

ANALIZA PRZEŻYCIA

Statystyka biomedyczna

Artur Klepaczko

CO ZAWIERA ZBIÓR ECHOCARDIOGRAM?



- *All the patients suffered heart attacks at some point in the past. **Some are still alive and some are not**. The survival and still-alive variables, when taken together, indicate whether a patient survived for at least one year following the heart attack.*
- Badanie obejmuje grupę osób, z których każda doznała zawału mięśnia sercowego.
- Osoby dołączały do badania w różnym czasie, ale o każdej z nich wiadomo, kiedy w przeszłości wystąpił zawał.
- Część z osób zmarła w trakcie badania. Pozostała część żyła nadal po zakończeniu badania. O tej podgrupie osób nie wiemy, jak długo żyły one po zawale.



RODZAJE BADAŃ

- Badanie *Echocardiogram* to tzw. badanie **prospektywne** (lub kohortowe):
 - ▶ osoby na początku badania są zazwyczaj „zdrowe”
 - ▶ w obserwacji długofalowej poszukuje się określonych zmian chorobowych
- Drugim rodzajem badań są badania **retrospektywne** (kliniczno-kontrolne):
 - ▶ pula osób poddanych badaniu dzieli się na chore i zdrowe (grupa kontrolna)
 - ▶ poszukuje się czynników, które mogły w przeszłości spowodować rozwój choroby

O CO CHODZI W BADANIU PROSPEKTYWNYM?

- W przypadku zbioru *Echocardiogram*, celem jest określenie szansy **przeżycia** co najmniej roku po zawale.
- Ogólniej, chodzi o prawdopodobieństwo przeżycia określonego czasu do śmierci (*time-to-death*).
- „Śmiercią” może być także inne **zdarzenie**, np.:
 - ▶ wystąpienie kolejnego zawału,
 - ▶ nawrót choroby nowotworowej,
 - ▶ wystąpienie objawów choroby GVHD (ang. *graft-versus-host disease*) po przeszczepie.

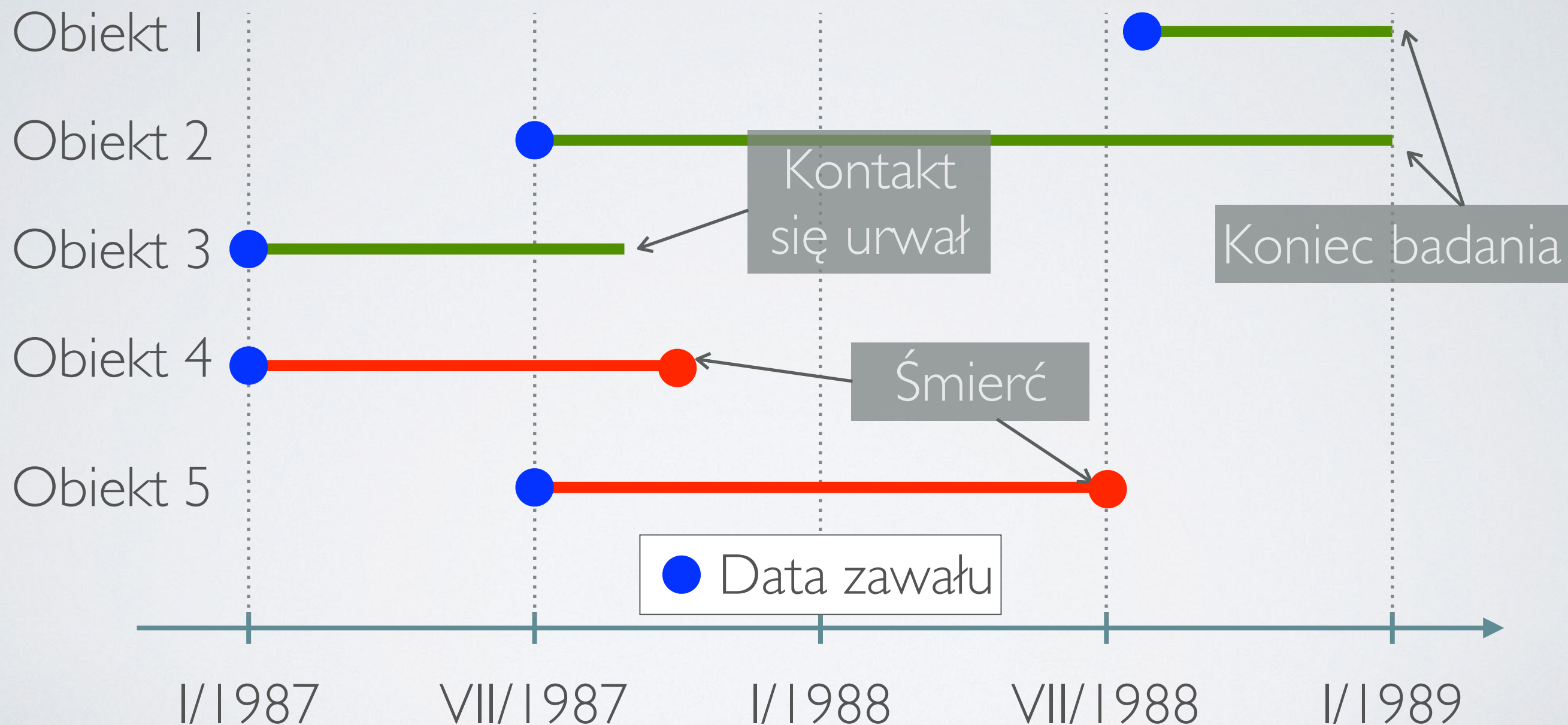
ANALIZA DANYCH *TIME-TO-EVENT*

- Dziedzina statystyki, która zajmuje się tego typu danymi nazywa się **analizą przeżycia** (ang. *survival analysis*).
- Celem analizy przeżycia jest:
 - estymacja najbardziej prawdopodobnego czasu do zdarzenia dla grupy osób
 - porównanie czasów do zdarzenia dla dwóch lub więcej grup (np. poddanych różnym rodzajom terapii)
 - ocena relacji pomiędzy czasem do zdarzenia, a jedną ze zmiennych towarzyszących (np. poziom dawki leku, inne czynniki, które mogą mieć wpływ na rozwój/zahamowanie choroby).

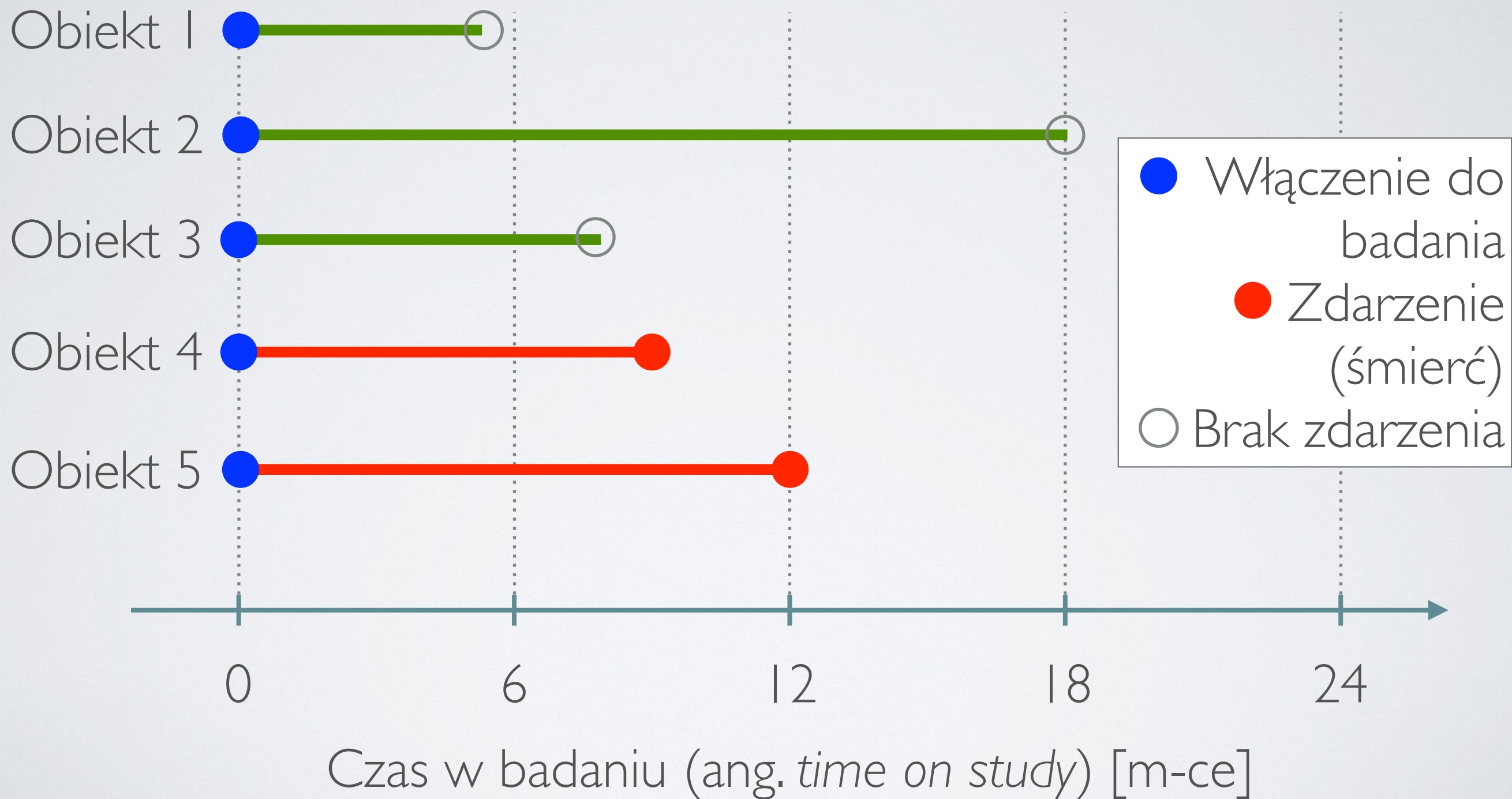
Dlaczego potrzebujemy
specjalnych narzędzi do
danych typu
time-to-event?



SPECYFIKA DANYCH TYPU *TIME-TO-EVENT*



CZAS W BADANIU



NASZE DANE SĄ UCIĘTE!

- Mogłoby się wydawać, że zmienna time-to-event jest zmienną ciągłą i można zastosować metody analizy statystycznej dla ciągłych rozkładów prawdopodobieństwa (test Studenta, regresja liniowa)... (A)
- Może też kusić proste obliczanie stosunku liczby zdarzeń (śmierci) do liczności całej próby losowej... (B)
- Ale wówczas nie bierzemy pod uwagę:
 - ▶ faktu ucięcia danych (ang. *data censoring*) (A)
 - ▶ upływu czasu (B)



PRZYKŁADY UCINANIA

- Koniec badania przed wystąpieniem zdarzenia
- Rezygnacja z udziału w badaniu
- Brak kontaktu z pacjentem
- Śmierć pacjenta w przypadku, gdy zdarzeniem miał być ewentualny rozwój choroby (GVHD)
- W każdym z powyższych przypadków, rzeczywiste czasy przeżycia zakończyłyby się później niż można to było zmierzyć. O takich danych mówimy, że są **prawostronnie ucięte**.



JEST TEŻ LEWA STRONA...

- Może także się zdarzyć, że do badania zostanie włączona osoba, o której wiadomo, że kiedyś w przeszłości nastąpił początek mierzonego okresu, ale nie wiemy, kiedy dokładnie to nastąpiło.
 - ▶ Przykładowo, pacjent stosuje od jakiegoś czasu określony lek, ale wcześniej stosował inny i nie pamięta, kiedy nastąpiła zmiana.
- O takich danych mówimy, że są **lewostronnie ucięte**.

CZY MOŻNA SIĘ POZBYĆ DANYCH UCIĘTYCH?

- Usuwamy ze zbioru danych, wszystkie przypadki, które żyją w chwili zakończenia badania:
 - ▶ bo nie wiemy, jak długo będą jeszcze żyć
 - ▶ nie możemy policzyć średniej długości życia po zawale
- Ale jeśli chcemy znać prawdopodobieństwo dożycia do 12 miesięcy, to usunięcie informacji o przypadkach, które żyły co najmniej 18 miesięcy fałszuje nasze oszacowania :-)
- Nie możemy usuwać takich przypadków tym bardziej, jeśli testujemy skuteczność jakiejś nowej terapii np. w porównaniu z inną istniejącą.

ANI CIAĞŁE, ANI DYSKRETNE

- Dane typu *time-to-event* są zasadniczo odmienne od cech nominalnych lub ciągłych – zdarzenie nie zawsze ma miejsce!
- Jako takie, wymagają one innego podejścia:
 - ▶ do analizy statystycznej
 - ▶ wizualizacji
 - ▶ wnioskowania

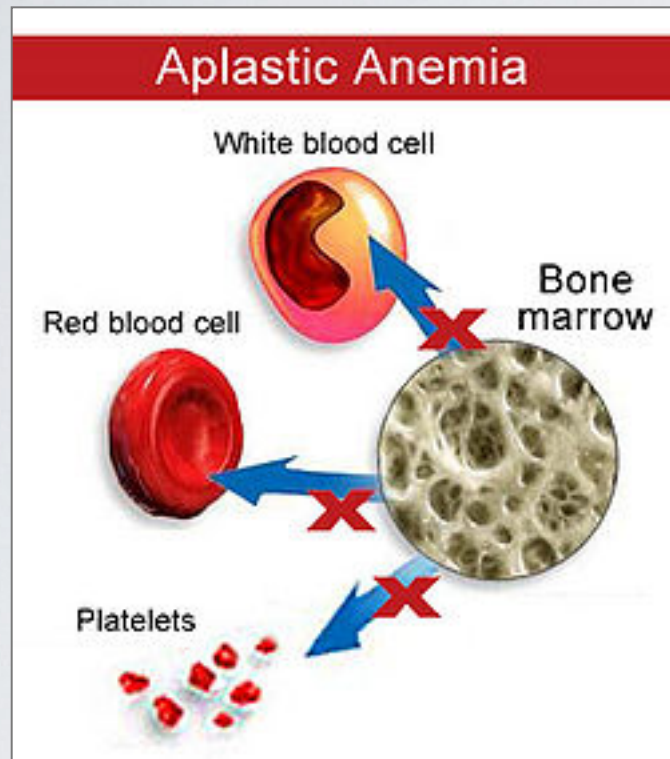


NIPIES, NI WYDRA

PANTA RHEI

- W analizie przeżycia dane są funkcją czasu (w badaniu).
- Zmienną losową zależną od czasu nazywamy *procesem*.
- Z biegiem czasu, mają miejsce
 - ▶ zdarzenia (t_i – czas zdarzenia, $d(t_i)$ – liczba zdarzeń w chwili t_i)
 - ▶ ucinanie danych (c_i – zmienna ucięcia = 0 lub 1)
 - ▶ zmienia się liczba osób biorących udział w badaniu (n)

PRZYKŁADOWE DANE



<http://www.bonemarinara.com/2014/02/aplastic-anemia-qa/>

Która terapia jest skuteczniejsza?

- Niedokrwistość aplastyczna – niewydolność szpiku kostnego do produkcji komórek krwi
- Leczenie: przeszczep szpiku – początek pomiaru czasu
- Komplikacja po przeszczepie: system odpornościowy traktuje przeszczepiony szpik jako ciało obce (choroba GVHD) – zdarzenie
- W zbiorze danych znajdują się dwie grupy pacjentów:
 - ▶ MTX – pacjenci otrzymują tylko metotreksat
 - ▶ MTX + CSP – pacjenci otrzymują także cyklosporynę

RZUT OKA NA DANE # 1

Grupa	Czas [l. dni]	Status GVHD
MTX+CSP	3	Nie
MTX+CSP	10	Tak
MTX+CSP	12	Nie
MTX+CSP	16	Tak
MTX+CSP	22	Tak
...		

t	$n(t)$	$d(t)$
0	22	0
3	22	0
4	21	0
10	21	1
11	20	0
12	20	0
16	19	1
22	18	1
...		

RZUT OKA NA DANE #2

Grupa	Czas [l. dni]	Status GVHD
MTX	9	Tak
MTX	11	Tak
MTX	12	Tak
MTX	20	Tak
MTX	20	Tak
MTX	25	Tak
MTX	25	Tak
MTX	25	Tak
MTX	25	Nie
MTX	28	Tak
MTX	28	Tak
...		

t	$n(t)$	$d(t)$
0	24	0
9	22	0
11	21	1
12	20	0
20	20	0
25	19	1
28	18	1
...		

FUNKCJA PRZEŻYCIA

- Podstawowym narzędziem analizy statystycznej danych typu time-to-event jest tzw. **funkcja przeżycia**.
- Jest ona zdefiniowana jako prawdopodobieństwo, że zdarzenie nie zostanie zaobserwowane do chwili t (pacjent przeżyje przynajmniej do t).
- Funkcję przeżycia oznaczamy $S(t)$.
- Przykładowo, jeśli $S(10) = 95\%$, to oznacza to, że jest tylko 5% prawdopodobieństwa, że zaobserwujemy objawy GVHD 10-ego dnia po przeszczepie.

WYZNACZANIE $S(T)$

- W chwili t , obserwowane prawdopodobieństwo wystąpienia zdarzenia wynosi:

$$\frac{d(t)}{n(t)}$$

- Z drugiej strony, prawdopodobieństwo, że zdarzenie nie będzie miało miejsca danej jest wzorem:

$$\frac{n(t) - d(t)}{n(t)}$$

WYZNACZANIE $S(T)$ - CD.

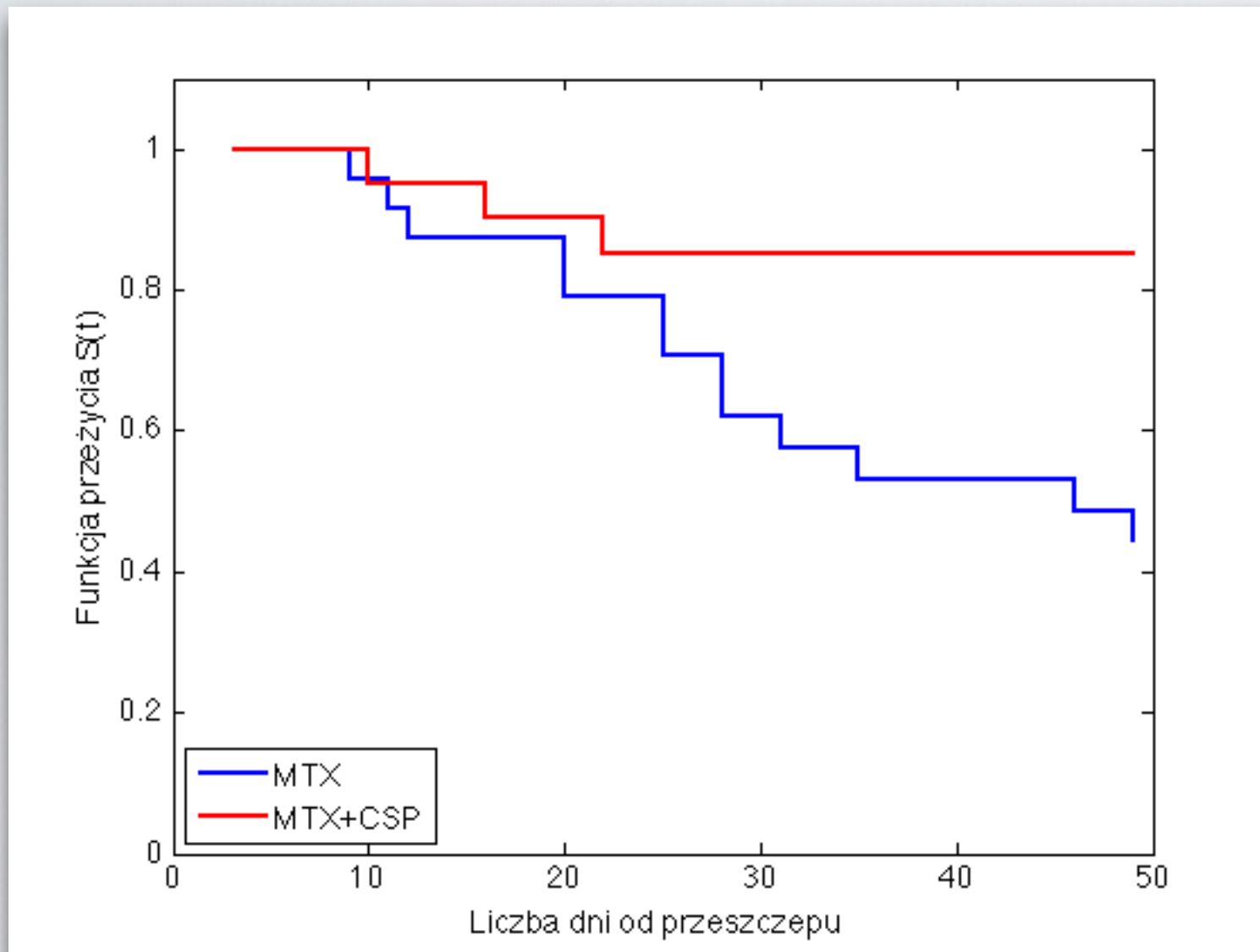
- Przyjmujemy na początku, że $S(t=0) = 1$.
- Z biegiem czasu $S(t)$ maleje o iloraz $d(t)/n(t)$, czyli:

$$\hat{S}(t) = \prod_i \frac{n(t_i) - d(t_i)}{n(t_i)}$$

gdzie t_i oznacza wszystkie chwile czasowe, w których zdarzenie miało miejsce przed chwilą t .

- Tak wyznaczoną funkcję $S(t)$ nazywamy estymatorem **Kaplana-Meiera**.

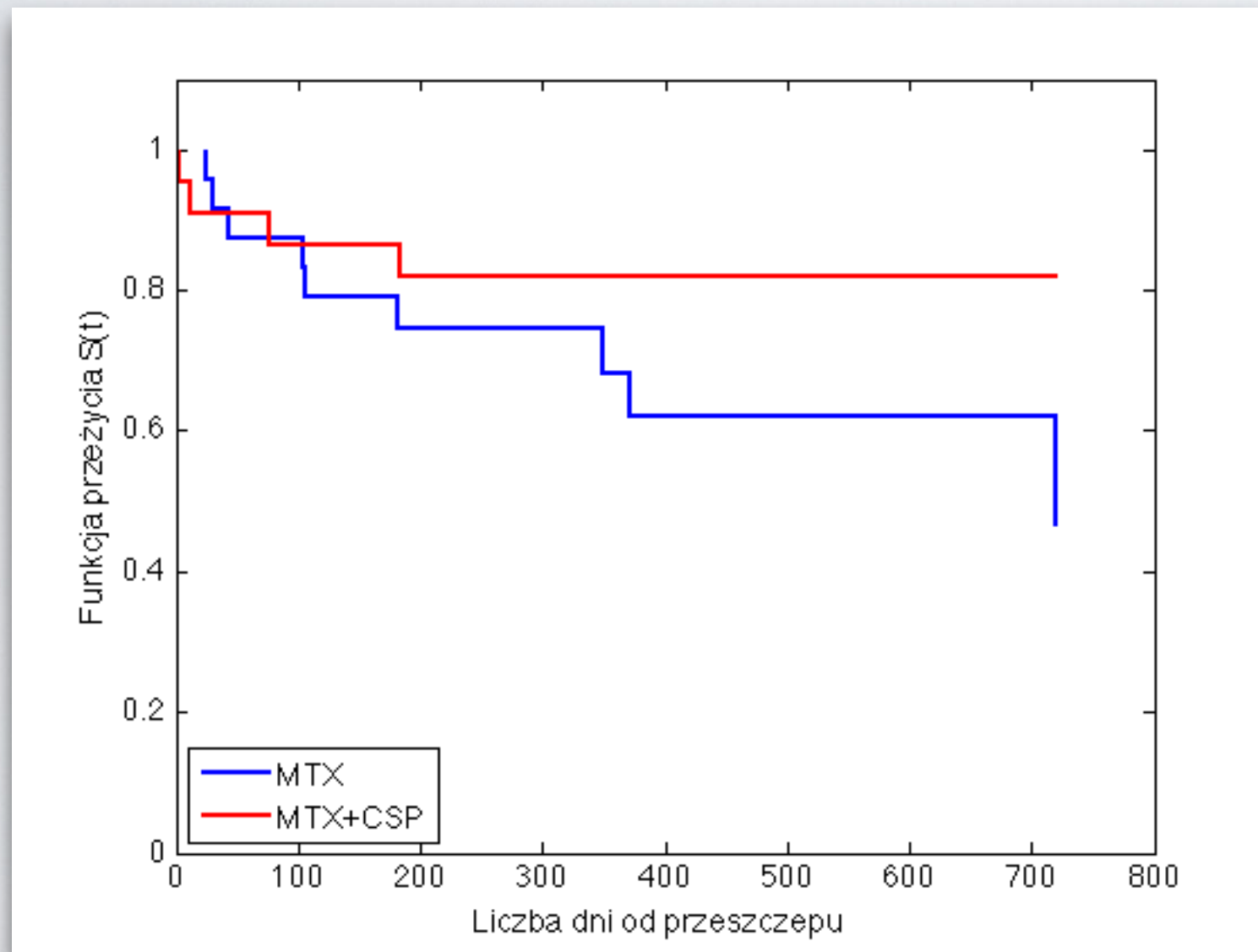
WYKRES FUNKCJI PRZEŻYCIA



Od 22 dnia
prawdopodobieństwo
przeżycia (nie wystąpienia
objawów choroby GVHD)
wynosi ok. 85% dla grupy
MTX+CSP

Funkcję przeżycia
możemy też traktować
jako **empiryczną**
skumulowaną
funkcję rozkładu
prawdopodobieństwa.

WYKRES FUNKCJI PRZEŻYCIA



Tutaj zdarzeniem jest śmierć pacjenta. Wskaźnik przeżycia co najmniej 2 lat po przeszczepie wynosi 82% i 60% odpowiednio dla grup MTX+CSP i MTX

PRZEDZIAŁY UFNOŚCI

- Jak zwykle można wyznaczyć granice przedziału ufności, w którym z prawdopodobieństwem $1-\alpha$ znajduje się prawdziwe prawdopodobieństwo przeżycia.
- Pakiety statystyczne posiadają wbudowane funkcje do obliczania zarówno $S(t)$, jak i przedziałów ufności.
- Przykładowo, w Matlabie możemy wydać polecenie:

```
[S, t, low, up] = ecdf(Times, 'censoring', Cens, 'function',  
'survivor', 'alpha', 0.5);
```

A JEŚLI NIE MA ŻADNEJ RÓŻNICY?



- Czy różnice w estymowanych funkcjach przeżycia są statycznie wiarygodne?
- Na tego typu pytania w statystyce odpowiadają testy statystyczne.
- W przypadku analizy przeżycia korzystamy z testu *log-rank*.

TEST LOG-RANK

- Dla każdej z porównywanych grup notujemy zdarzenia w czasach t_i , z pominięciem czasów ucięcia. Wyniki zapisujemy w tabeli:

	Group A	Group B	Total
Number of deaths observed	a_i	b_i	$a_i + b_i$
Number of patients alive	c_i	d_i	$c_i + d_i$
Number of patients "at risk"	$a_i + c_i$	$b_i + d_i$	$n_i = a_i + b_i + c_i + d_i$

W.W. Daniel, C.L. Cross, *Biostatistics. A Foundation for Analysis in the Health Sciences*, Wiley, 2013

TEST LOG-RANK

- Następnie, dla każdego czasu t_i obliczamy wyrafinowane wzory:

- ▶ oczekiwana częstotliwość zdarzenia w grupie A:

$$e_i = \frac{(a_i + b_i)(a_i + c_i)}{n_i}$$

- ▶ oczekiwane ryzyko międzygrupowe:

$$\nu_i = \frac{(a_i + b_i)(c_i + d_i)(a_i + c_i)(b_i + d_i)}{n_i^2(n_i - 1)}$$

- ▶ test statystyczny Mantela-Heanszel (czyli log-rank test)

$$\chi_{\text{LR}}^2 = \frac{(\sum_{i=1}^k a_i - \sum_{i=1}^k e_i)^2}{\sum_{i=1}^k \nu_i}$$

TEST LOG-RANK

- Przykładowo, zebraliśmy takie dane:

TABLE 14.4.3 Intermediate Calculations for the Log-Rank Test, Example 14.4.1

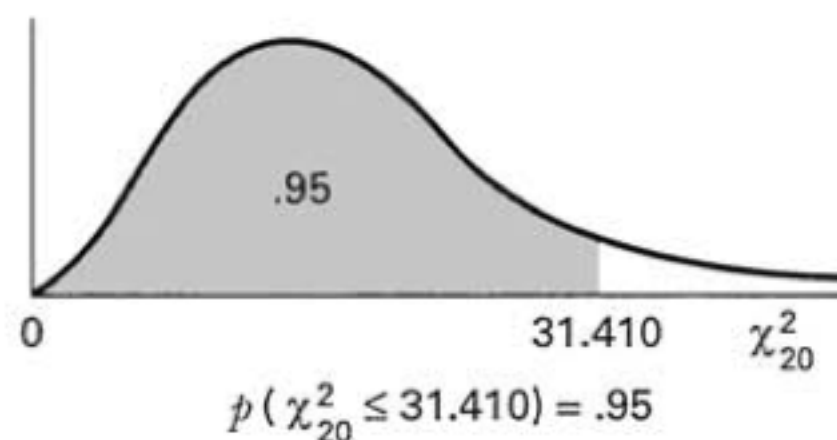
Time, t_j	a_j	c_j	$a_j + c_j$	b_j	d_j	$b_j + d_j$	n_j	e_j	y_j
0	0	25	25	1	13	14	39	0.641	0.230
2	0	25	25	1	12	13	38	0.658	0.225
3	0	25	25	1	11	12	37	0.676	0.219
4	0	25	25	1	10	11	36	0.694	0.212
29	1	17	18	0	1	1	19	0.947	0.050
33	1	16	17	0	1	1	18	0.944	0.052
79	1	13	14	0	1	1	15	0.933	0.062
102	1	10	11	0	1	1	12	0.917	0.076
212	1	1	2	0	0	0	2	1.000	0.000
Totals	9							17.811	3.140

TEST LOG-RANK

- Jeśli obliczona wartość testu jest równa lub większa od wartości krytycznej rozkładu χ^2 dla 1 stopnia swobody i wybranego poziomu ufności to odrzucamy hipotezę zerową o braku statystycznej różnicy pomiędzy obserwowanymi funkcjami przeżycia.
- W naszym przykładzie otrzymujemy:
$$\chi^2 = (9 - 17.811)^2 / 3.140 = 24.724$$

TEST LOG-RANK

TABLE F Percentiles of the Chi-Square Distribution



d.f.	$\chi^2_{.005}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$
1	.0000393	.000982	.00393	2.706	3.841	5.024	6.635	7.879
2	.0100	.0506	.103	4.605	5.991	7.378	9.210	10.597
3	.0717	.216	.352	6.251	7.815	9.348	11.345	12.838
4	.207	.484	.711	7.779	9.488	11.143	13.277	14.860

- Przyjmując poziom ufności aż 99.5%, wartość testu $\chi^2 = 24.724 > 7.879$, co nie upoważnia do odrzucenia hipotezy zerowej (p wartość < 0.005).

CDN.