

**PRACA POGLĄDOWA**

## **Badania typu *noninferiority* w praktyce klinicznej**

### Noninferiority trials in clinical practice

Monika Marko, Rafał Pawliczak

Zakład Immunopatologii, Wydział Lekarski, Oddział Nauk Biomedycznych, Uniwersytet Medyczny w Łodzi

#### **STRESZCZENIE**

Wraz z rozwojem nowych strategii terapeutycznych coraz trudniej jest opracować leki, które cechują się lepszą skutecznością. Dlatego poszukuje się nowych metod leczenia, które mają zbliżone działanie, ale pod innymi względami oferują lepszą jakość. Trend ten stymuluje rozwój badań klinicznych, których celem jest udowodnienie, że leczenie eksperymentalne nie jest (znacznie) gorsze od leczenia kontrolnego. Takie badania nazywane są badaniami typu *noninferiority* (nie mniejszej skuteczności) (*noninferiority trials*). W niniejszej pracy opisujemy je w odniesieniu do badań typu *superiority* (wyższości) (*superiority trials*) i badań typu *equivalence* (równoważności) (*equivalence trials*). Zwracamy również uwagę na problemy, jakie można napotkać podczas przeprowadzania badań typu *noninferiority*.

#### **SŁOWA KLUCZOWE**

badania typu *noninferiority*, praktyka kliniczna, skuteczność leków.

#### **ABSTRACT**

With the development of new therapeutic strategies, it is becoming increasingly difficult to develop drugs that show better efficacy. That is why new treatments are being sought that have similar efficacy, but in other respects show better quality. This trend stimulates the development of clinical trials aimed at demonstrating that experimental treatment is not (much) worse than control treatment. Such studies are called noninferiority trials, which we describe in this paper in relation to superiority trials and equivalence trials. We also pay attention to the problems you may encounter when conducting noninferiority trials.

#### **KEY WORDS**

noninferiority trials, clinical practice, drugs' efficacy.

#### **ADRES DO KORESPONDENCJI**

prof. dr hab. n. med. Rafał Pawliczak, Zakład Immunopatologii, Wydział Lekarski, Oddział Nauk Biomedycznych, Uniwersytet Medyczny w Łodzi, ul. Żeligowskiego 7/9, bud. II, kl. II, pok. 177,  
e-mail: rafal.pawliczak@csk.umed.lodz.pl

## WSTĘP

Wraz z rozwojem nauk medycznych powstaje coraz więcej alternatywnych metod leczenia tych samych chorób. W wyniku badań mamy do dyspozycji coraz to nowsze interwencje terapeutyczne, co powoduje, że często stajemy przed koniecznością wyboru, która z nich będzie skuteczniejsza. Poszukuje się metody, która przyniesie największe korzyści pacjentom i zminimalizuje koszty leczenia. Lekarz lub inna osoba podejmująca decyzje medyczne oczekuje natomiast od badań klinicznych odpowiedzi na wiele ważnych pytań (tab. 1) [1]. W tym celu można przeprowadzić badania, które porównują nową interwencję z placebo lub skuteczność dwóch interwencji – nowej i obecnie uważanej za najskuteczniejszą [2].

Nauki medyczne opierają się na danych badawczych i wymagają syntezy danych z badań w celu zidentyfikowania najlepszych dostępnych dowodów do opracowania strategii terapeutycznych. Przeglądy systematyczne mają na celu identyfikację wszystkich danych z badań istotnych dla odpowiedzi na pytanie kliniczne. Poszczególne badania rzadko mają wystarczającą moc, aby udzielić ostatecznych odpowiedzi, dlatego stosuje się metody metaanalizy do połączenia danych z podobnych badań [1, 3]. Badania kliniczne mają na celu sprawdzenie tezy postawionej na podstawie aktualnie dostępnej wiedzy poprzez przeprowadzenie eksperymentu medycznego. W takim przypadku teza odnosi się do interwencji medycznej, którą badacz ma zamiar oceniać. Najczęściej dokonuje się oceny skuteczności i bezpieczeństwa danej interwencji w odniesieniu do jednej lub większej liczby metod terapeutycznych [1]. Zazwyczaj badania te mają na celu udowodnienie, że nowa interwencja, czyli np. nowy lek, jest lepsza od dotychczas stosowanych metod lub równoważna.

## TYPY BADAŃ KLINICZNYCH – BADANIA RÓWNOWAŻNOŚCI ORAZ BADANIA WYŻSZOŚCI

Za skuteczność leczenia uważa się zdolność danej interwencji do uzyskania korzystnego efektu terapeutycznego.

Zastosowanie grup kontrolnych w badaniach klinicznych jest ważne dla różnicowania wyników leczenia pacjentów pochodzących z badań eksperymentalnych. W tabeli 2 przedstawiono zestawienie oraz opis aktualnie przeważających typów badań klinicznych stosowanych w badaniach porównawczych leków [2].

W tradycyjnym badaniu klinicznym nową terapię porównuje się ze standardowym leczeniem lub placebo w celu wykazania, że ma ona większą skuteczność. Hipoteza zerowa dla takiego badania będzie mówiła, że dwa zabiegi dają taki sam efekt. Odrzucenie tej hipotezy będzie sugerować, że efekty są różne. Chociaż nowe leczenie może dawać lepsze lub gorsze efekty, typowe badanie ma na celu stwierdzenie większej skuteczności nowego leczenia i jest znane jako badanie typu *superiority* (badanie wyższości) (*superiority trials*) (ryc. 1). Ponieważ badanie wyższości jest w stanie zidentyfikować zarówno szkodliwe, jak i korzystne skutki nowej terapii w porównaniu z kontrolą (obecną terapią), można w tym przypadku zastosować 95% CI (*confidence interval*) do wskazania górnej i dolnej granicy różnicy w działaniu leczniczym zgodnym z obserwowanymi danymi [4].

Badania typu *superiority* wskazujące, że badana interwencja jest lepsza, przeprowadzane są najczęściej w badaniach klinicznych produktów leczniczych we wczesnych fazach, zwykle gdy w grupie kontrolnej stosowane jest placebo. Badania te mają na celu udowodnienie, że oceniana metoda terapeutyczna jest skuteczniejsza od interwencji zastosowanej w grupie kontrolnej. Wyniki badania klinicznego poddawane są ocenie statystycznej i koniecz-

**TABELA 1.** Najczęściej spotykane problemy podczas analizy skuteczności leków

- Która z obecnie stosowanych interwencji jest skuteczniejsza?
- Którą interwencję zastosować u danego pacjenta?
- Która interwencja jest bardziej opłacalna?
- Którą interwencję należy refundować?
- Jaka jest skuteczność badanego leku w przypadku występowania czynników zakłócających (np. wpływ chorób współistniejących)?

**TABELA 2.** Typy badań klinicznych stosowanych w porównywaniu skuteczności leków

| Typ badania  | Co porównuje  | Zadawane pytanie  |
|--|---|---|
| badana interwencja jest lepsza – badanie wyższości ( <i>superiority trial</i> )                          | badana interwencja porównywana jest z placebo   | Czy interwencja jest bardziej skuteczna niż placebo?  |
| badana interwencja nie jest gorsza – badanie nie mniejszej skuteczności ( <i>non-inferiority trial</i> ) | porównywana jest skuteczność dwóch interwencji: badanej i obecnie stosowanej, uważanej za skuteczną | Czy badana interwencja nie jest mniej skuteczna niż alternatywna, aktualnie stosowana w praktyce? |
| badana interwencja jest równoważna – badanie równoważności ( <i>equivalence trial</i> )                  |   | Czy badana interwencja jest równie skuteczna jak alternatywna, aktualnie stosowana w praktyce?    |



**RYCINA 1.** Badania typu *noninferiority* a badania równoważności i wyższości – porównanie

ne jest określenie parametrów takiej oceny. Standardowo przyjmuje się wartość  $p$  na poziomie niższym niż 0,05. Oznacza to, że prawdopodobieństwo otrzymania w prawidłowo przeprowadzonym badaniu wyników przypadkowych, nie będzie większe niż 5% [4, 5].

Badania typu *equivalence* (badania równoważności) (*equivalence trials*) udowadniają, że badana interwencja jest równoważna (ryc. 1). Wykorzystywane są zazwyczaj w kontekście równoważności biologicznej. Przykładem może być sytuacja, kiedy chcemy wykazać, że badany lek jest równoważny pod względem parametrów farmakokinetycznych z lekiem, dla którego wygasła ochrona patentowa. W tym przypadku najważniejszym porównywanym parametrem będzie całkowita ilość wchłoniętego leku mierzona jako pole pod krzywą wykresu stężenia leku w surowicy na osi czasu [6].

Chociaż lepszą lub gorszą skuteczność nowego leczenia można wykazać w badaniu wyższości, ogólnie błędne byłoby stwierdzenie, że brak znaczącej różnicy w tym badaniu wskazuje, że terapie mają podobne skutki. Brak różnicy nie jest wiarygodnym dowodem. Potrzebne jest kontrolowane aktywnie badanie typu *noninferiority* (badanie nie mniejszej skuteczności) (*noninferiority trials*) (ryc. 1), aby ustalić, czy nowa interwencja, która oferuje inne korzyści, takie jak mniejsza toksyczność lub niższy koszt, nie ma mniejszej skuteczności niż dotychczasowe leczenie [4].

### **BADANIA TYPU NONINFERIORITY (NIE MNIEJSZEJ SKUTECZNOŚCI)**

Zespół przeprowadzający badania kliniczne często zainteresowany jest tym, czego dane badanie może dowieść, a co jedynie zasugerować. Badania mogą dowodzić wyższości (*superiority trials*), równoważności (*equivalence trials*) lub potwierdzać, że badana interwencja nie jest gorsza niż ta, z którą jest porównywana (*noninferiority trials*) [2].

Czasami celem porównania nowego leczenia ze standardowym nie jest znalezienie bardziej skutecznego podejścia, lecz terapii, która ma inne zalety, takie jak niższy koszt, mniej działań niepożądanych lub większa wygoda stosowania z co najmniej podobną skutecznością do standardowego leczenia. Dodatkowe zalety będzie miała terapia, która jest prawie tak samo skuteczna jak standardowa i może mieć zastosowanie w praktyce lub tylko u niektórych pacjentów. Celem badań typu *noninferiority* jest rygorystyczna ocena nowej interwencji w stosunku do obecnie stosowanego i skutecznego leczenia w celu stwierdzenia, że jest co najmniej prawie tak samo dobra (tj. nie gorsza) [4, 7]. Przykłady wykorzystania modelu badań nie mniejszej skuteczności przedstawiono w tabeli 3 [8–12].

Badania nie mniejszej skuteczności oparte są na konstrukcji typu „badana interwencja nie jest gorsza”, przy założeniu, że terapia wybrana do porównania nie będzie lepsza od testowanego leku. W badaniach tego typu grupą kontrolną jest standardowy, uważany za najbardziej skuteczny obecnie stosowany lek. Jako kontrole wykorzystuje się znane skuteczne metody leczenia, ponieważ niewiele można zyskać, wykazując, że nowa terapia nie jest gorsza od leczenia pozorowanego lub placebo [4].

W tym przypadku nie można oczekiwać, że badana interwencja będzie lepsza od leku porównawczego. Zakłada się, że satysfakcjonującym wynikiem będzie nie mniejsza skuteczność nowej metody terapii niż leku wybranego do porównania (charakteryzującego się największą skutecznością). Należy jednak określić, jaka różnica w skuteczności porównywanych terapii będzie zadowalająca na poziomie klinicznym. Ta różnica, oznaczana jako  $\Delta$  (delta), zależy od dziedziny terapeutycznej, skuteczności interwencji kontrolnej, wybranych punktów końcowych oraz przejętej miary skuteczności. Wartość  $\Delta$  określa próg istotności wyznaczający maksymalną dopuszczalną różnicę między badanymi grupami na niekorzyść badanej metody terapeutycznej. Można wówczas uznać, że rzeczywistość nie jest ona gorsza z klinicznego punktu widzenia [13]. Aby oszacować wartość  $\Delta$ , można zastosować różne metody statystyczne. Należy pamiętać, że im mniejsza jest  $\Delta$ , tym większą liczbę pacjentów należy zrekrutować do badania. Jednak często skutkuje to ryzykiem zastosowania zbyt dużej wartości  $\Delta$ , która również wymaga interpretacji klinicznej [13].

Badania nie mniejszej skuteczności mogą być przeprowadzane również w celu wykazania, że nowe leczenie jest lepsze niż zakładane placebo w sytuacjach, w których przeprowadzenie próby kontrolnej placebo jest nieetyczne. Mogą być także przydatne, gdy nowe leczenie może zaoferować ważne zalety w stosunku do obecnie dostępnych standardowych zabiegów pod względem poprawy bezpieczeństwa, wygody lub kosztów. Ponadto coraz częściej wymagane są badania kliniczne w celu stwierdze-

TABELA 3. Badania typu *noninferiority* – przykłady

| Przykład | Autor                    | Opis badania   | Wyniki i wnioski   |
|----------|--------------------------|--|--|
| 1        | Diercks i wsp. 2019 [8]  | zbadanie wpływu ibuprofenu na krwawienie po tonsillektomii, wymagające interwencji chirurgicznej u dzieci w porównaniu z acetaminofenem  | częstość krwawień wymagających interwencji operacyjnej wyniosła 1,2% w grupie przyjmującej acetaminofen i 2,9% w grupie otrzymującej ibuprofen; nie stwierdzono istotnych zdarzeń niepożądanych ani zgonów                           |
| 2        | Ginsburg i wsp. 2019 [9] | ocena skuteczności leczenia przeciwbólowego w przypadku ciężkiego zapalenia płuc w porównaniu z amoksyycyliną  | stosowanie placebo było znacznie mniej skuteczne niż leczenie amoksyycyliną  |
| 3        | Hood i wsp. 2019 [10]    | ocena skuteczności kwasu acetylosalicylowego w porównaniu z profilaktyką przeciwzakrzepową w celu zmniejszenia ryzyka wystąpienia pooperacyjnej choroby zakrzepowo-zatorowej u pacjentów poddawanych alloplastyce stawu kolanowego | kwas acetylosalicylowy nie był gorszy od innych leków przeciwzakrzepowych; kwas acetylosalicylowy może zapewnić podobną ochronę przed pooperacyjną chorobą zakrzepowo-zatorową jak inne leki przeciwzakrzepowe                       |
| 4        | Tanum i wsp. 2017 [11]   | ocena skuteczności naltreksonu o przedłużonym uwalnianiu w porównaniu z chlorowodorkiem naloksonu u pacjentów uzależnionych od opioidów  | naltrekson działał równie skutecznie jak chlorowodorek naloksonu podczas krótkotrwałej abstynencji od heroiny i innych nielegalnych substancji; naltrekson może być uznawany za opcję terapeutyczną u osób uzależnionych od opioidów |
| 5        | Garin i wsp. 2014 [12]   | ocena skuteczności $\beta$ -laktamu w porównaniu z kombinacją $\beta$ -laktamu i makrolidu w umiarkowanym zapaleniu płuc   | nie stwierdzono gorszej skuteczności $\beta$ -laktamu u pacjentów hospitalizowanych z powodu umiarkowanego zapalenia płuc  |

nia korzyści w klinicznych punktach końcowych, a nie zastępczych punktach końcowych, mimo że malejące są dodatkowe korzyści wynikające z nowych metod leczenia, co jest również ważnym czynnikiem przy określaniu wielkości próby. Może to skutkować również pojawianiem się tendencji do projektowania badań klinicznych, które mają na celu wykazanie, że terapie eksperymentalne mają działanie podobne do kontroli o udowodnionej skuteczności, a nie lepsze. Jednak w badaniach typu *noninferiority* projektowanie próby i interpretacja wyników są trudniejsze niż w badaniach wyższości [14].

Istnieje wiele różnych okoliczności, które mogą uzasadniać podjęcie próby przeprowadzenia badania typu *noninferiority*. Najprostszym przykładem jest stwierdzenie (jeśli jest to prawda), że skuteczność nowego leku jest taka sama jak istniejącego leku i nie przewiduje się żadnych innych różnic. Będzie to szczególnie prawdopodobne, jeżeli leki należą do tej samej klasy. W takim przypadku opracowanie leku *me too* może prowadzić do powstania podobnego produktu rynkowego, ale nie do znaczącej poprawy opieki. Zalety tak wąskiej różnorodności produktów w ramach określonej klasy leków mogą wydawać się niewielkie, z wyjątkiem zysków firmy. Jeśli jednak nawet średnie korzyści i profile bezpieczeństwa dwóch leków w klasie wydają się identyczne, możliwe jest, że poszczególni pacjenci mogą odnieść więcej korzyści z zastosowania jednego leku niż drugiego [6].

Badania nie mniejszej skuteczności mogą być wykorzystywane w wielu przypadkach (tab. 4) [15]. Implikują kompromis między utratą skuteczności standardowego leczenia a innymi korzyściami nowej terapii. Są coraz częściej wykorzystywane do porównań nowatorskich metod leczenia ze standardowymi aktywnymi kontrolami. Jednak ich konstrukcja jest złożona – klinicyści, pacjenci, a nawet badacze często mają trudności z pełnym zrozumieniem tej metodologii [16].

## WYZNACZANIE MARGINESU PORÓWNYWALNOŚCI

Badania kliniczne z randomizacją, czasami nazywane badaniami bezpośrednimi, mogą ocenić hipotezy typu *superiority* (większej skuteczności) lub typu *inferiority* (mniejszej skuteczności) interwencji eksperymentalnej

TABELA 4. Kiedy wybrać model badań *noninferiority*? [15]

- Przeprowadzenie badania z kontrolą placebo (*placebo-controlled trial*) nie jest etycznie możliwe
- Nie oczekuje się, że leczenie eksperymentalne będzie lepsze niż kontrola w pierwotnym punkcie końcowym (*primary end point*), ale będzie lepsze w drugorzędowych punktach końcowych lub jest bezpieczniejsze
- Nie oczekuje się, że leczenie eksperymentalne będzie lepsze niż kontrola w pierwotnym punkcie końcowym, ale jest tańsze w produkcji lub łatwiejsze w dawkowaniu

w porównaniu z najlepszym standardem leczenia. W badaniach typu *superiority* podstawową hipotezą jest lepsza skuteczność nowych interwencji potwierdzona poprzez wykluczenie jakiegokolwiek liczby gorszej skuteczności interwencji eksperymentalnych w porównaniu z istniejącym najlepszym standardem opieki. Hipoteza o nie mniejszej skuteczności badanego leku pozwala na pewną, klinicznie dopuszczalną „gorszą skuteczność” nowych interwencji w porównaniu ze znanymi skutecznymi interwencjami w zamian za potencjalne korzyści. Leki o nie mniejszej skuteczności mogą być mniej toksyczne, mniej inwazyjne, mniej kosztowne, wymagać mniejszych dawek lub ich stosowanie może się wiązać z poprawą jakości życia pacjentów. „Mniejsza skuteczność” jest określana poprzez wykluczenie pewnej ilości „gorszej skuteczności” wybranej przed badaniem, zwanej „marginesem porównywalności” (braku niższości) (*noninferiority margin* –  $\Delta$ NI) [17].

Zwyczaj do wyznaczenia  $\Delta$ NI stosuje się zarówno metody statystyczne, jak i osąd kliniczny. Jedną z powszechnych metod jest oszacowanie skuteczności standardowego leczenia w porównaniu z placebo, a następnie, na podstawie opinii eksperta, określenie części wielkości efektu standardowego leczenia, którą należy zachować (zwykle 50%) [16]. Determinanty marginesu porównywalności obejmują znany efekt standardowego leczenia w porównaniu z placebo, ciężkość choroby, toksyczność leku, efekty uboczne lub koszt standardowego leczenia i główny punkt końcowy. Mniejszy margines porównywalności jest prawdopodobnie odpowiedni, jeśli badana choroba jest ciężka lub głównym punktem końcowym jest zgon [4].

Wielkość próby wymagana do rzetelnego wykazania nie mniejszej skuteczności zależy zarówno od marginesu porównywalności, jak i założonych rzeczywistych względnych efektów porównywanych terapii. Badanie typu *noninferiority* często wymaga większej liczebności próby niż badania wyższości, ponieważ  $\Delta$ NI stosowane w badaniach typu *noninferiority* są na ogół mniejsze niż różnice poszukiwane w badaniach wyższości. Równie ważny jest zakładany efekt leczenia eksperymentalnego w stosunku do aktywnej kontroli.

W celu wyznaczenia marginesu porównywalności można kierować się zasadami określonymi w wytycznych *International Conference on Harmonisation* (ICH) [18, 19]. Pierwsza z nich mówi, że ustalenie porównywalności opiera się zarówno na uzasadnieniu statystycznym, jak i na ocenie klinicznej i powinno odzwierciedlać niepewność dowodów, na których opiera się wybór. Druga natomiast sugeruje, że margines ten nie może być większy niż najmniejsza wielkość efektu, jakiej można by oczekiwać od leku aktywnego w porównaniu z placebo [2].

Ponieważ badania typu *noninferiority* mają jedynie na celu stwierdzenie nie mniejszej skuteczności, są analizowane przy użyciu jednostronnego przedziału ufności lub testu hipotezy.

Interpretując  $\Delta$ NI, badacze muszą często podejmować samodzielne decyzje [16].

## CZY PACJENCI POWINNI BYĆ INFORMOWANI O UDZIALE W BADANIU TYPU *NONINFERIORITY*?

Badania typu *noninferiority* mają na celu ustalenie, czy interwencja nie jest nieakceptowalnie gorsza od leczenia kontrolnego – o więcej niż wcześniej określona różnica, zwaną marginesem porównywalności. Amerykańska Agencja ds. Żywności i Leków (*Food and Drug Administration* – FDA) [20] podała wytyczne określające, kiedy projekt badania nie mniejszej skuteczności musi być zastosowany w celu ustalenia skuteczności nowego leku, zwykle dlatego że stosowanie placebo lub „brak leczenia” nie byłoby etyczne. Inne wytyczne dotyczące badań nie mniejszej skuteczności obejmują instrukcje dotyczące interwencji z bardziej korzystnymi profilami bezpieczeństwa lub wygodniejszym dawkowaniem, mniejszym obciążeniem pacjenta, mniejszymi kosztami lub innymi korzyściami, dla których pacjenci i klinicyści mogą zaakceptować zmniejszoną skuteczność jako kompromis [21].

Doshi i wsp. [17] przeanalizowali formularze świadomej zgody pacjenta stosowane w badaniach nie mniejszej skuteczności i wykazali, że rzadko są one wystarczająco jasne, aby poinformować pacjenta, że bierze udział w badaniu przeciwnym do badania wyższości. Wnioskują, że jest to poważny błąd etyczny. Można stwierdzić, że prawdziwym problemem nie jest ujawnianie celów badań, ale raczej pewność, że potencjalni uczestnicy otrzymali odpowiednie informacje dotyczące ryzyka. Autorzy ci proponują konkretny schemat formularza zgody. Podkreślałby on, że badanie ma na celu ustalenie, czy nowy lek jest nie więcej niż np. 10% mniej skuteczny niż standardowe leczenie. Zawierałby też oświadczenie, że nowy lek może być lepszy dla niektórych pacjentów (np. z powodu mniejszej liczby działań niepożądanych). Niemniej wielu pacjentów w takiej sytuacji może dojść do wniosku, że nowy lek w przeważającej części nie jest tak dobry jak lek porównawczy lub że być może nie powinni poświęcać swojego czasu i wysiłku, nie mówiąc już o ryzyku dla zdrowia, aby wesprzeć to z pozoru mało użyteczne przedsięwzięcie [17].

Prawdopodobnie informowanie o udziale w badaniu nie mniejszej skuteczności może dezinformować potencjalnych uczestników. W rzeczywistości głównym uzasadnieniem dla badań nie mniejszej skuteczności jest to, że są pomocne w korygowaniu pewnego rodzaju stronniczości badań, a także pozwalają na zatwierdzenie niektórych rodzajów leków [22].

Zakładając, że udział w takim badaniu wiąże się z dodatkowym ryzykiem, nie ma wątpliwości, że ryzyko to powinno zostać wyraźnie wskazane potencjalnym

uczestnikom. Właściwym sposobem jest uwzględnienie w formularzach zgody tzw. sekcji ryzyka. W rzeczywistości nie wiadomo, co opisują formularze zgody na temat ryzyka – przydatne byłoby ustalenie, o czym informują formularze zgody w badaniach nie mniejszej skuteczności w porównaniu z badaniami wyższości [17, 22].

## BADANIA NIE MNIEJSZEJ SKUTECZNOŚCI – PUŁAPKI

W badaniach typu *noninferiority* szczególnie ważne jest przestrzeganie dobrze zdefiniowanego protokołu badania, a także jego udokumentowanie. Do głównych trudności zalicza się: wybór pacjentów, zgodność leczenia, sposoby pomiarów wyników, czas trwania leczenia i oceny oraz analizę statystyczną [6].

W celu wykazania, że nowe leczenie jest lepsze niż placebo lub aktywna kontrola, analizuje się statystycznie dane z badania, aby ustalić, czy wynik dostarcza wystarczających dowodów, aby odrzucić hipotezę zerową, tj. że dwa zabiegi mają taki sam efekt. Przykładowo, jeśli w kontrolowanym badaniu z randomizacją zaobserwowano obniżenie ciśnienia krwi o 15 mm Hg dla nowo opracowanego leku przeciwnadciśnieniowego, podczas gdy aktywna kontrola indukuje zmniejszenie tylko o 5 mm Hg, błędne jest stwierdzenie, że nowy lek jest bardziej skuteczny na podstawie zaobserwowanej różnicy 10 mm Hg, ponieważ taka różnica może wynikać z możliwości pobrania próbki [2, 14]. Zamiast tego obliczamy przedział ufności (CI) dla obserwowanej różnicy, co pozwala na zachowanie pewnego stopnia niepewności związanego z obserwowaną wartością 10 mm Hg, i to właśnie w tym CI prawdopodobnie leży prawdziwa różnica. Na przykład, jeśli oblicza się 95% CI, oznacza to, że przy wielokrotnym pobieraniu próbek i szacowaniu różnic można oczekiwać, że 95% szacowanych różnic będzie miało wartość od 3 do 17, a jedna z nich może być prawdziwą różnicą [5, 14].

Niektóre wartości mogą również wykraczać poza przedział ufności i chociaż mogą to być również praw-

dziwe wartości, prawdopodobieństwo to jest bardzo małe. Dlatego w tym przykładzie dochodzimy do wniosku, że prawdziwa różnica prawdopodobnie nie będzie wynosić 0. Istnieją wystarczające dowody, aby stwierdzić, że nowe leczenie przewyższa kontrolę, gdy wartość  $p$  wynosi  $< 0,05$ . Należy zauważyć, że jeśli z powodu małej wielkości próby CI jest bardzo szeroki, badanie prawdopodobnie nie będzie w stanie wykazać wyższości. Stwierdzenie równoważności na podstawie nieistotnego testu hipotezy zerowej, tj. braku różnicy między leczeniem eksperymentalnym a aktywną kontrolą, jest niewłaściwe [14].

Klinicyści powinni szczególnie pamiętać, że marginesy porównywalności są często zbyt duże, aby mieć znaczenie kliniczne, oraz że twierdzenie o równoważności może być również mylące, jeśli badanie nie zostało przeprowadzone zgodnie z obowiązującymi standardami [5].

Naiwnym podejściem do analizy danych eksperymentalnych jest porównywanie nowych i kontrolnych zabiegów w standardowy sposób i deklarowanie równoważności leczenia, jeśli nie wykryto żadnej różnicy. Problem z takim podejściem polega na tym, że jeśli wielkość próby jest zbyt mała, aby CI był zbyt szeroki, równoważność może zostać niewłaściwie stwierdzona. Jeśli celem jest ocena prawdziwej „równoważności”, hipoteza zerowa powinna opierać się na fakcie, że metody leczenia różnią się i należy zbadać, czy istnieją wystarczające dowody do stwierdzenia ich równoważności i odrzucenia hipotezy zerowej. Ponieważ praktycznie niemożliwe jest, aby dwa zabiegi miały dokładnie równoważne efekty, „równoważność” pod względem oceny klinicznej oznacza, że efekty tych zabiegów różnią się nie więcej niż dopuszczalną ilością, znaną jako margines porównywalności [14].

Problemem może być również fakt, że badania typu *noninferiority* mogą tylko wykazać, że nowa interwencja nie jest gorsza od kontroli o więcej niż określoną wcześniej ilość. Producenci leków mogą jednak nie chcieć oświadczyć, że ich produkt nie był gorszy od standardowego produktu w odniesieniu do z góry określonego marginesu najmniejszej istotnej klinicznie różnicy [5].

**TABELA 5.** Wytyczne dotyczące badań typu *noninferiority*

|   |  |
|---|--|
| CONSORT ( <i>Consolidated Standards of Reporting Trials</i> ) [23]            | instrukcje dotyczące raportów z badań typu <i>noninferiority</i>   |
| FDA 2010 ( <i>Food and Drug Administration</i> ) [20]                         | dokument koncentrujący się na wszystkich zagadnieniach związanych z badaniami typu <i>noninferiority</i> podający ogólne wytyczne                              |
| EMA 2006 i 2010 ( <i>Evaluation of Medicinal Products</i> ) [24]              | instrukcje z 2010 roku dotyczą badań typu <i>noninferiority</i> i <i>superiority</i> , natomiast wytyczne z 2006 roku dotyczą wyboru marginesu porównywalności |
| ICH E9, ICH E10 ( <i>International Conference on Harmonisation</i> ) [18, 19] | dokumenty zawierające wytyczne statystyczne dotyczące problemów napotykanym we wszystkich badaniach klinicznych i ich projektach                               |

**TABELA 6.** Badania typu *noninferiority* – podsumowanie informacji

- Badania nie mniejszej skuteczności oceniają, czy nowa interwencja nie jest znacznie gorsza w porównaniu ze standardowym leczeniem
- Badania te odpowiadają na pytanie, czy jesteśmy gotowi zaakceptować nową interwencję, która może być klinicznie gorsza, ale nadal jest korzystna dla pacjentów ze względu na takie zalety, jak mniej intensywne leczenie, niższe koszty lub mniej skutków ubocznych
- Badania nie mniejszej skuteczności oraz badania równoważności są czasami błędnie stosowane zamiennie
- Badania równoważności mają na celu stwierdzenie, że nowa interwencja działa niewiele gorzej i niewiele lepiej niż standardowe leczenie
- Badania równoważności i nie mniejszej skuteczności różnią się od badań wyższości, które mają na celu wykazanie, że nowa interwencja działa lepiej w porównaniu z kontrolą

Opublikowano kilka wytycznych, które mogą pomóc naukowcom w stosowaniu badań typu *noninferiority* (tab. 5) [18–20, 23, 24]. Istnieje jednak pewna niespójność między tymi wytycznymi, co może mieć negatywny wpływ na ogólną jakość badania. Badania nie mniejszej skuteczności wymagają większej uwagi w przypadku niektórych zagadnień, dlatego konieczne są jasne instrukcje dotyczące ich projektowania [13].

Margines porównywalności to wartość, która pozwala, aby nowe leczenie było „akceptowalnie gorsze”. Służy jako odniesienie do wniosków na temat nie mniejszej skuteczności leczenia. We wszystkich wytycznych zaleca się, aby margines ten był wybierany klinicznie, co oznacza maksymalny akceptowalny zakres, w którym nowy lek może być mniej skuteczny niż standardowy i nadal wykazywać pożądane działanie. Nie jest jednak jasne, czy względy statystyczne powinny również wpływać na wybór odpowiedniego marginesu, zgodnie z zaleceniami projektu FDA 2010 [20], ICH E10 (*International Conference on Harmonisation*) [19] i wytycznych EMEA 2006 (*Evaluation of Medicinal Products*) [24]. Ignorowanie dowodów statystycznych z metaanaliz lub przeglądów systematycznych może skutkować wyborem nierealnego marginesu przez badaczy [13].

Negatywny wynik badania nie mniejszej skuteczności zasadniczo nie wykazuje gorszej jakości leczenia eksperymentalnego, podobnie jak negatywny wynik badania wyższości nie wykazuje równoważności dwóch zabiegów. Badanie typu *noninferiority* jest podobne do badania równoważności, ponieważ celem obu jest stwierdzenie, że interwencja odpowiada działaniu ustalonej terapii w określonym wcześniej zakresie. Jednak celem badania nie mniejszej skuteczności jest jedynie wykazanie, że leczenie eksperymentalne nie jest zasad-

niczo gorsze niż leczenie standardowe, podczas gdy badanie równoważności ma wykazać, że leczenie eksperymentalne nie jest ani gorsze, ani lepsze od leczenia standardowego [5].

## PODSUMOWANIE

Coraz częstsze stosowanie badań typu *noninferiority* oraz związane z nimi trudności napotymane przez badaczy sugerują, że ważne jest zrozumienie problemów związanych z jakością w projektowaniu i analizie tych badań. W badaniu nie mniejszej skuteczności nowe leczenie jest porównywane z istniejącym, już dostępnym leczeniem, ale analiza statystyczna nie ocenia możliwej wyższości nowego leczenia. Zamiast tego sprawdza, czy nowe leczenie nie jest gorsze od istniejącego, stosując wstępnie ustalony próg oceny (margines porównywalności lub margines braku niższości). Badania kliniczne z randomizacją zaprojektowane z zastosowaniem hipotezy nie mniejszej skuteczności sprawdzają, czy leczenie eksperymentalne jest gorsze od standardowego leczenia o dany margines. Jeśli nowy lek lub strategia zmniejszy koszty lub liczbę działań niepożądanych albo poprawi łatwość podawania, pacjenci mogą w uzasadniony sposób zaakceptować możliwość niewielkiego zmniejszenia skuteczności terapii. W projektowaniu i interpretacji badań klinicznych nie mniejszej skuteczności należy wziąć pod uwagę preferencje pacjentów, dlatego że pacjenci i klinicyści mają różne wymagania dotyczące leczenia.

Istnieją jednak liczne obawy związane z definiowaniem, projektowaniem, przeprowadzaniem, raportowaniem i interpretowaniem badań typu *noninferiority*. Bez odpowiedzi pozostaje pytanie, do jakiego stopnia są one uzasadnione. Pojawiają się również problemy etyczne związane ze świadomą zgodą pacjenta. Podsumowanie najważniejszych informacji przedstawiono w tabeli 6.

## KONFLIKT INTERESÓW

Autorzy nie zgłaszają konfliktu interesów.

## PIŚMIENNICTWO

1. Lee AW. Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. *J Clin Epidemiol* 2014; 67: 138-43.
2. D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues – the encounters of academic consultants in statistics. *Stat Med* 2003; 22: 169-86.
3. Edwards SJ, Clarke MJ, Wordsworth S, et al. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Pract* 2009; 63: 841-54.
4. Kaji AH, Lewis RJ. Noninferiority Trials. Is a new treatment almost as effective as another? *JAMA* 2015; 23: 2371-2.

5. Gotzsche PC. Lessons from and cautions about noninferiority and equivalence randomized trials. *JAMA* 2006; 295: 1172-4.
6. Pocock SJ. The pros and cons of noninferiority trials. *Fundam Clin Pharmacol* 2003; 17: 483-90.
7. Gyawali B, Tessema FA, Jung EH, et al. Assessing the justification, funding, success, and survival outcomes of randomized noninferiority trials of cancer drugs a systematic review and pooled analysis. *JAMA Network Open* 2019; 2: e199570.
8. Diercks GR, Comins J, Bennett K, et al. Comparison of ibuprofen vs acetaminophen and severe bleeding risk after pediatric tonsillectomy a noninferiority randomized clinical trial. *JAMA Otolaryngol Head Neck Surg* 2019; 145: 494-500.
9. Ginsburg AS, Mvalo T, Nkwopara E, et al. Placebo vs amoxicillin for nonsevere fast-breathing pneumonia in malawian children aged 2 to 59 months a double-blind, randomized clinical noninferiority trial. *JAMA Pediatr* 2019; 173: 21-8.
10. Hood BR, Cowen ME, Zheng HT, et al. Association of aspirin with prevention of venous thromboembolism in patients after total knee arthroplasty compared with other anticoagulants. A noninferiority analysis. *JAMA Surg* 2019; 154: 65-72.
11. Tanum L, Solli KK, Latif ZE, et al. Effectiveness of injectable extended-release naltrexone vs daily buprenorphine-naloxone for opioid dependence. a randomized clinical noninferiority trial. *JAMA Psychiatry* 2017; 74: 1197-205.
12. Garin N, Genné D, Carballo S, et al. Beta-lactam monotherapy vs beta-lactam-macrolide combination treatment in moderately severe community-acquired pneumonia. a randomized noninferiority trial. *JAMA Intern Med* 2014; 174: 1894-901.
13. Rehal S, Morris TP, Fielding K, et al. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open* 2016; 6: e012594.
14. Hahn SK. Understanding noninferiority trials. *Korean J Pediatr* 2012; 55: 403-7.
15. Lesaffre E. Superiority, equivalence, and non-inferiority trials. *Bull NYU Hosp Jt Dis* 2008; 66: 150-4.
16. Acuna AS. Incorporating patient preferences in noninferiority trials. *JAMA* 2019; 322: 305-6.
17. Doshi P, Hur P, Jones M, et al. Informed consent to study purpose in randomized clinical trials of antibiotics, 1991 through 2011. *JAMA Intern Med* 2017; 177: 1452-9.
18. International Conference on Harmonisation. Statistical principles for clinical trials (ICH E 9). Food and Drug Administration, DHHS, 1998.
19. International Conference on Harmonisation. Guidance on choice of control group and related design and conduct issues in clinical trials (ICH E 10). Food and Drug Administration, DSSH, July 2000.
20. Food and Drug Administration, HHS. Draft Guidance for Industry Non-Inferiority Clinical Trials. 2010.
21. Doshi P, Spence O, Kuzucan A, Powers JH III. Communication of nonefficacy benefits of new drugs approved on the basis of noninferiority trials alone: cohort study of FDA and sponsor communication, 2011-2017. *JAMA Intern Med* 2019; 179: 719-21.
22. Menikoff J. What should patients be told about noninferiority studies? *JAMA Int Med* 2017; 177: 1459-60.
23. Piaggio G, Elbourne DR, Altman DG, et al. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006; 295: 1152-60.
24. Committee for Medicinal Products for Human Use, Efficacy Working Party, Committee for Release for Consultation. Committee for Medicinal Products for Human Use (CHMP) guideline on the choice of the non-inferiority margin. *Stat Med* 2006; 25: 1628-38.