

**SVEUČILIŠTE U SPLITU
MEDICINSKI FAKULTET**

Jelena Šuto Pavičić

**JEZIČNA ANALIZA LAIČKIH I ZNANSTVENIH SAŽETAKA COCHRANE
SUSTAVNIH PREGLEDA ONKOLOŠKIH INTERVENCIJA I UPORABA
UMJETNE INTELIGENCIJE U IZRADI LAIČKIH SAŽETAKA**

Doktorski rad

Split, 2025.

**SVEUČILIŠTE U SPLITU
MEDICINSKI FAKULTET**

Jelena Šuto Pavičić

**JEZIČNA ANALIZA LAIČKIH I ZNANSTVENIH SAŽETAKA COCHRANE
SUSTAVNIH PREGLEDA ONKOLOŠKIH INTERVENCIJA I UPORABA
UMJETNE INTELIGENCIJE U IZRADI LAIČKIH SAŽETAKA**

Doktorski rad

Mentor: prof. dr. sc. Ana Marušić, dr. med.

Split, 2025.

Doktorski rad izrađen je na Medicinskom fakultetu Sveučilišta u Splitu pod voditeljstvom prof. dr. sc. Ane Marušić i dio je formalnog istraživačkog projekta Hrvatske zaklade za znanost ProDeM, pod naslovom “Profesionalnost u zdravstvu: odlučivanje u praksi i znanosti” (projekt IP-2019-04-4882).

Voditeljica rada: prof. dr. sc. Ana Marušić, dr. med

Objavljeni znanstveni radovi na kojemu se temelji doktorski rad:

Šuto J, Marušić A, Buljan I. Linguistic analysis of plain language summaries and corresponding scientific summaries of Cochrane systematic reviews about oncology interventions. *Cancer Med.* 2023;12(9):10950-10960. (čimbenik odjeka 4,711)

Šuto Pavičić J, Marušić A, Buljan I. Using ChatGPT to Improve the Presentation of Plain Language Summaries of Cochrane Systematic Reviews About Oncology Interventions: Cross-Sectional Study. *JMIR Cancer* 2025;11:e63347. (čimbenik odjeka 3,3).

ZAHVALA

Hvala mojoj mentorici prof. dr. sc. Ani Marušić-ne samo što mi je pokazala što znači biti istinski znanstvenik, nego još i više, što mi je kroz svaki korak ovog puta pokazivala što znači biti čovjek. Od studentskih dana pa sve do danas, bila je moj tihi oslonac i primjer te je svojim djelima pokazivala što znači kad se spoje integritet, strpljivost i nesebičnost. Riječi ne mogu izraziti zahvalnost koju osjećam što nikad nije prestajala vjerovati u ono što ja ponekad sama više nisam mogla vidjeti.

Hvala doc. dr. sc. Ivanu Buljanu na kontinuiranoj podršci, razumijevanju i što je uvijek imao vremena za svaku moju dilemu i pitanje.

Hvala prof. dr. sc. Liviji Puljak, na predanosti, entuzijazmu i iskrenoj želji da nas već kao studente približi svijetu znanosti, kao i na dostupnosti i podršci tijekom cijelog doktorskog studija.

Hvala svim djelatnicima Klinike za onkologiju i radioterapiju KBC Split, s predstojnikom prof. dr. sc. Eduardom Vrdoljakom, koji su bili dostupni za razgovor i spremni podijeliti svoja iskustva i savjete kroz specijalizaciju i doktorski studij.

Hvala Tini Aljinović Baleti za sve one trenutke kad je bila tu. Kad su dani bili teški, tvoje prisustvo je bilo tu, tiho, ali nepokoljebljivo. Hvala ti što si uvijek imala vremena i nikad nisi ustuknula pred mojim pitanjima i neizvjesnostima te što si s lakoćom tu ne samo za mene, nego i za moju užu obitelj, potom širu, kao i najbliže prijatelje-hvala ti što si takva kakva jesi. Hvala mojim stijenama-Petri, Eli, Ani, Ivi, Denisu, Dori, Antoniji-teško je u nekoliko rečenica sročiti što znači imati ljude koji će s tobom ploviti i po najburnijem moru, veseliti se tvojim uspjesima više nego ti sam, a vjerovati u tebe kad i ti sam sumnjaš u sebe. Hvala vam što ste oslonac koji ne moram tražiti, jer ste tu i prije nego što pomislim da vas trebam.

Hvala mom bratu Dominiku što uvijek zna kako mi vratiti osmijeh na lice i podsjetiti me kako se lakše diše. I što si uvijek tu— bez velikih riječi, ali s najvećim srcem.

Hvala mojim roditeljima -mojoj mami Mariji i pokojnom ocu Viktoru, kojem posvećujem ovaj rad, kao i svaki svoj uspjeh. Hvala mojoj mami što me razumije i poznaje više nego ja sama sebe, a tati na svakom razgovoru-tvoje riječi, snaga i ljubav koju si dao našoj obitelji je ono što mene vodi danas i svaki sljedeći dan. Ljubav i podrška mojih roditelja je iza svega dobroga što napravim.

Hvala mom Ivanu-ne postoje riječi koje bi bile dovoljne za tebe. Ča je brodu feral na provi...

SADRŽAJ

1.	UVOD.....	1
1.1	Medicina temeljena na dokazima	2
1.2	Sustavni pregled	3
1.3	Cochrane	4
1.4	Cochrane laički sažetak	5
1.5	Zdravstvena pismenost	6
1.6	Dostupnost onkoloških informacija općoj populaciji	8
1.7	Umjetna inteligencija i njena primjena u medicini.....	10
2.	CILJ RADA I HIPOTEZE.....	15
2.1	Jezična analiza laičkih i pripadajućih znanstvenih sažetaka Cochrane sustavnih pregleda onkoloških intervencija	16
2.2	Korištenje umjetne inteligencije za poboljšanje prezentacije laičkih sažetaka Cochraneovih sustavnih pregleda o onkološkim intervencijama.....	16
3.	METODE.....	18
3.1	Jezična analiza laičkih i pripadajućih znanstvenih sažetaka Cochraneovih sustavnih pregleda onkoloških intervencija	19
3.1.1	Ustroj.....	19
3.1.2	Ishodi istraživanja	19
3.1.3	Prikupljanje podataka.....	19
3.1.4	Probir sažetaka i vađenje podataka	19
3.1.5	Obrada podataka i mjerni instrumenti.....	20
3.1.5.1	SMOG indeks	20
3.1.5.2	Jezična obilježja	21
3.1.5.3	Zaključivost	21
3.1.6	Statistički postupci.....	22
3.2	Korištenje umjetne inteligencije u poboljšanju laičkih sažetaka Cochrane sustavnih pregleda onkoloških informacija.....	23
3.2.1	Ustroj studije	23
3.2.2	Uzorak	23
3.2.3	Generiranje PLS-ova pomoću ChatGPT-a.....	24
3.2.4	Obrada podataka	26
3.2.4.1	SMOG indeks	26

3.2.4.2	Jezična obilježja	27
3.2.4.3	Kategorija zaključivosti.....	27
3.2.5	Statistička analiza	28
3.2.5.1	Deskriptivna statistika	28
3.2.5.2	Usporedba grupa	29
3.2.5.3	Statistički alati.....	29
4.	REZULTATI.....	30
	Jezična analiza laičkih i pripadajućih znanstvenih sažetaka Cochrane sustavnih pregleda onkoloških intervencija.....	31
4.1.1	Opis uzorka.....	31
4.1.2	Analiza čitkosti	32
4.1.3	Analiza jezičnih obilježja	33
4.1.4	Analiza zaključivosti.....	35
4.1.5	Analiza čitkosti u odnosu prema zaključivosti	39
4.1.6	Analiza jezičnih obilježja prema zaključivosti	40
4.1.7	Sumarizacija usporedbe laičkih i znanstvenih sažetaka.....	42
4.2	Korištenje umjetne inteligencije za poboljšanje prezentacije laičkih sažetaka Cochraneovih sustavnih pregleda o onkološkim intervencijama.....	44
4.2.1	SMOG indeks i lingvističke karakteristike	44
4.2.2	Zaključivost	46
5.	RASPRAVA	47
5.1	Jezična analiza Cochrane laičkih i znanstvenih sažetaka sustavnih pregleda onkoloških intervencija.....	48
5.1.1	Glavni rezultati i njihov značaj.....	48
5.1.2	Ograničenja istraživanja.....	51
5.2	Uporaba umjetne inteligencije u stvaranju laičkih sažetaka	51
5.2.1	Glavni rezultati	51
5.2.2	Ograničenja istraživanja.....	53
5.3	Zaključci i smjerovi za budućnost.....	54
6.	ZAKLJUČCI.....	56
7.	LITERATURA.....	58
8.	SAŽETAK	71
9.	LAIČKI SAŽETAK	73
10.	SUMMARY	75

11.	PLAIN LANGUAGE SUMMARY	77
12.	ŽIVOTOPIS	79
13.	DODATAK.....	81
13.1	Prilog 1. Tablica laičkih sažetaka	82

POPIS OZNAKA I KRATICA

AI – umjetna inteligencija (engl. *artificial intelligence*)

ARI – automatski indeks čitkosti (engl. *Automated Readability Indeks*)

CDSR – Cochrane baza podataka sustavnih pregleda (engl. *Cochrane Database of Systematic Reviews*)

ChatGPT – razgovorni generativni prethodno istrenirani transformer (engl. *Chat generative pre-trained transformer*)

CI – raspon pouzdanosti (engl. *confidence interval*)

CLI – Coleman–Liau Index (engl. *Coleman-Liau indeks*)

CNN – konvolucijska neuronska mreža (engl. *convolutional neural network*)

CRG – Cochrane pregledna skupina (engl. *Cochrane Review Group*)

CSR – Cochrane sustavni pregled (engl. *Cochrane systematic review*)

DL – duboko učenje (engl. *deep learning*)

DLT – tehnologija decentraliziranih evidencija (engl. *distributed ledger technology*)

EBM – medicina utemeljena na dokazima (engl. *evidence based medicine*)

FKGL – Flesh-Kincaid razina (engl. *Flesch-Kincaid Grade Level*)

GFI – Gunning Fog indeks (engl. *Gunning Fog Indeks*)

IQR – interkvartilni raspon (engl. *interquartile range*)

LIWC – Jezična analiza i brojanje riječi (engl. *Linguistic Inquiry and Word Count*)

LLM –veliki jezični model (engl. *Large language model*)

ML – strojno učenje (engl. *machine learning*)

NN – neuronska mreža (engl. *neural network*)

PICO – pacijent/populacija, intervencija, usporedna skupina, ishod (engl. *patient/population, intervention, comparison, outcome*)

PLEACS – Standardi za pisanje laičkih sažetaka u novim Cochrane pregledima intervencija (engl. *Standards for the reporting of Plain Language Summaries in new Cochrane Intervention Reviews*)

PLS – laički sažetak (engl. *plain language summary*)

RCT – randomizirano kontrolirano ispitivanje (engl. *randomized controlled trial*)

SA – znanstveni sažetak (engl. *scientific abstract*)

SMOG – Jednostavna mjera čitkosti (engl. *Simple Measure of Gobbledygook*)

STROBE – Smjernice za pisanje izvješća opažajnih studija (engl. *Strengthening the Reporting of Observational Studies in Epidemiology*)

SZO – Svjetska zdravstvena organizacija; odnosno WHO – engl. *World Health Organization*

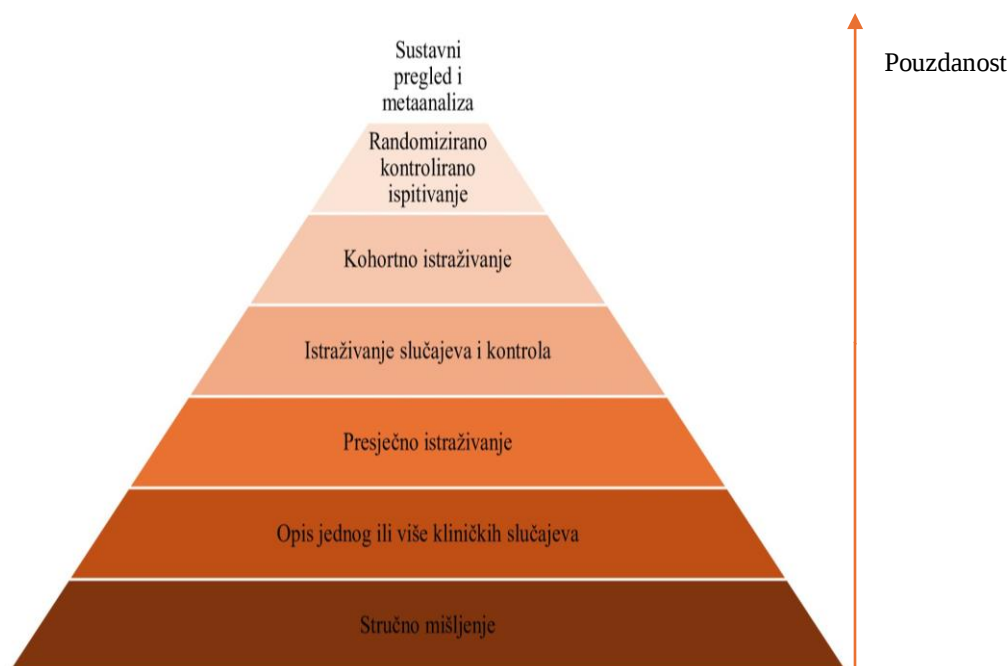
1. UVOD

1.1 Medicina temeljena na dokazima

Medicina temeljena na dokazima (engl. *Evidence-based medicine-EBM*) je pristup koji se temelji na integraciji najboljih dostupnih znanstvenih dokaza, kliničkog iskustva liječnika i mišljenja pacijenata, u svrhu poboljšanja donošenja odluka u liječenju (1,2). Ovaj pristup počeo se razvijati tijekom 1970-ih i 1980-ih, kada su znanstvenici poput Davida Sacketta, Gordon Guyatta i Archieja Cochranea naglasili važnost oslanjanja medicinske prakse na čvrstim empirijskim dokazima te predložili početna pravila za usmjeravanje kliničkih odluka (3–6). Kad se zdravstveni djelatnik u radu s pacijentom nađe pred dilemom, u medicini utemeljenoj na dokazima se započne s kliničkim pitanjem, koje oslikava problem s kojim se zdravstveni djelatnik suočava (1,7). Nakon što se kliničko pitanje formulira, traže se relevantni znanstveni dokazi koji se odnose njega i koji bi pomogli u njegovom rješavanju (1). Ti znanstveni dokazi mogu uključivati rezultate studija, ali isto tako i stručna mišljenja (1). Međutim, nisu svi podaci jednako pouzdani (1,5). Temelj medicine utemeljene na dokazima počiva na kliničkim studijama (5). Ovisno o snazi različitih kliničkih istraživanja, razvijene su različite hijerarhijske piramide (5), ali pet razina dokaza koje je definirao Sackett i dalje čine osnovu većine smjernica (8). Pet razina dokaza uključuje:

- (i) razina I – opsežna randomizirana kontrolirana istraživanja (engl. *randomized controlled trial-RCT*) ispitivanja s preciznim rezultatima;
- (ii) razina II – manja RCT ispitivanja s nejasnim rezultatima;
- (iii) razina III – kohortne studije i studije slučaja i kontrole;
- (iv) razina IV – povijesne kohortne studije ili studije slučaja i kontrole;
- (v) razina V – serije slučajeva bez kontrolne skupine (8).

Slikovito, hijerarhiju dokaza možemo prikazati i kao piramidu-bazu piramide čine istraživanja s najnižom razinom dokaza, dok svaka sljedeća razina piramide predstavlja sve pouzdanije dokaze (Slika 1). Sva navedena istraživanja pripadaju kategoriji primarnih istraživanja. Penjući se prema gore, dolazimo do najpouzdanijih dokaza - sustavnih pregleda, koji su oblik sekundarnih istraživanja jer sažimaju dokaze iz primarnih studija (9,10). Sustavni pregledi randomiziranih kontroliranih ispitivanja i visokokvalitetni RCT-ovi ostaju ključni temelji EBM-a (6,11,12).



Slika 1. Razine pouzdanosti dokaza u području medicine.

1.2 Sustavni pregled

Za razliku od pojedinačnih studija, sustavni pregled obuhvaća širi skup podataka jer uključuje i uspoređuje rezultate iz više relevantnih izvora. Za cilj ima pregledati sve studije na danu temu i po dovršenoj obradi, donijeti zaključak o učinkovitosti ispitivane intervencije. Provođenju ovakvih studija prethode detaljno razvijene upute za provođenje istih te se definiraju kriteriji kako bi se smanjio rizik od sustavne pogreške (10,13). Objedinjeni rezultati više primarnih istraživanja se često dodatno kvantitativno analiziraju statističkim metodama te na taj način dobivamo meta-analizu (13). Proces stvaranja sustavnog pregleda se sastoji od nekoliko koraka, a prvi je definiranje kliničkog pitanja i njegova formulacija prema PICO modelu (pacijenti, intervencija, usporedba, ishod; engl. *patient, intervention, comparison, outcome*) (7,10). Prilikom analize razmatraju se svi relevantni ishodi povezani s postavljenim kliničkim pitanjem. Sljedeći korak uključuje pretraživanje i evaluaciju postojećih sustavnih pregleda koji se bave istom temom kako bi se utvrdilo hoće li novi pregled donijeti dodatnu vrijednost znanstvenoj zajednici (10,14). Ako se zaključi da je novi sustavni pregled opravdan, osmišljava se i provodi strategija pretraživanja (10). Slijedi pregled literature, koji uključuje pretraživanje elektroničkih baza podataka, registara kliničkih ispitivanja, ne-engleskih izvora i sive literature (10,13,15). Također je važno uzeti u obzir neobjavljene studije i izvore koje prepoznaju stručnjaci u tom području (13,16). Nakon

identifikacije relevantnih studija slijedi postupak selekcije (10,17). Prvo se procijeni koje studije imaju potencijal za uključivanje, nakon čega slijedi detaljnija selekcija na temelju kriterija poput dizajna studije, veličine uzorka te cjelovitosti podataka (10,13). Nakon odabira studija, potrebno je procijeniti njihovu kvalitetu, što uključuje unaprijed definirane kriterije i sustav bodovanja (10,15,18). Važni čimbenici u procjeni kvalitete uključuju nasumičnost dodjele u skupine, sličnost skupina na početku istraživanja te primjenu zaslijepljivanja kod istraživača, zdravstvenih djelatnika i pacijenata (18,19). Nakon procjene kvalitete studija, izračunavaju se i kombiniraju ishodi (18), a metoda analize ovisi o vrsti podataka (20). Završni korak je interpretacija rezultata, koja se najčešće prikazuje pomoću forest dijagrama (*engl. forest plot*) (13,19). Ovaj grafički prikaz omogućuje jednostavan pregled rezultata pojedinačnih studija te ukupne procjene kombiniranog učinka, čineći ga dobrim alatom za vizualizaciju rezultata meta-analize (19). Iz svega navedenog je razvidno da sustavni pregledi zahtijevaju vremensko ulaganje i trud u izradu istih, kako bi pružili najbolji mogući dokaz o ispitivanoj temi. Najpoznatija institucija koja pruža podršku u izradi sustavnih pregleda je Cochrane (21,22). Cochrane je i formulirao posebne smjernice za izradu sustavnih pregleda, a sustavni radovi koji se izrade prema tim uputama se onda nazivaju Cochrane sustavni pregledi (*engl. Cochrane systematic review-CSR*) (23,24). Dostupni su preko mrežne platforme Cochrane Library (21,23).

1.3 Cochrane

Začeci Cochrane organizacije sežu u sedamdesete godine 20. stoljeća, kada je britanski epidemiolog Archie Cochrane istaknuo važnost donošenja kliničkih odluka temeljenih na dokazima (23,25). Naglasio je potrebu za izradom sažetih pregleda relevantnih randomiziranih kliničkih ispitivanja na određene teme. Tijekom sedamdesetih i osamdesetih godina koncept medicine utemeljene na dokazima dodatno se razvijao, a Cochraneove ideje, zajedno s doprinosima drugih znanstvenika, dovele su do toga da je 1992. godine britanski Nacionalni program za istraživanje i razvoj u zdravstvu osigurao sredstva za osnivanje „Cochrane Centra“. Cilj ovog centra bio je globalno unaprijediti izradu sustavnih pregleda u svim područjima medicine (23,25,26). Već sljedeće godine, 1993., uslijedio je međunarodni odgovor na Cochraneovu inicijativu, što je rezultiralo osnivanjem organizacije Cochrane Collaboration (23,27). Ova neprofitna organizacija posvećena je pomaganju ljudima u donošenju informiranih odluka o zdravlju kroz sustavne preglede koji su svima dostupni (28). Priprema i ažuriranje Cochraneovih pregleda odgovornost je međunarodnih suradničkih

skupina za preglede- Cochraneove skupine za preglede (engl. *Cochrane Review Groups – CRGs*) (29). Skupine podržavaju autore u izradi visokokvalitetnih sustavnih pregleda unutar njihovih stručnih područja (23,30). Svaka skupina usmjerena je na specifično područje istraživanja i vođena uredničkim timom (31). Trenutno postoji 28 takvih skupina, koje autorima pružaju metodološku i uredničku podršku te upravljaju uredničkim procesom (31,32). Članovi određenih Cochrane preglednih skupina nemaju sukob interes koji bi bio vezan uz područje djelovanja i proučavanja (32). Sustavni pregledi istraživanja iz područja zdravstvene skrbi i zdravstvene politike objavljuju se elektronički u Cochraneovoj bazi sustavnih pregleda (33). U trenutku provođenja ovih studija, Cochraneova mreža za rak (engl. *Cochrane Cancer Network*) bila je usmjerena na osiguravanje pouzdanih dokaza koji pomažu u donošenju ključnih odluka o dijagnostici i liječenju raka (31). U trenutku provođenja studije, Cochrane mreža za rak se sastojala od sedam skupina koje su bile usmjerene na liječenje zloćudnih tumora dojke, kolorektuma, pluća, hematoloških zloćudnih bolesti, ginekoloških tumora, tumora središnjeg živčanog sustava, rijetkih tumora te zloćudnih tumora u djece. U toku 2023. je došlo do promjena te su sada aktivne skupine relevantne za ovu disertaciju Cochrane skupina za Rak dojke (*Breast Cancer*); Cochrane skupina za urologiju (*Urology*); Cochrane skupina koja se odnosi na bolesti debelog crijeva (*Colorectal Group*); Cochrane skupina za ginekologiju i plodnost (*Gynaecology and Fertility Group*), Cochrane skupina za hematologiju (*Hematology Group*), Cochrane skupina Rak pluća (*Lung Cancer*). (34). Važno je napomenuti i da se u trenutku pisanja disertacije Cochraneova mreža za rak više ne postoji kao takva (35), nego su aktivne gore spomenute Cochrane skupine.

1.4 Cochrane laički sažetak

Sustavni pregledi predstavljaju najveću razinu dokaza jer pružaju temeljitu i sustavnu analizu literature na određeno istraživačko pitanje (19). Time olakšavaju pristup relevantnim dokazima i njihovu primjenu u praksi, posebno kliničarima s ograničenim vremenom te stručnjacima koji izrađuju kliničke smjernice u različitim granama medicine. Kao i većina znanstvenih radova, posebno uzimajući u obzir da su sustavni pregledi često opsežni i složeni, ključne informacije koje prenosi rad su sadržane u sažetku, tj. abstraktu (36). U praksi, sažetak je često jedini dio znanstvenog članka koji je lako dostupan tijekom pretraživanja, što ga čini najčešće čitanim dijelom i ima značajnu ulogu (36). Kao način prezentiranja znanstvenih informacija iz sustavnog pregleda, Cochrane sustavni pregledi, iznose dva tipa sažetaka: standardni znanstveni te sažetak namijenjen općoj populaciji-laički

sažetak (engl. *Plain language summary*-PLS) (37). Ovi sažeci trebali bi biti napisani na jasan i razumljiv način te lako dostupni osobama koje nisu eksperti za dano područje medicine (37). Kako bi se osigurala ujednačenost i kvaliteta ovih sažetaka, Cochrane je uspostavio smjernice pod nazivom *Standards for the reporting of Plain Language Summaries in new Cochrane Intervention Reviews* (PLEACS), koje definiraju standardizirani format njihova pisanja (38). Međutim, pokazano je da se autori članaka često ne pridržavaju ovih standarda te da postoji još mjesta za napredak (37). Engleski jezik je primarni jezik u pisanju sustavnih pregleda. Kako bi se poboljšala dostupnost visokokvalitetnih znanstvenih dokaza općoj populaciji, Cochrane laički sažeci se prevode na ostale strane jezike, uz pomoć volontera (39,40). Trenutno globalno postoji 20 prevoditeljskih timova koji osiguravaju prijevode na kineski (jednostavni i tradicionalni), španjolski, poljski hrvatski, ruski, nizozemski, francuski, njemački, indonezijski, japanski, farsi, hindi, korejski, mađarski, malajski, portugalski, rumunjski, , tajlandski i tamil (40).

1.5 Zdravstvena pismenost

Postoji više definicija zdravstvene pismenosti, ali sve u svojoj osnovi naglašavaju da se radi o znanju i sposobnostima koje osobi omogućavaju da stekne, razumije i koristi relevantne medicinske informacije kako bi mogao donositi odluke temeljena na istima i odgovorno upravljati vlastitim zdravstvenim stanjem (41). Svjetska zdravstvena organizacija (SZO, engl. *World Health Organization*, (WHO)) naglašava da je zdravstvena pismenost ključni faktor za poboljšanje zdravstvenih ishoda (42). Zdravstvena pismenost obuhvaća više dimenzija: i) Funkcionalna zdravstvena pismenost – osnovna sposobnost čitanja i razumijevanja zdravstvenih informacija, kao što su upute o lijekovima i zdravstveni letci (43), ii) Interaktivna zdravstvena pismenost – sposobnost primjene stečenih znanja kroz komunikaciju sa zdravstvenim djelatnicima i donošenje informiranih odluka (44), iii) Kritička zdravstvena pismenost – sposobnost evaluacije zdravstvenih informacija, prepoznavanja dezinformacija i donošenja odluka temeljenih na pouzdanim izvorima (45,46).

Čitkost (engl. *readability*) označava razinu lakoće kojom čitatelj može razumjeti neki tekst, a posebno je važna u zdravstvenoj komunikaciji gdje se informacije prenose općoj populaciji (47). Mjeri se pomoću različitih formula koje analiziraju jezične elemente poput broja riječi, slogova i rečenica (48). Formule za čitkost omogućuju objektivnu procjenu pisanih zdravstvenih informacija procjenjujući koliko je godina formalnog obrazovanja potrebno čitatelju da razumije sadržaj (48). Među najčešće korištenim dostupnim alatima nalaze se *Gunning Fog Index* (GFI), *Simple Measure of Gobbledygook* (SMOG), *Flesch-*

Kincaid Grade Level (FKGL), *Coleman–Liau Index* (CLI) i *Automated Readability Index* (ARI) (48). Svaka od ovih formula koristi različite kriterije za određivanje razine čitanja (48,49). SMOG formula se smatra najpouzdanijom za medicinske tekstove jer je posebno osjetljiva na složene riječi (višesložne riječi, s tri ili više slogova), koje su česte u zdravstvenoj terminologiji, a njezin rezultat precizno odražava broj godina obrazovanja potreban za potpuno razumijevanje teksta (49). SMOG se jednostavno primjenjuje i u svom izračunu prvo odabere tri uzorka iz teksta, pri čemu svaki uzorak sadrži po deset rečenica. U tim rečenicama prebroje se riječi koje imaju tri ili više slogova (višesložne riječi). Zatim se računa razina čitkosti pomoću ove formule:

$$\text{razina čitkosti} = 1.0430 \times \sqrt{(\text{broj polisilabičnih riječi} \times 30 / \text{broj rečenica})} + 3.1291.$$

Ova verzija formule ima i pojednostavljeni način izračuna koji se lako može koristiti i bez potrebe za kalkulatorom: prebroje se višesložne riječi u tri uzorka po deset rečenica, potom se uzme kvadratni korijen najbližeg savršenog kvadrata tog broja te se na taj broj doda 3. Dobiveni broj je procjena potrebne razine obrazovanja (u godinama školovanja) za razumijevanje teksta (50).

Postoji i rad na temu razvijanja SMOG indeksa za tekstove napisane na hrvatskom jeziku- SMOG-Cro, čija formula je razina čitkosti= $2 + \sqrt{4 + \text{slogova}}$, pri čemu rezultat, kao i u originalnom SMOG indeksu, pokazuje broj godina obrazovanja potreban da bi osoba mogla razumjeti određeni tekst. Predložena metodologija mogla bi pomoći primjeni u praksi, tj. procjeni zdravstvene pismenosti na hrvatskom jeziku, edukaciji pacijenata i poboljšanju komunikacije u hrvatskom zdravstvu (51).

Niska razina zdravstvene pismenosti povezana je sa slabijim pridržavanjem terapijskih uputa, povećanim opterećenjem zdravstvenog sustava i sveukupno lošijim zdravstvenim ishodima (52,53). Osobe s ograničenom zdravstvenom pismenošću često imaju poteškoća s razumijevanjem medicinskih uputa, što može rezultirati nepravilnom primjenom lijekova, izbjegavanjem preventivnih pregleda i kasnim prepoznavanjem simptoma bolesti (53–56). S druge strane, visoka razina zdravstvene pismenosti omogućava pojedincima da aktivno sudjeluju u svom liječenju, bolje komuniciraju s liječnicima te donose odluke koje pozitivno utječu na njihovo zdravlje (57). Poboljšana zdravstvena pismenost također može smanjiti zdravstvene troškove, jer omogućava rano prepoznavanje i liječenje bolesti, čime se smanjuje potreba za hospitalizacijom i hitnim intervencijama (58–61).

Unatoč važnosti zdravstvene pismenosti, mnogi pojedinci se suočavaju s nizom prepreka: složeni medicinski jezik – stručni izrazi često nisu prilagođeni pacijentima, što otežava razumijevanje informacija; digitalni jaz – nejednak pristup digitalnim zdravstvenim

resursima može dodatno produbiti zdravstvene nejednakosti; dezinformacije i lažne vijesti – širenje netočnih medicinskih informacija putem društvenih mreža može negativno utjecati na odluke pacijenata; kulturne i jezične barijere – osobe koje ne govore službeni jezik zemlje mogu imati poteškoće u pristupu pouzdanim informacijama (62–64). Nekoliko studija pokazalo je slične rezultate, otkrivajući da mnogi pacijenti osjećaju sram zbog svojih ograničenih vještina čitanja (65). Zbog toga rijetko traže pomoć i često razvijaju strategije kako bi prikrili poteškoće u razumijevanju pisanih zdravstvenih materijala (66). Istraživanja su također pokazala povezanost između starije životne dobi i niže razine pismenosti, što ukazuje na to da su starije osobe posebno izložene riziku (67). Osim toga, neki su istraživači dodatno izrazili zabrinutost zbog kulturne prikladnosti obrazovnih materijala za pacijente, posebno u kontekstu različitih zdravstvenih uvjerenja, naglašavajući potrebu za materijalima koji su istovremeno razumljivi i kulturno osjetljivi (68,69).

Kako bi se poboljšala zdravstvena pismenost, potrebne su multidisciplinarne strategije koje uključuju: pojednostavljenje zdravstvenih informacija putem jasnog i razumljivog jezika u komunikaciji s pacijentima; edukacijske inicijative poput radionica i javnozdravstvenih kampanja koje naglašavaju važnost zdravstvene pismenosti; korištenje digitalnih alata poput mobilnih aplikacija, interaktivnih portala i virtualnih savjetnika za lakši pristup informacijama; osnaživanje uloge zdravstvenih radnika kroz njihovu edukaciju o učinkovitom komuniciranju i prepoznavanju pacijenata s niskom zdravstvenom pismenošću; te partnerstvo s medijima i društvenim mrežama radi suzbijanja dezinformacija i pružanja točnih, stručnih sadržaja (62).

Zdravstvena pismenost igra ključnu ulogu u unaprjeđenju zdravlja pojedinaca i društva u cjelini. Povećanje razine zdravstvene pismenosti može dovesti do boljih zdravstvenih ishoda, smanjenja troškova zdravstvene skrbi i većeg povjerenja u medicinske institucije (62). Značajan broj zemalja je to prepoznao (SAD, Kanada, Australija, Europska unija, Kina) i uvrstio zdravstvenu pismenost među svoje ključne prioritete u politikama i praksama, a neke i u nacionalnim planovima (62,70,71).

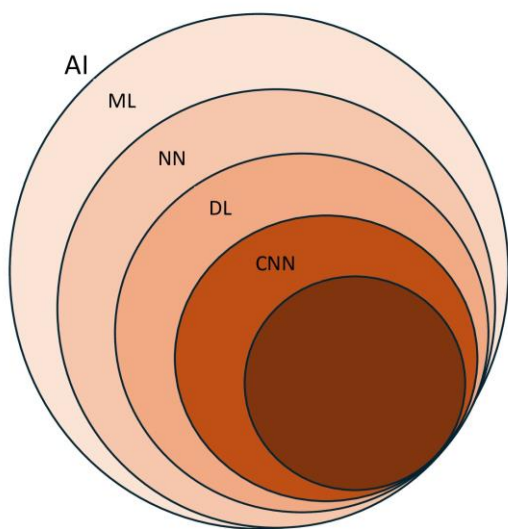
1.6 Dostupnost onkoloških informacija općoj populaciji

Dostupnost onkoloških informacija općoj populaciji igra ključnu ulogu u prevenciji, ranom otkrivanju i liječenju malignih bolesti (72–74). Razvoj digitalnih tehnologija, internetskih platformi i umjetne inteligencije omogućio je širu dostupnost pouzdanih medicinskih podataka, edukativnih materijala i interaktivnih alata koji pomažu građanima da bolje razumiju rizike, simptome i mogućnosti liječenja raka (75,76). Informacijske kampanje,

online baze podataka i virtualni asistenti mogu olakšati pristup stručnim smjernicama i najnovijim znanstvenim saznanjima, čime se motivira pojedince na pravovremene preglede i zdravije životne navike (75,77). Prema dostupnim podacima, više od trećine američkih stanovnika ima nisku razinu zdravstvene pismenosti. Ako govorimo o onkologiji, gotovo polovica pacijenata bolesnika nailazi na probleme u shvaćanju i razumijevanju bolesti s kojom su dijagnosticirani i terapijskim opcijama (62). Kao što je gore spomenuto, niska zdravstvena pismenost povezana je sa lošijim zdravstvenim ishodima, što je rezultat nekoliko faktora: slabija adherencija terapiji, veća učestalost propuštanja terapijskih termina, kao i povećan rizik od hospitalizacije (62,78). Ono što je bitno naglasiti jest da niska razina zdravstvene pismenosti može negativno utjecati na sposobnost informiranog donošenja odluka, što je posebno važno s obzirom na sve veću složenost onkološke terapije i aktivniju ulogu bolesnika i njihovih obitelji u procesu liječenja (79). Dostupnost onkoloških informacija općoj populaciji ključan je aspekt suvremene zdravstvene skrbi koji ima značajan utjecaj na prevenciju, rano otkrivanje i liječenje raka (80,81). Ovaj trend postaje sve važniji u globalnom kontekstu, posebno kada se ima na umu da je Internet postao primarni izvor zdravstvenih informacija za većinu ljudi (82,83). Društvene mreže igraju posebno važnu ulogu u širenju informacija o prevenciji raka, nudeći interaktivne i personalizirane pristupe (84,85). Međutim, ova dostupnost donosi i izazove, posebno u vidu potencijalnog širenja dezinformacija koje mogu narušiti povjerenje u zdravstveni sustav (86,87). Mnoge zemlje razvijaju nacionalne strategije za borbu protiv raka koje naglašavaju važnost edukacije o rizičnim čimbenicima i promocije zdravih životnih stilova (70,71). Unatoč napretku, još uvijek postoje značajni izazovi. Jedan od njih je stigmatizacija, odnosno nedovoljno poznavanje činjenica o raku, koja često rezultira izbjegavanjem razgovora o temi i odbijanjem preventivnih pregleda. Drugi izazov je kompleksnost informacija: onkološke informacije mogu biti složene za razumijevanje općoj populaciji, što zahtijeva pažljiv pristup u njihovom pojednostavljivanju bez gubitka točnosti (88,89). Zdravstveni profesionalci imaju ključnu ulogu u pružanju točnih informacija o raku (82). Iako se globalno bilježi napredak u poboljšanju dostupnosti onkoloških informacija, još uvijek postoji značajan prostor za unapređenje. Ključno je osigurati da informacije budu točne, lako dostupne i razumljive općoj populaciji, uz istovremeno jačanje povjerenja u zdravstveni sustav i poticanje preventivnih mjera.

1.7 Umjetna inteligencija i njena primjena u medicini

Primjena medicinskih tehnologija je uvijek bila značajan dio dijagnostičkih i terapijskih postupaka u medicini (90). Kroz posljednjih nekoliko godina svjedočimo značajnom uzletu uporabe umjetne inteligencije u ovom području, koja je sada već integrirana i široko primjenjena u određenim medicinskim sektorima (90,91). Umjetna inteligencija obuhvaća skup teorija, metoda i tehnologija čiji je cilj omogućiti računalnim sustavima da obavljaju zadatke koji zahtijevaju ljudsku inteligenciju, kao što su prepoznavanje uzoraka, učenje, rješavanje problema (91). Važno je napomenuti da je preciznu definiciju teško dati s obzirom da je pojam umjetne inteligencije širok i može se odnositi na različite pristupe kojima se koristi. Tako uz pojam umjetne inteligencije vezemo: umjetnu inteligenciju u užem smislu, strojno učenje (engl. *machine learning*), duboko učenje (engl. *deep learning*), tehnologija decentraliziranih evidencija (engl. *distributed ledger technology* (DLT)), *natural language processing* (NLP; algoritme temeljene na umjetnim neuronskim mrežama, koji pokušavaju oponašati način na koji funkcionira ljudski mozak), Metaverzum (engl. *Metaverse*), konvolucijske neuronske mreže (engl. *convolutional neural network* (CNN), odnosno *convneti*), Chat Generative Pretrained Transformer (ChatGPT), Transformer (92), kao i ostale različite termine. Slika 2. prikazuje odnose između različitih područja umjetne inteligencije.



Slika 2. Odnosi različitih područja umjetne inteligencije. AI-artificial intelligence, ML-machine learning, NN-neural network, DL-deep learning, CNN-convolutional neural network. Adaptirano prema Razavian et al (93).

Strojno učenje trenutačno je dominantan pristup umjetnoj inteligenciji, temeljen na načelu učenja iz podataka. Svaki algoritam strojnog učenja započinje obradom ulaznih podataka — informacija koje model prima za analizu. Na primjer, ako pričamo o medicinskom snimanju, ulaz može biti jedna slika ili niz slika. Isto tako ulaz može biti skup simptoma. Model zatim generira izlaz — informaciju, odnosno ono što pokušava predvidjeti. U kontekstu medicinskog snimanja, izlaz može biti dijagnoza temeljena na zadanom ulazu (slici), kao i ako pričamo o tome da je zadani ulaz bio skup simptoma. Svaki algoritam učenja također ima cilj učenja, tj. vrijednost koja pokazuje koliko je model uspješan u predviđanju izlaza na temelju ulaza. Kod većine modela treniranje uključuje višekratno predstavljanje podataka modelu kako bi mogao prilagoditi svoje parametre (92,93).

Neuronske mreže su modeli učenja koji se sastoje od više slojeva. Općenito, neuronsku mrežu možemo opisati kao niz funkcija s parametrima koji se mogu učiti, primijenjenih na ulaz. Izlaz mreže računa se kao niz primjena funkcija. Sve funkcije osim posljednje predstavljaju skrivene slojeve, dok zadnja funkcija čini izlazni sloj. Taj se sloj razlikuje jer daje konačnu predikciju mreže. Područje strojnog učenja koje se bavi neuronskim mrežama, posebno onima s mnogo slojeva, obično se naziva duboko učenje (engl. *deep learning*) (92,94,95).

Među arhitekturama dubokih neuronskih mreža, izdvajaju se konvolucijske neuronske mreže (CNN-ovi, odnosno engl. *convolutional neural network-convneti*). CNN-ovi imaju posebnu strukturu povezanosti slojeva, nadahnutu ljudskim vidnim korteksom, koja ih čini iznimno pogodnima za percepcijske zadatke poput prepoznavanja slika. Convneti imaju manji broj parametara, što ih čini manje sklonima preučenosti (92,96,97).

Tehnologija decentraliziranih evidencija (engl. *distributed ledger technology*) predstavlja siguran, nepromjenjiv i dostupan način bilježenja i dijeljenja podataka među više sustava. Omogućuje pacijentima upravljanje vlastitim zdravstvenim podacima, čime se potiče povjerenje u zdravstveni sustav. U kombinaciji s umjetnom inteligencijom, DLT nudi napredan pristup za inteligentno, pouzdano i sigurno upravljanje elektroničkim zdravstvenim zapisima (92,98).

Metaverzum (engl. *Metaverse*) je trodimenzionalno virtualno okruženje koji omogućava komunikaciju korisnika putem avatara. U zdravstvu omogućuje personalizirano, interaktivno i privlačno korisničko iskustvo. Primjenjuje se u edukaciji, istraživanju, praćenju pacijenata, prevenciji i liječenju bolesti (92,99).

ChatGPT je konverzijski alat temeljen na umjetnoj inteligenciji koji koristi obradu prirodnog jezika i algoritme strojnog učenja za simulaciju razgovora nalik ljudskima.

Treniran je na milijardama tekstualnih elemenata te je brzo privukao veliku pozornost zahvaljujući svojoj sposobnosti da jasno i informirano odgovara na pitanja iz različitih područja znanja. Ove tehnologije velikih jezičnih modela (engl. *large language model*-LLM), imaju širok spektar mogućih primjena. Kao tekstualni chatbot, razumije korisničke upite i generira odgovore. Chatbotovi u zdravstvu mogu pružati podršku pacijentima, pomoći u administraciji i praćenju te sudjelovati u dijagnostici, probiru i liječenju tumora (92,100,101).

Transformer je model dubokog učenja koji se široko primjenjuje u obradi prirodnog jezika i govora. U zdravstvu se koristi za obradu elektroničkih zdravstvenih zapisa, medicinsku dijagnostiku putem slika i detekciju bolesti poput COVID-19 (92,102,103).

Ovo je bio kratki prikaz najčešće korištenih vidova umjetne inteligencije, ali postoji još nekoliko različitih podjela umjetne inteligencije pa tako ako umjetnu inteligenciju možemo podijeliti i prema sposobnostima:

1. Uska umjetna inteligencija (Narrow AI)

Radi se o umjetnoj inteligenciji koja je trenutno jedina koja danas stvarno postoji (104). Specijalizirana je za jedan zadatak, koji može izvršavati brže i preciznije od čovjeka, ali ne može raditi izvan svog zadatka (105). Primjer je ChatGPT.

2. Opća umjetna inteligencija (General AI)

Teoretski koncept koji bi mogao učiti i rješavati različite zadatke slično čovjeku, bez dodatnog treniranja. Takva AI zasad ne postoji (105).

3. Superinteligentna AI (Super AI)

Još jedan teorijski oblik AI koji bi nadmašio ljudsku inteligenciju — uključujući osjećaje, svijest, i sposobnost samostalnog zaključivanja i donošenja odluka (105,106).

Dalje, umjetna inteligencija se može podijeliti i prema funkcionalnosti:

1. Reaktivna AI

Umjetna inteligencija koja nema memoriju i može reagirati samo na trenutno dostupne podatke. Koristi se u vrlo specifičnim zadacima, npr. kao *Netflix* algoritam preporuka (107).

2. AI s ograničenom memorijom

Može koristiti prošle podatke za donošenje odluka, ali ih ne pohranjuje dugoročno.

Primjeri su ChatGPT, virtualni asistenti, autonomna vozila (107).

3. AI s razumijevanjem uma (engl. *Theory of Mind AI*)

Zasad teorijska, ova vrsta AI bi mogla razumjeti emocije i namjere ljudi te na temelju toga personalizirati interakcije (107,108).

4. Samosvjesna AI (engl. *Self-Aware AI*)

Najnapredniji i najdalji oblik AI – također teorija. Takva bi AI imala svijest o sebi, vlastite emocije, potrebe i uvjerenja, uz razumijevanje ljudskih misli i osjećaja (106,107).

Uz ove podjele, neki autori dijele umjetnu inteligenciju na virtualnu i fizičku (109). Virtualni dio koristi sustave elektroničkih zdravstvenih zapisa i različite senzore za donošenje odluka o liječenju, dok fizički dio omogućuje robotima izvođenje kirurških zahvata, ugradnju implantata, zamjenu organa, skrb za starije osobe i druge medicinske zahvate (109).

Povijest umjetne inteligencije u medicini seže unatrag nekoliko desetljeća – od prvih sustava temeljenih na logici i pravilima do današnjih naprednih sustava koji koriste ogromne količine podataka (tzv. *big data*) i algoritme sposobne za samostalno učenje i prilagodbu. Razvoj računalne snage, dostupnost velikih medicinskih baza podataka i napredak u algoritmima doveli su do eksponencijalnog rasta primjene AI u zdravstvu (91,110). Shodno tome, razvijeni algoritmi mogu se koristiti kroz nekoliko područja: dijagnostika, personaliziranje terapije, robotska asistencija pri zahvatima, upravljanje i automatizacija administracijom, vođenje baza podataka, poboljšanje skrbi pacijenata kroz virtualne asistente i chatbotove (91,92,94). Ipak, unatoč ogromnim prednostima, razvoj i primjena umjetne inteligencije u medicini nose i niz izazova. Jedno od ključnih pitanja odnosi se na privatnost i sigurnost podataka – tko ima pristup osjetljivim osobnim i zdravstvenim informacijama pacijenata te na koji način se ti podaci pohranjuju i koriste (111). Osim toga, transparentnost algoritama – odnosno mogućnost objašnjavanja na temelju kojih kriterija je neka odluka donesena – ključna je za povjerenje pacijenata i medicinskih stručnjaka (112). Postoji i tehnički izazov u integraciji različitih sustava i izvora podataka, s obzirom na to da medicinske ustanove često koriste različite elektroničke sustave koji nisu međusobno kompatibilni. Kvaliteta i reprezentativnost podataka također igraju važnu ulogu – ako su ulazni podaci pristrani ili nepotpuni, i rezultati algoritama mogu biti netočni ili diskriminirajući (92).

Primjena umjetne inteligencije u onkologiji donosi značajne inovacije u ranom otkrivanju, dijagnostici i liječenju karcinoma, čime se poboljšavaju ishodi liječenja i kvaliteta života pacijenata (113,114). Algoritmi strojnog učenja i dubokog učenja omogućuju analizu velikih količina medicinskih podataka. Primjenu toga nalazimo u ranoj detekciji i dijagnozi tumora putem radioloških snimki: algoritmi poboljšavaju interpretaciju radioloških snimki, povećavajući točnost i brzinu dijagnoze raka dojke, pluća i drugih oblika tumora čime se postiže preciznija i brža dijagnoza malignih oboljenja (113,115). Dalje, analiza velikih količina medicinskih podataka je korisna u obradi patohistoloških uzoraka i genomskih informacija te se na taj način pomaže u personalizaciji terapijskih pristupa, predviđajući

odgovor pacijenata na određene tretmane i optimizirajući planove liječenja. Osim toga, primjena umjetne inteligencije u analizi kliničkih ispitivanja i otkrivanju novih biomarkera ubrzava razvoj inovativnih terapija (116,117). Dodatno, umjetna inteligencija može optimizirati kliničke procese i upravljanje pacijentima. Sustavi zasnovani na AI mogu analizirati elektroničke zdravstvene zapise kako bi predvidjeli tijek bolesti, preporučili optimalne terapijske strategije te unaprijedili planiranje medicinskih resursa (118,119). Automatizacija administrativnih zadataka također omogućava zdravstvenim djelatnicima da više vremena posvete pacijentima. Što se tiče radioterapije, umjetna inteligencija pomaže u automatizaciji iste-predlaže konture organa od rizika te optimizira doze zračenja (120). Unatoč brojnim prednostima, izazovi poput interpretabilnosti modela i integracije AI u kliničku praksu zahtijevaju daljnja istraživanja i razvoj te i dalje ostaju prisutni svi gore navedeni izazovi koji se odnose na samu primjenu umjetne inteligencije u medicini (121). Iako AI donosi značajne prednosti u onkologiji, važno je naglasiti da ona ne zamjenjuje liječnike, već im služi kao moćan alat za poboljšanje učinkovitosti i točnosti u njihovom radu.

2. CILJ RADA I HIPOTEZE

Cilj ovog istraživanja je analizirati sažetke Cochrane sustavnih pregleda koji se bave intervencijama u onkologiji u smislu čitkosti, duljine teksta i korištenja različitih emocionalnih tonova te utvrditi može li korištenje umjetne inteligencije pomoći u poboljšanju stvaranja laičkih sažetaka.

2.1 Jezična analiza laičkih i pripadajućih znanstvenih sažetaka Cochrane sustavnih pregleda onkoloških intervencija

Glavni cilj ove studije bio je analizirati laičke sažetke sustavnih pregleda koji su istraživali intervencije u onkologiji prema njihovim jezičnim karakteristikama te ih usporediti s odgovarajućim Cochraneovim znanstvenim sažecima. Jezične karakteristike koje su se utvrđivale su bile duljina teksta, čitkost te emocionalna obilježja teksta.

Ostali ciljevi su bili:

1. Usporediti laičke sažetke između različitih Cochrane onkoloških skupina s obzirom na gore navedene karakteristike;
2. Usporediti laičke i pripadajuće znanstvene sažetke s obzirom na zaključke koji su doneseni u određenom sustavnom pregledu na istraživanu intervenciju.

Hipoteze u ovom istraživanju su sljedeće:

1. Laički sažeci Cochrane sustavnih pregleda onkoloških intervencija su napisani iznad razine čitkosti preporučene za opću populaciju.
2. Laički sažeci različitih Cochrane onkoloških skupina se ne razlikuju značajno s obzirom na jezične karakteristike.
3. Laički sažeci Cochrane sustavnih pregleda napisani su na pristupačniji i jednostavniji način u usporedbi sa znanstvenim sažecima koje prate.
4. Laički i pripadajući znanstveni sažeci su sukladni u zaključcima koji su doneseni u određenom sustavnom pregledu na danu intervenciju.
5. Znanstveni sažeci koriste više analitičkog tona u odnosu na pripadajuće laičke sažetke.

2.2 Korištenje umjetne inteligencije za poboljšanje prezentacije laičkih sažetaka Cochraneovih sustavnih pregleda o onkološkim intervencijama

Glavni cilj ovog istraživanja je bio ustvrditi može li se korištenjem umjetne inteligencije, tj. chatbota temeljenog na naprednom jezičnom modelu poboljšati prezentacija laičkih sažetaka Cochraneovih sustavnih pregleda o onkološkim intervencijama.

Ostali ciljevi su bili:

1. Ustvrditi obilježja sažetaka generiranih od strane ChatGPT-a u smislu postotka korištenja analitičkog, autentičnog, emocionalnog tona i govora s pozicije moći i usporediti ih s originalnim laičkim sažecima.
2. Ustvrditi kategorije zaključivosti sažetaka generiranih od strane ChatGPT-a i usporediti ih s originalnim laičkim sažecima.

Hipoteze u ovom istraživanju su sljedeće:

1. Indeks čitkosti sažetaka koje generira ChatGPT bit će u prosjeku niži nego kod izvornih laičkih sažetaka koje su izradili autori sustavnih pregleda.
2. Sažeci generirani od strane ChatGPT-a će koristiti manje analitičkog tona.
3. Zaključci u sažecima generiranim pomoću ChatGPT-a neće se razlikovati od onih u izvornim laičkim sažecima.

3. METODE

3.1 Jezična analiza laičkih i pripadajućih znanstvenih sažetaka Cochraneovih sustavnih pregleda onkoloških intervencija

3.1.1 Ustroj

Ovo presječno istraživanje fokusiralo se na analizu jezičnih obilježja korištenih u sažecima Cochraneovih sustavnih pregleda vezanih uz onkološke intervencije.

3.1.2 Ishodi istraživanja

Primarni ishod istraživanja je bilo određivanje razina čitkosti laičkih i pripadajućih znanstvenih sažetaka. Sekundarni ishodi istraživanja su bili analiza jezičnih obilježja (korištenje emocionalnog, autentičnog, analitičkog tona i govora s pozicije moći) u laičkim i pripadajućim znanstvenim sažecima te analiza kategorije zaključivosti.

3.1.3 Prikupljanje podataka

Ovo istraživanje presječnog dizajna ispitalo je jezične karakteristike sažetaka sustavnih pregleda koji su analizirali onkološke terapije, iz baze Cochrane. Uzorak je obuhvaćao sve laičke i znanstvene sažetke sustavnih pregleda onkoloških intervencija dostupne u Cochrane knjižnici, do veljače 2019. godine. Sažeci koji nisu obuhvaćali intervencijske studije nisu bili uključeni u završnu analizu. U analizu su uvrštene isključivo one Cochrane skupine čiji je primarni fokus onkološka tematika: Skupina za rak pluća (engl. *Lung Cancer Review Group*), Skupina za rak dojke (engl. *Breast cancer Review Group*); Skupina za hematološke maligne bolesti (engl. *Haematological Malignancies Review Group*), Skupina za maligne bolesti u djece (engl. *Childhood cancer Review Group*); Skupina za ginekološke tumore, neuroonkologiju i rijetke tumore (engl. *Gynaecological, Neuro-oncology and Orphan Cancer Review Group*); Skupina za kolorektalni karcinom (engl. *Colorectal cancer Review Group*). I u drugim Cochrane skupinama mogu se pronaći studije vezane uz onkološke terapije, no u ovoj analizi obuhvaćene su isključivo one skupine koje su potpuno posvećene području onkologije. Promatrani laički i znanstveni sažeci bili su napisani na engleskom jeziku, a autori sustavnih pregleda izradili su obje verzije sažetaka.

3.1.4 Probir sažetaka i vađenje podataka

Neovisno pregledavanje sažetaka napravila su dva autora (J.Š.P. i I.B.), a nesuglasice su razriješene konzultacijom s trećim članom istraživačkog tima (A.M.). Za potrebe ekstrakcije podataka izrađen je obrazac koji je prethodno ispitan na uzorku od deset laičkih i

znanstvenih sažetaka. Dva autora (J.Š.P. i I.B.) su sudjelovala u izradi tablice za bilježenje podataka iz ekstrakcije, s podacima o izvlačenju podataka. Kako bi se osigurala dosljednost u prikupljanju podataka, provedena je kalibracijska vježba. Jedan autor je izvršio ekstrakciju podataka, dok je drugi ponovno izveo postupak na 10% uzorka radi provjere točnosti unosa u tablicu. Treći autor (A.M.) provjerio je unose iz ekstrakcije dvaju autora na uzorku od 20 laičkih i znanstvenih sažetaka na početku studije. Iz svakog sažetka prikupili smo sljedeće informacije: broj autora, godina kada je objavljen rad, Cochrane skupina koja je objavila analizirani Cochrane sustavni pregled, indeks čitkosti, jezične karakteristike teksta, kategoriju zaključka.

3.1.5 Obrada podataka i mjerni instrumenti

3.1.5.1 SMOG indeks

Čitkost je procijenjena uz pomoć SMOG indeksa (*Simple Measure of Gobbledygook*), koji pokazuje razinu obrazovanja potrebnu za razumijevanje određenog teksta. SMOG indeks smatra se najprikladnijim u određivanju razine čitkosti zdravstvenih informacija (49). Za izračun SMOG indeksa engleskih laičkih sažetaka korišten je internetski alat dostupan na adresi <https://www.webfx.com/tools/read-able/>. SMOG indeks svoju metodu procjene čitkosti temelji na brojanju višesložnih riječi, a dobiveni rezultat izražava potreban broj godina školovanja za njegovo razumijevanje. (49). Za tekstove s više od trideset rečenica, uzima se po deset rečenica s početnog, srednjeg i završnog dijela za procjenu čitkosti. Zatim se utvrđuje broj višesložnih riječi unutar tih rečenica, nakon čega se iz tog broja računa kvadratni korijen. Dobivena vrijednost se zaokružuje na najbliži cijeli broj i uvećava za tri, čime se dobiva SMOG indeks za promatrani tekst. U slučajevima kada tekst sadrži manje od 30 rečenica, uzimaju se sve rečenice, izračuna se ukupan broj višesložnih riječi, zatim kvadratni korijen najbližeg savršenog kvadrata tog broja, kojem se također dodaje tri (50). Može se koristiti i sljedeća formula: razina čitkosti = $1.0430 \times \sqrt{(\text{broj polisilabičnih riječi} \times 30 / \text{broj rečenica})} + 3.1291$ (50). Prema preporukama koje je izdalo Američko Medicinsko Udruženje i Nacionalni Institut za Zdravlje, zdravstveni tekstovi trebali bi biti napisani na razini 6, odnosno razumljivi osobama koje imaju 6 godina školovanja, što bi odgovaralo učenicima viših razreda osnovne škole (ako govorimo o hrvatskom obrazovnom sustavu - 6.razred osnovne škole) (122).

3.1.5.2 Jezična obilježja

Laički i znanstveni sažetci su analizirani prema svojim jezičnim obilježjima. Analiza se odnosila na jezične značajke sažetaka, uključujući duljinu rečenica, broj riječi i emocionalni ton.

Za procjenu broja riječi i prosječne dužine rečenica korišten je računalni alat Linguistic Inquiry and Word Count – LIWC (<https://liwc.wpengine.com/>). Program LIWC također je upotrijebljen za analizu učestalosti izraza povezanih s različitim jezičnim tonovima. Ova analiza temelji se na rječniku i klasificira riječi iz teksta u četiri glavne kateogrije: analitičnost, emocionalni ton, govor s pozicije moći (engl. *clout*) I autentičnost teksta. Te se varijable prikazuju kao postotak riječi iz teksta koje pripadaju određenoj kategoriji (123).

Kategorija analitičkog razmišljanja temelji se na prepoznavanju riječi povezanih s logikom te s povezivanjem i uspostavljanjem odnosa među pojmovima. Veći udio takvih riječi povezan je s kognitivnom složenošću i apstraktnim razmišljanjem (124,125).

Govor s pozicije moći (*clout*) odnosi se na uporabu izraza koji označavaju samopouzdanje ili lidersku poziciju. Veća učestalost takvih riječi sugerira da autor piše iz pozicije stručnosti i sigurnosti u izneseno, dok njihova manja zastupljenost upućuje na suzdržanije izlaganje informacija (126).

Autentičnost se određuje postotkom riječi povezanih s osobnošću, poput uporabe osobnih zamjenica u prvom licu (npr. „ja“, „moj“, „moje“), glagola u sadašnjem vremenu i odnosnih priloga (npr. „blizu“, „sada“). Korištenje tih riječi povezano je s osobnijim i iskrenijim stilom pisanja (127).

Emocionalni ton odnosi se na to koliko je ton teksta pozitivan na temelju upotrijebljenih riječi. Rezultat od 100 označava maksimalno pozitivan ton, dok rezultat od 50 označava ravnotežu između pozitivnih i negativnih emocionalnih izraza (128,129).

3.1.5.3 Zaključivost

Dvoje neovisnih procjenitelja (J.Š.P. i I.B.) svrstalo je zaključke sažetaka u jednu od unaprijed definiranih kategorija:

1. Uvjerljivo:

- Pozitivno uvjerljivo: Postojali su umjereni ili visokokvalitetni dokazi koji su ukazivali na učinkovitost ili sigurnost intervencije, tj. lijek je dokazan kao učinkovit/siguran.
- Negativno uvjerljivo: Postojali su umjereni ili visokokvalitetni dokazi koji su ukazivali da je intervencija neučinkovita ili štetna, ili su autori savjetovali protiv intervencije/usporedbe.
- Jednako uvjerljivo: Analizirane intervencije bile su jednako učinkovite i sigurne.

2. Neuvjerljivo:

- Pozitivno neuvjerljivo: Postojali su dokazi koji sugeriraju učinkovitost ili sigurnost, ali su bili niske kvalitete te autori sugeriraju da je potrebno više istraživanja.
- Negativno neuvjerljivo: Postojali su dokazi o neučinkovitosti ili šteti (dokazi koji pokazuju da nije bilo učinka ili da intervencija nije sigurna) ili su autori upozorili protiv intervencije ili usporedbe, no dokazi su bili niske kvalitete, odnosno autori navode da je potrebno više istraživanja.
- Jednako neuvjerljivo: Intervencije su se činile jednako učinkovitim i sigurnim, no dokazi su bili niže kvalitete te autori sugeriraju potrebu za dodatnim istraživanjima.

3. Nejasno:

- Nema dokaza: Nije bilo dokaza jer pretraživanje nije rezultiralo pronalaskom nikakvih randomiziranih kontroliranih studija (tj. prazni pregledi).
- Nema mišljenja: Autori nisu iznijeli nikakvo mišljenje ili sud.
- Nejasno: Autori nisu dali jasan zaključak ili su naveli potrebu za dodatnim istraživanjima (130,131).

3.1.6 Statistički postupci

Distribucija sažetaka po kategorijama zaključivosti i pripadajućim Cochrane skupinama izražena je u učestalostima i postocima. Zbog utvrđene nejednake distribucije podataka (Kolmogorov-Smirnov test) sljedeći podaci su prikazani kao medijani s 95%-tnim intervalima pouzdanosti (95% CI, od engl. *Confidence Interval*, CI): jezične karakteristike

(emocionalni ton, govor s pozicije moći, analitičnost, autentičnost), broj riječi, čitkost. Statistička usporedba među onkološkim skupinama provedena je korištenjem Kruskal-Wallis testa, uz dodatnu analizu Conover-Iman post hoc testom. Prag statističke značajnosti bio je postavljen na p vrijednost manju od 0,05. Statistička obrada podataka provedena je pomoću JASP softvera (verzija 0.12.1.0) (JASP Team, 2020).

3.2 Korištenje umjetne inteligencije u poboljšanju laičkih sažetaka Cochrane sustavnih pregleda onkoloških informacija

3.2.1 Ustroj studije

U ovoj studiji analizirali smo čitkost, jezične karakteristike i kategoriju zaključivosti PLS-ova generiranih pomoću umjetne inteligencije - ChatGPT-a (engl. *Chat generative pre-trained transformer*), koristeći kao temelj i podlogu informacija odgovarajuće znanstvene sažetke sustavnih pregleda iz Cochrane knjižnice koji su se odnosili na intervenciju u području onkologije. Zatim smo ih usporedili s čitkošću i jezičnim karakteristikama originalnih PLS-ova, kao i sa kategorijom zaključivosti pripadajućih znanstvenih sažetaka.

3.2.2 Uzorak

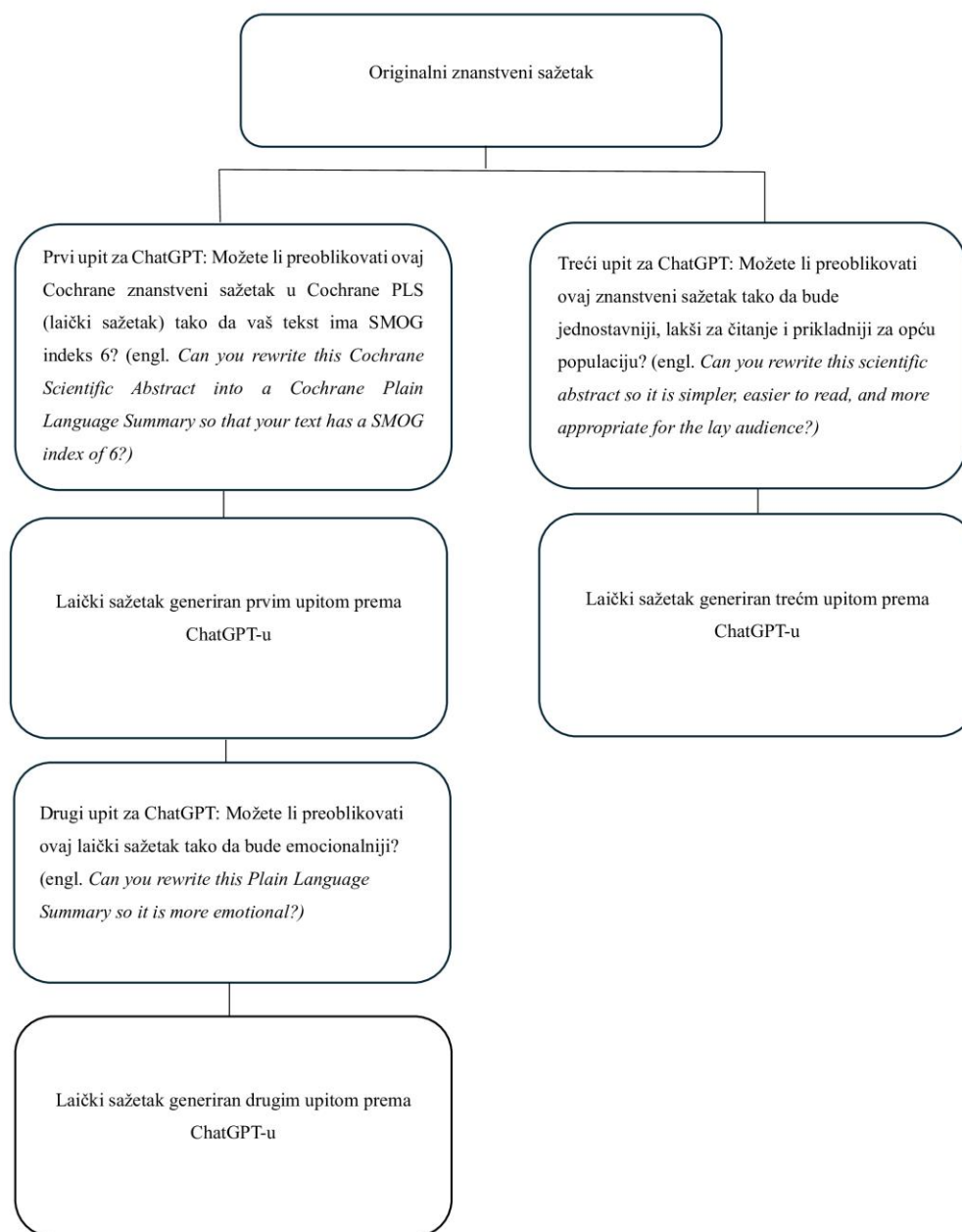
Uključeno je 275 PLS-ova i pripadajućih znanstvenih sažetaka Cochrane sustavnih pregleda o onkološkim intervencijama do veljače 2019. iz naše prethodne studije (132). U toj studiji procijenili smo jezične karakteristike PLS-ova Cochrane sustavnih pregleda o onkološkim intervencijama u usporedbi s pripadajućim Cochrane znanstvenim sažecima. Koristili smo ovaj skup podataka jer je sadržavao znanstvene sažetke za koje je već procijenjena kategorija zaključivosti intervencija, što nam je omogućilo usporedbu zaključivosti AI-generiranog PLS-a s onom iz izvornog znanstvenog sažetka. Cochrane sustavni pregledi uključeni u skup podataka potječu iz Cochrane skupina usmjerenih isključivo na onkologiju, tj. intervencije u onkologiji. Te skupine predstavljale su različite kliničke tipove raka, odnosno obuhvaćene su skupine koje se bave rakom pluća, rakom dojke, kolorektalnim rakom, hematološkim zloćudnim bolestima, tumorima središnjeg živčanog sustava, ginekološkim tumorima, rijetkim tumorima te zloćudnim tumorima pedijatrijske populacije. Sustavni pregledi koji nisu obrađivali intervencijske studije bili su isključeni. Sažetci su analizirani u smislu njihovog SMOG indeksa, jezičnih karakteristika (broj riječi i postotak riječi povezanih s različitim emocionalnim tonovima) te kategorije zaključivosti. Cijeli skup podataka javno je dostupan putem Open Science Framework (133), kao i Prilog 1. u sklopu ove disertacije.

Koristili smo Smjernice za pisanje izvješća opažajnih studija (engl. *Strengthening the Reporting of Observational Studies in Epidemiology*) (STROBE) kontrolni popis za izvještavanje ove studije.

3.2.3 Generiranje PLS-ova pomoću ChatGPT-a

Koristili smo pretplatnički chatbot, ChatGPT verziju 4 (Open AI). U trenutku našeg istraživanja, podaci za treniranje ChatGPT-a uključivali su informacije do listopada 2023. Formirali smo tri upita, a jedan autor (J.Š.P.) postavio je ChatGPT-u svaki od njih za svaki pojedini PLS zasebno. Prije postavljanja ovih pitanja, zamolili smo ChatGPT da objasni što je SMOG indeks. ChatGPT je ispravno opisao mjeru čitkosti, njezinu formulu i interpretaciju rezultata, čime je potvrdio da je dovoljno upoznat s SMOG indeksom. Zatim smo koristili sljedeće upite, koje smo postavili na engleskom jeziku (Slika 3):

1. Možete li preoblikovati ovaj Cochrane znanstveni sažetak u Cochrane PLS (laički sažetak) tako da vaš tekst ima SMOG indeks 6? (engl. *Can you rewrite this Cochrane Scientific Abstract into a Cochrane Plain Language Summary so that your text has a SMOG index of 6?*)
2. Možete li preoblikovati ovaj laički sažetak tako da bude emocionalniji? (engl. *Can you rewrite this Plain Language Summary so it is more emotional?*)
3. Možete li preoblikovati ovaj znanstveni sažetak tako da bude jednostavniji, lakši za čitanje i prikladniji općoj populaciji? (engl. *Can you rewrite this scientific abstract so it is simpler, easier to read, and more appropriate for the lay audience?*).



Slika 3. Generiranje PLS-ova od strane ChatGPT-a. SMOG-Simple Measure of Gobbledygook.

U prvom upitu tražili smo od ChatGPT-a da preoblikuje znanstveni sažetak u PLS sa SMOG indeksom 6, sukladno preporuci Američkog medicinskog udruženja i Nacionalnog instituta za zdravlje koje predlažu da sadržaj namijenjen za javnost bude napisan na tom stupnju čitkosti (122).

Za drugi upit nastavio se razgovor s ChatGPT-om započet u prvom upitu. Budući da smo u prethodnoj studiji utvrdili da PLS-ovi sadrže jezik koji je emocionalniji u usporedbi s pripadajućim znanstvenim sažecima i da PLS-ovi s većim postotkom emocionalnog tona imaju bolju čitkost (132), koristili smo PLS koji je ChatGPT pružio u prvom upitu i ponovno ga zamolili da ga preoblikuje tako da bude emocionalniji. Ovim upitom željeli smo procijeniti može li dodavanje emocionalne rezonancije poboljšati angažman čitatelja i povezanost.

U trećem upitu, slično kao u prvom, tražili smo od ChatGPT-a da pojednostavi znanstveni sažetak, ali bez definiranja SMOG indeksa. Ovim upitom istraživali smo može li korištenje jednostavnijeg jezika (bez specificiranog indeksa čitkosti) rezultirati većom pristupačnošću, uz zadržavanje bitnih informacija i nijansi potrebnih za točno razumijevanje onkoloških tema. Prilikom zadavanja ovog upita započeli smo novi razgovor s ChatGPT-om, kako se ne bi oslanjao na prethodni odgovor.

Nakon toga, za svaki PLS koji je pružio ChatGPT, izmjerili smo njegovu čitkost (izraženu kao SMOG indeks) te sljedeće jezične karakteristike: broj riječi i postotak riječi povezanih s autentičnošću, govorom s pozicije moći, emocionalnim i analitičkim tonom. Analizirali smo samo prve odgovore generirane od ChatGPT-a za svaki od tri upita i nismo tražili dodatne revizije teksta. Kako bismo osigurali da sadržaj PLS-a prenosi iste činjenice kao pripadajući znanstveni sažetak, također smo provjerili kategoriju zaključivosti svakog generiranog PLS-a, odnosno jesu li se kategorije zaključivosti podudarale s onima iz izvornog znanstvenog sažetka, kako je utvrđeno u našoj prethodnoj studiji.

3.2.4 Obrada podataka

3.2.4.1 SMOG indeks

Čitkost formata sažetaka na engleskom jeziku ocijenjena je korištenjem Simple Measure of Gobbledygook (SMOG) indeksa. SMOG indeks mjeri čitkost zadanog teksta brojanjem višesložnih riječi, a odgovor tog računanja dolazi je broj koji izražava koliko je godina školovanja potrebno da bi se shvatio određeni tekst (50). Smatra se prikladnim za zdravstvene informacije zbog svojih dosljednih rezultata te jednostavnosti korištenja. SMOG indeks za engleske laičke sažetke izračunat je pomoću online alata WebFX Readability Test Tool. Pouzdanost i pristupačnost alata učinili su ga prikladnim izborom za procjenu čitkosti sažetaka jednostavnog jezika u našoj studiji. Kao što je prethodno spomenuto, za interpretaciju SMOG indeksa treba imati na umu službenu preporuku vodećih američkih zdravstvenih institucija koje savjetuju da bi bilo poželjno da je medicinski sadržaj

namijenjen za širu javnost biti napisan na razini čitanja koja je shvatljiva osobi koja ima 6 godina školovanja (122). Viši SMOG indeks označava da tekst teži za pročitati, dok niže vrijednosti upućuju na jednostavnost i lakšu čitkost.

3.2.4.2 Jezična obilježja

PLS-ovi generirani od ChatGPT-a analizirani su u pogledu njihovih jezičnih karakteristika korištenjem alata *Linguistic Inquiry and Word Count* – LIWC, softverskog alata dizajniranog za analizu zadanog teksta uspoređujući svaku riječ s unaprijed definiranim rječnikom. Alat klasificira riječi u tekstu prema unaprijed definiranom rječniku u jednu od četiri moguće varijable: emocionalni ton, autentični ton, analitičnost, govor s pozicije moći, izražene kao postotak riječi iz teksta koje spadaju u određenu kategoriju.

3.2.4.3 Kategorija zaključivosti

Kategorija zaključivosti za svaki PLS generiran od strane ChatGPT-a analizirana je od strane dva autora (J.Š.P. i I.B.). Tablica za ekstrakciju podataka testirana je od strane dva autora (J.Š.P. i I.B.). Jedan autor je prikupljao podatke, a drugi je neovisno pregledao podatke u slučajnom uzorku od 10% PLS-ova i pripadajućih znanstvenih sažetaka. Zatim je provjereno jesu li unosi u tablici ispravni. Međupromatračka suglasnost bila je visoka (kappa raspon 0,80 do 1,00, 95% interval pouzdanosti (CI)=0,84–1,00). Razlike u ocjenjivanju riješene su raspravom s trećim autorom (A.M.) prije potpune ekstrakcije podataka.

Zaključci sažetaka o učinkovitosti i sigurnosti intervencija kategorizirani su kao:

- Uvjerljivo:

- Pozitivno uvjerljivo: Postojali su umjereni ili visokokvalitetni dokazi koji su ukazivali na učinkovitost ili sigurnost intervencije, tj. lijek je dokazan kao učinkovit/siguran.
- Negativno uvjerljivo: Postojali su umjereni ili visokokvalitetni dokazi koji su ukazivali da je intervencija neučinkovita ili štetna, ili su autori savjetovali protiv intervencije/usporedbe.
- Jednako uvjerljivo: Analizirane intervencije bile su jednako učinkovite i sigurne.

- Neuvjerljivo:
 - Pozitivno neuvjerljivo: Postojali su dokazi koji sugeriraju učinkovitost ili sigurnost, ali su bili niske kvalitete te autori sugeriraju da je potrebno više istraživanja.
 - Negativno neuvjerljivo: Postojali su dokazi o neučinkovitosti ili šteti (dokazi koji pokazuju da nije bilo učinka ili da intervencija nije sigurna) ili su autori upozorili protiv intervencije ili usporedbe, no dokazi su bili niske kvalitete, odnosno autori navode da je potrebno više istraživanja.
 - Jednako neuvjerljivo: Intervencije su se činile jednako učinkovitim i sigurnim, no dokazi su bili niže kvalitete te autori sugeriraju potrebu za dodatnim istraživanjima.
- Nejasno:
 - Nema dokaza: Nije bilo dokaza jer pretraživanje nije rezultiralo pronalaskom nikakvih randomiziranih kontroliranih studija (tj. prazni pregledi).
 - Nema mišljenja: Autori nisu iznijeli nikakvo mišljenje ili sud.
 - Nejasno: Autori nisu dali jasan zaključak ili su naveli potrebu za dodatnim istraživanjima (131).

Na temelju ovih kriterija definirali smo kategoriju zaključivosti za svaki izvedeni PLS, a zatim smo usporedili tu kategoriju s onom iz izvornog znanstvenog sažetka kako bismo provjerili podudaraju li se i daju li isti zaključak o učinkovitosti ili sigurnosti intervencije.

3.2.5 Statistička analiza

3.2.5.1 Deskriptivna statistika

Podaci o čitkosti, broju riječi i jezičnim karakteristikama ocijenjeni su kao numeričke varijable. Budući da su podaci odstupali od normalne distribucije, rezultati su predstavljeni kao medijani i 95% intervali pouzdanosti (CI) te su prikazani za originalne PLS-ove i unutar tri grupe upita. Podaci o kategoriji zaključivosti opisani su kao frekvencije i postoci i prikazani u stupčastom grafikonu za tri različite grupe upita.

3.2.5.2 Usporedba grupa

Rezultati analize PLS-ova generiranih pomoću ChatGPT-a uspoređeni su s već objavljenim podacima za originalne PLS-ove i znanstvene sažetke (132). Budući da su sve verzije izvedene iz istog PLS-a, rezultati su tretirani kao unutargrupni podaci pod različitim uvjetima. Kako unutargrupna ANOVA nije bila prikladna zbog odstupanja od normalne distribucije podataka, usporedba između grupa izvršena je korištenjem Friedmanovog neparametrijskog testa za ponovljena mjerenja, a post hoc analiza provedena je Conoverovim post hoc testom, budući da je to jedna od preporučenih metoda.

3.2.5.3 Statistički alati

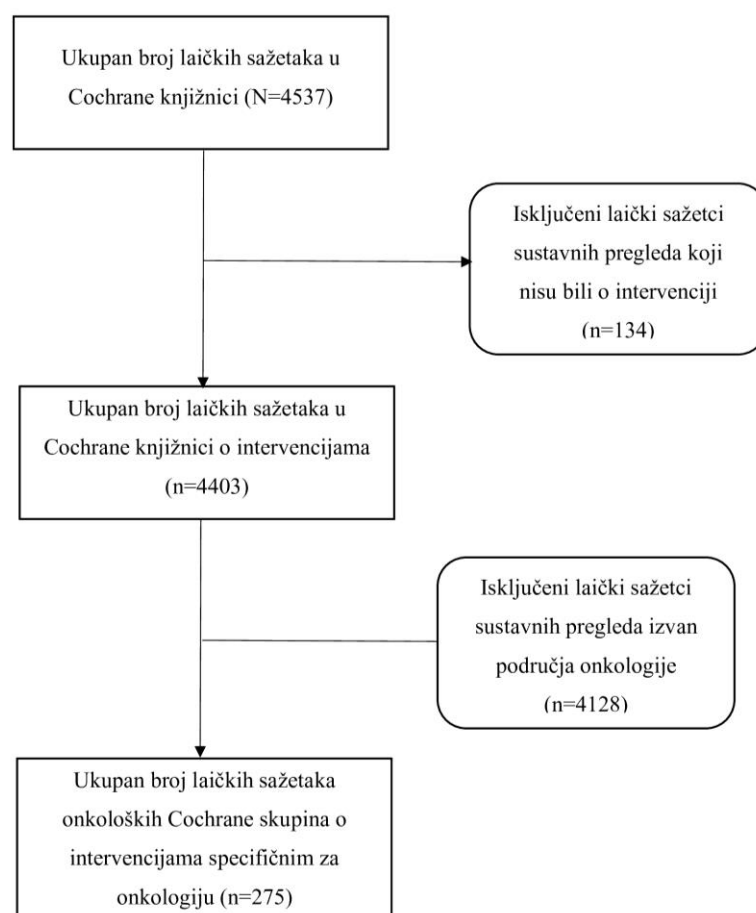
Analize su provedene korištenjem JASP v.0.18.1.0 (Jasp Team 2023) i R v4.3.3. (R core team, 2024).

4. REZULTATI

Jezična analiza laičkih i pripadajućih znanstvenih sažetaka Cochrane sustavnih pregleda onkoloških intervencija

4.1.1 Opis uzorka

Iz šest Cochrane skupina koje su uvrštene u analizu, obrađeno je ukupno 275 laičkih, uz jednak broj odgovorajućih znanstvenih sažetaka. Analizirani su sažeci sustavnih pregleda koji su istraživali određenu onkološku intervenciju (Slika 4). Dio rezultata je prikazan u diplomskom radu autorice disertacije (134). Medijan godine objave sustavnih pregleda je 2015 (IQR 2012-2017).



Slika 4 . Dijagram selekcije laičkih sažetaka.

4.1.2 Analiza čitkosti

Laički sažeci su bili kraći za čitati u usporedbi s odgovarajućim znanstvenim sažecima (Tablica 1, Tablica 2) – s medijanom od 364 riječi po sažetku (95% CI 339- 388) (Tablica 1). Naspram tome, broj riječi po znanstvenom sažetku je bio značajno veći u odnosu na laičke ($p < 0,001$; Wilcoxonov neparametrijski test uparenih uzoraka), s medijanom od 604 riječi po sažetku (95% CI 566- 653) (Tablica 2). Ukoliko bi radili analizu po skupinama, sažeci s najmanjim brojem riječi su bili iz Skupine koja je obrađivala maligne bolesti debelog crijeva (Tablica 1, Tablica 2). Što se tiče čitkosti, određivana je korištenjem SMOG indeksa. Laički sažeci su bili čitkiji u komparaciji sa znanstvenim sažecima. SMOG indeks za PLS-ove bio je veći od preporučene razine od 6 godina obrazovanja, odnosno medijan PLS-ova je bio 13,0 (95% CI = 12,8–13,3), dok je SMOG indeks znanstvenih sažetaka imao medijan od 16,6 (95% CI = 16,4–16,8) (Tablica 1, Tablica 2). Analizom sažetaka ovisno o tome kojoj Cochrane skupini pripadaju, odnosno koji tip tumora se istraživao, uočeno je da nema značajne razlike u čitkosti laičkih, kao ni znanstvenih sažetaka, između šest analiziranih skupina malignih bolesti (Tablica 1, Tablica 2).

Tablica 1. Usporedba obilježja laičkih sažetaka po broju autora i čitkosti u šest onkoloških grupa (n=275)

	Cochrane skupina (medijan, 95% raspon pouzdanosti)						P*
	Skupina za rak dojke (n=35)	Skupina za maligne bolesti u djece (n=29)	Skupina za kolorektalni karcinom (n=79)	Skupina za ginekološke, neuroonkološke i rijetke tumore (n=43)	Skupina za hematološke maligne bolesti (n=67)	Skupina za rak pluća (n=22)	
Broj autora	5 (4-6)	5(4-5)†‡	5 (4-5)†‡	6 (5-6)	6 (5-6)	5 (4-5.5)	0.002
Broj riječi po sažetku	400 (367- 433)	314 (241- 393)§	221 (166 - 268)§	410 (372-440)	475 (427- 526)	279 (200- 376)§‡	<0.001
Čitkost (SMOG indeks)	12.8 (11.6- 13.1)	12.5 (11.8- 13.1)	13.3 (12.6- 13.4)	13.3 (12.7-13.6)	13.0 (12.7- 13.2)	12.4 (11.8- 13.0)	0.053

*Kruskal Wallis test i Conover Iman *post-hoc* test.

†Značajno različito od Skupine za kolorektalni karcinom.

‡Značajno različito od Skupine za ginekološke, neuroonkološke i rijetke tumore.

Tablica 2. Usporedba obilježja znanstvenih sažetaka po indeksu čitkosti među 6 različitim onkoloških skupina (n=275)

Cochrane skupina (medijan, 95% raspon pouzdanosti)							P*
	Skupina za rak dojke(n=35)	Skupina za maligne bolesti u djece (n=29)	Skupina za kolorektalni karcinom(n=79)	Skupina za ginekološke, neuroonkološke i rijetke tumore (n=43)	Skupina za hematološke maligne bolesti (n=67)	Skupina za rak pluća (n=22)	
Broj riječi	575 (450-723) ^{†‡}	643 (538-715) [†]	458 (420-492)	715 (642-765) [†]	714 (626-793) [†]	604 (497-694) ^{†‡}	<0.001
Čitkosit (SMOG indeks)	16.5 (16.0-17.1)	16.4 (15.4-16.5)	16.4 (16.0-16.8)	16.9 (16.2-17.4)	16.7 (16.4-16.8)	16.7 (15.7-17.3)	0.462

*Kruskal Wallis test i Conover Iman *post-hoc* test.

[†]Značajno različito od Skupine za kolorektalni karcinom.

[‡]Značajno različito od Skupine za ginekološke, neuroonkološke i rijetke tumore.

4.1.3 Analiza jezičnih obilježja

U laičkim sažecima svih onkoloških skupina je primjećen visok postotak korištenja analitičkog tona (Tablica 3). Takav trend je zabilježen i u analizi znanstvenih sažetaka (Tablica 4), ali u većem omjeru nego u sažecima pisanim za opću populaciju (Tablica 10). Najmanje autentičnosti u tonu je bilo izraženo u laičkim sažecima Skupine za hematološke maligne bolesti. Skupine koje se bave zloćudnim tumorima pluća, rakom dojke i tumorima debelog crijeva i rektuma su bile emocionalnije u izražavanju svojih rezultata u odnosu na sažetke sustavnih pregleda pedijatrijskih maligniteta (Tablica 3). Najmanja učestalost korištenja govora s pozicije moći je zabilježena u sažecima koji su istraživali liječenja u karcinomima kolorektuma - i u laičkim i u znanstvenim sažecima (Tablica 3, Tablica 4). Skupina za hematološke zloćudne bolesti je imala najmanji udio autentičnosti u predstavljanju sustavnih pregleda, ali isto tako i najmanje riječi koje su bile vezane uz analitički ton (Tablica 4).

Tablica 3. Usporedba obilježja laičkih sažetaka po jezičnim obilježjima u šest onkoloških grupa (n=275)

	Cochrane skupina (medijan, 95% raspon pouzdanosti)					P*
	Skupina za rak dojke (n=35)	Skupina za maligne bolesti u djece (n=29)	Skupina za kolorektalni karcinom (n=79)	Skupina za ginekološke, neuroonkološke i rijetke tumore (n=43)	Skupina za hematološke maligne bolesti (n=67)	Skupina za rak pluća (n=22)
Korištenje analitičkog tona	95.4 (94.8- 96.2)	95.5 (93.8- 96.4)	96.5 (95.5- 96.8)	93.6 (91.5-94.8)	95.8 (94.6- 96.4)	93.8 (90.9- 95.5)
Korištenje govora s pozicije moći	50.8 (45.4- 57.8)†	50.0 (43.3- 55.4)†	43.8 (40.4- 47.5)	55.3 (51.1- 56.2)†	49.0 (46.5- 53.8)†‡	52.3 (46.8- 53.6)†
Udio autentičnosti u tonu	23.3 (17.0- 31.3)¶	25.4 (20.5- 30.7)¶	27.3 (19.0- 30.3)¶	27.2 (21.8- 31.2)¶	19.8 (15.0- 22.8)	34.8 (18.9- 40.5)¶
Korištenje emocionalnog tona	28.1 (17.5- 39.3)¶¶	10.5 (3.7- 25.8)	25.8 (16.8- 25.8)¶¶	18.0 (12.9-30.7)	16.7 (11.2- 19.9)	25.8 (13.7- 30.1)¶¶

*Kruskal Wallis test i Conover Iman *post-hoc* test.

†Značajno različito od Skupine za kolorektalni karcinom.

‡Značajno različito od Skupine za ginekološke, neuroonkološke i rijetke tumore.

§Značajno različito od Skupine za rak dojke.

¶Značajno različito od Skupine za maligne bolesti u djece.

¶¶Značajno različito od Skupine za hematološke maligne bolesti.

Tablica 4. Usporedba karakteristika znanstvenih sažetaka po jezičnim obilježjima među 6 različitih onkoloških grupa (n=275)

	Cochrane skupina (median, 95% raspon pouzdanosti)						P*
	Skupina za rak dojke(n=35)	Skupina za maligne bolesti u djece (n=29)	Skupina za kolorektalni karcinom(n=79)	Skupina za ginekološke, neuroonkološke i rijetke tumore (n=43)	Skupina za hematološke maligne bolesti (n=67)	Skupina za rak pluća (n=22)	
Korištenje analitičkog tona	96.8 (96.0-97.3)	97.4 (97.1-97.6)	96.3 (95.9-96.6)¶	96.4 (95.4-96.6)¶	96.8 (96.1-97.1)¶	97.1 (95.8-97.6)†	<0.001
Korištenje govora s pozicije moći	53.5 (48.4-55.9)†	51.6 (44.6-54.2)†	45.9 (43.1-47.8)	55.0 (52.4-56.4)†	51.9 (48.6-53.8)†‡	49.7 (46.7-57.6)†	<0.001
Udio autentičnosti u tonu	17.1 (14.9-19.8)	19.7 (15.1-26.8)¶	20.0 (16.9-2.0)¶	18.4 (15.7-21.5)¶	14.6 (13.8-17.2)	23.7 (13.7-29.5)¶	<0.020
Korištenje emocionalnog tona	25.8 (17.3-37.2)¶	17.3 (11.5-22.2)	25.8 (17.6-25.8)¶	18.2 (10.8-24.0)	11.6 (9.5-16.6)	26.1 (14.0-37.8)¶	<0.001

*Kruskal Wallis test i Conover Iman *post-hoc* test.

†Značajno različito od Skupine za kolorektalni karcinom.

‡Značajno različito od Skupine za ginekološke, neuroonkološke i rijetke tumore.

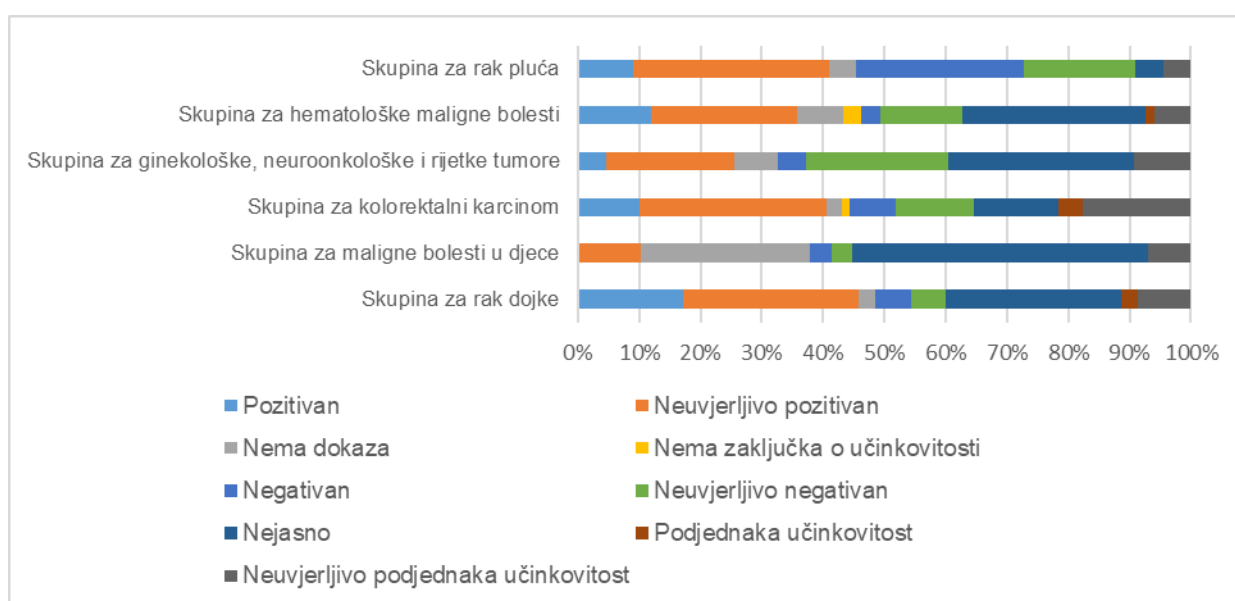
§Značajno različito od Skupine za rak dojke.

¶Značajno različito od Skupine za hematološke maligne bolesti.

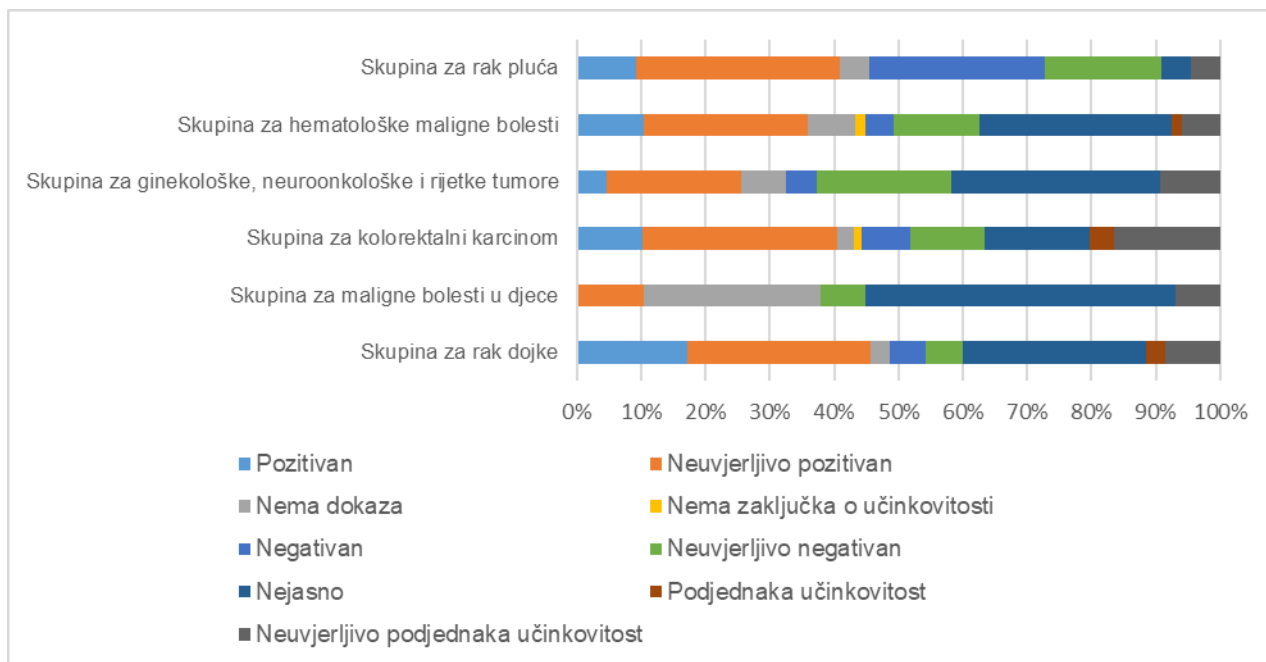
4.1.4 Analiza zaključivosti

Cochrane onkološke skupine su analizirane i podijeljene ovisno o tome što bi se zaključilo iz određenog sažetka iz dane skupine o učinkovitosti onkološke intervencije. Prema tome su sažeci mogli biti raspoređeni u jednu od devet mogućih kategorija zaključivosti. Primjećeno je da je u većem broju laičkih sažetaka donesen zaključak “Nejasno”, odnosno upućuje se da autori nisu imali dovoljno dokaza da bi donijeli konačan sud o učinkovitosti intervencije (Slika 4, Tablica 5). Ta kategorija zaključka je bila najviše

zastupljena u četiri od šest Cochrane skupina (Tablica 5). Analizom po skupinama, zapaženo je da je u laičkim sažecima koji su se bavili intervencijama za liječenje kolorektalnog karcinoma najviše bio zastupljen “Neuvjerljivo pozitivan” zaključak. Nijedan sustavni pregled u istraživanjima dječjih zloćudnih tumora nije dao definitivno pozitivan zaključak (Slika 5). Ista skupina je, u odnosu na ostale ispitivane tipove tumora, imala najveći postotak “Nema dokaza” i “Nejasno” zaključaka. (Tablica 5). Skupina za rak pluća je u odnosu na ostale tipove malignih bolesti imala najviše negativnih zaključaka, odnosno onih koji nisu išli u prilog učinkovitosti ili sigurnosti ispitivanog liječenja (Slika 5). Znanstveni sažeci su donijeli zaključke koji su bili sukladni onima u laičkim sažecima (Slika 6).



Slika 5. Raspodjela kategorija zaključaka u laičkim sažecima Cochrane sustavnih pregleda onkoloških intervencija iz šest Cochrane onkoloških skupina.



Slika 6. Raspodjela kategorija zaključaka u znanstvenim sažecima Cochrane sustavnih pregleda onkoloških intervencija.

Tablica 5. Raspodjela kategorija zaključaka među 6 grupa Cochrane sustavnih pregleda onkoloških intervencija u Cochrane laičkim sažecima (n=275).

Skupine (n, %)						
Kategorija	Skupina za rak dojke(n=35)	Skupina za maligne bolesti u djece (n=29)	Skupina za kolorektalni karcinom (n=79)	Skupina za ginekološke, neuroonkološke i rijetke tumore (n=43)	Skupina za hematoloških maligne bolesti (n=67)	Skupina za rak pluća (n=22)
Pozitivan	6 (17%)	0 (0%)	8 (10%)	2 (5%)	8 (12%)	2 (9%)
Neuvjerljivo pozitivan	10 (28%)	3 (10%)	24 (30%)	9 (21%)	16 (24%)	7 (32%)
Nema dokaza	1 (3%)	8 (28%)	2 (2,53%)	3 (7%)	5 (8%)	1 (4%)
Nema zaključka o učinkovitosti	0 (0%)	0 (0%)	1 (1%)	0 (0%)	2 (3%)	0 (0%)
Negativan	2 (6%)	1 (3%)	6 (8%)	2 (5%)	2 (3%)	6 (27%)
Neuvjerljivo negativan	2 (6%)	1 (3%)	10 (13%)	10 (23%)	9 (13%)	4 (18%)
Nejasno	10 (28%)	14 (49%)	11 (14%)	13 (30%)	20 (30%)	1 (5%)
Podjednaka učinkovitost	1 (3%)	0 (0%)	3 (4%)	0 (0%)	1 (1%)	0 (0%)
Neuvjerljivo podjednaka učinkovitost	3 (9%)	2 (7%)	14 (18%)	4 (9%)	4 (6%)	1 (5%)

4.1.5 Analiza čitkosti u odnosu prema zaključivosti

Laički sažeci su podijeljeni u skupine prema kategorijama zaključaka i potom su analizirana njihova jezična obilježja. Laički sažeci koji su donijeli definitivno pozitivan zaključak o učinkovitosti određenog načina liječenja su bili teži za čitati od sažetaka koji su kategorije zaključka “Neuvjerljivo pozitivan“. Pokazano je da su najlakši za čitati sažeci kategorije “Nema dokaza“, odnosno oni koji su iznijeli da nema odgovarajuće studije koja bi dala odgovor na istraživačko pitanje. (Tablica 6).

Što se tiče znanstvenih sažetaka, tu nije utvrđena razlika u čitkosti u različitim kategorijama zaključka (Tablica 7).

Tablica 6. Usporedba jezičnih obilježja laičkih sažetaka po kategorijama zaključaka

	Medijan (95% raspon pouzdanosti)						P*
	Pozitivna n (n=26)	Neuvjerljivo pozitivan (n=69)	Nema dokaza (n=20)	Negativna n (n=19)	Neuvjerljivo negativan (n=36)	Nejasno (n=69)	Neuvjerljivo o podjednaka učinkovitost (n=28)
Čitkost	13.6	12.8 (12.1-	12.0	13.5	12.6 (12.1 -	12.9	13.4 (13.2-
(SMOG	(12,4-	13.0)	(11.3-	(12.5-	13.7)†	(12.1-	13.9)†
-indeks)	14.3)‡		12.4)	14.1)†		12.9)†	

Isključene kategorije: “Nema zaključka o učinkovitosti” i “Podjednaka učinkovitost” kategorije zbog male veličine uzorka.

*Kruskal Wallis test i Conover Iman *post-hoc* test.

†Značajno različito od “Nema dokaza”.

‡ Značajno različito od “Neuvjerljivo pozitivan”.

Tablica 7. Usporedba jezičnih obilježja među znanstvenim sažecima podijeljenima s obzirom na kategoriju zaključka

	Median, 95% raspon pouzdanosti)						P*	
	Pozitiva n (n=25)	Neuvjerljiv o pozitivan (n=70)	Nema dokaz a (n=20)	Negativa n (n=19)	Neuvjerljiv o negativan (n=35)	Nejasn o (n=72)	Neuvjerljiv o podjednaka učinkovitost (n=27)	
Čitkost	16.7		16.2	16.3		16.7		
(SMOG	(16.0-	16.5 (16.2-	(15.6-	(15.8-	16.6 (15.5-	(16.3-	16.5 (15.9-	0.75
-indeks)	17.2)	17.0)	16.8)	17.2)	17.0)	16.9)	17.4)	3

Isključene kategorije: “Nema zaključka o učinkovitosti” i “Podjednaka učinkovitost” kategorije zbog male veličine uzorka.

*Kruskal Wallis test i Conover Iman *post-hoc* test.

4.1.6 Analiza jezičnih obilježja prema zaključivosti

Ponovno je napravljena podjela sažetaka prema kategoriji zaključka te su analizirana jezična obilježja: broj riječi i korištenje različitim emocionalnih tonova. Broj riječi za sažetke različitih skupina tumora se nije signifikantno razlikovao. Također, nije bilo značajne razlike u korištenju autentičnosti, emocionalnoga ton i govora s pozicije moći u različitim skupinama. Što se tiče analitičkog tona, značajno više je bio upotrebljen u kategoriji „Pozitivan“ u odnosu na ostale zaključke (Tablica 8). U znanstvenim sažecima to nije bio slučaj-svi analizirani tonovi se nisu značajno razlikovali među različitim skupinama. Što se tiče duljine teksta, pokazalo se da su znanstveni sažeci čije su analizirane intervencije donijele kategoriju zaključka “Negativan” i “Nema dokaza” koristili u prosjeku manje riječi u odnosu na ostale kategorije zaključka (Tablica 9).

Tablica 8. Usporedba jezičnih obilježja među laičkim sažecima po kategorijama zaključaka

Medijan, 95% raspon pouzdanosti)							P*	
	Pozitivna n (n=26)	Neuvjerljiv o pozitivan (n=69)	Nema dokaz a (n=20)	Negativna n (n=19)	Neuvjerljiv o negativan (n=36)	Nejasn o (n=69)	Neuvjerljiv o podjednak a učinkovitost (n=28)	
Broj riječi	384 (22.9- 478)	343 (313- 386)	382 (216- 416)	200 (150- 281)	374 (261- 426)	395 (351- 414)	353 (242- 392)	0.05 2
Analitički ton	97.4 (96.2 - 98.2)	95.3 (94.4- 96.3)§	94.1 (91.4- 95.8)§	95.0 (92.0- 95.6)§	94.2 (93.4- 95.4)§	95.6 (94.6- 96.1)§	95.6 (93.2- 96.9)§	0.00 8
Govor s pozicije moći	46.6 (40.1- 52.2)	50.0 (43.0- 52.8)	52.4 (44.9- 56.0)	46.1 (32.8- 48.7)	49.7 (44.5- 52.3)	55.3 (47.0- 57.3)	47.6 (40.9- 51.4)	0.05 8
Autentičnost	25.5 (17.5- 27.3)	27.2 (19.0- 30.1)	19.4 (13.3- 25.7)	31.2 (17.1- 37.2)	21.2 (16.4- 30.7)	23.3 (20.2- 27.0)	27.5 (19.9- 32.7)	0.49 8
Emocionalni i ton	25.8 (12.1- 27.5)	20.4 (15.4- 25.8)	25.8 (7.2- 42.6)	17.2 (8.7- 31.2)	22.8 (12.5- 30.1)	19.9 (13.6- 25.8)	25.8 (12.0- 29.4)	0.95 3

Isključene kategorije: “Nema zaključka o učinkovitosti” i “Podjednaka učinkovitost” kategorije zbog male veličine uzorka.

*Kruskal Wallis test i Conover Iman *post-hoc* test.

§ Značajno različito od Pozitivno.

Tablica 9. Usporedba jezičnih obilježja među znanstvenim sažecima podijeljenima s obzirom na kategoriju zaključka

medijan, (95% raspon pouzdanosti)							<i>P*</i>
	Pozitivna n (n=25)	Neuvjerljiv o pozitivan (n=70)	Nema dokaz a (n=20)	Negativna n (n=19)	Neuvjerljiv o negativan (n=35)	Nejasn o (n=72)	Neuvjerljiv o podjednak a učinkovitost t (n=27)
Broj riječi	603 (407- 746)†‡	692 (581- 743)†‡	459 (346to 527)	425 (374- 485)	645 (526- 696)†‡	633 (577- 724)†‡	566 (437- 670)‡ <0.00 1
Analitički ton	97.0 (96.2- 97.3)	96.2 (95.6- 97.0)	96.4 (94.1to 97.3)	95.6 (94.6- 96.7)	96.7 (96.1- 97.3)	96.8 (96.4- 97.2)	97.1 (96.2- 97.2) 0.164
Govor s pozicije moći	48.8 (44.5- 53.2)	48.7 (46.0- 52.4)	52.2 (50.0to 55.6)	47.4 (41.0- 51.4)	51.7 (45.5- 54.8)	53.1 (48.7- 55.0)	50.7 (47.2- 52.6) 0.293
Autentično st	16.1 (12.8- 23.2)	20.0 (18.0- 22.8)	17.6 (14.1- 19.2)	23.5 (18.7- 28.0)	14.2 (11.8- 17.3)	17.4 (14.7- 22.0)	19.2 (14.1- 22.5) 0.050
Emocional ni ton	25.8 (14.2- 31.4)	22.1 (11.8- 27.7)	27.4 (16.5- 39.1)	20.5 (11.5- 40.6)	19.4 (12.2- 29.0)	17.3 (14.4- 20.8)	22.2 (11.5- 23.4) 0.364

Isključene kategorije: “Nema zaključka o učinkovitosti” i “Podjednaka učinkovitost” kategorije zbog male veličine uzorka.

*Kruskal Wallis test i Conover Iman *post-hoc* test.

†Značajno različito od “Nema dokaza”.

‡ Značajno različito od “Negativan”.

4.1.7 Sumarizacija usporedbe laičkih i znanstvenih sažetaka

U analizi čitkosti, laički sažeci su bili lakši za čitati u usporedbi sa sukladnim znanstvenim sažecima. Što se tiče jezičnih obilježja, laički sažeci su koristili manje analitičkog, a više autentičnog i emocionalnog tona u odnosu na znanstvene sažetke (Tablica 10).

Tablica 10. Usporedba jezičnih obilježja između laičkih i pripadajućih znanstvenih sažetaka

	Medijan (95% raspon pouzdanosti)		<i>P*</i>
	Laički sažetci	Znanstveni sažetci	
Broj riječi	364 (339- 388)	604 (566- 653)	<0.001
Čitkost	13.0 (12.8-13.3)	16.6 (16.4-16.8)	<0.001
Analitički ton	95.5 (95.0-95.8)	96.6 (96.4-97.0)	<0.001
Govor s pozicije moći	50.0 (47.7-51.8)	50.7 (48.8-52.4)	0.181
Autentičnost	24.1 (21.8-27.2)	17.9 (17.0-19.5)	<0.001
Emocionalni ton	22.1 (18.0-25.8)	20.4 (17.5-23.4)	0.014

*Wilcoxon neparametrijski upareni test uzoraka.

4.2 Korištenje umjetne inteligencije za poboljšanje prezentacije laičkih sažetaka Cochraneovih sustavnih pregleda o onkološkim intervencijama

Generirali smo ukupno 275 PLS-ova za svaki od tri ChatGPT upita (engl.*prompt*). U prosjeku, svi su imali statistički manje riječi od originalnih PLS-ova (Tablica 11).

4.2.1 SMOG indeks i lingvističke karakteristike

PLS-ovi generirani prvim upitom (napiši PLS sa SMOG indeksom 6) bili su najlakši za čitanje, s medijanom SMOG indeksa od 8,2 (95% interval pouzdanosti (CI)=8–8,4) i najkraći. Što se tiče lingvističkih karakteristika, ovi PLS-ovi imali su manje analitičkog tona, a više autentičnosti, govora s pozicije moći i emotivnog tona u usporedbi s originalnim PLS-ovima koje su napisali autori (Tablica 11).

PLS-ovi generirani drugim upitom (učini PLS iz prvog upita emotivnijim) imali su medijan SMOG indeks od 11,4 (95% CI=11,1–12). Ti PLS-ovi također su imali više analitičkog tona i govora s pozicije moći u usporedbi s PLS-ovima generiranim prvim upitom, ali nije bilo razlike u emotivnom tonu. Također su koristili manje riječi povezanih s autentičnošću nego PLS-ovi generirani prvim upitom (Tablica 11).

PLS-ovi generirani trećim upitom (napiši jednostavniji PLS od originala) imali su medijan SMOG indeksa od 8,7 (95% CI=8,4–8,8). Lingvističke karakteristike nisu se razlikovale od onih kod PLS-ova generiranih prvim upitom, ali su imali manje analitičkog, a više autentičnog tona nego PLS-ovi generirani drugim upitom (Tablica 11).

Kod svih GPT upita rezultati su bili konzistentni, bez značajnih odstupanja.

Tablica 11. Usporedba jezičnih karakteristika (medijan, 95% raspon pouzdanosti) između različitih grupa jezičnih sažetaka.

Izvorni sažetak¶		Upit ChatGPT-u			
		Prvi upit – napiši PLS na temelju SA na SMOG razini 6	Drugi upit – učini PLS iz prvog upita emocionalnijim	Treći upit – napiši jednostavniji PLS	P*
Broj riječi	364 (339-388)	240 (232- 248)	285 (278-292)	273 (266-278)	<0.001
SMOG indeks †	13.1 (12.9-13.4)	8.2 (8.0-8.4)	11.4 (11.1-12)	8.7 (8.4-8.8)	<0.001
Jezične karakteristike‡:					
Analitički ton	95.5 (95.0-95.8)	55.9 (53.6-57.9)§	85.7 (84.0-86.7)	60.9 (57.7-63.1)	<0.001
Govor s pozicije moći	50.0 (47.7-51.8)	67.2 (64.9-70.9)§	80.3 (77.8-83.6)	70.5 (68.0-72.8)	<0.001
Autentičnost	28.6 (26.2-31.3)	50.5 (47.0-53.5)§	38.0 (34.8-40.1)	49.4 (45.8-54.2)	<0.001
Emocionalni ton	22.1 (18.5-26.2)	54.8 (51.3-58.7)§,	63.9 (58.6-69.6)§	54.4 (51.9-56.4)	<0.001

*Friedman neparametrijski test. Sve post hoc razlike su bile statistički značajne osim onih koje su označene simbolima.

†Čitkost je mjerena kao SMOG index (135). Viši SMOG index označava lošiju čitkost.

‡Jezična obilježja teksta su bila mjerena programom Linguistic Inquiry and Word Count – LIWC (136).

Varijable su prezentirane kao postoci riječi koje su povezane s određenom kategorijom.

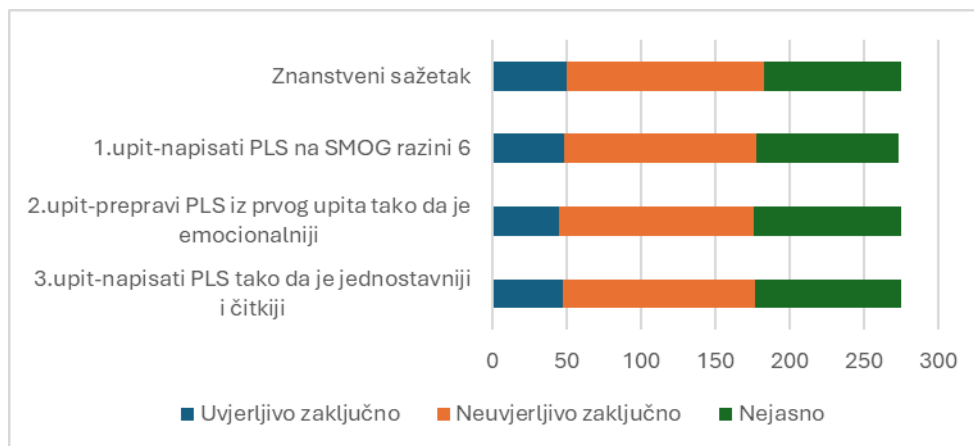
§Ne razlikuje se od Upita 3.

||Ne razlikuje se od Upita 2.

¶Rezultati za izvorne laičke sažetke su iz prve studije ove disertacije (132).

4.2.2 Zaključivost

Kategorija zaključivosti svih triju ChatGPT-generiranih PLS-ova nije se razlikovala od one originalnog znanstvenog sažetka (Slika 7).



Slika 7. Raspodjela laičkih sažetaka prema kategoriji zaključivosti o učinkovitosti intervencija opisanih u sustavnim pregledima unutar četiri različite skupine-originalni znanstveni sažetak i laički sažeci generirani od tri upita prema ChatGPT-u.

5. RASPRAVA

5.1 Jezična analiza Cochrane laičkih i znanstvenih sažetaka sustavnih pregleda onkoloških intervencija

5.1.1 Glavni rezultati i njihov značaj

Rezultati ovog istraživanja ističu izazove i moguća poboljšanja u razumljivosti i dostupnosti laičkih sažetaka Cochrane sustavnih pregleda onkoloških intervencija.

Prvo istraživanje pokazalo je da je indeks čitkosti postojećih PLS-ova previsok za opću populaciju, zahtijevajući prosječno 13 godina obrazovanja za razumijevanje danog teksta. Prosječan broj godina obrazovanja potreban za razumijevanje znanstvenih sažetaka je 16,6 godina. Iako možemo zaključiti da su PLS-ovi čitkiji od znanstvenih sažetaka, i dalje su prekomplikirani za opću javnost, posebno imajući na umu da je preporučena razina čitkosti koju savjetuju krovne američke zdravstvene institucije 6 godina obrazovanja (137). Broj riječi laičkih sažetaka u odnosu na znanstvene sažetke je bio značajno manji (prosječno 364 riječi za laički sažetak naspram 604 riječi za znanstveni), međutim to smanjenje teksta nažalost ne prati i poboljšanje čitkosti. Presložen jezik koji se koristi u laičkim sažecima može otežati donošenje informiranih odluka i dovodi do upitnosti svrhe takvih laičkih sažetaka. Stoga je nužno osigurati da PLS-ovi budu dovoljno jednostavni za razumijevanje, bez kompromisa u točnosti sadržaja. Treba imati na umu da su u onkologiji nazivi terapija često kompleksni pa je upitno i koliko možemo očekivati da će tekstovi koji sumiraju sustavni pregled biti napisani na razini 6. razreda osnovne škole. Pretjerano pojednostavljivanje može dovesti do netočnog tumačenja ili izostavljanja važnih informacija te se, npr., neki stručni izrazi kao što je ime lijeka trastuzumab-deruxtecan ne mogu zamijeniti jednostavnijim nazivom bez gubitka značenja. Dalje, velik broj PLS-ova nastaje bez suradnje s laicima ili pacijentima. To znači da nema sustavnog uvida u to što je ljudima stvarno razumljivo, a što nije, te nam je i dalje glavni oslonac mjerenje razine čitkosti pomoću različitih alata. Isto tako, ne postoji univerzalna metodologija koja bi osigurala dosljednu čitkost i pristupačnost PLS-ova među različitim publikacijama i terapijskim područjima. Stoga je potreban multimodalitetan pristup u poboljšanju PLS-ova koji bi djelovao na nekoliko razina: 1) rad autora znanstvenih članaka u predstavljanju informacija javnosti, ne samo kroz praćenje smjernice, nego i kroz potencijalnu suradnju s novinarima i udrugama pacijenata; 2) uključiti i pacijente i laike u proces izrade i testiranja PLS-ova, primjerice kroz fokus grupe, kako bi se osigurala relevantnost, razumljivost i korisnost

sadržaja; 3) istražiti različite načine prikaza informacija-infografike, ikone ili jednostavni dijagrami mogu pomoći u dodatnom pojašnjenju kompleksnih koncepata.

Jezične karakteristike poput autentičnosti, govora s pozicije moći, emocionalnosti i analitičkog tona imaju značajan utjecaj na čitkost teksta. Autentičan jezik, koji djeluje iskreno i prirodno, povećava pristupačnost teksta i olakšava povezivanje čitatelja s porukom, dok govor s pozicije moći, iako može ojačati dojam stručnosti i autoriteta, često smanjuje razumljivost ako koristi prekompleksan ili distanciran stil izražavanja. Emocionalnost može dodatno angažirati čitatelja i naglasiti važnost poruke, ali pretjerana emocionalna obojenost može umanjiti jasnoću ako odvlači pažnju od suštine. S druge strane, analitičan ton doprinosi logičkoj strukturi i preciznosti, što je korisno za prenošenje složenih informacija, no ako nije uravnotežen s primjerima ili ljudskim elementom, može djelovati hladno i apstraktno. Ono što je pokazano u našim istraživanjima jest da laički sažeci i dalje imaju velik udio analitičkog tona, odnosno, ako sumiramo rezultate laičkih i znanstvenih sažetaka primjećujemo da su omjeri korištenja emocionalnog i analitičkog tona, autentičnosti, kao i govora s pozicije moći bez značajnih razlika, odnosno da se jezična obilježja znanstvenog sažetka prelijevaju i u pripadajući laički sažetak. Ukoliko analiziramo podatke prema onkološkim Cochrane skupinama, opet primjećujemo da sve skupine imaju visok udio analitičkog tona, dok je udio autentičnosti i emocionalnog tona značajno manji. U skupini za maligne bolesti djece samo 10% riječi je bilo povezano s pozitivnim emocionalnim tonovima, 25% s autentičnošću, naspram 95.5% udjela analitičkog tona u riječima. Treba imati na umu da su autori laičkih sažetaka autori samog sustavnog pregleda i da je većini njih analitički, više znanstveni dio izražavanja bliži i da nisu svi jednako vješti u približavanju informaciju kroz personaliziranje, tj. korištenje više autentičnog tona, kao i toga da se zadrži točan omjer pozitivnosti, odnosno da se eksprimira emocionalni ton kako informacija ne bi krivo bila shvaćena kao negativna. U tome bi znanstvenicima mogle pomoći ili obučene osobe za predstavljanje informacija javnosti (npr. novinari, komunikolozi, PR stručnjaci), kao i korištenje umjetne inteligencije u balansiranju jezičnih tonova.

Zaključak "Nejasno", odnosno konstatacija autora da se ne može donijeti jasan sud o učinkovitosti intervencije, je bio najčešći u četiri Cochrane onkološke skupine, dok je u jednoj skupini bio drugi najčešći. Ono što je posebno zanimljivo u rezultatima je da je u Cochrane skupini koja se bavi istraživanjima maligniteta u pedijatrijskoj populaciji, kategorije zaključka „Nejasno“ činila skoro polovinu (49%) zaključaka svih sažetaka sustavnih pregleda onkoloških intervencija, dok nijedna intervencija nije pokazala uvjerljivo pozitivan rezultat naglašavajući poteškoće u provođenju randomiziranih kontroliranih

ispitivanja u pedijatrijskoj onkologiji. U Skupini za rak dojke se prati najveći postotak uvjerljivo pozitivnih zaključaka za danu intervencije, koji iznosi 17%, što nas upućuje na relativno skromne rezultate u ostalim skupinama, ali i odražava visoku učestalost kliničkih ispitivanja i napredak u liječenju raka dojke. Suprotno tome, PLS-ovi iz skupine za rak pluća sadržavali su najveći udio negativnih zaključaka, odražavajući stalne izazove u liječenju raka pluća i njegovu visoku stopu smrtnosti. Ovi nalazi nam jasno ukazuju na potrebu za dodatnim istraživanjima kako bi se omogućilo donošenje čvrstih zaključaka o učinkovitosti ispitivanih terapija ili intervencija, ali isto tako ukazuju i da očito postoje prepreke u provođenju istih. Ono što moramo imati na umu je da su u razvoju novog lijeka protiv raka, klinička ispitivanja ključni dio procjene učinkovitosti nove terapijske opcije. Međutim, procjenjuje se da manje od 5 % odraslih oboljelih od raka sudjeluje u kliničkim ispitivanjima (138,139). S jedne strane postoji potreba za većim brojem kliničkih ispitivanja, ali s druge strane, u njima sudjeluje mali broj pacijenata, iako se čini da je 70 % oboljelih od raka u Sjedinjenim Američkim Državama voljno sudjelovati u kliničkim ispitivanjima (138,140). Tijekom posljednja dva desetljeća zabilježen je nagli porast broja kliničkih ispitivanja u onkologiji, što odgovara rastućim javnozdravstvenim potrebama (141). Rak je danas drugi vodeći uzrok smrtnosti u svijetu, s porastom incidencije i stalnom potražnjom za inovativnim terapijama (142). Iako je onkološko istraživanje u porastu u mnogim zemljama, i dalje su prisutne izražene nejednakosti između različitih skupina zemalja prema dohotku (141,143). Zemlje s visokim dohotkom i dalje čine većinu kliničkih ispitivanja, zahvaljujući razvijenoj istraživačkoj infrastrukturi, dostupnosti stručnog kadra te boljim mogućnostim financiranja iz javnog i privatnog sektora (141,144). Nasuprot tome, zemlje s niskim dohotkom suočavaju se s brojnim preprekama: slabom infrastrukturom, manjkom obučenog osoblja, nedovoljnom regulacijom i oskudnim financijskim sredstvima, što značajno usporava njihov razvoj u području kliničkih ispitivanja (141,144,145) te se tek 5 % svjetskih sredstava za istraživanja raka troši se u zemljama niskog i srednjeg dohotka (146,147). Potrebna je pravednija raspodjela istraživačkih resursa kako bi se učinkovito odgovorilo na globalni izazov raka (141,148). Veliki problem je što zemlje u kojima nisu dostupna klinička ispitivanja, po iskorištavanju svih standardnih terapijskih opcija, nemaju što dalje ponuditi pacijentima, a velika većina pacijenata nema dovoljno financijskih sredstava za odlazak na liječenje u inozemstvo. Taj problem je naglašeniji ukoliko se radi o rijetkim sijelima raka, u kojih ne postoji mnogo terapijskih linija i kliničko istraživanje nekada bude najbolje šansa za pacijenta, a isto tako pridonosi i poboljšanju znanja o toj bolesti. Uključivanje raznolikih populacija ima i znanstveno i etičko značenje. U kontekstu rastuće svijesti o važnosti

uključivanja ranjivih skupina (poput trudnica), uključivanje geografski i genetski raznolikih populacija, ključno je za razvoj učinkovitijih terapija i individualizacije pristupa svakom pacijentu. Određeni oblici raka imaju specifične karakteristike u pojedinim populacijama, što dodatno naglašava potrebu za uključivanjem različitih demografskih skupina u klinička ispitivanja, poput nazofaringealnog karcinoma u Aziji (149), raka vrata maternice u Africi (150) ili hepatobilijarnih karcinoma u Indiji i Jugoistočnoj Aziji (149,150) te ovo, uz sve gore navedeno, naglašava potrebu za većim uključenjem pacijenata u istraživanja. Što se tiče sažetaka, nejasan zaključak na kraju sažetka može biti izvor zabune za širu javnost te dovesti do nepovjerenja u onkologa koji bi se odlučio za određenu terapiju. Kako bi se to spriječilo, autori bi mogli navesti da sažetak ne daje jasan odgovor te preporučiti konzultaciju s kliničkim onkologom, osobito zato što se svaka odluka o liječenju treba temeljiti na procjeni onkološke vrijednosti terapije.

Ono što je još analizirano u studiji je koliko su čitke bile skupine sažetaka koje su bile podijeljene ovisno o tome kakav su zaključak donijele o danoj intervenciji. Laički sažeci koji su iznijeli da nema dokaza o određenom kliničkom pitanju su očekivano bili najlakši za pročitati, jer su zahtijevali manje objašnjenja. Čitkost u znanstvenim sažecima se nije razlikovala s obzirom na to koji zaključak donio sustavni pregled. Što se tiče jezičnih obilježja, laički sažeci koji su donijeli uvjerljivo pozitivan zaključak za intervenciju su koristili veći udio analitičkog tona. Sve navedeno je u skladu s prethodnim nalazima da znanstvena literatura, općenito, postaje sve teža za čitanje zbog uporabe tehničkog žargona.

5.1.2 Ograničenja istraživanja

Jedno od ograničenja istraživanja je to što su se rezultati oslanjali isključivo na Cochrane knjižnicu kao izvor sažetaka zdravstvenih informacija. Međutim, treba imati na umu prednost ovakvog pristupa jer su istraživanja iz različitih grana onkologije koristila ujednačen način prikazivanja informacija. Također, sustavni pregledi iz Cochrane baze pokazali su višu metodološku i izvještajnu kvalitetu u odnosu na one izvan Cochranea (153). Isto tako, na to koliko će tekst biti razumljiv i pristupačan mogu utjecati i vizualni aspekti, poput oblikovanje teksta i strukture, što nismo obuhvatili u ovoj studiji (154,155).

5.2 Uporaba umjetne inteligencije u stvaranju laičkih sažetaka

5.2.1 Glavni rezultati

Drugo istraživanje ispitalo je mogućnost korištenja umjetne inteligencije, konkretno ChatGPT-a, za generiranje PLS-ova koji su razumljiviji laičkoj publici. Rezultati su pokazali

da su PLS-ovi generirani ChatGPT-om kraći te imaju niži SMOG indeks (što ukazuje na bolju čitkost). Što se tiče jezičnih obilježja, dok su izvorni PLS-ovi imali visok udio analitičkog tona i često negativnu emocionalnu komponentu, AI-generirani sažeci pokazali su smanjenje analitičkog tona uz povećanje autentičnosti, pozitivnog emocionalnog tona i govora s pozicije moći. Ključno je da su unatoč ovim jezičnim prilagodbama, zaključci izvornog znanstvenog pregleda ostali nepromijenjeni u AI-generiranim sažecima, čime se osigurala vjerodostojnost prijenosa točnih informacija unatoč pojednostavljenju sadržaja.

Umjetna inteligencija, u najširem smislu, označava razvoj računalnih sustava sposobnih za obavljanje zadataka koji inače zahtijevaju ljudsku inteligenciju – poput razumijevanja jezika, zaključivanja, učenja i donošenja odluka. Veliki jezični modeli poput ChatGPT-a, trenirani na golemim količinama podataka, koriste se za generiranje teksta koji može simulirati ljudsku komunikaciju. U kontekstu zdravstvene komunikacije, AI se sve češće koristi za prevođenje složenih medicinskih informacija u jednostavnije, čitkije forme (156,157). Najnovija istraživanja u području primjene AI u medicini pokazuju kako AI može značajno poboljšati angažiranost korisnika bez kompromitiranja znanstvene točnosti (158,159). Naši nalazi podržavaju takve spoznaje, budući da su AI-generirani PLS-ovi bili dosljedno kraći, lakši za čitanje i imali su angažiraniji ton. Kroz različite upite, AI-generirani PLS-ovi imali su niže SMOG indekse (u rasponu od 8,2 do 11,2) u usporedbi s originalnim PLS-ovima (medijan SMOG indeksa od 13,1), iako su još uvijek iznad idealne razine od 6.

Još jedno važno otkriće bilo je da su AI-generirani PLS-ovi imali viši pozitivan emocionalni ton, veću sigurnost u izražavanju (engl. *clout*) i autentičnost u odnosu na originalne PLS-ove, čineći ih privlačnijima i osobnijima. Važno je napomenuti da je povećanje emocionalnog sadržaja poboljšalo angažman čitatelja, ali je također blago povećalo jezičnu složenost, što ukazuje na potrebu za ravnotežom između čitkosti i emocionalnog utjecaja. Dalje, bitno je istaknuti da su AI-generirani PLS-ovi zadržali istu kategoriju zaključivosti o učinkovitosti intervencija kao i originalni SA-ovi. Ova dosljednost sugerira da AI može biti vrijedan alat u približavanju zdravstvenih informacija javnosti, bez iskrivljavanja znanstvenih zaključaka.

Unatoč tim prednostima, AI-generirani PLS-ovi ne bi trebali biti smatrani samostalnim rješenjem. Korištenje AI-a u zdravstvenoj komunikaciji donosi izazove poput potencijalne pristranosti, nedostatka ažuriranja u stvarnom vremenu i potrebe za ljudskim nadzorom kako bi se osigurala točnost.

Dalje, kako se alati temeljeni na velikim jezičnim modelima, poput ChatGPT-a, sve više približavaju stvarnoj kliničkoj primjeni, otvaraju se nova regulatorna i etička pitanja.

Trenutačno, bez zakonskih izmjena, ne postoji jasna pravna ovlast koja bi regulirala većinu LLM-ova koji generiraju kliničke sažetke. Regulatorna pojašnjenja mogli bi značajno doprinijeti zaštiti pacijenata, a pritom zadržati prednosti koje ti alati donose. Bilo bi odgovorno provesti testiranja kliničke učinkovitosti u stvarnim uvjetima prije njihove široke primjene.

5.2.2 Ograničenja istraživanja

Ovi rezultati moraju se razmatrati u svjetlu nekoliko ograničenja. Prvo, koristili smo laičke sažetke, tj. PLS-ove iz jednog izvora – Cochrane Library – a te su sažetke pisali različiti autori. Međutim, sažeci iz Cochrane Library imaju ujednačen format predstavljanja zdravstvenih informacija i specifične smjernice za pisanje PLS-a, što ih čini usporedivima. Drugo, analizirali smo isključivo PLS-ove na engleskom jeziku jer je to bio jedini zajednički jezik sažetaka u sustavnim pregledima iz onkologije. Fokus na engleski jezik osigurao je dosljednost u jezičnoj analizi, izbjegavajući nedosljednosti koje bi nastale u procesu prevođenja. Treće, značajno ograničenje ove studije je što se od ChatGPT-a tražilo da generira PLS-ove na temelju znanstvenih sažetaka, a ne iz cjelovitih tekstova Cochraneovih sustavnih pregleda. Uobičajeno se PLS-ovi izrađuju na temelju cjelokupnog sadržaja pregleda, što pruža dublje razumijevanje nalaza, metodologije i konteksta studije. Oslanjanje isključivo na znanstvene sažetke može dovesti do površnih i nedovoljno detaljnih PLS-ova. Četvrto, ChatGPT-4 razvijen je korištenjem raznovrsnog skupa javno dostupnih podataka iz više domena, a u vrijeme provođenja naše studije, njegovi podaci bili su ograničeni na informacije dostupne do 2023. godine. OpenAI ne navodi točan sadržaj uključenih medicinskih podataka niti precizni vremenski okvir korištenih podataka. Koristili smo ChatGPT-4, verziju s pretplatom (za razliku od prethodne verzije ChatGPT-3.5), koja je bila najnaprednija dostupna verzija ChatGPT-a tijekom trajanja studije, čime smo osigurali pristup njegovim najnovijim mogućnostima, ali je dostupna samo korisnicima koji su spremni platiti mjesečnu pretplatu. Za drugi upit u našoj studiji oslanjali smo se na odgovore koje je ChatGPT dao u prvom upitu. Kod sva tri upita nismo tražili dodatno preformuliranje teksta, već smo analizirali isključivo prvi generirani odgovor. Različiti AI modeli razlikuju se po načinu treniranja, algoritmima i mogućnostima, stoga se korištenje i rezultati jednog modela ne mogu univerzalno primijeniti na druge. Nismo koristili druge AI alate poput Microsoft Bing AI, Bard, Jasper ili ChatSonic, koji su mogli dati drugačije rezultate. Microsoft Bing AI može obraditi do 2000 znakova, što nije prikladno za sažetke. Bard nije bio dostupan u Hrvatskoj u vrijeme provođenja studije, a Jasper i ChatSonic nudili su samo pretplatničke

verzije za koje nismo imali dovoljno sredstava. Također, nismo koristili ni specijalizirane AI alate razvijene isključivo za izradu PLS-ova, poput Sorcero ili pristupa Putnam Associates u korištenju generativne AI za PLS. Odluku da ne koristimo te specijalizirane alate temeljili smo na fokusu na opće AI modele koji su dostupni širem krugu korisnika, kao i na namjeri da procijenimo učinkovitost široko dostupne, nespecijalizirane AI u generiranju PLS-ova. Peto, treba imati na umu da, iako je naša studija procjenjivala jezične karakteristike poput uvjerljivosti, analitičkog tona, autentičnosti i emocionalnog tona, jednako je važno uzeti u obzir i kulturološke i emocionalne osjetljivosti ciljane publike. AI modeli poput ChatGPT-4 trenirani su na opsežnim skupovima podataka koji možda ne obuhvaćaju u potpunosti nijanse različitih kulturnih pozadina. Kao rezultat toga, generirani PLS-ovi mogu nedostajati kulturološke relevantnosti ili osjetljivosti potrebne za učinkovitu komunikaciju sa svim dijelovima laičke javnosti. Buduća istraživanja trebala bi uključivati PLS-ove iz više izvora kako bi se procijenila opća primjenjivost PLS-ova generiranih AI-em u različitim formatima i stilovima pisanja, te istražiti potencijal AI alata u generiranju PLS-ova na jezicima osim engleskog, kako bi se podržali Cochraneovi napori u pružanju zdravstvenih informacija na 20 različitih jezika.

5.3 Zaključci i smjerovi za budućnost

Osim tehničkih i kliničkih aspekata, važno je postaviti pitanje: Što su PLS-ovi i kome su namijenjeni? PLS-ovi predstavljaju jeftin i jednostavan način širenja rezultata istraživanja. Prednost je što ih pišu sami istraživači – informacije dolaze "iz prve ruke", no istodobno, mnogi istraživači nemaju edukaciju iz komunikacije prema laicima. Jezik znanosti često je ekskluzivan – kompleksan, zatvoren i teško razumljiv onima izvan struke. Time se stvara nejednakost u pristupu znanju i smanjuje angažman javnosti.

Pisanje PLS-ova podrazumijeva obraćanje određenoj publici, no publika nije homogena. "Opća javnost" obuhvaća različite razine interesa, znanja i povjerenja u znanost. Stil komunikacije uvjetuje tko će poruku uopće moći razumjeti. Ako je sadržaj prepun žargona, jasno je da nije namijenjen svima. U tom kontekstu, PLS-ovi često ne ispunjavaju očekivanja stvarnog javnog uključivanja.

Istraživanja učinkovitosti PLS-ova i dalje su u razvoju. Dosadašnji pokušaji poboljšanja, primjerice uz pomoć smjernica ili profesionalnih urednika, donijeli su mješovite rezultate.

Za zaključak rezultata ovih studija, razvidno je da je potrebno poboljšati dostupnost i kvalitetu laičkih sažetaka. Nekoliko je područja na kojima se može/treba djelovati: 1)

znanstvenici bi trebali surađivati s novinarima i komunikacijskim stručnjacima radi izrade razumljivijih sažetaka; 2) podrška umjetne inteligencije: AI alati poput ChatGPT-a mogu poboljšati čitkost, ali je potreban ljudski nadzor kako bi se osigurala točnost; 3) inicijative za zdravstvenu pismenost: potrebno je povećati zdravstvenu pismenost opće populacije, javnozdravstvenim akcijama potaknuti ljudi na veći angažman o zdravlju, reklamama ili uz pomoć javnih osoba upozoriti na rizična zdravstvena ponašanja; 4) suradnja s udrugama pacijenata; 5) višekanalna komunikacija: informacije bi trebale biti distribuirane putem različitih medija, uz regulaciju širenja dezinformacija; na ovaj način se bi se pokrilo više dobnih skupina, kao i smanjila pristranost zbog eventualnih kulturnih razlika; 6) trebalo bi procijeniti stvarno razumijevanje AI-generiranih PLS-ova među pacijentima. Zaključno, PLS-ovi imaju potencijal biti ono što su u suštini i trebali biti-most između znanstvene zajednice i opće populacije, no to zahtijeva promišljen, uključiv i višedimenzionalan pristup.

6. ZAKLJUČCI

1. Laički sažeci Cochrane sustavnih pregleda koji su obrađivali onkološke intervencije pokazali su nisku razinu čitkosti.
2. U usporedbi sa znanstvenim sažecima, laički sažeci su bili lakši za čitanje, imali manji broj riječi, sadržavali manje analitičkog tona, ali su bili emocionalniji i djelovali autentičnije.
3. Među različitim Cochrane onkološkim skupinama nije pronađena statistički značajna razlika u čitkosti laičkih sažetaka.
4. Kategorija zaključka “Nejasno”, odnosno autori ne bi konstatirali jasan zaključak pitanja sustavnog pregleda ili bi sugerirali da je potrebno više istraživanja, je bila najčešće donesena kategorija zaključka u četiri Cochrane skupine koje se bave područjem onkologije.
5. Laički sažeci kreirani pomoću umjetne inteligencije na temelju znanstvenih sažetaka su imali niži indeks čitkosti u usporedbi sa originalnim laičkim sažecima, napisanim od strane autora sustavnog pregleda.
6. Laički sažeci kreirani pomoću umjetne inteligencije su se podudarali u kategorijama zaključka u odnosu na originalne laičke sažetke.
7. ChatGPT može biti vrijedan alat u pojednostavljivanju PLS-ova kao medicinski povezanih formata za širu publiku. Potrebno je više istraživanja, uključujući i o mehanizmima nadzora kako bi se osiguralo da su informacije točne, pouzdane i kulturno relevantne za različite publike.

7. LITERATURA

1. Tenny S, Varacallo MA. Evidence-Based Medicine. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2025.
2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996 Jan 13;312(7023):71–2.
3. Sackett DL. Clinical epidemiology. *Am J Epidemiol*. 1969 Feb;89(2):125–8.
4. Deborah J. Cook GG. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice, 3rd ed.
5. Ratnani I, Fatima S, Abid MM, Surani Z, Surani S. Evidence-Based Medicine: History, Review, Criticisms, and Pitfalls. *Cureus*. 2023 Feb;15(2):e35266.
6. Sur RL, Dahm P. History of evidence-based medicine. *Indian J Urol IJU J Urol Soc India*. 2011 Oct;27(4):487–9.
7. Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *J Med Libr Assoc JMLA*. 2018 Oct;106(4):420–31.
8. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 1989 Feb;95(2 Suppl):2S–4S.
9. Wallace SS, Barak G, Truong G, Parker MW. Hierarchy of Evidence Within the Medical Literature. *Hosp Pediatr*. 2022 Aug 1;12(8):745–50.
10. Caldwell PH, Bennett T. Easy guide to conducting a systematic review. *J Paediatr Child Health*. 2020 Jun;56(6):853–6.
11. Manchikanti L, Hirsch JA, Smith HS. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 2: Randomized controlled trials. *Pain Physician*. 2008 Dec;11(6):717–73.
12. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg*. 2011 Jul;128(1):305–10.
13. Gopalakrishnan S, Ganeshkumar P. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. *J Fam Med Prim Care*. 2013 Jan;2(1):9–14.
14. Siddaway AP, Wood AM, Hedges LV. How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. *Annu Rev Psychol*. 2019 Jan 4;70:747–70.
15. Phillips V, Barker E. Systematic reviews: Structure, form and content. *J Perioper Pract*. 2021 Sep;31(9):349–53.
16. Trespidi C, Barbui C, Cipriani A. Why it is important to include unpublished data in systematic reviews. *Epidemiol Psychiatr Sci*. 2011 Jun;20(2):133–5.

17. White A, Schmidt K. Systematic literature reviews. *Complement Ther Med*. 2005 Mar;13(1):54–60.
18. Brignardello-Petersen R, Santesso N, Guyatt GH. Systematic reviews of the literature: an introduction to current methods. *Am J Epidemiol*. 2025 Feb 5;194(2):536–42.
19. Al-Khabori M, Rasool W. Introduction to Systematic Reviews and Meta-analyses of Therapeutic Studies. *Oman Med J*. 2022 Sep 30;37(5):e428.
20. Ahn E, Kang H. Introduction to systematic review and meta-analysis. *Korean J Anesthesiol*. 2018 Apr;71(2):103–12.
21. Henderson LK, Craig JC, Willis NS, Tovey D, Webster AC. How to write a Cochrane systematic review. *Nephrol Carlton Vic*. 2010 Sep;15(6):617–24.
22. Scholten RJPM, Clarke M, Hetherington J. The Cochrane Collaboration. *Eur J Clin Nutr*. 2005 Aug;59 Suppl 1:S147-9; discussion S195-196.
23. Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of The Cochrane Database of Systematic Reviews. *Int J Technol Assess Health Care*. 2009 Jul;25 Suppl 1:182–95.
24. Cipriani A, Furukawa TA, Barbui C. What is a Cochrane review? *Epidemiol Psychiatr Sci*. 2011 Sep;20(3):231–3.
25. Stavrou A, Chalkou D, Dimitrakakis G. Archibald Cochrane (1909-1988): the father of evidence-based medicine. *Interact Cardiovasc Thorac Surg*. 2014 Jan;18(1):121–4.
26. Cochrane Switzerland. History. [Internet]. Cochrane Switzerland.; [cited 2025 Apr 8]. Available from: <https://swiss.cochrane.org/our-history-0>
27. Levin A. The Cochrane Collaboration. *Ann Intern Med*. 2001 Aug 21;135(4):309–12.
28. Jadad AR, Haynes RB. The Cochrane Collaboration--advances and challenges in improving evidence-based decision making. *Med Decis Mak Int J Soc Med Decis Mak*. 1998 Mar;18(1):2–9; discussion 16-18.
29. Allen C, Richmond K. The Cochrane Collaboration: International activity within Cochrane Review Groups in the first decade of the twenty-first century. *J Evid-Based Med*. 2011;4(1):2–7.
30. Cochrane.org. Cochrane Review Groups [Internet]. [cited 2025 Apr 8]. Available from: <https://www.cochranelibrary.com/about/cochrane-review-groups>
31. Cochrane Iran. Networks of Cochrane Review Groups [Internet]. [cited 2025 Apr 8]. Available from: <https://iran.cochrane.org/news/networks-cochrane-review-groups>
32. Cochrane.org. Review Groups [Internet]. [cited 2025 Apr 8]. Available from: <https://www.cochrane.org/about-us/our-global-community/review-groups>
33. Shah HM, Chung KC. Archie Cochrane and his vision for evidence-based medicine. *Plast Reconstr Surg*. 2009 Sep;124(3):982–8.

34. Cochrane.org. List of Cochrane Review Groups. [cited 2025 Apr 8]; Available from: <https://www.cochrane.org/about-us/our-global-community/review-groups>
35. Cochrane.org. Cochrane Thematic Groups [Internet]. [cited 2025 Apr 8]. Available from: <https://www.cochrane.org/about-us/our-global-community/thematic-groups>
36. Kumar A. Writing an abstract: Revealing the essence with eloquence. *J Indian Soc Periodontol*. 2022 Feb;26(1):1–2.
37. Jelicic Kadic A, Fidahic M, Vujcic M, Saric F, Propadalo I, Marelja I, et al. Cochrane plain language summaries are highly heterogeneous with low adherence to the standards. *BMC Med Res Methodol*. 2016 May 23;16:61.
38. Maurer M, Siegel JE, Firminger KB, Lowers J, Dutta T, Chang JS. Lessons Learned from Developing Plain Language Summaries of Research Studies. *Health Lit Res Pract*. 2021 Apr;5(2):e155–61.
39. Behmen D, Marušić A, Puljak L. Capacity building for knowledge translation: A survey about the characteristics and motivation of volunteer translators of Cochrane plain language summaries. *J Evid-Based Med*. 2019 May;12(2):147–54.
40. Cochrane.org. Knowledge translation in different languages. [Internet]. [cited 2025 Apr 9]. Available from: <https://community.cochrane.org/review-development/knowledge-translation/knowledge-translation-different-languages>
41. Rowlands G. Health literacy. *Hum Vaccines Immunother*. 2014;10(7):2130–5.
42. Barańska A, Kłak A. Recent Trends in Health Literacy Research, Health Status of the Population and Disease Prevention: An Editorial. *Int J Environ Res Public Health*. 2022 Jul 10;19(14).
43. Chehuen Neto JA, Costa LA, Estevanin GM, Bignoto TC, Vieira CIR, Pinto FAR, et al. Functional Health Literacy in chronic cardiovascular patients. *Cienc Saude Coletiva*. 2019 Mar;24(3):1121–32.
44. Campbell JK, Erickson JM. Interactive Health Literacy and Symptom Self-management in Patients With Lung Cancer: A Critical Realist Analysis. *Cancer Nurs*. 2024 Oct 1;47(5):397–407.
45. Abel T, Benkert R. Critical health literacy: reflection and action for health. *Health Promot Int*. 2022 Aug 1;37(4).
46. Liu C, Wang D, Liu C, Jiang J, Wang X, Chen H, et al. What is the meaning of health literacy? A systematic review and qualitative synthesis. *Fam Med Community Health*. 2020 May;8(2).
47. Noblin AM, Zraick RI, Miller AN, Schmidt-Owens M, Deichen M, Tran K, et al. Readability and Suitability of Information Presented on a University Health Center Website. *Perspect Health Inf Manag*. 2022 Fall;19(4):1f.
48. Nash E, Bickerstaff M, Chetwynd AJ, Hawcutt DB, Oni L. The readability of parent information leaflets in paediatric studies. *Pediatr Res*. 2023 Sep;94(3):1166–71.

49. Wang LW, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. *Res Soc Adm Pharm RSAP*. 2013 Oct;9(5):503–16.
50. Mc Laughlin GH. SMOG grading-a new readability formula. *J Read*. 1969;12(8):639–46.
51. Brangan S. Development of SMOG-Cro readability formula for healthcare communication and patient education. *Coll Antropol*. 2015 Mar;39(1):11–20.
52. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*. 2011 Jul 19;155(2):97–107.
53. Davis TC, Wolf MS, Bass PF 3rd, Thompson JA, Tilson HH, Neuberger M, et al. Literacy and misunderstanding prescription drug labels. *Ann Intern Med*. 2006 Dec 19;145(12):887–94.
54. Aljassim N, Ostini R. Health literacy in rural and urban populations: A systematic review. *Patient Educ Couns*. 2020 Oct;103(10):2142–54.
55. Weng YP, Chen SC, Hong RM, Huang W, Wang YH, Yang YH, et al. Psychological Journey of Delayed Medical Consultation in Patients With Breast Cancer Who Have Giant Tumors. *Cancer Nurs*. 2025 Apr 7;
56. Morris NS, Field TS, Wagner JL, Cutrona SL, Roblin DW, Gaglio B, et al. The association between health literacy and cancer-related attitudes, behaviors, and knowledge. *J Health Commun*. 2013;18 Suppl 1(Suppl 1):223–41.
57. Altin SV, Stock S. The impact of health literacy, patient-centered communication and shared decision-making on patients' satisfaction with care received in German primary care practices. *BMC Health Serv Res*. 2016 Aug 30;16(1):450.
58. Pleasant A, Cabe J, Patel K, Cosenza J, Carmona R. Health Literacy Research and Practice: A Needed Paradigm Shift. *Health Commun*. 2015;30(12):1176–80.
59. Pleasant A. Advancing health literacy measurement: a pathway to better health and health system performance. *J Health Commun*. 2014 Dec;19(12):1481–96.
60. Haun JN, Patel NR, French DD, Campbell RR, Bradham DD, Lapcevic WA. Association between health literacy and medical care costs in an integrated healthcare system: a regional population based study. *BMC Health Serv Res*. 2015 Jun 27;15(1):249.
61. Greene JC, Haun JN, French DD, Chambers SL, Roswell RH. Reduced Hospitalizations, Emergency Room Visits, and Costs Associated with a Web-Based Health Literacy, Aligned-Incentive Intervention: Mixed Methods Study. *J Med Internet Res*. 2019 Oct 17;21(10):e14772.
62. Health Literacy: Improving Health, Health Systems, and Health Policy Around the World: Workshop Summary. Washington (DC); 2013.

63. Schiavo R, Wye G, Manoncourt E. COVID-19 and health inequities: the case for embracing complexity and investing in equity- and community-driven approaches to communication. *J Commun Healthc*. 2022 Jan;15:1–6.
64. Sentell T, Braun KL. Low health literacy, limited English proficiency, and health status in Asians, Latinos, and other racial/ethnic groups in California. *J Health Commun*. 2012;17 Suppl 3(Suppl 3):82–99.
65. Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV. Shame and health literacy: the unspoken connection. *Patient Educ Couns*. 1996 Jan;27(1):33–9.
66. Wolf MS, Williams MV, Parker RM, Parikh NS, Nowlan AW, Baker DW. Patients' shame and attitudes toward discussing the results of literacy screening. *J Health Commun*. 2007 Dec;12(8):721–32.
67. Baker DW, Gazmararian JA, Sudano J, Patterson M. The association between age and health literacy among elderly persons. *J Gerontol B Psychol Sci Soc Sci*. 2000 Nov;55(6):S368-374.
68. McCray AT. Promoting health literacy. *J Am Med Inform Assoc JAMIA*. 2005 Apr;12(2):152–63.
69. Ho EY, Tran H, Chesla CA. Assessing the cultural in culturally sensitive printed patient-education materials for Chinese Americans with type 2 diabetes. *Health Commun*. 2015;30(1):39–49.
70. Baur C, Harris L, Squire E. The U.S. National Action Plan to Improve Health Literacy: A Model for Positive Organizational Change. *Stud Health Technol Inform*. 2017;240:186–202.
71. Schaeffer D, Griesse L, Hurrelmann K. [Establishing and implementing the National Action Plan Health Literacy in Germany]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2025 Mar;68(3):231–9.
72. Goerling U, Faller H, Hornemann B, Hönig K, Bergelt C, Maatouk I, et al. Information needs in cancer patients across the disease trajectory. A prospective study. *Patient Educ Couns*. 2020 Jan;103(1):120–6.
73. McCaffery KJ, Smith SK, Wolf M. The challenge of shared decision making among patients with lower literacy: a framework for research and development. *Med Decis Mak Int J Soc Med Decis Mak*. 2010 Feb;30(1):35–44.
74. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021 May;71(3):209–49.
75. Kinnane NA, Milne DJ. The role of the Internet in supporting and informing carers of people with cancer: a literature review. *Support Care Cancer Off J Multinatl Assoc Support Care Cancer*. 2010 Sep;18(9):1123–36.
76. Kaidar-Person O, Antunes M, Cardoso JS, Ciani O, Cruz H, Di Micco R, et al. Evaluating the ability of an artificial-intelligence cloud-based platform designed to

- provide information prior to locoregional therapy for breast cancer in improving patient's satisfaction with therapy: The CINDERELLA trial. *PloS One*. 2023;18(8):e0289365.
77. Martinez-Gutierrez J, Soto MG, Rioseco A, Bienzobas C, Fowler M, Ulloa G, et al. Are we ready? assessing effectiveness and implementation of cancer control strategies in primary care: a comprehensive review of systematic reviews. *Fam Pract*. 2025 Feb 7;42(2).
 78. Oldach BR, Katz ML. Health literacy and cancer screening: a systematic review. *Patient Educ Couns*. 2014 Feb;94(2):149–57.
 79. Hasannejadasl H, Roumen C, Smit Y, Dekker A, Fijten R. Health Literacy and eHealth: Challenges and Strategies. *JCO Clin Cancer Inform*. 2022 Sep;6:e2200005.
 80. Halbach SM, Ernstmann N, Kowalski C, Pfaff H, Pförtner TK, Wesselmann S, et al. Unmet information needs and limited health literacy in newly diagnosed breast cancer patients over the course of cancer treatment. *Patient Educ Couns*. 2016 Sep;99(9):1511–8.
 81. Bandini L, Schatz A, Hood V, Clark N, Hall MJ, Carlson RW. NCCN Policy Summit: Reducing the Cancer Burden Through Prevention and Early Detection. *J Natl Compr Cancer Netw JNCCN*. 2023 Sep;21(9):910-914.e1.
 82. Hyatt A, Shelly A, Cox R, Humphries E, Lock G, Varlow M. How can we improve information for people affected by cancer? A national survey exploring gaps in current information provision, and challenges with accessing cancer information online. *Patient Educ Couns*. 2022 Aug 1;105(8):2763–70.
 83. Bujnowska-Fedak MM, Waligóra J, Mastalerz-Migas A. The Internet as a Source of Health Information and Services. *Adv Exp Med Biol*. 2019;1211:1–16.
 84. Han CJ, Lee YJ, Demiris G. Interventions Using Social Media for Cancer Prevention and Management: A Systematic Review. *Cancer Nurs*. 2018 Dec;41(6):E19–31.
 85. Plackett R, Kaushal A, Kassianos AP, Cross A, Lewins D, Sheringham J, et al. Use of Social Media to Promote Cancer Screening and Early Diagnosis: Scoping Review. *J Med Internet Res*. 2020 Nov 9;22(11):e21582.
 86. Ghenai A, Mejova Y. Fake Cures: User-centric Modeling of Health Misinformation in Social Media. 2018.
 87. Loeb S, Langford AT, Bragg MA, Sherman R, Chan JM. Cancer misinformation on social media. *CA Cancer J Clin*. 2024 Oct;74(5):453–64.
 88. Rosenberg SA, Francis DM, Hullet CR, Morris ZS, Brower JV, Anderson BM, et al. Online patient information from radiation oncology departments is too complex for the general population. *Pract Radiat Oncol*. 2017 Feb;7(1):57–62.
 89. Friedman DB, Hoffman-Goetz L, Arocha JF. Readability of cancer information on the internet. *J Cancer Educ Off J Am Assoc Cancer Educ*. 2004 Summer;19(2):117–22.

90. Bitkina OV, Park J, Kim HK. Application of artificial intelligence in medical technologies: A systematic review of main trends. *Digit Health*. 2023 Dec;9:20552076231189331.
91. Bekbolatova M, Mayer J, Ong CW, Toma M. Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. *Healthc Basel Switz*. 2024 Jan 5;12(2).
92. Al Kuwaiti A, Nazer K, Al-Reedy A, Al-Shehri S, Al-Muhanna A, Subbarayalu AV, et al. A Review of the Role of Artificial Intelligence in Healthcare. *J Pers Med*. 2023 Jun 5;13(6).
93. Razavian N, Knoll F, Geras KJ. Artificial Intelligence Explained for Nonexperts. *Semin Musculoskelet Radiol*. 2020 Feb;24(1):3–11.
94. Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. *J Fam Med Prim Care*. 2019 Jul;8(7):2328–31.
95. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthc J*. 2021 Jul;8(2):e188–94.
96. Duecker K, Idiart M, van Gerven M, Jensen O. Oscillations in an artificial neural network convert competing inputs into a temporal code. *PLoS Comput Biol*. 2024 Sep;20(9):e1012429.
97. Loke J, Seijdel N, Snoek L, Sörensen LKA, van de Klundert R, van der Meer M, et al. Human Visual Cortex and Deep Convolutional Neural Network Care Deeply about Object Background. *J Cogn Neurosci*. 2024 Mar 1;36(3):551–66.
98. Carlos Ferreira J, Elvas LB, Correia R, Mascarenhas M. Enhancing EHR Interoperability and Security through Distributed Ledger Technology: A Review. *Healthc Basel Switz*. 2024 Oct 2;12(19).
99. Zoccali C, Mallamaci F. Exploring the metaverse: opportunities for nephrology in patient care and education. *Nephrol Dial Transplant Off Publ Eur Dial Transpl Assoc - Eur Ren Assoc*. 2025 Apr 28;40(5):865–73.
100. Kleib M, Darko EM, Akingbade O, Kennedy M, Majekodunmi P, Nickel E, et al. Current trends and future implications in the utilization of ChatGPT in nursing: A rapid review. *Int J Nurs Stud Adv*. 2024 Dec;7:100252.
101. Xu T, Weng H, Liu F, Yang L, Luo Y, Ding Z, et al. Current Status of ChatGPT Use in Medical Education: Potentials, Challenges, and Strategies. *J Med Internet Res*. 2024 Aug 28;26:e57896.
102. Jordan DM, Vy HMT, Do R. A deep learning transformer model predicts high rates of undiagnosed rare disease in large electronic health systems. *medRxiv : the preprint server for health sciences*. United States; 2023. p. 2023.12.21.23300393.
103. Yang Z, Mitra A, Liu W, Berlowitz D, Yu H. TransformEHR: transformer-based encoder-decoder generative model to enhance prediction of disease outcomes using electronic health records. *Nat Commun*. 2023 Nov 29;14(1):7857.

104. Coiera E. The Price of Artificial Intelligence. *Yearb Med Inform.* 2019 Aug;28(1):14–5.
105. Abonamah AA, Tariq MU, Shilbayeh S. On the Commoditization of Artificial Intelligence. *Front Psychol.* 2021;12:696346.
106. IBM. Understanding the different types of artificial intelligence [Internet]. [cited 2025 Apr 17]. Available from: <https://www.ibm.com/think/topics/artificial-intelligence-types>
107. Raza MA, Aziz S, Noreen M, Saeed A, Anjum I, Ahmed M, et al. Artificial Intelligence (AI) in Pharmacy: An Overview of Innovations. *Innov Pharm.* 2022;13(2).
108. Cuzzolin F, Morelli A, Cîrstea B, Sahakian BJ. Knowing me, knowing you: theory of mind in AI. *Psychol Med.* 2020 May;50(7):1057–61.
109. Bindra S, Jain R. Artificial intelligence in medical science: a review. *Ir J Med Sci.* 2024 Jun;193(3):1419–29.
110. Bhuyan SS, Sateesh V, Mukul N, Galvankar A, Mahmood A, Nauman M, et al. Generative Artificial Intelligence Use in Healthcare: Opportunities for Clinical Excellence and Administrative Efficiency. *J Med Syst.* 2025 Jan 16;49(1):10.
111. Virk A, Alasmari S, Patel D, Allison K. Digital Health Policy and Cybersecurity Regulations Regarding Artificial Intelligence (AI) Implementation in Healthcare. *Cureus.* 2025 Mar;17(3):e80676.
112. Weiner EB, Dankwa-Mullan I, Nelson WA, Hassanpour S. Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLOS Digit Health.* 2025 Apr;4(4):e0000810.
113. Lotter W, Hassett MJ, Schultz N, Kehl KL, Van Allen EM, Cerami E. Artificial Intelligence in Oncology: Current Landscape, Challenges, and Future Directions. *Cancer Discov.* 2024 May 1;14(5):711–26.
114. Kolla L, Parikh RB. Uses and limitations of artificial intelligence for oncology. *Cancer.* 2024 Jun 15;130(12):2101–7.
115. Mahoro E, Akhloufi MA. Applying Deep Learning for Breast Cancer Detection in Radiology. *Curr Oncol Tor Ont.* 2022 Nov 16;29(11):8767–93.
116. Prelaj A, Miskovic V, Zanitti M, Trovo F, Genova C, Viscardi G, et al. Artificial intelligence for predictive biomarker discovery in immuno-oncology: a systematic review. *Ann Oncol Off J Eur Soc Med Oncol.* 2024 Jan;35(1):29–65.
117. Kamya P, Ozerov IV, Pun FW, Tretina K, Fokina T, Chen S, et al. PandaOmics: An AI-Driven Platform for Therapeutic Target and Biomarker Discovery. *J Chem Inf Model.* 2024 May 27;64(10):3961–9.
118. Tsai ML, Chen KF, Chen PC. Harnessing Electronic Health Records and Artificial Intelligence for Enhanced Cardiovascular Risk Prediction: A Comprehensive Review. *J Am Heart Assoc.* 2025 Mar 18;14(6):e036946.

119. Lavoie-Gagne O, Woo JJ, Williams RJ 3rd, Nwachukwu BU, Kunze KN, Ramkumar PN. Artificial Intelligence as a Tool to Mitigate Administrative Burden, Optimize Billing, Reduce Insurance- and Credentialing-Related Expenses, and Improve Quality Assurance Within Health Care Systems. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2025 Mar 20;S0749-8063(25)00216-6.
120. Krishnamurthy R, Mummudi N, Goda JS, Chopra S, Heijmen B, Swamidas J. Using Artificial Intelligence for Optimization of the Processes and Resource Utilization in Radiotherapy. *JCO Glob Oncol*. 2022 Nov;8:e2100393.
121. Verlingue L, Boyer C, Olgiati L, Brutti Mairesse C, Morel D, Blay JY. Artificial intelligence in oncology: ensuring safe and effective integration of language models in clinical practice. *Lancet Reg Health Eur*. 2024 Nov;46:101064.
122. Eltorai AEM, Ghanian S, Adams CAJ, Born CT, Daniels AH. Readability of patient education materials on the american association for surgery of trauma website. *Arch Trauma Res*. 2014 Jun;3(2):e18161.
123. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language. use: our words, our selves. *Annu Rev Psychol*. 2003;54:547–77.
124. Smith-Keiling BL, Hyun HIF. Applying a Computer-Assisted Tool for Semantic Analysis of Writing: Uses for STEM and ELL. *J Microbiol Biol Educ*. 2019;20(1).
125. Pennebaker JW, Chung CK, Frazee J, Lavergne GM, Beaver DI. When small words foretell academic success: the case of college admissions essays. *PloS One*. 2014;9(12):e115844.
126. Kacewicz E, Pennebaker JW, Davis M, Jeon M, Graesser AC. Pronoun Use Reflects Standings in Social Hierarchies. *J Lang Soc Psychol*. 2014;33(2):125–43.
127. Newman ML, Pennebaker JW, Berry DS, Richards JM. Lying words: predicting deception from linguistic styles. *Pers Soc Psychol Bull*. 2003 May;29(5):665–75.
128. Cohn MA, Mehl MR, Pennebaker JW. Linguistic markers of psychological change surrounding September 11, 2001. *Psychol Sci*. 2004 Oct;15(10):687–93.
129. Vaughn LA. Contents of Hopes and Duties: A Linguistic Analysis. *Front Psychol*. 2018;9:757.
130. Radman M, Babic A, Runjic E, Jelcic Kadic A, Jeric M, Moja L, et al. Revisiting established medicines: An overview of systematic reviews about ibuprofen and paracetamol for treating pain in children. *Eur J Pain Lond Engl*. 2019 Jul;23(6):1071–82.
131. Bralić N, Mijatović A, Marušić A, Buljan I. Conclusiveness, readability and textual characteristics of plain language summaries from medical and non-medical organizations: a cross-sectional study. *Sci Rep*. 2024 Mar 12;14(1):6016.
132. Šuto J, Marušić A, Buljan I. Linguistic analysis of plain language summaries and corresponding scientific summaries of Cochrane systematic reviews about oncology interventions. *Cancer Med*. 2023 May;12(9):10950–60.

133. Šuto Pavičić, Jelena, Buljan I, Marušić A. Using ChatGPT to improve the presentation of plain language summaries of cochrane systematic reviews about oncology interventions: cross-sectional study [Data set]. [Internet]. Open Science Framework; [cited 2025 Apr 14]. Available from: https://osf.io/92tnv/?view_only=36d0b17c0ed547929fcd8ce93e8bea30
134. Šuto J. Jezična analiza sažetaka Cochraneovih sustavnih pregleda onkoloških intervencija [Diplomski rad] [Internet]. [Split]: Sveučilište u Splitu, Medicinski fakultet; 2020 [cited 2025 Jun 5]. Available from: <https://urn.nsk.hr/urn:nbn:hr:171:390584>
135. Wang LW, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. *Res Soc Adm Pharm RSAP*. 2013;9(5):503–16.
136. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*. 2003.
137. Weis B. Health literacy: a manual for clinicians. American Medical Association Foundation and American Medical Association. 2003;
138. Kumar NB, Bahl S, Dhillon J, Poch M, Manley B, Li R, et al. Challenges to Recruiting Men on Active Surveillance for Prostate Cancer in Clinical Chemoprevention Trials. *Cancers*. 2023 Feb 16;15(4).
139. Dudipala H, Burns L, Jani CT, Radwan A, Al Omari O, Patel M, et al. Disparities in Lung Cancer Clinical Trial Discussion and Enrollment at a Safety Net Hospital. *Community Health Equity Res Policy*. 2023 Dec 12;2752535X231221394.
140. Unger JM, Shulman LN, Facktor MA, Nelson H, Fleury ME. National Estimates of the Participation of Patients With Cancer in Clinical Research Studies Based on Commission on Cancer Accreditation Data. *J Clin Oncol Off J Am Soc Clin Oncol*. 2024 Jun 20;42(18):2139–48.
141. Izarn F, Henry J, Besle S, Ray-Coquard I, Blay JY, Allignet B. Globalization of clinical trials in oncology: a worldwide quantitative analysis. *ESMO Open* [Internet]. 2025 Jan 1 [cited 2025 Apr 14];10(1). Available from: <https://doi.org/10.1016/j.esmoop.2024.104086>
142. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024 Jun;74(3):229–63.
143. Halaseh SA, Al-Karadsheh A, Mukherji D, Alhjahaja A, Farkouh A, Al-Ibraheem A, et al. Prostate cancer clinical trials in low- and middle-income countries. *Ecancermedalscience*. 2023;17:1629.
144. Wells JC, Sharma S, Del Paggio JC, Hopman WM, Gyawali B, Mukherji D, et al. An Analysis of Contemporary Oncology Randomized Clinical Trials From Low/Middle-Income vs High-Income Countries. *JAMA Oncol*. 2021 Mar 1;7(3):379–85.

145. Bizuayehu HM, Ahmed KY, Kibret GD, Dadi AF, Belachew SA, Bagade T, et al. Global Disparities of Cancer and Its Projected Burden in 2050. *JAMA Netw Open*. 2024 Nov 4;7(11):e2443198.
146. Desai A, Sirohi B, Mathew A. Clinical Trial Access in Low- and Middle-Income Countries: A Case Study on India. *Cancer Invest*. 2021 Oct;39(9):685–9.
147. Organization WH. Technical report: pricing of cancer medicines and its impacts: a comprehensive technical report for the World Health Assembly Resolution 70.12: operative paragraph 2.9 on pricing approaches and their impacts on availability and affordability of medicines for the prevention and treatment of cancer. World Health Organization; 2018.
148. Rubagumya F, Hopman WM, Gyawali B, Mukherji D, Hammad N, Pramesh CS, et al. Participation of Lower and Upper Middle-Income Countries in Clinical Trials Led by High-Income Countries. *JAMA Netw Open*. 2022 Aug 1;5(8):e2227252.
149. Cui M, Cheng H, Yuan L, Zhang M, Chen J, Jia J, et al. Burden of nasopharyngeal carcinoma in Asia from 1990 to 2021. *J Dent*. 2025 Mar;154:105583.
150. Dzinamarira T, Moyo E, Dzobo M, Mbunge E, Murewanhema G. Cervical cancer in sub-Saharan Africa: an urgent call for improving accessibility and use of preventive services. *Int J Gynecol Cancer Off J Int Gynecol Cancer Soc*. 2023 Apr 3;33(4):592–7.
151. Mak LY, Liu K, Chirapongsathorn S, Yew KC, Tamaki N, Rajaram RB, et al. Liver diseases and hepatocellular carcinoma in the Asia-Pacific region: burden, trends, challenges and future directions. *Nat Rev Gastroenterol Hepatol*. 2024 Dec;21(12):834–51.
152. Jafri W, Kamran M. Hepatocellular Carcinoma in Asia: A Challenging Situation. *Euroasian J Hepato-Gastroenterol*. 2019 Jun;9(1):27–33.
153. Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *J Clin Epidemiol*. 2009 Oct;62(10):1021–30.
154. Wong B. Visual representation of scientific information. *Sci Signal*. 2011 Feb 15;4(160):pt1.
155. Traboco L, Pandian H, Nikiphorou E, Gupta L. Designing Infographics: Visual Representations for Enhancing Education, Communication, and Scientific Research. *J Korean Med Sci*. 2022 Jul 11;37(27):e214.
156. Yang X, Xiao Y, Liu D, Shi H, Deng H, Huang J, et al. Enhancing Physician-Patient Communication in Oncology Using GPT-4 Through Simplified Radiology Reports: Multicenter Quantitative Study. *J Med Internet Res*. 2025 Apr 17;27:e63786.
157. Maroncelli R, Rizzo V, Pasculli M, Cicciarelli F, Macera M, Galati F, et al. Probing clarity: AI-generated simplified breast imaging reports for enhanced patient comprehension powered by ChatGPT-4o. *Eur Radiol Exp*. 2024 Oct 30;8(1):124.

158. Kim M, Kim E, Lee H, Piao M, Rosen B, Allison JJ, et al. A Culturally Tailored Artificial Intelligence Chatbot (K-Bot) to Promote Human Papillomavirus Vaccination Among Korean Americans: Development and Usability Study. *AsianPacific Isl Nurs J*. 2025 Apr 7;9:e71865.
159. Ghorashi N, Ismail A, Ghosh P, Sidawy A, Javan R. AI-Powered Chatbots in Medical Education: Potential Applications and Implications. *Cureus*. 2023 Aug;15(8):e43271.

8. SAŽETAK

Ciljevi: Cilj prve studije u sklopu ove disertacije bio je istražiti laičke, kao i znanstvene sažetke iz područja onkologije, odnosno sustavnih pregleda u Cochrane knjižnici koji su istraživali onkološke intervencije. Karakteristike koje su se ispitivale su bile duljina teksta, indeks čitkosti i emocionalna obilježaj teksta. Potom je cilj bio ustvrditi može li se korištenjem umjetne inteligencije, tj. chatbota temeljenog na naprednom jezičnom modelu poboljšati prezentacija laičkih sažetaka Cochraneovih sustavnih pregleda o onkološkim intervencijama.

Materijali i metode: Provedene su dvije studije. Prva je ispitala čitkost i jezične karakteristike 275 Cochraneovih laičkih sažetaka (engl. *plain language summary*-PLS) iz područja onkologije u usporedbi s njihovim znanstvenim sažecima. Za procjenu čitkosti je korišten Simple Measure of Gobbledygook (SMOG) index. Isto tako, analizirana je kategorija zaključivosti sažetka svakog sustavnog pregleda. Druga studija koristila je Chat generative pre-trained transformer (ChatGPT-4) za preoblikovanje dijela tih PLS-ova, s ciljem pojednostavljenja jezika i poboljšanja tona, uz očuvanje izvornih zaključaka.

Rezultati: Prva studija pokazala je da je prosječan SMOG indeks za PLS-ove bio 13,0, što ukazuje na visoku razinu potrebnog obrazovanja za razumijevanje teksta, a većina sažetaka imala je nejasne zaključke. Nasuprot tome, PLS-ovi generirani pomoću umjetne inteligencije imali su znatno bolju čitkost (SMOG indeks i do 8,2) te su pokazivali izraženije emocionalne i autentične tonove. Ključno je da su sačuvali točnost i suštinu izvornih zaključaka.

Zaključci: Trenutni Cochraneovi PLS-ovi iz područja onkologije često nadilaze preporučenu razinu čitkosti i nerijetko imaju nejasne zaključke, što može ograničiti razumijevanje među pacijentima. Primjena alata poput ChatGPT-4 može znatno unaprijediti pristupačnost ovih sažetaka, što ukazuje na potencijalnu vrijednost umjetne inteligencije u poboljšanju komunikacije usmjerene na pacijenta u okviru medicine utemeljene na dokazima.

9. LAIČKI SAŽETAK

Ova disertacija proučavala je kako su napisani Cochraneovi sažeci o liječenju raka. Osim stručnih, znanstvenih sažetaka, Cochrane uz svoja istraživanja prezentira i sažetak namijenjen za opću populaciju, tzv. laički sažetak. Usporedili smo te laičke sažetke sa znanstvenim verzijama da vidimo koliko su dugi, kakve riječi koriste, koliko su jasni i kakvi su emocionalni tonovi teksta. Zatim smo isprobali može li umjetna inteligencija – točnije *Chat-generative pretrained transformer* (ChatGPT) – poboljšati te sažetke i učiniti ih lakšima za čitanje, bez da se izgubi značenje. Proveli smo dva istraživanja. U prvom smo analizirali 275 laičkih sažetaka iz područja onkologije i usporedili ih sa stručnim verzijama. Provjerili smo koliko su teški za čitanje (koristeći *Simple Measure of Gobbledygook* (SMOG) indeks) i jesu li zaključci jasni. U drugom dijelu koristili smo ChatGPT-4 za preoblikovanje sažetaka. Cilj nam je bio pojednostaviti ih, ali zadržati točnost informacija. Originalni laički sažeci bili su teški za razumijevanje. Za njih je bila potrebna visoka razina obrazovanja (prosječan SMOG indeks 13 – što znači barem 13 godina školovanja). Mnogi su imali nejasne poruke o onkološkom liječenju. S druge strane, sažeci koje je preoblikovao ChatGPT-4 bili su mnogo lakši za čitanje (SMOG indeks oko 8 – što znači da je potrebno 8 godina obrazovanja za razumijeti dani tekst). Zvučali su prirodnije i suosjećajnije, a zadržali su točnost. U zaključku, mnogi sažeci o liječenju raka preteški za opću populaciju i često nejasni. To otežava pacijentima pristup važnim informacijama. Alati poput ChatGPT-4 mogu pomoći da takvi tekstovi budu jednostavniji, jasniji i korisniji, što je važno za bolju informiranost pacijenata koji se liječe od raka.

10. SUMMARY

Title: Linguistic analysis of plain language summaries and scientific summaries of Cochrane systematic reviews of oncology interventions and the use of artificial intelligence in the creation of plain language summaries

Objectives: The aim of the first study within this dissertation was to identify the linguistic characteristics and readability of plain language summaries (PLSs) of Cochrane systematic reviews of oncology interventions, in comparison with the corresponding scientific summaries. This included examining aspects such as text length, proportions of different emotional tones, and readability index. The subsequent objective was to determine whether the use of artificial intelligence—specifically a chatbot based on an advanced language model (chat generative pre-trained transformer (ChatGPT))—could improve the presentation of PLSs of Cochrane systematic reviews on oncology interventions.

Materials and methods: Two studies were conducted. The first examined the readability and linguistic characteristics of 275 Cochrane oncology PLSs compared to their scientific abstracts. The Simple Measure of Gobbledygook (SMOG) index was used to assess readability. Additionally, the conclusiveness category of each systematic review summary was analyzed. The second study used ChatGPT-4 to reformulate these PLSs, aiming to simplify the language and improve tone while preserving the original conclusions.

Results: The first study showed that the average SMOG index for the PLSs was 13.0, indicating a high level of required education, and most summaries had unclear conclusions. In contrast, PLSs generated using artificial intelligence demonstrated significantly better readability (with SMOG scores as low as 8.2) and conveyed more emotional and authentic tones. Crucially, they maintained the accuracy and core messages of the original conclusions.

Conclusions: Current Cochrane PLSs in oncology often exceed the recommended readability level and frequently present unclear conclusions, potentially limiting understanding among patients. The use of tools such as ChatGPT-4 can significantly enhance the accessibility of these summaries, highlighting the potential value of artificial intelligence in improving patient-centered communication within evidence-based medicine.

11. PLAIN LANGUAGE SUMMARY

This dissertation analyzed how Cochrane summaries on cancer treatment are written. Cochrane creates two types of summaries: one for experts and one for the general public, called the plain language summary. We compared these two versions to see how they differ in length, word use, clarity, and emotional tone. We also tested whether artificial intelligence—specifically Chat Generative Pretrained Transformer (ChatGPT)—could help improve the plain language summaries. Our goal was to make them easier to read without changing their meaning. We did two studies. First, we looked at 275 plain language summaries about cancer. We compared each one to its matching scientific version. We checked how hard they were to read using the Simple Measure of Gobbledygook (SMOG) index, how emotional they are and whether their messages were clear. In the second study, we used ChatGPT-4 to rewrite the plain language summaries. We aimed to make them simpler while keeping the information accurate. We found that the original plain language summaries were hard to read. On average, they needed at least 13 years of education to understand. Many also had unclear messages about treatment. In contrast, the summaries rewritten by ChatGPT-4 were much easier to read. They had a SMOG score of around 8, which corresponds to 8 years of education. These versions sounded more natural and caring while staying true to the original message. Many cancer treatment summaries are too complex for most people to understand. This makes it harder for patients to get the information they need. Tools like ChatGPT-4 can help make health information clearer and more useful. This is important for the patients to make better decisions about their care.

12. ŽIVOTOPIS

OSOBNİ PODATCI

Ime i prezime: Jelena Šuto Pavičić

Adresa: Split

Elektronička pošta: jsuto@kbsplit.hr

Državljanstvo: hrvatsko

Datum i mjesto rođenja: 7.6.1995., Split, Hrvatska

IZOBRAZBA

Rujan 2014. – Srpanj 2020. Doktor medicine; Medicinski fakultet Sveučilišta u Splitu, Split, Hrvatska

Adresa: Šoltanska 2, 21000 Split, Hrvatska; web stranica: <http://www.mefst.unist.hr>

Listopad 2020. – trenutno doktorand; Doktorski studij Translacijska istraživanja u biomedicini, Medicinski fakultet Sveučilišta u Splitu, Split, Hrvatska

Adresa: Šoltanska 2, 21000 Split, Hrvatska; web stranica: <http://www.mefst.unist.hr>

RADNO ISKUSTVO

Listopad 2020. – Lipanj 2021. doktor medicine; Dom zdravlja Splitsko-dalmatinske županije, Split, Hrvatska

Kolovoz 2021. – trenutno specijalizant onkologije i radioterapije; Klinika za onkologiju i radioterapiju, KBC Split, Split, Hrvatska

MATERINSKI JEZIK

- hrvatski jezik

OSTALI JEZICI

- engleski jezik
- talijanski jezik

13. DODATAK

13.1 Prilog 1. Tablica laičkih sažetaka

Dostupno na <https://osf.io/92tnv/>.