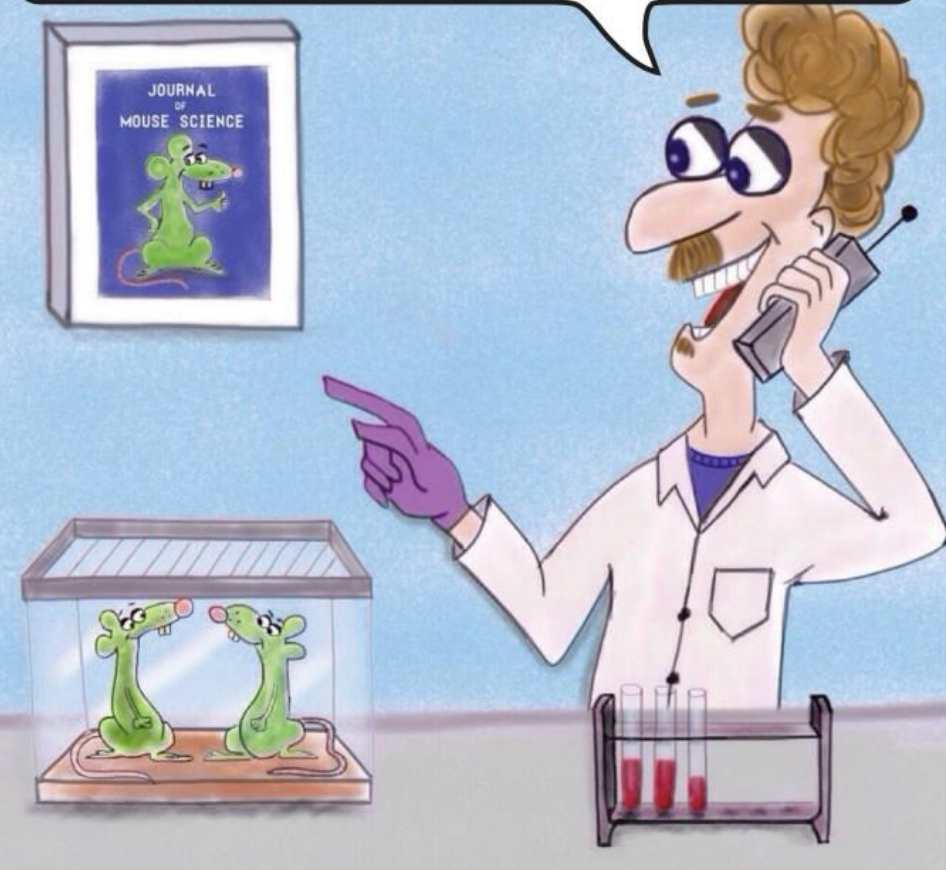


I followed the methods in your paper to the last detail... This experiment doesn't work!



Well.... When did you do the experiment? On a Friday? Oooh... that's why! It never works on Fridays!

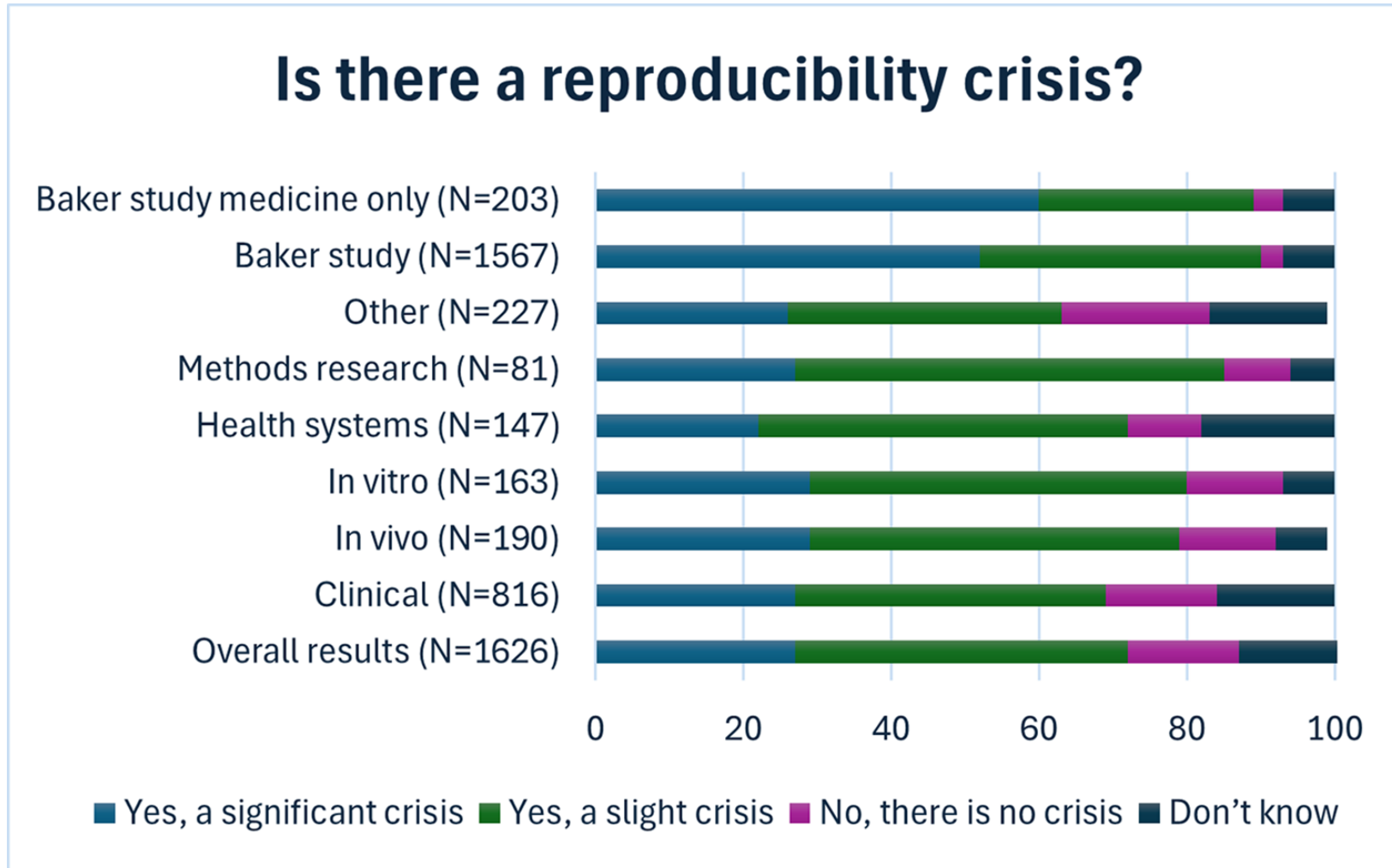


Dobra istraživačka praksa

Izv.prof. Rosa Karlič
Predavanje 10, MZIRuB 2024/2025

29.01.2025.

Kriza reproducibilnosti znanosti



Cobey et al. (2024). *PLoS biology*, <https://doi.org/10.1371/journal.pbio.3002870>

Kriza reproducibilnosti znanosti

	N(%)					
	Always contributes	Very often Contributes	Sometimes Contributes	Does not Contribute	Unsure	Missing data
Selective reporting of the published literature	131 (8)	638 (40)	714 (45)	43 (3)	73 (5)	31
Selective publication of entire studies	182 (11)	698 (44)	577 (36)	71 (4)	71 (4)	31
Pressure to publish	300 (19)	693 (43)	473 (30)	75 (5)	57 (4)	32
Low statistical power	185 (12)	706 (44)	579 (36)	76 (5)	48 (3)	36
Poor statistical analysis	197 (12)	615 (38)	649 (41)	99 (6)	44 (3)	26
Not enough internal replication (E.g., by the original lab/authors)	132 (8)	539 (34)	697 (44)	93 (6)	142 (9)	27
Insufficient study oversight	86 (5)	376 (24)	799 (50)	194 (12)	143 (9)	32
Lack of training in reproducibility	153 (10)	522 (33)	622 (39)	168 (11)	135 (8)	30
Failure to make materials openly available	141 (9)	449 (28)	722 (45)	191 (12)	99 (6)	28
Failure to make original study data openly available	137 (9)	476 (30)	685 (43)	205 (13)	94 (6)	33
Poor study design	208 (13)	584 (36)	678 (42)	96 (6)	38 (2)	26
Fraud	185 (12)	120 (8)	624 (40)	320 (20)	330 (21)	51
Poor quality peer review	140 (9)	437 (27)	755 (47)	192 (13)	72 (5)	34
Problems in the design of replication studies	103 (6)	406 (25)	809 (51)	162 (10)	123 (8)	27
Technical expertise required for replication	96 (6)	429 (27)	743 (46)	190 (12)	144 (9)	28
Variability of standard reagents	82 (5)	288 (18)	617 (39)	229 (14)	380 (24)	34
Bad luck	23 (1)	70 (4)	461 (29)	568 (36)	466 (29)	42

<https://doi.org/10.1371/journal.pbio.3002870.t002>

Kako osigurati da je istraživanje reproducibilno

- Potrebno je osigurati reproducibilnost istraživanja na različitim razinama
 - Planiranje istraživanja
 - Provedba istraživanja
 - Predstavljanje rezultata istraživanja

PLOS COMPUTATIONAL BIOLOGY

EDITORIAL

Ten simple rules for good research practice

Simon Schwab^{1,2*}, Perrine Janiaud³, Michael Dayan⁴, Valentin Amrhein⁵,
Radoslaw Panczak⁶, Patricia M. Palagi⁷, Lars G. Hemkens^{3,8,9}, Meike Ramon¹⁰,
Nicolas Rothen¹¹, Stephen Senn¹², Eva Furrer^{1,2}, Leonhard Held^{1,2}

1. Specificirajte istraživačko pitanje (planiranje istraživanja)

- Specifično i jasno postavljeno istraživačko pitanje.
- Treba sustavno procijeniti prethodna istraživanja kako bi se identificirale „praznina” u trenutnom znanju (pregledni radovi, sistematske recenzije, komentari)
- Kod dobro postavljenog istraživačkog pitanja, lakše je odrediti sljedeće korake u istraživanju (točna definiciju populacije, ishod, podaci koje treba prikupiti i potrebna veličinu uzorka)
- Potvrдна (konfirmatorna) istraživanja moraju imati dobro definiranu hipotezu
- Eksploratorna istraživanja ne zahtijevaju dobro definiranu hipotezu od samog početka (omics i multi-omics studije)

2. Napišite protokol istraživanja (planiranje istraživanja)

- Treba sadržavati barem:
 - istraživačko pitanje i hipotezu
 - opis istraživane populacije
 - veličinu uzorka
 - kriterije za uključivanje/isključivanje iz studije
 - eksperimentalni dizajn
 - plan transformacije i procesiranja podataka
 - planirane statističke analize
- Za kliničke studije protokole je potrebno i predregistrirati
- Moguće je prilagoditi istraživanje nakon predregistracije (npr. pojavile su se nove i napradnije eksperimentalne metode)
- Projektne prijave očekuju da detaljno opišete metode

3. Odredite potrebnu veličinu uzorka (planiranje istraživanja)

- Mala veličina uzorka smanjuje statističku snagu istraživanja (može dovesti i do lažno pozitivnih i lažno negativnih rezultata)
- Može doći do preoptimistične procjene veličine učinka (jer su samo veliki učinci statistički značajni zbog male veličine uzorka)
- Veličinu uzorka treba procijeniti prije početka istraživanja

4. Osmislite plan upravljanja podacima (planiranje istraživanja)

- Dokument koji sadržava informacije o tome kako će podaci biti organizirani, pohranjeni, analizirani i zaštićeni tijekom i nakon završetka istraživanja
- Načela podataka koji se mogu pronaći, dostupni, interoperabilni i višekratno (FAIR: Findable, Accessible, Interoperable, and Reusable) promiču maksimalnu upotrebu podataka i omogućuju računalima pristup i ponovnu upotrebu podataka uz minimalnu ljudsku intervenciju
- FAIR načela zahtijevaju da se podaci zadržavaju, čuvaju i dijele po mogućnosti s nepromjenjivim jedinstvenim identifikatorom i jasnom licencom za korištenje.
- Odgovarajući metapodaci pomoći će drugim istraživačima (ili računalima) da otkriju, obrade i razumiju podatke.

5. Izbjegnite pristranost (planiranje istraživanja)

Neki učestali tipovi pristranosti:

- Pristranost odabira – sustavna razlika u dodjeljivanju ispitanika eksperimentalnim skupinama
- Konfirmacijska pristranost - sklonost da primijetimo, usmjerimo pažnju i pridamo veću vjerodostojnost dokazima koji se uklapaju u naša postojeća uvjerenja
- Informacijska pristranost - sustavne razlike u prikupljanju, sjećanju, bilježenju ili rukovanje informacijama
- Pristranost u objavljivanju – sustavno objavljivanje samo jedne vrste rezultata (npr. Statistički značajnih)
- Zbunjujuća pristranost – umjetna povezanost zavisne i nezavisne varijable jer postoji treća varijabla koja je povezana i sa zavisnom i sa nezavisnom varijablom

6. Izbjegavajte loše istraživačke prakse (provedba istraživanja)

- U loše istraživačke prakse ubrajaju se **niska statistička snaga istraživanja** i **pseudoreplikacija**
- Mogu se izbjeći dobrim planiranjem i dizajnom istraživanja (izračun potrebnog broja uzoraka i upotreba prikladne statističke metodologije)
- Primjerena statistička snaga potrebna je kako bi se osigurala reproducibilnost istraživanja (mala snaga može uzrokovati lažno pozitivne i lažno negativne rezultate)
- Postoje i dodatne problematične istraživačke prakse:



<https://www.nature.com/nature-index/news/the-seven-deadly-sins-of-research>

p-hakiranje (*p-hacking*)

- Manipuliranje analizom podataka ili selektivnim izvješćivanjem o rezultatima kako bi se postigli statistički značajni rezultati (obično $p < 0.05$), čak i kada ne postoji pravi učinak
- Najčešće se provodi opetovanim testiranjem različitih statističkih modela, varijabli ili podskupova podataka dok se ne dobije željeni rezultat, *post hoc* isključivanje stršćih vrijednosti (outlier) ili observacija
- Ako studije nisu pre-registrirane, neke oblike p-hakiranja vrlo je teško otkriti
- Prilikom recenzije treba procijeniti jesu li analize ispravno odabrane, je li isti plan analize upotrijebljen u prijašnjim publikacijama, jesu li istraživači prikupili veliki broj mjerenja i jesu li ih sve analizirali



Mijenjanje hipoteze nakon dobivanja rezultata (*HARK-ing, Hypothesizing After Results Are Known*)

- Formuliranje ili modificiranje hipoteze nakon što su poznati rezultati analize
- Smanjuje reproducibilnost i povećava vjerojatnost lažno pozitivnih rezultata
- Može se izbjeći:
 - Predregistriranjem studije
 - Razlikovanjem konfirmatornih i eksplorativnih studija
 - Prezentacijom svi testiranih hipoteza (ne samo značajnih)

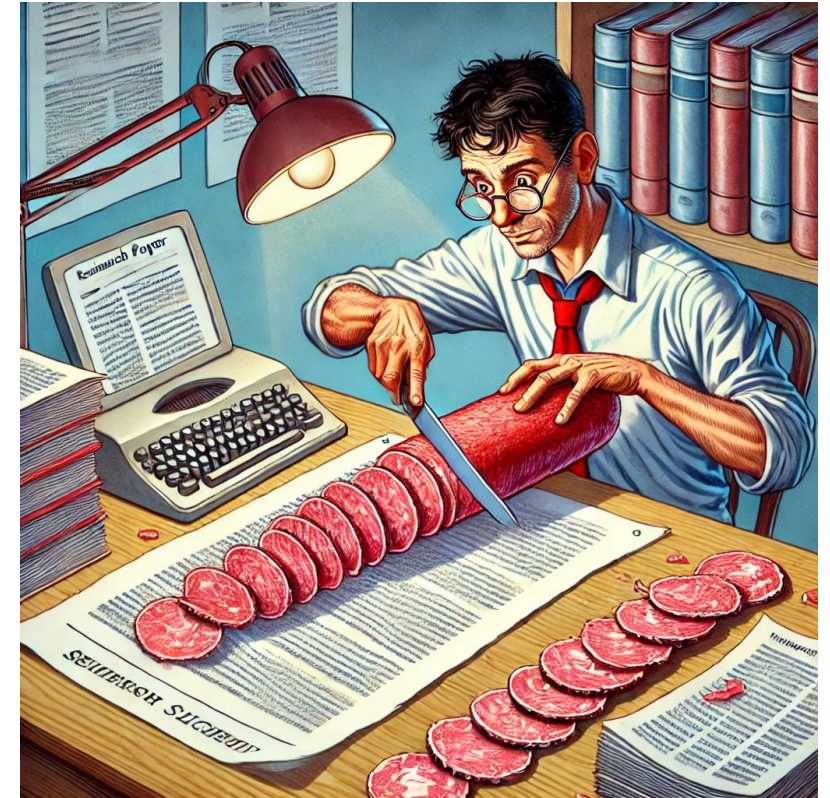


Dijeljenje jednog seta podataka na analize opisane u različitim publikacijama (*salami slicing*)

- Dijeljenje jednog, opsežnog skupa podataka ili studije u više manjih publikacija, često bez dodavanja bitnih novih uvida.
- Umjetno povećava broj publikacija, dovodi do redundancije radova i povećava rizik lažno pozitivnih rezultata (zbog opetovanog testiranja)

Primjeri:

- Dijeljenje velikog skupa podataka u više radova – Umjesto analize cijelog skupa podataka u jednoj studiji, objavljuju se zasebni radovi o manjim podskupovima.
- Ponovna analiza istih podataka s malim varijacijama – Korištenje različitih statističkih testova ili varijabli ishoda iz istog skupa podataka za generiranje više publikacija.
- Odvojeno objavljivanje privremenih rezultata – Izvještavanje o ranim, nepotpunim nalazima kao o samostalnim studijama umjesto njihovog integriranja u konačnu, sveobuhvatnu analizu.



salami slicing - primjer

Istraživač provodi veliku proteomičku studiju kako bi istražio kako novi lijek protiv raka (lijek Y) utječe na ekspresiju proteina u stanicama tumora. U eksperimentu je proizveden jedan skup podataka koji sadrži razine ekspresije 5000 proteina prije i poslije tretmana.

- Dobar pristup: Napisati jedan sveobuhvatan rad u kojem su analizirani svi relevantni putevi proteina na koje utječe lijek Y, raspravljajući i o reguliranim i smanjenim proteinima u različitim biološkim kontekstima.
- Loš pristup (*salami slicing*): Umjesto objavljivanja jedne smislene studije, istraživač dijeli skup podataka u više manjih radova bez dodavanja značajnih novih uvida:

Rad 1: “Lijek Y mijenja ekspresiju metaboličkih proteina u stanicama raka” – Izvještava o samo 50 proteina povezanih s metabolizmom.

Rad 2: “Učinak lijeka Y na upalne puteve kod raka” – Izvještava o samo 40 proteina povezanih s upalom iz istog skupa podataka.

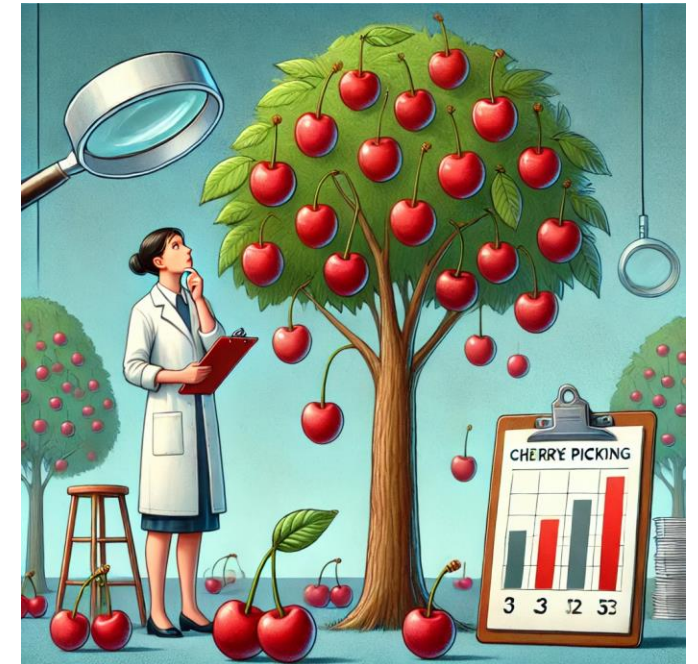
Rad 3: “Utjecaj lijeka Y na proteine za popravak DNA” – koristi 30 proteina povezanih s popravkom DNA, koji su već dio izvornog skupa podataka.

Rad 4: “Proteini oksidativnog stresa reagiraju na liječenje lijekom Y” – Još jedna selektivna analiza iz istog skupa podataka.

- Svaki od ovih radova predstavljen je kao neovisna studija, unatoč tome što se svi oslanjaju na isti skup podataka.

Selektivan odabir značajnih rezultata (*cherry picking*)

- Selektivni odabir podataka, rezultata ili statističkih analiza koje podupiru željeni zaključak dok se ignoriraju ili izostavljaju podaci koji su u suprotnosti s njim.
- Posljedica: pristrano tumačenje i pogrešni zaključci
- Problematične istraživačke prakse:
 - Odabir samo povoljnih podatkovnih točaka
 - Selektivno ignoriranje stršećih vrijednosti (outliera)
 - Korištenje ograničenog vremenskog okvira.
 - Isticanje samo statistički značajnih rezultata



Cherry picking - primjer

- Istraživač proučava učinak novog lijeka (lijek X) na ekspresiju gena u stanicama raka. Cilj je utvrditi da li lijek X inhibira gene povezane s upalom. Istraživač mjeri ekspresiju 10 upalnih gena nakon tretmana.

Umjesto da prijavi sve rezultate, istraživač:

- Uključuje samo četiri gena (A, B, H, J) koji su pokazali značajno smanjenje, podupirući hipotezu da lijek X smanjuje upalu.
- Izostavlja gene C, D, F, G, I, koji nisu pokazali značajne promjene ili čak povećanje.
- Ignorira Gen E, koji se značajno povećao, što je u suprotnosti s tvrdnjom da lijek X smanjuje upalu.

Gen	Promjena u ekspresiji (Lijek X vs. kontrola)	p-vrijednost
Gen A	↓ (značajan)	0.02
Gen B	↓ (značajan)	0.03
Gen C	↓ (nije značajan)	0.08
Gen D	↑ (nije značajan)	0.07
Gen E	↑ (značajan)	0.01
Gen F	↓ (nije značajan)	0.09
Gen G	↑ (nije značajan)	0.12
Gen H	↓ (značajan)	0.04
Gen I	↑ (nije značajan)	0.10
Gen J	↓ (značajan)	0.01

Cirkularno zaključivanje (*double dipping*)

- Retrospektivni odabir značajki za karakterizaciju zavisnih varijabli, što rezultira krivim rezultatima statističke analize
- Često koristi podjelu (npr. podgrupiranje, grupiranje) ili smanjenje (npr. definiranje interesne regije, uklanjanje 'outliera') cijelog skupa podataka pomoću kriterija odabira koji je retrospektivan i inherentno relevantan za ishod statističke analize



Double dipping - primjer

Pokušavate trenirati model višestruke linearne regresije za predviđanje razine ekspresije određenog gena na temelju različitih koncentracija biomarkera (npr. proteina ili metabolita) u krvi. Imate skup podataka gdje ste za svaki uzorak izmjerili ekspresije gena (ciljna varijabla) i mjerenja različitih biomarkera (nezavisne varijable). Provodite analizu u sljedećim koracima:

1: Podijelite podatke dijelite na skup za treniranje (80% uzoraka) i testni skup (20% uzoraka).*

2: Provodite odabir značajki kako biste odabrali najrelevantnije biomarkere (prediktore) za model. Međutim, slučajno koristite i trening skup i testni skup za odabir značajki. Na primjer, identificirate pojedinačne biomarkere koji su u korelaciji s ekspresijom gena, koristeći cijeli skup podataka.

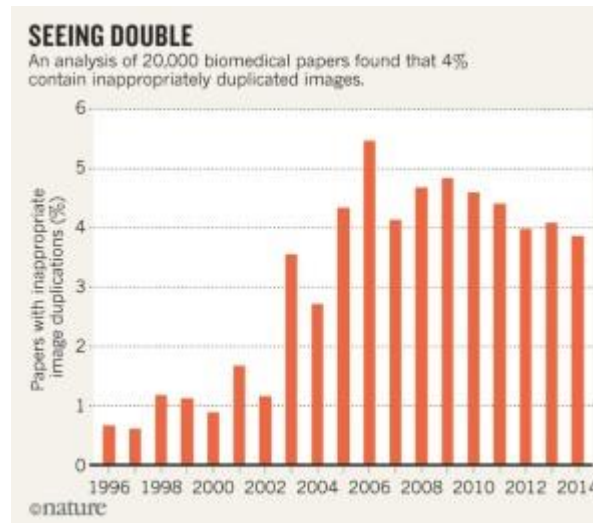
3: Nakon odabira relevantnih biomarkera, **opet trenirate svoj model pomoću cijelog skupa podataka** (trening i test skup podataka) s odabranim značajkama. Cilj je minimizirati pogrešku i maksimizirati snagu predviđanja, ali budući da ste već koristili testne podatke tijekom odabira značajki, **model je "vidio" informacije iz testnog skupa.**

4: Procjenjujete izvedbu svog modela pomoću testnog (koji sada uključuje biomarkere koji su odabrani na temelju cijelog skupa podataka). **Budući da je model neizravno imao pristup informacijama iz testnog skupa tijekom odabira značajki, to vam daje preoptimističnu procjenu izvedbe modela, što dovodi do pristranih rezultata.**

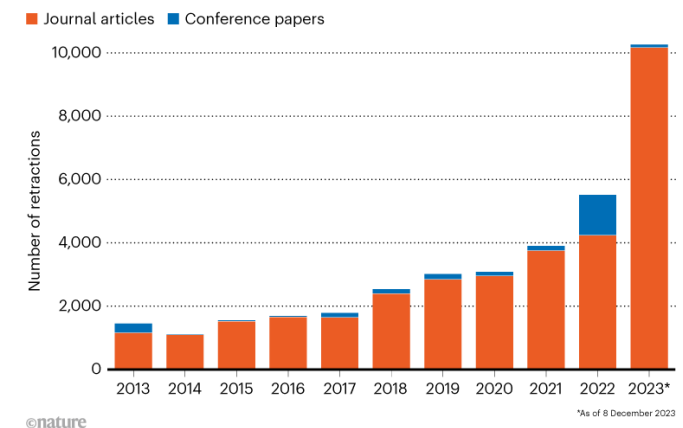
*Modeli se uvijek trebaju trenirati na trening skupu, a testirati na testnom skupu podataka.

Izmišljanje podataka

- Prema studiji iz 2009. (PLoS One) prosječno 2% znanstvenika priznalo je da su barem jednom lažirali podatke, no oko 14% istraživača svjedočilo je lažiranju rezultata od strane kolega.
- 2023 – preko 10,000 povučenih znanstvenih članaka



A BUMPER YEAR FOR RETRACTIONS
Retraction notices in 2023 have passed 10,000, largely because of more than 8,000 retractions by Hindawi.



Neobjavljanje negativnih rezultata

- Časopisi imaju pristranost prema objavljivanju statistički značajnih rezultata
- Znanstvenici rijetko pokušavaju objaviti negativne rezultate, što znači da se većina rezultata ne objavljuje
- Studija 221 istraživanja u društvenim znanostima (Franco et al., Science, 2014)
 - Samo 21% negativnih rezultata objavljeno
 - 95% objavljenih rezultata su statistički značajni



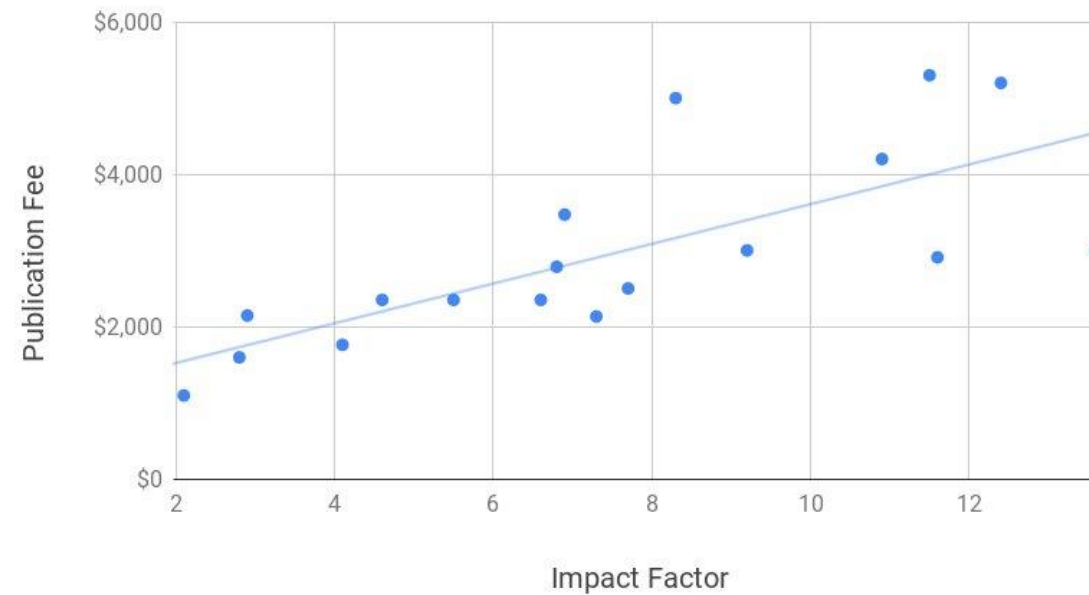
7. Pažljiva interpretacija statističke značajnosti (provedba istraživanja)

- Statistički značajni rezultati ne znače biološku značajnost
- Višestruko testiranje podataka povećava vjerojatnost za lažno pozitivne rezultate

8. Provodite otvorena (transparentna istraživanja) (provedba istraživanja)

- Pohranjujte sirove podatke i kod korišten za analizu
- Objavljujte rezultate u otvorenom pristupu

Publication Fee vs. Impact Factor



9. Presentirajte sve rezultate (prezentacija istraživanja)

- Objavljajte i negativne rezultate kako biste smanjili pristranost u objavljivanju
- Smanjuje vjerojatnost da će drugi znanstvenici ponavljati analizu koja neće dati rezultate

 **BMC** Part of Springer Nature

BMC Research Notes

Aims and scope

BMC Research Notes is an open access journal publishing peer-reviewed contributions from across all scientific and clinical disciplines, including intriguing initial observations, updates to previous work and established methods, valid negative results, and scientific data sets and descriptions. We are committed to making it easier to find, cite and share your short reports by providing an inclusive forum for valuable data and research observations.

10. Sljedite smjernice za objavu rezultata (prezentacija istraživanja)

Table 2. Examples of reporting guidelines for different study types.

Guideline name	Study type
ARRIVE	Animal experiments
CONSORT	Randomized trials
STROBE	Observational studies
PRISMA	Systematic reviews
SPIRIT	Study protocols
STARD/TRIPOID	Diagnostic/prognostic studies

The EQUATOR Network is a library with more than 400 reporting guidelines in health research (www.equator-network.org).

Literatura

- Cobey, K. D., Ebrahimzadeh, S., Page, M. J., Thibault, R. T., Nguyen, P. Y., Abu-Dalfa, F., & Moher, D. (2024). Biomedical researchers' perspectives on the reproducibility of research. *PLoS biology*, 22(11), e3002870. <https://doi.org/10.1371/journal.pbio.3002870>
- Schwab, S., Janiaud, P., Dayan, M., Amrhein, V., Panczak, R., Palagi, P. M., Hemkens, L. G., Ramon, M., Rothen, N., Senn, S., Furrer, E., & Held, L. (2022). Ten simple rules for good research practice. *PLoS computational biology*, 18(6), e1010139. <https://doi.org/10.1371/journal.pcbi.1010139>
- Conroy, G. (2019). The 7 deadly sins of research. <https://www.nature.com/nature-index/news/the-seven-deadly-sins-of-research>
- Tamar R Makin, Jean-Jacques Orban de Xivry (2019) Science Forum: Ten common statistical mistakes to watch out for when writing or reviewing a manuscript eLife 8:e48175 <https://doi.org/10.7554/eLife.48175>

Za one koji žele znati više...

- Ten simple rules collection – kolekcija metodoloških članaka koju je objavljivao PLoS Computational Biology s raznim korisnim temama o kojima nećete čuti puno tijekom studija...

<https://collections.plos.org/collection/ten-simple-rules/>

PLOS COMPUTATIONAL BIOLOGY

Ten simple rules to make the most out of your undergraduate research career

May 4, 2017
Megan Yu, Yu-Min Kuo

PLOS COMPUTATIONAL BIOLOGY

Ten Simple Rules for Effective Statistical Practice

June 9, 2016
Robert E. Kass, Brian S. Caffo, Marie Davidian, Xiao-Li Meng, Bin Yu, Nancy Reid

PLOS COMPUTATIONAL BIOLOGY

Ten Simple Rules for Better Figures

September 11, 2014
Nicolas Rougier, Michael Droettboom, Bourne Philip

PLOS COMPUTATIONAL BIOLOGY

Ten Simple Rules for Graduate Students

November 30, 2007
Jenny Gu, Philip E Bourne

PLOS COMPUTATIONAL BIOLOGY

Ten simple rules for developing good reading habits during graduate school and beyond

October 11, 2018
Marcos Méndez

PLOS COMPUTATIONAL BIOLOGY

Ten simple rules for innovative dissemination of research

April 16, 2020
Tony Ross-Hellauer, Jonathan P. Tennant, Viltė Banelytė, Edit Gorogh, Daniela Luzi, Peter Kraker, Lucio Pisacane, Roberta Ruggieri, Electra Sifacaki, Michela Vignoli