne w różnych publikacjach

Obliczenia zastały wykonane przy użyciu metody tablic przeżycia Kaplan'a-Meier'a dla 30 przypadków z grupy Christensen'a. Rycina pokazuje *skumulowane prawdopodobieństwo przeżycia* (określenie Christensen, przymiotnik *skumulowany* nie jest standardową terminologią dla wykresów tablic przeżycia) oraz ilustruje, jak krzywa przeżycia zależy od *skumulowanego ryzyka*.

Test ilorazu wiarygodności

Model Cox'a dla danych Christensen'a zawierających trzy zmienne: białko ALB, \log_{10} bilirubiny LGB i alkoholizm ALC, przyjmuje postać:

$$\lambda(t, Z) = \lambda_0(t) \exp[\beta_1 Z_1 + \beta_2 Z_2 + \beta_p Z_3]$$

gdzie: $Z_1 = \text{ALB}$, $Z_2 = \text{LGB}$ i $Z_3 = \text{ALC}$

W procesie dopasowania modelu można rozważać w sumie siedem możliwości (Tab. III). Trzy modele zawierające po jednej zmiennej, trzy zawierające po dwie zmienne i jeden zawierający wszystkie trzy zmienne, co w sumie daje siedem możliwych modeli.

Szacowanie parametrów i testowanie istotności danego modelu wykorzystuje koncepcję *wiarygodności* (*likelihood*), która określa prawdopodobieństwo wyjaśnienia przez model zaobserwowanych danych. Określenie całkowitej istotności modelu opiera się na stosunku

funkcji wiarygodności $L(0)$ dla modelu, w którym wpływ czynników na czas przeżycia jest zerowy (w tym przykładzie $\beta_1 = \beta_2 = \beta_3 = 0$) i funkcji wiarygodności $L(b)$ dla modelu z nieznanymi współczynnikami b . Te współczynniki są szacowane w taki sposób, aby funkcja $L(b)$ miała wartość maksymalną. Takie szacowanie parametrów (funkcję bazową i współczynniki b) nazywa się szacowaniem metodą największej wiarygodności.

Im większa wartość $L(b)$ lub mniejsza wartość *ilorazu wiarygodności* $L(0)/L(b)$, tym lepiej model wyjaśnia zaobserwowane dane [4].

Istotność każdego modelu może być testowana statystycznie przy użyciu statystyki o rozkładzie χ^2 podanej poniżej, gdzie liczba stopni swobody jest równa liczbie czynników w modelu [4]. Wyniki Christensen'a [19] dla siedmiu modeli Cox'a, które są typowym spektrum danych uzyskiwanych w wydrukach komputerowych dla modelu Cox'a, chociaż nie zawsze są prezentowane w takim układzie, podano w Tabeli III.

$$\begin{aligned} \chi^2 &= -2 \times \log_e [L(0)/L(b)] = -2 \times [\log_e L(0) - \log_e L(b)] = \\ &= 2 \times [\log_e L(b) - \log_e L(0)] \end{aligned}$$

Standardowe odchylenie normalne Z , test Wald'a i względna istotność zmiennych

W Tabeli III, Z oznacza *standardowe odchylenie normalne*, które równe jest $b/SE(b)$. Istotność współczynników regresji b może być szacowana przez porównanie wartości Z^2 z rozkładem χ^2 dla $DF = 1$, co czasem jest nazywane *testem Wald'a* [19]. Jeżeli $Z > 1,96$, to b jest statystycznie różne od zera na poziomie istotności $\alpha = 0,05$. Jednakże, jeśli $Z < 1,96$ dla jakiegoś współczynnika b , to nie musi to oznaczać, że dany czynnik nie ma wpływu na prognozę. Może też oznaczać, że jego efekt jest zbyt mały, aby być wykryty w badaniu z daną liczbą pacjentów. *Względna istotność zmiennych* określa numeryczna wielkość Z . Im większa wartość Z , tym większa istotność zmiennej w modelu. Jak pokazuje Tabela III dla modelu 7, istotność zmiennych spada w następującej kolejności: ALB, LGB, ALC, gdzie ALC jest statystycznie nieistotne.

Ryzyko względne: iloraz hazardów

Korzystając z wartości współczynników regresji b , można oszacować *ryzyko względne*, np. zachorowalności w populacji narażonej do zachorowalności w populacji nie-narażonej.

Ryzyko względne jest użyteczną statystyką, na przykład kiedy analizujemy ryzyko zachorowania na leukemii czy powstawania nowotworów litych w populacji, która przeżyła wybuch bomby atomowej w Hiroszimie i Nagasaki, w zależności od wieku napromienienia i otrzymanej dawki.

Wartość ryzyka względnego równa 1,1 oznacza wzrost ryzyka o 10%, w porównaniu z ryzykiem dla grupy odniesienia. W analizie Cox'a ryzyko względne to ilorazy hazardów przypisane różnym poziomom danego

Tab. III. Wyniki dla siedmiu modeli regresji Cox'a, podane przez Christensen'a [19]

Numer Modelu	χ^2	DF	P	Włączone czynniki	b	SE(b)	Z	P***
1	30,99	1	< 0,0001	ALB	-0,42	0,089	-4,71	< 0,0001
2	21,24	1	< 0,0001	LGB	4,44	1,06	4,17	< 0,0001
3	8,79	1	0,003	ALC	1,55	0,55	2,82	0,005
4	35,89	2	< 0,0001	ALB LGB	-0,35 2,36	0,10 1,11	-3,43 2,12	0,0006 0,03
5	32,50	2	< 0,0001	ALB ALC	-0,9 0,79	0,094 0,64	-4,16 1,23	< 0,0001 0,22
6	25,13	2	< 0,0001	LGB ALC	3,88 1,14	1,06 0,59	3,66 1,93	0,0002 0,056
7	37,04	3	< 0,0001	ALB LGB ALC	-0,32 2,25 0,71	0,11 1,11 0,66	-3,07 2,03 1,08	0,002 0,04 0,28

Oznaczenia

χ^2 – wartość obliczona na podstawie podanej formuły

$\log_e L(0)$ – wartość taka sama dla wszystkich modeli, np. -52,319

$L(b)$ i χ^2 – wartości zależne od włączonych czynników. Na przykład dla modelu 1

$\log_e L(b) = -36,825$, a $\chi^2 = 2(-36,825 - (-52,319)) = 30,99$ (patrz druga kolumna)

DF – liczba stopni swobody (*degrees of freedom*)

P – wartość krytyczna testu (wartość-p) χ^2 i P*** wartość-p dla współczynnika b

Analizowane czynniki: ALB – białko, LGB – \log_{10} bilirubiny i ALC – alkoholizm

b – współczynnik regresji, SE(b) błąd standardowy b

czynnika, podczas gdy wszystkie inne czynniki pozostają niezmiennione.

Rozważmy Tabelę III, model 3 ze zmienną ALC. Ryzyko względne (RR) ALC do nie-ALC można obliczyć z równania:

$$RR = e^{1,55x1}/e^{1,55x0} = 4,7/1 = 4,7$$

Dla tej jednej zmiennej (ALC) mamy $b=1,55$. Dla modelu 4 zmniejszenie poziomu białka w surowicy (ALB) o 1 mg, np. z 30 mg/l do 29 mg/l, jest związane ze wzrostem ryzyka względnego o wartość:

$$RR = e^{-0,35x29}/e^{-0,35x(-1)} = 1,42$$

ponieważ w tym wypadku $b=0,35$. Ryzyko względne związane ze spadkiem poziomu ALB o 1 mg/l jest niezależne od poziomu ALB.

Krokowa selekcja zmiennych

Istnieją dwa typy selekcji zmiennych: krokowa – postępująca (*forward*) lub wsteczna (*Backward*) [19], zwana też *krokową eliminacją zmiennych*.

W metodzie krokowej postępującej (*forward stepwise method*) model jest budowany krokowo, na drodze kolejnego włączania zmiennych, dla których spadek ilorazu wiarygodności jest największy lub równoważnie największy jest wzrost statystyki χ^2 . I tak, w pierwszym kroku, zostałaby włączona zmienna ALB (patrz model 1), ponieważ daje najwyższy, statystycznie istotny wzrost χ^2 ze wszystkich modeli z jedną zmienną (modele 1-3).

W następnym kroku zostałaby dodana zmienna LGB (model 4), ponieważ ona istotnie zwiększa wartość χ^2

$$(35,89 - 30,99 = 4,90 \text{ dla } DF=1, p < 0,05)$$

w przeciwieństwie do ALC (model 5), gdzie wzrost χ^2 jest nieistotny.

$$(32,50 - 30,99 = 1,51 \text{ dla } DF=1, p > 0,20)$$

Liczba *stopni swobody* (DF) jest obliczana jako różnica liczby szacowanych współczynników w porównywanych modelach.

Włączenie ALC do modelu zawierającego ALB i LGB (model 7) nie powoduje statystycznie istotnego wzrostu statystyki χ^2

$$(37,05 - 35,89 = 1,15 \text{ dla } DF=1, p > 0,20)$$

i dlatego, przy zastosowaniu metody selekcji krokowej, model 4 byłby modelem końcowym. Stosując metodę *krokowej eliminacji zmiennych*, zaczynamy od modelu zawierającego wszystkie zmienne. Następnie statystycznie nieistotne zmienne są kolejno usuwane poprzez wyłączenie w każdym kroku zmiennej najmniej istotnej. I tak, moglibyśmy rozpocząć od modelu 7, następnie usunąć ALC, ponieważ zmienna ta jest statystycznie nieistotna. To doprowadziłoby nas do modelu 4, który byłby modelem końcowym, ponieważ obie zmienne ALB i LGB są statystycznie istotne.

Kiedy analiza nie jest zbyt skomplikowana, jak w przypadku danych Christensen'a z trzema czynnikami,

obie metody selekcji – postępującej i wstecznej – prowadzą do jednego modelu końcowego. Jednak w przypadku bardziej złożonej analizy, z wieloma zmiennymi, metody te mogą prowadzić do różnych modeli.

Bernard Asselain MD, PhD

Unité de Biostatistique
Institut Curie
26, rue d'Ulm
75248 Paris, Cedex 05
France
e-mail: bernard.asselain@curie.net

Richard F. Mould MSc PhD

4 Town End Meadow
Cartmel
Grange-over-Sands
Cumbria LA11 6QG
United Kingdom

Piśmiennictwo

1. Cox DR. Regression models and life tables. *J Royal Stat Soc* 1972; 34B: 187-220.
2. Breslow NE. Analysis of survival data under the proportional hazards model. *Int Stat Rev* 1975; 44: 45-58.
3. Norwegian Multicentre Study Group. Timolol-induced reduction in mortality and reinfarction. *N Eng J Med* 1981; 304: 801-7.
4. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. New York: John Wiley, 1980.
5. O'Quigley J. Regression models and survival prediction. *Statistician* 1982; 31: 107-16.
6. Breslow NE, Day NE, Halvorsen KT i wsp. Estimation of multiple relative risk functions in matched case-control studies. *Amer J Epidemiol* 1978; 108: 299-307.
7. Breslow NE, Day NE (red.). *Statistical Methods in Cancer Research. 1. The Analysis of Case-Control Studies*. IARC Scientific Publications Nr. 32. Lyon: International Agency for Research on Cancer, 1980.
8. Cox DR. *Renewal Theory*. Methuen Monograph on Applied Probability & Statistics. London: Methuen, 1962.
9. Hastings NA, Peacock JB. *Statistical Distributions*. London: Butterworths, 1975.
10. Nelson W. Theory and application of hazard plotting for censored failure data. *Technometrics* 1972; 14: 945-66.
11. Gross AJ, Clark VA. *Survival Distributions: Reliability Applications in the Biomedical Sciences*. New York: John Wiley, 1975.
12. Mould RF. *Cancer Statistics*. Bristol: Adam Hilger; 1983: 170-5.
13. Rubens RD, Armitage P, Winter PJ i wsp. Prognosis in inoperable stage III carcinoma of the breast. *Eur J Cancer* 1977; 13: 805-11.
14. Gore SM, Pocock SJ. The statistical modelling of survival in breast cancer. Lecture at the Institute of Statisticians Conference, *Statistics in Medicine*, Cambridge 1981.
15. Langlands AO, Pocock SJ, Kerr GR i wsp. Long-term survival of patients with breast cancer: a study of the curability of the disease. *Br Med J* 1979; 2: 1247-51.
16. Willner J, Kiricuta IC, Kolbl O. Locoregional recurrence of breast cancer following mastectomy: always a fatal event? *Int J Radiat Oncol Biol Phys* 1997; 37: 853-63.
17. Halverson KJ, Perez CA, Kuske RR i wsp. Survival following locoregional recurrence of breast cancer: univariate analysis & multivariate analysis. *Int J Radiat Oncol Biol Phys* 1992; 23: 285-91.
18. Schwaibold F, Fowble BL, Solin LJ i wsp. The results of radiation therapy for isolated local regional recurrence after mastectomy. *Int J Radiat Oncol Biol Phys* 1991; 21: 299-319.
19. Christensen E. Multivariate survival analysis using Cox's regression model. *Hepatology* 1987; 7: 1346-58.

Otrzymano i przyjęto do druku: 9 września 2009 r.