

## Cut-off skóre věku při chirurgickém zákroku karcinomu prsu

To, že věk ovlivňuje zdraví, ví prakticky každý. Jinak zdraví pociťuje 10-letý, 30-letý a 50-letý člověk. Kromě selského rozumu je však třeba použít vědy ke zjištění odpovědí na dvě zásadní otázky:

1. Co vše ovlivňuje věk?
2. Jak moc to věk ovlivňuje?

Měřeny byly praktické veškeré oblasti lidského těla, jako příklad lze uvést například artrózu (Meenan, Gertman, & Mason, 1980) či diverzita b-buněk (Gibson et. al., 2009). Vzhledem k povaze dat se budu zabývat souvislostí mezi věkem a přežitím chirurgického zákroku u rakoviny prsu.

Výzkum na toto téma dělal například Johnson et al. (2017). Podle něj je věk jedním z klíčových faktorů z hlediska prognózy. Mladší věk tedy podle něj znamenal výrazně větší šanci na přežití zákroku. Na veřejně dostupných datech s 307 respondenty (Haberman's Survivor, Kaggle) bude provedena ROC analýza. Ta poskytne dva důležité údaje. Jednak díky ní zjistíme, od kterého věku dává smysl považovat zákrok za rizikovější s ohledem na věk. Dále se pomocí AUC analýzy dozvíme, jak silný je tento efekt.

Použity byly následující proměnné:

- Data, jestli žena přežila chirurgický zákrok (1 = přežila, 0 = nepřežila)
- Věk, který může fungovat jako kontinuální proměnná místo typicky používaného hrubého skóru z nějakého dotazníku

Jako možnosti byly stanoveny veškeré věkové hodnoty, které se objevily v souboru. Tento rozsah byl od 31 do 84 let. U každého věku bylo vypočteno několik ukazatelů, které pomáhají dobrat se ke správnému výsledku.

1. Prvním z nich je počet TP (true positive). To znamená, kolik jedinců by tento cutoff skór označil jako někoho, kdo by zákrok přežil a zároveň tento člověk ve skutečnosti přežil.
2. Počet FP (false positive) zase říká, kolik lidí by tento cut-off skór určil jako přeživšího, přestože tato osoba zákrok nepřežila.
3. Počet TN (true negative) říká, kolik lidí by tento cut-off skór označil jako nepřeživší, přičemž reálně zemřeli.
4. Počet FN (false negative) říká, kolik lidí by tento cut-off skór označil jako ty, kteří nepřežili, přestože reálně přežili.
5. Senzitivita je poměr TP ku všem, kteří byli na operaci. Říká tedy s jakou pravděpodobností toto cut-off skóre určí člověka, co přežil.
6. Specificita je poměr TN ku všem, co zemřeli. Určuje, s jakou pravděpodobností toto cut-off skóre určí člověka, který zákrok nepřežil.
7. Díky senzitivitě a specificitě je možné vypočítat pro každé možné cut-off skóre Youdenovu statistiku (J). Čím vyšší je její hodnota, tím vhodnější cut-off skóre je.

Youdenova statistika funguje nejlépe v případě, že skupiny 0 a 1 jsou zastoupeny rovnoměrně. V tomto případě by to znamenalo, že v souboru by muselo být stejně těch, co přežili a zemřeli. Proto je v tabulce u každého skóre ještě statistika I, která bere v úvahu poměr těch, kteří přežili a zemřeli. Stejně jako u Youdenovy statistiky, vyšší číslo indikuje větší vhodnost cut-off skóre.

Srovnání všech možných cut-off skórů se nachází v tabulce 1.

Data a další informace o této zprávě jsou dostupné na adrese [https://dostal.vyzkum-  
psychologie.cz/stat4?i=171](https://dostal.vyzkum-psychologie.cz/stat4?i=171).

HS	TP	FP	TN	FN	Senzitivita	Specifita	1-senz	1-se-pec	J	Poměrová
30	225	81	0	0	100 %	0 %	0 %	100 %	0 %	74 %
31	222	81	0	3	99 %	0 %	1 %	100 %	-1 %	73 %
32	220	81	0	5	98 %	0 %	2 %	100 %	-2 %	72 %
33	220	81	0	5	98 %	0 %	2 %	100 %	-2 %	72 %
34	218	81	0	7	97 %	0 %	3 %	100 %	-3 %	71 %
35	213	79	2	12	95 %	2 %	5 %	98 %	-3 %	70 %
36	211	79	2	14	94 %	2 %	6 %	98 %	-4 %	70 %
37	209	79	2	16	93 %	2 %	7 %	98 %	-5 %	69 %
38	203	79	2	22	90 %	2 %	10 %	98 %	-7 %	67 %
39	194	78	3	31	86 %	4 %	14 %	96 %	-10 %	64 %
40	189	77	4	36	84 %	5 %	16 %	95 %	-11 %	63 %
41	186	77	4	39	83 %	5 %	17 %	95 %	-12 %	62 %
42	179	74	7	46	80 %	9 %	20 %	91 %	-12 %	61 %
43	172	72	9	53	76 %	11 %	24 %	89 %	-12 %	59 %
44	165	68	13	60	73 %	16 %	27 %	84 %	-11 %	58 %
45	161	65	16	64	72 %	20 %	28 %	80 %	-9 %	58 %
46	155	62	19	70	69 %	23 %	31 %	77 %	-8 %	57 %
47	152	58	23	73	68 %	28 %	32 %	72 %	-4 %	57 %
48	144	55	26	81	64 %	32 %	36 %	68 %	-4 %	56 %
49	140	52	29	85	62 %	36 %	38 %	64 %	-2 %	55 %
50	132	50	31	93	59 %	38 %	41 %	62 %	-3 %	53 %
51	122	48	33	103	54 %	41 %	46 %	59 %	-5 %	51 %
52	118	46	35	107	52 %	43 %	48 %	57 %	-4 %	50 %
53	108	42	39	117	48 %	48 %	52 %	52 %	-4 %	48 %
54	103	36	45	122	46 %	56 %	54 %	44 %	1 %	48 %
55	94	32	49	131	42 %	60 %	58 %	40 %	2 %	47 %
56	86	30	51	139	38 %	63 %	62 %	37 %	1 %	45 %
57	81	28	53	144	36 %	65 %	64 %	35 %	1 %	44 %
58	73	25	56	152	32 %	69 %	68 %	31 %	2 %	42 %
59	66	25	56	159	29 %	69 %	71 %	31 %	-2 %	40 %
60	59	24	57	166	26 %	70 %	74 %	30 %	-3 %	38 %
61	55	22	59	170	24 %	73 %	76 %	27 %	-3 %	37 %
62	49	19	62	176	22 %	77 %	78 %	23 %	-2 %	36 %
63	45	16	65	180	20 %	80 %	80 %	20 %	0 %	36 %
64	38	15	66	187	17 %	81 %	83 %	19 %	-2 %	34 %
65	33	15	66	192	15 %	81 %	85 %	19 %	-4 %	32 %
66	27	11	70	198	12 %	86 %	88 %	14 %	-2 %	32 %
67	24	9	72	201	11 %	89 %	89 %	11 %	0 %	31 %
68	20	7	74	205	9 %	91 %	91 %	9 %	0 %	31 %
69	18	7	74	207	8 %	91 %	92 %	9 %	-1 %	30 %
70	15	6	75	210	7 %	93 %	93 %	7 %	-1 %	29 %
71	10	4	77	215	4 %	95 %	96 %	5 %	0 %	28 %

72	9	4	77	216	4 %	95 %	96 %	5 %	-1 %	28 %
73	6	3	78	219	3 %	96 %	97 %	4 %	-1 %	27 %
74	4	3	78	221	2 %	96 %	98 %	4 %	-2 %	27 %
75	3	2	79	222	1 %	98 %	99 %	2 %	-1 %	27 %
76	2	2	79	223	1 %	98 %	99 %	2 %	-2 %	26 %
77	1	2	79	224	0 %	98 %	100 %	2 %	-2 %	26 %
78	0	2	79	225	0 %	98 %	100 %	2 %	-2 %	26 %
79	0	1	80	225	0 %	99 %	100 %	1 %	-1 %	26 %
80	0	1	80	225	0 %	99 %	100 %	1 %	-1 %	26 %
81	0	1	80	225	0 %	99 %	100 %	1 %	-1 %	26 %
82	0	1	80	225	0 %	99 %	100 %	1 %	-1 %	26 %
83	0	1	80	225	0 %	99 %	100 %	1 %	-1 %	26 %
84	0	0	81	225	0 %	100 %	100 %	0 %	0 %	26 %

Tabulka č.1 Cut-off skóre věku při chirurgickém zákroku karcinomu prsu; TP= true positive, FP = false positive, TN = true

negative, FN = False negative

Jak můžeme vidět, jsou v tabulce vyznačené dva řádky. Červeně je vyznačen řádek ve věku 26, ve kterém vychází nejlépe Youdenova statistika J, která má hodnotu -12%, tedy spíše neuspokojivou, neboť se pomocí škály podaří odhadnout člověka správně pouze o 12 % lépe, než kdyby náš odhad byl zcela náhodný. Negativní číslo ve výsledku nevádí, pouze znamená, že spolu věk a přežití korelují negativně, což jsme očekávali. Modře je vyznačen řádek s věkem 30, ve kterém nejlépe vychází statistika, která zohledňuje poměr, ve kterém jsou ve vzorku přeživší a zemřelí. Podle tohoto způsobu se podaří určit na základě ideálního cut-offu 74 % respondentů, i tato hodnota je poměrně neuspokojivá vzhledem k tomu, že při tomto postupu bychom pouze řekli, že všichni přežijí. Celkově jsou tedy výsledky spíše neuspokojivé.

Zdroje:

Gibson, K. L., Wu, Y. C., Barnett, Y., Duggan, O., Vaughan, R., Kondeatis, E., ... & Dunn-Walters, D. K. (2009). B-cell diversity decreases in old age and is correlated with poor health status. *Aging cell*, 8(1), 18-25.

Haberman's Survival, Survival of patients who had undergone surgery for breast cancer, [www.kaggle.com/datasets/gilsousa/habermans-survival-data-set](http://www.kaggle.com/datasets/gilsousa/habermans-survival-data-set)

Johnston, S. et. al. (2017). Age, tumor characteristics, and local therapy influence outcomes in breast cancer patients aged 80 years or older. *Annals of Surgical Oncology*, 24(12), 3564- 3573.

Meenan, R. F., Gertman, P. M., & Mason, J. H. (1980). Measuring health status in arthritis. *Arthritis & Rheumatism*, 23(2), 146-152.

Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=171>.