

Interval pouzdanosti

Confidence interval

Ana-Maria Šimundić

Klinički zavod za kemiju, Klinička bolnica „Sestre milosrdnice“, Zagreb

University Department of Chemistry, Sestre milosrdnice University Hospital, Zagreb, Croatia

Sažetak

Interval pouzdanosti za bilo koju statističku mjeru predstavlja raspon mogućih vrijednosti unutar kojega se s izvjesnom vjerojatnosti nalazi ta statistička mjera populacije. Kao takav, interval pouzdanosti je objektivna procjena (ne)preciznosti i veličine uzorka nekog istraživanja. Stoga, na interval pouzdanosti posredno možemo gledati i kao na mjerilo kvalitete uzorka i istraživanja. Interval pouzdanosti je omeđen granicama. Ovisno o razini pouzdanosti koju biramo, mijenja se i raspon, tj. granice intervala. Najčešće korišteni intervali pouzdanosti u biomedicinskoj literaturi su 90%-tni, 95%-tni, 99%-tni i nešto rjeđe 99,9%-tni interval pouzdanosti. Što su granice intervala uže, preciznost procjene je veća. Tradicionalno se u literaturi najčešće koristi 95%-tni interval pouzdanosti, koji je u svezi s opće prihvaćenom razinom statističke značajnosti $P < 0,05$. Za uzorak iste veličine vrijedi pravilo: što je manja razina pouzdanosti – veća je preciznost procjene. Samo istraživanja provedena na velikom uzorku će dati vrlo uski interval pouzdanosti koji ukazuje na veliku preciznost procjene, s visokom razinom pouzdanosti. Interval pouzdanosti moguće je pridružiti gotovo svakom statističkom pokazatelju. Iako postoje i neki drugi načini, interval pouzdanosti najčešće računamo pomoću standardne greške. Interval pouzdanosti je komplementaran statistički pokazatelj s P vrijednosti. Oni govore o istoj stvari na dva različita načina i međusobno se nadopunjuju. P vrijednost nam govori o vjerojatnosti s kojom je moguće da se uočeni fenomen (razlika) dogodio slučajno, dok interval pouzdanosti nudi granice unutar kojih je moguće očekivati vrijednost tog fenomena. Posljednjih dvadesetak godina sve je više časopisa u kojima je obveza autora prikazati intervale pouzdanosti za svoje ključne rezultate. Prikaz intervala pouzdanosti pruža dodatnu informaciju o našem uzorku i rezultatima, i nadasve je korisna i nezamjenjiva nadopuna klasičnom testiranju hipoteze i opće prihvaćenoj P vrijednosti. Prikaz ključnih rezultata uz pripadajuće intervale pouzdanosti trebao bi postati standard svih znanstvenih časopisa, jer zainteresiranom čitatelju omogućuje bolje razumijevanje prikazanih podataka.

Ključne riječi: interval pouzdanosti, uzorak, standardna greška

Abstract

Confidence interval presents a range of possible values within which, with some certainty, we can find the statistical measure of the population. As such, it is an objective estimate of (in)precision and sample size of certain research. Therefore, we can consider confidence interval also as a measure of the sample and research quality. Confidence interval is defined by its margins of error. Depending on the confidence level that we choose, the interval margins of error and respective range also change. The most used confidence intervals in the biomedical literature are the 90%, 95%, 99% and not so often 99.9% one. The narrower the margins of an interval are, the higher is the estimate accuracy. The 95% confidence interval is traditionally the most used interval in the literature and this relates to the generally accepted level of statistical significance $P < 0.05$. There is a rule for same sized samples: the smaller the confidence level is, the higher is the estimate accuracy. Only the studies with a large sample will give a very small confidence interval, which points to high estimate accuracy with a high confidence level. A confidence interval can be attributed to almost every statistical measure. Although there are some other ways of calculating it, the confidence interval is generally and most frequently calculated using standard error. The P value and the confidence interval are two complementary statistical indicators. They describe the same thing, but in two different ways. The P value describes probability that the observed phenomenon (difference) occurred by chance, whereas the confidence interval provides margins of error within which it is possible to expect the value of that phenomenon. In the last twenty years, increasing number of journals require reporting of the confidence intervals for each of their key results. Reporting of this confidence interval provides additional information about the sample and the results. It is, moreover, very useful and irreplaceable supplement to a classical hypothesis testing and to the generally accepted P value. It should become a standard of all scientific journals to report key results with respective confidence intervals because it enables better understanding to the interested reader.

Key words: confidence interval, sample, standard error

Pristiglo: 7. veljače 2008.

Prihvaćeno: 1. travnja 2008.

Received: February 7, 2008

Accepted: April 1, 2008

Uvod

Svakom znanstvenom istraživanju prethodi temeljita analiza područja od interesa, postavljanje jasnog cilja i hipoteze te promišljeno planiranje ustroja istraživanja, načina prikupljanja i analize podataka kao i prikaza rezultata. Pritom je izbor reprezentativnog uzorka od presudne važnosti za uspješnost istraživanja (1). Dobar uzorak osigurava vjerodostojnost naših rezultata i zaključaka koji iz tih rezultata proizlaze. Ispitivanjem izvjesnih značajki uzorka, mi ustvari želimo steći uvid u stanje populacije. Populaciju je u cjelini najčešće nemoguće u cijelosti ispitati. Dakle, temeljem onoga što deskriptivnom analizom prikupljenih podataka saznamo o našem uzorku, donosimo zaključke za populaciju u cjelini (2). To je osnovni koncept tzv. inferencijske statistike. Ono što pri tome podrazumijevamo jest činjenica da sve ono što saznamo o uzorku pouzdano vrijedi za cijelu populaciju.

Pretpostavimo, primjerice, da nas zanima kolika je prosječna koncentracija kolesterola u populaciji. Da bismo odgovorili na to pitanje izabiremo uzorak koji držimo reprezentativnim za tu populaciju i u tom uzorku odredimo srednju vrijednost koncentracije kolesterola. Slučajnim izborom odlučili smo se za uzorak ($N = 121$) u kojemu dobijemo srednju vrijednost koncentracije kolesterola $5,7 \pm 1,4$ mmol/L. Srednja vrijednost, tj. aritmetička sredina nam u ovom slučaju služi kao procjena (engl. *point estimate*) koncentracije kolesterola u populaciji. Ono što želimo znati jest odgovor na pitanje: možemo li dobivenu srednju vrijednost smatrati dobrom procjenom koncentracije kolesterola u populaciji. Koji statistički pokazatelj nam na to ukazuje? O čemu ovisi pouzdanost te procjene? Odgovor na ta pitanja daje nam interval pouzdanosti (engl. *confidence interval*, CI).

Što je interval pouzdanosti?

Interval pouzdanosti za bilo koju statističku mjeru predstavlja raspon mogućih vrijednosti unutar kojega se s izvjesnom vjerojatnosti nalazi ta statistička mjera populacije.

Vratimo se na primjer prosječne koncentracije kolesterola u populaciji. Izabrali smo slučajnim izborom jedan uzorak od 121 ispitanika i na njemu odredili srednju vrijednost koncentracije kolesterola. Potom ponovimo taj postupak još jednom i dobijemo neki novi uzorak s novom aritmetičkom sredinom. Uzorkovanje ponavljamo 100 puta i svaki put dobijemo neku aritmetičku sredinu koncentracije kolesterola. Svako aritmetičkoj sredini uzorka pridružujemo i pripadajući interval pouzdanosti. Od ukupno stotinu pripadajućih 95%-tnih intervala pouzdanosti, njih 95% će sadržavati pravu aritmetičku sredinu populacije (μ). To ujedno i jest najtočnija definicija intervala pouzdanosti. Kao takav, interval pouzdanosti je objektivna procjena (ne)preciznosti i veličine uzorka nekog istraživanja (3). Sto-

Introduction

Every science research starts with a detailed analysis of the field of interest, defining the clear aim and hypothesis, by thorough planning of the research design and the way of data collecting and data analysis, as well as the reporting of the results. For a successful research the choice of the representative sample is crucial (1). A good sample assures reliability of our results and conclusions that arise from them. By analysing some sample characteristics we actually want to have an overview of the situation in the population. In the most cases it is impossible to examine the population as a whole. Therefore, based on what we find out about our sample using the descriptive analysis we make conclusions for the population as a whole (2). This is the basic concept of, so called, inferential statistics. What is hereby understood is the fact that everything we conclude about the sample is reliably applicable to the whole population.

Let us presume, for example, that we want to know what the average cholesterol concentration in the population is. To answer this question we select a sample for which we believe that represents the population and in this sample we calculate the mean of cholesterol concentration. By random selection we have decided for the sample ($N = 121$) and we determine the mean of cholesterol concentration: 5.7 ± 1.4 mmol/L. In this case we use the arithmetic mean as a point estimate of cholesterol concentration in the population. The thing we want to know is the answer to the question: can we consider this calculated mean as a good estimate for cholesterol concentration in the population. Which statistical indicator points to that? What does the reliability of this estimate depend on? The confidence interval (CI) gives the answer to all these questions.

What is confidence interval?

In statistics, for any statistical measure, a confidence interval presents a range of possible values within which, with some certainty, we can find the statistical measure of the population.

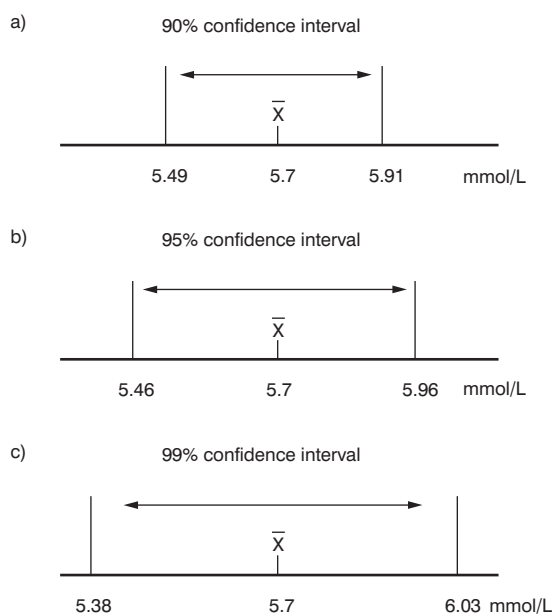
Let us go back to the example of the average cholesterol concentration in the population. By random selection we have chosen one sample of 121 individuals and on this sample we have calculated the mean cholesterol concentration. Subsequently, we repeat the procedure and we take a new sample with a new arithmetic mean. We repeat sampling for 100 times and each time we calculate a certain arithmetical mean of cholesterol concentration. To each arithmetical mean we give an appropriate confidence interval. Out of a total of hundred related 95%-confidence intervals, 95% of them will contain the actual arithmetical mean of the population (μ). This is also the most accurate definition of a confidence interval.

ga, na interval pouzdanosti posredno možemo gledati i kao na mjerilo kvalitete uzorka i istraživanja. Upravo zbog toga mnogi časopisi zahtijevaju od autora da svoje ključne rezultate u članku izraze uz pripadajuće intervale pouzdanosti (4,5).

Interval pouzdanosti je omeđen granicama. Ovisno o razini pouzdanosti koju biramo, mijenja se i raspon, tj. granice intervala. Najčešće korišteni intervali pouzdanosti u biomedicinskoj literaturi su 90%-tni, 95%-tni, 99%-tni i nešto rjeđe 99,9%-tni interval pouzdanosti. Što su granice intervala uže, preciznost procjene je veća. Tradicionalno se u literaturi najčešće koristi 95%-tni interval pouzdanosti, koji je u svezi s opće prihvaćenom razinom statističke značajnosti $P < 0,05$. Za uzorak iste veličine vrijedi pravilo: što je manja razina pouzdanosti – veća je preciznost procjene. Pogledajmo kako se raspon i granice intervala pouzdanosti mijenjaju u ovisnosti o razini pouzdanosti, na našem primjeru procjene koncentracije kolesterola populacije (Slika 1).

As such, confidence interval is a realistic estimate of (in)precision and sample size of certain research (3). Therefore, we can consider confidence interval also as a measure of a sample and research quality. Many journals therefore require providing the key results with respective confidence intervals (4,5).

Confidence interval is defined by its margins of error. Depending on the confidence level that we choose, the interval margins of error also change. The most used confidence intervals in the biomedical literature are the 90%, 95%, 99% and not so often 99.9% one. The narrower the margins of an interval are, the higher is the estimate accuracy. The 95% confidence interval is traditionally the most used interval in the literature and this relates to the generally accepted level of statistical significance $P < 0.05$. There is a rule for same sized samples: the smaller the confidence level is, the higher is the estimate accuracy. Let us now see how the range of confidence interval and its margins of error change depending on the confidence level in our example of cholesterol concentration estimate in the population (Figure 1).



SLIKA 1. a) 90%-tni interval pouzdanosti; b) 95%-tni interval pouzdanosti; c) 99%-tni interval pouzdanosti srednje vrijednosti koncentracije kolesterola ($N = 121$).

FIGURE 1. a) 90% confidence interval; b) 95% confidence interval; c) 99% confidence interval of mean cholesterol concentration ($N = 121$)

S 90%-tnom pouzdanosti možemo tvrditi da je koncentracija kolesterola u populaciji unutar granica intervala 5,49–5,91 mmol/L. Drugim riječima: ako sto puta slučaj-

We can claim with 90% confidence that the cholesterol concentration in the population lies within the interval margins of 5.49–5.91 mmol/L. In other words: if we select

nim izborom odaberemo uzorak od 121 osobe i na tom uzorku odredimo srednju koncentraciju kolesterola i interval pouzdanosti te procjene, onda u deset od tih sto uzoraka interval pouzdanosti neće obuhvatiti pravu srednju vrijednost populacije. Ono što ne znamo jest – kojih je to deset uzoraka? Upravo to i jest ono što našu procjenu čini (ne)pouzdanom.

Ukoliko se odlučimo za 95%-tnu pouzdanost, granice intervala su 5,46–5,96 mmol/L i u pet slučajnih uzoraka neće obuhvatiti stvarnu srednju vrijednost populacije, dok najveći raspon pripada 99%-tnoj pouzdanosti (5,38–6,03 mmol/L). Što je širi interval pouzdanosti, veća je vjerojatnost da taj interval obuhvaća i srednju vrijednost koncentracije kolesterola u populaciji.

Samo istraživanja provedena na velikom uzorku će dati vrlo uski interval pouzdanosti koji ukazuje na veliku preciznost procjene, s visokom razinom pouzdanosti.

Kako izračunati interval pouzdanosti?

Interval pouzdanosti se može pridružiti gotovo svakoj statističkoj mjeri: koeficijentu korelacije (6), omjeru rizika (engl. *odds ratio*, OR) te primjerice mjerama dijagnostičke točnosti kao što su osjetljivost, specifičnost i neke druge (7). Iako postoje i neki drugi načini, interval pouzdanosti najčešće računamo pomoću standardne greške. Standardna greška je standardna devijacija srednjih vrijednosti niza uzoraka, koju bismo dobili temeljem stotinu slučajnih uzoraka iz neke populacije (8).

Najprije valja odrediti razinu pouzdanosti s kojom želimo procijeniti srednju vrijednost nekog parametra u populaciji. Drugim riječima, zapravo se pitamo u kojoj mjeri si možemo/želimo dozvoliti pogrešnu procjenu? Najčešće se odlučujemo za 95%-tnu pouzdanost, što znači da ćemo dozvoliti da samo u 5% slučajeva naš interval pouzdanosti ne obuhvati pravu srednju vrijednost populacije.

Granice intervala pouzdanosti računamo koristeći Z vrijednost, standardnu devijaciju uzorka (SD) i veličinu uzorka (n) prema formuli:

$$Z \times \frac{SD}{\sqrt{N}}$$

Donju granicu intervala pouzdanosti dobijemo tako da od srednje vrijednosti oduzmemo izračunati iznos iz formule, a gornju granicu tako da srednjoj vrijednosti dodamo taj isti iznos. Interval pouzdanosti definiramo kao:

$$95\%CI = (\bar{X} - Z \times \frac{SD}{\sqrt{N}}) - (\bar{X} + Z \times \frac{SD}{\sqrt{N}})$$

Vrijednost Z ovisi o razini pouzdanosti za koju smo se odlučili. Valja istaknuti kako je interval pouzdanosti točan samo za uzorke koji slijede normalnu raspodjelu bez obzira

randomly hundred times a sample of 121 individuals and calculate the mean cholesterol concentration and the confidence interval of that estimate in the sample, then in ten out of these hundred samples the confidence interval will not include the actual mean of the population. The thing we do not know is – which 10 samples are we talking about? Exactly that is what makes our estimate (un)reliable.

If we decide for the 95% confidence, the interval margins are 5.46–5.96 mmol/L and in five randomly selected samples the mean of population will not be included. The largest range belongs to the 99% confidence (5.38–6.03 mmol/L). The larger the confidence interval is, the higher is the possibility that this interval includes also the mean of cholesterol concentration in the population.

Only the studies with a large sample will give a very narrow confidence interval, which points to high estimate accuracy with a high confidence level.

How to calculate a confidence interval?

A confidence interval can be attributed to almost every statistical measure: to a correlation coefficient (6), to odds ratio (OR) and e.g. to the measures of diagnostic accuracy such as sensitivity, specificity and some others (7). Although there are some other ways of calculating it, the confidence interval is generally and most frequently calculated using standard error. Standard error is a standard deviation of sample means, calculated out of hundred random population samples (8).

First we have to determine the confidence level for estimating the mean of a parameter in a population. In other words, we ask ourselves to what extent we accept a wrong estimate. Most often we decide for the 95% confidence what means that we will allow that only in 5% cases the actual mean of population does not fall into our interval.

Margins of error for a confidence interval are calculated using Z value, standard sample deviation (SD) and sample size (N) according to the formula:

$$Z \times \frac{SD}{\sqrt{N}}$$

The lower margin of a confidence interval is calculated by deducting the previous formula result from the mean. The upper margin is calculated by adding the formula result to the mean. The definition of a confidence interval is therefore:

$$95\%CI = (\bar{X} - Z \times \frac{SD}{\sqrt{N}}) - (\bar{X} + Z \times \frac{SD}{\sqrt{N}})$$

The Z value depends on the chosen confidence level. It should be kept in mind that a confidence interval is

na veličinu uzorka, dok je približno točan za velike uzorke koji nisu raspodijeljeni normalno. Ako se radi o malom uzorku ($N < 30$) u formulu za interval pouzdanosti valja umjesto Z unijeti t vrijednost. Vrijednost t proizlazi iz Studentove t raspodjele i ovisi o veličini uzorka, uz uvažavanje stupnjeva slobode $N-1$ (9). Za male uzorke t vrijednost je veća od Z vrijednosti, iz čega logično proizlazi da je interval pouzdanosti za manje uzorke uz istu razinu pouzdanosti širi. U Tablici 1 nalaze se Z vrijednosti za pripadajuće 90%, 95%, 99% i 99,9-tne razine pouzdanosti. Mnogi statistički udžbenici sadrže tablice s t vrijednostima za odgovarajuću razinu pouzdanosti i različite stupnjeve slobode (1).

accurate only for the samples that follow a normal distribution, whereas it is approximately accurate for large samples that are not distributed normally. For small samples ($N < 30$) the t value should be used instead of the Z value in the formula for confidence interval, with $N-1$ degrees of freedom (9). The t value comes from the Student's t-distribution and depends on the sample size. For small samples the t value is higher than the Z value what logically means that the confidence interval for smaller samples with the same confidence level is larger. Z values for matching 90%, 95%, 99% and 99.9% confidence levels are listed in the Table 1. Many statistical textbooks contain tables with t values for matching confidence level and different degrees of freedom (1).

TABLICA 1. Z vrijednosti za najčešće razine pouzdanosti

Confidence level	Z value
90%	1.65
95%	1.96
99%	2.58
99,9%	3.291

TABLE 1. Z values for the most frequent confidence levels

Za naš primjer uzorka u kojem želimo odrediti prosječnu koncentraciju kolesterola u populaciji, interval pouzdanosti bismo računali koristeći Z vrijednost budući da se radi o velikom uzorku ($N = 121$) koji slijedi normalnu raspodjelu.

U našem primjeru srednju vrijednost koncentracije kolesterola uz pripadajući 95%-tni interval pouzdanosti bismo naveli kao: 5,7 mmol/L (95% CI = 5,46–5,96).

Kako danas mnogi statistički programi nude mogućnost izračuna i iskaza intervala pouzdanosti za većinu statističkih pokazatelja, malo tko će uistinu ručno računati interval pouzdanosti za svoje podatke. No, važno je poznavati ulazne varijable iz kojih računamo interval pouzdanosti kako bismo mogli bolje razumjeti njegovo značenje i interpretaciju.

Imaju li vrijednost P i interval pouzdanosti isto značenje?

P vrijednost i interval pouzdanosti su dva komplementarna statistička pokazatelja. Oni govore o istoj stvari na dva različita načina i međusobno se nadopunjuju. P vrijednost nam govori o vjerojatnosti s kojom je moguće da se uočeni fenomen (razlika) dogodio slučajno, dok interval pouzdanosti nudi granice unutar kojih je moguće očekivati vrijednost tog fenomena.

In our sample example, with the mean cholesterol concentration in the population, the confidence interval would be calculated by using the Z value because of the large size of the sample ($N=121$) that is normally distributed.

In our example we would report the mean cholesterol concentration with appropriate 95% confidence interval as 5.7 mmol/L (95% CI = 5.46–5.96).

Due to the fact that today there are a lot of statistical softwares that calculate and provide confidence intervals for the majority of statistical indicators, we shall rarely calculate a confidence interval manually. However, it is important to know the input based on which the confidence interval is calculated, so we could better understand its meaning and interpretation.

Do the P value and the confidence interval have the same meaning?

The P value and the confidence interval are two complementary statistical indicators. They describe the same thing, but in two different ways and are complementary to each other. The P value describes probability that the observed phenomenon (deviation) occurred by chance, whereas the confidence interval provides margins of error within which it is possible to expect the value of that phenomenon.

Interval pouzdanosti možemo računati za razliku ili omjer između bilo koja dva statistička pokazatelja, da bismo ispitali možemo li toj razlici ili omjeru pripisati statističku značajnost. Pogledajmo na primjeru s koncentracijom kolesterola u populaciji, kako interval pouzdanosti može poslužiti u procjeni statističke značajnosti razlike između dvije srednje vrijednosti.

Razlika u koncentraciji kolesterola između muškaraca i žena u našem uzorku je 0,22 mmol/L. Je li ta razlika statistički značajna? Imaju li žene u našem uzorku uistinu niže koncentracije kolesterola od muškaraca ili je opažena razlika samo slučajna? Odgovor na to pitanje daje nam t-test. Vrijednost P dobivena t-testom iznosi 0,426, što nam ukazuje da se koncentracije kolesterola u muškaraca i žena ne razlikuju statistički značajno. To isto, samo na drugačiji način nam govori i 95%-tni interval pouzdanosti razlike srednjih vrijednosti koncentracije kolesterola muškaraca i žena, koji iznosi od -0,322 do 0,757 (Tablica 2).

Confidence interval can be calculated for difference or ratio between any two statistical indicators, so we could examine if this difference or ratio is of any statistical significance. Let us go back to our example of cholesterol concentration in the population to see how the confidence interval can be used for estimating statistical significance of the difference between two means.

The difference in the cholesterol concentration between men and women in our sample is 0.22 mmol/L. Is this difference statistically significant? Do women in our sample have indeed a lower cholesterol concentration than men, or did the observed difference only occur by chance? The answer to this question gives us the Student t-test. The P value calculated by the t-test is 0.426 proving that the difference of the cholesterol concentration between men and women is not statistically significant. The same thing, only in a different way, we can see in 95% confidence interval of difference of mean cholesterol concentration between men and women, which is (-0.322) to (+0.757) (Table 2).

TABLICA 2. Razlika u koncentraciji kolesterola između muškaraca i žena (N = 121)

	Women	Men
N	82	39
Mean	5,63	5,85
95% confidence interval of the mean	5.318-5.954	5.432-6.275
Difference	0,22	
SE	0,27	
95% confidence interval of the difference	(-0.322) to (+0.757)	
t-test	P = 0.426	

TABLE 2. Difference in the cholesterol concentration between men and women (N = 121)

Što možemo zaključiti iz tog intervala pouzdanosti? Prisjetimo se interpretacije intervala pouzdanosti: on definira granice unutar kojih s 95% pouzdanosti možemo očekivati pravu vrijednost. Naš interval pouzdanosti sadrži i nulu (0), što znači da je sasvim moguće da je prava vrijednost razlike ravna nuli, tj. da razlika koncentracije kolesterola muškaraca i žena ne postoji.

Kako interval pouzdanosti interpretiramo kad se on odnosi na omjer, kao što je primjerice OR? Recimo da smo svim našim ispitanicima iz uzorka (N = 121) slikovnim tehnikom procijenili prohodnost karotidnih arterija i prema nalazu ih podijelili u dvije skupine: one s prohodnim arterijama i one sa stenozom barem jedne karotidne arterije koja iznosi > 50% lumena. Zanima nas razlikuju li se te dvije

Which conclusion can we draw from that confidence interval? Let us remember what the definition of confidence interval was: it defines margins of error within which we can expect the actual value with 95% confidence. Our confidence interval contains also a zero (0) meaning that it is quite possible that the actual value of the difference will equal zero, namely, that there is no difference between the cholesterol concentration between men and women.

How do we define the confidence interval when it is related to a ratio as, for example, in OR? Let us imagine that we have assessed the degree of carotid artery stenosis for all our individuals from the sample (N = 121) using echo-colour Doppler sonography analysis. We divide them

skupine po prosječnoj koncentraciji kolesterola i je li koncentracija kolesterola čimbenik rizika za nastanak stenozu karotidnih arterija. Odgovor na to pitanje dati će nam OR (Tablica 3).

OR je veći od 1, no interval pouzdanosti za OR obuhvaća brojku 1. Što to znači? To znači da je moguće da je jednako izgledno da koncentracija kolesterola jest i nije čimbenik rizika za stenozu karotidnih arterija. Odnosno, za bilo koju koncentraciju kolesterola izgledi da neka osoba ima ili nema stenozu karotidnih arterija su jednaki. O tome nam govori i postotak ispravno razvrstanih ispitanika s obzirom na koncentraciju kolesterola (50,41%). Tek svaki drugi ispitanik je razvrstan u odgovarajuću skupinu – dakle izbor ne ovisi o koncentraciji kolesterola, već je isključivo posljedica slučajnosti.

into two groups: the ones with the absence of carotid artery stenosis and the others with stenosis (> 50% lumen) of at least one carotid artery. We are interested whether those two groups are different in average cholesterol concentration and is the cholesterol concentration a risk factor for development of carotid artery stenosis. The answer to this question gives us the OR (Table 3).

OR is higher than 1, but the confidence interval for OR includes the number 1 as well. What does it mean? It means that the odds of cholesterol concentration being and not being a risk factor for carotid artery stenosis are even. That is, for any cholesterol concentration the odds of that person having or not having carotid artery stenosis are the same. This also confirms the percentage of correctly classified individuals considering the cholesterol concentration (50.41%). Only half of the individuals are correctly categorised in an adequate group – so the selection does not depend on cholesterol concentration but on a pure chance.

TABLICA 3. OR i 95%-tni interval pouzdanosti koncentracije kolesterola u diskriminaciji osoba sa stenozom karotidnih arterija

Parameter	OR	95% confidence interval	Correctly classified individuals, %
Cholesterol	1.204	0.777-1.678	50.41

TABLE 3. OR and 95% confidence interval of cholesterol concentration in individuals with/without carotid artery stenosis

Zaključak

Interval pouzdanosti moguće je pridružiti gotovo svakom statističkom pokazatelju. Posljednjih dvadesetak godina sve je više časopisa u kojima je obveza autora prikazati intervale pouzdanosti za svoje ključne rezultate (nalaze). Prikaz intervala pouzdanosti pruža dodatnu informaciju o našem uzorku i rezultatima, i nadalje je korisna i nezamjenjiva nadopuna klasičnom testiranju hipoteze i opće prihvaćenoj P vrijednosti. Prikaz ključnih rezultata uz pripadajuće intervale pouzdanosti trebao bi postati standard svih znanstvenih časopisa, jer zainteresiranom čitatelju omogućuje bolje razumijevanje prikazanih podataka.

Conclusion

Confidence interval can be attributed to almost every statistical measure. In the last twenty years, every day there are more journals that require reporting of the confidence intervals for each of their key results. Reporting of this confidence interval provides additional information about the sample and the results. It is, moreover, above all useful and irreplaceable supplement to a classical hypothesis testing and to the generally accepted P value. It should become a standard of all scientific journals to report key results with respective confidence intervals because it enables better understanding of the data to an interested reader.

Adresa za dopisivanje:

Ana-Maria Šimundić, Pomoćni urednik
Klinički zavod za kemiju
Klinička bolnica „Sestre milosrdnice“
Vinogradska 29
10 000 Zagreb
Hrvatska
e-pošta: am.simundic@gmail.com
mob: 098 554-708
tel: 01 3787-432
faks: 01 3768-280

Corresponding author:

Ana-Maria Simundic, Assistant editor
University Department of Chemistry
Sestre milosrdnice University Hospital
Vinogradska 29
10 000 Zagreb
Croatia
e-mail: am.simundic@gmail.com
mob: +385 98 554-708
tel: +385 1 3787-432
faks: +385 1 3768-280

Literatura/References:

1. Glantz SA. *Primer of biostatistics*. 5th ed. New York (NY): McGraw-Hill; 2002.
2. Simundic AM. *Types of variables and distribution*. *Acta Med Croat* 2006;60(Suppl 1):17-35.
3. Altman DG. *Why we need confidence intervals*. *World J Surg* 2005;29:554-6.
4. Sterne JAC, Smith GD. *Sifting the evidence – what's wrong with significance tests?* *BMJ* 2001;322:226-31.
5. Russel I. *Statistics – with confidence?* *Brit J Gen Pract* 1991; 41: 179–80.
6. Udovičić M, Baždarić K, Bilić-Zulle L, Petrovečki M. *What we need to know when calculating the coefficient of correlation*. *Biochemia Medica* 2007;17:10-5.
7. Raslich MA, Markert RJ, Stutes SA. *Selecting and interpreting diagnostic tests*. *Biochemia Medica* 2007;17:151-61.
8. McHugh M. *Standard error*. *Biochemia Medica* 2008;18:7-13.
9. Bland M. *An introduction to medical statistics*. 3rd ed. New York (NY): Oxford University Press; 2000.