

# Kako Monte Carlo metodom izmjeriti KEP-indeks čitavoj studentskoj populaciji Sveučilišta u Splitu u jedno poslijepodne?

---

**Dodig, Dora**

**Master's thesis / Diplomski rad**

**2023**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Split, School of Medicine / Sveučilište u Splitu, Medicinski fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:171:876527>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-07-14**



*Repository / Repozitorij:*

[MEFST Repository](#)



**SVEUČILIŠTE U SPLITU  
MEDICINSKI FAKULTET**

**Dora Dodig**

**KAKO *MONTE CARLO* METODOM IZMJERITI KEP-INDEKS ČITAVOJ  
STUDENSKOJ POPULACIJI SVEUČILIŠTA U SPLITU U JEDNO POSLIJEPODNE?**

**Diplomski rad**

**Akadska godina:**

**2022./2023.**

**Mentor:**

**izv. prof. dr. sc. Darko Kero**

**Split, rujan 2023.**

*Hvala mom mentoru, izv. prof. dr. sc. Darku Keri na strpljenju, brizi, stručnosti te nesebičnoj pomoći koju mi je ukazao za vrijeme pisanja ovog diplomskog rada*

*Ovaj rad posvećujem svima koji su tu bili za mene i koji me vole, od samog rođenja, pa do ovog trenutka.*

## SADRŽAJ

<b>1. UVOD</b> .....	1
<b>2. CILJ ISTRAŽIVANJA</b> .....	4
<b>3. ISPITANICI I POSTUPCI</b> .....	6
3.1. Prikupljanje podataka.....	7
3.2. Statistička analiza.....	7
<b>4. REZULTATI</b> .....	9
4.1. Demografski parametri ispitanika i obilježja KEP-indeksa iz uzorka.....	10
4.2. Obilježja KEP-indeksa studentskih populacija generiranih <i>Monte Carlo</i> simulacijom.....	11
4.3. Procjena prosječnog KEP-indeksa za studentsku populaciju prema raspodjeli prosječnih vrijednosti KEP-indeksa uzoraka simuliranih <i>Monte Carlo</i> simulacijom.....	15
<b>5. RASPRAVA</b> .....	18
<b>6. ZAKLJUČAK</b> .....	21
<b>7. LITERATURA</b> .....	23
<b>8. SAŽETAK</b> .....	26
<b>9. SUMMARY</b> .....	28
<b>10. ŽIVOTOPIS</b> .....	31

## **1. UVOD**

Jedan od ključnih preduvjeta za valjanost statističke analize u znanstvenim istraživanjima se odnosi na reprezentativnost uzorka s obzirom na populaciju koja se istražuje (1). Ako je uzorak reprezentativan, onda se i parametri mjerene varijable (ili više varijabli) iz uzorka mogu s većom točnošću poopćiti na ukupnu populaciju iz koje je uzorak uzet. Jedan od postupaka kojim se može osigurati reprezentativnost uzorka je randomizacija što znači da prilikom uzorkovanja za istraživanje svaki pojedinac iz određene populacije ima jednake izgleda da bude uključen u istraživanje (1-3). U slučaju randomiziranih kliničkih pokusa, to bi značilo da svaki pacijent koji uzima određeni lijek ili prima neku vrstu terapije, ima jednake izgleda biti uključen u istraživanje o učinkovitosti tog lijeka ili terapije.

Osim randomizacije ispitanika, jednako tako je važno odrediti i potrebnu veličinu uzorka (4). Istraživači se mogu voditi logikom da što je uzorak veći je ujedno i reprezentativniji za populaciju iz koje se uzima, što bi prema principima statističke analize svakako bilo opravdano. Međutim, uzimanje i obrada velikog uzorka (primjerice u laboratorijskim eksperimentalnim istraživanjima i kliničkim istraživanjima) zbog financijskih i organizacijskih zahtjeva često nisu izvedivi. Kako bi se doskočilo tome, osmišljeni su statistički postupci u sklopu analize snage studije (eng. *Power analysis*) kojima se *a priori* može odrediti potrebna veličina uzorka (5). Također, reprezentativnost uzorka se lakše postiže ako se, prema cilju istraživanja i glavnom istraživačkom pitanju, ciljana populacija „smanjuje" uvođenjem dodatnih kriterija za uzorkovanje koji se odnose primjerice na dobnu ili društvenu skupinu, ili oboljele od rijetke bolesti i sl. Osim toga, i određene varijable mogu se jednostavnije prikazati indeksima što je prikladno za epidemiološka i klinička istraživanja (6). U dentalnoj medicini se za mjeru oralnog zdravlja (koje je samo po sebi složena varijabla) koriste različiti klinički indeksi, od kojih je najpoznatiji KEP-indeks (kao zbroj karioznih, izvađenih i saniranih zubi u pojedincu) s teoretski zadanim rasponom od 0 do 64 (7, 8).

Nedostatak svih postupaka koji se koriste za postizanje reprezentativnosti uzorka je što, ukoliko nije poznat određeni populacijski parametar, nikad nije moguće sa sigurnošću tvrditi da je njihovom provedbom ostvarena reprezentativnost uzorka (9, 10). Primjerice, raspodjela vrijednosti za pojedinu varijablu u uzorku (neovisno o tome radi li se o uniformnoj, normalnoj ili nekom tipu asimetrične/ukošene raspodjele poput eksponencijalne) ne mora nužno odgovarati raspodjeli te varijable u populaciji (11). Čak ni pravilno provedena randomizacija pri odabiru ispitanika za

istraživanje ne jamči da je dobiveni uzorak reprezentativan za populaciju iz koje je uzet. U slučaju multifaktorijalnih kroničnih bolesti, iluzorno je očekivati da će se odabirom ispitanika u uzorku adekvatno pokriti pojedinačni i/ili relativni utjecaj svih čimbenika odgovornih za nastanak i razvoj bolesti. Kod oralnih bolesti poput karijesa, uz multifaktorijalnu patogenezu, dodatni problem kod uzorkovanja predstavlja i općenito velika prevalencija te bolesti te raznolika zastupljenost karijesa ovisno o dobnim skupinama. S obzirom na to, logično je pretpostaviti da bi procjene o zastupljenosti karijesa i njegovim posljedicama u populaciji bilo opravdano ili raditi samo na velikim uzorcima kako bi se rezultati mogli smatrati koliko-toliko pouzdanim, ili opetovano replicirati isto istraživanje na novim uzorcima. Svrha i jednog i drugog pristupa je okrupnjavanje uzorka zbog bolje statističke procjene populacijskih parametara bolesti.

Treći pristup okrupnjavanju uzorka se temelji na stohastičkom algoritamskom postupku prema *Monte Carlo* metodi koju su 1949. godine razvili John von Neumann, Stanislaw Ulam i Nicholas Metropolis, znanstvenici koji su prethodno radili na projektu razvoja nuklearnog oružja za američku vladu poznatog kao „Manhattan Project“ (12). *Monte Carlo* metodom se pomoću slučajnih ili kvazislučajnih brojeva i velikog broja proračuna može računalno simulirati opetovano uzorkovanje uzevši u obzir: (i) da su parametri simuliranih uzoraka međusobno različiti čime se odražava raznolikost ispitivane populacije; (ii) da su prema središnjem graničnom teoremu prosječne vrijednosti mjerene varijable u simuliranim uzorcima bliže normalnoj raspodjeli što je  $n$  simuliranih uzoraka veći; (iii) a središnje mjesto u toj normalnoj raspodjeli pripada prosječnoj vrijednosti mjerene varijable koja je bliska pravoj populacijskoj vrijednosti (11).

## **2. CILJ ISTRAŽIVANJA**



Cilj ovog istraživanja bio je generirati podatke za KEP-indeks čitave studentske populacije Sveučilišta u Splitu iz tri uzorka različite veličine korištenjem *Monte Carlo* metode.

Hipoteze istraživanja su sljedeće:

1. Prosječne vrijednosti KEP-indeksa studentskih populacija Sveučilišta u Splitu generiranih *Monte Carlo* simulacijom (MCS) će se za minimalno jedan bod razlikovati od prosječne vrijednosti KEP-indeksa pripadajućeg uzorka iz koje su te populacije generirane.
2. Prema povezanosti broja karioznih, izvađenih i saniranih zubi s KEP-indeksom, sastav KEP-indeksa studentskih populacija Sveučilišta u Splitu generiranih MCS-om će biti različit.
3. Prosječne vrijednosti KEP-indeksa simuliranih uzoraka u svakom od tri MCS-a će biti normalno raspodijeljene.
4. Prosječne vrijednosti KEP-indeksa studentskih populacija Sveučilišta u Splitu generiranih MCS-om će se međusobno razlikovati za minimalno jedan bod.

### **3. ISPITANICI I POSTUPCI**

### 3.1. Prikupljanje podataka

Podatci za potrebe istraživanja prikupljeni su tijekom akademskih godina 2021./22. i 2022./23., u sklopu kliničkih vježbi na kolegijima „Restaurativna dentalna medicina 2“ te „Endodoncija 2“. Svi ispitanici su bili studenti Sveučilišta u Splitu koji su došli na pregled ili liječenje u ordinacije Ustrojbene jedinice Centra za zdravstvenu djelatnost u dentalnoj medicini Medicinskog fakulteta Sveučilišta u Splitu „Dental academicus“ u prostorijama studentskog doma „Dom Kampus dr. Franjo Tuđman“ u Splitu. Ispitanici su potpisali informirani pristanak na liječenje i obradu osobnih podataka za potrebe istraživanja, uzeta im je osobna, obiteljska i stomatološka anamneza, obavljen klinički pregled usne šupljine i zabilježen oralni status pomoću KEP-indeksa. Preglede su obavili studenti četvrte i pete godine Integriranog studija dentalne medicine Medicinskog fakulteta Sveučilišta u Splitu uz nadzor voditelja vježbi. Svi sirovi podatci su pohranjeni u objedinjenu tablicu u programu Microsoft Office Excel 2016 (Microsoft Corporation, Redmon, WA, USA).

### 3.2. Statistička analiza

Demografski parametri ispitanika poput dobi i spola analizirani su opisnom statistikom, izraženi bročanim vrijednostima, postotcima i/ili prosječnim vrijednostima uz standardnu devijaciju (SD) i raspon pouzdanosti (eng. *Confidence Interval* – CI) od 99%. Vrijednosti za KEP-indeks i pod-indekse K, E i P također su prikazan bročano (diskretni tip podataka), te kao prosjeci uz SD i 99% CI. Za provjeru raspodjele podataka svih varijabli (osim spola) korišteni su histogrami.

Iz prikupljenih podataka o KEP-indeksu izvedena su tri uzorka – mali uzorak od prvih 50 ispitanika, uzorak srednje veličine od narednih 100 ispitanika i cjelokupni uzorak od 200 ispitanika. Postavljena su dva glavna kriterija za simuliranje uzoraka MCS-om i generiranje podataka za čitavu studentsku populaciju, a to su vjerojatnosti pojedinih ishoda te međusobna povezanost KEP-indeksa i njegovih sastavnica (K, E i P). Vjerojatnosti pojedinih ishoda prikazane su funkcijom gustoće vjerojatnosti (eng. *Probability Density Function*) i funkcijom kumulativne raspodjele (eng. *Cumulative Distribution Function*). S obzirom da su vrijednosti KEP.indeksa i njegovih sastavnica cijeli brojevi, samo potonji prikaz je korišten kao pretpostavka za MCS. Iako je teoretski raspon KEP-indeksa od 0 do 64, vjerojatnosti ishoda su prikazane za raspon od 0 do 41 nakon što je proveden predviđeni broj iteracija za MCS iz sva tri uzorka. Kako bi se i međusobna

povezanost KEP-indeksa i njegovih sastavnica uvrstila kao dodatni kriterij za MCS, napravljene su korelacijske matrice za navedene varijable u sva tri uzorka (Pearson-ova korelacija), a potom i dekompozicija korelacijskih matrica (dekompozicija Cholesky). Kako bi se uklonile decimalne i/ili negativne vrijednosti varijabli u simuliranim uzorcima, napravljene su dvije korekcije generiranih vrijednosti – zaokruživanjem decimalnih brojeva na cijele brojeve i zamjenom negativnih vrijednosti nulom. Vrijednosti i parametri varijabli u simuliranim uzorcima i generiranim populacijama (prosjeci, SD, 99% CI, rasponi i korelacijski koeficijenti) zabilježeni su u zasebnim tablicama. Raspodjela prosječnih vrijednosti za KEP-indeks u simuliranim uzorcima svake od generiranih populacija prikazana je grafički.

Povezanost KEP-indeksa i njegovih sastavnica za generirane populacije analizirana je jednostavnom linearnom regresijom s KEP-indeksom kao glavnim ishodom, te K, E i P podindeksima kao prediktorima. Usporedba prosječnih vrijednosti KEP-indeksa početnih uzoraka, te KEP-indeksa među studentskim populacijama generiranim MCS-om, napravljena je višestrukom linearnom regresijom s KEP-indeksom kao glavnim ishodom, te prediktorima koji se odnose na pripadnost odgovarajućem početnom uzorku ili generiranoj studentskoj populaciji kodiranim kao kategorijske „dummy“ varijable. Kodiranje „dummy“ varijabli napravljeno je kako je prethodno opisano (13-15).

Statistička analiza je u potpunosti provedena u Microsoft Excel-u 2016 (Microsoft Corporation, Redmond, WA, USA). Razina značajnosti je postavljena na  $\alpha = 0,01$  ( $P < 0,01$ ).

## **4. REZULTATI**

#### 4.1. Demografski parametri ispitanika i obilježja KEP-indeksa iz uzorka

Za istraživanje su prikupljeni podatci o KEP-indeksu od ispitanika ( $n = 200$ ) prosječne dobi  $22,89 \pm 2,81$  godina (99% CI [22,62; 23,16]). Raspodjela po spolu u uzorku je bila relativno ujednačena s 45,60% (93/200) muškaraca i 54,40% (107/200) žena. Prosječni KEP-indeks je iznosio  $9,79 \pm 5,03$  (99% CI [9,3; 10,27]), uz prosječni broj karioznih zubi (K)  $2,6 \pm 2,75$  (99% CI [2,36; 2,85]), izvađenih zubi (E)  $0,83 \pm 1,92$  (99% CI [0,65; 1,01]), odnosno saniranih zubi (P)  $6,35 \pm 4,55$  (99% CI [5,93; 6,76]).

Prosječna dob ispitanika u malom uzorku ( $n = 50$ ) je bila  $26,36 \pm 2,24$  godina (99% CI [25,93; 26,79]) s raspodjelom spola od 58% (29/50) muškaraca i 42% (21/50) žena. Prosječni KEP-indeks je iznosio  $9,68 \pm 5,9$  (99% CI [8,56; 10,79]) uz prosječni broj karioznih zubi (K)  $2,42 \pm 2,31$  (99% CI [1,98; 2,86]), izvađenih zubi (E)  $0,96 \pm 2,12$  (99% CI [0,56; 1,36]), odnosno saniranih zubi (P)  $6,3 \pm 4,8$  (99% CI [5,39; 7,21]).

Prosječna dob ispitanika u srednjem uzorku ( $n = 100$ ) je bila  $22,29 \pm 0,84$  godine (99% CI [22,18; 22,41]) s ravnomjernom raspodjelom spola od 50% muškaraca (50/100) i žena (50/100). Prosječni KEP-indeks je iznosio  $9,77 \pm 5,57$  (99% CI [9,03; 10,5]), uz prosječni broj karioznih zubi (K)  $2,66 \pm 2,32$  (99% CI [2,35; 2,97]), izvađenih zubi (E)  $1,01 \pm 2,13$  (99% CI [0,73; 1,29]) i saniranih zubi (P)  $6,1 \pm 4,44$  (99% CI [5,52; 6,68]).

Rezultati korelacijske analize su potvrdili da je sastav KEP-indeksa bio sličan u sva tri uzorka (**Tablica 1**). Od sastavnica KEP-indeksa, pod-indeks P najjače je povezan s KEP-indeksom i to  $R = 0,87$  (mali uzorak),  $R = 0,86$  (srednji uzorak) i  $R = 0,80$  (cjelokupni uzorak). Ovi nalazi upućuju na to da se radi o ispitanicima koji su relativno dobro sanirani.

**Tablica 1.** Korelacijske matrice KEP-indeksa i K, E i P pod-indeksa za mali (n = 50), srednji (n = 100) i cjelokupni uzorak (n = 200) prema podacima o KEP-indeksu prikupljenim od ispitanika.

Korelacija (Pearson)	Uzorci											
	n = 50				n = 100				n = 200			
Varijable	K	E	P	KEP	K	E	P	KEP	K	E	P	KEP
K	1	-0,15	-0,20	0,17	1	-0,16	-0,16	0,23	1	-0,08	-0,28	0,25
E	-0,15	1	0,39	0,62	-0,16	1	0,34	0,58	-0,08	1	0,23	0,51
P	-0,20	0,39	1	0,87	-0,16	0,34	1	0,86	-0,28	0,23	1	0,80
KEP	0,17	0,62	0,87	1	0,23	0,58	0,86	1	0,25	0,51	0,8	1

**Korelacijski koeficijenti (R)**

#### 4.2. Obilježja KEP-indeksa studentskih populacija generiranih *Monte Carlo* simulacijom

U svakom od tri postupka za MCS veličina studentske populacije je pretpostavljena na n = 20000. U podatke o KEP-indeksu za svaku od tri generirane studentske populacije uključeni su i izvorni podatci iz malog, srednjeg i cjelokupnog uzorka – za generiranje populacije iz malog uzorka provedeno je 400 iteracija (MCS 50), iz srednjeg uzorka 200 iteracija (MCS 100), a iz cjelokupnog uzorka 100 iteracija (MCS 200).

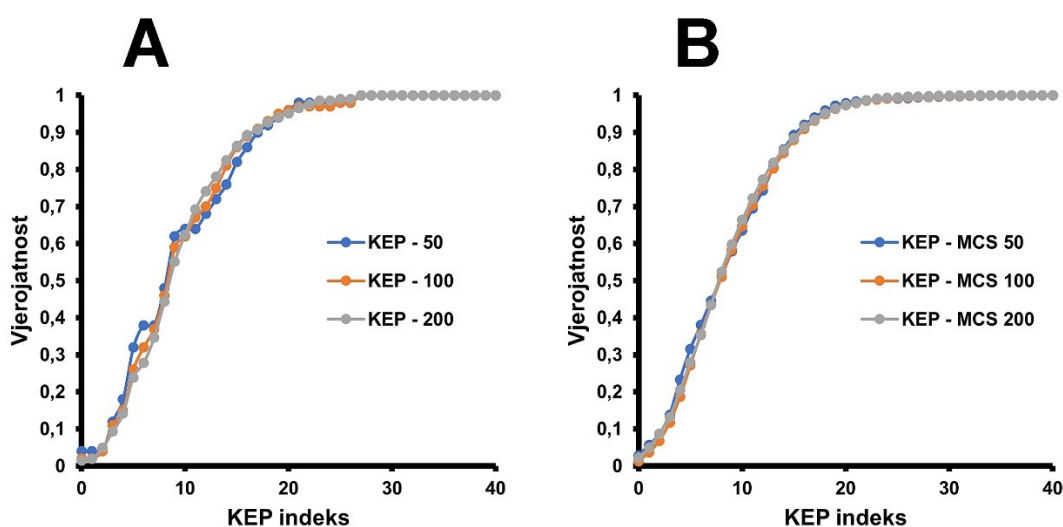
Za studentsku populaciju generiranu u MCS 50, prosječni KEP-indeks iznosio je  $8,92 \pm 5,33$  (99% CI [8,87; 8,97]), prosječni broj karioznih zubi (K)  $2,44 \pm 2,3$  (99% CI [2,41; 2,46]), izvađenih zubi (E)  $0,82 \pm 1,92$  (99% CI [0,8; 0,84]), a saniranih zubi  $5,66 \pm 4,35$  (99% CI [5,62; 5,7]).

Za studentsku populaciju generiranu u MCS 100, prosječni KEP-indeks iznosio je  $9,15 \pm 5,22$  (99% CI [9,09; 9,19]), prosječni broj karioznih zubi (K)  $2,63 \pm 2,95$  (99% CI [2,61; 2,65]), izvađenih zubi (E)  $0,89 \pm 2,01$  (99% CI [0,87; 0,91]) te saniranih zubi (P)  $5,63 \pm 4,16$  (99% CI [5,59; 5,67]).

Za studentsku populaciju generiranu u MCS 200, prosječni KEP-indeks iznosio je  $8,93 \pm 5,19$  (99% CI [8,88; 8,98]), prosječni broj karioznih zubi (K)  $2,62 \pm 2,77$  (99% CI [2,59; 2,64]),

izvađenih zubi (E)  $0,81 \pm 1,89$  (99% CI [0,79; 0,82]), a saniranih zubi (P)  $5,51 \pm 4,24$  (99% CI [5,47; 5,54]).

Prema izgledu funkcije kumulativne raspodjele za KEP-indeks, vjerojatnost pojedinih ishoda za KEP-indeks u studentskim populacijama generiranim MCS-om većim dijelom se preklapala s odgovarajućom funkcijom kumulativne raspodjele iz izvornih uzoraka uz iznimku maksimalnih vrijednosti za KEP-indeks (**Slika 1**). Raspon za KEP-indeks u sva tri izvorna uzorka se kretao od 0 do 27, a u generiranim populacijama i do 38 (MCS 50), odnosno 37 (MCS 100 i MCS 200).



**Slika 1.** Usporedba pojedinačnih ishoda za KEP-indeks prema krivuljama kumulativne raspodjele KEP-indeksa u izvornim uzorcima (A) i studentskim populacijama generiranim MCS-om (B) iz malog (MCS 50), srednjeg (MCS 100) i cjelokupnog uzorka (MCS 200).

Rezultati korelacijske analize su potvrdili sličan sastav KEP-indeksa u sva tri studentske populacije generirane MCS-om (**Tablica 2**). Od sastavnica KEP-indeksa, pod-indeks P najjače je povezan s KEP-indeksom i to  $R = 0,84$  (MCS 50),  $R = 0,82$  (MCS 100) i  $R = 0,77$  (MCS 200). Nalazi korelacije KEP-indeksa i njegovih sastavnica u studentskim populacijama generiranim MCS-om ne odstupaju bitno od korelacijskih obrazaca dobivenih prema podacima iz izvornih uzoraka.



**Tablica 2.** Korelacijske matrice KEP-indeksa i K, E i P pod-indeksa studentskih populacija generiranih MCS-om iz izvornih uzoraka.

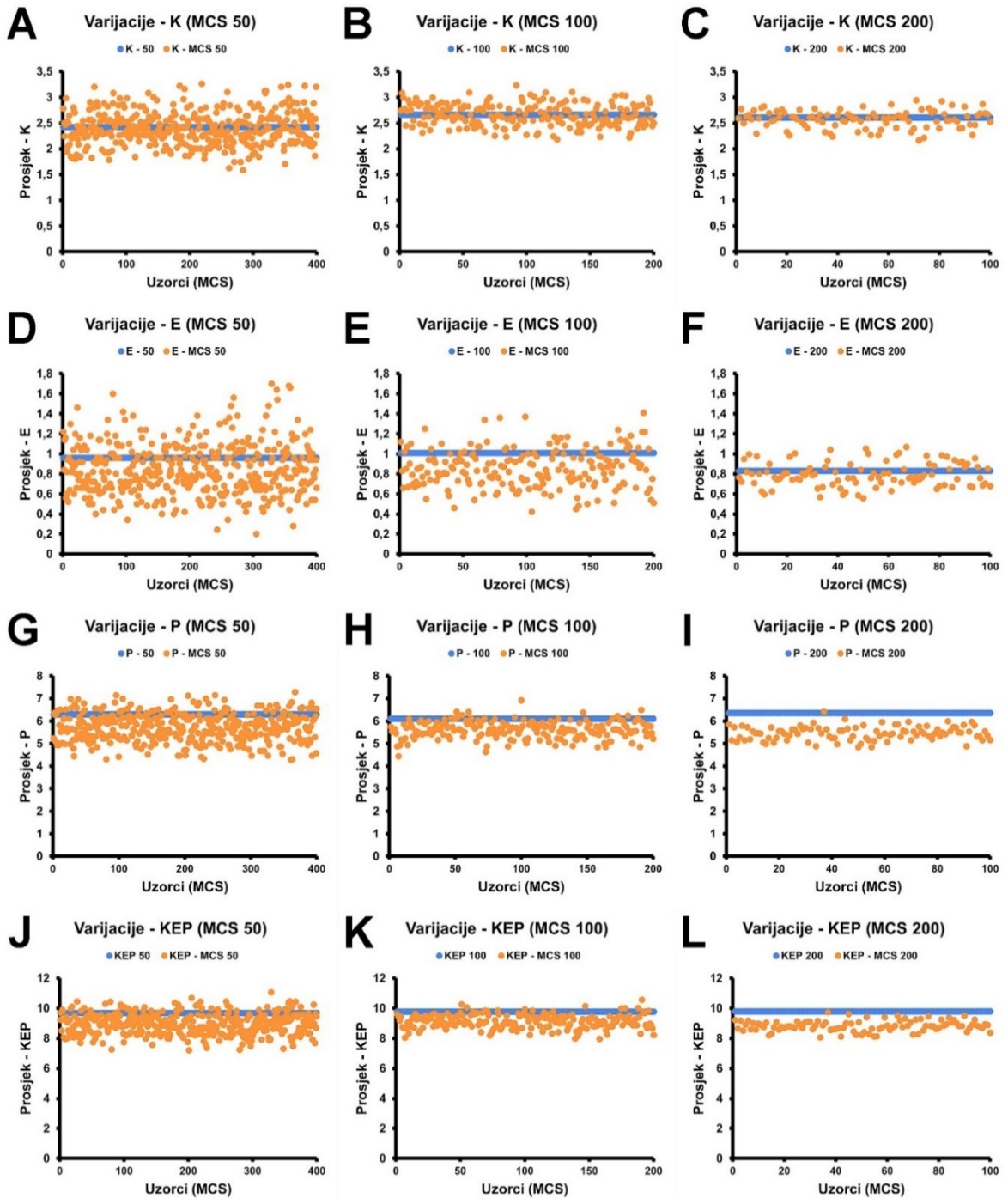
Korelacija (Pearson)	Studentske populacije											
	MCS 50*				MCS 100 <sup>†</sup>				MCS 200 <sup>‡</sup>			
Varijable	K	E	P	KEP	K	E	P	KEP	K	E	P	KEP
K	1	-0,06	-0,10	0,33	1	-0,06	-0,08	0,35	1	-0,02	-0,15	0,40
E	-0,06	1	0,18	0,48	-0,06	1	0,17	0,49	-0,02	1	0,09	0,43
P	-0,10	0,18	1	0,84	-0,08	0,17	1	0,82	-0,15	0,09	1	0,77
KEP	0,33	0,48	0,84	1	0,35	0,49	0,82	1	0,40	0,43	0,77	1

**Korelacijski koeficijenti (R)**

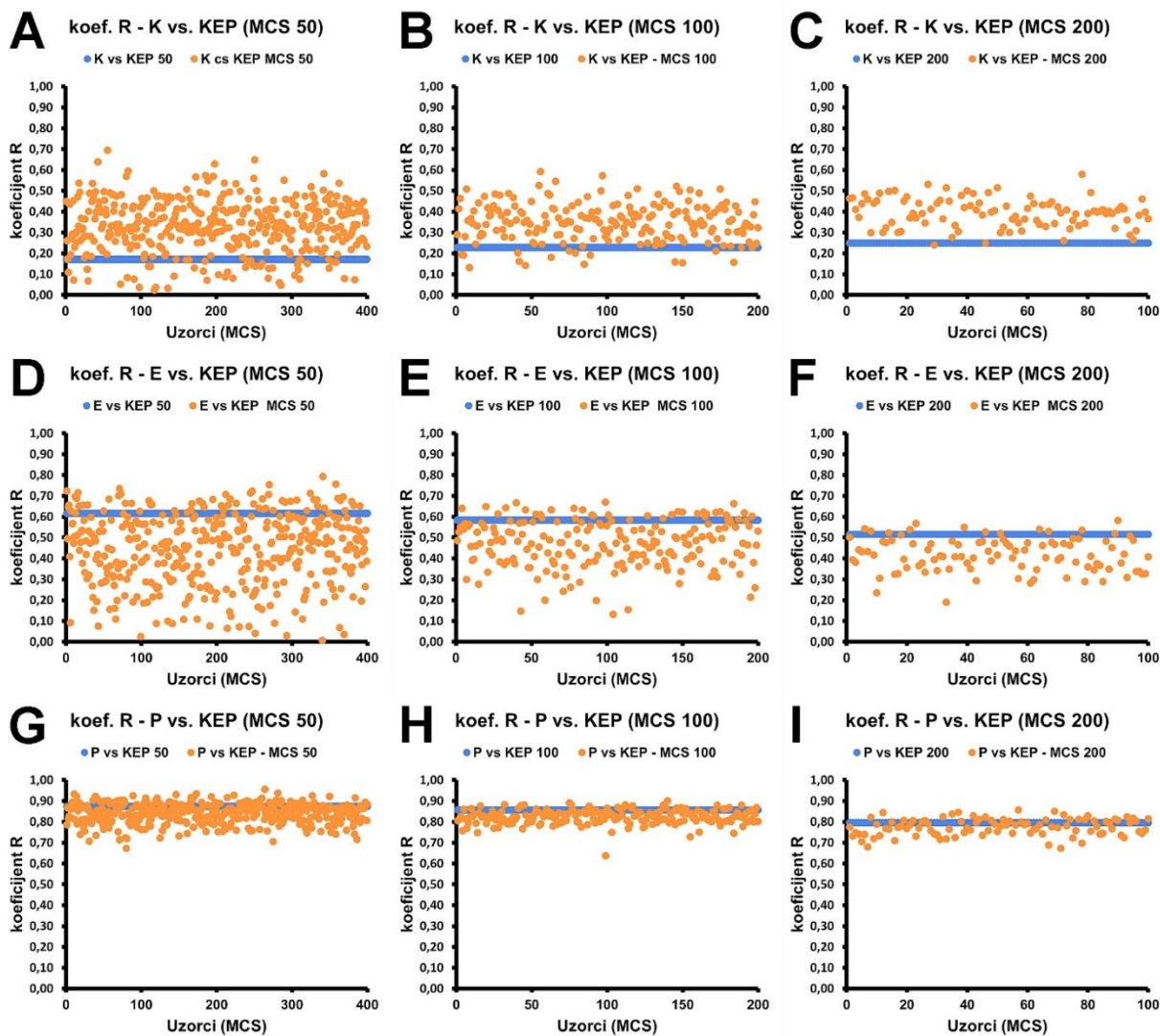
\*Studentska populacija generirana MCS-om prema malom izvornom uzorku (n = 50) <sup>†</sup>Studentska populacija generirana MCS-om prema srednjem izvornom uzorku (n = 100) <sup>‡</sup>Studentska populacija generirana MCS-om prema cjelokupnom izvornom uzorku (n = 200)

Analizirani su i obrasci varijabilnosti parametara u simuliranim uzorcima koji su uklopljeni u studentske populacije generirane MCS-om. Od parametara prikazana je varijabilnost prosječnih vrijednosti KEP-indeksa i njegovih sastavnica (**Slika 2**), odnosno korelacijskih koeficijenata (**Slika 3**). Pokazalo se da je varijabilnost KEP-indeksa i njegovih sastavnica, odnosno korelacijskih koeficijenata općenito izraženija što je izvorni uzorak prema kojemu se izvodi MCS manji. Među KEP-indeksom i njegovim sastavnicama, najmanju varijabilnost u MCS-u prikazao je sam KEP-indeks kao složena varijabla koja se izravno izvodi iz tri sastavnice – samo su u manjem dijelu simuliranih uzoraka prosječne vrijednosti za KEP-indeks bile više u odnosu na one izmjerene u pripadajućim uzorcima. Štoviše, u MCS 200 nijedna prosječna vrijednost KEP-indeksa nije prebacivala onu izmjerenu u cjelokupnom izvornom uzorku.

fsdfsd



**Slika 2.** Varijabilnost prosječnih vrijednosti pod-indeksa K (A-C), E (D-F), P (G-I) i KEP-indeks (J-L) u simuliranim uzorcima izvedenim MCS-om iz malog (A, D, G, J), srednjeg (B, E, H, K) i cjelokupnog (C, F, I, L) izvornog uzorka. Plave vodoravne crte označavaju prosječne vrijednosti izmjerene u izvornim uzorcima.

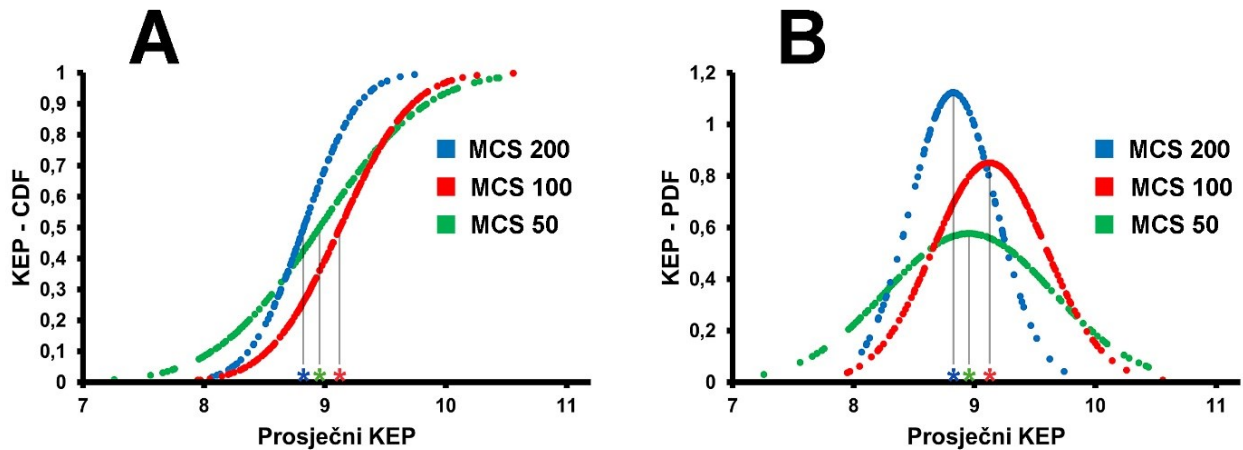


**Slika 3.** Varijabilnost vrijednosti korelacijskog koeficijenta R u simuliranim uzorcima izvedenim MCS-om iz malog (A, D, G), srednjeg (B, E, H) i cjelokupnog (C, F, I) izvornog uzorka. Plave vodoravne crte označavaju korelacijske koeficijente izmjerene u izvornim uzorcima.

### 4.3. Procjena prosječnog KEP-indeksa za studentsku populaciju prema raspodjeli prosječnih vrijednosti KEP-indeksa uzoraka simuliranih *Monte Carlo* simulacijom

Prema središnjem graničnom teoremu prosječne vrijednosti KEP-indeksa simuliranih uzoraka trebale bi biti normalno raspodijeljene pod uvjetom da je broj simuliranih uzoraka dovoljno velik. Prema postulatima istog teorema, središnje mjesto u toj raspodjeli trebalo bi pripadati stvarnoj prosječnoj vrijednosti KEP-indeksa studentske populacije. Raspodjele

prosječnih vrijednosti KEP-indeksa svih simuliranih uzoraka prikazane su zasebnim krivuljama za pojedini MCS (**Slika 4**). Procjena za prosječni KEP-indeks studentske populacije prema MCS 50 je  $8,96 \pm 0,69$  (99% CI [8,91; 9]), prema MCS 100  $9,12 \pm 0,47$  (99% CI [9,08; 9,21]), te prema MCS 200  $8,82 \pm 0,36$  (99% CI [8,77; 8,87]).



**Slika 4.** Raspodjela prosječnih vrijednosti KEP-indeksa u simuliranim uzorcima prema MCS 50 (zeleno), MCS 100 (crveno) i MCS 200 (plavo) prikazana funkcijom kumulativne raspodjele (A) i funkcijom gustoće vjerojatnosti (B). Procjene za stvarne prosječne vrijednosti KEP-indeksa studentske populacije prema pojedinom MCS-u označene su zvjezdicama.

Od tri procjene prosječnog KEP-indeksa za studentsku populaciju statistički značajno odstupa jedino procjena prema MCS 100 u odnosu na MCS 50 i MCS 200. Međutim, ta odstupanja su manja od jednog boda (0,16 bodova u odnosu na procjenu prema MCS 50, odnosno 0,30 bodova u odnosu na procjenu prema MCS 200) (**Tablica 3**) što se može smatrati zanemarivim odstupanjem.

**Tablica 3.** Usporedba procjene prosječnih vrijednosti KEP-indeksa za studentsku populaciju prema prosječnim vrijednostima KEP-indeksa uzoraka simuliranih MCS-om iz malog (MCS 50), srednjeg (MCS 100) i cjelokupnog (MCS 200) izvornog uzorka.

ISHOD	„Dummy“ varijable (kategorije)	PREDIKTORI		Parametri modela		
		Koeficijent*	P-vrijednost	R	P- vrijednost <sup>†</sup>	df <sup>‡</sup>
Prosječni KEP <sup>¶</sup>	KEP – MCS 200 <sup>§</sup>	/	/	0,16	< 0,001	2
	KEP – MCS 100	0,13	0,049			
	KEP – MCS 50	0,30	< 0,001			

\*Izraženi kao razlika (pozitivna/negativna) od prosjeka prosječnih vrijednosti KEP-indeksa svih simuliranih uzoraka iz MCS 200 kao referentne kategorije

<sup>†</sup>Višestruka linearna regresija ( $\alpha = 0.01$  ( $P < 0.01$ ); <sup>‡</sup>stupnjevi slobode)

<sup>§</sup>Referentna kategorija ("dummy" kodiranje), n = 100

<sup>¶</sup>Zavisna varijabla – uključuje prosječne vrijednosti KEP-indeksa svih simuliranih uzoraka iz MCS 50, MCS 100 i MCS 200 (n = 700)

## **5. RASPRAVA**

U ovom istraživanju generirani su podaci te napravljena procjena prosječne vrijednosti i sastava KEP-indeksa čitave studentske populacije Sveučilišta u Splitu prema uzorcima različite veličine iz iste populacije korištenjem *Monte Carlo* metode. Prema rezultatima, parametri KEP-indeksa iz izvornih uzoraka i simuliranih studentskih populacija dobivenih MCS-om (prosječne vrijednosti, SD i korelacijski koeficijenti za procjenu povezanosti sastavnica KEP-indeksa s KEP-indeksom) nisu se bitno razlikovali, osim CI za prosječne vrijednosti koji su bili znatno uži za simulirane studentske populacije. Potonje proizlazi iz veličine simuliranih studentskih populacija koje su bile od 100 do 400 puta veće od izvornog uzorka.

Procjene prosječnog KEP-indeksa studentske populacije iz ovog istraživanja su uglavnom u skladu s procjenama prosječnog KEP-indeksa iz sličnih istraživanja koja su prethodno provedena na uzorcima različite veličine (od 50 do 1000 ispitanika) adolescenata i studentskih populacija Sveučilišta u Splitu i Sveučilišta u Zagrebu (8, 16-20). Procjene prosječnog KEP-indeksa u tim istraživanjima su se kretale od 7 do 11 bodova za KEP-indeks što je unutar raspona prosječnog KEP-indeksa za uzorke koji su u ovom istraživanju simulirani *Monte Carlo* metodom.

Ograničenja procjene prosječnog KEP-indeksa MCS-om kako je prikazano u ovom istraživanju proizlaze iz činjenice da, iako ta procjena može biti statistički opravdana, i dalje ne mora odgovarati stvarnom stanju na terenu. Naime, preciznost MCS-a zadana je prije svega raspodjelom podataka u uzorku te ulaznim pretpostavkama prema kojima se MCS izvodi. Što je raspodjela podataka u uzorku sličnija stvarnoj raspodjeli u populaciji, odnosno, što je uzorak reprezentativniji za populaciju, to će i procjena parametara MCS-om biti preciznija. Poklapanje raspodjele podataka u uzorku i populaciji samo djelomično ovisi o veličini uzorka (21-23). Kod varijabli s uniformnom raspodjelom, precizne procjene parametara se mogu raditi i na temelju malih uzoraka jednako kao što je u tom slučaju i za MCS potreban manji broj iteracija. U slučaju varijabli s normalnom raspodjelom ili nekom vrstom asimetrične/ukošene raspodjele (kao što je ekspancijalna), za precizniju procjenu parametara potreban je veći uzorak, odnosno mora se provesti MCS s većim brojem iteracija da bi procjena prosječne vrijednosti parametara za populaciju bila u skladu s postulatima središnjeg graničnog teorema (11). Veličina uzorka u ovom slučaju je važna samo iz razloga što je kod većih uzoraka vjerojatnije da će biti reprezentativniji, tj. da će se vjernije obuhvatiti stvarna raspodjela podataka za neku varijablu kakva je u populaciji.

To, naravno, ne znači da se isto ne može postići i malim uzorkom samo je manje izgledno (11, 24, 25).

Uvođenje korelacije među varijablama kao dodatne pretpostavke za MCS je važan postupak jer su varijable rijetko zaista nezavisne jedna od druge. KEP-indeks je složena varijabla koja izravno proizlazi (i time je izravno povezana) s brojem karioznih, izvađenih i saniranih zubi. Povezanost također postoji i među sastavnicama KEP-indeksa, a primjer za to je što se u broj karioznih zubi ubrajaju i zubi sa sekundarnim karijesom koji se razvija uz ispune, što znači da su zubi prethodno morali biti sanirani. Također, nedostatak zubi može biti posljedica karijesa, a istovremeno izvađene zube se ne može ubrajati pod kariozne ili sanirane zube. Prosječni KEP-indeks studentske populacije prema MCS-u temeljenom samo na vjerojatnosti pojedinih ishoda za KEP-indeks i njegovih sastavnica procijenjen je na oko 15 što za 6 bodova prebacuje procjenu prosječnog KEP-indeksa prikazanog u ovom istraživanju. U ovom istraživanju iz korelacijskih matrica izostavljena je dob s obzirom da su ispitanici bili iz jedne dobne skupine i varijabilnost u dobi nije bila izražena. Međutim, da su u uzorku bili ispitanici iz nekoliko različitih dobnih skupina, povezanost KEP-indeksa i dobi bi se morala ugraditi kao jedna od pretpostavki za MCS.

*Monte Carlo* metoda može biti korisna i trebala bi se češće primjenjivati u procjeni prosječnih populacijskih vrijednosti kliničkih indeksa u dentalnoj medicini uključujući i KEP-indeks. Najveća prednost *Monte Carlo* metode je što se računalno u kratkom vremenu može simulirati nasumičnost parametara kakva se susreće prilikom opetovanog uzimanja uzoraka. Iako se MCS može uspješno izvesti i iz malih uzoraka, okrupnjavanje uzorka je važno s obzirom na broj varijabli i, shodno tome, broj pretpostavki koje se mogu ugraditi u MCS. Za potrebe diplomskih radova, to znači da je opravdano repliciranje istraživanja jer se tako mogu prikupiti klinički podaci čija daljnja obrada MCS-om (kao što je prikazano i u ovom istraživanju na primjeru procjene KEP-indeksa za studentsku populaciju Sveučilišta u Splitu) može biti korisna za uvid u stanje oralnog zdravlja određene populacije.



## **6. ZAKLJUČAK**

Prema rezultatima analize izvedeni su sljedeći zaključci:

1. Prosječne vrijednosti KEP-indeksa studentskih populacija Sveučilišta u Splitu generiranih *Monte Carlo* simulacijom (MCS) razlikuju se za manje od jednog boda od prosječne vrijednosti KEP-indeksa pripadajućeg uzorka iz koje su te populacije generirane.
2. Prema povezanosti broja karioznih, izvađenih i saniranih zubi s KEP-indeksom, nema bitne razlike u sastavu KEP-indeksa studentskih populacija Sveučilišta u Splitu generiranih MCS-om.
3. Prosječne vrijednosti KEP-indeksa simuliranih uzoraka u svakom od tri MCS-a su bile normalno raspodijeljene.
4. Prosječne vrijednosti KEP-indeksa studentskih populacija Sveučilišta u Splitu generiranih MCS-om razlikovale su se za manje od jedanog boda.

## **7. LITERATURA**

1. Rudolph JE, Zhong Y, Duggal P, Mehta SH, Lau B. Defining representativeness of study samples in medical and population health research. *BMJ Med.* 2023;2(1):e000399.
2. Lim CY, In J. Erratum: Randomization in clinical studies. *Korean J Anesthesiol.* 2019;72(4):396.
3. Lim CY, In J. Randomization in clinical studies. *Korean J Anesthesiol.* 2019;72(3):221-32.
4. Gupta KK, Attri JP, Singh A, Kaur H, Kaur G. Basic concepts for sample size calculation: Critical step for any clinical trials! *Saudi J Anaesth.* 2016;10(3):328-31.
5. Serdar CC, Cihan M, Yucel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochem Med (Zagreb).* 2021;31(1):010502.
6. Pincus T, Bergman MJ, Maclean R, Yazici Y. Complex measures and indices for clinical research compared with simple patient questionnaires to assess function, pain, and global estimates as rheumatology "vital signs" for usual clinical care. *Rheum Dis Clin North Am.* 2009;35(4):779-86, ix.
7. Radic M, Benjak T, Vukres VD, Rotim Z, Zore IF. Presentation of DMFT/dmft Index in Croatia and Europe. *Acta Stomatol Croat.* 2015;49(4):275-84.
8. Cvitanović B. Procjena oralnog zdravlja studenata Medicinskog fakulteta Sveučilišta u Splitu temeljem KEP indeksa [diplomski rad]. Split: Medicinski fakultet Sveučilišta u Splitu; 2022.
9. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol.* 2013;42(4):1012-4.
10. Taherdoost H. Sampling methods in research methodology; how to choose a sampling technique for research. *Int J Acad Res in Manag.* 2016;5(2):18-27.
11. MIT OpenCourseWare. Introduction to computational thinking and data science - Confidence intervals [Internet]. [citirano 18 kolovoza 2023]. Dostupno na: [https://www.youtube.com/watch?v=rUxP7TM8-wo&list=PLU14u3cNGP619EG1wp0kT-7rDE\\_Az5TNd&index=8&t=9s](https://www.youtube.com/watch?v=rUxP7TM8-wo&list=PLU14u3cNGP619EG1wp0kT-7rDE_Az5TNd&index=8&t=9s).
12. Dirk P. Kroese TB, Thomas Taimre, Zdravko i. Botev. Why the Monte Carlo method is so important today? *Wiley Interdiscip Rev Comput Stat.* 2014;6(6):386-92.
13. Cohen J. Multiple regression as a general data-analytic system. *Psychol Bull.* 1968;70(6):426-43.

14. Delić I., Usklađenost diplomskih radova na Studiju dentalne medicine u Splitu s IMRaD strukturom izvornog znanstvenog članka [diplomski rad]. Split: Medicinski fakultet Sveučilišta u Splitu; 2022.
15. Delić I, Kero D. How much are the IMRaD structures of diploma theses and original research articles similar? A cross-sectional analysis of diploma theses from the Dental Medicine study in Split. *ST-OPEN*. 2023;4:e2023:2208:6.
16. Ivica A, Galic N. Attitude towards Oral Health at Various Colleges of the University of Zagreb: A Pilot Study. *Acta Stomatol Croat*. 2014;48(2):140-6.
17. Vranjković H. Radiografska procjena dentalnog statusa adolescenata Splitsko-dalmatinske županije [diplomski rad]. Split: Medicinski fakultet Sveučilišta u Splitu; 2016.
18. Tepić A. Oralno zdravlje stomatologa [diplomski rad]. Zagreb: Stomatološki fakultet Sveučilišta u Zagrebu; 2017.
19. Jurković A. Epidemiološka obilježja karijesa kod studenata Stomatološkog fakulteta u Zagrebu [diplomski rad]. Zagreb: Stomatološki fakultet Sveučilišta u Zagrebu; 2018.
20. Domazet I. Oralno zdravlje studenata prve i pete godine studija Dentalne medicine Medicinskog fakulteta Sveučilišta u Splitu [diplomski rad]. Split: Medicinski fakultet Sveučilišta u Splitu; 2022.
21. Yan F, Robert M, Li Y. Statistical methods and common problems in medical or biomedical science research. *Int J Physiol Pathophysiol Pharmacol*. 2017;9(5):157-63.
22. Hannigan A, Lynch CD. Statistical methodology in oral and dental research: pitfalls and recommendations. *J Dent*. 2013;41(5):385-92.
23. Pandis N, Polychronopoulou A, Eliades T. An assessment of quality characteristics of randomised control trials published in dental journals. *J Dent*. 2010;38(9):713-21.
24. Limpert E, Stahel WA. Problems with using the normal distribution--and ways to improve quality and efficiency of data analysis. *PLoS One*. 2011;6(7):e21403.
25. Szucs D. A Tutorial on Hunting Statistical Significance by Chasing N. *Front Psychol*. 2016;7:1444.

## **8. SAŽETAK**

### **Ciljevi istraživanja:**

Cilj ovog istraživanja bio je generirati podatke za KEP-indeks čitave studentske populacije Sveučilišta u Splitu iz tri uzorka različite veličine korištenjem *Monte Carlo* metode.

### **Materijali i metode:**

Podatci za potrebe istraživanja prikupljeni su u sklopu kliničkih vježbi na kolegijima „Restaurativna dentalna medicina 2“ te „Endodoncija 2“. Svi ispitanici (n =200) su bili studenti Sveučilišta u Splitu koji su došli na pregled ili liječenje u ordinacije fakultetske poliklinike za dentalnu medicinu „Dental academicus“. Iz prikupljenih podataka o KEP-indeksu izvedena su tri uzorka – mali uzorak (n =50), uzorak srednje veličine (n = 100) i cjelokupni uzorak (n = 200). Iz navedenih uzoraka *Monte Carlo* simulacijama (MCS) su generirani podatci o KEP-indeksu za tri studentske populacije (MCS 50, MCS 100 i MCS 200) od kojih je svaka bila veličine 20 000 pojedinaca. Vjerojatnosti pojedinih ishoda za KEP-indeks i njegove sastavnice, kao i međusobna povezanost KEP-indeksa i sastavnica KEP-indeksa su ubačeni kao pretpostavke za MCS.

### **Rezultati:**

Procjena za prosječni KEP-indeks studentske populacije prema MCS 50 je  $8,96 \pm 0,69$  (99% CI [8,91; 9]), prema MCS 100  $9,12 \pm 0,47$  (99% CI [9,08; 9,21]), te prema MCS 200  $8,82 \pm 0,36$  (99% CI [8,77; 8,87]). Od sastavnica KEP-indeksa, broj saniranih zubi (P) je u sve tri simulacije bio najjače povezan s KEP-indeksom i to  $R = 0,84$  (MCS 50),  $R = 0,82$  (MCS 100) i  $R = 0,77$  (MCS 200). Procjene prosječnog KEP-indeksa MCS-om odstupale su za manje od 1 boda u odnosu na procjene prosječnog KEP-indeksa iz pripadajućih uzoraka i međusobno.

### **Zaključci:**

*Monte Carlo* metoda može biti korisna u procjeni prosječnih populacijskih vrijednosti kliničkih indeksa u dentalnoj medicini uključujući i KEP-indeks. Prema postavkama MCS u ovom istraživanju, veličina izvornih uzoraka nije bitnije utjecala na završnu procjenu parametara KEP-indeksa.

## **9. SUMMARY**



**Diploma thesis title:** How to use *Monte Carlo* method to measure DMFT index of the entire student population of the University of Split in one afternoon?

**Objectives:**

The aim of this study was to generate data for the DMFT (Decayed, Missing, and Filled Permanent Teeth) index of the entire student population of the University of Split from three samples of different sizes using *Monte Carlo* method.

**Materials and methods:**

Data for this study were collected during clinical exercises in the courses "Restorative Dental Medicine 2" and "Endodontics 2". All participants (n = 200) were students of the University of Split who came for an examination or treatment in the offices of the faculty polyclinic for dental medicine "Dental academicus". Three samples were derived from the collected data on the DMFT index - a small sample (n = 50), a medium-sized sample (n = 100) and the entire sample (n = 200). *Monte Carlo* simulations (MCS) were used to obtain DMFT index data for three student populations (MCS 50, MCS 100, and MCS 200) of 20,000 individuals each from the above samples. The probabilities of individual outcomes for the DMFT index and its components, as well as the mutual correlation between the DMFT index and the components of the DMFT index, were entered as assumptions for the MCS.

**Results:**

The estimated mean DMFT index of the student population by MCS 50 was  $8.96 \pm 0.69$  (99% CI [8.91; 9]), by MCS 100 was  $9.12 \pm 0.47$  (99% CI [9.08; 9.21]), and by MCS 200 was  $8.82 \pm 0.36$  (99% CI [8.77; 8.87]). Regarding the components of the DMFT index, the number of repaired teeth (F) in MCS 50, MCS 100, and MCS 200 was most strongly correlated with the DMFT index, namely  $R = 0.84$ ,  $R = 0.82$ , and  $R = 0.77$ , respectively. The estimated mean DMFT indices by MCS differed from the mean DMFT indices of the corresponding samples and from each other by less than 1 point.

**Conclusions:**

The *Monte Carlo* method may be useful in estimating the population means of clinical indices in dental medicine, including the DMFT index. According to the assumptions made for MCS in this

study, the size of the original samples did not significantly affect the final estimate of the parameters of the DMFT index.

## **10. ŽIVOTOPIS**

#### OSOBNI PODATCI:

Ime i prezime: Dora Dodig

Datum i mjesto rođenja: 4. prosinca 1998. godine, Split

Državljanstvo: hrvatsko

Elektronska pošta: dora-dodig@hotmail.com

#### OBRAZOVANJE:

2005. – 2013. Osnovna škola „Spinut“, Split

2013. – 2017. Prva gimnazija Split, jezični smjer

2017. – 2023. Medicinski fakultet Sveučilišta u Splitu, Studij dentalne medicine, Split

MATERINSKI JEZIK: hrvatski

STRANI JEZICI: engleski, njemački, španjolski

DRUGA POSTIGNUĆA: osvojen kviz Potjera (2017.), peto mjesto na Svjetskom prvenstvu za kvizaše do 23. godine (2019.)