

Interpretacija rezultatov statističnih testov

Marija Petek Šter

Ljubljana, februar/marec 2016

Vsebina

- Interpretacija diagnostičnih testov
- Ocenjevanje učinkovitosti zdravljenja
- Statistična in klinična pomembnost
- Zavajanja v statistiki

Ocena diagnostičnih testov

Katere mere so klinično najbolj uporabne?

A. Senzitivnost (občutljivost)

Je verjetnost, da bo test pozitiven, pri osebah z boleznijo.

B. Specifičnost

Je merilo verjetnosti, da bo test negativen pri osebah brez bolezni.

C. Pozitivna napovedna vrednost — Verjetnost, da ima **nekdo** bolezen, ko ima znak

D. Negativna napovedna vrednost — Verjetnost, da je bolezen pri **osebi** odsotna, ko ni znaka

TABELA 2X2

		Bolezen	
		+	-
Test	+	a Resnično pozitiven	b Lažno pozitiven
	-	c Lažno negativen	d Resnično negativen

Senzitivnost

- **Senzitivnost** (občutljivost) je verjetnost, da bo **test pozitiven** pri osebah, ki so **resnično bolne**
- Test je senzitiv (Se), če zazna bolezen, ko ta dejansko obstaja
- **$Se = RP / (RP + LN)$** (max 1 oz. 100%)
- **Kazalec, ki ustrezno razvršča osebe z boleznijo**

Specifičnost

- **Specifičnost** je merilo verjetnosti, da bo **test negativen pri osebah, ki niso bolne**
- Specifičnost (Sp) je sposobnost testa, da **eliminira bolezen**, ko ta dejansko ne obstaja
- **$Sp = RN / (RN + LP)$** (max. 1 oz. 100%)
- **Kazalec, ki ustrezno razvršča osebe brez bolezni**

Primer

- Z UZ smo pregledali 345 bolnikov s sumom na globoko vensko trombozo. Med njimi jih je imelo 243 pozitiven D-dimer, 102 pa negativen D-dimer. Globoko vensko trombozo smo ugotovili pri 168 bolnikih s pozitivnim D-dimerom ter 2 bolnikih z negativnim D-dimerom.
- Kakšna je senzitivnost in specifičnost testa?

Računanje senzitivnosti in specifičnosti

	Prisotna bolezen GVT	Odstotna bolezen GVT
Prisotnost znaka – D – dimer pozitiven	168	75
Odsotnost znaka – D-dimer negativen	2	100

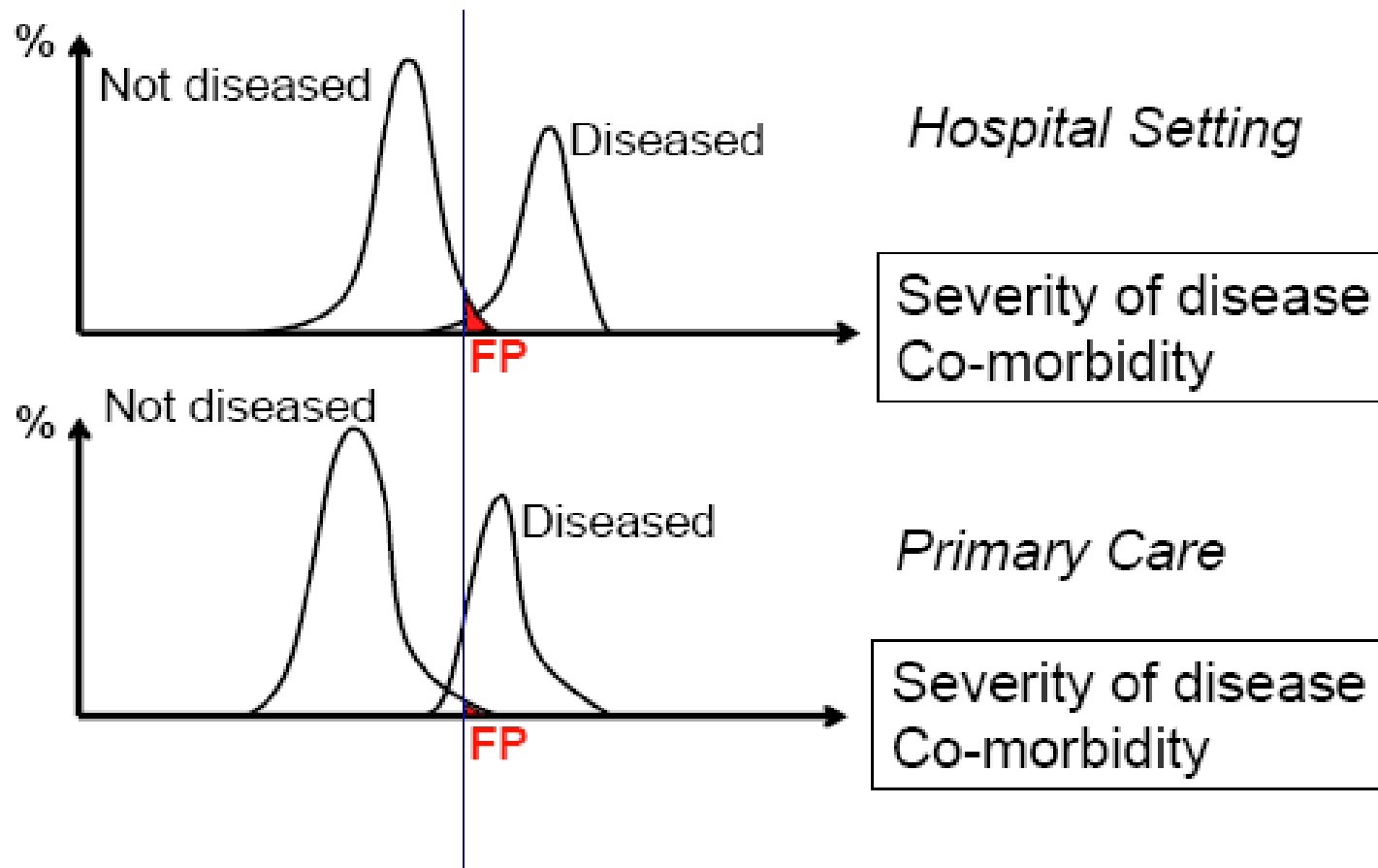
$$Se = RP / RP + LN$$

$$Se = 168 / 168 + 2 = 0,99$$

$$Sp = RN / RN + LP$$

$$Sp = 100 / 100 + 75 = 0,57$$

Effects of patient spectrum: Specificity



Senzitivnost in specifičnost

- Antagonizem med kazalcema: s povečanjem senzitivnosti zmanjšamo specifičnost in obratno
- **Kaj je pomembno:**
- **Zaznati vse osebe z boleznijo (Se)**
- **Izločiti tiste brez bolezni? (Sp)**

Napovedna vrednost – vezana na osebo

- **Odvisna od prevalece bolezni**
- **Pozitivna napovedna vrednost:**

Verjetnost, da ima nekdo bolezen, ko ima znak

$$NV+ = RP / (RP + LP)$$

- **Negativna napovedna vrednost**

Verjetnost biti brez bolezni, ko znak ni navzoč

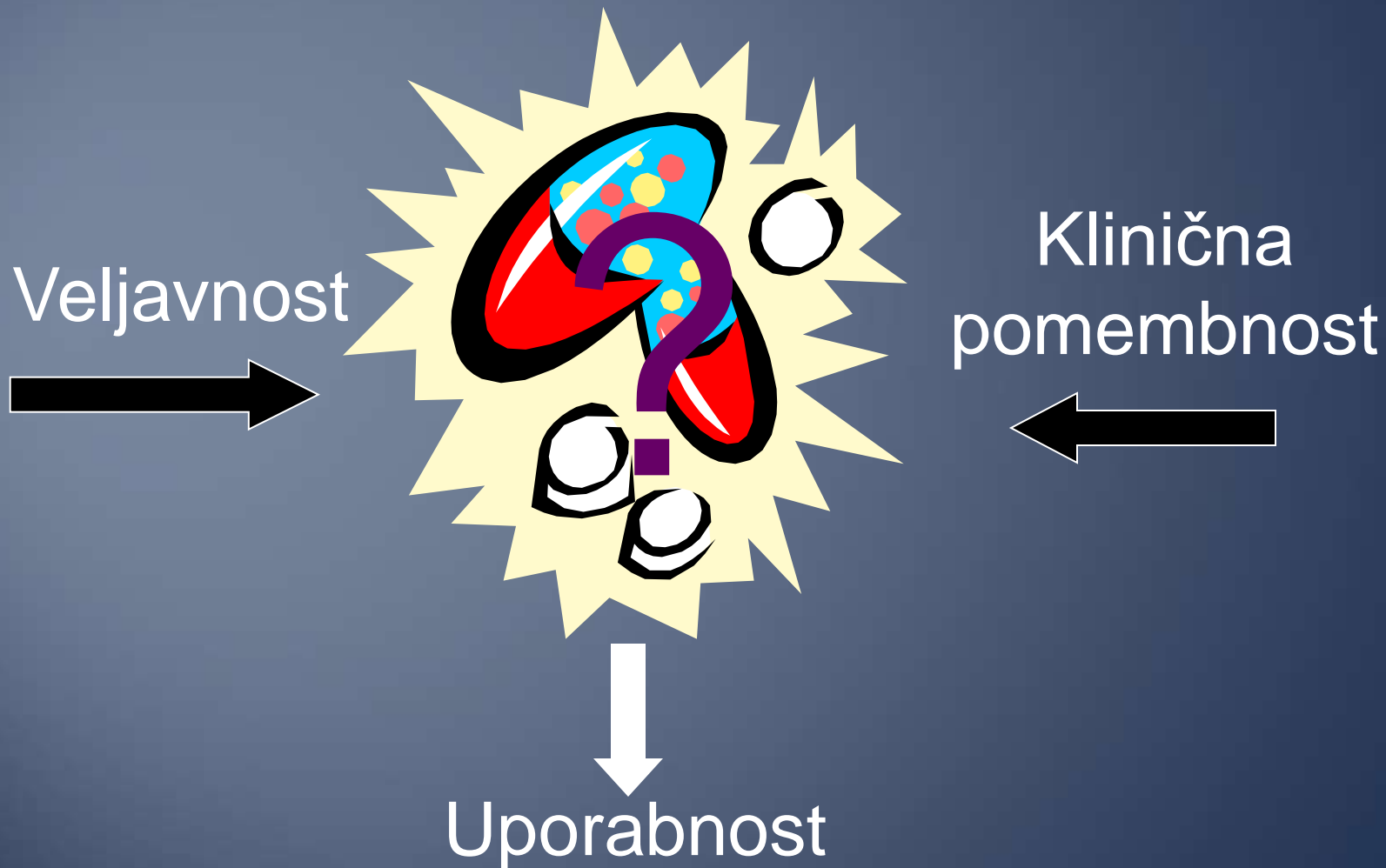
$$NV- = RN / (RN + LN)$$

Napovedna vrednost glede na prevalenco bolezni

$Se = 99\%$, $Sp = 98\%$

Prevalenca bolezni	NV+	NV-
1% ₀	4,7%	99,99%
1%	33,3%	99,98%
2%	50,0%	99,98%
3%	60,0%	99,97%
4%	67,0%	99,96%
5%	72,0%	99,95%
10%	84,0%	99,89%
20%	92,0%	99,75%
30%	95,0%	99,56%

Kritična presoja rezultatov raziskav o zdravljenju



Veljavnost (validnost)

1. Ali so raziskovalci **jasno definirali skupini bolnikov**, ki se med seboj **nista raziskovali** v ničemer drugem, kot v izpostavljenosti **opazovanemu dejavniku**?
2. Ali sta bila izpostavitveni dejavnik in izid obravnave **merjen na enak način pri obeh skupinah bolnikov**? (objektivno, slepo)?
3. Ali je bilo opazovanje **zadosti dolgo** in tekom raziskave ni prišlo do **znatnega osipa** števila sodelujočih?
4. Ali so rezultati raziskav, kjer so ugotavljali **škodljivost določenega dejavnika**, izpolnjujejo kriterije za **sklepanje na vzročnost**?

Ocena klinične pomembnosti z dvema vprašanjema:

1. Kakšen je terapevtski učinek?

RRR (zmanjšanje relativnega tveganja) = Ugotovljeno tveganje v kontrolni skupini - ugotovljeno tveganje v testni skupini / ugotovljeno tveganje v kontrolni skupini

RR (relativno tveganje) je razmerje incidenca med izpostavljenimi in neizpostavljenimi

ARR (zmanjšanje absolutnega tveganja) = Ugotovljeno tveganje v kontrolni skupini - ugotovljeno tveganje v testni skupini

NNT (število oseb, ki jih je potrebno zdraviti, da preprečimo en dogodek) = $1/ARR$

NNH = $1/ARI$ ($1/povečanje$ absolutnega tveganja)

2. Kako natančna je ocena učinka?

95% CI – območje, za katerega lahko s 95% verjetnostjo trdimo, da vsebuje pravo vrednost za populacijo

Zmanjšanje relativnega tveganja RRR

- **RRR (zmanjšanje relativnega tveganja)=**
Ugotovljeno tveganje v kontrolni skupini-
ugotovljeno tveganje v testni skupini/ ugotovljeno
tveganje v kontrolni skupini

Primer: V vsaki od skupin A in B smo imeli 4500 bolnikov z Barrettovim požiralnikom. V skupino A, kjer smo klasični terapiji dodali zdravilo X, je karcinom požiralnika razvilo 18 bolnikov, kontrolni skupini B pa 34 bolnikov. Kolikšno je zmanjšanje relativnega tveganja za razvoj karcinoma požiralnika v skupini A?

$$\text{RRR} = (34/4500 - 18/4500)/34/4500 = 0,47 \times 100 = (47\%)$$

Relativno tveganje (RR)

- Je razmerje incidence med izpostavljenimi in neizpostavljenimi (npr. tveganje za opazovani dogodek v testni skupini/tveganje za opazovani dogodek v kontrolni skupini)
- $RR = R_t / R_c$

Primer

- Kolikšno je relativnega tveganja za pojav karcinoma požiralnika pri bolnikih z Barretovim požiralnikom v skupini A (ki so ob klasični terapiji prejeli še zdravilo X?
- $RR = 18 \times 4466 / 34 \times 4482 = 0,527$
- Verjetnost, za karcinom požiralnika pri bolnikih na standardni terapiji (OR) je $(1/RR)$ za **1,9X** večja, kot pri bolnikih, zdravljenih z zdravilom X

Absolutno tveganje

- Predstavlja razliko v pojavnosti opazovanega dogodka med dvema skupinama (npr. izpostavljeni : ne-izpostavljeni)
- Primer: V vsaki od skupin A in B smo imeli 4500 bolnikov z Barretovim požiralnikom. V skupino A, kjer smo klasični terapiji dodali zdravilo X, je karcinom požiralnika razvilo 18 bolnikov, kontrolni skupini B pa 34 bolnikov. Kolikšno je zmanjšanje absolutnega tveganja za razvoj karcinoma požiralnika v skupini A?
- **Absolutno tveganje:** $34/4500 - 18/4500 = 0,0076 - 0,004 = 0,0036$

NNT (number needed to treat)

- **NNT** (število oseb, ki jih moramo zdraviti, da preprečimo en dogodek) = $1/ARR$ (kontrolno-eksperimentalno)
- **NNH**: $1/ARI$
- V vsaki od skupin A in B smo imeli 4500 bolnikov z Barrettovim požiralnikom. V skupino A, kjer smo klasični terapiji dodali zdravilo X, je karcinom požiralnika razvilo 18 bolnikov, kontrolni skupini B pa 34 bolnikov. Kolikšno število bolnikov moramo zdraviti, da preprečimo en karcinom požiralnika?
- **NNT** = $1/ARR = 1/0,0036 = 228$

NNH

- NNH: $1 / \text{ARI}$
- Raziskava o zdravilu X za preprečevanje karcinoma požiralnika je pokazala, da 5% bolnikov, ki je prejelo zdravilo X razvije depresijo, medtem, ko se je depresija razvila pri 3% bolnikov, ki so prejeli placebo. Kakšen je NNH?
- $\text{NNH} = 1 / \text{ARI} = 1 / (0.05 - 0.03) = 50$

PRIMER IZ PRAKSE

	terapija	placebo	Razlika v številu dogodkov	Zmanjšan je absolutnega tveganja	Zmanjšan je relativnega tveganja	NNT
Primerov	2365	2371				
CVI	94	149	55	0,022	35,5%	45
TIA	62	82	20	0,009	25,1%	120
AMI	50	74	24	0,010	32,3%	100
smrt	213	242	29	0,012	12,0%	84

SHEP Cooperative Research Group. Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the Systolic Hypertension in the Elderly Program (SHEP). JAMA 1991 Jun 26;265(24):3255-3264

RT in NNT – vpliv ogroženosti za

bolezen na NNT: Koliko oseb moramo zdraviti s statinom, da preprečimo en srčno žilni dogodek: primerjava bolnikov s koronarno boleznijo in povišanim holesterolom in zdravih posameznikov z normalnim holesterolom

Dejavnik	RT (učinek zdravljenja s statini)	NNT
Povišan holesterol pri bolnikih s koronarno boleznijo	0,66	63
Normalen holesterol in odsotnost koronarne bolezni	0.63	256

Interval zaupanja (95% CI)

● Območje v katerem se z 95% stopnjo zaupanja nahaja prava vrednost učinka zdravljenja:

Npr: Zmanjšanje tveganja za možgansko kap (RRR) pri bolnikih z atrijsko fibrilacijo, ki jemljejo warfarin je 79%:

- 95% CI je od 52% do 90%

Interval zaupanja (CI)

- Interval zaupanja je odvisen od:
 - **velikosti vzorca:** večji je vzorec, ožji je interval zaupanja ob enaki stopnji tveganja
 - **stopnje zaupanja (p vrednosti)** – večja je stopnja zaupanja (manjši p) - širši je interval zaupanja
- **Ožji je interval zaupanja, večja je natančnost ocene**

Klinična uporabnost (aplikabilnost)

1. Ali so naši bolniki drugačni od tistih, vključenih v raziskavo in je zato uporabnost rezultatov raziskave omejena?
2. Ali je način zdravljenja praktično izvedljiv?
3. Kakšne so koristi in kakšno je morebitno tveganje?
4. Kakšne so vrednote ter prepričanja bolnikov glede zdravljenja in pričakovanih izidov?

Nekaj najpogostejših trikov, s katerimi si pomagajo predstavniki farmacevtske industrije-1

- Izbira nadomestnega (surogatnega izida), ki nima neposredne povezave z dejanskim izidom za katerega je zdravilo promovirano
- Izbor populacije, kliničnih parametrov ter trajanja raziskave, ki omogočata kar najboljši možen rezultat za zdravilo
- Primerjava produkta s placebom ali primerljivim zdravilom, vendar v nižji (subterapevtski) dozi
- Ne omenijo pomembnih stranskih učinkov zdravila, ki so bili zaznani, vendar pogosto neobjavljen

Nekaj najpogostejših trikov, s katerimi si pomagajo predstavniki farmacevtske industrije-2

- Grafično prikažejo učinek zdravila na način, da podkrepijo svoje sporočilo: ne označilo skal, ne povedo, ali gre za linearno oz. logaritmično skalo
- Svoje zdravilo navedejo kot boljše, čeprav je med primerljivima zdraviloma razlika ni statistično pomembna, razliko predstavijo v grafični obliki
- Omenijo mnenja lokalnih “mnenjskih voditeljev”, ki podpirajo uporabo njihovega zdravila (obrnjena piramida hierarhije dokazov)
- Zdravilo, ki je v resnici dražje od konkurence predstavijo v luči ekonomske učinkovitosti

Zaključek

- Razumevanje osnov statističnega sklepanja je nujno potrebno, da lahko pravilno interpretiramo pomen rezultatov raziskav
- Statistična pomemben rezultat ni vedno klinično pomemben
- Bodite kritični, ko vam nekdo želi s pomočjo uporabe statističnih podatkov in prikazov predstaviti svoj produkt v najboljši luči