



## ZROZUMIEĆ DANE I WYNIKI (NA PRZYKŁADACH BADAŃ EPIDEMIOLOGICZNYCH)

*Ewa Kawalec, Zakład Epidemiologii i Badań Populacyjnych, Instytut Zdrowia Publicznego,  
Wydział Ochrony Zdrowia, Collegium Medicum Uniwersytetu Jagiellońskiego*

### Wprowadzenie

Co mieści się w sformułowaniu „zrozumieć dane i wyniki”? Będzie ono obejmowało wyjaśnienie zagadnień związanych z przeprowadzeniem badania statystycznego w naukach medycznych, ze szczególnym uwzględnieniem gromadzenia danych, sposobów ich weryfikacji oraz interpretacji wyników. Gromadzenie danych w sposób prawidłowy i rzetelny, a także poprawna interpretacja wyników są tak samo ważną częścią procesu badawczego, jak umiejętność przeprowadzenia właściwych analiz. Od sposobu zebrania danych i ich jakości, a także od wnikliwej analizy uzyskanych wyników zależeć będzie jakość badania naukowego. Dlatego tak ważne jest, aby badacz zrozumiał problemy związane z projektowaniem badania i jego praktycznymi konsekwencjami. W prezentacji wykorzystane zostaną wyniki badań epidemiologicznych, które realizowane były w Zakładzie Epidemiologii i Badań Populacyjnych Instytutu Zdrowia Publicznego CM UJ w Krakowie. Wszystkie analizy statystyczne wykonywane były z wykorzystaniem pakietu statystycznego *STATISTICA*.

### Projektowanie badania statystycznego

Statystyka pełni w badaniach medycznych szczególną rolę, gdyż nie tylko pozwala na gromadzenie danych, ich grupowanie, prezentowanie, analizę i interpretację wyników, ale przede wszystkim bierze pod uwagę wewnętrzną zmienność nieodłącznie związaną z procesami biologicznymi. Medyczny projekt badawczy (np. badanie epidemiologiczne) niejednokrotnie wymaga zastosowania analizy danych oraz użycia pakietu statystycznego. Z pomocą przyjdzie tu zaplanowanie odpowiedniego badania statystycznego.

Etapy projektowania badania statystycznego:

1. Przygotowanie badania.
2. Gromadzenie danych i ich weryfikacja.
3. Analizy statystyczne.



4. Prezentacja wyników.
5. Interpretacja wyników.
6. Publikacja wyników.

## Przygotowanie badania

Przygotowanie badania obejmuje:

- ◆ ocenę aktualnego stanu wiedzy,
- ◆ powołanie zespołu badawczego, w którego skład powinni wchodzić lekarze, pielęgniarki, laboranci, personel pomocniczy, a także, co bardzo często jest pomijane na tym etapie, statystycy – którzy projektują analizy, dobierają próbę do badania,
- ◆ ustalenie celów (ogólnych i szczegółowych) badania,
- ◆ postawienie hipotez badawczych,
- ◆ postawienie hipotez statystycznych,
- ◆ ustalenie listy cech badania,
- ◆ opracowanie kwestionariuszy badawczych,
- ◆ opracowanie instrukcji wypełniania kwestionariuszy,
- ◆ przeprowadzenie badania pilotażowego,
- ◆ korektę i druk kwestionariuszy oraz instrukcji.

## Gromadzenie danych i ich weryfikacja

Źródłem danych o uczestnikach badania są m.in.:

- ◆ kwestionariusz badawczy,
- ◆ wywiad lekarski,
- ◆ wyniki badań laboratoryjnych,
- ◆ wyniki badań EKG, Rtg, MRI, CT,
- ◆ karty szpitalne
- ◆ karty zgonów,
- ◆ roczniki statystyczne i inne publikacje.

Szczególne uwagi należy zwrócić na opracowanie kwestionariusza badawczego, który powinien być przygotowany do przetwarzania komputerowego oraz powinien zawierać:

- ◆ na stronie tytułowej nazwę badania, numer identyfikacyjny badanego, datę badania,
- ◆ pytania niezbędne do realizacji celów badania,
- ◆ pytania w postaci zamkniętej,



- ◆ pytania sformułowane jasno i precyzyjnie z możliwością odpowiedzi „nie wiem”, „nie pamiętam”.

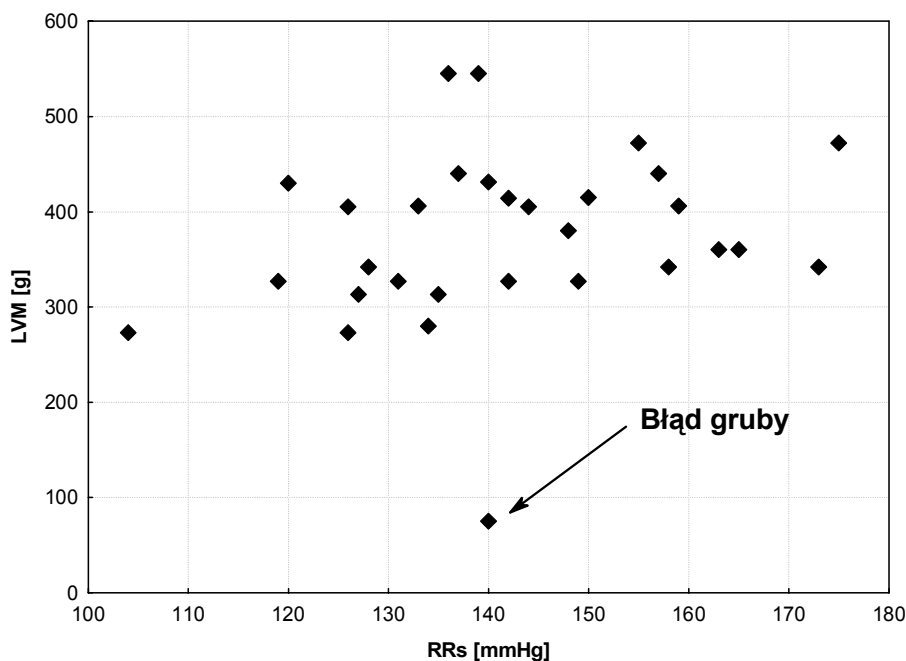
Ponadto pytania powinny być ponumerowane. Zaleca się stosowanie pytań dychotomicznych (z odpowiedziami tak – nie).

Pozyskiwanie danych wymaga odpowiedniego przygotowania, które obejmuje m.in.:

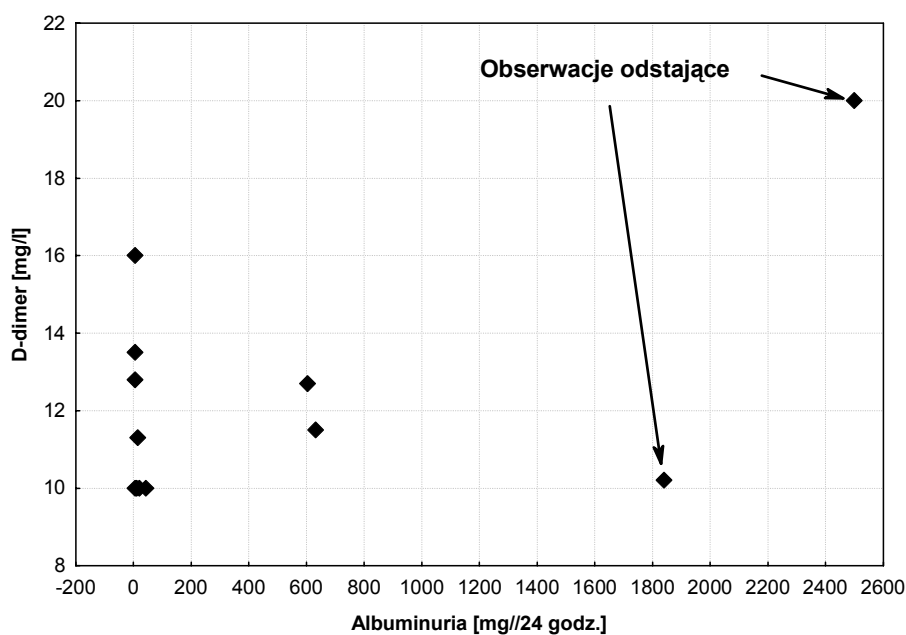
- ◆ organizację zespołów badawczych,
- ◆ dobór próby do badania (w tym ustalenie, w zależności od typu badania, wielkości próby, sposobu losowania),
- ◆ zgromadzenie odpowiedniego sprzętu badawczego, określenie czasu i miejsca badania,
- ◆ szkolenie personelu badawczego,
- ◆ zorganizowanie systemu kontroli uzyskiwanych danych.

Przed przystąpieniem do analiz statystycznych należy bardzo dokładnie przyjrzeć się zebranym danym. Wymagają one bowiem naszej szczególnej troski i koniecznie trzeba im poświęcić sporo czasu. Nasza omylność, zawodność urządzeń pomiarowych, a nierzadko brak dokładności i rzetelności w przeprowadzeniu wywiadu lub zapisie danych w bazie komputerowej są źródłem błędów. W celu ich wykrycia można zastosować metodę podwójnego kodowania, która polega na wprowadzaniu danych do bazy przez dwie niezależne osoby. Zbiory te są porównywane, a niezgodności wyjaśniane i korygowane. Metoda ta jest co prawda czasochłonna i kosztowna, ale efektywna. Inne metody pozwalające na zidentyfikowanie błędów, to m. in.:

- ◆ ustalenie zakresu dopuszczalnych wartości (np. dla BMI: min=16 kg/m<sup>2</sup>, max=45 kg/m<sup>2</sup>),
- ◆ zbadanie powiązań logicznych (np. wartość ciśnienia skurczowego powinna być większa niż rozkurczowego),
- ◆ zbadanie spójności danych (np. poziom wykształcenia i liczba lat spędzonych w szkole),
- ◆ sprawdzenie wartości minimalnej i maksymalnej (np. błędne wartości dla minimalnej masy ciała=0 kg, maksymalnego tętna=400 ud./min),
- ◆ ustalenie sposobu zapisu braków danych (np. spacja lub 999 dla braku informacji o masie ciała),
- ◆ przyglądnięcie się obserwacjom odstającym oraz ustalenie sposobu postępowania z nimi. Obserwacja odstająca to wartość badanej cechy znacznie odbiegająca od wyznaczonego zakresu wartości, leżąca poza tzw. ciałem danych. Źródłem takiej obserwacji może być błąd (np. wprowadzania do bazy danych) – wtedy jest to tzw. błąd gruby, ale może też to być wartość rzeczywista, znacznie przekraczająca normę (np. pomiar OB po 2 godz. wynoszący 120). Błąd gruby można skorygować jedynie w oparciu o dane źródłowe. Jeśli dostęp do nich nie jest możliwy, to obserwacja traktowana jest jako brak danych, natomiast rzeczywiste obserwacje odstające wymagają szczególnej uwagi przy opisie danych i stosowaniu analiz statystycznych. Bardzo użyteczną metodą zidentyfikowania obserwacji odstających jest narysowanie wykresu punktowego rozrzutu (ang. *scatter-plot*) (rys. 1 i rys. 2).



Rys. 1. Wykres rozrzutu: RRs [mmHg] vs LVM [g].



Rys. 2. Wykres rozrzutu: albuminuria [mg/24 godz.] vs D-dimer [mg/l].

Na rys. 1 zaznaczono błąd grubý, który wymaga skorygowania (u badanego, którego ciśnienie skurczowe krwi (RRs) = 140 mmHg, masa lewej komory serca (LVM) wyniosła ok. 80 g, co jest niemożliwe). Na rys. 2 wartości zarówno albuminurii, jak i D-dimer znacznie przekraczają ustalone przez badacza zakresy wartości. Do prowadzącego badanie należy podjęcie decyzji, co do dalszego postępowania. Niezależnie od tego, czy obserwacje te pozostaną w bazie danych, czy nie (będą potraktowane jako braki danych) należy liczyć się z konsekwencjami przeprowadzanych analiz. Zarówno



w jednym, jak i drugim przypadku należy, przy omawianiu wyników analiz, szczegółowo opisać i uzasadnić sposób postępowania.

- ◆ zbudowanie dla wszystkich cech ilościowych histogramów i przeprowadzenie statystyki opisowej.

## Typy danych

Zanim przejdziemy do statystyki opisowej, należy określić typ danej. Podstawowy podział obejmuje:

- ◆ dane ilościowe (mieralne) – dane, które uzyskuje się w wyniku pomiaru i wyraża w skali liczbowej. Mogą być ciągłe, gdy przyjmują wartości z pewnego przedziału liczbowego (można je podać z dowolną dokładnością, np. wiek, wzrost, stężenie cholesterolu całkowitego we krwi) lub dyskretne, gdy przyjmują wartości liczb całkowitych (np. liczba zachorowań na gripę w ciągu roku, częstość akcji serca, liczba dzieci w rodzinie),
- ◆ dane jakościowe (niemierzalne) – określane są jedynie przez nazwę. Mogą być podawane w skali nominalnej (np. płeć, stan cywilny, grupa krwi) lub porządkowej, gdy oprócz nazwy możemy ustalić naturalny porządek (np. poziom wykształcenia, samoocena stanu zdrowia).

## Statystyka opisowa

Łatwy dostęp do pakietów statystycznych jest pokusą do przeskoku od razu po zebraniu danych do analiz. Należy tego unikać. Błędy popełnione w trakcie zbierania i wprowadzania danych do baz komputerowych, w zwielokrotnionym natężeniu, ujawniają się w końcowej fazie badania i są najczęściej źródłem fałszywych wyników. A zatem krokiem pośrednim, ale niezbędnym, jest „zaprzyjaźnienie” się z danymi i zrozumienie ich natury poprzez charakterystykę i prezentację, co stanowi bezcenną pomoc w „wyłapywaniu” błędów i pierwszy krok w zaplanowaniu odpowiednich analiz.

Dla danych ilościowych charakterystyka obejmuje podanie ogólnej liczby badanych ( $N$ ), liczby obserwacji ważnych ( $n$ ) oraz średniej arytmetycznej ( $\bar{x}$ ) i odchylenia standardowego ( $sd$ ) w przypadku cech o niewielkim współczynniku skośności (według różnych autorów uznaje się współczynnik skośności jako mały, gdy co do wartości bezwzględnej jest mniejszy od 1,5 (lub 2), albo mediany (i ewentualnie pierwszego oraz trzeciego kwartyła), gdy współczynnik skośności jest duży ( $|\text{wsp. skośności}| \geq 1,5$  (lub 2); rozkład badanej cechy wykazuje wtedy silną asymetrię). Podawanie średniej arytmetycznej w przypadku rozkładów silnie skośnych jest nieuprawnione, gdyż dla tych rozkładów nie jest ona rzetelną miarą.



## Przykład 1.

W wielośrodkowym badaniu CASCADE oceniającym zaburzenia funkcji poznawczych u osób w wieku 64–78 lat przebadano próbę 882 osób (367 mężczyzn i 515 kobiet) wylosowaną z populacji mieszkańców byłego województwa tarnobrzesckiego. Rys. 3 przedstawia statystykę opisową kilku wybranych zmiennych: MMSE – wynik testu oceniającego zaburzenia funkcji poznawczych [0 do 30 punktów], HR – częstość akcji serca [liczba uderzeń/min], TG – stężenie trójglicerydów we krwi [0,40 mmol/l – 10,00 mmol/l], RRs – ciśnienie skurczowe krwi [60 mmHg – 280 mmHg], BMI – wskaźnik masy ciała [15,00 kg/m<sup>2</sup> – 48,00 kg/m<sup>2</sup>].

Zmienna	Statystyki opisowe (przykład 1.STA) Włącz warunek: płeć='mężczyźni'							
	N ważnych	Średnia	Mediana	Min	Max	Odch.Std.	Skośność	Kurtoza
<b>MMSE</b>	367	25,1	26	2	30	4,04	-1,6287	4,6935
<b>HR</b>	367	73,8	72	0	104	9,80	-0,7353	8,6030
<b>TG</b>	362	1,33	1,10	0,43	8,39	0,9018	3,5571	20,3626
<b>RRs</b>	361	135,4	131	90	219	19,37	0,9034	1,5896
<b>BMI</b>	361	25,38	25,0	15,40	45,80	3,761	0,7759	1,8428

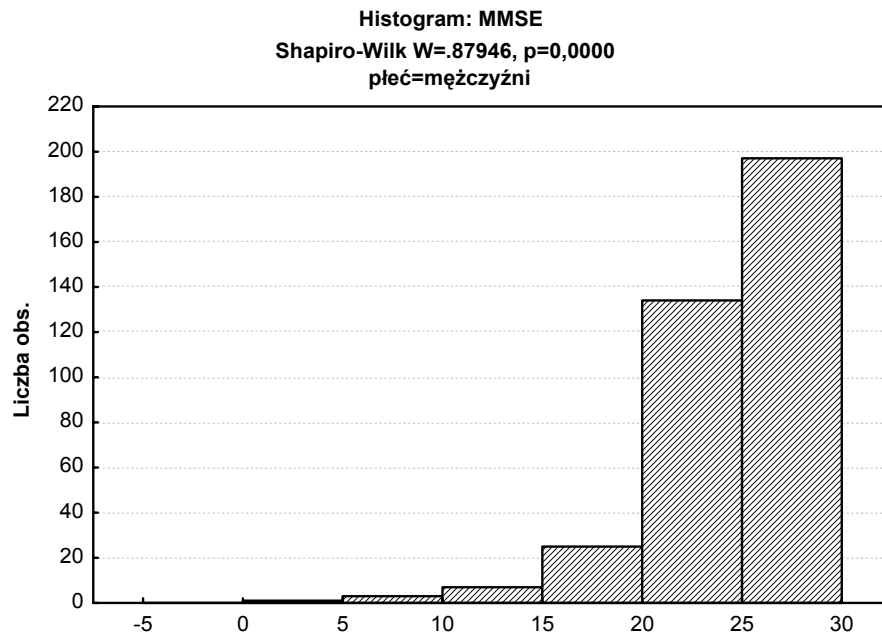
Rys. 3. Statystyka opisowa wybranych zmiennych.

Na podstawie powyższego wydruku budowana jest wynikowa tabela charakteryzująca badaną grupę pod względem poszczególnych cech (tabela 1).

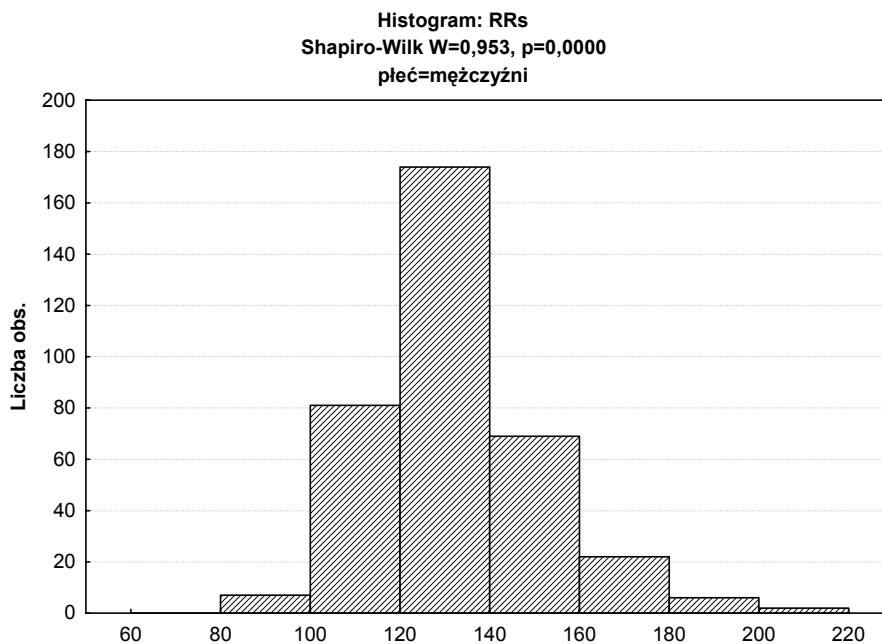
Tabela 1. Charakterystyka badanej grupy.

Badana cecha	Mężczyźni N = 367				Komentarz
	n	x	sd	Me	
MMSE	367			26	duża skośność i kurtoza
HR [ud./min]					błąd danych → weryfikacja
TG [mmol/l]	362			1,33	duża skośność i kurtoza
RRs [mmHg]	361	135,4	19,37		mała skośność i kurtoza
BMI [kg/m <sup>2</sup> ]	361	25,38	3,761		mała skośność i kurtoza

Uzupełnieniem charakterystyki liczbowej danych ilościowych są histogramy, przedstawiające rozkład badanej cechy (rys. 4 i rys. 5). Na ich podstawie możemy ocenić, czy rozkład jest w przybliżeniu normalny, czy też nie.



Rys. 4. Rozkład wyników testu MMSE.



Rys. 5. Rozkład ciśnienia skurczowego krwi.

Charakterystyka danych jakościowych obejmuje podanie ogólnej liczby badanych (N), liczbę jednostek wyróżnionych (n) oraz procent =  $(n/N) \cdot 100$ .

#### Przykład 2.

W badaniu CASCADE, przytoczonym w przykładzie 1, przebadano próbę losową 882 osób – 367 mężczyzn i 515 kobiet. Rys. 6 przedstawia tabelę dwudzielczą, która dostarcza informacji o rozpowszechnieniu nadciśnienia wśród mężczyzn i kobiet w starszym wieku.



Podsumowująca tabela dwudzielcza			
nadciśnienie	pleć		Wiersz Razem
	mężczyźni	kobiety	
<b>nie występuje</b>	172	153	325
% z kolumny	48,87%	29,71%	
% z wiersza	52,92%	47,08%	
% z całości	19,50%	17,35%	36,85%
<b>występuje</b>	195	362	557
% z kolumny	53,13%	70,29%	
% z wiersza	35,01%	64,99%	
% z całości	22,11%	41,04%	63,15%
Ogół	367	515	882
% z całości	41,61%	58,39%	100,00%

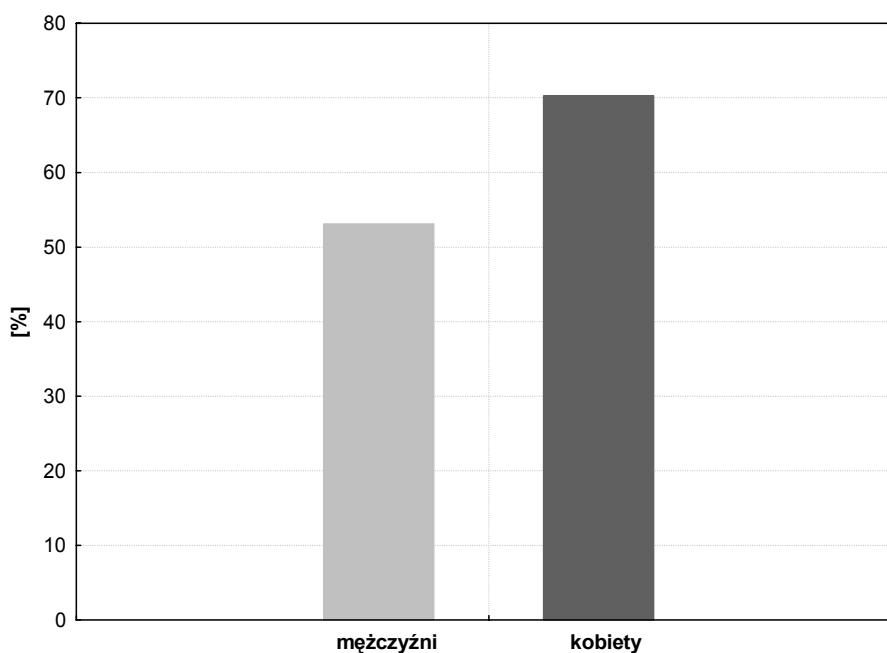
Rys. 6. Tabela dwudzielcza: nadciśnienie vs płeć.

Na podstawie powyższego wydruku konstruowana jest tabela wynikowa podająca charakterystykę badanej próby pod względem rozpowszechnienia nadciśnienia (tabela 2).

Tabela 2. Rozpowszechnienie nadciśnienia wśród mężczyzn i kobiet.

Badana cecha	Mężczyźni N = 367		Kobiety N = 515	
	n	%	n	%
Występowanie nadciśnienia	195	53,1	362	70,3

Dopełnieniem charakterystyki zmiennej jakościowej, jaką jest w tym przypadku występowanie nadciśnienia, może być wykres słupkowy (rys. 7).



Rys. 7. Częstość występowania nadciśnienia u mężczyzn i kobiet.



## Analizy statystyczne

Kolejnym etapem badania statystycznego jest przeprowadzenie analiz statystycznych. Temat ten wykracza poza zakres tego opracowania, jednak warto zwrócić uwagę na kryteria doboru właściwych analiz. Wybór analizy lub testu statystycznego zależy m.in. od:

- ◆ typu zmiennej i skali pomiarowej (porządkowa, nominalna),
- ◆ typu rozkładu (np. czy zmienna pochodzi z populacji o rozkładzie normalnym),
- ◆ liczby porównywanych grup,
- ◆ liczebności grup,
- ◆ modelu badania (grupy niezależne, powiązane),
- ◆ spełnienia wymaganych założeń.

### Przykład 3.

*„Materiał i metody – ...Stopień otyłości określano na podstawie względnej masy ciała (BMI), rozmieszczenie tkanki tłuszczowej – na podstawie stosunku obwodu pasa do obwodu bioder (WHR – waist-hip ratio). Ocenę współzależności między badanymi parametrami przeprowadzono, opierając się na współczynnikach korelacji  $r$  Pearsona...*

*Wyniki – ...Stwierdzono dodatnią korelację WHR z odpowiedzią glukozową ( $r=0,43$ ;  $p<0,001$ ) i insulinową ( $r=0,31$ ;  $p<0,01$ ) oraz stężeniem trójglicerydów ( $r=0,29$ ;  $p<0,01$ )...*”

Powyższy cytat pochodzi z abstraktu zamieszczonego w „Medycynie Metabolicznej”, t. I, 3-4/97, str. 49. Z akapitu dotyczącego metod wiemy, że zastosowano analizę korelacji Pearsona, która wymaga założenia o normalności rozkładu obu badanych cech, a rozkład glukozy, insuliny i stężenia trójglicerydów nie jest zgodny z rozkładem normalnym. Należało zatem zastosować analizę korelacji Spearmana.

## Prezentacja wyników

Zwieńczeniem analiz statystycznych jest prezentacja i interpretacja wyników. Mogą one być opisane słownie albo przedstawione w postaci tabel lub wykresów.

Jeśli porównywane są średnie wartości badanych cech pomiędzy dwiema (lub więcej) grupami, podajemy, dla każdej z grup, ogólną liczbę przebadanych osób (N), liczbę ważnych obserwacji (n), średnią arytmetyczną ( $\bar{x}$ ), odchylenie standardowe (sd), w przypadku rozkładu silnie skośnego – medianę (Me) oraz poziom krytyczny testu  $p$ .

### Przykład 4.

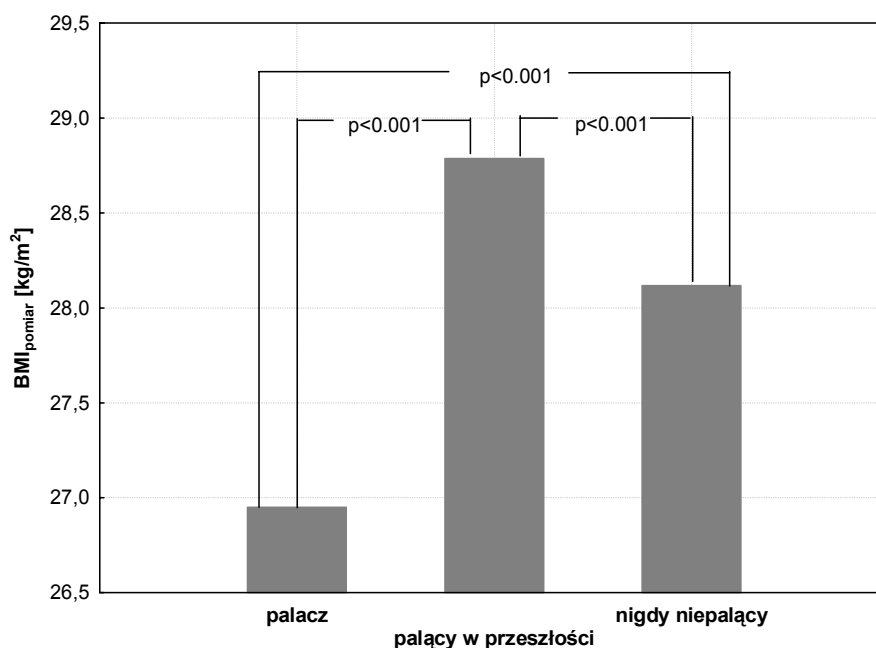
W wieloośrodkowym badaniu HAPIEE, którego celem było poznanie wpływu czynników społeczno-ekonomicznych i psychologicznych oraz stylu życia na stan zdrowia w państwach Europy Środkowo-Wschodniej, wylosowano z populacji stałych mieszkańców Krakowa próbę 10728 osób wieku 45-69 lat. Ze względu na braki danych w ocenie



i pomiarach wzrostu i masy ciała do analiz, dotyczących wskaźnika masy ciała (BMI), włączono 9204 osób (49% mężczyzn i 51% kobiet). Analizie poddano m.in. wskaźnik masy ciała BMI [kg/m<sup>2</sup>] wyliczony na podstawie samooceny masy ciała i wzrostu (BMI<sub>ocena</sub>) oraz pomiaru (BMI<sub>pomiar</sub>) u mężczyzn, w poszczególnych kategoriach palenia papierosów (palący, palący w przeszłości, nigdy niepalący). Stwierdzono, na podstawie przeprowadzonej analizy wariancji (przy spełnionych założeniach normalności rozkładu BMI w każdej z grup i jednorodności wariancji), że poszczególne grupy istotnie statystycznie różniły się średnią wartością BMI<sub>pomiar</sub> (tabela 3). Ponadto, po przeprowadzeniu analizy *post hoc* Tukeya, okazało się, że bardzo silna istotna różnica występowała pomiędzy wszystkimi grupami (rys. 8).

Tabela 3. Średnie wartości i odchylenia standardowe BMI<sub>pomiar</sub> w kategoriach palenia papierosów.

Badana cecha	Palacze N = 1533			Palący w przeszłości N = 1655			Nigdy niepalący N = 1296			p
	n	$\bar{x}$	sd	n	$\bar{x}$	sd	n	$\bar{x}$	sd	
BMI <sub>pomiar</sub> [kg/m <sup>2</sup> ]	1533	26,95	4,284	1655	28,79	3,762	1296	28,12	3,822	< 0,001



Rys. 8. Średnie wartości BMI<sub>pomiar</sub> [kg/m<sup>2</sup>] w kategoriach palenia papierosów u mężczyzn.

Gdy porównywane są częstości między dwiema (lub więcej) grupami, to dla każdej z grup podajemy ogólną liczbę badanych (N), liczbę osób wyróżnionych (n), % oraz *p*.

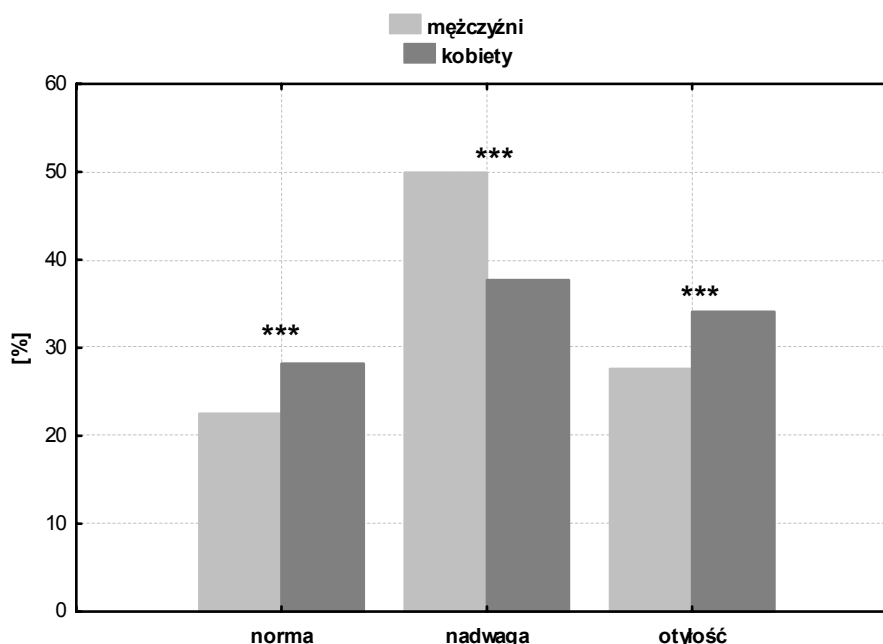
#### Przykład 5.

W badaniu opisanym w przykładzie 4 oceniano również rozpowszechnienie występowania nadwagi i otyłości u mężczyzn i kobiet na podstawie BMI uzyskanego z pomiaru masy ciała i wzrostu (BMI<sub>pomiar</sub> [kg/m<sup>2</sup>]), przy czym za normę przyjęto BMI < 25 kg/m<sup>2</sup>, nadwagę: 25 ≤ BMI < 30 kg/m<sup>2</sup>, otyłość BMI ≥ 30 kg/m<sup>2</sup> (tabela 4, rys. 9). Oceny dokonano, stosując

test  $\chi^2$ . Stwierdzono istotne statystycznie różnice w rozpowszechnieniu nadwagi i otyłości pomiędzy mężczyznami i kobietami. Mężczyźni częściej mają nadwagę, a kobiety częściej niż mężczyźni są otyłe.

Tabela 4. Rozkład częstości występowania kategorii masy ciała u mężczyzn i kobiet.

Kategorie BMI <sub>pomiar</sub>	Mężczyźni N = 4485		Kobiety N = 4719		p
	n	%	n	%	
norma	1011	22,5	1332	28,2	< 0,001
nadwaga	2238	49,9	1777	37,7	
otyłość	1236	27,6	1610	34,1	



Rys. 9. Rozpowszechnienie należnej masy ciała, nadwagi i otyłości u mężczyzn i kobiet.

Wyniki analizy korelacji podawane są w postaci liczby obserwacji (n), współczynnika korelacji (r) oraz p.

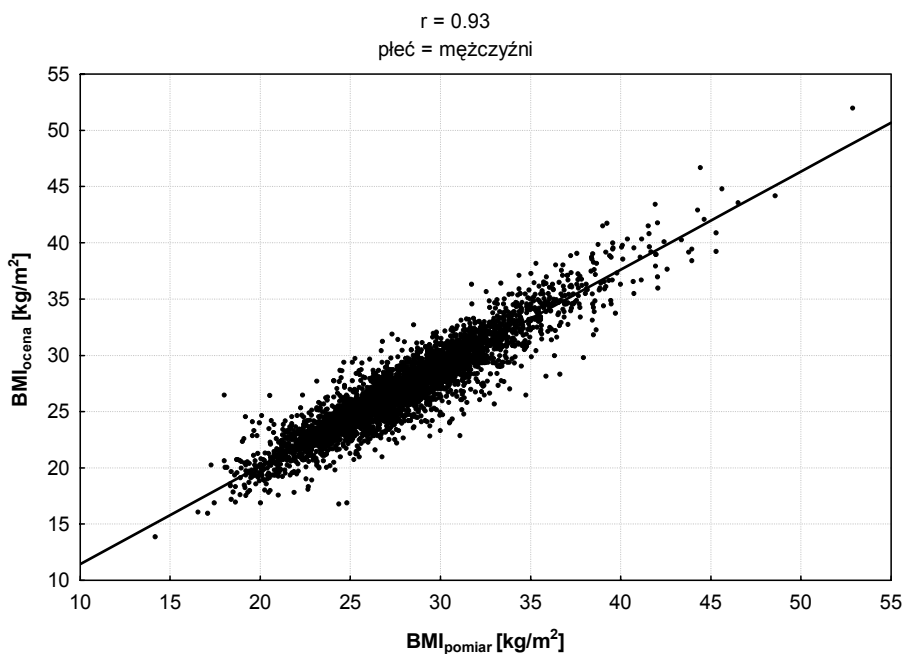
#### Przykład 6.

W badaniu z przykładu 4 oceniano także związek pomiędzy wskaźnikiem BMI uzyskanym z pomiaru (BMI<sub>pomiar</sub> [kg/m<sup>2</sup>]) i samooceny (BMI<sub>ocena</sub> [kg/m<sup>2</sup>]) u mężczyzn i kobiet. Wobec normalności rozkładów BMI zastosowano analizę korelacji Pearsona. Wyniki zestawiono w tabeli 5 oraz na rys. 10 i rys. 11, stwierdzając bardzo silny istotny statystycznie związek pomiędzy badanymi cechami, zarówno u mężczyzn, jak i u kobiet.

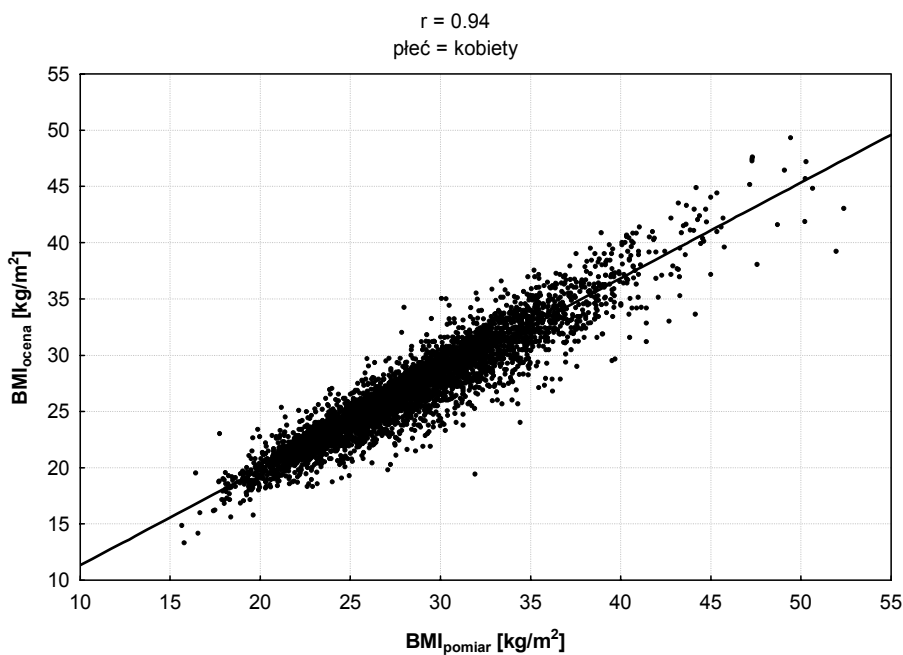


Tabela 5. Współczynniki korelacji pomiędzy BMI bazującym na pomiarze i na samoocenie, u mężczyzn i kobiet.

Badana cecha	Mężczyźni N=4485		Kobiety N=4719	
	r	p	r	P
BMI <sub>pomiar</sub> vs BMI <sub>ocena</sub> [kg/m <sup>2</sup> ]	0,93	< 0,001	0,94	< 0,001



Rys. 10. Związek pomiędzy BMI<sub>pomiar</sub> vs BMI<sub>ocena</sub> u mężczyzn.



Rys. 11. Związek pomiędzy BMI<sub>pomiar</sub> vs BMI<sub>ocena</sub> u kobiet.



Wyniki analizy regresji podawane są w postaci współczynnika regresji (b), błędu standardowego dla współczynnika regresji (se), współczynnika determinacji  $R^2$  oraz  $p$ .

Przykład 7.

W badaniu POL-MONICA Kraków, mającym na celu określenie czynników ryzyka chorób układu krążenia (ChUK) i zbadanie trendów umieralności, śmiertelności i zachorowalności z powodu ChUK w populacji mieszkańców byłego województwa tarnobrzeskiego w wieku 25-64 lat, oceniano m.in. związek pomiędzy wiekiem, wskaźnikiem masy ciała (BMI  $\text{kg/m}^2$ ) i stężeniem cholesterolu całkowitego we krwi (TC [mmol/l]) a ciśnieniem skurczowym krwi (RRs [mmHg]) u 693 mężczyzn, którzy brali udział w trzecim badaniu w latach 1992-1993. Wyniki analizy regresji zestawiono w tabeli 6. Okazało się, że u mężczyzn wiek i wskaźnik masy ciała są dodatnio, istotnie statystycznie związane z ciśnieniem skurczowym krwi, przy czym wiek, po uwzględnieniu BMI i stężenia cholesterolu we krwi, w 13% wyjaśnia całkowitą zmienność ciśnienia skurczowego.

Tabela 6. Współczynniki regresji dla wieku, BMI i TC w modelu wielokrotnej regresji liniowej.

zmienne w modelu	b	se	R2	p
wiek [lata]	1,04	0,098	0,13	< 0,001
BMI $\text{kg/m}^2$	0,99	0,140	0,05	< 0,001
TC [mmol/l]	-1,26	0,859	0,10	NS

Wyniki analizy regresji logistycznej podawane są w postaci ilorazu szans (OR) oraz przedziału ufności dla OR.

Przykład 8.

W badaniu CASCADE, opisanym w przykładzie 1, spośród 882 uczestników badania wylosowano 154 osoby (78 mężczyzn i 76 kobiet), u których przeprowadzono badanie MRI. W tej grupie oceniano m.in. związek pomiędzy płcią a wystąpieniem nadciśnienia. Zastosowano analizę regresji logistycznej, której wyniki przedstawia tabela 7. Szansa wystąpienia nadciśnienia u kobiet, po uwzględnieniu wpływu wieku, jest 2,38 raza wyższa niż u mężczyzn i jest ona istotna statystycznie, o czym świadczy 95% przedział ufności, który nie obejmuje wartości 1.

Tabela 7. Surowy i standaryzowany do wieku iloraz szans (OR) dla związku pomiędzy płcią i występowaniem nadciśnienia.

		surowy OR	95%CI	standaryzowany OR	95%CI
Płeć:	mężczyźni	1		1	
	kobiety	2,41	1,16-5,01	2,38	1,13-5,00

Wyniki analizy regresji proporcjonalnych hazardów Coxa podawane są w postaci liczby obserwacji cenzorowanych i kompletnych, ryzyka względnego (RR) oraz przedziału ufności dla RR.



### Przykład 9.

W pierwszym (1983-1984) i drugim badaniu (1987-1988) POL-MONICA Kraków, wspomnianym w przykładzie 7, oceniano w 10-15 letniej obserwacji prospektywnej m.in. związek cech psychologicznych z umieralnością z powodu chorób układu krążenia (ChUK). Do analiz wykorzystano metodę proporcjonalnych hazardów Coxa, stwierdzając, że u mężczyzn wzrost poziomu cechy lęku o jedno odchylenie standardowe (4 punkty) związany jest z 15% wzrostem ryzyka wystąpienia zgonu z powodu ChUK. Wynik ten pozostał niezmienny po uwzględnieniu wpływu wieku (tabela 8).

Tabela 8. Ryzyko względne ( $RR_{sd}$ ) obliczone na przyrost cechy lęku o jedno odchylenie standardowe ( $sd=4$ ) i 95% przedział ufności ( $95\%CI_{sd}$ ) dla  $RR_{sd}$ , dla cechy lęku u mężczyzn – umieralność z powodu ChUK.

obserwacje cenzorowane: żyją lub zgony inne niż ChUK       $n = 1640$  (90,3%)  
 obserwacje kompletne: zmarli z powodu ChUK               $n = 76$  (9,7%)

zmienne w modelu	$RR_{sd}$	95% $CI_{sd}$
lęk	1,15	1,01 - 1,31
lęk, wiek	1,15	1,01 - 1,31

Kilka uwag dotyczących prezentacji graficznych:

- ◆ histogramy – pokazują rozkład badanej cechy ilościowej (słupki „sklejone”),
- ◆ wykresy słupkowe:
  - należy ustalić kategorie dla słupków,
  - nie więcej niż 3 słupki w porównywanych grupach (zalecane 2),
  - należy zachować odstęp pomiędzy porównywanymi grupami,
  - w obrębie grupy słupki powinny być sklejone,
  - kolory słupków powinny być odpowiednio dobrane,
- ◆ wykres rozrzutu – pokazuje zależność pomiędzy dwiema zmiennymi ilościowymi: w układzie współrzędnych XY, na osi X – zmienna niezależna, na osi Y – zmienna zależna,
- ◆ każdy wykres powinien być zatytułowany, a osie opisane wraz z podaniem skali i jednostek.

## Interpretacja wyników

Przy interpretacji uzyskanych wyników należy zwrócić uwagę na pewne problemy. Kilka, najbardziej istotnych, przedstawiono poniżej:

- ◆ Istotność statystyczna vs istotność kliniczna.



Istotność statystyczna wcale nie musi oznaczać istotności klinicznej. Przykładowo, mimo iż w badaniu wykazano istotną statystycznie różnicę w obniżeniu ciśnienia skurczowego krwi pomiędzy dwiema grupami pacjentów leczonych różnymi lekami hipotensyjnymi (o 1 mmHg), to ta różnica nie jest istotna klinicznie. Błędna też może być konkluzja, gdy wynik nieistotny statystycznie interpretowany jest jako nie mający znaczenia klinicznego. Brak istotności statystycznej nie wyklucza istotności klinicznej.

- ◆ Brak istotności statystycznej nie oznacza, że wynik jest nieważny.  
„Brak dowodów nie jest dowodem braku związku” (Altman DG, Bland M. Absence of evidence is not evidence of absence. Br. Med J 1995; 311:485).
- ◆ Mała wartość poziomu krytycznego testu  $p$  nie świadczy o wielkości efektu.  
Poziom  $p$  oznacza jedynie prawdopodobieństwo odrzucenia hipotezy zerowej, mimo że jest prawdziwa. Mała wartość  $p$  mówi tylko (np.  $p < \alpha$ , gdzie  $\alpha$  jest arbitralnie ustalonym przez badacza poziomem istotności; najczęściej  $\alpha = 0,05$  lub  $0,01$  lub  $0,001$ ), że jest mało prawdopodobne, aby badany efekt (różnica, związek) wynikał z przypadku (odrzucając hipotezę zerową i przyjmujemy alternatywną). Gdy  $p \geq \alpha$  (nie ma powodu do odrzucenia hipotezy zerowej; popełnilibyśmy zbyt duży błąd w stosunku do założonego, odrzucając hipotezę zerową), to jest bardzo prawdopodobne, że analizowany efekt jest przypadkowy. A zatem na podstawie poziomu  $p$  nie można nic powiedzieć o wielkości zaobserwowanego efektu. Dlatego też, gdy chcemy ocenić precyzję oszacowania, zaleca się wyznaczenie i podanie przedziału ufności. Im węższy przedział ufności, tym dokładniejsze oszacowanie. Szeroki przedział ufności oznacza małą wiarygodność wyników.
- ◆ Mały efekt, a istotny statystycznie.  
Nawet bardzo mała różnica (patrz np. przykład 4) albo bardzo niski współczynnik korelacji (np.  $r = 0,01$ ) mogą wykazywać silną istotność statystyczną, co jest związane z dużą liczebnością próby.
- ◆ Błędnie zastosowane testy – brak weryfikacji założeń.  
Przed przystąpieniem do analiz należy sprawdzić, czy są spełnione założenia (np. normalność rozkładu – patrz przykład 3, jednorodność wariancji).
- ◆ Odpowiednia liczba obserwacji w analizach regresji.  
W analizach regresji należy zwrócić uwagę na odpowiednią liczbę obserwacji, która powinna być uwzględniona w analizie. W regresji liniowej na jedną zmienną niezależną w modelu powinno przypadać co najmniej 10 obserwacji, a w regresji logistycznej – co najmniej 20 obserwacji na jedną zmienną niezależną. Minimalna liczebność próby w analizach regresji związana jest z rozwiązywaniem układów równań i odwracaniem macierzy współczynników. Przy zbyt małej próbie oszacowania obciążone są dużym błędem.



◆ Uogólnianie wyników.

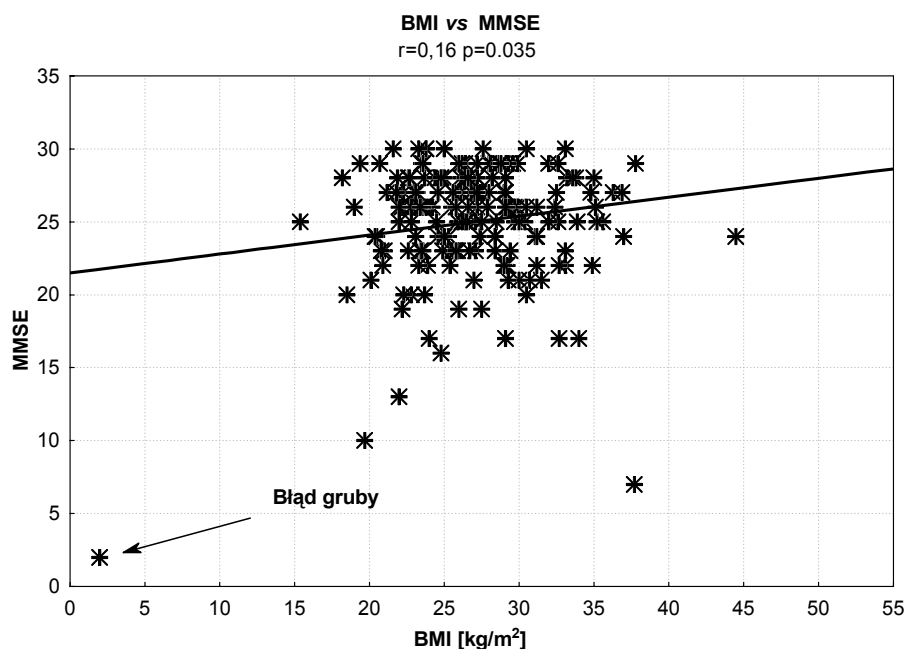
Wyniki uzyskane z analiz mogą być uogólniane jedynie w przypadku, gdy badana próba była reprezentatywna dla całej populacji. W przeciwnym razie wyniki odnoszą się tylko do badanej grupy.

◆ Wysoka korelacja, zakres predykcji:

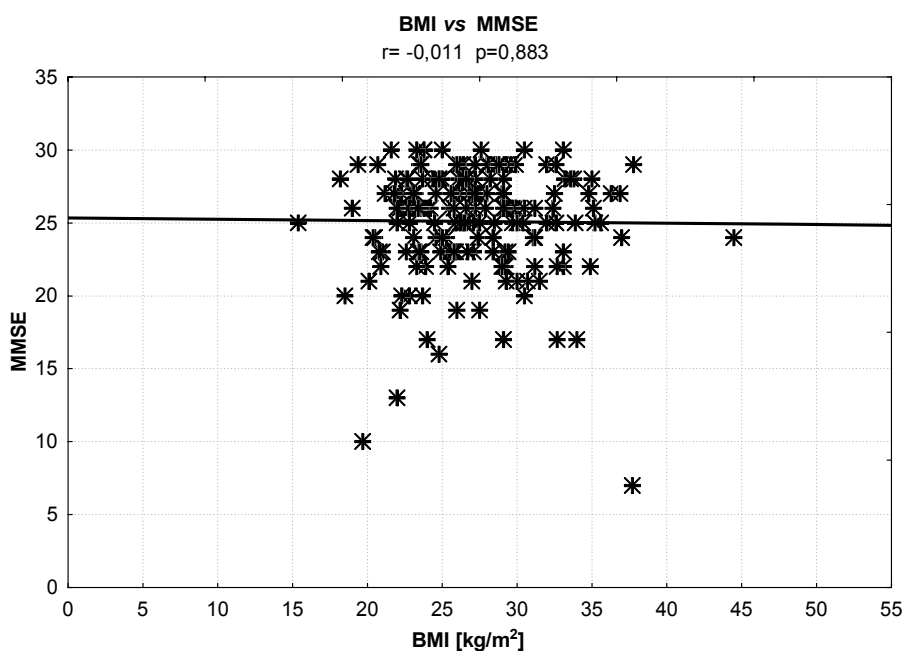
- wysoka korelacja pomiędzy zmiennymi X i Y wcale nie musi świadczyć o związku przyczynowo-skutkowym (patrz: poniżej),
- współczynnik korelacji nie może być ekstrapolowany poza wartości zmiennych X i Y. Związek pomiędzy X i Y, poza zakresem zmienności wartości X i Y, może się zmieniać,
- współczynnik korelacji Pearsona określa siłę i kierunek liniowego związku pomiędzy X i Y,
- współczynnik korelacji jest bardzo czuły na obserwacje odstające (rys. 12 i 13).

Przykład 10.

W badaniu opisanym w przykładzie 1 oceniano także związek pomiędzy wskaźnikiem masy ciała BMI [ $\text{kg}/\text{m}^2$ ] i wynikami testu MMSE u 78 mężczyzn. Analiza korelacji Spearmana (wyniki testu MMSE nie były zgodne z rozkładem normalnym) wykazała dodatni, istotny statystycznie związek pomiędzy BMI i MMSE. Należy jednak zwrócić uwagę na wystąpienie błędu grubego (rys. 12), którego korekta na podstawie danych źródłowych, doprowadziła do uzyskania związku ujemnego, nieistotnego statystycznie (rys. 13).



Rys. 12. Wykres rozrzutu BMI vs MMSE, gdy występuje błąd gruby.



Rys. 13. Wykres rozrzutu BMI vs MMSE, po skorygowaniu błędu grubego.

◆ Ocena związku przyczynowo-skutkowego:

- wiarygodność,
- zgodność,
- siła,
- zależność dawka-odpowiedź,
- odwracalność,
- typ badania,
- ocena dowodu.

## Publikacja wyników

Ostatnim etapem badania statystycznego jest publikacja wyników. Ogólny schemat publikacji powinien obejmować: tytuł, streszczenie, wprowadzenie, z którego ma wynikać cel pracy, materiał i metody, wyniki, dyskusję i wnioski [3, 5, 11-15].

## Literatura

1. Altman D.G. *Practical Statistics for Medical Research*. Chapman & Hall London 1991.
2. Beaglehole R., Bonita R. et al. *Podstawy epidemiologii*. Łódź 2002.
3. BMJ Wydanie polskie. *Rady dla autorów*. Maj 1996, 50-56.



4. Centers for Disease Control. *Principles of Epidemiology*. 2nd Ed. U.S. Department of Health and Human Services, Atlanta GA 1992.
5. Hall M. (red.) *Publikacje naukowe w medycynie. Jak pisać*. BMJ Warszawa 1996.
6. HAPIEE (*Health, Alcohol and Psychosocial Factors in Eastern Europe*), <http://www.ucl.ac.uk/easteurope/hapiee.html>.
7. Kirkwood B.R., Sterne J.A.C. *Essential Medical Statistics*. Blackwell Science 2003.
8. Pająk A., Kawalec E., Szczudlik A. *Zaburzenia funkcji poznawczych a czynniki ryzyka chorób układu krążenia. Projekt CASCADE Kraków. Część I: Projekt badania oraz wyjściowe narażenie na czynniki ryzyka chorób układu krążenia w badanej próbie*. *Przegl. Lek* 1998; 55(12), p. 676-682.
9. POL-MONICA Project: cykl artykułów w: *Przegląd Lekarski* 1996 tom 53.
10. Stanisław A. *Podstawy statystyki dla prowadzących badania naukowe*. *Medycyna Praktyczna*. 9/2002, 189-194.
11. *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication* (International Committee of Medical Journal Editors. Updated October 2007).
12. randomised controlled trials, <http://www.consort-statement.org>.
13. systematic reviews and meta-analyses, <http://www.consort-statement.org/Initiatives/MOOSE/moose.pdf>.
14. observational studies in epidemiology, <http://www.strobe-statement.org>.
15. meta-analyses of observational studies in epidemiology, <http://jama.ama-assn.org/cgi/content/full/283/15/2008>.