

Statistike dijagnostičkih testova

Damjan Krstajić

(objavljeno 10. aprila 2020. u Politikinom Kulturnom dodatku)

Sa pojavom dijagnostičkih testova na virus korona, u javnosti su se postavila pitanja u vezi njihove pouzdanosti. Nemam nikakve podatke u vezi tih testova, ali na osnovu dugogodišnjeg bavljenja medicinskom statistikom mogu da primetim da problematika nije uopšte jednostavna. Ovde mi nije cilj da dovodim u pitanje njihovu validnost. Naprotiv, namera mi je da ukažem na šta sve naučnici paze pre nego što puste medicinski dijagnostički test u promet.

Zamislimo da imamo test na oboljenje X koji daje binarne rezultate - pacijent je pozitivan (ima oboljenje X) ili negativan (bez oboljenja X). Neka test daje pozitivne rezultate za 95 od 100 pacijenata sa oboljenjem X. Tehnički rečeno, u tom slučaju *senzitivnost testa* je 95%. Neka takođe test daje negativne rezultate za 90 od 100 pacijenata bez oboljenja X i tada kažemo da je *specifičnost testa* 90%. E sad dolazimo do prakse gde za našeg pacijenta dobijemo pozitivan nalaz. Kolika je verovatnoća da je on oboleo od X? Ako kažete 95%, onda je vaš odgovor isti kao i moj kad sam se prvi put suočio sa ovim problem, a on je, nažalost, pogrešan!

Prvo da objasnimo kako izgleda tabela za binarne predikcije (slika 1). Među pacijentima koji su dobili pozitivan rezultat razlikujemo one koji stvarno imaju oboljenje X (*tačno pozitivni* TP) i one koji nemaju oboljenje X (*lažno pozitivni* LP). Slično i među pacijentima koji su dobili negativan rezultat razlikujemo lažno negativne (LN) i tačno negativne (TN).

	ima oboljenje X	nema oboljenje X
pozitivan nalaz testa	TP	LP
negativan nalaz testa	LN	TN

Slika 1

Zanima nas verovatnoća da je pacijent koji je dobio pozitivan rezultat na testiranju stvarno oboleo od X. Ta vrednost predstavlja udeo tačno-pozitivnih među svim pacijentima sa pozitivnim nalazom. Ova verovatnoća se naziva *pozitivna prediktivna vrednost*. Slično se definiše i *negativna*

prediktivna vrednost, kao udeo tačno-negativnih među svim pacijentima sa negativnim nalazom.

$$\text{senzitivnost} = TP / (TP + LN)$$

$$\text{specifičnost} = TN / (TN + LP)$$

$$\text{pozitivna prediktivna vrednost} = TP / (TP + LP)$$

$$\text{negativna prediktivna vrednost} = TN / (TN + LN)$$

Test A

	ima oboljenje X	nema oboljenje X
pozitivan nalaz testa	380	1000
negativan nalaz testa	20	9000

Test B

	ima oboljenje X	nema oboljenje X
pozitivan nalaz testa	950	150
negativan nalaz testa	50	1350

Slika 2

Na slici 2 su prikazani rezultati dva dijagnostička testa izvedena za isto oboljenje X. Test A je obavljen na uzorku od 10400 pacijenata, a test B na 2500. Oba dijagnostička testa imaju istu senzitivnost (95%) i specifičnost (90%). Provera. Kod testa A imamo TP=380 i LN=20, a kod testa B TP=950 i LN =50 i senzitivnost jeste 95% (prema formuli TP/(TP+LN)). Takođe, kod testa A imamo TN=9000 i LP=1000, a kod testa B TN=1350 i LP=150 i specifičnost jeste 90% (prema formuli TN/(TN+LP)).

U testu A imamo uzorak od 10400 pacijenata od kojih je 1380 sa pozitivnim nalazima. Među pacijentima sa pozitivnim nalazima imamo 27,5% (380/1380) koji su oboleli. Dakle, samo na osnovu rezultata testa A mogli bismo reći da je 27,5% verovatnoća da je pozitivno testiran pacijent oboleo od X. Slično kod testa B to je 86,4% (950/1100). Koji test daje veću pozitivnu prediktivnu vrednost? Nijedan. Oba daju istu pozitivnu prediktivnu vrednost. Kako?

Polovinom 70-ih godina prošlog veka počeli su da se pojavljuju naučni radovi u medicinskim časopisima koji ukazuju na činjenicu da se prediktivne vrednosti, pozitivna i negativna, ne smeju zaključiti na osnovu uzorka, jer one zavise od *prevalense*, odnosno od procenta obolelih u populaciji.

U prošlom članku sam objasnio neophodnost provere da li je uzorak reprezentativan da bismo bili sigurni da možemo nešto da zaključimo na osnovu njega. U ovom slučaju, da nam tabela bude reprezentativna, odnosno da uzorak možemo da smatramo reprezentativnim, potrebno je da procenat obolelih od X u celokupnoj populaciji bude isti kao u našem uzorku. Procenat obolelih od X u tabeli A je 3,8% (400/10400) dok je u tabeli B 40% (1000/2500). Postavlja se pitanje kako to da im je pozitivna prediktivna vrednost ista?

Pokazalo se da se na osnovu specifičnosti, senzitivnosti i prevalense (u populaciji, a ne u uzorku!) mogu izračunati prediktivne vrednosti.

$$\text{pozitivna prediktivnavrednost} = \frac{\text{senzitivnost} \cdot \text{prevalensa}}{(\text{senzitivnost} \cdot \text{prevalensa}) + (1 - \text{specifičnost}) \cdot (1 - \text{prevalensa})}$$

$$\text{negativna prediktivnavrednost} = \frac{\text{specifičnost} \cdot (1 - \text{prevalensa})}{(1 - \text{senzitivnost}) \cdot \text{prevalensa} + \text{specifičnost} \cdot (1 - \text{prevalensa})}$$

Dakle, pošto je u pitanju isto oboljenje (prevalensa je ista u oba testa) i kako imaju istu specifičnosti i senzitivnost, onda su im identične i negativna i pozitivna prediktivna vrednost. Test A i Test B za oboljenje X daju identične statistike dijagnostike, ali ne možemo samo na osnovu njih da zaključimo kolika je verovatnoća da je pacijent sa pozitivnim nalazom stvarno oboleo. Za to nam je potrebna informacija o prevalensi (procentu obolelih u populaciji).

Britanski statističari Daglas Altman (Douglas G. Altman) i Martin Bland (J Martin Bland) su serijom mini članaka o dijagnostičkim testovima približili ovu problematiku lekarima. Članci su objavljeni 1994. godine u časopisu BMJ (British Medical Journal) i dostupni su besplatno preko interneta.

U svojoj praksi sam prihvatio upozorenja u vezi prediktivnih vrednosti, ali dugi niz godina mi nije bilo jasno zašto su specifičnost i senzitivnost pouzdane statistike na osnovu uzorka, a prediktivne vrednosti nisu. Sve dok nisam pročitao uvodni članak „*Standing Statistics Right Side Up*“ Franka Davidofa (Frank Davidoff) objavljenog 1999. godine u časopisu *Annals of Internal Medicine*, čiji je on tada bio urednik. Frank Davidof je objasnio da su specifičnosti i senzitivnost deduktivne mere, a pozitivne i negativne

prediktivne vrednosti induktivne mere. Omakla mi se tanana razlika između dedukcije i indukcije!

Pre nego što se bilo koji dijagnostički test pusti u opticaj, izračunaju se sve gore navedene statistike i izabere se onaj test koji ima kliničku vrednost za dato oboljenje. Ako se ne zna prevalensa, onda se simuliraju razne vrednosti. Ne postoji fiksno pravilo za odabir dijagnostičkih testova, jer je za neka oboljenja veći problem postojanje velikog broja lažno pozitivnih rezultata, a za druga pak veći broj lažno negativnih.

Da li sad razumete kako su javna i neargumentovana licitiranja sa pouzdanošću medicinskih dijagnostičkih testova ustvari igranje sa neznanjem? Problematika je kompleksna i ima još drugih statistika koja se računaju u praksi. Trebalo mi je dosta vremena i vežbe na papiru da razumem sve gore navedene finese. Iz mog iskustva, cilj naučnika koji rade na medicinskim dijagnostičkim testovima jeste da naprave test koji će najviše pomoći lekarima. Nažalost, nema idealnih rešenja.

Reference koje podržavaju činjenice spomenute u članku

1. Dijagnostički testovi u medicini

https://en.wikipedia.org/wiki/Medical_test

2. Senzitivnost i specifičnost

https://en.wikipedia.org/wiki/Sensitivity_and_specificity

3. Pozitivna i negativna prediktivna vrednost

https://en.wikipedia.org/wiki/Positive_and_negative_predictive_values

4. Galen i Gambino su 1975. godine objavili u knjizi problematiku u vezi prediktivnih vrednosti, ali ne znam da li su oni prvi.

Galen, R. S., and S. R. Gambino. "The predictive value and efficiency of medical diagnoses." *J. Willey & Sons, New York* (1975).

5. Mini članci od Dagleasa Altmana i Martina Blanda

Diagnostic tests. 1: Sensitivity and specificity.

<https://www.bmj.com/content/bmj/308/6943/1552.full.pdf>

Diagnostic tests. 2: predictive values

<https://www.bmj.com/content/bmj/309/6947/102.1.full.pdf>

6. „Standing Statistics Right Side Up“ Franka Davidofa

<https://annals.org/aim/article-abstract/712766/standing-statistics-right-side-up>