



PRZYSTĘPNIE O STATYSTYCZNYM PODEJŚCIU DO TESTOWANIA HIPOTEZ BADAWCZYCH I SZACOWANIA LICZEBNOŚCI PRÓBY

Janusz Wątroba, StatSoft Polska Sp. z o.o.

Potrzeba dobrego rozumienia statystycznego podejścia do analizy i interpretacji wyników badań klinicznych wynika bezpośrednio z obowiązku dbałości o wiarygodność wniosków z badań, który w całości spoczywa na badaczu. Wymaga to oparcia badań na danych empirycznych, odpowiedniego sformułowania pytań lub hipotez badawczych oraz ich przełożenie na język hipotez statystycznych. Kolejną ważną kwestią jest dobór stosownego testu statystycznego oraz poprawna interpretacja jego wyników. Badacz powinien jednocześnie zdawać sobie sprawę z grożących mu błędów wnioskowania oraz odróżniać dwa ważne pojęcia: istotność statystyczną oraz istotność praktyczną (np. kliniczną). Jedną z ważnych, praktycznych korzyści, jakie daje znajomość podstaw wnioskowania statystycznego jest możliwość oszacowania liczebności jednostek, które należy uwzględnić w badaniach.

W trakcie wystąpienia w oparciu o przykładowe dane medyczne zostaną pokazane praktyczne przykłady testowania hipotezy badawczej oraz szacowania liczebności próby w programie *STATISTICA*. Dodatkowe przykłady oraz ćwiczenia praktyczne z komputerami zostaną przeprowadzone w trakcie warsztatów.

Projektowanie badań klinicznych

Poprawne zaplanowanie i przeprowadzenie badań klinicznych stanowi zazwyczaj spore wyzwanie i wymaga uwzględnienia szeregu aspektów metodologicznych, medycznych, epidemiologicznych, statystycznych, etycznych i wielu innych. Ze względu na tę złożoność pomocnym zabiegiem może okazać się oddzielenie dwóch punktów widzenia. Z jednej strony na badania kliniczne składają się konkretne elementy składowe („anatomia” badania): pytanie badawcze, plan badania, badane obiekty, pomiary i obserwacje, szacowanie wymaganej liczebności obiektów itd. Głównym celem badacza jest dobór takiej postaci elementów składowych, aby badanie było praktycznie wykonalne, pozwalało skutecznie ocenić badane efekty przy uwzględnieniu posiadanych środków. Drugi punkt widzenia uwzględnia wielorakie powiązania występujące pomiędzy elementami składowymi („fizjologia” badania). O praktycznych korzyściach z badania decyduje stopień poprawności wnioskowania w odniesieniu do tego, co zostało stwierdzone w poddanej badaniu



zbiorowości obiektów oraz w odniesieniu do zbiorowości obiektów, na które badacz chciałby uogólnić wyniki. Oczywiście wyróżnienie dwóch powyższych aspektów jest tak samo sztucznym zabiegiem jak oddzielenie budowy organizmu od jego funkcji, ale czasami umożliwia łatwiejsze zrozumienie tego złożonego procesu [4, 5].

Odpowiednie sformułowanie celów badania (w postaci stwierdzeń odnoszących się do czytelnych, zwięzłych, precyzyjnych, naukowo poprawnych i sformułowanych w sposób ilościowy pytań klinicznych) na etapie jego planowania ułatwia podjęcie niektórych decyzji związanych ze statystycznymi aspektami badania. Zazwyczaj ułatwia to przekształcenie pierwszoplanowych i drugoplanowych celów badania do postaci weryfikowalnych hipotez statystycznych, oszacowanie niezbędnej liczebności badanych jednostek (pacjentów) oraz wstępny dobór odpowiednich metod statystycznych stosowanych do opracowania wyników badania.

Bardzo ważną kwestią jest określenie zbiorowości jednostek, którą faktycznie badano, oraz zbiorowości, której ma dotyczyć wnioskowanie [3]. To decyduje o tym, czy badanie będzie miało charakter *badania opisowego* i w związku z tym będzie wymagało użycia metod opisu statystycznego (np. obliczenie podstawowych charakterystyk liczbowych rozkładów badanych cech w porównywanych grupach badanych jednostek lub proporcji jednostek, u których wystąpiło interesujące badacza zjawisko), czy też będzie to tzw. *badanie uogólniające* i będzie wymagało zastosowania metod *wnioskowania statystycznego* (np. ocena statystycznej istotności różnic wartości przeciętnych badanych cech pomiędzy dwoma populacjami pacjentów).

Dwa najczęściej stawiane statystykowi (analitykowi, konsultantowi) przez badacza pytania to: „Jak przeanalizować dane?” oraz „Ile przebadac jednostek?”. Odpowiedzi na te pytania wymagają podobnego podejścia jak w przypadku stawiania diagnozy przez lekarza. Trzeba przeprowadzić wywiad, który pozwoli osobie doradzającej badaczowi przynajmniej w części zrekonstruować proces badawczy. Chodzi głównie o to, żeby umieścić te pytania w szerszym kontekście projektu badawczego i dowiedzieć się, w jaki sposób dane otrzymano. Dzięki temu osoba współpracująca z badaczem będzie w stanie „postawić trafną diagnozę”. I nie chodzi tu o to, żeby ta wiedza była dokładnie taka sama jak wiedza badacza, bo to przecież nie jest możliwe.

Najlepiej jeśli pytania badawcze zostaną od razu postawione w języku zmiennych zależnych (wyjściowych) i niezależnych (wejściowych). Wówczas łatwo jest ocenić, ile zmiennych zależnych i niezależnych wchodzi w grę, oraz jaki jest ich charakter (skala pomiarowa).

Pytanie badawcze stanowi punkt wyjścia przy formułowaniu hipotezy badawczej. Hipoteza badawcza może być traktowana jako specyficzna wersja pytania badawczego, która podsumowuje główne elementy badania – badaną zbiorowość obiektów oraz predyktory i zmienne zależne – w postaci stanowiącej podstawę dla testów istotności statystycznej. Hipotezy nie są potrzebne w przypadku badań opisowych, które opisują rozkład badanych charakterystyk w populacji (np. badania dotyczące występowania określonego genotypu wśród pacjentów). Najkrócej rzecz ujmując, dobra hipoteza badawcza powinna być prosta, sformułowana w precyzyjny sposób oraz postawiona przed przystąpieniem do badań.



Przykładem prostej hipotezy badawczej postawionej z użyciem jednej zmiennej niezależnej i jednej zmiennej zależnej byłyby hipoteza:

Siedzący tryb życia jest powiązany ze zwiększonym ryzykiem wystąpienia białkomoczu u pacjentów z cukrzycą

Hipoteza złożona może zawierać więcej niż jedną zmienną niezależną i więcej niż jedną zmienną zależną. Poniżej przykład:

Spożycie alkoholu jest powiązane ze zwiększonym ryzykiem wystąpienia białkomoczu i neuropatii u pacjentów z cukrzycą

Złożone hipotezy nie mogą być testowane za pomocą pojedynczego testu statystycznego, lecz wymagają zazwyczaj postawienia kilku hipotez prostych.

Jak to zostało podane wyżej hipoteza badawcza powinna również zostać sformułowana w sposób jak najbardziej czytelny, tak aby można było zorientować się, jaki charakter mają brane pod uwagę zmienne (jakościowe czy ilościowe). Poniżej przykład takiej hipotezy:

Spożycie alkoholu (w mg/dzień) jest powiązane ze zwiększonym ryzykiem wystąpienia białkomoczu (>300 mg/dzień) u pacjentów z cukrzycą

Ponadto hipoteza badawcza powinna zostać postawiona już na etapie planowania badania. Pozwala to skupić uwagę badacza na realizacji pierwszoplanowego celu badań. Wstępnie ustalona pojedyncza hipoteza tworzy mocną podstawę do interpretacji wyników badania w porównaniu do kilku hipotez, które wyłaniają się na podstawie przeglądania danych. Hipotezy, które są formułowane, po przeglądaniu danych stanowią formę testowania wielokrotnego, które może prowadzić do nadinterpretacji ważności ustaleń.

Poniżej przedstawiono wyniki przykładowych analiz w przypadku zbiorowości, dla której brak informacji, czy reprezentuje ona jakąś szerszą populację. W takim przypadku należy poprzestać na statystykach opisowych (nie są to badania uogólniające).

Przykłady analizy wyników badań opisowych

Pierwszy z zaprezentowanych przykładów dotyczy badania zbiorowości 206 pacjentów ze zdiagnozowaną chorobą niedokrwienną serca [6]. U części z badanych (132 przypadki) wystąpił zawał serca. Badacz chce porównać, jak często obserwowano zawał serca u osób palących i niepalących.

W tak sformułowanym celu badania mamy do czynienia z dwiema zmiennymi typu jakościowego (nominalne). Dla otrzymania odpowiedzi na postawione pytanie należy zestawić dane w postaci tabeli dwudzielczej. W zamieszczonej poniżej tabeli znajdują się wyniki.

Podsumowująca tabela dwudzielcza: częstości obserwowane (Zawały.sta)			
Palenie	Zawał		Wiersz Razem
	nie wystąpił	wystąpił	
nie pali	26	32	58
%wiersza	44,83%	55,17%	
pali	48	100	148
%wiersza	32,43%	67,57%	
Ogół	74	132	206

Rys. 1. Tabela dwudzielcza dla porównania częstości wystąpienia zawału u palących i niepalących.

Na podstawie przedstawionych wyników można stwierdzić, że zawał serca występował częściej u osób palących. Szacowane na tej podstawie ryzyko wystąpienia zawału (liczone za pomocą wskaźnika Ryzyka względnego) u palących było o około 23% wyższe w stosunku do osób niepalących (RW=1,23).

Przyjmujemy, że dla tej samej zbiorowości badanych chcemy następnie porównać przeciętny poziom wyników pomiaru cholesterolu całkowitego oraz frakcji HDL i LDL u pacjentów, u których nie wystąpił zawał, w porównaniu do pacjentów z zawałem. Tym razem zmienne mają charakter ilościowy, w związku z tym należy obliczyć podstawowe charakterystyki liczbowe rozkładu branych pod uwagę cech w grupach osób różniących się faktem wystąpienia zawału. Przedstawiono je poniżej w tabelach.

Statystyki opisowe w grupach (Zawały.sta)						
Zawał	Cholesterol całkowity N Ważnych	Cholesterol całkowity Średnie	Cholesterol całkowity Odch.std	Cholesterol całkowity Mediana	Cholesterol całkowity Q25	Cholesterol całkowity Q75
nie wystąpił	73	222,4	34,5	218,0	204,0	245,0
wystąpił	130	262,0	35,3	260,5	239,0	289,0
Ogół	203	247,8	39,8	249,0	218,0	277,0

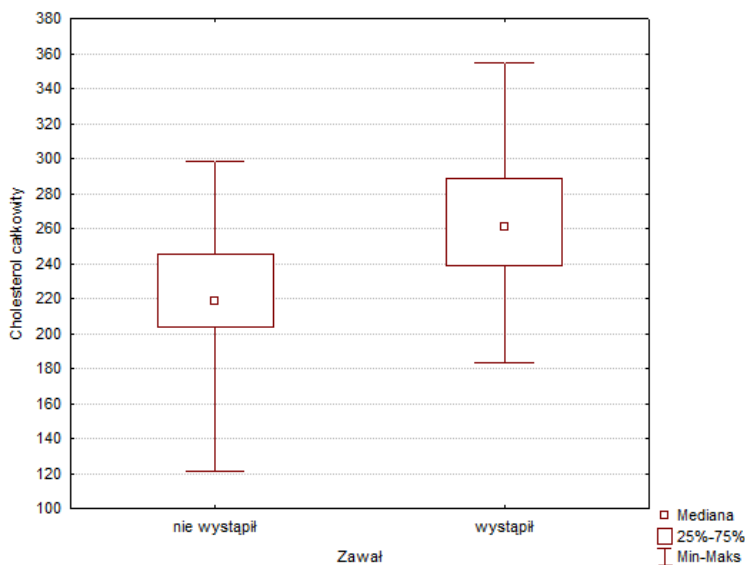
Statystyki opisowe w grupach (Zawały.sta)						
Zawał	HDL N Ważnych	HDL Średnie	HDL Odch.std	HDL Mediana	HDL Q25	HDL Q75
nie wystąpił	72	45,4	10,9	46,0	37,0	52,0
wystąpił	129	44,0	12,1	42,0	35,0	49,0
Ogół	201	44,5	11,7	43,0	37,0	50,0

Statystyki opisowe w grupach (Zawały.sta)						
Zawał	LDL N Ważnych	LDL Średnie	LDL Odch.std	LDL Mediana	LDL Q25	LDL Q75
nie wystąpił	72	143,2	33,2	140,7	120,8	168,0
wystąpił	129	183,8	37,9	185,2	154,8	210,2
Ogół	201	169,3	41,1	165,6	138,2	201,0

Rys. 2. Podstawowe charakterystyki liczbowe dla cholesterolu całkowitego, HDL i LDL w grupach osób bez zawału i z zawałem.

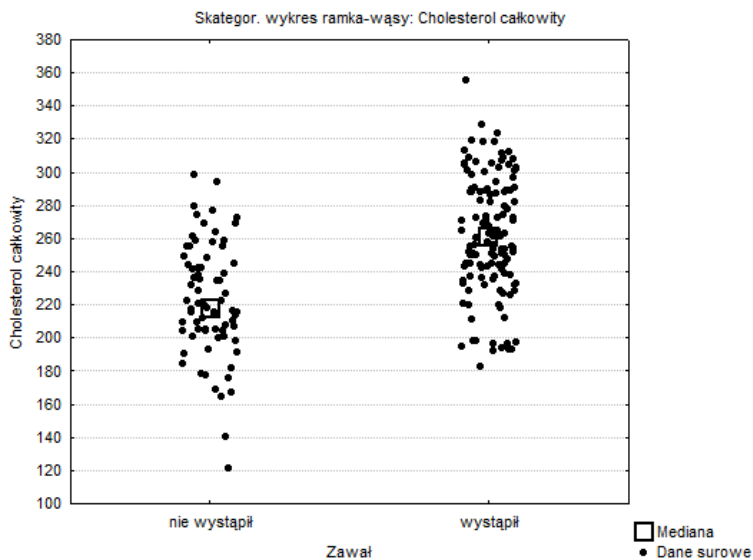
Przedstawione powyżej wyniki wskazują na występowanie dość sporych różnic w przeciętnym poziomie w przypadku cholesterolu całkowitego i LDL. W przypadku HDL różnica okazała się niewielka. Dla łatwiejszego uchwycenia występujących różnic zamieszczone

w tabeli wyniki dla cholesterolu całkowitego przedstawiono również na wykresie ramkowym oraz wykresie surowych danych.



Rys. 3. Wykres typu ramka-wąsy dla zmiennej Cholesterol całkowity.

Wykres pozwala zauważyć różnice w przeciętnym poziomie analizowanej cechy w porównywanych grupach. Jednocześnie zmienność wyników mierzona zakresem od pierwszego do trzeciego kwartyla jest w obu grupach zbliżona.



Rys. 4. Wykres surowych danych dla zmiennej Cholesterol całkowity.



Na wykresie surowych danych widać wyraźne przesunięcie obu rozkładów. Przy okazji wykres ten umożliwia sprawdzenie, czy pojedyncze obserwacje rozłożyły się równomiernie w całym zakresie wartości oraz czy występują obserwacje odstające. W opisywanym przykładzie występuje kilka takich przypadków.

W praktycznych zastosowaniach metod statystycznych w badaniach medycznych zdecydowanie większe znaczenie mają metody wnioskowania statystycznego, a wśród nich szczególnie metody weryfikacji hipotez statystycznych. Są one wykorzystywane zawsze wtedy, gdy badania mają mieć charakter uogólniający.

Korzyści z dobrej znajomości zagadnienia testowania hipotez statystycznych są rozliczne: dobrze dobrane metody i poprawnie zinterpretowane wyniki analiz statystycznych dają dobrą podstawę do interpretacji merytorycznej. Oprócz tego w przypadku różnego rodzaju publikacji (np. w czasopismach naukowych) coraz częściej obok recenzenta merytorycznego występuje również recenzent statystyczny, którego opinia ma wpływ na wydanie opracowania drukiem.

Statystyczne podejście do testowania hipotez badawczych

Dla celów testowania istotności statystycznej hipoteza badawcza musi zostać przełożona do postaci, która umożliwi odróżnienie spodziewanej różnicy pomiędzy badanymi grupami. Wymaga to postawienia dwóch hipotez wykluczających się nawzajem: zerowej oraz alternatywnej. Hipoteza zerowa zawiera stwierdzenie o braku powiązania pomiędzy zmienną niezależną i zależną w populacji. Stanowi ona formalną podstawę procedury testowania istotności statystycznej. Zakładając brak powiązania w populacji, test statystyczny umożliwia oszacowanie prawdopodobieństwa, że powiązanie zaobserwowane w badaniu jest dziełem przypadku. Hipoteza zakładająca powiązanie jest nazywana hipotezą alternatywną. Niestety nie może ona być bezpośrednio testowana, jest natomiast akceptowana domyślnie, jeśli test statystyczny prowadzi do odrzucenia hipotezy zerowej.

Kolejne potrzebne rozróżnienie to postać hipotezy alternatywnej. Hipoteza jednostronna określa kierunek powiązania pomiędzy zmienną niezależną i zależną. Z kolei hipoteza dwustronna określa tylko, że zależność występuje.

W przypadku badań, które mają charakter uogólniający, należy pamiętać, że procedura testowania istotności statystycznej dotyczy hipotezy, która została sformułowana dla populacji, której ma dotyczyć wnioskowanie. W związku z tym badacz musi brać pod uwagę możliwe błędy wnioskowania. W grę wchodzi dwa rodzaje błędów: I rodzaju (błędne odrzucenie sprawdzanej hipotezy zerowej, w sytuacji kiedy w populacji jest ona prawdziwa) oraz drugiego rodzaju (błędne przyjęcie hipotezy zerowej w sytuacji, gdy w populacji jest ona fałszywa).

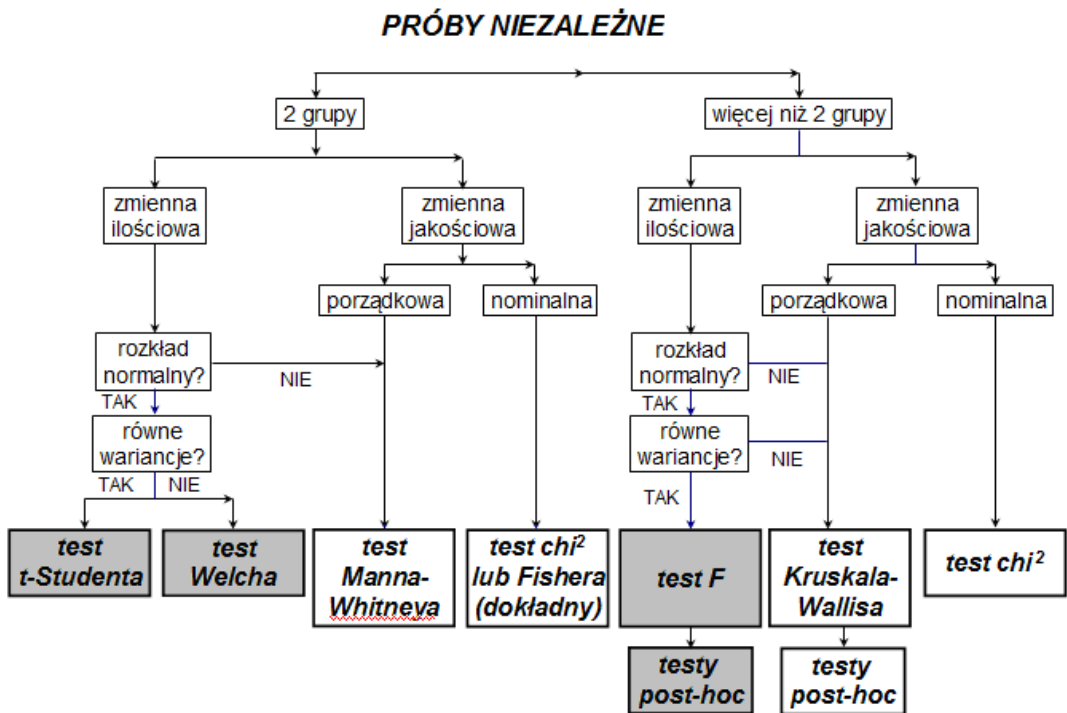
Jak to zostało już wcześniej zasygnalizowane, procedura oceny prawdziwości postawionej przez badacza hipotezy statystycznej (zerowej) polega na oszacowaniu prawdopodobieństwa, że zaobserwowany w badaniu efekt jest dziełem przypadku. Z punktu widzenia badacza oznacza to konieczność podjęcia decyzji o prawdziwości hipotezy odnoszącej się do

populacji w oparciu o wyniki otrzymane w wyniku badania częściowego. Sytuację tę ilustruje poniższa tabela.

Decyzja \ H ₀	prawdziwa	falszywa
	odrzuć	<i>błąd I rodzaju</i> α
nie odrzucać	<i>decyzja trafna</i>	<i>błąd II rodzaju</i> β

Rys. 5. Podejmowanie decyzji przy testowaniu hipotez statystycznych

Przy opracowywaniu wyników badań empirycznych najczęściej stosowane są testy istotności różnic. Wybierając odpowiedni test statystyczny, bierze się pod uwagę dwie grupy warunków. Pierwsza z nich jest związana ze sposobem, w jaki badania zostały zaplanowane (jaki schemat badań został wybrany; dla prób zależnych czy niezależnych, ile było porównywanych grup oraz na jakiej skali pomiarowej zmienne zależne były mierzone), natomiast druga dotyczy zgromadzonych danych (czy dane spełniały określone założenia, np. normalność rozkładu i jednorodność wariancji). W wielu podręcznikach prezentowane są schematy ułatwiające badaczowi wybranie odpowiedniego testu istotności.



Rys. 3. Schemat wyboru testów istotności różnic w przypadku prób niezależnych (niepowiązanych).

Zamieszczony powyżej schemat został zaczerpnięty z podręcznika do metodologii badań Brzezińskiego, przeznaczonej dla psychologów i pedagogów [2].

Przykłady analizy wyników badań uogólniających

Dla zilustrowania oceny statystycznej istotności zróżnicowania wartości przeciętnych w dwóch populacjach oraz sposobu wykorzystania powyższego schematu wykorzystamy ponownie dane 206 pacjentów ze zdiagnozowaną chorobą niedokrwienną serca [6].

Przyjmując tym razem, że badana zbiorowość pacjentów stanowi reprezentatywną próbę z populacji pacjentów ze zdiagnozowaną chorobą niedokrwienną serca, sprawdzimy, czy zróżnicowanie proporcji pacjentów niepalących z zawałem i palących z zawałem jest istotne statystycznie. Ponieważ mamy do czynienia z próbami pochodzącymi z dwóch niezależnych populacji, a zmienna zależna (wystąpienie zawału) ma charakter nominalny, na podstawie przytoczonego schematu odpowiednim testem będzie więc *test chi-kwadrat*. Jego wyniki zamieszczono poniżej.

Różnica między dwoma wskaźnikami struktury

‰ 1: .675700 N1: 148

‰ 2: .551700 N2: 58

p= .0952

Jednostronny

Dwustronny

Oblicz

Rys. 4. Wyniku testu chi-kwadrat.

Jak widać, wyniki testowania nie dają wystarczających podstaw do stwierdzenia, że proporcje przypadków zawału wśród pacjentów niepalących i palących różnią się statystycznie istotnie ($p < \alpha$).

Kolejny przykład dotyczy oceny istotności zróżnicowania przeciętnego poziomu cholesterolu całkowitego pomiędzy pacjentami, u których zawał nie wystąpił, oraz pacjentami, u których stwierdzono wystąpienie zawału. Ponownie mamy przykład z dwoma próbami pochodzącymi z dwóch niezależnych populacji, jednak tym razem zmienna zależna jest wyrażona na skali ilościowej. W związku z tym przy wyborze odpowiedniego testu należy sprawdzić założenia normalności rozkładu oraz równości wariancji (w naszym przykładzie obydwa założenia były spełnione). Odpowiednim testem w takiej sytuacji jest *test t-Studenta* w wersji dla prób niezależnych (niepowiązanych). Jego wyniki przedstawiono poniżej.

Zmienna	Testy t; Grupująca: Zawał: =(v4=1 or v4=2) (Zawały.sta)				
	Srednia nie wystąpił	Srednia wystąpił	t	df	p
Cholesterol całkowity	222.40	262.04	-7.74232	201	0.000000

Rys. 5. Wyniku testu *t-Studenta*.

Na podstawie otrzymanych wyników możemy stwierdzić, że istnieją mocne podstawy do odrzucenia hipotezy zakładającej brak różnic. Umożliwia to przeprowadzenie merytorycznej oceny występujących różnic.

Zagadnienie mocy testu oraz szacowanie odpowiedniej liczebności badanych jednostek

Przy projektowaniu badań empirycznych (zwłaszcza w medycynie) badacze coraz częściej nie ograniczają się jedynie do formułowania hipotez badawczych, których prawdziwość ma zostać oceniona, oraz przyjęcia poziomu istotności wyniku, lecz zastanawiają się także, jak duży efekt kliniczny uznać za zadowalający oraz jaka ma być moc testu [1, 4].

Moc testu statystycznego to $1-\beta$, czyli prawdopodobieństwo odrzucenia hipotezy zerowej, gdy jest ona fałszywa. W praktyce moc testu oznacza prawdopodobieństwo podjęcia na podstawie testu decyzji o istnieniu efektu, gdy on rzeczywiście istnieje. Moc testu zależy od konkretnych ustawień testu (w większości testów od obszaru krytycznego statystyki testowej, który to z kolei zależy od liczebności próby, rozproszenia wyników w próbie, przyjętego poziomu istotności α i przyjętych w teście założeń) oraz od wielkości efektu, którego wykrycie interesuje badacza. Podobnie jak β , moc można policzyć tylko dla konkretnej alternatywy, czyli dla konkretnej wielkości efektu.

Analiza mocy testu daje badaczowi wymierne korzyści, przede wszystkim umożliwiając optymalne zaplanowanie odpowiedniej liczebności jednostek, które powinny zostać uwzględnione w badaniach. Szacowanie liczebności próby stanowi ważny etap planowania eksperymentu, gdyż próba może być za duża lub za mała. Jeśli próba jest zbyt mała eksperyment może mieć za małą dokładność albo dawać wnioski ze zbyt dużym błędem. Jeśli próba będzie zbyt duża, badania wymagać będą dużego nakładu środków i czasu, zazwyczaj przy minimalnym zysku statystycznym.

W praktycznych zastosowaniach przy szacowaniu niezbędnej liczebności próby badacz musi na wstępie określić pewne parametry lub wykorzystać wyniki wcześniej przeprowadzonych badań. Jak to zostało wcześniej zasygnalizowane, należy ustalić minimalną akceptowaną przez badacza wielkość efektu, która może zostać potraktowana jako znacząca klinicznie, określić poziom istotności testu, który będzie wykorzystywany przy analizie wyników badania oraz podać wymaganą moc badania.

W przykładach jeszcze raz wykorzystamy dane 206 pacjentów ze zdiagnozowaną chorobą niedokrwinną serca [6]. W pierwszej kolejności sprawdzimy, jaka liczebność próby byłaby potrzebna, aby różnica w proporcji pacjentów niepalących z zawałem i pacjentów palących z zawałem była statystycznie istotna przy przyjęciu poziomu istotności $\alpha=0,05$ oraz mocy testu $1-\beta=0,80$. Poniżej tabela prezentująca wyniki.

Jak widać (na rys. 7), aby można było różnicę w ocenianych proporcjach uznać za statystycznie istotną, potrzebne byłoby przebadanie 241 pacjentów (przy dodatkowym przyjęciu jednakowej liczebności pacjentów w obu porównywanych próbach).



	Liczność próby (Zawały.sta) Dwie frakcje, test Z H0: $Pi1 = Pi2$
	Wartość
Frakcja w populacji $Pi1$	0,5517
Frakcja w populacji $Pi2$	0,6757
Prawdop. bł. I rodzaju (Alfa)	0,0500
Moc docelowa	0,8000
Moc (bez poprawki)	0,8002
Liczność próby $N1$	241,0000
Liczność próby $N2$	241,0000

Rys. 7. Wyniku szacowania wymaganej liczebności próby.

W drugim przykładzie sprawdzimy, jaka liczebność badanych wystarczyłaby dla stwierdzenia statystycznie istotnej różnicy w przeciętnych poziomach cholesterolu całkowitego u pacjentów z zawałem i bez zawału. Przy szacowaniu liczebności przyjmujemy wyniki otrzymane w jednym z poprzednich przykładów.

	Liczność próby (Zawały.sta) Dwie średnie, test t, próby niezależne H0: $Mi1 = Mi2$
	Wartość
Średnia populacyjna $Mi1$	222,40
Średnia populacyjna $Mi2$	262,00
Odch. std. w populacji (Sigma)	35,00
Efekt standaryzowany (Es)	-1,13
Prawdop. bł. I rodzaju (Alfa)	0,05
Wartość krytyczna t	2,06
Moc docelowa	0,80
Moc dla wymaganej liczebności próby N	0,82
Wymagane N (w grupie)	14,00

Rys. 8. Wyniku szacowania wymaganej liczebności próby.

Okazało się, że przy przyjętej wielkości efektu zróżnicowania wystarczyłoby objąć badanymi grupy pacjentów o liczebności 14.

Podsumowanie

W praktycznych zastosowaniach statystyki najczęściej zadawane statystykowi pytania dotyczą wyboru odpowiednich metod opracowania danych pod kątem postawionych przez badacza pytań lub hipotez oraz doboru odpowiedniej liczebności jednostek, które mają zostać objęte badaniami.

Ze strony statystyka wymaga to przynajmniej częściowego odtworzenia procesu badawczego dla dobrego zrozumienia postawionych celów badawczych oraz przełożenia pytań lub hipotez badawczych do postaci hipotez statystycznych.

W wielu sytuacjach wybór stosownych narzędzi statystycznych dokonywany jest dwukrotnie. Wstępny wybór odbywa się na etapie planowania badań i jest potrzebny do



oszacowania liczebności badanych, którzy mają zostać objęci badaniami. Z kolei po zrealizowaniu badań i zebraniu danych dobór metod ich opracowania jest weryfikowany w oparciu o empiryczne dane.

Przy merytorycznej interpretacji wyników należy wyraźnie rozróżniać wnioski statystyczne od wniosków badawczych.

Oszacowanie liczebności jednostek, które mają zostać uwzględnione w badaniu, wymaga ustalenia wielkości spodziewanego efektu, przyjęcia poziomu istotności i określenia mocy testu.

Literatura

1. Bausell R. B., Li Y.-F. (2002) Power Analysis for Experimental Research. A Practical Guide for the Biological, Medical and Social Science, Cambridge University Press.
2. Brzeziński J. (1996) Metodologia badań psychologicznych, PWN Warszawa
3. Chow, S.C., Liu, J. P. (2004) Design and Analysis of Clinical Trials. Concepts and Methodologies, (2nd ed.), John Wiley and Sons, NJ: Erlbaum.
4. Designing Clinical Research (2007) ed. S.B. Hulley at all. (3rd ed.), Lippincott Williams & Wilkins.
5. Podstawy EBM, czyli medycyny opartej na danych naukowych dla lekarzy i studentów medycyny, pod red. P. Gajewskiego, R. Jaeschke i J. Brożka, Medycyna Praktyczna, Kraków 2008.
6. Watała C. (2002) Biostatystyka – wykorzystanie metod statystycznych w pracy badawczej w naukach biomedycznych, α -medica Press.