PRACA POGLĄDOWA

Jak interpretować wyniki badań porównawczych leków? Metody analizy statystycznej – standaryzowana różnica średnich SMD oraz różnica średnich obliczona metodą najmniejszych kwadratów LSMD

How to interpret the results of comparative drug tests? Statistical analysis methods: standardized SMD mean difference and mean difference calculated by the LSMD least squares method

Monika Marko, Rafał Pawliczak

Zakład Immunopatologii, Katedra Alergologii, Immunologii i Dermatologii, Wydział Nauk Biomedycznych i Kształcenia Podyplomowego, Uniwersytet Medyczny w Łodzi

STRESZCZENIE

Odpowiednie narzędzia statystyczne oraz ich właściwe zastosowanie umożliwiają wiarygodną ocenę bezpieczeństwa i skuteczności badanych leków. Do tej pory opisano kilka metod statystycznych, które pomagają ocenić wielkość różnic między dwiema lub więcej interwencjami w badaniach klinicznych i zapewniają lepsze oszacowanie efektów leczenia. Nauki medyczne oraz biomedyczne intensywnie wykorzystują dane badawcze. Wskazuje to, jak ważna jest wiedza statystyczna oraz posiadanie umiejętności jej wykorzystania w zakresie badań medycznych i biomedycznych. Na klinicyście spoczywa obowiązek wybrania odpowiedniej metody w celu wykonania analizy statystycznej przeprowadzonych przez siebie badań. Niniejszy artykuł opisuje i tłumaczy metody statystyczne SMD (*standardized mean difference*) oraz LSMD (*least squares mean difference*) służące do porównywania skuteczności interwencji w badaniach porównawczych leków.

SŁOWA KLUCZOWE

standaryzowana różnica średnich, różnica średnich obliczona metodą najmniejszych kwadratów, analiza statystyczna.

ABSTRACT

Adequate statistical tools and their proper application enable reliable assessment of the safety and effectiveness of the tested drugs. To date, several statistical methods have been described that help assess the magnitude of differences between two or more interventions in clinical trials and that provide a better estimate of treatment effects. Medical and biomedical sciences have become fields intensively using research data. It shows the importance of statistical knowledge and the ability to use it in the field of medical/biomedical research.

It is the clinician's responsibility to choose the appropriate method to perform statistical analysis of the tests they have carried out. This article describes and explains SMD statistical methods (standardized mean difference – SMD) and LSMD (least squares mean difference) for comparing the effectiveness of interventions in comparative drug trials.

KEY WORDS

standardized mean difference, least squares mean difference, statistical analysis.

ADRES DO KORESPONDENCJI

prof. dr hab. n. med. Rafał Pawliczak, Zakład Immunopatologii, Katedra Alergologii, Immunologii i Dermatologii, Uniwersytet Medyczny w Łodzi, ul. Żeligowskiego 7/9, 90-752 Łódź,
e-mail: rafal.pawliczak@csk.umed.lodz.pl

WSTĘP

Nauki medyczne oraz biomedyczne to dziedziny intensywnie wykorzystujące dane badawcze, które potrzebują zaawansowanych metod analizy i eksploracji tych danych [1]. Zastosowanie odpowiednich narzędzi statystycznych pozwala w sposób wiarygodny ocenić bezpieczeństwo i skuteczność leku [2]. Aby wybrać właściwe leczenie spośród wielu terapii uznanych za skuteczne, należy oprzeć się na dowodach z badań systematycznych dotyczących skuteczności oraz działań niepożądanych różnych terapii. Najbardziej efektywne wydaje się porównanie różnych metod leczenia na podstawie danych uzyskanych w badaniach z randomizacją (*randomized*) przeprowadzonych metodą podwójnej ślepej próby (*double-blind*) oraz w badaniach bezpośrednich (*head-to-head trials*) [1, 2].

Opisane do tej pory metody statystyczne pomagają ocenić wielkość różnic między dwiema lub więcej interwencjami w badaniach klinicznych i zapewniają lepsze oszacowanie efektów leczenia niż opieranie się jedynie na wartościach *p* (wynik istotny statystycznie) [3]. Należą do nich takie metody, jak standaryzowana różnica średnich (*standardized mean difference* – SMD) oraz różnica średnich obliczona metodą najmniejszych kwadratów (*least squares mean difference* – LSMD) [2, 4]. Na klinicyście spoczywa obowiązek zebrania uzyskanych wyników oraz wyboru odpowiednich metod w celu wykonania analizy statystycznej [3].

Celem niniejszego artykułu jest zapoznanie czytelników z metodami SMD oraz LSMD, które służą porównywaniu skuteczności interwencji w badaniach, oraz przedstawienie przykładów, w jaki sposób te metody można wykorzystać przy podejmowaniu decyzji o leczeniu.

METODY STATYSTYCZNE I ICH INTERPRETACJA W BADANIACH PORÓWNAWCZYCH LEKÓW

Gdy wyniki badań klinicznych są istotne statystycznie, nie należy wybierać metody leczenia na podstawie porównań istotności statystycznej, ponieważ na jej wielkość duży wpływ ma liczba badanych pacjentów. Dlatego niewielka próba wysoce skutecznej terapii może mieć statystycznie istotny wynik, który jest mniejszy niż wynik dużej próby umiarkowanie skutecznego leczenia [2].

Chociaż wyniki analiz statystycznych dostarczają kluczowych informacji, wielkość istotności statystycznej niekoniecznie wskazuje na wielkość efektu leczenia. W związku z tym niemożliwe jest ustalenie na podstawie stopnia istotności statystycznej, jak na przykład skuteczność nowej terapii ocenianej w jednym badaniu wypada w porównaniu ze skutecznością innych ustalonych lub nowych metod leczenia tego samego przypadku [5].

Problem interpretacji znaczenia statystycznego można rozwiązać, jeśli zastosuje się koncepcję wielkości efektu, która została opracowana, aby umożliwić porównanie skuteczności między badaniami klinicznymi. Wielkość efektu może pomóc w podjęciu decyzji, czy często niewielki wzrost skuteczności nowszych terapii jest wystarczający, aby uzasadnić wprowadzenie nowej metody leczenia [2]. W tym celu należy dokładnie przeanalizować efektywność powszechnie stosowanych terapii. Jednak porównanie badań klinicznych dotyczących różnych leków lub metod terapeutycznych może być trudne. W wielu przypadkach stosuje się miary wielkości efektu. Jak sama nazwa wskazuje, jej oszacowanie dostarcza możliwość do interpretacji wartości wielkości efektu leczenia. Tę miarę można następnie wykorzystać do porównania uzyskanej skuteczności terapii z podobnie obliczonymi miarami skuteczności leczenia w innych badaniach, w któ-

rych stosuje się pozornie nieporównywalne miary [3, 5]. W związku z tym miary wielkości efektu są niezbędne, aby dokonać racjonalnych ocen.

PRZYKŁAD 1

Zespół klinicystów przeprowadził badanie oceniające skuteczność terapii bólu pleców za pomocą 5-punktowej skali oceny intensywności bólu. Równolegle trwało badanie z zastosowaniem 10-punktowej skali oceny. W takim przypadku nie można porównać wyników, ponieważ nawet 1-punktowy spadek ma inne znaczenie dla każdej skali [2]. Nawet jeśli dwa badania wykorzystują tę samą miarę, nie można porównać zmienionych wyników między leczeniem a placebo albo alternatywnym leczeniem, ponieważ badania te mogą się różnić standardami dokładności pomiaru. Stosowanie różnych skal pomiaru i różnice w jego precyzji utrudniają porównanie badań. Trudności te można omiąć, jeśli oszacuje się wielkość efektu [2, 5].

STANDARYZOWANA RÓŻNICA ŚREDNICH JAKO MIARA WIELKOŚCI EFEKTU

Miary wielkości efektu stosuje się do mierzenia wpływu pewnego czynnika na wynik ogólny grupy, czyli siły zwią-

ku między zmienną niezależną (np. dawka leku) a zmienną zależną (np. stężenie cholesterolu we krwi) [6]. Należy pamiętać, że interpretacja wielkości efektu opiera się na założeniu normalności rozkładów wyników porównywanych grup i nie zależy od wielkości próby [5]. Wśród miar wielkości efektu wyróżnia się między innymi rodzinę d , do której zalicza się SMD, w programach statystycznych zamiennie określana jako d Cohena i stosowana dla testu t dla prób zależnych. Dla testu prób niezależnych stosuje się g Hedgesa, zaliczaną również do rodziny d [5, 6].

SMD jest stosowana, gdy badania podają skuteczność w kategoriach pomiaru ciągłego, takiego jak wynik na skali oceny intensywności bólu. Informacje dotyczące SMD przedstawiono w tabeli 1.

PRZYKŁAD 2

Przeprowadzono liczne badania dotyczące stanu zdrowia psychicznego pacjentów. We wszystkich badaniach mierzono depresję, używając różnych skal psychometrycznych. W tej sytuacji konieczne jest ujednolicenie wyników badań według określonej skali przed ich połączeniem. SMD wyraża wielkość efektu interwencji w każdym badaniu w stosunku do zmienności zaobserwowanej w tym badaniu. SMD jest używana jako statystyka podsu-

TABELA 1. Standaryzowana różnica średnich (SMD)

<p>PRZYKŁAD</p> <p>Grupa naukowców prowadząca badania dotyczące atopowego zapalenia skóry chciała w wiarygodny sposób ocenić skuteczność i bezpieczeństwo preparatów emolientowych. W tym celu zastosowano odpowiednie narzędzia do statystycznego opracowania zebranych wyników. Istotnym elementem dalszej analizy danych było określenie miar efektu, świadczące o racjonalności łączenia wyników. W badaniach pierwotnych za miary efektu przyjmuje się m.in. standaryzowaną różnicę średnich (<i>standardized mean difference</i> – SMD) w przypadku skali ciągłej (np. nasilenie świądu w grupie stosującej badany emolient vs placebo lub alternatywny preparat).</p> <p>Definicja: Miara różnicy między grupą badaną a grupą kontrolną uzyskanej w skali ciągłej, gdy w badaniach pierwotnych włączonych do metaanalizy stosowano różne skale pomiarowe (np. zmniejszenia natężenia świądu ocenianego w różnych skalach).</p> <p>Zamienna nazwa: d Cohena, wielkość efektu</p> <p>Wzór:</p> $SMD = \frac{\text{średnia z grupy pacjentów stosujących interwencję} - \text{średnia z grupy pacjentów stosujących placebo}}{\text{odchylenie standardowe wyników wśród pacjentów}^*}$ <p><small>*odchylenie standardowe w populacji.</small></p> <p>Interpretacja: Wskazuje różnice między działaniem leku a placebo lub innego alternatywnego sposobu leczenia ze względu na skalę i niedokładność pomiarów. Może być przełożony na prawdopodobieństwo poprawy stanu pacjentów przyjmujących lek w większym stopniu niż pacjentów przyjmujących placebo lub alternatywny lek.</p> <p>Typowe wartości SMD:</p> <p>0 – nie ma różnic w działaniu między lekiem a placebo lub alternatywnym leczeniem</p> <p>Wartości SMD większe od 0 wskazują stopień, w jakim jedno leczenie jest bardziej skuteczne od drugiego. Konwencjonalnie przyjmuje się następującą interpretację wartości SMD:</p> <p>0–0,2 – mały efekt terapeutyczny</p> <p>0,2–0,5 – umiarkowany efekt terapeutyczny</p> <p>0,5–0,8 – duży efekt terapeutyczny, istotny statystycznie</p> <p>> 0,8 – bardzo duży efekt, wskazujący na pewny efekt terapeutyczny [1, 2]</p>
--

mowująca w metaanalizie, gdy wszystkie badania oceniają ten sam wynik, ale mierzą go na różne sposoby.

Wartość SMD oblicza się, dzieląc obserwowane różnice w średnich przez odpowiednie odchylenie standardowe w każdej próbie (tab. 1). Znormalizowane efekty leczenia są wyrażone jako jednostki odchylenia standardowego i powinny zapewnić, że efekty obserwowane w różnych badaniach mogą być statystycznie połączone niezależnie od rodzaju narzędzia stosowanego do oceny wyniku klinicznego [2].

Warto wiedzieć, że metoda SMD nie koryguje różnic w kierunku skali. Jeśli niektóre skale rosną wraz z nasileniem choroby, a inne maleją, konieczne jest pomnożenie średnich wartości z jednego zestawu badań przez –1 (lub alternatywnie odjęcie średniej od maksymalnej możliwej wartości dla skali), aby zapewnić, że wszystkie

skale wskazują ten sam kierunek. Wszelkie takie korekty należy opisać w sekcji przeglądu dotyczącej metod statystycznych. Odchylenie standardowe nie wymaga modyfikacji [7, 8].

WADY METODY SMD

Należy pamiętać, że oszacowane wielkości efektu z badań klinicznych, zwykle oparte na porównaniach aktywnych leków z placebo lub alternatywnym leczeniem, nie korelują bezpośrednio z reakcjami pacjentów w praktyce klinicznej. Niemniej większa świadomość wśród klinicystów doprowadziłaby do bardziej szczegółowej oceny literatury dotyczącej badań klinicznych niż poleganie wyłącznie na wartościach *p* [8]. Z kolei wiedzę na temat szacowania efektu leczenia można wykorzystać w celu

TABELA 2. SMD w badaniach klinicznych – wybrane przykłady

Autor	Opis badania	Wielkość efektu SMD (95% CI) (wybrane przykłady)
Zeng i wsp. 2018 [9]	ocena skuteczności i bezpieczeństwa stosowania miejscowego niesteroidowego leku przeciwzapalnego w porównaniu z placebo w leczeniu pacjentów z chorobą zwyrodnieniową stawów w okresie 2–12 tygodni	natężenie bólu: diklofenak plastry vs placebo: –0,94 (od –1,20 do –0,68) ibuprofen vs placebo: –0,66 (od –1,06 do –0,25) nimesulid vs placebo: –0,39 (od –0,90 do 0,13) piroksykam vs placebo: –0,50 (od –0,82 do –0,19) indometacyna vs diklofenak żel: 0,29 (–0,11 do 0,69)
Helfer i wsp. 2016 [10]	wpływ terapii skojarzonej z użyciem leków przeciwdepresyjnych i leków przeciwpsychotycznych (LPP) na nasilenie objawów lub na jakość życia u chorych z zaburzeniami schizofrenii (w porównaniu z LPP stosowanymi w monoterapii lub w skojarzeniu z placebo)	nasilenie objawów depresyjnych: –0,25 (od 0,38 do –0,12) nasilenie objawów negatywnych: –0,30 (od –0,44 do –0,16) nasilenie objawów pozytywnych: –0,17 (od –0,33 do –0,01) ogólne nasilenie objawów psychotycznych: –0,24 (od –0,39 do –0,09) jakość życia: –0,032 (od –0,57 do –0,06)
Ford i wsp. 2014 [11]	ocena skuteczności probiotyków i synbiotyków w porównaniu z placebo u dorosłych z zespołem jelita drażliwego (IBS) w ciągu ≥ 7 dni leczenia	probiotyki vs placebo: nasilenie objawów IBS lub bólu brzucha: o 0,25 mniejsze (0,36–0,14) nasilenie wzdęcia brzucha: o 0,15 mniejsze (0,27–0,03) nasilenie oddawania gazów jelitowych: o 0,23 mniejsze (0,38–0,07) nasilenie gwałtownych parć na stolec: o 0,10 mniejsze (od o 0,30 mniejsze do o 0,10 większe) synbiotyki (probiotyki i prebiotyki łącznie) vs placebo: nasilenie objawów lub ból brzucha: o 1,73 mniejsze (od o 3,37 mniejsze do o 0,27 większe)

TABELA 2. Cd.

Autor	Opis badania	Wielkość efektu SMD (95% CI) (wybrane przykłady)
Schmidt-Hansen i wsp. 2015 [12]	ocena skuteczności oksykodonu w porównaniu z innym lekiem opioidowym lub oksykodonom w innej postaci w leczeniu bólu u dorosłych chorych na nowotwór	oksykodon IR vs morfina IR: –0,15 (od –0,79 do 0,49) porównanie z oksykodonom CR: hydromorfon CR: –0,13 (od –0,63 do 0,37) morfina CR: 0,14 (od –0,04 do 0,32) oksykodon IR: 0,10 (od –0,06 do 0,26) oksymorfon ER: 0,23 (od –0,23 do 0,69) tapentadol ER: –0,06 (od –0,3 do 0,18)
Hartling i wsp. 2011 [13]	ocena skuteczności i bezpieczeństwa różnych leków rozszerzających oskrzela i glikokortykosteroidów (stosowanych w monoterapii lub leczeniu skojarzonym) w terapii pierwszego epizodu wirusowego zapalenia oskrzelików u dzieci < 2. roku życia	dzieci leczone ambulatoryjnie: – w grupie otrzymującej adrenalinę w porównaniu z grupą otrzymującą placebo SMD po 1 godz.: 0,45, SMD po 2 godz.: 0,83 – w grupie otrzymującej adrenalinę w porównaniu z grupą otrzymującą GKS SMD po 1 godz.: 0,31 – w grupie otrzymującej adrenalinę w porównaniu z grupą otrzymującą salbutamol SMD po 3–10 dniach: 0,50 – w grupie otrzymującej adrenalinę i deksametazon w porównaniu z grupą otrzymującą placebo SMD po 1 godz.: 0,34 – w grupie otrzymującej adrenalinę i deksametazon w porównaniu z grupą otrzymującą salbutamol SMD po 3–10 dniach: 1,22 – w grupie otrzymującej salbutamol w porównaniu z grupą otrzymującą placebo SMD po 1 godz.: 0,49 – w grupie otrzymującej salbutamol w porównaniu z grupą otrzymującą GKS SMD po 1 godz.: 0,65, SMD po 3–6 godz.: 0,70 dzieci hospitalizowane: – w grupie otrzymującej adrenalinę w porównaniu z grupą otrzymującą salbutamol SMD po 1 godz.: 0,79 SMD po 2 godz.: 0,52 – w grupie otrzymującej GKS w porównaniu z grupą otrzymującą placebo SMD po 3–6 godz.: 1,03, SMD po 6–12 godz.: 0,62 – w grupie otrzymującej salbutamol lub terbutalinę w porównaniu z grupą otrzymującą placebo SMD po 6–12 godz.: 0,81
Kalish i wsp. 2012 [14]	ocena skuteczności glikokortykosteroidów stosowanych miejscowo u chorych na przewlekłe zapalenie błony śluzowej nosa i zatok przynosowych ze współistniejącymi polipami nosa	GKS miejscowe vs placebo: nasilenie objawów: –0,46 (od –0,65 do –0,27) niedrożność nosa: –0,81 (od –1,01 do –0,62) zaawansowanie zmian: –0,49 (od –0,77 do –0,21) rozległość zmian: –0,73 (od –1 do –0,46) przepływ powietrza przez nos: –0,57 (–0,85 do –0,29) GKS miejscowo mniejsza dawka vs większa dawka: nasilenie objawów: –0,29 (od –0,73 do 0,15) rozległość zmian: –0,04 (od –0,44 do 0,36)

IR (*immediate-release*) – postać leku o szybkim uwalnianiu, CR (*controlled-release*) – postać leku o kontrolowanym uwalnianiu, ER (*extended-release*) – postać leku o przedłużonym uwalnianiu, CI (*confidence interval*) – przedział ufności.

promowania przestrzegania zaleceń przez pacjenta poprzez wykorzystanie dowodów epidemiologicznych opisujących ryzyko lub korzyści z leczenia lub interwencji [3]. W tabeli 2 przedstawiono wybrane przykłady badań klinicznych, w których analiza statystyczna obejmowała metodę SMD [9–14].

RÓŻNICA ŚREDNICH OBLICZONA METODĄ LSMD

Do najstarszych i najważniejszych metod obliczeniowych wykorzystywanych w analizach statystycznych należy metoda najmniejszych kwadratów, która polega na wyznaczaniu linii regresji (linii trendu) dla zebranych danych. Metoda ta służy do oszacowania zależności zarówno liniowej, jak i nieliniowej [4].

PRZYKŁAD 3

W jednym z ośrodków badawczych w Polsce prowadzono badania dotyczące wpływu suplementacji witaminą D₃ na poziom markerów prozapalnych we krwi. Metoda najmniejszych kwadratów ma na celu dopasowanie do uzyskanych par wyników (poziom markerów prozapalnych oraz dawka witaminy D₃) takiej linii prostej (model liniowy), która jest do nich najlepiej dopasowana (obliczeniowo), aby ogólny błąd oszacowania (dla wszystkich danych) był jak najmniejszy [4].

Pojęcie regresji jest bardzo szerokie. Przykłady opisywane w niniejszym artykule dotyczą postaci liniowej. W tabeli 3 przedstawiono informacje dotyczące regresji liniowej oraz jej powiązania z LSMD.

W celu lepszego zrozumienia metody LSMD należy zapoznać się z samą ideą analizy regresji.

TABELA 3. Różnica średnich obliczona metodą najmniejszych kwadratów (LSMD) [4]

PRZYKŁAD

Grupa naukowców przeprowadziła badania dotyczące wpływu danego leku na poprawę stanu zdrowia pacjentów w czasie 25 tygodni w porównaniu z placebo. Wyniki przedstawiono w postaci różnicy średnich obliczonej metodą najmniejszych kwadratów, aby określić różnicę w działaniu badanego leku i placebo.

W tym celu w programie służącym do analizy statystycznej wyników przygotowano bazę otrzymanych danych oraz utworzono wykresy. Rycina 1 przedstawia przykład otrzymanego wykresu.

Aby przedstawić dane w postaci różnicy średnich obliczonej metodą najmniejszych kwadratów, należy wyznaczyć linię regresji liniowej. Regresja liniowa mówi nam o kształcie zależności pomiędzy zmiennymi. Punkty na wykresie to wartości zmiennych. Do tych punktów na wykresie trzeba dopasować prostą (linię regresji), która będzie najlepiej charakteryzować zależność między x i y . Otrzymana funkcja pozwala na analizę powiązań między zmiennymi oraz prognozowanie y na podstawie nowo poznanego x .

Aby lepiej zobrazować pojęcie regresji liniowej, na rycinie 2 przedstawiono prostą przechodzącą przez dwa punkty i wyznaczony wzór. Gdy na wykresie mamy więcej niż dwa punkty (ryc. 1), nie jest możliwe, żeby prosta przeszła przez każdy punkt. Metoda, która wyznacza optymalną prostą (linię regresji) i która najlepiej obrazuje zależności pomiędzy x i y , to metoda najmniejszych kwadratów. Polega ona na wyznaczeniu minimum dla sumy kwadratów różnic wartości obserwowanych i wartości teoretycznej.

Ryciny 3 i 4 przedstawiają wykresy zawierające linię regresji.

Parametry regresji można wyznaczyć na utworzonym wykresie poprzez kliknięcie prawym przyciskiem myszy na jeden z punktów i zaznaczenie opcji dodawania linii trendu oraz równania regresji. W omawianym przykładzie stosuje się wykres regresji liniowej. Równania widoczne na wykresach są równaniami matematycznymi wyznaczonych linii. Służą do wyliczenia średniej (LSM).

LSMD – jak liczyć?

Za x należy podstawić tydzień (np. 15) i wykonać obliczenia według wzoru.

RYCINA 1. Przykładowy wykres punktowy otrzymany podczas analizy wyników

RYCINA 2. Wykres przedstawiający prostą przechodzącą przez dwa punkty

RYCINA 3. Przykładowy wykres przedstawiający linię regresji (LSM) otrzymany podczas analizy wyników

RYCINA 4. Przykładowy wykres przedstawiający linię regresji (LSM)

Klinicyści często są zainteresowani prognozowaniem: chcą wiedzieć, który pacjent zachoruje na dane schorzenie, a który nie, który pacjent poradzi sobie dobrze, a który będzie musiał pozostać dłużej w szpitalu, jak będą wyglądały postępy leczenia po zastosowaniu odpowiedniej terapii. Analiza regresji jest przydatna w rozwiązywaniu tego rodzaju problemów. Jak wiadomo, model regresyjny należy do najbardziej znanych metod statystycznych służących do analizy danych. Opiera się on na prognozowaniu lub przewidywaniu danych dla pewnej zmiennej na podstawie innych zmiennych. Krótko mówiąc, odpowiada na pytanie, jaką wartość przyjmie interesująca nas zmienna przy znanej nam wartości innej zmiennej. W tym celu należy za pomocą analizy regresji utworzyć model regresyjny, który przy założeniu błędu statystycznego będzie prognozował poziom i wartość cechy, którą chcemy zbadać [4, 6].

PRZYKŁAD 4

Pracownicy firmy farmaceutycznej produkującej suplementy diety dla kobiet przeprowadzili analizę rynku

w miastach w Polsce, która miała na celu prognozę efektów przyszłej sprzedaży. Na podstawie zebranych danych skonstruowali za pomocą analizy regresji model regresyjny, który wykazywał zależność liniową pomiędzy kobietami zamieszkującymi małe, średnie oraz duże miasta w wieku powyżej 30 lat a ich opinią na temat zażywania suplementów diety w celu poprawy stanu skóry, włosów i paznokci. Dzięki temu firma wprowadzająca nowy produkt może przewidywać lub prognozować z pewnym błędem, jakie będą efekty sprzedaży.

Opisany przykład pokazuje, że analiza regresji służy stwierdzeniu, czy bazując na danej zmiennej lub zestawie zmiennych, można prognozować inną zmienną, czyli zmienną zależną (objaśnianą), i jaki model będzie najdokładniejszy. Należy pamiętać, że model prognozy przewiduje pewien błąd oszacowania. Ten błąd będzie tym mniejszy, im model będzie lepszy. Zadaniem regresji jest zmniejszenie błędu oszacowania do tego stopnia, aby model regresji był użyteczny w swoich prognozach. Innymi słowy – tylko te modele będą spełniały zadanie, które będą cechowały się niskim błędem oszacowania [6, 15].

TABELA 4. LSMD w badaniach klinicznych – wybrane przykłady

Autor	Opis badania	LSMD (95% CI) (wybrane przykłady)
Ghofrani i wsp. 2013 [18]	ocena skuteczności riociguatu w porównaniu z placebo u chorych z przewlekłym zakrzepowo-zatorowym nadciśnieniem płucnym w okresie 16 tygodni	dystans uzyskany w 6MWT (m): 46 (25–67) płucny opór naczyniowy ($\text{dyn} \times \text{s} \times \text{cm}^{-5}$): –246 (od –303 do –190) stężenie NT-proBNP (pg/ml): –444 (od –843 do –45) ciśnienie w tętnicy płucnej (mm Hg): –5 (od –7 do –3) rzut minutowy (l/min): 0,9 (0,6–1,1) objawy niepożądane (%): 1,06 (0,97–1,17) zaprzymanie leczenia z powodu objawów niepożądanych: 1,27 (0,25–6,42)
Cherian i wsp. 2015 [19]	ocena skuteczności genetycznie zmodyfikowanych chondrocytów GEC-TGF- β 1 w porównaniu z placebo u dorosłych z chorobą zwyrodnieniową stawu kolanowego	GEC-TGF-β1 vs placebo w 12. tygodniu: 13,8 (25,0–2,6) GEC-TGF-β1 vs placebo w 52. tygodniu: 13,1 (25,1–1,1)
Calabrese i wsp. 2015 [20]	ocena skuteczności, bezpieczeństwa i tolerancji kariprazyny w niskiej i wysokiej dawce u pacjentów spełniających kryteria DSM-IV-TR dla ostrych epizodów maniakałnych lub mieszanych związanych z zaburzeniem dwubiegunowym typu I	kariprazyna vs placebo: 3–6 mg/dobę –6,1 (od –8,4 do –3,8) 6–12 mg/dobę –5,9 (od –8,2 do –3,6)
Barabássi i wsp. 2017 [21]	codzienne i społeczne funkcjonowanie pacjentów z objawami negatywnymi schizofrenii: analizy <i>post-hoc</i> badania klinicznego fazy III z zastosowaniem kariprazyny w monoterapii i rysperydonu	kariprazyna vs rysperydon: od początku badania do końca leczenia: 4,632 (2,71–6,56) relacje osobiste i społeczne –0,2 (od –0,4 do –0,1) społecznie użyteczne działania –0,4 (od –0,5 do –0,2)
Tuttle i wsp. 2018 [22]	ocena skuteczności baricytynibu (0,75 mg/dobę, 0,75 mg 2 razy dziennie, 1,5 mg/dobę lub 4 mg/dobę)	łączna liczba epizodów hipoglikemii (w ciągu 30 dni): placebo: 0,16 baricytynib (0,75 mg raz dziennie): 0,10 baricytynib (0,75 mg dwa razy dziennie): 0,39 baricytynib (1,5 mg raz dziennie): 0,14 baricytynib (4 mg raz dziennie): 0,44
Donohue i wsp. 2016 [23]	ocena długoterminowego bezpieczeństwa i tolerancji aklidyny 400 mg/formoterolu 2 razy na dobę w porównaniu z formoterolem 12 mg u pacjentów z POChP	FEV₁ i FVC: aklidyna 400 mg/formoterol 12 mg vs formoterol 12 mg poprawa w 1. tygodniu: 87,4 ml i 157,8 ml pod koniec badania: 81,5 ml i 185,4 ml

Kolejnym ważnym zagadnieniem w modelach regresji są zmienne (*variables*), czyli cechy, które badacz może rejestrować, mierzyć, kontrolować lub nimi manipulować (do pewnego stopnia). Należy pamiętać, że modele regresji charakteryzują się brakiem manipulacji zmiennymi niezależnymi. Regresja opiera się na wykonaniu pomiaru zmiennych niezależnych i zmiennych zależnych [16]. Zmienne zależne (objaśniające) opisują czynniki ryzyka – manipulowane przez badacza, np. stężenie roztworu, oznaczane jako x – i są zależne od woli badacza. Zmienne niezależne (wyjaśniane, objaśniane) mogą być jedynie mierzone i rejestrowane przez badacza, który nie ma wpływu na to, jakie wartości przyjmują, np. masa ciała, wiek, stężenie glukozy we krwi. Są oznaczane jako y i są zależne od x . Nie dochodzi tu do manipulacji wartościami zmiennych, dlatego w regresji nie mówi się o wpływie jednej zmiennej na drugą [16, 17]. Chodzi o wyjaśnienie za pomocą jednej zmiennej lub zestawu zmiennych jakiejś innej zmiennej, a nie o wpływanie na nią. Istotna zależność pomiędzy zmienną niezależną a zmienną zależną może dotyczyć jedynie współwystępowania zmiennych, a nie rzeczywistego wpływu jednej zmiennej na drugą [16].

W tabeli 4 przedstawiono zestawienie wybranych przykładów badań klinicznych, w których analiza statystyczna obejmowała metodę LSMD [18–23].

WADY METODY NAJMNIEJSZYCH KWADRATÓW

Metoda najmniejszych kwadratów nie toleruje w zbiorze danych wartości odstających. Wynika to z faktu, że wartość odstająca wpływa na linię regresji. Dlatego w analizie regresji eliminuje się wartości odstające z utworzonych baz danych, aby nie zaburzały postaci linii regresji. Pozostawienie wartości odstających skutkuje zmianą kierunku linii i powoduje, że taki model traci swoją funkcję, jaką jest przewidywanie [4, 15].

LSMD CZY SMD?

Wybór metody zależy od „charakteru” uzyskanych danych. Jeżeli dane układają się w widoczny trend (widać, że wyniki rosną lub nie), wtedy metoda najmniejszych kwadratów będzie bardziej przejrzysta (LSMD). Gdy dane są rozrzucone na skali ocen, lepiej nie wybierać metody LSMD, ponieważ jest to metoda predykcyjna, czyli próbująca przewidzieć wartości na podstawie istniejących danych. Każdy rozrzut danych spowoduje, że przewidywanie (lub predykcja) będzie o wiele mniej dokładne. SMD to wskaźnik liczony na podstawie średnich z wszystkich danych, co zmniejsza wpływ danych odstających, ponieważ przy liczeniu średnich mają one o wiele mniejszy wpływ na średnią [4, 15].

Wartości SMD oraz LSMD należy przedstawiać wraz z podaniem zastosowanego przedziału ufności (tab. 3 i 4), który pokazuje, na ile można ufać danej wartości. Przedział ufności informuje, czy poszukiwana przez nas wartość zawiera się (z założonym prawdopodobieństwem) w określonym przedziale, czyli dostarcza wartości zakresu (od – do), w którym z założonym prawdopodobieństwem znajduje się interesująca nas wartość. Przedział ufności oznacza się jako CI (*confidence interval*), np. 95% CI [24]. Taki przedział oznacza, że mamy 5% szans na pomylenie się w obliczeniach badawczych, czyli 5% szans na to, że prawdziwa wartość średnia znajduje się poza przedziałem, który został wyznaczony. Prawdopodobieństwo może być również większe lub mniejsze. Jeżeli zwiększy się prawdopodobieństwo np. do 99%, wówczas wyznaczony zakres się rozszerzy, natomiast jeśli zmniejszy się prawdopodobieństwo np. do 90%, wtedy zakres się zmniejszy [8].

PODSUMOWANIE

Do prawidłowego i skutecznego stosowania metod analizy danych oraz ich eksploracji w przypadku podejmowania decyzji niezbędna jest wiedza statystyczna. To na badaczu spoczywa obowiązek wyboru odpowiedniej metody opracowania uzyskanych danych. Medycyna i nauki biomedyczne szeroko wykorzystują dane badawcze, dlatego wymagają zastosowania precyzyjnie dobranych metod i podejść, aby analizować coraz bardziej złożone zbiory danych. W przypadku porównywania skuteczności leków doskonale sprawdza się metoda SMD, która jest miarą wielkości efektu, oraz LSMD, zaliczana do metod predykcyjnych. Obie metody umożliwiają prawidłowe oszacowanie efektów leczenia, co wynika z przeglądu literatury dotyczącej badań klinicznych poświęconych skuteczności różnych interwencji terapeutycznych.

KONFLIKT INTERESÓW

Autorzy nie zgłaszają konfliktu interesów.

PIŚMIENNICTWO

1. Bellazzi R, Diomidous M, Sarkar IN, et al. Data analysis and data mining: current issues in biomedical informatics. *Methods Inf Med* 2011; 50: 536-544.
2. Faraone SV. Interpreting estimates of treatment effects implications for managed care. *P&T* 2008; 33: 700-711.
3. McGough JJ, Faraone SV. Estimating the size of treatment effects: moving beyond p values. *Psychiatry* 2009; 6: 21-29.
4. Lenth RV. Least-squares means: the r package ls means. *J Statist Software* 2016; 69: 1-33.
5. Livingston EH, Elliot A, Hynan L, et al. Effect size estimation. a necessary component of statistical analysis. *Arch Surg* 2009; 144: 706-722.

6. Murad MH, Drake MT, Mullan RJ, et al. Comparative effectiveness of drug treatments to prevent fragility fractures: a systematic review and network meta-analysis. *J Clin Endocrinol Metab* 2012; 97: 1871-1880.
7. handbook-51.cochrane.org/chapter_9/9_2_3_2_the_standardized_mean_difference.htm
8. Chavalarias D, Wallach JD, Li AH, et al. Evolution of reporting P values in the biomedical literature, 1990-2015. *JAMA* 2016; 315: 1141-1148.
9. Zeng C, Wei J, Persson MSM, et al. Relative efficacy and safety of topical non-steroidal anti-inflammatory drugs for osteoarthritis: a systematic review and network meta-analysis of randomised controlled trials and observational studies. *Br J Sports Med* 2018; 52: 642-650.
10. Helfer B, Samara MT, Huhn M, et al. Efficacy and safety of antidepressants added to antipsychotics for schizophrenia: a systematic review and meta-analysis. *AJP in Advance* 2016; 173: 876-886.
11. Ford AC, Quigley EMM, Lacy BE, et al. Efficacy of prebiotics, probiotics, and synbiotics in irritable bowel syndrome and chronic idiopathic constipation: systematic review and meta-analysis. *Am J Gastroenterol* 2014; 109: 1547-1561.
12. Schmidt-Hansen M, Bennett MI, Arnold S, et al. Oxycodone for cancer-related pain. *Cochrane Database Syst Rev* 2015; 27: CD003870.
13. Hartling L, Fernandes RM, Bialy L, et al. Steroids and bronchodilators for acute bronchiolitis in the first two years of life: systematic review and meta analysis. *BMJ* 2011; 342: d1714.
14. Kalish L, Snidvongs K, Sivasubramaniam R, et al. Topical steroids for nasal polyps (Review). *Cochrane Database Syst Rev* 2012; 12: CD006549.
15. Nick TG, Campbell KM. Logistic regression. *Methods Mol Biol* 2007; 404: 273-301.
16. Guyatt G, Walter S, Shannon H, et al. Basic statistics for clinicians: correlation and regression. *CMAJ* 1995; 152: 497-504.
17. Mesa JL. Understanding data in clinical research: a simple graphical display for plotting data (up to four independent variables) after binary logistic regression analysis. *Med Hypotheses* 2004; 62: 228-232.
18. Ghofrani HA, D'Armini AD, Grimminger F, et al. Riociguat for the treatment of chronic thromboembolic pulmonary hypertension. *N Engl J Med* 2013; 369: 319-329.
19. Cherian JJ, Parvizi J, Bramlet D, et al. Preliminary results of a phase II randomized study to determine the efficacy and safety of genetically engineered allogeneic human chondrocytes expressing TGF- β 1 in patients with grade 3 chronic degenerative joint disease of the knee. *Osteoarthritis Cartilage* 2015; 23: 2109-2118.
20. Calabrese JR, Keck PE, Starace A, et al. Efficacy and safety of low- and high-dose cariprazine in acute and mixed mania associated with bipolar I disorder: a double-blind, placebo-controlled study. *J Clin Psychiatry* 2015; 76: 284-292.
21. Barabási A, Laszlovszky I, Szatmári B, et al. Day-to-day and social functioning of patients with negative symptoms of schizophrenia: post-hoc analyses of a phase 3 clinical trial with cariprazine monotherapy and risperidone. *25th European Congress of Psychiatry/ European Psychiatry* 2017; 41: 238-302.
22. Tuttle KR, Brosius FC, Adler SG, et al. JAK1/JAK2 inhibition by baricitinib in diabetic kidney disease: results from a phase 2 randomized controlled clinical trial. *Nephrol Dial Transplant* 2018; 33: 1950-1959.
23. Donohue JF, Soong W, Wu X, et al. Longterm safety of aclidinium bromide/formoterol fumarate fixed-dose combination: results of a randomized 1-year trial in patients with COPD. *Respir Med* 2016; 116: 41-48.
24. Murad MH, Montori VM, Ioannidis JPA. How to read a systematic review and meta-analysis and apply the results to patient care users' guides to the medical literature. *JAMA* 2014; 312: 171-179.