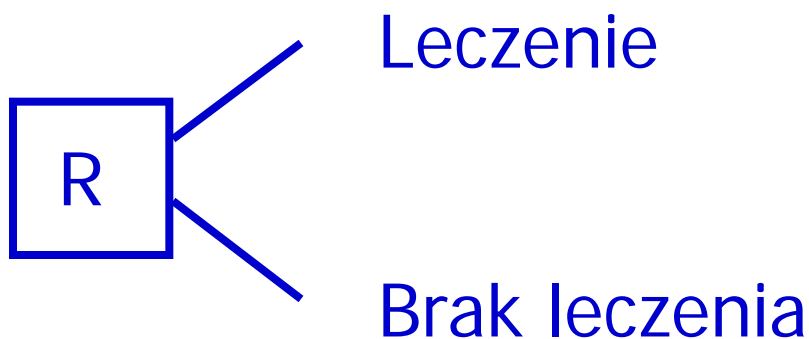


Grupy w próbach klinicznych

- Grupa kontrolna
- Grupy eksperymentalne (leczone)
- Metody randomizacji
- Warstwowanie
- Kiedy randomizować?
- Okres „rozbiegowy”
- Losowanie skupień (cluster randomization)

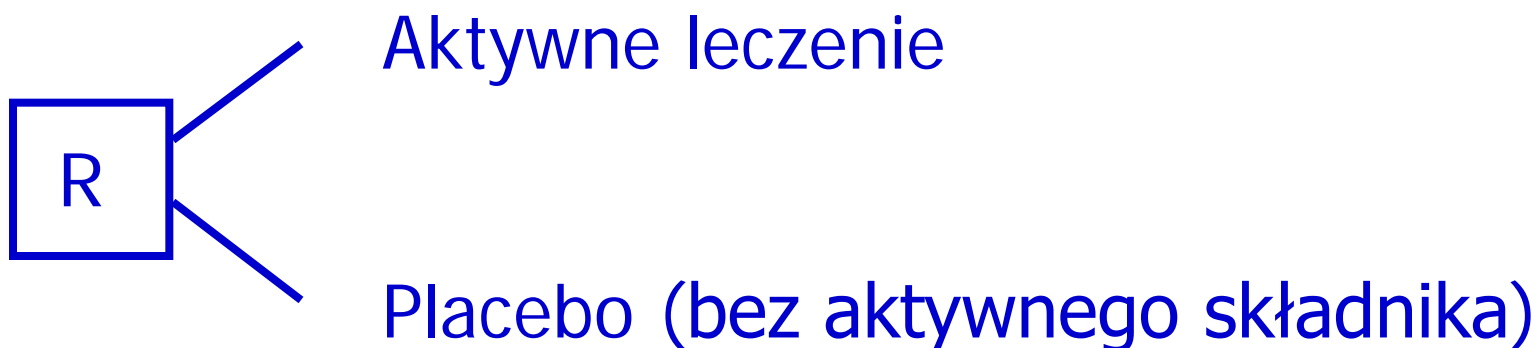
Nieleczona grupa kontrolna



Etycznie dopuszczalna jeśli:

- Brak skutecznego leczenia
- Choroba nie ma „poważnych” skutków

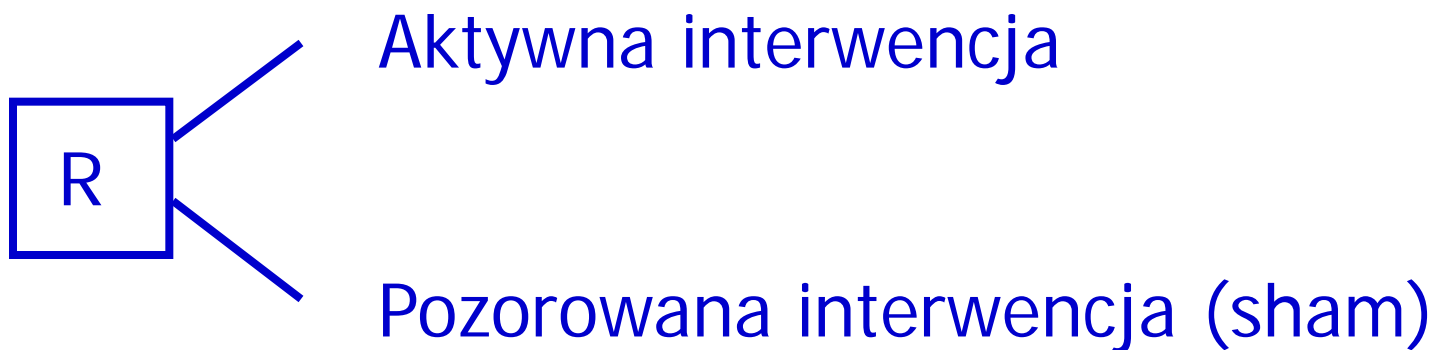
Placebo i „zaślepianie”



Użycie placebo umożliwia „zaślepianie” próby:

- *Pojedynczo ślepa próba*: chory zna rodzaj leczenia, jakie mu przydzielono
- *Podwójnie ślepa próba*: ani chory, ani lekarz nie zna rodzaju leczenia przydzielonego choremu

„Placebo” dla interwencji



Przykład:

- Pozorowana akupunktura (w losowych punktach a nie w 腧穴)
- Pozorowana iniekcja śródgałkowa (chory czuje nacisk strzykawki ale oko nie jest nakłuwane)
- Może być trudne do zastosowania

Placebo jako grupa kontrolna

Wskazane z naukowego punktu widzenia jeśli:

- Informacja o leczeniu może wpływać na opiekę nad pacjentem
- Informacja o leczeniu może wpływać na ocenę wyniku leczenia pacjenta

Placebo jako grupa kontrolna

Przykłady trudne z etycznego punktu widzenia:

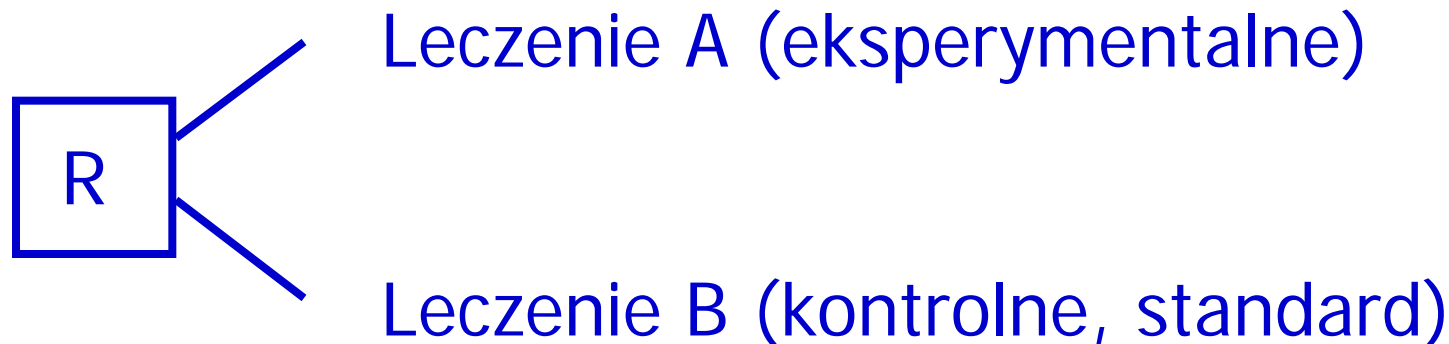
- Próby w zaawansowanej depresji (istnieją skuteczne leki dla chorych)
- Wczesne próby kliniczne AZT w AIDS (AIDS to śmiertelna choroba)
- Iniekcje śródgałkowe w leczeniu ARMD (jeśli placebo byłoby wstrzykiwane do oka)

Grupa kontrolna: szacowanie efektów



Ze statystycznego punktu widzenia,
preferowany jest największy kontrast; tak więc
grupa leczona i kontrolna powinny być tak
różne, jak to możliwe

„Aktywna” grupa kontrolna

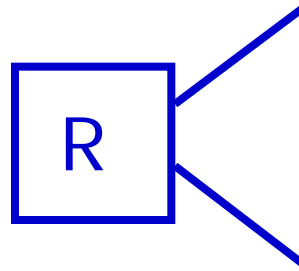


Pytania:

- Czy można wykazać nadrzędność, czy raczej wystarczy nie-podrzędność (np. jeśli nowe leczenie jest mniej toksyczne)?
- Nie-podrzędność jest trudniejsza do wykazania (większa liczebność próbki, problemy z wrażliwością próby)

„Aktywna” grupa kontrolna z „zaślepieniem”

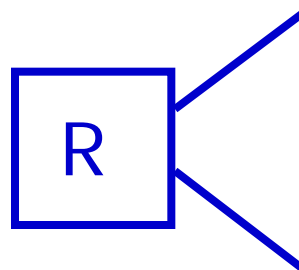
Simple
blinding



Blinded treatment A

Blinded treatment B

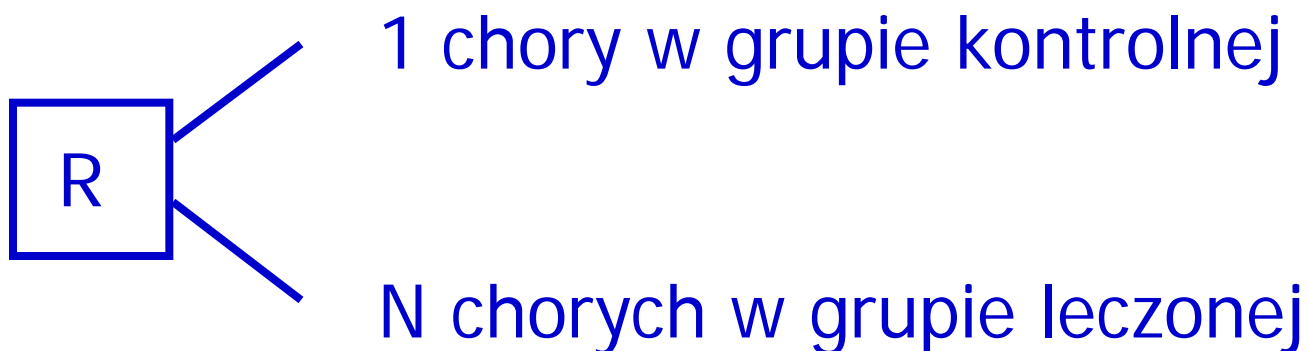
Double
dummy



Blinded treatment A +
placebo for B

Blinded treatment B +
placebo for A

Układy doświadczalne faworyzujące grupę leczoną

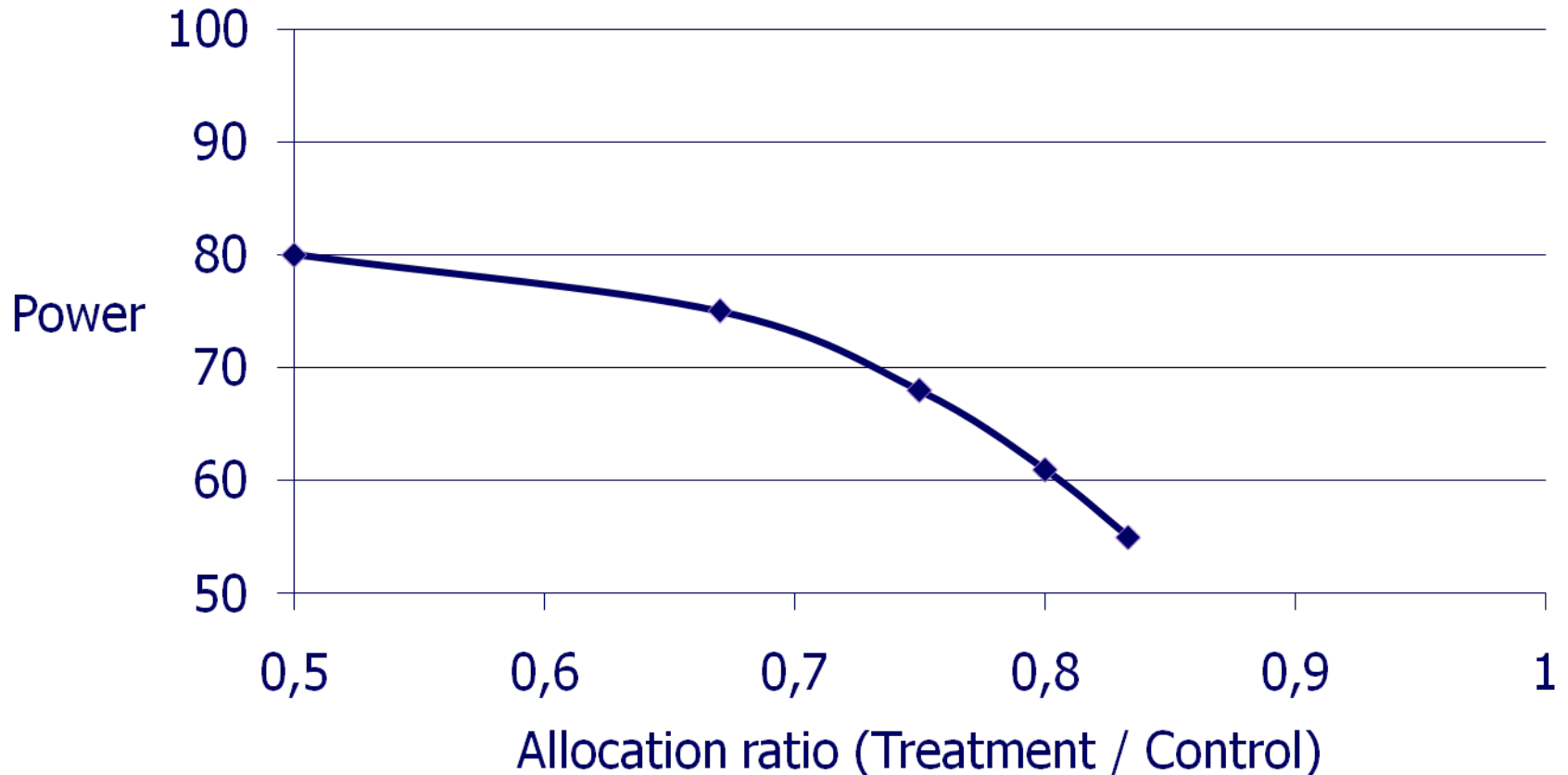


Zalety: Niewielka strata mocy jeśli N małe (< 3)
Więcej informacji o nowym leczeniu

Wady: Etycznie akceptowalne? (lack of equipoise)

Strata mocy dla niezrównoważonych układów doświadczalnych

Power for comparison of 45% vs 60%
among N = 340 subjects

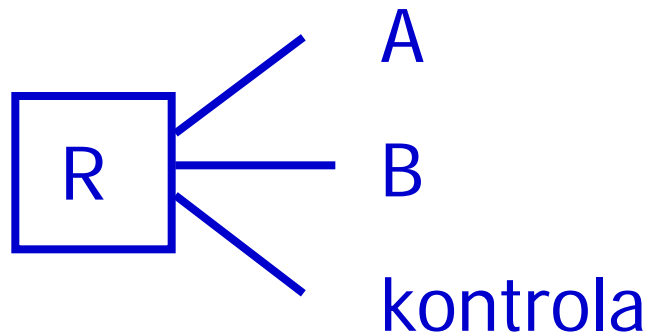


Grupy leczone: ile ?

Generalnie, próba z dwoma grupami (two-arm) jest preferowana z racji :

- efektywności
- łatwości interpretacji

Przykład : próba z trzema grupami (3-arm)



Hipotezy w próbie trziramiennej

- Obie metody leczenia są skuteczne

$$H_0 : r_{(A+B)} = r_0$$

- Jedna z metod jest skuteczna

$$H_0 : r_A = r_0 \quad \text{lub} \quad H_0 : r_B = r_0$$

- Jedna z metod jest skuteczniejsza

$$H_0 : r_A = r_B$$

- Metody się różnią

$$H_0 : r_0 = r_A = r_B \quad (a \ 2 \ d.f. \ test, \ \underline{not} \ efficient)$$

Nieefektywność prób wieloramiennych

Przyjmijmy, że interesuje nas porównania wszystkich par.

Dla K takich porównań, użycie „poprawionego” poziomu istotności skutkuje stratą mocy

Poprawka Bonferroniego (konserwatywna!)

$$\alpha_i = \alpha/K \quad i = 1, \dots, K$$

Użyteczne układy dla prób wieloramiennych

- *Określanie dawki (dose finding)*: placebo vs niska vs średnia vs wysoka dawka
- *Dwa (lub więcej) niezależne pytania*: układ czynnikowy
- Leczenie kombinowane: A vs B vs AB
- „*Gold standard*” : nowy lek vs standard vs kontrola

Randomizacja...

- eliminuje źródła błędów systematycznych (obciążeń) za wyjątkiem przypadkowego niezrównoważenia (accidental bias)
- dąży do zrównoważenia porównywanych grup ze względu na znane i nieznane czynniki prognostyczne
- zapewnia teoretyczne podstawy do użycia statystyk testowych i oszacowań opartych na rozkładach p-stwa

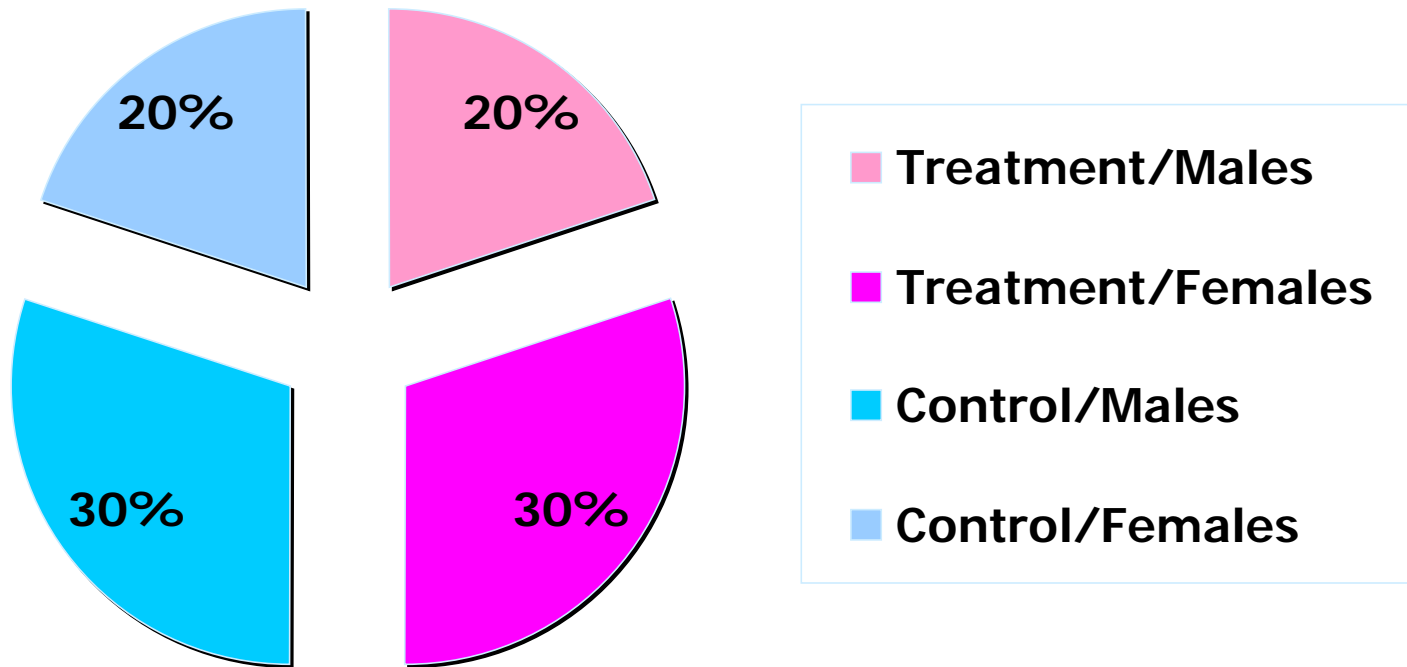
Przypadkowe niezrównoważenie rozkładu czynników prognostycznych

W randomizowanej próbie klinicznej, leczenie jest przydzielane każdemu z chorych losowo.

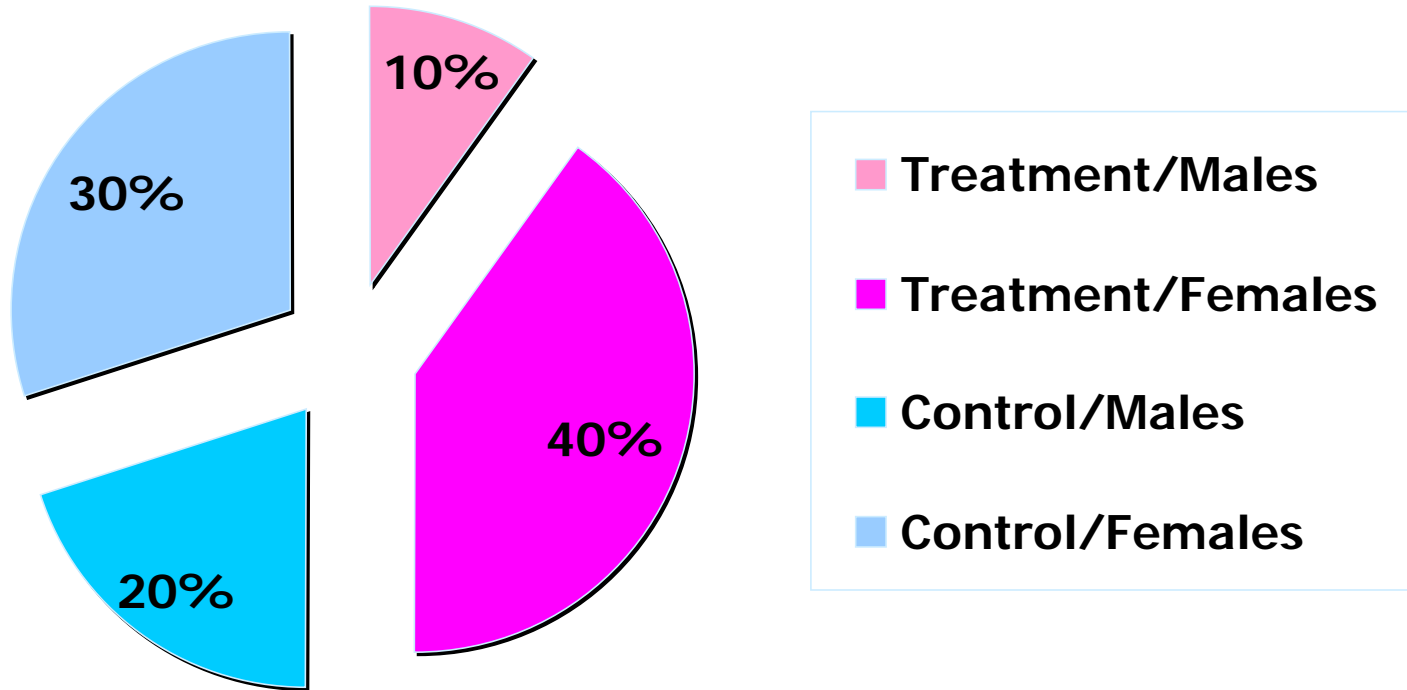
Przypadkowe różnice w rozkładzie czynników prognostycznych mogą się zdarzyć, zwłaszcza w próbkach o niskiej liczebności.

Przykład: rozkład płci w porównywanych grupach

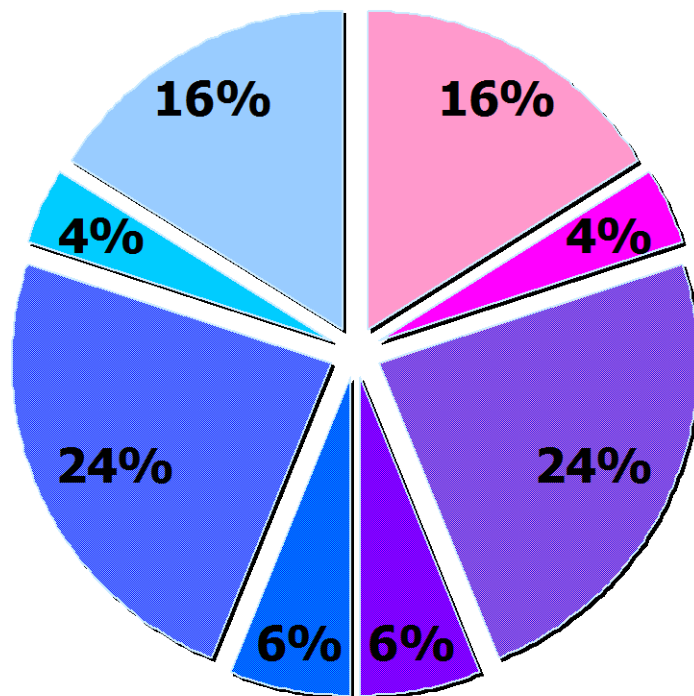
Doskonałe zrównoważenie rozkładu płci



Pryzpadkowe niezrównoważenie rozkładu płci



Doskonałe zrównoważenie rozkładu płci i mierzalności guza



- Treatment/ Males/ Measurable
- Treatment/ Males/ Non-measurable
- Treatment/ Females/ Measurable
- Treatment/ Females/ Non-measurable
- Control/ Males/ Measurable
- Control/ Males/ Non-measurable
- Control/ Females/ Measurable
- Control/ Females/ Non-measurable

Przypadkowe niezrównoważenie rozkładu czynników prognostycznych

Może skutkować błędem systematycznym jeśli dotyczy czynników o dużym znaczeniu prognostycznym.

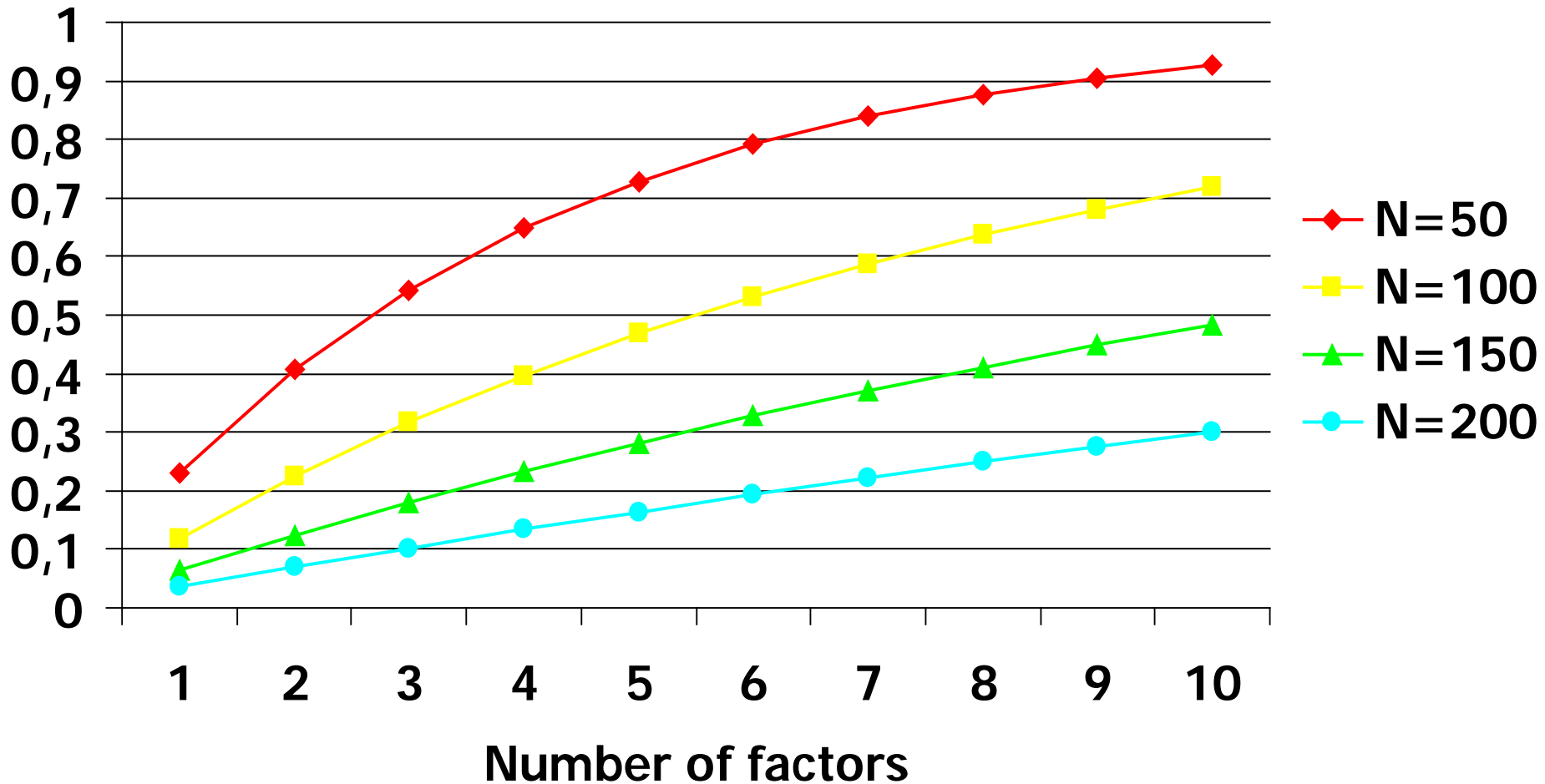
Nawet jeśli, ze statystycznego punktu widzenia, niezrównoważenie jest ignorowalne, może rodzić wątpliwości co do zaobserwowanych różnic w skuteczności leczenia.

Przypadkowe niezrównoważenie rozkładu czynników prognostycznych

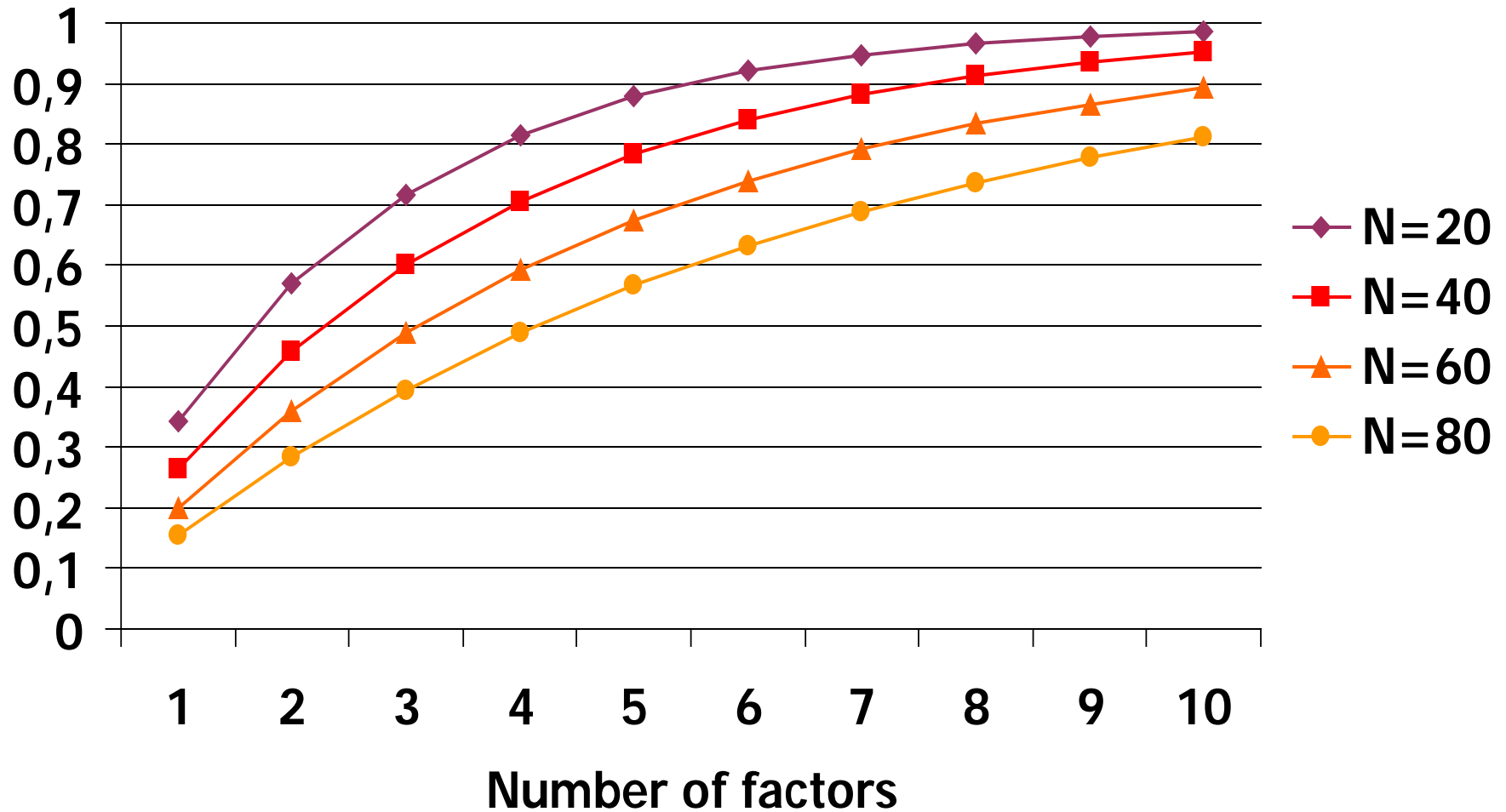
Prawdopodobieństwo rośnie z

- malejącą liczebnością próbki
- rosnącą liczbą czynników prognostycznych

Probability of $> 40/60$ imbalance



Probability of $> 40/60$ imbalance



Obciążenie wynikające z selekcji chorych (selection bias)

Nie zrównoważenie rozkładu czynników prognostycznych wynikające z wiedzy o rodzaju leczenia, które ma być przydzielone choremu

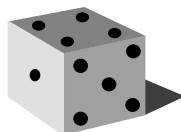
Obciążenie wynikające z selekcji chorych (selection bias)

Może się pojawić jeśli randomizacja jest źle przeprowadzona :

- drukowane listy liczb pseudo-losowych
(można przeczytać)
- użycie kopert z informacją o leczeniu
(można otworzyć)
- przewidywalne algorytmy oparte o „statyczne” metody
(np. ostatni element permutowanego bloku)

Randomizacja prosta

Użycie monety, kostki lub innego procesu *losowego*.



Orzeł

Reszka



Leczenie

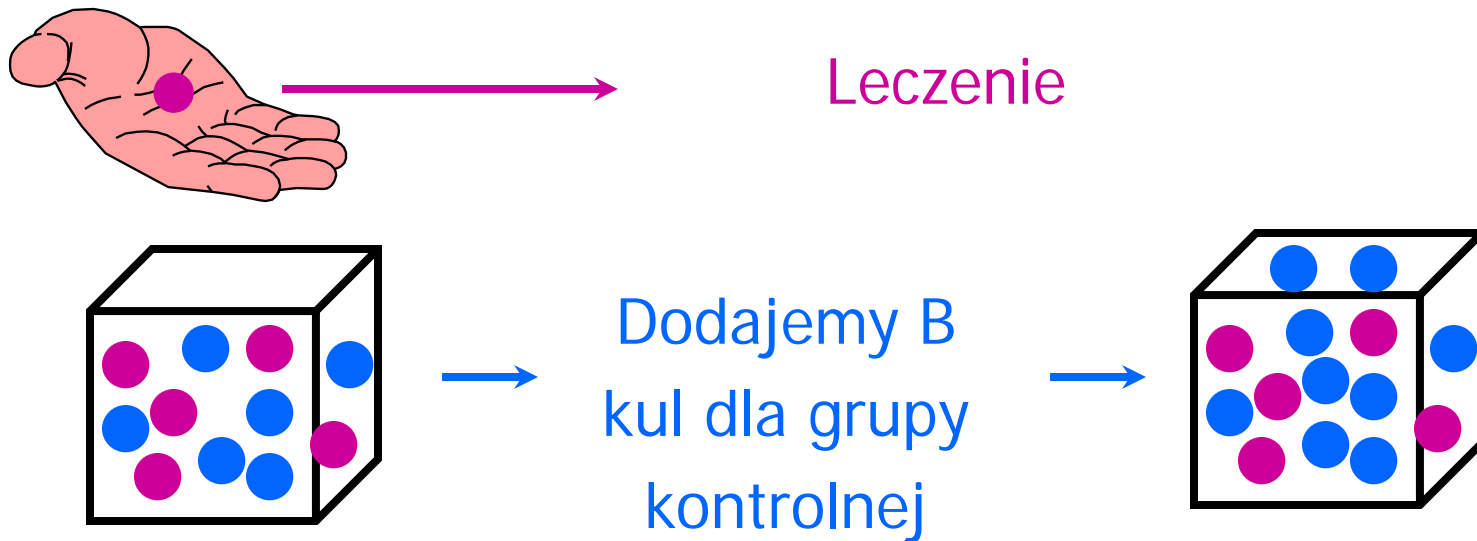
Kontrola

Wady i zalety prostej randomizacji

- + Prosta 😊
- + Nieprzewidywalna
- Nie uwzględnia znanych czynników prognostycznych
- Możliwość przypadkowego niezrównoważenia

Randomizacja urnowa

Urnę początkowo zawiera A kul każdego koloru.
Jedna kula jest losowana i przydzielane jest leczenie.
 B kul drugiego koloru jest dodawane.



Wady i zalety randomizacji urnowej

- + Elastyczność:

 - A duże, B małe \Rightarrow faworyzuje losowość

 - B duże, A małe \Rightarrow zrównoważenie (początkowo)

- + Może być uwzględniona w analizie (test permutacji)

- Staje się losowa gdy N rośnie

Warstwowanie

Czynnik warstwujący to zmienna kategoryzowana, która dzieli populację chorych na kategorie

Przykłady :

- płeć, 2 poziomy : kobieta, mężczyzna
- wiek, np. 3 poziomy : < 40 , $40-59$, ≥ 60 lat

Warstwowanie

Dla 2 i więcej czynników, warstwy zdefiniowane kombinacjami kategorii

Przykład :

Płeć i wiek (jak poprzednio) dają 6 warstw :

Warstwa 1: Kobiety < 40

Warstwa 2: Kobiety 40-59

Warstwa 3: Kobiety ≥ 60

Warstwa 4: Mężczyźni < 40

Warstwa 5: Mężczyźni 40-59

Warstwa 6: Mężczyźni ≥ 60

Uwzględnienie czynników prognostycznych w randomizacji

Czynniki prognostyczne dla terapii
adjuwantowej raka piersi:

- Stan menopauzalny (pre/post)
- Liczba zajętych węzłów chłonnych (1-3/4+)
- Receptory estrogenu (+/-)
- Receptory progesteronu (+/-)
- mutacja genu her2neu (under/overexpressed)
- ośrodek kliniczny (np. 20)

$$\Rightarrow 2^5 \times 20 = 640 \text{ różnych warstw !}$$

Uwzględnienie czynników prognostycznych w randomizacji

Permutowane bloki
w warstwach

Tylko ośrodek i stan
menopauzalny mogą
być uwzględnione
(40 warstwa)

Minimizacja

Wszystkie sześć
czynników może być
uwzględnionych

Permutowane bloki w warstwach

Przyjmijmy, że kolejna chora jest zgłoszona z ośrodka 02 i ma następującą charakterystykę

- przed menopauzą
- 1-3 zajętych węzłów
- dodatni poziom receptorów estrogenu
- dodatni poziom receptorów progesteronu
- obniżona ekspresja genu her2neu

Permutowane bloki w warstwach

<u>Center</u>	<u>Menopause</u>	<u>Lists of permuted blocks</u>			
01	Pre	C T C TT C C T	T C T C	T T C C	
01	Post	C C T TC T T C	C T C T	T T C C	
→ 02	Pre	C T C TT C T C	T T C C	C T T C	
02	Post	C T T CT C T C C C T T	C C T T		
03	Pre	T C C T	C T C T	C T T C	T C T C
03	Post	C C T T	C T T C	T C C T	C T C T

The next treatment in the patient's stratum is T

Wady i zalety randomizacji przy użyciu permutowanych bloków

- + Losowa procedura uwzględniająca informację o czynnikach prognostycznych
- + Łatwa w implementacji → często używana przy pakowaniu leków
- Ograniczona liczba warstw
- Przewidywalny wynik pod koniec bloku
- Informacja o czynnikach prognostycznych musi być dostępna przed randomizacją

Minimizacja brzegowego rozkładu czynn timerów prognostycznych

	<u>I</u>	<u>C</u>
Menopausal status = pre	21	22
Nr of nodes = 4+	18	19
ER status = positive	31	32
PR status = positive	26	23
Her2neu gene = overexpressed	13	12

Minimizacja brzegowego rozkładu czynn timerów prognostycznych: kolejna chora otrzymuje T

	<u>I</u>	<u>C</u>	<u>δ</u>
Menopausal status = pre	22	22	0
Nr of nodes = 4+	19	19	0
ER status = positive	32	32	0
PR status = positive	27	23	-4
Her2neu gene = overexpressed	14	12	-2

"Range of imbalances": $\sum |\delta| = 6$

"Variance of imbalances": $\sum \delta^2 = 20$

Minimizacja brzegowego rozkładu czynn timerów prognostycznych: kolejna chora otrzymuje C

	<u>I</u>	<u>C</u>	<u>δ</u>
Menopausal status = pre	21	23	+2
Nr of nodes = 4+	18	20	+2
ER status = positive	31	33	+2
PR status = positive	26	24	-2
Her2neu gene = overexpressed	13	13	0

“Range of imbalances”: $\sum |\delta| = 8$

“Variance of imbalances”: $\sum \delta^2 = 16$

Minimizacja brzegowego rozkładu czynn timerów prognostycznych

Przydzielamy T jeśli "zakres niezerównoważeń" ma być minimalizowany

Przydzielamy C jeśli "wariancja niezerównoważeń" ma być minimalizowana

Jeśli ma być minimalizowana wariancja, prostsza implementacja jest możliwa

Minimizacja brzegowego rozkładu czynn timerów prognostycznych: prostsza implementacja

	<u>I</u>	<u>C</u>
Menopausal status = pre	21	22
Nr of nodes = 4+	18	19
ER status = positive	31	32
PR status = positive	26	23
Her2neu gene = overexpressed	13	12
<u>Total</u>	<u>109</u>	<u>108</u>
	↓	↓

Allocate C to minimize "total" imbalance

Minimizacja brzegowego rozkładu czynników prognostycznych

δ : różnica w sumarycznej liczbie dla C i T

- jeśli $\delta = 0$, T lub C mogą być przydzielone
- jeśli $\delta < 0$, T powinno być faworyzowane
- jeśli $\delta > 0$, C powinno być faworyzowane

Niech $\delta < 0$, i niech p_x będzie p-stwem przydzielenia leczenia X ($X = T, C$)

Możliwe wybory w minimizacji

$p_T = p_C$: prosta randomizacja

$p_T = p$ i $p_C = 1-p$ dla $\frac{1}{2} < p < 1$

(np. $p = .8$): minimizacja stochastyczna
(analog rzutu obciążoną monetą, „biased coin” design)

$p_T = 1$ i $p_C = 0$ jeśli $\delta > S$, dla małego S : semi-deterministyczna minimizacja

$p_T = 1$ i $p_C = 0$: deterministyczna minimizacja

Minimizacja w skomplikowanych próbach klinicznych

Przykład: adjuwantowe leczenie raka piersi
(A = adriamycin; C = cyclophosphamide; T = taxotere)

Dwie grupy kontrolne

$C_1 : A$

$C_2 : A \ C$

Dwie grupy leczone

$T_1 : A \rightsquigarrow T$

$T_2 : A \ T$

Minimizacja w skomplikowanych próbach klinicznych

Próba ma ocenić efekt

1. Taxotere podanego po Adriamycinie (sekwencja)

$C_1 : A$ vs $T_1 : A \Rightarrow T$

2. Taxotere podanego z Adriamyciną (kombinacja)

$C_1 : AC$ vs $T_2 : AT$

3. Taxotere (ogólnie)

$C_1 + C_2$ vs $T_1 + T_2$

Minimizacja z niezrównoważonym przydziałem leczenia

Chcemy przydzielić dwa razy więcej chorych do grup leczonych:

C_1 control group 1: $n/2$ patients

C_2 control group 2: $n/2$ patients

T_1 treatment group 1: n patients

T_2 treatment group 2: n patients

Minimizacja z niezrównoważonym przydziałem leczenia

Dodatkowo, chcemy uwzględnić następujące czynniki prognostyczne :

- liczbę zajętych węzłów
- stan receptorów estrogenu

oraz

- ośrodek

Minimizacja z niezrównoważonym przydziałem leczenia

<i>Allocation ratio</i>	1	1	2	2
<i>Treatment arm</i>	<u>C₁</u>	<u>C₂</u>	<u>I₁</u>	<u>I₂</u>
Center = 02	2	1	4	3
Nr of nodes = 4+	9	10	19	20
ER status = positive	15	14	30	28
×coefficient	<u>×2</u>	<u>×2</u>	<u>×1</u>	<u>×1</u>
<u>Adjusted total</u> =	<u>52</u>	<u>50</u>	<u>53</u>	<u>51</u>

Minimizacja z niezrównoważonym przydziałem leczenia

\mathcal{M} : zbiór metod leczenia z najmniejszą sumą liczb

\mathcal{N} : pozostałe metody leczenia

Wybieramy zbiór \mathcal{M} z p-stwem $p_{\mathcal{M}}$ (np. 0.8), z zbiór \mathcal{N} z p-stwem $1 - p_{\mathcal{M}}$ (np. 0.2)

W wybranym zbiorze przydzielamy metody leczenia z różnym p-stwem

Wady i zalety minimizacji

- + Dąży do zrównoważenia rozkładu czynników prognostycznych (użyteczna zwłaszcza w małych próbach lub gdy koszt leku jest wysoki)
- + Liczba czynników prognostycznych może być znaczna
- + Redukcja rozmiaru testu istotności (eliminacja przypadkowego niezrównoważenia)
- Niewielki zysk na mocy testu
- Zastrzeżenia do analizy
- Bardziej skomplikowana implementacja

Analiza prób używających randomizacji z „wiązaniami”

Czy w takich próbach do porównania grup nie należałoby używać testów “permutacyjnych” (“randomizacyjnych”)?

Takie zalecenie jest czasem formułowane w przypadku użycia randomizacji „z wiązaniami” (constrained randomization), np. schematów urnowych, permutowanych bloków, lub minimizacji. Testy te dają poziomy krytyczne bliskie asymptotycznym dla dużych liczebnie próbek.

Testy permutacyjne dla permutowanych bloków

Rozważmy jeden blok, gdzie P_X = p-stwo przydziału X ($X=T, C$)

Patient	Arm	P_T	P_C	Outcome
Pt-1	T	0.5	0.5	Success
Pt-2	T	0.5	0.5	Failure
Pt-3	C	0	1	Failure
Pt-4	C	0	1	Failure

Testy permutacyjne

„Dokładny” poziom krytyczny testu (exact p-value) wyznaczony jest przez liczbę permutacji, dla których statystyka testowa (w naszym przykładzie, różnica w odsetku odpowiedzi dla grupy leczonej i kontrolnej) jest co najmniej tak duża jak zaobserwowana.

Testy permutacyjne dla permutowanych bloków

Wszystkie permutacje zgodne z randomizacją, gdzie S_X = odsetek odpowiedzi dla grupy X (X=T,C)

Permutation	S_T	S_C	$S_T - S_C$
T T C C	1/2	0/2	0.5
T C T C	1/2	0/2	0.5
T C C T	1/2	0/2	0.5
C T T C	0/2	1/2	-0.5
C T C T	0/2	1/2	-0.5
C C T T	0/2	1/2	-0.5

← Actual allocation
p-value = 0.5
(= 3 / 6)

Testy permutacyjne dla permutowanych bloków

Ten sam blok dla randomizacji prostej

Patient	Arm	P_T	P_C	Outcome
Pt-1	T	0.5	0.5	Success
Pt-2	T	0.5	0.5	Failure
Pt-3	C	0.5	0.5	Failure
Pt-4	C	0.5	0.5	Failure

Testy permutacyjne dla permutowanych bloków

Permutation	S_T	S_C	$S_T - S_C$
T C C C	1/1	0/3	1
T C T C	1/2	0/2	0.5
T C C T	1/2	0/2	0.5
T T C C	1/2	0/2	0.5
T T T C	1/3	0/1	0.33
T T C T	1/3	0/1	0.33
T C T T	1/3	0/1	0.33
T T T T	1/4	0/0	0.25
C C C C	0/0	1/4	-0.25
C T C C	0/1	1/3	-0.33
C C T C	0/1	1/3	-0.33
C C C T	0/1	1/3	-0.33
C C T T	0/2	1/2	-0.5
C T T C	0/2	1/2	-0.5
C T C T	0/2	1/2	-0.5
C T T T	0/3	1/1	-1

← Actual allocation
p-value = 0.25
(= 4 / 16)

Testy permutacyjne dla permutowanych bloków

W naszym przykładzie, poziom krytyczny testu permutacyjnego $p=0.25$ dla randomizacji prostej, a $p=0.5$ dla permutowanych bloków.

Generalnie, dla randomizacji „z więzami” testy permutacji dają konserwatywne poziomy krytyczne, bowiem oparte są na mniejszej liczbie punktów w przestrzeni permutacji (niż testy asymptotyczne).

Testy permutacyjne

Liczba permutacji rośnie wykładniczo (2^N dla N chorych przy użyciu randomizacji prostej) i prawdopodobieństwo każdej z nich jest trudne do wyznaczenia dla minimizacji

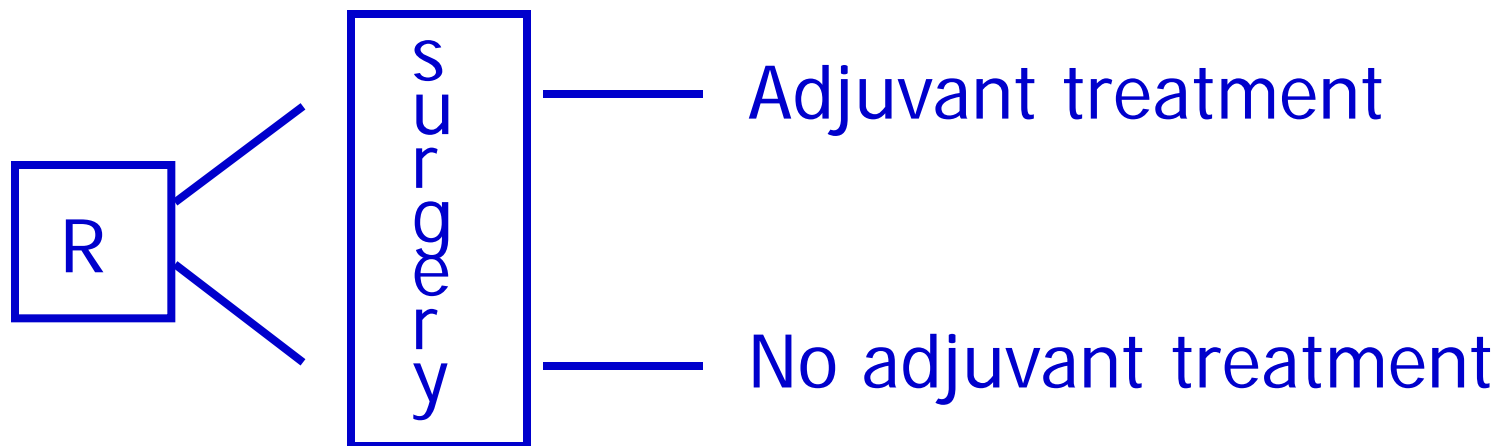
⇒ Symulacje:

- Randomizujemy tych samych chorych w symulowanych próbach
- Obliczamy poziom krytyczny (lub statystykę) testu dla każdej próby
- Porównujemy zaobserwowany poziom krytyczny (lub statystykę) testu do rozkładu otrzymanego z symulacji

Kiedy randomizować ?

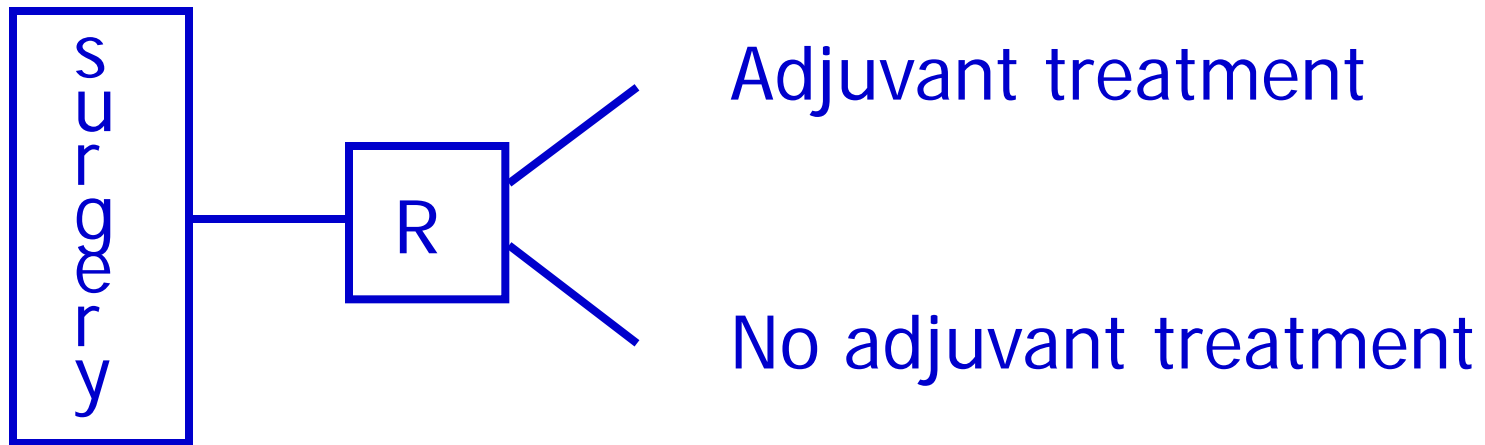
Jak najpóźniej jak jest to możliwe.

*Przykład: pooperacyjne leczenie adjuwantowe
wczesnego raka jelita grubego*



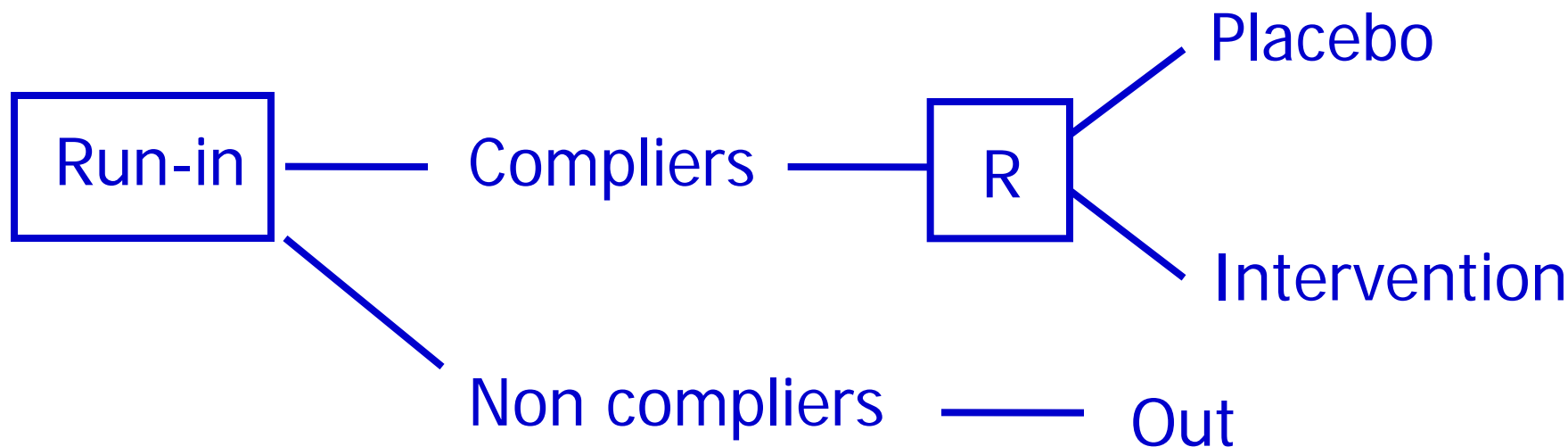
Kiedy randomizować ?

Przykład: lepiej randomizować chorych po operacji, by uniknąć randomizowania pacjentów, którzy nie otrzymają leczenia adjuwantowego



Okres „rozbiegowy” (run-in)

Czasem używany przed randomizacją. W tym okresie wszyscy chorzy są leczeni w ten sam sposób. Pozwala to na wykluczenie chorych, którzy nie stosują się do wymogów leczenia.



Losowanie skupień (cluster randomization)

Polega na randomizacji całych populacji chorych.

Przykłady:

- *Fabryki (np. próby badań przesiewowych)*
- *Miasta (np. próby kampanii na rzecz rzucenia palenia)*
- *Rodziny (np. próby zmian sposobu żywienia)*

Losowanie skupień

Wymagane gdy interwencja organizowana jest na poziomie grupy/populacji.

Obserwacje w tym samym skupieniu są zależne, co wymaga odpowiedniego uwzględnienia w analizie i powoduje redukcję mocy statystycznej.

Przykład losowania skupień

Próba COMMIT

Cel

Dotarcie do palaczy (zwłaszcza dużo palących) z kampanią na rzecz trwałego rzucenia palenia

Interwencja

Poprzez lokalne organizacje, media, lekarzy, przedsiębiorstwa, itp.

Losowanie skupień

Randomizacja 11 par gmin o średniej wielkości (50,000 – 250,000 mieszkańców) dopasowanych geograficznie

Przykład losowania skupień

Próba COMMIT

**Pierwszoplanowe kryterium oceny skuteczności:
odsetek osób, które rzuciły palenie**

Dla każdej z m par gmin, różnica w odsetkach

$$d_j = P_{1j} - P_{2j} \quad (j = 1 \dots m)$$

Wariancja rozkładu dwumianowego musi być
zwiększona ze względu na zależność obserwacji

$$\text{Var}(d_j) = P_{1j}(1 - P_{1j}) / n_{1j} + P_{2j}(1 - P_{2j}) / n_{2j} + 2\sigma^2$$

gdzie σ^2 jest między-gminną wariancją (oszacowaną z poprzednich prób)

Przykład losowania skupień

Próba COMMIT - analiza

Porównanie różnic $d_j = P_{1j} - P_{2j}$ przy pomocy testu permutacyjnego

Procedura randomizacji umożliwia przypisanie rozkładu różnicom d_j przy założeniu H_0 :

- Zakładając H_0 , $E(d_j) = 0$
- Wyznaczamy średnią d_j dla każdej z 2^m permutacji prowadzących do różnych przydziałów interwencji w parach ($m = 11$)
- Porównujemy obserwowaną średnią do rozkładu otrzymanego z permutacji

Przykład losowania skupień

Próba COMMIT

Wyniki: odsetki rzucających palenie

Extent of smoking	Intervention	Control	Difference
Heavy	N = 4,976	N = 5,043	
	0.180	0.187	- 0.007 (NS)
Light to moderate	N = 5,177	N = 5,151	
	0.306	0.276	+ 0.030 (<i>P</i> = 0.004)