Pogotowie®
Statystyczne

# Wskaźniki siły efektu

## Teoria, wzory, wyjaśnienia
## symboli, interpretacja

**Autorzy:**
Andrzej Jankowski | ajankowski@pogotowiestatystyczne.pl
Marta Formela | mformela@pogotowiestatystyczne.pl

**Korekta merytoryczna:**
Paweł Iwankowski
Paweł Krasa

**Skład i opracowanie graficzne:**
Pogotowie Statystyczne

**Dane firmowe:**
Pogotowie Statystyczne
Paweł Iwankowski
ul. prof. Stefana  Hausbrandta 34/88
80-126 Gdańsk
NIP: 7412032970
REGON: 280490493


**Kontakt:**
info@pogotowiestatystyczne.pl
tel. 501 599 278



Gdańsk 2026

Publikacja dostępna bezpłatnie w formacie PDF na stronie:
www.pogotowiestatystyczne.pl

**Cytowanie publikacji:**
Pogotowie Statystyczne (2026). *Wskaźniki siły efektu - teoria, wyjaśnienia symboli, interpretacja.* https://pogotowiestatystyczne.pl/

# WPROWADZENIE

Siła efektu to jedna z najbardziej istotnych statystyk, które zgodnie z wytycznymi zawartymi w stylu APA należy uwzględnić w raportowaniu wyników analizy statystycznej.

## DLATEGO TEŻ W TYM ARTYKULE

przedstawimy najważniejsza założenia teoretyczne, wskazówki dotyczące interpretacji oraz wzory na poszczególne wskaźniki siły efektu dla różnych testów statystycznych.

*„Siła efektu (ang. effect size) to miara statystyczna, która służy do oceny wielkości uzyskanego efektu”*

**Czym jest siła efektu?**

W paradygmacie częstościowym (NHST) wnioskowanie statystyczne opieramy przede wszystkim na interpretacji wartości *p* (*p-value*), którą zestawiamy z przyjętym poziomem istotności statystycznej (α). Z racji tego, że zazwyczaj α = 0,05, przyjęło się, że wynik $p < 0,05$ określamy jako „istotny statystycznie”, a $p > 0,05$ jako „nieistotny statystycznie”. Na tej podstawie podejmujemy decyzję o odrzuceniu bądź nieodrzuceniu hipotezy zerowej czy też przyjęciu lub odrzuceniu hipotezy badawczej.

Warto jednak pamiętać, że samo określenie tego, czy wynik testu jest „istotny statystycznie”, nie jest wystarczające do dokonania pełnej interpretacji. Konieczne jest uzupełnienie raportowania o inne statystyki. Jedną z nich jest wartość wskaźnika siły efektu (ang. *effect size*) – miara statystyczna, która służy do oceny skali (wielkości) uzyskanego efektu, np. różnicy między grupami lub siły związku między zmiennymi. Warto podkreślić, że raportowanie *effect size* powinniśmy potraktować nie jako możliwość, z której warto skorzystać, a jako konieczność – jest to bowiem niezbędny element raportu wymagany w standardzie APA.

**Jaki jest związek między siłą efektu a wartością *p*?**

Raportowanie siły efektu obok wartości *p* jest o tyle istotne, że interpretacja *p-value* pozwala „jedynie” na weryfikację postawionej hipotezy. Pojęcie „istotności statystycznej”, związane z *p*-value, nie jest bowiem tożsame z „istotnością praktyczną”, czyli tym, na ile dany wynik jest ważny z teoretycznego punktu widzenia. Możliwe jest zatem uzyskanie wyniku „istotnego statystyczne”, który ma niewielkie

**DEFINICJA**

*Siła efektu (ang. effect size) to wskaźnik statystyczny służący do oceny skali (wielkości) uzyskanego efektu, np. różnicy między grupami czy korelacji. Siłę efektu zwykle interpretuje się w trzech przedziałach: słaba, umiarkowana, silna. Jej analiza pozwala ocenić praktyczne znaczenie wyników - w odróżnieniu od wartości p, która określa jedynie statystyczną „istotność" wyniku.*

znaczenie praktyczne lub efektu o ważnym znaczeniu z punktu widzenia teoretycznego, ale o wartości *p* wyższej od ustalonego umownie progu 0,05. Jak zatem ocenić czy znaczenie praktyczne danego efektu jest niewielkie, umiarkowane lub duże? Między innymi na podstawie oceny wartości *effect size*.

Okazuje się zatem, że *p*-value i effect size możemy traktować jako dwa odrębne wskaźniki, które pozwalają na ocenę dwóch różnych aspektów uzyskanego rezultatu. Jak wynika z kontekstu, wartości siły efektu oraz *p*-value nie są ze sobą bezpośrednio powiązane. Wyjaśnijmy dlaczego tak jest. To istotne, bo rozumienie tych zależności jest kluczem do poprawnej interpretacji wyników i wyciągania właściwych wniosków.

## Od czego zależy wartość siły efektu i wartość *p*?

Wyobraźmy sobie, że przeprowadzamy dwa osobne, choć podobne badania. W obu, pod względem nasilenia takiej samej zmiennej zależnej, porównujemy między sobą dwie takie same grupy, stosując test *t* Studenta dla prób niezależnych. Przyjmijmy, że w obu badaniach uzyskaliśmy analogiczne wyniki, w sensie - te same średnie, a co za tym idzie tę samą różnicę średnich między grupami, a także te same wartości odchylenia standardowego.

Na pierwszy rzut oka moglibyśmy uznać, że skoro obliczone statystyki są takie same, to można przyjąć, że uzyskaliśmy takie same wyniki. I w pewnym sensie jest to prawda, skoro uzyskane różnice rzeczywiście są sobie równe. Co więcej, wniosek ten potwierdzi również wartość obliczonej statystyki *d* Cohena, która jest jedną z miar siły efektu w teście *t* Studenta dla prób niezależnych – w obu przypadkach będzie ona taka sama.

Załóżmy jednak, że oba te badania różni jeden aspekt – wielkość próby,

gdzie jedno badanie przeprowadzono na próbie $N = 60$, a drugie na próbie $N = 120$ osób. W żaden sposób nie wpłynie to na uzyskaną wartość $d$ Cohena, ponieważ różnice między grupami pozostają nadal takie same. Okazuje się jednak, że wartości $p$ uzyskane w obu tych badaniach będą różne! W badaniu wykonanym na większej próbie wartość $p$ będzie niższa niż w badaniu wykonanym na mniejszej próbie. Może okazać się nawet, że jeden z tych wyników okaże się istotny statystycznie, a drugi nieistotny statystycznie!

Skąd wynikają te rozbieżności? Okazuje się, że wartość $p$ (przy założeniu stałej wielkości efektu) spada wraz ze wzrostem liczebności próby. Wynika to ze sposobu działania testów statystycznych, których „precyzja" pomiaru (mierzona wielkością błędu standardowego) wzrasta wraz z liczebnością próby (wtedy wspomniany błąd standardowy maleje). Efekt ten możemy odnieść też do pojęcia mocy testu statystycznego, czyli jego „czułości" do wykrycia danego efektu. Ta, analogicznie, wzrasta wraz z liczebnością próby, co w praktyce przekłada się na uzyskiwanie wyników istotnych statystycznie przy znacznie mniejszych realnych efektach (mierzonych wartością danego *effect size*).

Z kolei wartość siły efektu nie zależy w taki sposób od liczebności próby jak wartość $p$. Co prawda pewna zależność istnieje (pomijamy tu szczegóły), ale jego skala jest dużo mniejsza w porównaniu do relacji między liczebnością próby a $p$-value. Siła efektu w większym stopniu odzwierciedla zatem rozmiar realnego efektu, który mierzymy.

**Dlaczego warto raportować i interpretować wartość siły efektu?**

Odmienny charakter zależności pomiędzy liczebnością próby a $p$-value i *effect size* ma wyraźne przełożenie na uzyskiwane wyniki wykonywanych testów statystycznych.

**DEFINICJA**

*Moc testu to „czułość" danego testu statystycznego do wykrycia danego efektu (np. różnicy czy korelacji). Moc testu wzrasta wraz z liczebnością próby, co przykłada się na uzyskiwanie wyników istotnych statystycznie (określanych przy pomocy wartości p) przy co raz mniejszych realnych efektach (mierzonych poprzez siłę efektu).*

Z jednej strony, wartość *p* jest kluczowa przy weryfikacji danej hipotezy statystycznej. Z drugiej, wniosek z takiej analizy warto uzupełnić o interpretację wartości wskaźnika siły efektu. Ma to duże znaczenie praktyczne, o którym warto pamiętać dokonując interpretacji wyników przeprowadzonej analizy. Poniżej przeanalizujemy dwa przykłady, które obrazują na jakiej zasadzie taka interpretacja może przebiegać.

W badaniach przeprowadzanych na dużych próbach (rzędu setek obserwacji) często uzyskuje się wiele wyników istotnych statystycznie, z których wiele (a czasami większość) charakteryzuje niewielki efekt, wyrażony wartością danego *effect size*. W takiej sytuacji, interpretując samą wartość *p*, pozbawiamy się szansy na wyciągnięcie pełnowartościowych wniosków, ponieważ na równi traktujemy wyniki odzwierciedlające różne skale efektu. W rzeczywistości, bardziej istotne bowiem jest to, że korelacja liniowa między jedną parą zmiennych jest silna i wyniosła $r = 0{,}67$ a między drugą parą zmiennych jest słaba i wyniosła $r = 0{,}17$, niż fakt, że oba związki były „istotne statystycznie". Interpretacja takich wyników powinna obejmować więc zarówno wartości *p* jak i *effect size*.

**PAMIĘTAJ**

***Interpretacja siły efektu to jeden z najważniejszych elementów wnioskowania statystycznego w oparciu o klasyczne testy statystyczne. Raportowanie siy efektu jest również wymagane w standardzie APA.***

Analogicznie przedstawia się sprawa w badaniach wykonywanych na niewielkich próbach (rzędu np. 20 czy 30 obserwacji). W takich sytuacjach nieistotne statystycznie mogą okazać się wyniki, dla których wartość *effect size* jest względnie wysoka. Jeśli jednak rozumiemy relacje między wielkością próby, siłą efektu i wartością *p*, taki fakt możemy odnotować i uwzględnić w interpretacji. Przykładowo, jeśli dla testu *t* dla prób niezalenych uzyskaliśmy wynik nieistotny statystycznie, choć wartość *d* Cohena wyniosła 0,67 (co interpretujemy jako umiarkowaną różnicę), możemy śmiało taki rezultat opisać, szczególnie jeśli efekt ten ma duże znaczenie dla postawionego problemu badawczego. Takie sytuacje nie rzadkie, pojawiają się np. w badaniach pilotażowych lub w sytuacjach, gdy uzyskanie większej próby jest trudne np. ze względu na wysoki koszt badań lub specyfikę populacji (sytuacja spotykana w badanich klinicznych).

## Podsumowanie

Podsumowując, *p*-value jest miarą prawdopodobieństwa statystycznego i nie informuje o „praktycznym" znaczeniu wyniku. Jej wartość zależy od kilku wypadkowych związanych z „mechaniką" danego testu statystycznego, przy czym największe praktyczne znaczenie ma wielkość próby. W konsekwencji to, na ile adekwatna jest interpretacja tej wartości w odniesieniu do danego zjawiska zależy w dużej mierze od tego, na ile przeprowadzone badanie zostało poprawnie zaprojektowanie, np. od tego czy wielkość zebranej próby została oparta na wcześniejszych obliczeniach uwzględniających moc wykorzystywanego testu.

**REKOMENDACJA**

*Poza wskaźnikami siły efektu można też obliczyć przedział ufności dla tych wartości. Warto pamiętać też o raportowaniu dokładnej wartości p, odpowiedniej statystyki testowej, stopni swobody, oraz właściwych statystyk opisowych.*

Siła efektu z kolei jest w mniejszym stopniu zależna od tego typu czynników i w bardziej bezpośredni sposób odzwierciedla skalę uzyskanego efektu. Dlatego też raportowanie wartości *effect size* dla poszczególnych testów jest tak istotne. Niemniej, ważne jest jednak rozumienie zależności między liczebnością próby, mocą testu, wartością *p* i wartością siły efektu oraz wykorzystanie tej wiedzy do umiejętnej interpretacji uzyskanego wyniku, która uwzględnia wszystkie te składowe równocześnie.

Jako uzupełnienie warto dodać, że raportować możemy nie tylko pojedynczą wartość danego wskaźnika siły efektu, ale też odpowiadający mu przedział ufności. W ten sposób estymację punktową uzupełniamy o estymację przedziałową, dzięki czemu uzyskujemy kolejne dane, które możemy wykorzystać w interpretacji. Przykładowo, jeśli w dwóch badaniach otrzymujemy jednakową wartość *d* Cohena = 0,67 to uzyskanie informacji, że przedział ufności 95% w pierwszym przypadku wynosi [0,17; 1,14] a w drugim [0,62; 0,75] pozwala nam na wyciągnięcie dodatkowych wniosków – w skrócie, na przykład większy zakres wskazuje na mniejszą precyzję oszacowania.

## Jakie są popularne wskaźniki siły efektu i jak je obliczyć?

W tabeli zawartej na kolejnych stronach prezentujemy popularne wskaźniki siły efektu dla różnych testów statystycznych, wraz z przedziałami wartości pozwalającymi dokonać interpretacji uzyskanego wyniku. Przedstawiamy również wzory pozwalające je obliczyć - większość z nich pochodzi z podręcznika Ellisa (2010). Warto pamiętać, że zazwyczaj nie dokonujemy tych obliczeń „ręcznie", ponieważ większość z nich jest obliczana w popularnych pakietach statystycznych.

Warto podkreślić, że poszczególne wartości progowe służące do interpretacji różnych wskaźników siły efektu nie są obiektywne, lecz wynikają z konsensusu wypracowanego przez teoretyków i statystyków. Zdarza się, że niektóre wartości posiadają więcej niż jeden próg klasyfikacji, co wynika z różnych propozycji autorów.

## Literatura

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results.* Cambridge university press.

Fritz, C. O., Morris, P. E., Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General, 141*(1), 2-18.

Maher, J. M., Markey, J. C., Ebert-May, D. (2013). The other half of the story: effect size analysis in quantitative research. *CBE—Life Sciences Education, 12*(3), 345-351.

www.spss-tutorials.com/effect-size/ (dostęp: 11.02.2026r.)

| Test | Wskaźnik | Wzór | Objaśnienia symboli | Interpretacja |
|---|---|---|---|---|
| Test *t* dla prób niezależnych | *d* Cohena | $d = \frac{M_A - M_B}{SD_{pooled}}$ | $M_A$ *i* $M_B$ - średnie dwóch grup<br>$SD_{pooled}$ - łączne odchylenie standardowe | 0,2 - efekt słaby<br>0,5 - efekt umiarkowany<br>0,8 - efekt silny |
| | *g* Hedges'a | $g = \frac{M_A - M_B}{SD_{pooled*}}$ | $M_A$ *i* $M_B$ - średnie dwóch grup<br>$SD_{pooled*}$ - łączne odchylenie standardowe (ważone) | |
| | Δ Glassa | $\Delta = \frac{M_A - M_B}{SD_{control}}$ | $M_A$ *i* $M_B$ - średnie dwóch grup<br>$SD_{control}$ - odchylenie standardowe dla grupy kontrolnej | |
| Test *t* dla prób zależnych | *d* Cohena | $d = \frac{t}{\sqrt{N}}$ | *t* - statystyka testu *t*<br>*N* - liczebność próby | 0,2 - efekt słaby<br>0,5 - efekt umiarkowany<br>0,8 - efekt silny |
| Test *U* Manna Whitenya | $r_g$ | $r_g = \frac{Z}{\sqrt{N}}$ | *Z* - standaryzowana statystyka testu *U* Manna Whitneya<br>*N* - liczebność próby | 0,1 - efekt słaby<br>0,3 - efekt umiarkowany<br>0,5 - efekt silny |
| | $\eta^2$ | $\eta^2 = \frac{Z^2}{N}$ | | 0,01 - efekt słaby<br>0,06 - efekt umiarkowany<br>0,14 - efekt silny |
| Test Wilcoxona | $r_g$ | $r_g = \frac{Z}{\sqrt{N}}$ | *Z* - statystyka testu Wilcoxona<br>*N* - liczebność próby pomniejszona o liczbę rang wiązanych | 0,1 - efekt słaby<br>0,3 - efekt umiarkowany<br>0,5 - efekt silny |
| Korelacje | *r* | $r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2 \Sigma(y - \bar{y})^2}$ | *x* - wartość zmiennej x<br>$\bar{x}$ - średnia zmiennej x<br>*y* - wartość zmiennej y<br>$\bar{y}$ - średnia zmiennej y | 0,1 - efekt słaby<br>0,3 - efekt umiarkowany<br>0,5 - efekt silny |
| | $r_s$ lub ρ | $r_s = 1 - \frac{6\sum d^2}{N(N^2 - 1)}$ | *d* - różnica w rangach<br>*N* - liczebność próby | |

| | | | | |
|---|---|---|---|---|
| | τ | $\tau = \frac{n_c - n_d}{\frac{N(N-1)}{2}}$ | $n_c$ - liczba par zgodnych<br>$n_d$ - liczba par niezgodnych<br>$N$ - liczebność próby | |
| ANOVA, ANCOVA | $\eta_p^2$ | $\eta_p^2 = \frac{SS_{effect}}{SS_{effect} + SS_{error}}$ | $SS_{effect}$ - suma kwadratów dla efektu<br>$SS_{error}$- suma kwadratów dla błędu | 0,01 - efekt słaby<br>0,06 - efekt umiarkowany<br>0,14 - efekt silny |
| | $\eta^2$ | $\eta^2 = \frac{SS_{effect}}{SS_{total}}$ | $SS_{effect}$ - suma kwadratów dla efektu<br>$SS_{total}$- całkowita suma kwadratów | |
| | $\omega_p^2$ | $\omega_p^2 = \frac{df_{effect}(MS_{effect} - MS_{error})}{df_{effect}(MS_{effect}) + (N - df_{effect})(MS_{error})}$ | $df_{effect}$ - liczba stopni swobody dla efektu<br>$MS_{effect}$ - średni kwadrat dla efektu<br>$MS_{error}$- średni kwadrat dla błędu<br>$N$ - liczebność próby | |
| | $\omega^2$ | $\omega^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_{error}^2}$<br><br>$\omega^2 = \frac{df_{effect}(MS_{effect} - MS_{error})}{SS_{total} + MS_{error}}$ | $\sigma_A^2$ - wariancja dla czynnika A<br>$\sigma_{error}^2$ - wariancja błędu dla czynnika A<br>$df_{effect}$ - liczba stopni swobody dla efektu<br>$MS_{effect}$ - średni kwadrat dla efektu<br>$MS_{error}$- średni kwadrat dla błędu<br>$SS_{total}$- całkowita suma kwadratów | |
| | $f$ Cohena | $f = \frac{\sigma_A}{\sigma}$ | $\sigma_A$ - odchylenie standardowe dla średniej<br>$\sigma$ - odchylenie standardowe dla populacji | 0,10 - efekt słaby<br>0,25 - efekt umiarkowany<br>0,40 - efekt silny |

| | | | | |
|---|---|---|---|---|
| Test Kruskala Wallisa | $\eta^2$ | $\eta^2 = \frac{H - k + 1}{N - k}$ | $H$ - statystyka testu H Kruskala Wallisa<br>$k$ - liczba grup<br>$N$ - liczebność próby | 0,01 - efekt słaby<br>0,06 - efekt umiarkowany<br>0,14 - efekt silny |
| $\chi^2$ | $\phi$ | $\phi = \sqrt{\frac{\chi^2}{N}}$ | $\chi^2$ - statystka chi-kwadrat<br>$N$ - liczebność próby | 0,1 - efekt słaby<br>0,3 - efekt umiarkowany<br>0,5 - efekt silny |
| | $V$ Cramera | $V = \sqrt{\frac{\chi^2}{N(k-1)}}$ | $\chi^2$ - statystka chi-kwadrat<br>$N$ - liczebność próby<br>$k$ - liczba wierszy lub kolumn w tabeli | |
| Analiza regresji liniowej | $R^2$ | $R^2 = \frac{SS_{regression}}{SS_{total}}$ | $SS_{regression}$- suma kwadratów dla regresji<br>$SS_{total}$- całkowita suma kwadratów | 0,02 - efekt słaby<br>0,13 - efekt umiarkowany<br>0,26 - efekt silny |
| | Skorygowane $R^2$ | $R^2_{adj.} = 1 - (1 - R^2)\left(\frac{N-1}{N-k-1}\right)$ | $R^2$ - współczynnik determinacji | |
| | $f^2$ | $f^2 = \frac{R^2}{1 - R^2}$ | | 0,02 - efekt słaby<br>0,15 - efekt umiarkowany<br>0,35 - efekt silny |
| Analiza regresji logistycznej | OR (iloraz szans) | $\mathrm{OR} = \frac{\frac{p}{(1-p)}}{\frac{q}{(1-q)}}$ | $p$ - prawdopodobieństwo wystąpienia w grupie A<br>$q$ - prawdopodobieństwo wystąpienia w grupie B | 1,5 - efekt słaby<br>2,5 - efekt umiarkowany<br>4,0 - efekt silny |

# Od 2007 roku...

...świadczymy usługi analizy statystycznej w badaniach do prac dyplomowych i artykułów naukowych.

Nasza pomoc trafiła już do ponad 20 000 studentów i pracowników naukowych wszystkich uczelni w Polsce!

Zapoznaj się z naszą ofertą:
www.pogotowiestatystyczne.pl

Pogotowie® Statystyczne