

# Statystyka w badaniach klinicznych: najważniejsze pojęcia

## Wprowadzenie

Metody statystyczne umożliwiają przeprowadzanie formalnych obliczeń dotyczących zmienności odpowiedzi pacjentów na leczenie. Zastosowanie statystyki pozwala badaczom klinicznym na formułowanie uzasadnionych i dokładnych wniosków na podstawie zebranych informacji oraz na podejmowanie właściwych decyzji w razie braku pewności. Statystyki mają kluczowe znaczenie dla zapobiegania błędom, w tym systematycznym, w badaniach medycznych. W tym artykule przedstawiono niektóre główne pojęcia statystyczne i ich zastosowanie w badaniach klinicznych.

## Testowanie hipotezy

Hipoteza jest założeniem lub zestawem założeń, które a) stwierdza coś tymczasowo na potrzeby przeprowadzenia badań naukowych lub b) potwierdza coś jako wysoce prawdopodobne w świetle znanych faktów.

Dla celów tego artykułu posłużymy się hipotezą, która coś stwierdza – na przykład, że nowa terapia choroby jest lepsza niż istniejące leczenie standardowe. Jeśli nowa terapia zostanie określona jako „B”, a standardowa jako „A”, to hipoteza stwierdza, że „B” jest lepsza niż „A”.

Można przypuszczać, że naukowcy zajmą się udowodnieniem tej hipotezy, ale w tym wypadku tak nie jest. Cel jest realizowany metodą pośrednią. Zamiast próbować udowodnić hipotezę B, metoda naukowa zakłada, że w rzeczywistości prawdziwa jest

wersja A – a więc że nie ma różnicy między standardowym i nowym leczeniem. Określa się to jako hipotezę „zerową”. Naukowcy przystępują następnie do obalenia hipotezy A. Takie postępowanie nosi nazwę falsyfikacji hipotezy zerowej (stwierdzenia nieprawdziwości hipotezy zerowej). Jeśli falsyfikacja się powiedzie – dowiedzie, że hipoteza A jest fałszywa i że leczenie standardowe nie jest lepsze niż nowe – w konsekwencji zostanie potwierdzona prawdziwość hipotezy B, a więc tego, że nowe leczenie jest lepsze niż standardowe.

## **Jak się to robi?**

Nie ma prostej odpowiedzi – ta szeroko akceptowana metoda zmieniała się w nowoczesnej nauce – ale pomocne może okazać się zastosowanie analogii prawniczej. Hipoteza zerowa dotyczy naszej obecnej sytuacji lub wiedzy (analogicznie w sądzie zakłada się że „oskarżony jest niewinny”), o której prawdziwości musimy być przekonani, dopóki nie uzyskamy wystarczających dowodów, że jest inaczej. Jeśli jednak staralibyśmy się udowodnić „alternatywną hipotezę” (jak wiadomo, przeciwieństwo „hipotezy zerowej”), to zakładalibyśmy, że „oskarżony jest winny”.

Innym, może łatwiejszym sposobem wyjaśnienia tego rozumowania jest zacytowanie Alberta Einsteina:

*„Żadna liczba eksperymentów nie może dowieść, że mam rację: jeden eksperyment może pokazać, że jej nie mam”.*

Słowa te sugerują, że próba udowodnienia, że hipoteza zerowa jest fałszywa lub błędna, jest bardziej rygorystycznym i osiągalnym celem niż próba udowodnienia słuszności hipotezy alternatywnej. Należy pamiętać, że NIE wyjaśnia to poprawnie, dlaczego nauka korzysta z takiego podejścia, ale może pomóc łatwiej zrozumieć i przyswoić tę przewrotną koncepcję.

# Błędy I i II rodzaju

Poniższa tabela przedstawia różnicę pomiędzy błędami I rodzaju (fałszywie dodatnie) i błędami II rodzaju (fałszywie ujemne).

	<b>Hipoteza zerowa jest prawdziwa</b>	<b>Hipoteza zerowa jest fałszywa</b>
Odrzucenie hipotezy zerowej	Błąd rodzaju I „fałszywie dodatni”	Wynik prawidłowy „prawdziwie dodatni”
Niepowodzenie odrzucenia hipotezy zerowej	Wynik prawidłowy „prawdziwie ujemny”	Błąd rodzaju II „fałszywie ujemny”

Nadal jest to dość niejasne, więc w celu prostszego przedstawienia tego rozumowania użyjemy wyrazistego przykładu:

- Błędy rodzaju I mogą zabić pacjenta – wyobraźmy sobie badanie, w którym nieprawidłowo określono, że standardowe leczenie nie było lepsze od nowego i w efekcie zastosowano nowe leczenie u ludzi, co spowodowało tragiczne skutki. Popełnienie błędu rodzaju I spowodowało nieprawidłowe wykrycie nieistniejącego efektu.
- Błędy rodzaju II oznaczają zmarnowanie potencjalnie cennych badań. Być może te badania mogłyby się okazać naprawdę użyteczne, ale ponieważ nie były kontynuowane, pacjenci nie odnieśli żadnych szkód. Popełnienie błędu rodzaju II spowodowało niewykrycie istniejącego efektu.

Jasne jest, że w odniesieniu do pacjentów błędy rodzaju I są poważniejsze niż błędy rodzaju II.

## Poziom istotności

Poziom istotności jest prawdopodobieństwem popełnienia błędu rodzaju I. Ma na niego wpływ rozmiar próby i „moc

statystyczna” testu.

## Moc statystyczna

„Moc” testu statystycznego jest prawdopodobieństwem, że test będzie prawidłowo prowadzić do odrzucenia hipotezy zerowej – lub innymi słowy zdolnością testu do wykrycia efektu, jeśli ten rzeczywiście istnieje. Można także powiedzieć, że „moc” testu jest prawdopodobieństwem NIEPOPEŁNIENIA błędu rodzaju II.

## Wartości P

Wartości P lub „prawdopodobieństwa” wyznaczają siłę dowodu w skali od 0 do 1. Mała wartość P (zwykle mniejsza niż 0,05 lub 5%) wskazuje, że istnieją silne dowody *przeciw* hipotezie zerowej, które mogą prowadzić do jej odrzucenia, podczas gdy duża wartość P (większa niż 0,05) wskazuje odwrotną sytuację.

## Korelacja a przyczynowość

Analizując wyniki uzyskane w badaniu klinicznym, należy pamiętać, że korelacja to nie to samo, co przyczynowość. O korelacji mówimy, jeśli dwie zmienne są w jakiś sposób połączone, jednak nie znaczy to, że jedna powoduje drugą (między obiema zmiennymi istnieje powiązanie). Przykładem może być hormonalna terapia zastępcza (HTZ) i choroba niedokrwienna serca (ChNS) – kobiety stosujące hormonalną terapię zastępczą są mniej narażone na wystąpienie choroby niedokrwiennej serca. Nie wynika to jednak z faktu stosowania HTZ, a raczej stąd, że kobiety stosujące tę terapię należą zwykle do wyższej grupy społeczno-ekonomicznej, lepiej się odżywiają i ćwiczą.

O przyczynowości mówimy, jeśli dany czynnik powoduje wynik. Czynniki przyczynowe stanowią zwykle częściową przyczynę wyniku. Rozróżnienie między korelacją a przyczynowością wymaga zebrania możliwie jak największej ilości informacji

dotyczących uczestników badania. Konieczne jest także staranne zastosowanie metodologii naukowej w trakcie projektowania badań klinicznych i ocenienie możliwego błędu systematycznego w badaniu.

## **Manipulowanie danymi**

Manipulowanie danymi polega na nieprawidłowym wybiórczym zgłaszaniu danych lub na tworzeniu fałszywych wyników. Może to mieć miejsce, jeśli na przykład dane niezgodzące się z oczekiwanymi wynikami są celowo odrzucane, aby zwiększyć odsetek wyników potwierdzających przedstawioną hipotezę. ChNSJeśli badacz usuwa obserwacje odstające (wartości znacznie większe lub mniejsze od następnego najbliższego wyniku) z uzyskanych wyników, należy sprawdzić, czy są to rzeczywiście obserwacje odstające, a nie wyniki, które różnią się od oczekiwanych lub pożądaných. Innym przykładem manipulowania danymi może być losowe wygenerowanie pełnego zestawu danych pochodzących z pomiarów zebranych od jednego pacjenta.

## **Transformacja danych**

Transformacja danych polega na zastosowaniu wzoru matematycznego do pewnych danych zebranych w czasie badania. Jest często stosowana, aby przedstawiane dane były bardziej przejrzyste i łatwiejsze do zrozumienia. Na przykład w przypadku pomiarów ekonomiki paliwowej samochodów wydajność zwykle mierzy się jako „kilometry na litr”. Jeśli jednak oceniałoby się, ile dodatkowego paliwa potrzeba, aby zwiększyć przejechany dystans, wartość ta zostałaby wyrażona w „litrach na kilometr”. W tym wypadku zastosowanie nieodpowiedniego wzoru w celu uzyskania nowego obrazu danych mogłoby wpłynąć na całościowe wyniki badania.

# Scalanie danych

Scalanie danych polega na łączeniu danych z wielu badań w celu uzyskania lepszego oglądu sytuacji. Jedną z najczęstszych form takiego działania jest metaanaliza polegająca na zestawieniu wyników z kilku opublikowanych badań w celu ich połączenia i porównania. W przypadku przeprowadzania metaanalizy należy starannie sprawdzać, czy metodologie badań są takie same, czy porównywalne. Należy brać pod uwagę wszelkie różnice w projekcie, tak aby nie występowały żadne różne podstawowe zmienne (zmienne zakłócające). Przykładem nieprawidłowego scalania danych może być łączenie danych z kilku badań na różnych gatunkach myszy w przypadku badań na zwierzętach.

A2-4.33.3-V1.1