SKRYPT DLA STUDENTÓW | październik 2026

Statystyka w fizjoterapii

Od bazy danych z ankiety do tabeli wyników, abstraktu i posteru

Przygotowane dla Studenckiego Koła Naukowego przy Zakładzie Kinezyterapii i Metod Specjalnych

Wydział Nauk o Zdrowiu w Katowicach, Śląski Uniwersytet Medyczny w Katowicach

W skrypcie:

- dobór testu statystycznego do pytania badawczego i typu danych
- analiza danych z badań ankietowych: od bazy do tabeli wyników
- zmienne typowe dla fizjoterapii: natężenie bólu (NRS), BMI, aktywność, typ pracy
- iloraz szans surowy i skorygowany: dlaczego wynik może się odwrócić
- pomiar przed i po terapii: istotność statystyczna a kliniczna
- statystyka w pracy dyplomowej, abstrakcie i na posterze: co raportować, czego unikać
- test: czy poradzisz sobie sam

Opracowanie: Marek Porycki

analizastatystyczna.me | kontakt@analizastatystyczna.me

Udostępniono bezpłatnie do użytku członków koła

Spis treści

Jak korzystać ze skryptu 2

Część 1. Dobór testu do pytania badawczego i typu danych 2

Część 2. Od bazy danych do tabeli wyników 4

Część 3. Praca dyplomowa, abstrakt, poster: co raportować, czego unikać 8

Kiedy warto skonsultować analizę 10

Test: czy poradzisz sobie sam 10

Literatura 10

Wsparcie przy pracy dyplomowej i publikacji 11

Jak korzystać ze skryptu

Skrypt ma trzy części, odpowiadające trzem pytaniom, które w pracy z danymi zadaje sobie prawie każdy student fizjoterapii: jaki test wybrać, jak przejść od surowej bazy z ankiety do tabeli wyników i jak potem pokazać wyniki w pracy dyplomowej, w abstrakcie konferencyjnym i na posterze.

Wszystko pokazano na jednym **przykładzie przewodnim**: ankiecie wśród 303 osób czynnych zawodowo, w której zapytano o typ pracy (siedząca, stojąca, fizyczna), regularne ćwiczenia, tygodniowy czas aktywności fizycznej, ból w odcinku lędźwiowym kręgosłupa w ostatnich 12 miesiącach, jego natężenie, wzrost, masę ciała, wiek i płeć. Natężenie bólu mierzono numeryczną skalą oceny bólu NRS: 11 punktów od 0 (np. „brak bólu”) do 10 (np. „najsilniejszy ból, jaki można sobie wyobrazić”) (Hawker i in., 2011). W części 2 dochodzi drugi, mały zbiór: pomiar bólu przed i po programie ćwiczeń, typowy dla prac z kinezyterapii.

> **Uwaga o danych**
>
> Zbiory są **dydaktyczne**: wygenerowano je tak, żeby zawierały typowe problemy prawdziwych ankiet (duplikaty, ankiety wypełnione w kilkadziesiąt sekund, literówki, braki danych). Wszystkie liczby w skrypcie policzono na tych zbiorach, ale nie opisują one żadnej rzeczywistej populacji i nie wolno ich cytować jako wyników badań.

Część 1. Dobór testu do pytania badawczego i typu danych

Test wynika z dwóch rzeczy: z **pytania** (różnica między grupami, zmiana w czasie czy związek między zmiennymi) i z **typu zmiennych**. Zanim otworzysz program statystyczny, zapisz dla każdej hipotezy, która zmienna jest zależna, która niezależna i jakiego są typu.

Zmienna w przykładzie	Typ	Jak opisywać	Uwagi
płeć, regularne ćwiczenia	nominalna, 2 kategorie	*n* (%)	porównanie dwóch grup
typ pracy	nominalna, 3 kategorie	*n* (%)	trzy grupy niezależne
ból lędźwiowy w 12 mies. (tak/nie)	dychotomiczna	*n* (%)	wynik w regresji logistycznej

Zmienna w przykładzie	Typ	Jak opisywać	Uwagi
natężenie bólu NRS (0-10)	porządkowa lub ilościowa	*Me* [Q1; Q3]	tylko u osób z bólem; status skali bywa sporny
aktywność (min/tydzień)	ilościowa, skośna	*Me* [Q1; Q3]	długi prawy ogon
BMI (kg/m²)	ilościowa, wyliczona	*M* (*SD*)	liczona z masy i wzrostu, nie pytamy o nią
wiek	ilościowa	*M* (*SD*)	pytaj o lata, nie o przedziały

Czy skalę NRS traktować jako porządkową, czy jako ilościową, w literaturze bywa sporne. W skrypcie stosujemy testy rangowe, które nie wymagają założenia o równych odstępach między punktami skali. Niezależnie od wyboru, zapisz go i uzasadnij w rozdziale metodologicznym.

Mapa wyboru testu

Pytanie	Zmienna zależna	Test podstawowy	Wielkość efektu
różnica: 2 grupy niezależne	ilościowa, rozkład zbliżony do normalnego	*t* Welcha	*d* Cohena
różnica: 2 grupy niezależne	porządkowa lub skośna	U Manna-Whitneya	*r* rangowo-dwuseryjne
różnica: 3+ grupy niezależne	ilościowa, rozkład zbliżony do normalnego	ANOVA Welcha + Games-Howell	ω^2
różnica: 3+ grupy niezależne	porządkowa lub skośna	Kruskal-Wallis + test Dunna	ε^2
zmiana: ci sami badani przed i po	ilościowa, różnice zbliżone do normalnego	*t* dla prób zależnych	*d* dla różnic
zmiana: ci sami badani przed i po	porządkowa lub skośna	test Wilcoxona dla par	*r*; mediana zmiany
związek dwóch zmiennych	porządkowe lub skośne	ρ Spearmana	ρ
związek dwóch zmiennych	obie nominalne	χ^2 (lub dokładny test Fishera)	φ, *V* Craméra, OR
wynik tak/nie z kilku czynników	dychotomiczna	regresja logistyczna	OR skorygowane

Test U Manna-Whitneya sprawdza, czy wartości w jednej grupie są systematycznie większe niż w drugiej (Mann i Whitney, 1947), a nie wprost, czy różnią się mediany. O medianach można mówić tylko wtedy, gdy rozkłady w grupach mają podobny kształt. Test Kruskala-Wallisa uogólnia to podejście na więcej grup (Kruskal i Wallis, 1952). Gdy wynik jest istotny, porównania par wykonuje się testem Dunna, który korzysta z rang nadanych łącznie wszystkim grupom (Dunn, 1964). Dla dwóch pomiarów u tych samych osób odpowiednikiem rangowym jest test Wilcoxona dla par (Wilcoxon, 1945).

Wiele porównań: korekta Holma

Przy trzech grupach są trzy porównania par, przy pięciu już dziesięć. Każde kolejne porównanie zwiększa ryzyko fałszywego alarmu. Procedura Holma kontroluje prawdopodobieństwo co najmniej jednego błędu I rodzaju w całej rodzinie porównań i nigdy nie ma mniejszej mocy niż poprawka Bonferroniego: odrzuca co najmniej te same hipotezy, a zwykle więcej (Holm, 1979). Wartości *p* sortuje się od najmniejszej; najmniejszą mnoży się przez liczbę porównań m, następną przez $m - 1$ itd.

Poziom wyżej

W badaniach fizjoterapeutycznych ten sam problem pojawia się w innej postaci: kilka wyników (ból, zakres ruchu, siła, sprawność, jakość życia) mierzonych w kilku punktach czasu. Decyzja, który wynik jest główny, a które są dodatkowe, zapada na etapie planu badania i powinna być zapisana w rozdziale metodologicznym, zanim zobaczysz dane.

Gdzie możemy pomóc

Plan analizy pod Twój temat: przypisanie testu do każdej hipotezy, decyzja o korektach, opis metod statystycznych do pracy.

Ocena, ilu badanych potrzeba, zanim ankieta lub program ćwiczeń trafi do uczestników.

Część 2. Od bazy danych do tabeli wyników

Krok 1. Kontrola jakości ankiet

Zanim policzysz cokolwiek, ustal na piśmie kryteria wykluczenia i stosuj je przed obejrzeniem wyników. W ankietach internetowych typowe są duplikaty (ta sama osoba wysłała formularz dwa razy) i ankiety wypełnione zbyt szybko, żeby odpowiedzi mogły być przemyślane (Curran, 2016).

Etap	Liczba ankiet
pobrane z formularza	303
wykluczone: duplikaty (identyczne odpowiedzi we wszystkich pytaniach)	−3
wykluczone: czas wypełnienia poniżej 120 s	−5
włączone do analizy	**295**

W pracy opisz to jednym zdaniem i ewentualnie schematem przepływu: „Z 303 ankiet wykluczono 3 duplikaty i 5 ankiet wypełnionych w czasie krótszym niż 120 s; analizie poddano 295 ankiet.” Próg 120 s jest tu przykładem; w swoim badaniu ustal go na podstawie czasu wypełnienia w pilotażu.

Krok 2. Wartości niemożliwe i braki danych

Wzrost 1,72 to wzrost wpisany w metrach zamiast centymetrów, a masa 7,8 kg to zgubiona cyfra. Ustal granice wiarygodności (tu: wzrost 120-220 cm, masa 30-250 kg) i wartości spoza nich zamień na braki albo, gdy pomyłka jest jednoznaczna i opisana, popraw. W przykładzie dotyczyło to 1 wartości wzrostu i 1 wartości masy, które zamieniono na braki. Razem z osobami, które nie podały wzrostu, BMI nie da się policzyć dla 7 osób.

Usuwanie całych przypadków z brakami jest bezpieczne tylko wtedy, gdy braki są nieliczne i nie zależą od wartości badanych zmiennych. Mechanizm braków decyduje o tym, czy prosta analiza kompletnych przypadków jest uprawniona (Rubin, 1976), a gdy braków jest więcej, stosuje się imputację wielokrotną, która ma własne pułapki (Sterne i in., 2009).

Krok 3. Zmienne wyliczane i pytania warunkowe

```
BMI = masa_kg / (wzrost_cm / 100)²
nrs: pytanie zadawane tylko osobom z odpowiedzią „tak” na pytanie o ból (u pozostałych brak z założenia)
```

Zmienną wyliczaną tworzysz formułą w programie, nigdy ręcznie. Brak NRS u osób bez bólu nie jest brakiem danych, tylko konsekwencją budowy ankiety: analizy natężenia bólu dotyczą 134 osób, które zgłosiły ból, a nie całej próby. To trzeba jasno napisać przy każdej tabeli z NRS.

Krok 4. Tabela charakterystyki próby

Tabela 1

Charakterystyka próby według regularnych ćwiczeń

Zmienna	Ćwiczy regularnie (*n* = 184)	Nie ćwiczy (*n* = 111)	Ogółem (*n* = 295)
Kobiety, *n* (%)	85 (46,2)	63 (56,8)	148 (50,2)
Wiek, *M* (*SD*)	38,6 (10,0)	46,0 (9,8)	41,4 (10,5)
Praca siedząca, *n* (%)	89 (48,4)	37 (33,3)	126 (42,7)
Praca fizyczna, *n* (%)	52 (28,3)	38 (34,2)	90 (30,5)
Aktywność, min/tydz., *Me* [Q1; Q3]	220 [160; 300]	70 [40; 125]	170 [90; 250]
BMI, *M* (*SD*)	24,3 (3,6)	26,0 (3,6)	25,0 (3,7)
Ból lędźwiowy w 12 mies., *n* (%)	74 (40,2)	60 (54,1)	134 (45,4)
NRS u osób z bólem, *Me* [Q1; Q3]	3 [2; 5]	4 [4; 6]	4 [3; 5]

Uwaga. Zbiór dydaktyczny. BMI: brak danych dla 7 osób. NRS: tylko osoby zgłaszające ból (*n* = 74, 60 i 134). Q1, Q3: pierwszy i trzeci kwartyl.

Zwróć uwagę na wiek: osoby ćwiczące regularnie są w tej próbie wyraźnie młodsze i mają niższe BMI. Wrócimy do tego w kroku 6, bo właśnie te różnice zmieniają interpretację jednego z wyników.

Krok 5. Analizy i zdania wynikowe

Natężenie bólu a regularne ćwiczenia

Zapis wyniku

„Wśród osób zgłaszających ból lędźwiowy osoby ćwiczące regularnie oceniały jego natężenie niżej (*Me* = 3, *n* = 74) niż osoby niećwiczące (*Me* = 4, *n* = 60), *U* = 1486, *p* < ,001, *r* = ,33.”

Przy skali 0-10 wiele osób ma tę samą odpowiedź, więc program musi zastosować poprawkę na rangi wiązane (w trybie asymptotycznym robi to domyślnie). Wielkość efektu *r* rangowo-dwuseryjne to różnica między prawdopodobieństwem, że losowo wybrana osoba niećwicząca ma wyższy wynik niż losowo wybrana osoba ćwicząca, a prawdopodobieństwem sytuacji odwrotnej.

Natężenie bólu a typ pracy

Tabela 2

Porównania par typów pracy w natężeniu bólu (test Dunna z korektą Holma)

Porównanie	*z*	*p*	*p* po korekcie Holma
siedząca vs stojąca	−1,30	,194	,194
siedząca vs fizyczna	2,37	,018	,036
stojąca vs fizyczna	3,38	< ,001	,002

Uwaga. n = 134 osób zgłaszających ból. Dodatnie *z*: wyższe rangi w drugiej grupie pary.

Zapis wyniku

„Natężenie bólu różniło się w zależności od typu pracy, $H(2) = 12{,}37$, $p = ,002$, $\varepsilon^2 = ,09$. Osoby pracujące fizycznie (*Me* = 5) oceniały ból wyżej niż osoby z pracą siedzącą (*Me* = 4; $p_{Holm} = ,036$) i stojącą (*Me* = 3,5; $p_{Holm} = ,002$). Osoby z pracą siedzącą i stojącą nie różniły się istotnie ($p_{Holm} = ,194$).”

Wielkość efektu ε² dla testu Kruskala-Wallisa liczy się jako H / (n − 1) (Tomczak i Tomczak, 2014). Wartość ,09 oznacza, że typ pracy wiąże się z niewielką częścią zróżnicowania natężenia bólu.

Natężenie bólu a BMI

Zapis wyniku

„Wśród osób zgłaszających ból wyższe BMI wiązało się z wyższym natężeniem bólu, ρ = ,25, p = ,003, n = 132.”

To badanie przekrojowe: nie wiadomo, czy masa ciała nasila ból, czy ból ogranicza aktywność i sprzyja przyrostowi masy. Właściwe sformułowanie to „wiązało się z”, a nie „wpływało na”.

Krok 6. Iloraz szans surowy i skorygowany

Analiza dwuzmiennowa

Zapis wyniku

„Ból lędźwiowy w ostatnich 12 miesiącach zgłosiło 74 z 184 osób ćwiczących regularnie (40,2%) i 60 z 111 osób niećwiczących (54,1%), $\chi^2(1, N = 295) = 5,35$, p = ,021, φ = ,13; OR = 0,57 [0,36; 0,92].”

Iloraz szans porównuje szanse wystąpienia wyniku w dwóch grupach; OR = 1 oznacza brak różnicy (Szumilas, 2010). Przy częstym wyniku (tu ból zgłosiło blisko pół próby) OR jest wyraźnie dalej od 1 niż stosunek odsetków, więc nie czytaj go jako „ćwiczący mają o 43% rzadziej ból”. Na tym etapie wygląda to jak wynik do tytułu pracy: ćwiczenia chronią przed bólem krzyża.

Ten sam wynik po uwzględnieniu wieku, płci, BMI i typu pracy

Tabela 3

Czynniki związane z bólem lędźwiowym w ostatnich 12 miesiącach: ilorazy szans surowe i skorygowane (regresja logistyczna)

Zmienna	OR surowe [95% CI]	OR skorygowane [95% CI]	*p*
ćwiczy regularnie (vs nie)	0,57 [0,35; 0,92]	1,30 [0,72; 2,34]	,389
wiek (o 1 rok)	1,081 [1,053; 1,110]	1,086 [1,055; 1,118]	< ,001
kobieta (vs mężczyzna)	1,36 [0,86; 2,17]	1,74 [1,02; 2,97]	,040
BMI (o 1 kg/m²)	1,131 [1,057; 1,209]	1,104 [1,025; 1,188]	,009
praca stojąca (vs siedząca)	1,04 [0,58; 1,86]	1,44 [0,74; 2,82]	,283
praca fizyczna (vs siedząca)	1,80 [1,04; 3,12]	2,58 [1,37; 4,85]	,003

Uwaga. N = 288 (osoby z kompletnymi danymi o BMI); ból zgłosiło 132 osób. OR surowe: każda zmienna osobno; OR skorygowane: wszystkie zmienne w jednym modelu. Wartości p dotyczą modelu skorygowanego.

Po uwzględnieniu pozostałych zmiennych iloraz szans dla regularnych ćwiczeń zmienił się z 0,57 na 1,30, a przedział ufności [0,72; 2,34] obejmuje 1. Osoby ćwiczące były w próbie młodsze (średnio 38,6 vs 46,0 lat) i miały niższe BMI, a wiek i BMI silnie wiążą się z bólem. „Ochronny” efekt ćwiczeń z analizy dwuzmiennowej wynikał więc w dużej mierze z tego, kto ćwiczy, a nie z samych ćwiczeń. W badaniu przekrojowym nie da się też wykluczyć odwrotnego kierunku: część osób mogła zacząć ćwiczyć właśnie z powodu bólu.

Dwa zastrzeżenia. Po pierwsze, przedział obejmujący 1 oznacza brak istotności, a nie dowód braku związku (Szumilas, 2010). Po drugie, iloraz szans może się zmienić po dodaniu do modelu zmiennej związanej z wynikiem nawet wtedy, gdy nie jest ona czynnikiem

zakłócającym; część zmiany może więc wynikać z właściwości samej miary, a nie z zakłócenia.

Poziom wyżej

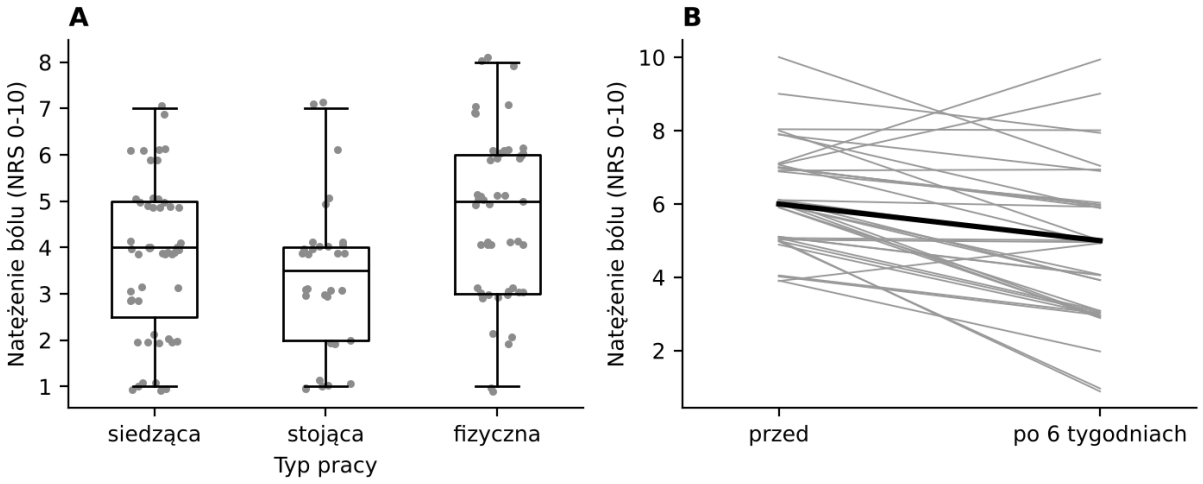
Tabela 3 to model w najprostszej postaci. W pracy, która ma przejść recenzję, trzeba jeszcze uzasadnić dobór zmiennych kontrolnych (na podstawie wiedzy merytorycznej, a nie istotności), sprawdzić, czy wiek i BMI wiążą się z wynikiem liniowo w skali logitu, rozważyć interakcje (np. czy związek z ćwiczeniami jest taki sam przy pracy siedzącej i fizycznej), ocenić dopasowanie modelu i liczbę zdarzeń w stosunku do liczby parametrów. Każdy z tych kroków może zmienić wniosek.

Krok 7. Pomiar przed i po terapii

Drugi zbiór dydaktyczny: 36 osób z bólem lędźwiowym, które ukończyły 6-tygodniowy program ćwiczeń. Natężenie bólu w skali NRS zmierzono przed programem i po nim. To ci sami ludzie, więc pomiary są zależne: test dla grup niezależnych byłby tu błędem.

Zapis wyniku

„Po 6 tygodniach programu natężenie bólu było niższe (*Me* = 5 [3; 6]) niż przed nim (*Me* = 6 [5; 7]), test Wilcoxona dla par: $z = -3{,}93$, $p < {,}001$. Mediana zmiany wyniosła 1 pkt. Poprawę o co najmniej 2 pkt uzyskało 17 z 36 osób (47,2%).”



Rycina 1

Natężenie bólu według typu pracy (A) i przed oraz po 6-tygodniowym programie ćwiczeń (B)

Uwaga. Zbiory dydaktyczne. Panel A: *n* = 134, punkty lekko rozsunięte. Panel B: *n* = 36; szare linie to poszczególne osoby, czarna linia to mediana.

Wynik jest istotny statystycznie, ale to nie koniec interpretacji. Dla skali NRS w bólu krzyża zaproponowano minimalną istotną zmianę równą 2 punkty, a przy uwzględnieniu poziomu wyjściowego poprawę o 30% (Ostelo i in., 2008). Mediana zmiany w przykładzie wynosi 1 pkt, a próg 2 pkt osiągnęło mniej niż pół grupy. W pracy pokaż oba wyniki: istotność statystyczną i odsetek osób z klinicznie istotną poprawą. Osoby bez żadnej zmiany (tu: 5) test Wilcoxona w klasycznej wersji pomija, co też warto odnotować.

Drugie zastrzeżenie dotyczy samego układu badania. Do programu trafiają zwykle osoby, którym w danym momencie dokucza silny ból, a skrajne wyniki przy kolejnym pomiarze mają tendencję do zbliżania się do średniej. To zjawisko regresji do średniej może wyglądać jak efekt terapii (Barnett i in., 2005). Bez grupy kontrolnej nie da się go oddzielić od działania ćwiczeń, dlatego właściwy wniosek brzmi „po programie ból był niższy”, a nie „program

zmniejszył ból”.

Poziom wyżej

Badanie, które ma odpowiedzieć na pytanie, czy terapia działa, wymaga grupy porównawczej, najlepiej z losowym przydziałem. Wynik po terapii analizuje się wtedy z uwzględnieniem wyniku wyjściowego (analiza kowariancji), przy trzech i więcej pomiarach stosuje się modele dla pomiarów powtarzanych, a osoby, które przerwały terapię, wymagają osobnej decyzji analitycznej. To etap, na którym większość prac interwencyjnych potrzebuje konsultacji statystycznej.

Gdzie możemy pomóc

Analiza statystyczna Twoich danych: od czyszczenia bazy po tabele i ryciny w APA 7 z opisem wyników.

Badania przed i po terapii oraz z grupą kontrolną: plan pomiarów, analiza zmian, istotność kliniczna.

Modele wieloczynnikowe (regresja logistyczna i liniowa) z uzasadnieniem doboru zmiennych i kontrolą założeń.

Część 3. Praca dyplomowa, abstrakt, poster: co raportować, czego unikać

Dla badań obserwacyjnych, w tym przekrojowych badań ankietowych, standardem raportowania jest lista STROBE: 22 punkty dotyczące tytułu, streszczenia, wstępu, metod, wyników i dyskusji (von Elm i in., 2007). Warto przejść ją punkt po punkcie przed złożeniem pracy, nawet jeśli wydział jej nie wymaga. Zapis statystyk w skrypcie (symbole kursywą, dokładne *p*, brak zera przed przecinkiem przy *p* i *r*) jest zgodny z wytycznymi APA 7 (American Psychological Association, 2020); jeśli wydział wymaga zapisu 0,05, stosuj go konsekwentnie.

W pracy dyplomowej

Raportuj	Unikaj
liczbę pobranych, wykluczonych i analizowanych ankiet z powodami	samego *N* końcowego bez historii
test, statystykę, stopnie swobody, dokładne *p*	*p* = ,000 przepisanego z programu (pisz *p* < ,001)
wielkość efektu i 95% CI przy każdym wyniku	samych gwiazdek istotności
dla ilu osób liczono NRS i dlaczego	tabel z NRS bez informacji, kogo dotyczą
OR surowe i skorygowane obok siebie	tylko tego modelu, który „wyszedł”
przy terapii: istotność statystyczną i odsetek poprawy klinicznej	samego *p* jako dowodu skuteczności
„wiązało się z” w badaniu przekrojowym	„wpływa”, „powoduje”, „chroni przed”
„nie wykazano istotnej różnicy” z przedziałem ufności	„brak różnicy”, „trend w kierunku istotności”

W abstrakcie konferencyjnym

Abstrakt ma ścisły limit słów lub znaków ustalony przez organizatorów, więc wyniki muszą się zmieścić w trzech-czterech zdaniach. Muszą w nich być: liczba badanych, główny wynik z miarą efektu i przedziałem ufności oraz wniosek, który nie wychodzi poza dane. Przykład dla wyników z tego skryptu:

Wyniki i wnioski (przykład)

„Ból lędźwiowy w ostatnich 12 miesiącach zgłosiło 40,2% osób ćwiczących regularnie i 54,1% osób niećwiczących (OR = 0,57 [0,36; 0,92]). Po uwzględnieniu wieku, płci, BMI i typu pracy związek nie był istotny (OR = 1,30 [0,72; 2,34]). Ból wiązał się ze starszym wiekiem, wyższym BMI, płcią żeńską i pracą fizyczną. Różnice w częstości bólu między osobami ćwiczącymi i niećwiczącymi mogą wynikać z różnic w wieku i masie ciała.”

Najczęstszy błąd w abstrakcie to podanie samego *p* („różnica była istotna, *p* = ,021”) bez informacji, jak duża była różnica. Drugi: wniosek o skuteczności terapii z badania bez grupy kontrolnej.

Na posterze

- Jedna główna wiadomość w tytule lub pod nim, napisana zdaniem, a nie hasłem.
- Wyniki jako wykres, tabela tylko dla charakterystyki próby; najwyżej 2-3 wykresy.
- Dla pomiaru przed i po: wykres z liniami dla poszczególnych osób (jak Rycina 1B) mówi więcej niż dwa słupki ze średnimi.
- Ilorazy szans na wykresie typu forest plot; bez efektów 3D i bez wykresów kołowych dla porównań grup.
- Liczby zaokrąglone: odsetki do 1 miejsca po przecinku, OR do 2 miejsc; nie przepisuj 4 miejsc z programu.
- Pod każdym wykresem *N* i nazwa testu; ograniczenia badania w jednym zdaniu.
- Tekst czytelny z odległości kilku kroków: tytuł, nagłówki i podpisy osi duże; mniej słów, więcej odstępu.

Gdzie możemy pomóc

Przygotowanie wyników do abstraktu i posteru: wykresy z przedziałami ufności, forest plot, tabela charakterystyki próby.

Recenzja rozdziałów metodologicznego i wynikowego przed oddaniem promotorowi, redakcja naukowa i korekta, przygotowanie do obrony.

Kiedy warto skonsultować analizę

Wiele analiz da się wykonać samodzielnie na podstawie tego skryptu. Są jednak sytuacje, w których ryzyko błędu, który wyjdzie dopiero w recenzji, jest duże:

- badasz skuteczność terapii: pomiar przed i po, grupa kontrolna, kilka punktów pomiaru;
- planujesz więcej niż jedną hipotezę główną albo kilka wyników (ból, zakres ruchu, siła, jakość życia);
- chcesz pokazać związek „po uwzględnieniu” wieku, płci lub innych zmiennych (regresja, ilorazy szans skorygowane);
- braki danych przekraczają kilka procent albo dotyczą kluczowej zmiennej;
- używasz kwestionariusza wielopozycyjnego i musisz policzyć rzetelność oraz wynik ogólny;
- wynik jest na granicy istotności i nie wiesz, jak go opisać;
- promotor lub recenzent odesłał wyniki z uwagami.

Test: czy poradzisz sobie sam

- ☐ Dla każdej hipotezy wiem, która zmienna jest zależna, która niezależna i jakiego są typu.
- ☐ Wiem, kiedy opisuję zmienną średnią, a kiedy medianą i kwartylami.
- ☐ Mam spisane z góry kryteria wykluczania ankiet.
- ☐ Wiem, co zrobić z wartościami niemożliwymi i brakami danych, i umiem to opisać.
- ☐ Potrafię wybrać test dla dwóch i trzech grup oraz dla pomiaru przed i po.
- ☐ Wiem, po co jest korekta Holma i kiedy ją stosować.
- ☐ Przy każdym teście umiem podać wielkość efektu i przedział ufności.
- ☐ Rozumiem, czym różni się iloraz szans surowy od skorygowanego.
- ☐ Umiem odróżnić istotność statystyczną od klinicznej i wiem, czym grozi brak grupy kontrolnej.
- ☐ Odpowiem komisji na pytanie: „dlaczego ten test, a nie inny?”

Wynik

9-10 razy „tak”: masz solidne podstawy; warto najwyżej sprawdzić wyniki przed złożeniem pracy.

5-8 razy „tak”: analiza jest w zasięgu, ale ryzyko błędu jest realne. Skonsultuj plan analizy przed jej wykonaniem.

0-4 razy „tak”: ta część zajmie prawdopodobnie kilka tygodni, a błędy wyjdą w recenzji. To dobry moment na wsparcie specjalisty.

Literatura

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). https://doi.org/10.1037/0000165-000

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34*(1), 215-220. https://doi.org/10.1093/ije/dyh299

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology, 66*, 4-19. https://doi.org/10.1016/j.jesp.2015.07.006

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics, 6*(3), 241-252. https://doi.org/10.1080/00401706.1964.10490181

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care & Research, 63*(S11), S240-S252. https://doi.org/10.1002/acr.20543

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70. https://www.jstor.org/stable/4615733

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583-621. https://doi.org/10.1080/01621459.1952.10483441

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics, 18*(1), 50-60. https://doi.org/10.1214/aoms/1177730491

Ostelo, R. W. J. G., Deyo, R. A., Stratford, P., Waddell, G., Croft, P., Von Korff, M., Bouter, L. M., & de Vet, H. C. (2008). Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine, 33*(1), 90-94. https://doi.org/10.1097/BRS.0b013e31815e3a10

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581-592. https://doi.org/10.1093/biomet/63.3.581

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ, 338*, Article b2393. https://doi.org/10.1136/bmj.b2393

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 19*(3), 227-229.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences, 21*(1), 19-25.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet, 370*(9596), 1453-1457. https://doi.org/10.1016/S0140-6736(07)61602-X

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80-83. https://doi.org/10.2307/3001968

Wsparcie przy pracy dyplomowej i publikacji

Wszystko, co opisano w skrypcie, da się zrobić samodzielnie. Jeśli brakuje czasu, analiza okazuje się bardziej złożona, niż zakładano, albo promotor odesłał wyniki do poprawy, zespół **analizastatystyczna.me** przejmie tę część pracy. Materiał przygotował Marek Porycki, psycholog, autor serwisu.

Usługa	Zakres
Plan badania i liczebność próby	problem, hipotezy, zmienne, dobór narzędzi, analiza mocy, opis do rozdziału metodologicznego
Realizacja badania	kwestionariusz, ankieta online, zbieranie danych, baza danych i księga kodowa
Analiza statystyczna	obliczenia, modele wieloczynnikowe, tabele i ryciny w APA 7, opis wyników i metod statystycznych
Badania interwencyjne	plan pomiarów przed i po terapii, grupa kontrolna, analiza zmian i istotności klinicznej
Sprawdzenie analizy	weryfikacja wykonanych obliczeń i opisu przed złożeniem pracy
Abstrakt i poster	wyniki do abstraktu konferencyjnego, wykresy i układ posteru

Usługa	Zakres
Recenzja rozdziałów	ocena rozdziałów przed oddaniem promotorowi: struktura, logika wywodu, zgodność z wynikami, uwagi i sugestie poprawek
Redakcja naukowa i korekta	język naukowy, spójność i styl, redakcja tabel i rycin, bibliografia w wymaganym stylu
Przygotowanie do obrony	metody i wyniki pod kątem pytań komisji, slajdy z wynikami

Zbiór dydaktyczny z przykładu przewodniego (303 ankiet, plik CSV) wysyłamy na prośbę, żeby każdy mógł przećwiczyć analizy w SPSS, Statistice, JASP, jamovi lub R i porównać wyniki z tabelami w skrypcie.

Jak zacząć: napisz na **kontakt@analizastatystyczna.me** albo zeskanuj kod QR. Podaj temat pracy, typ badania, przybliżoną liczbę osób badanych i termin.

analizastatystyczna.me
kontakt@analizastatystyczna.me

Materiał udostępniony bezpłatnie członkom Studenckiego Koła Naukowego przy Zakładzie Kinezyterapii i Metod Specjalnych Śląskiego Uniwersytetu Medycznego w Katowicach. Można go rozpowszechniać w całości, z zachowaniem informacji o autorze.