

## INTERPRETACIJA INTERVALA POUZDANOSTI

Nermin SALKIĆ

Klinika za interne bolesti  
Univerzitetsko - kliničkog  
centra Tuzla, Tuzla, Bosna i  
Hercegovina

Dok u testiranju hipoteze, rezultati studije kroz prikaz p-vrijednosti daju podatke da li je nulta hipoteza odbačena ili ne, čitatelj iz ovakvog prikaza ne može procijeniti kolika je zbilja veličina razlike u mjerenom efektu. Međutim, interval pouzdanosti omogućava da se procijeni veličina razlike u posmatranom efektu kao i njena snaga i pouzdanost. Interval pouzdanosti pruža raspon u kome se prava vrijednost nalazi sa 95% vjerovatnoćom, a također daje i podatke o tome da li je uzorak dovoljno velik da bi rezultat bio pouzdan.

**Ključne riječi:** Interval pouzdanosti ▪ Uzorak, Hipoteza ▪ P-vrijednost

Adresa za dopisivanje:  
Nermin Salkić  
Miroslava Krleže 15/26  
75000 Tuzla,  
Bosna i Hercegovina  
E-mail: snermin@gmail.com

### Uvod

Statističari imaju nezgodnu naviku da se uvijek ograđuju. U svojim izvještajima o rezultatima medicinskih istraživanja konstantno upozoravaju čitatelja o različitim pretpostavkama vezanim za upotrijebljeni test, a da se pri tome rijetko kada usude da neki rezultat okarakterišu nekim jačim atributom od npr. »moguće«. Dapače, poznata je izreka »Statistika je način da nikada ne morate da izjavite da ste sigurni«.

Razlog za ovakav stav leži u tome što statističari vrše kalkulacije na osnovu nesavršenih podataka. Vrlo često od njih se traži da se izjasne o populaciji (veliki broj) koristeći informacije iz određenog uzorka (malog broja pažljivo odabranih pripadnika populacije). Međutim, bez obzira na to koliko je uzorak pažljivo odabran, podaci koji se dobiju na osnovu njegove statističke analize, uvijek imaju određeni nivo nepouzdanosti.

**Primljeno:** 6. 11. 2008.  
**Prihvaćeno:** 22. 11. 2008.

Pedijatrija danas 2009;5(1):95-99

## Interval pouzdanosti

Sve doskoro, procjena ove nepouzdanosti rezultata se rutinski bazirala na testiranju hipoteze, koja proizvodi opštepoznatu p-vrijednost. Naime, zamislimo da smo postavili hipotezu da postoji razlika između tretmana A i tretmana B. U procesu testiranja hipoteze, mi formuliramo tzv. nultu hipotezu, odnosno da između tretmana A i tretmana B nema nikakve razlike u efikasnosti. P-vrijednost nam označava vjerovatnoću postojanja nulte hipoteze, dakle ako je p vrijednost 0,05, postoji samo 5 postotna vjerovatnoća (ili šansa 1:19) da je nulta hipoteza tačna. Arbitrarno je odlučeno da je ova vrijednost od 0,05 granica između »statistički signifikantnog« ( $p < 0,05$ ) i »statistički nesignifikantnog« ( $p > 0,05$ ). Međutim, p-vrijednosti nam pokazuju samo da li je razlika značajna ili ne. Dakle, ne pružaju nam numeričke podatke o tome kolika je razlika u efikasnosti. Tu na scenu stupa Interval pouzdanosti (prema engleskom: Confidence Interval – CI) koji je mnogo korisnija mjera. Iako je naizgled bitno različit od p-vrijednosti, sa istom ima mnogo sličnosti, a i sama teorija u pozadini, kao i način izračunavanja su vrlo slični (1).

CI je raspon vrijednosti koji pokušava da kvantificira nepouzdanost rezultata. Njegova formalna definicija je da je CI raspon vrijednosti neke varijable od interesa, ali napravljen tako da ima tačno određenu vjerovatnoću da uključuje i pravu vrijednost. Ova vjerovatnoća se naziva nivo pouzdanosti, a krajnje tačke CI nazivamo granicama pouzdanosti (2).

Jednostavnije, CI možemo smatrati rasponom vjerovatnih vrijednosti od kojih je jedna ona prava. Uzak interval pouzdanosti podrazumijeva veću preciznost rezultata, dakle, mali broj vjerovatnih vrijednosti. Nasuprot tome, široki interval pouzdanosti podrazumijeva i manju preciznost rezultata; vjerovatne vrijednosti pripadaju širokom i neinformativnom rasponu. Opšteprihvaćen nivo pouzdanosti za CI je 95%, a što praktično znači da za ovakav CI možemo reći da sa vjerovatnoćom od 95% sadrži i pravu vrijednost.

## Interpretacija intervala pouzdanosti

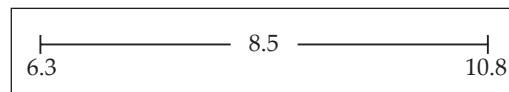
Kako ustvari interpretirati CI kada naiđemo na njega u nekome članku? Ilustrirajmo ovo sjajnim primjerom datim od strane Raoa (3). Zamislimo imaginarno istraživanje u kome istražujemo da li postoji razlika u krvnom pritisku između grupe ispitanika kojoj je dat neki novi antihipertenziv i grupe kojoj je dat placebo. Rezultati našeg istraživanja pokazuju da razlika postoji i ona se može prikazati koristeći klasični pristup testiranja hipoteze ovako:

$$\text{Prosječna razlika} = 8,5 \text{ mmHg}; p = 0,03$$

Međutim, malo je vjerovatno da je prava razlika u pritisku 8,5 mmHg – da ponovimo studiju, vrijednost bi bila barem malo drugačija. Ova mogućnost varijacije nije nam kvantitativno pokazana kroz p-vrijednost, ista samo govori da je razlika signifikantna, ali nam ne daje nikakve podatke o preciznosti i veličini naše procjene. Međutim, ako isti rezultat navedemo uz CI, on bi izgledao ovako:

$$\text{Prosječna razlika} = 8,5 \text{ mmHg}; \%95 \text{ CI} = 6,3-10,8 \text{ ili grafički (Slika 1)}.$$

Pojednostavljeno interpretirano, prava vrijednost prosječne razlike u krvnom pritisku najvjerovatnije leži u intervalu od 6,3 do 10,8. Postoji samo 5% vjerovatnoće da se nalazi izvan ovoga intervala.



**Slika 1** Grafički predstavljen rezultat sa intervalom pouzdanosti (objašnjenje u tekstu). Reproduciran sa dozvolom<sup>3</sup> (Copyright 2003 Dowden Health Media, Inc.)

**Figure 1** Graphical representation of estimate with confidence interval (detailed explanation provided in text). Translated with permission<sup>3</sup> (Copyright 2003 Dowden Health Media, Inc.)

Kada čitamo i zapazimo CI u nekom medicinskom članku, treba obratiti pažnju na

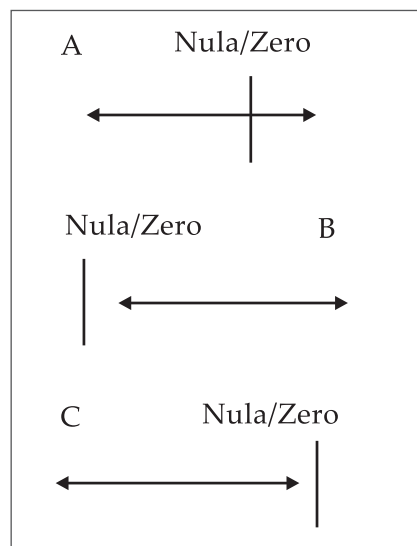
to da li CI uključuje vrijednost koja podrazumijeva odsustvo efekta odnosno odsustvo bilo kakve razlike (npr. da su dva tretmana jednake efikasnosti ili da je ispitivani tretman čak gori). Kada posmatramo numeričku razliku između dva tretmana, kao u navedenom primjeru sa antihipertenzivima, treba gledati da li CI uključuje i nulu. Naprimjer, zamislimo da je razlika u pritiscima 4,5 mmHg sa 95% CI od -3,4 mmHg do 5,8 mmHg. Ovakav rezultat upućuje da postoji realna mogućnost da tretmani ne samo da nisu različiti (zato što uključuje nulu), nego da je ispitivani tretman čak i gori (zato što uključuje i negativne vrijednosti).

Ukoliko istraživanje sadrži odnose šansi (prema engleskom: Odds Ratio – OR), onda treba gledati da li CI uključuje jedinicu, budući da  $OR < 1$  znači smanjenje šansi, dok  $OR > 1$  govori za povećavanje šansi. Dakle ako u rezultatima naiđemo na OR koji iznosi 0,82 i koji ima 95% CI od 0,56 do 1,24, vidimo da isti zahvata jedinicu i samim time podrazumijeva odsustvo statistički signifikantnog efekta.

Na slici 2A je dat CI koji uključuje nulu. Ovakav interval implicira odsustvo statistički signifikantne razlike. Na slici 2B je dat CI koji isključuje nulu. Ako pretpostavimo da je više u ovom slučaju bolje, ovakav interval upućuje na statistički signifikantno poboljšanje. Na slici 2C je dat interval koji takođe ne uključuje nulu, ali koji implicira statistički signifikantno pogoršanje (naravno, ako je i dalje na snazi pretpostavka da je više bolje). Za odnose šansi grafikoni su analogni, samo što uspravna linija nije nula, nego jedinica.

Preciznost i širina CI ovisi i od veličine uzorka. Upravo ova njegova osobina nam omogućuje da u medicinskom članku procijenimo da li je uzorak dovoljno veliki da bi dao odgovor na istraživačko pitanje. Ovo možemo ilustrovati primjerom bacanja novčića (4). Zamislimo da bacamo novčić – naša procjena je da ćemo npr. dobiti pismo u 50%

bacanja. Međutim, interesuje nas koji je to raspon vrijednosti, odnosno CI u kojem se nalazi prava vjerovatnoća da dobijemo pismo. Vidimo da je za mali broj bacanja ovaj raspon dosta velik (Tabela 1). Ali kako povećavamo broj bacanja, smanjuje se i raspon CI. Upravo ovu pojavu možemo iskoristiti da bismo procijenili pouzdanost podataka. Npr. ako pogledamo Tabelu 1, ukoliko smo imali 100 bacanja novčića, naša procjena leži unutar 10% od prave vrijednosti, a ako bacimo novčić 1000 puta, preciznost naše procjene se penje na 3%. Jednostavnije, sa povećanjem uzorka smanjuje se i raspon intervala pouzdanosti.



**Slika 2** Intervali pouzdanosti u odnosu na nultu vrijednost i njihova interpretacija (Objašnjenje dato u tekstu)

**Figure 2** Confidence intervals: relationship to zero value and corresponding interpretation (Detailed explanation provided in text)

Dakle, da se ponovno vratimo na naš eksperiment sa antihipertenzivima. Ukoliko bismo povećali broj ispitanika u našem istraživanju, naš interval pouzdanosti ne bi bio 6,3-10,8 mmHg, nego bi bio 7,8-9,1 mmHg, a što upućuje na to da bi naša procjena efekta bila preciznija i pouzdanija i da je ne bi trebalo olako shvatiti. Intervali pouz-

**Tabela 1** Interval pouzdanosti oko proporcije od 0,5 u eksperimentu bacanja novčića. Objašnjenje u tekstu.**Table 1** Confidence interval of proportion estimate of 0,5 in coin toss experiment. Explanation provided in text.

Broj bacanja novčića/ Number of coin tosses	Stvarni rezultat/ True result	95% interval pouzdanosti/ 95% confidence interval
2	1 Pismo/Tail 1 Glava/Head	0,01-0,99
10	5 Pismo/Tail 5 Glava/Head	0,19-0,81
50	25 Pismo/Tail 25 Glava/Head	0,36-0,65
100	50 Pismo/Tail 50 Glava/Head	0,40-0,60
1000	500 Pismo/Tail 500 Glava/Head	0,47-0,53

danosti iz velikih studija imaju tendencu da budu uži, te nam na taj način pokazuju preciznost sa kojom je studija u stanju da procijeni veličinu pravog efekta. Zbog toga se rezultati iz ovakvih istraživanja više i cijene.

Na kraju treba napomenuti da CI polako istiskuje iz upotrebe p-vrijednost i u većini kvalitetnijih medicinskih časopisa postaje sine qua non u navođenju dobijenih rezultata. Razlog za to je naveden u redovima iznad - interval pouzdanosti jednostavno, pruža

više informacija od p-vrijednosti, a urednici žele da omoguće svojim čitateljima što bolju procjenu rezultata istraživanja. S obzirom na ovakav trend, otuda je i namjera da se u ovome kratkom prikazu navedena problematika malo više osvijetli, kako bi se svima onima koji se bave medicinskim istraživanjima prvo omogućilo da kritičnije čitaju tuđe radove, a zatim i da kvalitetnije pišu svoje članke, za koje imaju aspiracije za publikaciju u nekome od kvalitetnijih časopisa.

## Literatura

1. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br Med J (Clin Res Ed)*. 1986;292:746-50.
2. Huw D. What are confidence intervals? 2001 [pristupljeno: 05.11.2008.]; Dostupno sa: <http://www.evidence-based-medicine.co.uk/ebmfiles/WhatareConfInter.pdf>
3. Rao G. Interpretation of confidence intervals. *J Fam Pract*. 2003;52:970.
4. Guyatt G, Jaeschke R, Heddle N, Cook D, Shannon H, Walter S. Basic statistics for clinicians: 2. Interpreting study results: confidence intervals. *CMAJ*. 1995;152:169-73.

## Summary

### INTERPRETATION OF CONFIDENCE INTERVAL

*Nermin SALKIĆ*

Department of Gastroenterology, Internal Medicine Hospital,  
University Clinical Center Tuzla Tuzla, Bosnia and Herzegovina

In hypothesis testing, study results expressed through p-values provide information whether to accept or reject a null hypothesis. However, expressed this way they do not provide sufficient information for a reader to assess the true difference in a measured effect. On the other hand, confidence interval is informative enough to assess both the difference in a measured effect and its strength and reliability. Confidence interval provides a reader with a span of values that contains the true value with a 95-percent probability and also tells the reader whether the sample is big enough for reliable results.

**Key words:** Confidence interval ▪ Sample ▪ Hypothesis ▪ P-value

**Received:** November 6, 2008

**Accepted:** November 22, 2008