PREGLEDNI RAD

UDK BROJEVI: 615.036
616-071
DOI: https://www.doi.org/10.5937/zdravzast52-47107

TUMAČENJE REZULTATA KLINIČKIH RANDOMIZOVANIH STUDIJA

Božana Nikolić[1]

[1] Katedra za farmaciju, Medicinski fakultet Univerziteta u Novom Sadu, Novi Sad, Republika Srbija

[*] Korespondencija: doc. dr Božana Nikolić, Katedra za farmaciju, Medicinski fakultet Univerziteta u Novom Sadu, Hajduk Veljkova 3, 21137 Novi Sad, Republika Srbija; e-mail: bozana.nikolic@mf.uns.ac.rs

SAŽETAK

Cilj ovog preglednog rada je da kroz primere objasni kako izabrati, računati i tumačiti mere efekta u kliničkim randomizovanim studijama. Kliničke randomizovane studije (engl. *Randomized Control Trials*, RCTs) pripadaju grupi eksperimentalnih studija i sprovode se nad obolelim osobama. U ovim studijama ispitanici se metodom randomizacije raspoređuju u eksperimentalnu grupu, koja dobija ispitivano sredstvo, i kontrolnu, koja prima placebo ili neko do tada korišćeno sredstvo. Rezultati RCTs mogu se prezentovati kroz relativne i apsolutne mere efekta. Relativne mere efekta su relativni rizik (engl. R*elative Risk*, RR) i relativno smanjenje rizika (engl. *Relative Risk Reduction*, RRR), a apsolutne mere efekta su apsolutno smanjenje rizika (engl. *Absolute Risk Reduction*, ARR) i broj pacijenata koje treba lečiti (engl. *Number Needed to Treat*, NNT). Najjači dokaz između uzroka i posledice dobija se u ovim studijama, ali se ove studije retko izvode zbog svoje cene. Mali broj ispitanika u ovim studijama, može se prevazići korišćenjem meta-analize i multicentričnih studija.

Ključne reči: epidemiološke studije, kliničke randomizovane studije, mere efekta

Uvod

Sve epidemiološke studije mogu biti eksperimentalne (interventne) i opservacione (Grafikon 1). U eksperimentalnim studijama istraživač ispituje efikasnost korišćenja leka ili nekog drugog sredstva, dok u opservacionim studijama istraživač ne interveniše, nego posmatra šta se dešava u populaciji (npr. ispitanici koriste lekove nezavisno od istraživača tj. istraživanja, odnosno kao deo rutinske medicinske nege). Nadalje, eksperimentalne studije se prema načinu razvrstavanja ispitanika u grupe dele na randomizovane i nerandomizovane. Randomizacija je proces kojim se ispitanici metodom slučajnog izbora raspoređuju u eksperimentalnu i kontrolnu grupu, da bi ispitanici u obe posmatrane grupe bili što sličniji po svojim karakteristikama.

Opservacione studije mogu biti deskriptivne i analitičke. Tako, deskriptivne studije (prikaz slučaja, serija slučajeva, ekološka studija) nemaju grupu za poređenje, dok analitičke studije (kohortna studija, studija slučajeva i kontrola, studija preseka) karakteriše poređenje dve grupe (1). Opservacionim studijama se ne ispituje efikasnost medicinske intervencije, kao što je lek ili medicinski uređaj, ali mogu pomoći da se identifikuju novi tretmani ili preventivne mere koji bi se potom testirali u kliničkim ispitivanjima.

Prema Uredbi Evropske unije br. 536/2014, kliničke studije se definišu kao istraživanja na ljudima čija je namena da otkriju ili potvrde kliničke, farmakološke i/ili druge farmakodinamske efekte lekova; i/ili utvrde neželjene reakcije na lekove; i/ili prouče apsorpciju, raspodelu, metabolizam i/ili eliminaciju lekova (1,2). Randomizovane kontrolisane studije (engl. *Randomized Controlled Trials*, RCTs) su zlatni standard za izradu smernica za lečenje, dok su sve druge studije komplementarne i korisne za generisanje hipoteza. Njihovi rezultati se izveštavaju kao mere efekta, koje mogu biti relativne (mere odnosa) i apsolutne (mere razlike) (3,4).

Cilj ovog preglednog rada je da kroz primere objasni kako izabrati, računati i tumačiti mere efekta u RCTs.

Health Care 52(4), 2023

REVIEW ARTICLE

UDK NUMBERS: 615.036
616-071
DOI: https://www.doi.org/10.5937/zdravzast52-47107

INTERPRETING THE RESULTS OF RANDOMIZED CONTROLLED TRIALS

Božana Nikolić[1]

[1] Department of Pharmacy, Faculty of Medicine, University of Novi Sad, Republic of Serbia

[*] Correspondence: Assoc. professor Božana Nikolić, MD, Department of Pharmacy, Faculty of Medicine, University of Novi Sad, Novi Sad Hajduk Veljkova 3, 21137 Novi Sad, Republic of Serbia; e-mail: bozana.nikolic@mf.uns.ac.rs

SUMMARY

The aim of this review article is to explain, with the help of examples, how to choose, calculate and interpret effect measures in randomized controlled trials. Randomized controlled trials (RCTs) belong to the group of experimental studies and they are conducted on sick persons. In these studies, participants are assigned to the experimental group, which receives the medicine which is being tested, and to the control group, which receives a placebo or a medicine, which has been previously used. The results of RCTs can be presented through relative or absolute effect measures. Relative effect measures include the relative risk (RR) and the relative risk reduction (RRR), while absolute effect measures include the absolute risk reduction (ARR) and the number of patients who should be treated (Number Needed to Treat – NNT). The strongest evidence of cause and effect is obtained in these studies, but these studies are rarely conducted because of their cost. The small number of participants in these studies can be overcome by using meta-analysis and multicenter studies.

Keywords: epidemiological studies, randomized controled trial, effect measures

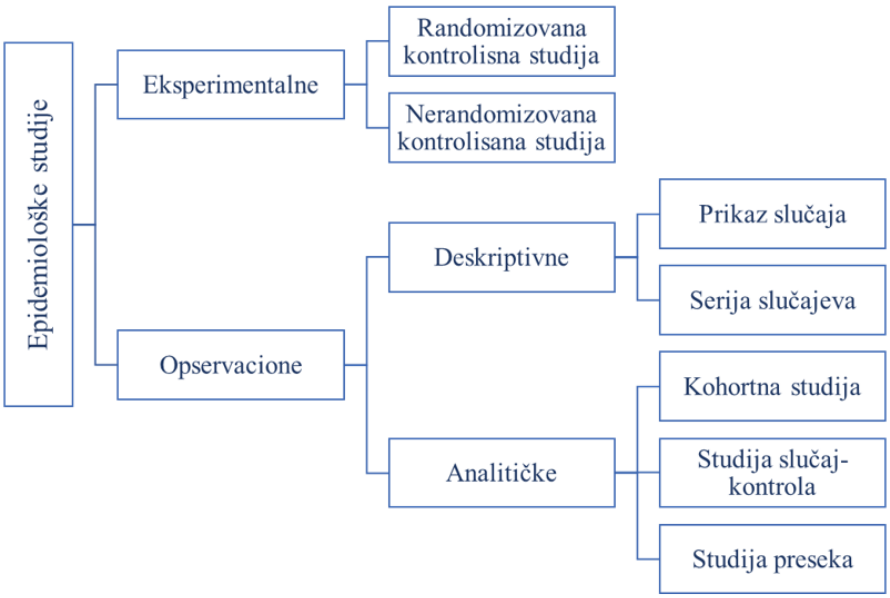
Introduction

All epidemiological studies can be classified as experimental (interventional) and observational (Figure 1). In experimental studies, the researcher examines the efficacy of using a drug or some other treatment, while in observational studies, the researcher does not intervene, but observes what is happening in that population (e.g. participants use drugs independently of the researcher, that is, as part of routine medical care). Furthermore, experimental studies are classified as randomized and non-randomized according to the ways of assigning the participants to groups. Randomization is the process which refers to a random assignment of participants to the experimental and control group, so that the participants in both observed groups are as similar as possible in terms of their characteristics.

Observational studies can be descriptive and analytical. Thus, descriptive studies (case report, case series, ecological study) do not have a control group, while analytical studies (cohort study, case-control study, cross-sectional study) are characterized by the comparison of two groups (1). Observational studies do not examine the efficacy of a medical intervention, such as a drug or medical device, but they can help identify new treatments or preventive measures that can be tested in clinical trials afterwards.

According to the Regulation of the European Union no. 536/2014, clinical studies are defined as research on humans whose purpose is to discover or verify clinical, pharmacological and/or other pharmacodynamic effects of drugs; and/or identify any adverse reactions; and/or study the absorption, distribution, metabolism and/or excretion of drugs (1,2). Randomized controlled trials (RCTs) are the gold standard for creating guidelines for the treatment, while all other studies are complementary and useful for generating hypotheses. Their results are reported as effect measures, which can be relative (ratio measures) and absolute (difference measures) (3,4).

The aim of this review article is to explain through examples how to choose, calculate and



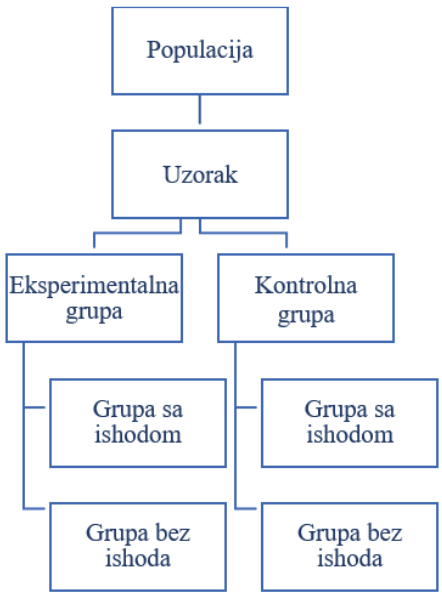
Grafikon 1. Podela epidemioloških studija

Metode

Na osnovu pregleda literature, korišćenjem MEDLINE bibliografske baze podataka, identifikovani su radovi koji se odnose na RCTs u smislu izbora, računanja i tumačenja mera efekata.

Randomizovani kontrolisani trajali (engl. *Randomized Controlled Trials* – RCTs)

U RCTs (Grafikon 2) ispitanici (obolele osobe) se metodom randomizacije raspoređuju u eksperimentalnu (interventnu) i kontrolnu grupu. Randomizacija može da bude prosta, stratifikovana, blok, itd. Eksperimentalna grupa može da dobije lek ili neko drugo ispitivano sredstvo (npr. novi lek, stari lek ali nova doza, kombinaciju dva ili više lekova i slično), dok kontrolna grupa dobija placebo ili standardno lečenje (lek koji se uobičajeno koristi za ispitivanu indikaciju), pa se prati ishod u obe grupe ispitanika (Grafikon 2). Ishodi, ako se koristi neki lek, mogu da budu: bolest da/ne, relaps bolesti da/ne, ili smrt nastupila/nije nastupila (1,4).



Grafik 2. Randomizovana klinička studija

Primer RCT u kojima lek smanjuje rizik od neželjenih ishoda

RCT je sproveden kako bi se ispitali efekti enalaprila na preživljavanje pacijenata sa srčanom insuficijencijom (n=2569) (5). Pacijenti su metodom randomizacije podeljeni u eksperimentalnu i kontrolnu grupu. Eksperimetalna grupa je dobijala enalapril (n=1285), a kontrolna grupa placebo (n=1284). RCT je trajao četiri godine i za to vreme smrt je nastupila kod 452 pacijenta koja su dobijala enalapril i 510 pacijenata koji su dobijali placebo (Tabela 1).

Rezultati RCTs mogu se prezentovati kroz relativne i apsolutne mere efekta. Relativne mere efekta su relativni rizik (engl. *Relative Risk, RR*) i relativno smanjenje rizika (engl. *Relative Risk Reduction*, RRR), a apsolutne mere efekta su apsolutno smanjenje rizika (engl. *Absolute Risk Reduction*, ARR) i broj pacijenata koje treba lečiti (engl. *Number Needed to Treat*, NNT) (4,6).

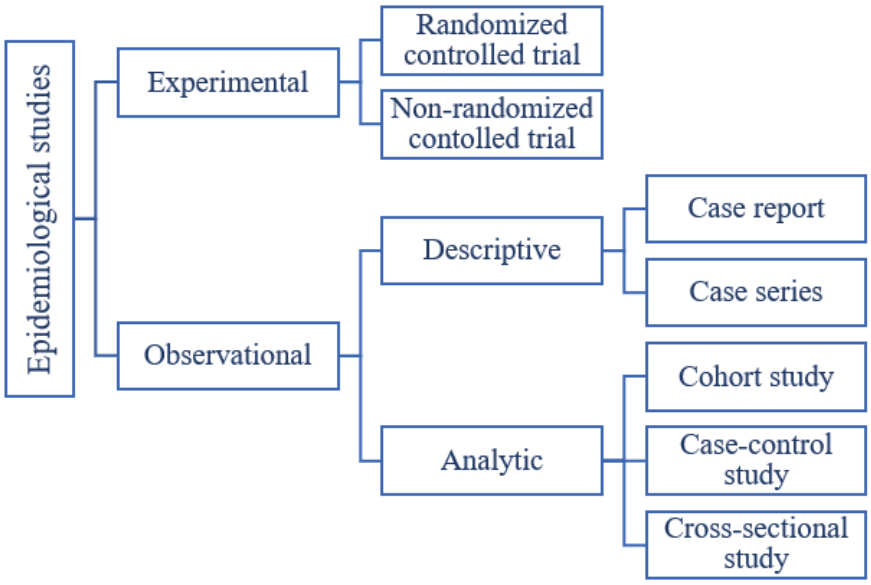


Figure 1. Classification of epidemiological studies

interpret effect measures in RCTs.

Methods

Based on the literature review, using the MEDLINE bibliographic database, studies relating to RTCs in terms of selection, calculation and interpretation of effect measures were identified.

Randomized controlled trials (RCTs)

In RCTs (Figure 2), participants (patients) are assigned to the experimental (interventional) group and to the control group with the help of the randomization method. Randomization can be simple, stratified, block, etc. The experimental group can receive a drug or some other tested medicinal product (e.g. a new drug, an old drug but a new dose, a combination of two or more drugs, etc.), while the control group receives a placebo or standard treatment (a drug that is commonly used for the examined indication), and then the outcome is observed in both groups of participants (Figure 2). Outcomes, if a drug is used, may be: disease yes/no, disease relapse yes/no, or death occurred/did not occur (1,4).

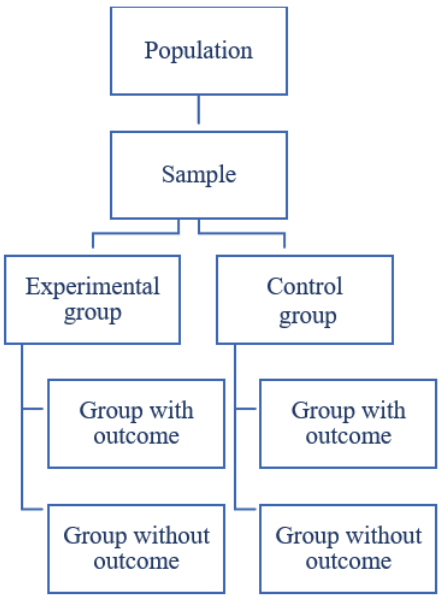


Figure 2. Randomized controlled trial

An example of RCTs in which a drug reduces the risk of adverse outcomes

RCT was conducted to examine the effects of enalapril on survival in patients with heart failure (n=2569) (5). The patients were divided into experimental and control groups using the method of randomization. The experimental group received enalapril (n=1285), and the control group received a placebo (n=1284). RCT lasted for four years, and during that period death occurred in 452 patients who received enalapril and 510 patients who received the placebo (Table 1).

The results of RCTs can be presented through relative and absolute effect measures. Relative effect measures are the relative risk (RR) and the relative risk reduction (RRR), while absolute effect measures are the absolute risk reduction (ARR) and the number of patients who need to be treated (NNT) (4,6).

Tabela 1. Rezultati četvorogodišnje randomizovane kontrolisane kliničke studije: enalapril (n=1285) naspram placebo (n=1284)

Grupa	Smrt		Ukupno
	Da	Ne	
Enalapril	452 a	833 b	1285 a + b
Placebo	510 c	774 d	1284 c + d

$$\text{Rizik od ishoda u eksperimentalnoj grupi}=\frac{a}{a+b}=\frac{452}{1285}=0{,}352$$

$$\text{Rizik od ishoda u kontrolnoj grupi}=\frac{c}{c+d}=\frac{510}{1284}=0{,}397$$

$$RR=\frac{a/(a+b)}{c/(c+d)}=\frac{0{,}352}{0{,}397}=0{,}89$$

$$RRR=\frac{[c/(c+d)-a/(a+b)]}{c/(c+d)}=1-RR=0{,}11$$

$$ARR=c/(c+d)-a/(a+b)=0{,}397-0{,}352=0{,}045$$

$$NNT=\frac{1}{ARR}=22{,}22$$

RR, relativni rizik; RRR, relativno smanjenje rizika; ARR, apsolutno smanjenje rizika; NNT, broj pacijenata koje treba lečiti.

Relativni rizik (RR) je odnos rizika od ishoda u eksperimentalnoj grupi i rizika od ishoda u kontrolnoj grupi (Tabela 1). U ovom primeru, rizik od smrti u eksperimentalnoj grupi je 0,352 ili 35,2%, a rizik od smrti u kontrolnoj grupi je 0,397 ili 39,7%, te je RR=0,89 ili 89% (Tabela 1), što znači da enalapril u odnosu na placebo smanjuje rizik od smrti kod pacijenata sa srčanom insuficijencijom za 11%.

Generalno, kada je RR < 1 to ukazuje da lečenje smanjuje rizik od umiranja (ishoda), a RR > 1 znači da lečenje povećava rizik od umiranja (ishoda). Sa kliničkog aspekta, vrednosti RR manje od 0,5 i veće od 2 se smatraju značajnim. Takođe, vrednosti RR koje su bliže jedinici se mogu smatrati značajnim ukoliko su ishodi ozbiljni ili od izuzetnog javnozdravstvenog interesa (7). Konačno, kada je RR=1, to znači da nema razlike u riziku od ishoda između eksperimentalne i kontrolne grupe.

Relativno smanjenje rizika (RRR) je relativno smanjenje rizika od neželjenog ishoda u eksperimentalnoj grupi u odnosu na rizik od neželjenog ishoda u kontrolnoj grupi (Tabela 1). Što je njegova vrednost veća, to je lečenje efikasnije. RRR u ovom primeru je 0,11 tj. 11% (Tabela 1); to znači da enalapril u odnosu na placebo smanjuje rizik od fatalnog ishoda za 0,11 puta odnosno 11%.

Apsolutno smanjenje rizika (ARR) je apsolutna razlika između rizika od neželjenog ishoda u kontrolnoj i eksperimentalnoj grupi (Tabela 1). Ukoliko je njegova vrednost nula, to ukazuje da između eksperimentalne i kontrolne grupe nema razlike u riziku od ishoda (tj. lečenje ne menja rizik od ishoda - umiranja). ARR u ovom primeru je 0,045 tj. 4,5% (Tabela 1). To znači da je razlika u riziku od smrtnog ishoda između grupe koja je dobijala placebo i grupe koja je dobijala enalapril 4,5%.

Za razliku od RRR, ARR uzima u obzir činjenicu da se neželjeni ishod (npr. smrt) javlja i u populaciji koja nije izložena leku odnosno placebu (3,6).

NNT je broj pacijenata koje treba lečiti u eksperimentalnoj grupi nekim lekom da bi se dobio jedan neželjeni ishod manje u odnosu na kontrolnu grupu koja obično dobija do tada poznato ispitivano sredstvo/placebo. Matematički, predstavlja recipročnu vrednost ARR. NNT u ovom primeru je 22,22 (Tabela 2), što znači da 23 pacijenta treba lečiti enalaprilom da bi se dobio jedan smrtni ishod manje u odnosu na lečenje placebom. Prema konvenciji, NNT se uvek tumači kao ceo broj (leči se ceo, a ne deo pacijenta) uz zaokruživanje na veću vrednost (izbegava se precenjivanje efikasnosti lečenja).

Tumačenje NNT je arbitrarno, i zavisi od kliničkog konteksta. Na primer, za profilkatičke postupke prihvatljiv NNT je između 10 i 100; za hronična stanja, lečenje je vrlo efikasno ako je NNT do 10; dok za akutna stanja (npr. infekcije koje se leče antibioticima), NNT ne bi smeo biti veći od 2. Dakle, poželjno je da NNT ima što manju vrednost (8).

NNT je specifičan za bolest, težinu bolesti, ishod bolesti i trajanje lečenja, pa te varijable treba navesti kako bi se NNT mogao pravilno tumačiti (2). Na primer, pacijenti sa hipertenzijom koriste antihipertenzive kako bi prevenirali kardiovaskularne događaje (infarkt miokarda, moždani udar); i NNT je 141 kada oni sa blagom do umerenom dijastolnom hipertenzijom (90-109 mmHg) koriste antihipertenzive pet godina, a NNT je tri kada oni sa teškom dijastolnom hipertenzijom (110-129 mmHg) koriste antihipertenzive pet godina (8). Dakle, mnogo je više pacijenata sa teškom dijastolnom hipertenzijom koji imaju korist od lečenja. Ili, na primer, duloksetin (60 mg/dan) se koristi u lečenju velikih depresivnih poremećaja i za remisiju je potrebno nekoliko nedelja. Zbog toga je

The relative risk (RR) is the ratio of the risks of an outcome in an experimental group to the risks of an outcome in the control group (Table 1). In this example, the risk of death in the experimental group is 0.352 or 35.2%, while the risk of death in the control group is 0.397 or 39.7%, so RR=0.89 or 89% (Table 1), which means that enalapril compared to placebo reduces the risk of death in patients with heart failure by 11%.

In general, when the RR < 1, it indicates that the treatment reduces the risk of death (outcome), while the RR > 1 means that the treatment increases the risk of death (outcome). From a clinical perspective, the values of RR smaller than 0.5 and greater than 2 are considered to be significant. Also, RR values, which are close to 1, can be considered significant if the outcomes are serious or have the exceptional public health significance (7). Finally, when the RR=1, it means that there is no difference in the risk of an outcome between the experimental and control group.

The RRR is the relative risk reduction in the risk of an adverse outcome in the experimental group compared to the risk of an adverse outcome in the control group (Table 1). The treatment is more efficient when its value is higher. The RRR in this example is 0.11, that is, 11% (Table 1), which means that enalapril reduces the risk of a fatal outcome by 0.11 times or 11% in comparison to placebo.

The ARR is the absolute difference between the risk of an adverse outcome in the control and experimental group (Table 1). If its value is zero, it means that there is no difference in the risk of an outcome between the experimental and control group (the treatment does not change the risk of the outcome – dying). The ARR in this example is 0.045 or 4.5% (Table 1). This means that the difference in the risk of deathly outcome between the group which received placebo and the group which received enalapril is 4.5%.

Unlike the RRR, the ARR takes into consideration the fact that the adverse outcome (e.g. death) occurs in the population that is not exposed to the drug or placebo (3,6).

The NNT is the number of patients who need to be treated in the experimental group with a drug in order to achieve one adverse outcome less compared to the control group that usually receives a previously known tested medicinal product/placebo. Mathematically, it represents a reciprocal value of ARR. The NNT in this example is 22.22 (Table 2), which means that 23 patients need to be treated with enalapril to obtain one deathly outcome less compared to the placebo treatment. According to the Convention, the NNT is always interpreted as a whole number (the whole patient is treated, and not a part of the patient), so the NNT is rounded up to the next higher whole number (overestimation of treatment efficacy is avoided).

The interpretation of the NNT is arbitrary, and it depends on the clinical context. For example, for prophylactic procedures, the acceptable NNT is between 10 and 100; for chronic conditions, the treatment is very efficient is the NNT is up to 10; while for acute conditions (e.g. infections treated with antibiotics), the NNT should not be higher than 2. Therefore, it is desirable that the NNT has the lowest possible value (8).

The NNT is specific for the disease, the severity of the disease, the outcome of the disease and the duration of treatment, and therefore, these variables should be specified so that the NNT could be interpreted correctly (2). For example, patients with hypertension use antihypertensive

Table 1. Results of a four-year randomized controlled trial: enalapril (n=1285) vs pacebo (n=1284)

Group	Death		Total
	Yes	No	
Enalapril	452 a	833 b	1285 a + b
Placebo	510 c	774 d	1284 c + d

$$\text{Risk of outcome in experimental group} = \frac{a}{a+b} = \frac{452}{1285} = 0.352$$

$$\text{Risk of outcome in control group} = \frac{c}{c+d} = \frac{510}{1284} = 0.397$$

$$RR = \frac{a/(a+b)}{c/(c+d)} = \frac{0.352}{0.397} = 0.89$$

$$RRR = \frac{[c/(c+d) - a/(a+b)]}{c/(c+d)} = 1 - RR = 0.11$$

$$ARR = c/(c+d) - a/(a+b) = 0.397 - 0.352 = 0.045$$

$$NNT = \frac{1}{ARR} = 22.22$$

RR, relative risk; RRR, relative risk reduction; ARR, absolute risk reduction; NNT, number needed to treat.

NNT u prvoj nedelji lečenja 79, a nakon toga se NNT postepeno smanjuje i u devetoj nedelji lečenja je šest (9).

Primer RCTs u kojima lek povećava rizik od neželjenih ishoda

U hipotetičkom RCT-u su ispitivani efekti antidepresiva venlafaksina na seksualnu funkciju. Pacijenti sa depresijom (n=152) su metodom randomizacije podeljeni u eksperimentalnu i kontrolnu grupu. Eksperimentalna grupa je dobijala venlafaksin (n=80), a kontrolna grupa placebo (n=72). RCT je trajao 12 nedelja i za to vreme seksualna disfunkcija se prvi put javila kod 16 pacijenata koji su dobijali venlafaksin i šest pacijenata koji su dobijali placebo (Tabela 2).

Rezultati RCT-a u kojima lek povećava rizik od neželjenih ishoda se izveštavaju kao relativni rizik (RR), relativno povećanje rizika (engl. *Relative Risk Increase*, RRI), apsolutno povećanje rizika (engl. *Absolute Risk Increase*, ARI) i broj pacijenata koji je potreban za neželjeni ishod (engl. *Number Needed to Harm*, NNH) (1).

Tabela 2. Rezultati hipotetičke, dvanaestonedeljne randomizovane kontrolisane studije: venlafaksin (n=80) naspram placebo (n=72)

Grupa	Seksualna disfunkcija		Ukupno
	Da	Ne	
Venlafaksin	16 a	64 b	80 a + b
Placebo	6 c	66 d	72 c + d

$$\text{Rizik od ishoda u eksperimentalnoj grupi} = \frac{a}{a+b} = \frac{16}{80} = 0{,}200$$

$$\text{Rizik od ishoda u kontrolnoj grupi} = \frac{c}{c+d} = \frac{6}{72} = 0{,}083$$

$$RR = \frac{a/(a+b)}{c/(c+d)} = \frac{0{,}200}{0{,}083} = 2{,}41$$

$$RRI = \frac{[a/(a+b) - c/(c+d)]}{c/(c+d)} = RR-1 = 2{,}41-1 = 1{,}41$$

$$ARI = a/(a+b) - c/(c+d) = 0{,}200-0{,}083 = 0{,}117$$

$$NNH = \frac{1}{ARI} = \frac{1}{0{,}117} = 8{,}55$$

RR, relativni rizik; RRI, relativno povećanje rizika; ARI, apsolutno povećanje rizika; NNH, broj pacijenata koji je potreban za neželjeni ishod

Kako je navedeno, RR je odnos rizika od ishoda u eksperimentalnoj grupi i rizika od ishoda u kontrolnoj grupi. U ovom primeru, rizik od seksualne disfunkcije u eksperimentalnoj grupi je 0,200 ili 20,0%, a u kontrolnoj grupi 0,083 ili 8,3%, te je RR=2,41 ili 241% (Tabela 2), što ukazuje da venlafaksin u odnosu na placebo povećava rizik od seksualne disfunkcije kod pacijenata sa depresijom.

Relativno povećanje rizika (RRI) je relativno povećanje rizika od neželjenog ishoda u eksperimentalnoj grupi u odnosu na rizik od neželjenog ishoda u kontrolnoj grupi (Tabela 2). U ovom primeru, RRI = 1,41 tj. 141% (Tabela 2), to znači da venlafaksin u odnosu na placebo povećava rizik od seksualne disfunkcije za 1,41 put odnosno 141%.

Apsolutno povećanje rizika (ARI) je apsolutna razlika između rizika od neželjenog ishoda u eksperimentalnoj i kontrolnoj grupi (Tabela 2). U ovom primeru, ARI = 0,117 tj. 11,7% (Tabela 2); dakle, apsolutna razlika u riziku od seksualne disfunkcije između grupe koja dobija venlafaksin i grupe koja dobija placebo je 11,7%.

Broj pacijenata koji je potreban za neželjeni ishod (NNH) je broj pacijenata koje treba lečiti u eksperimentalnoj grupi da bi se dobio jedan neželjeni ishod više u odnosu na kontrolno lečenje. Matematički, predstavlja recipročnu vrednost ARI -ja (Tabela 2). NNH u ovom primeru je 8,55 (Tabela 2), što znači da osam pacijenata treba lečiti venlafaksinom da bi se dobila jedna seksualna disfunkcija više u odnosu na grupu koja je dobijala placebo. Prema konvenciji, NNH se uvek tumači kao ceo broj (leči se ceo, a ne deo pacijenta) uz zaokruživanje na manju vrednost (izbegava se precenjivanje bezbednosti terapije).

Kao i NNT, tumačenje NNH je arbitrarno i zavisi od kliničkog konteksta. Za neželjene ishode koji ozbiljno kompromituju zdravstveni status pacijenta, prihvatljiv NNH bi trebalo da bude 1000; za neželjene ishode koji nisu ozbiljni, ali zahtevaju prekid lečenja, prihvatljiv NNH bi mogao da bude u opsegu od 10 do 100; konačno, za blage, prolazne neželjene ishode prihvatljiv je NNH ispod 10 (1,8). Dakle, za razliku od NNT, NNH je poželjno da ima što veću vrednost.

Relativne naspram apsolutne mere efekta – implikacije za kliničku praksu

Uprkos tome što je važno da se rezultati RCTs izveštavaju kao relativne i aposolutne mere efekta, oni se uglavnom izveštavaju kao relativne mere

drugs in order to prevent cardiovascular events (myocardial infarction, stroke); and the NNT is 141 when patients with mild to moderate diastolic hypertension (90-109 mmHg) use antihypertensive drugs for five years, while the NNT is three when those with severe diastolic hypertension (110-129 mmHg) use antihypertensive drugs for five years (8). Thus, there are many more patients with severe diastolic hypertension who benefit from the treatment. Or, for example, duloxetine (60 mg/day) is used for the treatment of depressive disorders and remission takes several weeks. Therefore, the NNT is 79 in the first week of treatment, after which the NNT gradually decreases and it amounts to six in the ninth week of treatment (9).

An example of RCTs in which a drug increases the risk of adverse outcomes

In a hypothetical RCT, the effects of the antidepressant venlafaxine on sexual function were examined. The patients with depression (n=152) were randomly divided into the experimental and control group. The experimental group received venlafaxine (n=80), and the control group received placebo (n=72). The RCT lasted for 12 weeks, and during that period sexual dysfunction first occurred in 16 patients who received venlafaxine and in six patients who received placebo (Table 2).

Table 2. Results of a hypothetical twelve-week randomized controlled trial: venlafaxine (n=80) vs placebo (n=72)

Group	Sexual dysfunction		Ukupno
	Da	Ne	
Venlafaksin	16 a	64 b	80 a + b
Placebo	6 c	66 d	72 c + d

$$\text{Risk of outcome in experimental group} = \frac{a}{a+b} = \frac{452}{1285} = 0.352$$

$$\text{Risk of outcome in control group} = \frac{c}{c+d} = \frac{510}{1284} = 0.397$$

$$RR = \frac{a/(a+b)}{c/(c+d)} = \frac{0.352}{0.397} = 0.89$$

$$RRR = \frac{[c/(c+d) - a/(a+b)]}{c/(c+d)} = 1 - RR = 0.11$$

$$ARR = c/(c+d) - a/(a+b) = 0.397 - 0.352 = 0.045$$

$$NNT = \frac{1}{ARR} = 22.22$$

RR, relative risk; RRI, relative risk increase; ARI, absolute risk increase; NNH, number needed to harm.

The results of RCTs in which a drug increases the risk of adverse outcomes are reported as the relative risk (RR), the relative risk increase (RRI), the absolute risk increase (ARI) and the number needed to harm (NNH) (1).

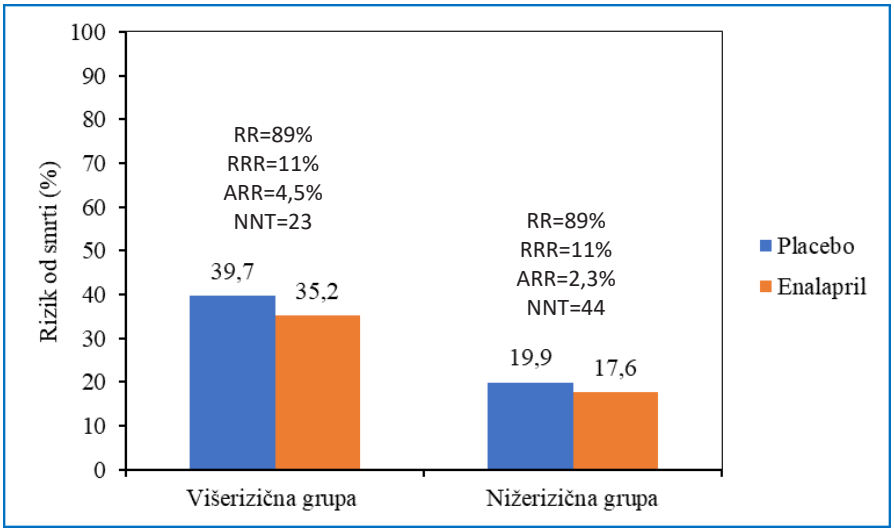
As it has been stated, the RR is the ratio of risk of an outcome in the experimental group to the risk of an outcome in the control group. In this example, the risk of sexual dysfunction in the experimental group is 0.200 or 20%, and in the control group it is 0.083 or 8.3%, so the RR=2.41 or 241% (Table 2), which indicates that venlafaxine in comparison to placebo increases the risk of sexual dysfunction in patients with depression.

The RRI is the relative increase in the risk of an adverse outcome in the experimental group compared to the risk of an adverse outcome in the control group (Table 2). In this example, the RRI=1.41, that is, 141% (Table 2), which means that venlafaxine compared to placebo increases the risk of sexual dysfunction by 1.41 times, or 141%.

The ARI is the absolute difference between the risk of an adverse outcome in the experimental group and control group (Table 2). In this example, the ARI =0.117, that is, 11.7% (Table 2); therefore, the absolute difference regarding the risk of sexual dysfunction between the group which receives venlafaxine and group which receives placebo is 11.7%.

The NNH is the number of patients who need to be treated in the experimental group in order to obtain one more adverse outcome compared to the control treatment. Mathematically, it represents the reciprocal value of ARI (Fiigure 3). The NNH in this example is 8.55 (Figure 3), which means that eight patients need to be treated with venlafaxine to get one more sexual dysfunction compared to the group which receives placebo. According to the convention, the NNH is interpreted as the whole number (the whole patient is treated, not a part of the patient) while it is rounded down to the smaller value (to avoid overestimating the safety of the therapy).

Like the NNT, the interpretation of NNH is arbitrary and depends on the clinical context. For adverse outcomes that seriously compromise the patient's health status, the acceptable NNH should be 1000; for adverse outcomes that are not



Grafik 3. Dve grupe pacijenata sa različitim rizikom od ishoda i posledično različitim vrednostima ARR i NNT.

RR, relativni rizik; RRR, relativno smanjenje rizika; ARR, apsolutno smanjenje rizika; NNT, broj pacijenata koje treba lečiti.

efekta (10,11). To za posledicu može imati precenjivanje efekata lečenja (3,6). Ilustracije radi, kada bi rizik od smrti u RCT koji se odnosio na primenu enalaprila (gore navedeni primer) bio manji za pola u obe grupe (226/1285=0,176 tj. 17,6% u grupi koja dobija enalapril i 255/1284=0,199 tj. 19,9% u grupi koja dobija placebo), RR (0,176/0,199=0,89 tj. 89%) i RRR (1–0,89=0,11 tj. 11%) bi ostali nepromenjeni, dok bi se ARR i NNT promenili i iznosili bi 2,3% i 44, redom (Grafikon 3). Promenjene vrednosti ARR i NNT bi mogle promeniti relevantnost i klinički značaj rezultata, te uticati na odluku kliničara da preporuči lečenje.

Navedeni primer ukazuje na činjenicu da se relativne mere efekta ne menjaju, dok se apsolutne mere efekta menjaju kada pacijenti imaju različite bazične rizike od ishoda. Zbog toga, apsolutne mere efekta treba koristiti kada se donose odluke o lečenju (12).

Prednosti i nedostaci RCTs

Kako je navedeno, RCTs daju najjače dokaze između uzroka i posledica pa su zlatni standard za izradu smernica za lečenje. To je njihova glavna prednost, a bazira se na postupcima randomizacije i slepe tehnike (13-14). Ta dva postupka značajno smanjuju pristrasnost istraživača koja bi mogla negativno uticati na dokaze odnosno rezultate istraživanja. Jedna od prednosti je i moguća multicentričnost RCTs što znači da mogu da se sprovode u više kliničkih centara u svetu istovremeno i tako se značajno povećava broj ispitanika te pouzdanost rezultata (13). Sa druge strane, RCTs imaju i određene nedostatake, kao npr. razlika između eksperimentalnih i uslova u kliničkoj praksi, cena, etički problemi (ponekad), te kratko trajanje (13-15). Eksperimentalni uslovi se razlikuju od onih u kliničkoj praksi zbog rigoroznih kriterijuma za uključenje i isključenje što može ograničiti ekstrapolaciju rezultata na širu populaciju. Cena je čest razlog zbog kojeg se određeni RCTs nikada ne sprovedu. Prema podacima koji su dostupni u literaturi, cena jednog RCT je između 0,2 i 611,5 miliona američkih dolara (16). Osim toga, farmaceutske kompanije često finansiraju RCTs što za posledicu može imati pristrasnost istraživača te precenjivanje efekata lečenja (13). Nadalje, dobro je poznato da etički problemi ne opravdavaju upotrebu lekova u trudnoći u svrhu unapređenja medicinskih znanja. Zbog toga za mnoge lekove nedostaju dokazi na osnovu kojih bi kliničari mogli da donesu odluke o njihovoj bezbednoj upotrebi u trudnoći. Konačno, RCTs uglavnom traju između nekoliko nedelja i nekoliko meseci pa odložene neželjene reakcije na lek mogu ostati neotkrivene (14).

Zaključak

Najjači dokaz između uzroka i posledice dobija se u RVTs, ali se ove studije retko izvode zbog svoje cene. Mali broj ispitanika u ovim studijama, može se prevazići korišćenjem meta-analize i multicentričnih studija. Relativne mere efekta koje se mogu koristiti u RCTs su relativni rizik i relativno smanjenje rizika, a od apsolutnih mera efekta apsolutno smanjenje rizika i broj pacijenata koje treba lečiti.

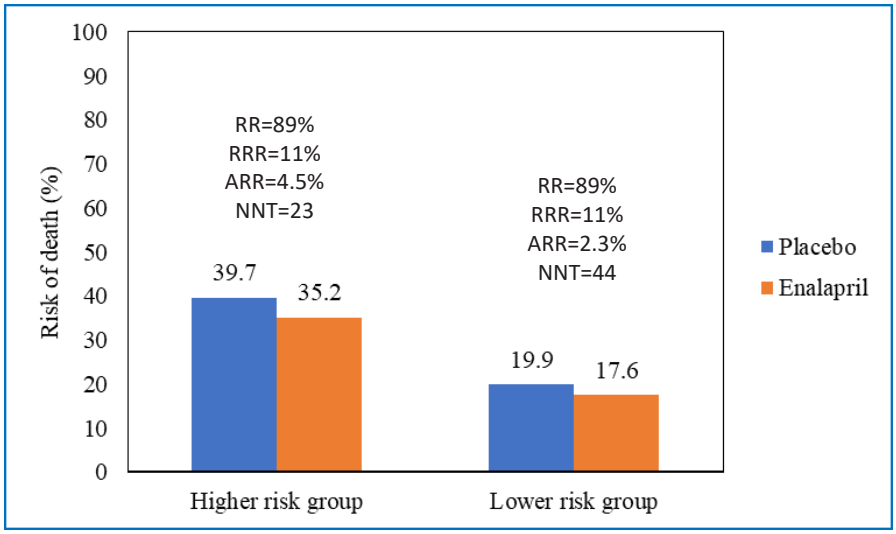


Figure 3. Two groups of patients at different risk of outcome and consequently different values of ARR and NNT.

RR, relative risk; RRR, relative risk reduction; ARR, absolute risk reduction; NNT, number needed to treat.

serious but require the termination of treatment, the acceptable NNH could be in the range between 10 and 100; finally, for mild, transient adverse outcomes, the acceptable NNH would be below 10 (1,8). Therefore, in contrast to the NNT, the desirable value of NNH should be as high as possible.

Relative versus absolute effect measures – implications for clinical practice

Although it is important that the results of RCTs are reported as relative and absolute effect measures, they are mostly reported as relative effect measures (10,11). This may result in an overestimation of treatment effects (3,6). For example, if the risk of death in the RCT related to the use of enalapril (the above mentioned example) was reduced by half in both groups (226/1285=0.176, that is 17.6% in the group which received enalapril and 255/1284=0.199, that is, 19.9% in the group which received placebo), the RR (0.176/0.199=0.89 or 89%) and RRR (1-0.89=0.11 or 11%) would remain unchanged and ARR and NNT would change and amount to 2.3% and 44%, respectively (Figure 3). The changed values of ARR and NNT would change the relevance and clinical significance of results, and therefore, influence the clinician's decision to recommend the treatment.

The above mentioned example points to the fact that relative effect measures do not change, while absolute effect measures change when patients have different main risks of an outcome. Therefore, absolute effect measures should be used when decisions about treatment are made (12).

The advantages and disadvantages of RCTs

As it has been stated, RCTs provide the strongest evidence between the cause and effect, and therefore, they are the gold standard for developing treatment guidelines. This is their main advantage, and it is based on the processes of randomization and blinding (13-14). These two procedures significantly reduce the researcher's bias, which could negatively affect the evidence or research results. One of the advantages is the possible multicenter nature of RCTs, which means that they can be simultaneously conducted in several clinical centers in the world, and thus the number of participants significantly increases and consequently the reliability of results, as well (13). On the other hand, RCTs have certain disadvantages, such as the difference between experimental conditions and conditions in clinical practice, cost, ethical problems (sometimes), and short duration (13-15). Experimental conditions are different from those in clinical practice due to rigorous inclusion and exclusion criteria that may limit the extrapolation of results to the wider population. Cost is a common reason why certain RCTs are never conducted. According to the data available in the literature, the cost of one RCT is between 0.2 and 611.5 million US dollars (16). In addition, pharmaceutical companies often finance RCTs, which can, consequently, cause researcher's bias and overestimation of treatment effects (13). Furthermore, it is well-known that ethical problems do not justify the use of drugs in pregnancy for the purpose of improving the medical knowledge.

Konflikt interesa

Autor je izjavio da nema konflikta interesa.

Reference

1. Gamulin S. Klinička istraživanja – klinička epidemiologija. Zagreb: Medicinska naklada, 2015. 144 p.
2. European Commission. Clinical trials – Regulation (EU) No 536/2014 [Internet]. [cited August 2023]. Avilable from https://www.health.ec.europa.eu
3. Etminan M, Wright JM, Carleton BC. Evidence-based pharmacotherapy: review of basic concepts and applications in clinical practice. Ann Pharmacother 1998;32(11):1193-200. doi: 10.1345/aph.17333.
4. Higgins JPT, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effects. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook
5. The SOLVD Investigators. Effect of enalapril on survival in patients with reduced ejection fraction and congestive heart failure. N Engl J Med 1991;325(5):293-302. doi: 10.1056/NEJM199108013250501.
6. Darzi AJ, Busse JW, Phillips M, Wykoff CC, Guymer RH, Thabane L, et al. Interpreting results from randomised controlled trials: What measures to focus on in clinical practice. Eye (Lond) 2023; doi: 10.1038/s41433-023-02454-7. Online ahead of print.
7. Andrade C. Understanding realtive risk, odds ratio, and related terms: as simple as it can get. J Clin Psychiatry 2015; 76(7):e857-61. doi: 10.4088/JCP.15f10150.
8. Citrome L, Ketter TA. When does a difference make a difference? Interpretation of number needed to treat, number needed to harm, and likelihood to be helped or harmed. Int J Clin Pract 2013;67(5): 407-11. doi: 10.4088/JCP.15f10150.
9. Cookson J, Gilaberte I, Desaiah D, Kajdasz DK. Treatment benefits of duloxetine in major depressive disorder as assessed by number needed to treat. Int Clin Psychopharmacol 2006;21(5):267-73. doi: 10.1097/00004850-200609000-00004.
10. Agarwal A, Johnston BC, Vernooij RWM, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, et al. Authors seldom report the most patient-important outcomes and absolute effect measures in systematic review abstracts. J Clin Epidemiol 2017;81:3-12. doi: 10.1016/j.jclinepi.2016.08.004.
11. Andersen LW. Absolute vs. Relative effects-implications for subgroup analyses. Trials 2021; 22(1):50 . doi: 10.1186/s13063-020-05005-7.
12. Sormani MP, Bruzzi P. Estimating a treatment effect: chosing between relative and absolute measures. Mult Scler 2017;23(2):197-200.
13. Štimac D, Lackovic A, Poropat G. Važnost provođenja randomiziranih kliničkih istraživanja. Medicina Fluminensis 2017; 53(4):448-53. doi: 10.21860/medflum2017_187364.
14. Saturni S, Bellini F, Braido F, Paggiaro P, Sanduzzi A, Scichilone N, et al. Randomized controlled trials and real-life studies. Approaches and methodologies: a clinical point of view. Pulm Pharmacol Ther 2014;27(2):129-38. doi: 10.1016/j.pupt.2014.01.005.
15. Hariton E, Locascio J. Randomised controlled trials – the gold standard for effectiveness research. BJOG 2018;125(13):1716. doi: 10.1111/1471-0528.15199.
16. Speich B, von Neiderhäusern B, Schur N, Hemkens LG, Fürst T, Bhatnagar N, et al. Systematic review on costs and resource use of randomized clinical trials shows a lack of transparent and comprehensive data. J Clin Epidemiol 2018:96:1-11. doi: 10.1016/j.jclinepi.2017.12.018.

Primljen: 14.10.2023. **Revizija:** 26.11.2023. **Prihvaćen:** 26.11.2023.

Therefore, for many drugs, there is a lack of evidence, based on which clinicians could make decisions about their safe use in pregnancy. Finally, RCTs commonly last between several weeks and several months, so delayed adverse drug reactions may remain undiscovered (14).

Conclusion

The strongest evidence of cause and effect is obtained in RCTs, but these studies are rarely conducted because of their cost. The small number of participants may be overcome by using meta-analysis and multicenter studies. Relative effect measures that can be used in RCTs are the relative risk and relative risk reduction, while absolute effect measures are the absolute risk reduction and the number of patients to be treated.

Competing interests

The author declared no competing interests.

References

1. Gamulin S. Klinička istraživanja – klinička epidemiologija. Zagreb: Medicinska naklada, 2015. 144 p.
2. European Commission. Clinical trials – Regulation (EU) No 536/2014 [Internet]. [cited August 2023]. Avilable from https://www.health.ec.europa.eu
3. Etminan M, Wright JM, Carleton BC. Evidence-based pharmacotherapy: review of basic concepts and applications in clinical practice. Ann Pharmacother 1998;32(11):1193-200. doi: 10.1345/aph.17333.
4. Higgins JPT, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effects. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook
5. The SOLVD Investigators. Effect of enalapril on survival in patients with reduced ejection fraction and congestive heart failure. N Engl J Med 1991;325(5):293-302. doi: 10.1056/NEJM199108013250501.
6. Darzi AJ, Busse JW, Phillips M, Wykoff CC, Guymer RH, Thabane L, et al. Interpreting results from randomised controlled trials: What measures to focus on in clinical practice. Eye (Lond) 2023; doi: 10.1038/s41433-023-02454-7. Online ahead of print.
7. Andrade C. Understanding realtive risk, odds ratio, and related terms: as simple as it can get. J Clin Psychiatry 2015; 76(7):e857-61. doi: 10.4088/JCP.15f10150.
8. Citrome L, Ketter TA. When does a difference make a difference? Interpretation of number needed to treat, number needed to harm, and likelihood to be helped or harmed. Int J Clin Pract 2013;67(5): 407-11. doi: 10.4088/JCP.15f10150.
9. Cookson J, Gilaberte I, Desaiah D, Kajdasz DK. Treatment benefits of duloxetine in major depressive disorder as assessed by number needed to treat. Int Clin Psychopharmacol 2006;21(5):267-73. doi: 10.1097/00004850-200609000-00004.
10. Agarwal A, Johnston BC, Vernooij RWM, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, et al. Authors seldom report the most patient-important outcomes and absolute effect measures in systematic review abstracts. J Clin Epidemiol 2017;81:3-12. doi: 10.1016/j.jclinepi.2016.08.004.
11. Andersen LW. Absolute vs. Relative effects-implications for subgroup analyses. Trials 2021; 22(1):50 . doi: 10.1186/s13063-020-05005-7.
12. Sormani MP, Bruzzi P. Estimating a treatment effect: chosing between relative and absolute measures. Mult Scler 2017;23(2):197-200.
13. Štimac D, Lackovic A, Poropat G. Važnost provođenja randomiziranih kliničkih istraživanja. Medicina Fluminensis 2017; 53(4):448-53. doi: 10.21860/medflum2017_187364.
14. Saturni S, Bellini F, Braido F, Paggiaro P, Sanduzzi A, Scichilone N, et al. Randomized controlled trials and real-life studies. Approaches and methodologies: a clinical point of view. Pulm Pharmacol Ther 2014;27(2):129-38. doi: 10.1016/j.pupt.2014.01.005.
15. Hariton E, Locascio J. Randomised controlled trials – the gold standard for effectiveness research. BJOG 2018;125(13):1716. doi: 10.1111/1471-0528.15199.
16. Speich B, von Neiderhäusern B, Schur N, Hemkens LG, Fürst T, Bhatnagar N, et al. Systematic review on costs and resource use of randomized clinical trials shows a lack of transparent and comprehensive data. J Clin Epidemiol 2018:96:1-11. doi: 10.1016/j.jclinepi.2017.12.018.

Received: 10/14/2023 **Revised:** 11/26/2023 **Accepted:** 11/26/2023