

Vrijednost tekstualnih podataka u opservacijskim epidemiološkim istraživanjima

Hercigonja-Szekeres, Mira

Doctoral thesis / Disertacija

2010

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

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:105:212637>

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

Download date / Datum preuzimanja: **2024-11-18**



Repository / Repozitorij:

[Dr Med - University of Zagreb School of Medicine Digital Repository](#)





Središnja medicinska knjižnica

Hercigonja-Szekeres, Mira (2010) *Vrijednost tekstualnih podataka u opservacijskim epidemiološkim istraživanjima [Value of textual data in observational epidemiological research]*. Doktorska disertacija, Sveučilište u Zagrebu.

<http://medlib.mef.hr/826>

University of Zagreb Medical School Repository

<http://medlib.mef.hr/>

**SVEUČILIŠTE U ZAGREBU
MEDICINSKI FAKULTET**

Mira Hercigonja-Szekeres

Vrijednost tekstualnih podataka
u opservacijskim
epidemiološkim istraživanjima

DISERTACIJA

Zagreb, 2010.

Disertacija je izrađena na Katedri za medicinsku statistiku, epidemiologiju i medicinsku informatiku Škole narodnog zdravlja „Andrija Štampar“ Medicinskog fakulteta Sveučilišta u Zagrebu.

Voditelj rada: prof. dr. sc. Davor Ivanković

HVALA

Prof. dr. sc. Davoru Ivankoviću, mentoru, na prvom poticaju i potpori, a zatim i zadnjem, presudnom za dovršenje ovog rada.

Prof. dr. sc. Silviju Vuletiću što mi je otvorio vrata u medicinsku statistiku i poticao me na istraživanja u analizi tekstnih podataka.

Prof. dr. sc. Josipi Kern što me je uvjerala da se započeto mora završiti.

Liječnicima obiteljske medicine: dr. sc. dr. Hrvoju Tiljku, dr. Rikardu Perinoviću, dr. Miroslavi Jarnjak i dr. Ireni Erman Baldini na pomoći oko *intensity sampling* ispitanika u ovom istraživanju.

Dr. sc. dr. Ingi Heim osim na pomoći oko *intensity sampling* ispitanika i na korisnim savjetima.

Dr. Ognjenu Brboroviću na analizi Upitnika SF-36.

Profesorici Bojani Zeljko-Lipovšćak na lekturi i jezičnim savjetima.

Mojoj obitelji: Jánosu, Ivanu i Idi te Ivani na nesebičnoj potpori, poticaju, pomoći i razumijevanju. Bez vas ne bih uspjela !

POPIS OZNAKA I KRATICA

SF-36	<i>Short Form 36</i> , Upitnik SF-36
PF	<i>Physical Functioning</i> , fizičko funkcioniranje
RP	<i>Role limitation due to physical problems</i> , ograničenje zbog fizičkih teškoća
BP	<i>Body pains</i> , tjelesni bolovi
GH	<i>General health perception</i> , percepcija općeg zdravlja
VT	<i>Vitality/energy</i> , vitalnost i energija
SF	<i>Social functioning</i> , socijalno funkcioniranje
RE	<i>Role limitation due to emotional problems</i> , ograničenje zbog emocionalnih teškoća
MH	<i>Mental health</i> , psihičko zdravlje
PCS	<i>Physical Component Summary Measure</i> , profil fizičkog zdravlja
MCS	<i>Mental Component Summary Measure</i> , profil psihičkog zdravlja
ICT	<i>Information and Communication Technologies</i> , informacijske i komunikacijske tehnologije
IT	<i>Information technologies</i> , informacijske tehnologije
Dtm-Vic	Data and Text Mining – Visualization, Inference, Classification, računalni program
<u>PCS</u> _{kv}	Srednja vrijednost profila fizičkog zdravlja za klustere dobivene interpretativnom tehnikom analize
<u>MCS</u> _{kv}	Srednja vrijednost profila psihičkog zdravlja za klustere dobivene interpretativnom tehnikom analize
mlad	Dob do 50 godina
sred	Dob od 50 do 64 godine
star	Dob 65 i više godina
m	Muškarci
z	Žene
neoz	Neoženjeni / neudane
brak	Oženjeni / udane
rast	Rastavljeni / rastavljene
udov	Udovci / udovice

osn	Završena osnovna škola
sred.šk.	Završena srednja škola
vis.šk.	Završena viša / visoka škola
bolje	Bolje - samoprocjena svojeg zdravlje u usporedbi sa zdravljem svojih vršnjaka
jedn	Jednako - samoprocjena svojeg zdravlje u usporedbi sa zdravljem svojih vršnjaka
lošije	Lošije - samoprocjena svojeg zdravlje u usporedbi sa zdravljem svojih vršnjaka
tv	Televizija, televizor
gimn	Završena srednja škola
faks	Završena viša / visoka škola
pcs1	Prvi kvartil profila fizičkog zdravlja
pcs2	Drugi kvartil profila fizičkog zdravlja
pcs3	Treći kvartil profila fizičkog zdravlja
mcs1	Prvi kvartil profila psihičkog zdravlja
mcs2	Drugi kvartil profila psihičkog zdravlja
mcs3	Treći kvartil profila psihičkog zdravlja
<u>PCS</u> _{kvan}	Srednja vrijednost profila fizičkog zdravlja za klastere dobivene računalnom tehnikom analize
<u>MCS</u> _{kvan}	Srednja vrijednost profila psihičkog zdravlja za klastere dobivene računalnom tehnikom analize
Clust	<i>Cluster</i> , klaster
EHR	<i>Electronic health record</i> , elektronički zdravstveni zapis

SADRŽAJ

	stranica
1. Uvod	1
2. Hipoteza i cilj rada	7
2.1. Hipoteza	7
2.2. Ciljevi istraživanja	7
3. Etički obziri	9
4. Ispitanici i metode	10
4.1. Ispitanici	10
4.2. Metode	14
4.2.1. Upitnik SF-36	14
4.2.1.1. Definicija i struktura Upitnika SF-36	14
4.2.1.2. Zašto Upitnik SF-36	19
4.2.1.3. Kada i zašto je nastao Upitnik SF-36	19
4.2.2. Kvalitativan pristup u znanstvenom istraživanju	21
4.2.2.1. Povijesne napomene	21
4.2.2.2. Kvalitativan pristup nasuprot kvantitativnom	22
4.2.2.3. Prikupljanje i obrada podataka	23
4.2.3. Obrada tekstnih podataka računalom	26
4.2.3.1. Analiza tekstnih podataka	26
4.2.3.2. Statistička obrada teksta	27
4.2.3.3. Posebni tekstovi – odgovori na otvorena pitanja	27
4.2.3.4. Postupci pri analizi tekstnih podataka	28
4.2.3.5. Računalni program za obradu tekstnih podataka	31
5. Rezultati	33
5.1. Osnovne značajke ispitanika	33
5.2. Rezultati analize Upitnika SF-36	37
5.3. Rezultati kvalitativne analize	43
5.3.1. Prikupljanje podataka	43
5.3.2. Rezultati kvalitativne analize: interpretativna tehnika	44
5.3.3. Rezultati kvalitativne analize: računalna tehnika – analiza tekstnih podataka	48
5.3.3.1. Kompjuterizacija teksta	48
5.3.3.2. Rezultati leksikometrijske analize	49
5.3.3.3. Rezultati korespondencijske i klaster analize	51
6. Rasprava	69
7. Zaključak	85
8. Sažetak	87
9. Summary	89
10. Popis literature	91
11. Životopis	99

1. UVOD

Glavni je izazov svih istraživanja kako obuhvatiti sveukupnost stvarnosti, a ipak ograničiti potencijalno neograničenu količinu informacija koju treba prikupiti za opis te stvarnosti. Ljudi nisu učinkoviti u obradi velikog broja podataka te se koriste raznim strategijama kako organizirati informacije i odlučiti koje su važne. Te se strategije razlikuju u dobro poznatom poretku, no krajnosti u tim pristupima istaknute su posebno, ne samo idejom nego i posebnim imenima, npr. muški i ženski, lijevi i desni, yin i yang. Na isti način slične krajnosti su vidljive u našem pristupu znanju, znanosti pa i našoj definiciji stvarnosti. To su više nego različite metode proučavanja: to su paradigme koje osim različitih metoda uključuju i različite pretpostavke, uvjerenja i vrijednosti. Tako su se kao krajnje suprotne metode iskristalizirale metode kvantitativnog i kvalitativnog istraživanja. Kvantitativne metode ili objektivne metode definirane su kao metode koje uključuju neovisnost, odvojenost i odmak od predmeta istraživanja, nasuprot kvalitativnim metodama koje uključuju bliskost, subjektivnost i podrobnost [1.].

Povijesno ta se dva pristupa izmjenjuju u primjeni u istraživanjima sukladno sa stupnjem razvitka društva i znanja općenito. Razvoj logike i znanosti u 17. i 18. stoljeću podupire vjerovanje da se svijet može tumačiti putem naših osjeta upravljanih znanstvenim metodama. To je vrijeme realizma i logičkog pozitivizma u filozofiji. Poslije su društvena previranja u revolucionarnoj Europi pobudila sumnje može li se ljudsko ponašanje tumačiti tom logikom i tom filozofijom. U 19. stoljeću filozofi koji su podupirali filozofiju idealizma predlažu drugu paradigmu koja je dokazivala da je ljudski um izvor znanja, a društvo je proizvod ljudi i ne može se analizirati poput mehaničkog proizvoda. Umjesto toga društvo treba promatrati kao fenomen koji raste i kreće se te uključuje međusobna ljudska djelovanja. Čovjek je aktivan sudionik postojeće stvarnosti i pokušava joj dati smisao, dakle ponašanje se ne može objasniti bez razumijevanja čovjeka koji je dio te stvarnosti. U tom su filozofskom pravcu osnove fenomenologije i ostalih kvalitativnih pristupa istraživanjima [2.].

Doseg istraživanja u epidemiologiji i zdravstvu obuhvaća dva osnovna područja: ljudsko ponašanje te učinke okoline i bioloških procesa na čovjeka. Dugo su vremena istraživači pretpostavljali, osobito u području zdravstva, samo jedan pristup istraživanjima, većinom kvantitativan. Danas se sve više preporučuje i provodi kombinirani pristup. Čini se gotovo nemogućim na istraživanja primijeniti istodobno oba pristupa koji su dugo smatrani antitetičnima.

Metodološki izazovi istraživanja zdravstva proizlaze iz njegova velikog opsega i različitosti među ljudima: biološke različitosti na osjetljivost na različite bolesti, razlike u ponašanju kao odgovor na različite životne okolnosti i slično. Problem je istraživanja kada se usredotočiti na općenitost, a kada na posebnost, jer se prema njima dva tradicijska pristupa u istraživanju odnose različito. Zdravstveni zahvati mogu biti propisani na svakoj razini; općenito je određeno javnozdravstvenim programima i promicanjem zdravlja. Epidemiologija je istraživačka disciplina, a paradigma je da se iz deduktivnog zaključivanja i kvantitativnih metoda evoluiranja programa dobivaju tipična ili prosječna otkrića i saznanja. Pojedinačni slučaj je određen susretom liječnik - pacijent, istraživačke discipline su klinička medicina ili psihologija, a znanstvena je paradigma indukcija i opis pojedinačnog slučaja, uzimajući krajnje primjere u Freudovim ili Jungovim radovima. Zbog pojedinačnih različitosti dvojba je kliničara koliko općeniti rezultat vrijedi u pojedinačnom, posebnom slučaju. Kvantitativna istraživačka paradigma ne rješava u potpunosti taj izazov.

Kvantitativne metode ostvaruju smanjivanje opsega podataka fokusirajući se na zajedničko i odbacujući pojedinačna odstupanja, prosjek postaje načelom deskriptivne statistike. Odstupanja među grupama su zanimljivija nego odstupanja unutar grupe, posebnosti se odbacuju. Pretpostavka je da su takvi analitički postupci preneseni iz istraživanja u temeljnim znanostima gdje su pojedinačni izabrani uzorci mnogo homogeniji, nego što su homogene odabrane grupe ljudi. Zdravstvena istraživanja pretpostavljaju da ljudi nisu homogeni, ali skloni su tu činjenicu zanemariti. Prednost je kvantitativnih istraživanja da su parsimonična i sažeta, ona daju srž karakteristika grupe s mogućim rizikom da se izgubi utjecaj ekstrema. Ostavljaju mogućnost rasprave o takvim slučajevima naknadno, nakon statističke analize.

Kvantitativne metode i statistički podaci daju nam parsimonični sažetak uzoraka, dok kvalitativna proučavanja mogu ponuditi duboko, detaljno i pojedinačno značenje. Kvantitativna istraživanja nameću vanjske standarde; rezultati su kodirani i analizirani kao numeričke vrijednosti, dok se kvalitativni podaci analiziraju koristeći se ponajprije jezikom, to su uglavnom tekstni podaci. Sažimajući stvarnost u nizove numeričkih podataka, kvantitativne metode mogu izbrisati osnovno značenje tih podataka i dati tim brojkama glavnu ulogu u analizi, bez obzira na osnovno značenje.

Kvalitativni podaci u obliku riječi i opažanja prethodili se kvantitativnima u akademskim promatranjima. Oni su bili glavno uporište biologije, povijesti, prava, a poslije sociologije i političkih znanosti. Kvalitativni podaci su interpretativni, oni generiraju koncepte, tipologije i teorije. Metode prikupljanja su općenito otvorene i iterativne, uključuju analizu pisanih dokumenata, otvorene upitnike (često u grupi) i neposredna promatranja uključujući i promatranje sudionika.

Velika primjena kvalitativnih metoda istraživanja javlja se kao odgovor na predominantnost i uočena ograničenja kvantitativnih istraživanja. To je uključivalo tendenciju kvalitativnih metoda da komprimiraju stvarnost, pojednostavljujući je na praktični sažetak složene cjeline [3., 4.].

Pravilno primijenjene, kvalitativne metode imaju prednost otvorenosti, bez unaprijed određenih kategorija podataka, stila pitanja, kodiranja ili analize. To pridonosi dubini i potankosti ispitivanja. Kvalitativni podaci prenose dublje osjećaje ispitanika, pomažu vidjeti svijet očima ispitanika te izazivaju empatiju kod promatrača.

Istraživač utječe na kvalitativna istraživanja i to od samog početka određujući predmet istraživanja. U kvantitativnim istraživanjima pogreške su vidljive odmah kod definiranja hipoteze ili određivanja sadržaja upitnika. Međutim u kvalitativnim istraživanjima pogreške se pojavljuju poslije, tek u analizi podataka ili u interpretaciji rezultata analize. Riječi imaju višesmisleno značenje i njihova se interpretacija može razlikovati od istraživača do istraživača. U kvalitativnim istraživanjima istraživač je instrument istraživanja, kvaliteta podataka ovisi o njegovoj vještini i odlučnosti da prikupi podatke te ih analizira. Naime, istraživač je u tim istraživanjima često u

situaciji da u prikupljanje podataka kreće s jednom idejom i tako izabire uzorak ispitanika, a tijekom analize podataka nameću se dodatne ideje, pa ponekad još i tijekom samog prikupljanja. Količina podataka dobivena kvalitativnim istraživanjima izuzetno je velika, obično su to mnogobrojne stranice tekstova i zapisa, sati videozapisa ili kutije dokumentacije. Zbog toga su veličine istraživanih uzoraka kod kvantitativnih istraživanja razmjerno manje nego kod kvalitativnih, a pogreške kod odabira uzorka veće. Često se tek poslije ustanovi da su u istraživanju zadržani neinformativni uzorci, a propušteni zanimljivi. Ipak, strategija izbora uzoraka uvijek mora biti dobro objašnjena bilo da se radi o kvantitativnim ili kvalitativnim istraživanjima, a manji rizik lošeg izbora uvjetovan je dobrim predznanjima istraživača te njegovom dobrom informiranosti [5., 6., 7.].

Prije su kvalitativne analize često bile obilježene kao nesistematične, pa čak i tajanstvene. Takva je obilježja jednostavno moguće izbjeći velikom transparentnošću analitičkih metoda koja će omogućiti i nekom drugom istraživaču da ih može ponoviti. Intuitivne, češće danas nazivane interpretativne, metode analize kvalitativnih podataka moraju biti objašnjene i validirane kako ne bi bile obilježene kako je prije rečeno.

Danas prevladava mišljenje da oba pristupa istraživanju u zdravstvu imaju podjednaku valjanost, dapače da ih treba kombinirati u svim fazama istraživanja; u prikupljanju podataka, njihovoj analizi i na kraju interpretaciji te konačnom izvještaju.

Razlikujemo dva osnovna načina kombiniranja kvalitativnog i kvantitativnog pristupa istraživanjima: hijerarhijski i partnerski.

Hijerarhijsko kombiniranje znači da ćemo najprije primijeniti jedan pristup, a zatim drugi. Bilo je uobičajeno da se odluka koji će pristup biti primijenjen prvi, a koji drugi donosi u postavlja istraživanja ovisno o sadržaju istraživanja i, donekle, ispitanicima. Tako da su se numerički podaci nakon obrade interpretirali nekom od kvalitativnih metoda, a podaci dobiveni kvalitativnim pristupom analizirali metodama svojstvenim za kvantitativni pristup.

Partnerski pristup kombiniranju obiju metoda znači primjenu tih metoda simultano, sekvencijalno ili ciklički. Kada će se koristiti koja primjena, ovisilo je o fazi

istraživanja. Kod prikupljanju podataka obično su se kombinirala ta dva pristupa istodobno; podaci su prikupljeni istodobno postupcima tipičnim za kvantitativni pristup, ali i razgovorom, promatranjem, snimanjem. Analiza podataka bila je prikladna za sekvencijalnu ili cikličku primjenu kombiniranja pristupa, ovisno o istraživanju, ali još više o istraživačima. Slična je situacija i s interpretacijom analiziranih podataka, gdje su se opet kombinirale obje primjene partnerskoga pristupa.

Danas, kada su kvalitativne i kvantitativne metode u istraživanjima u svim područjima, pa tako i u epidemiologiji i zdravstvu, potpuno izjednačene, nema ni potrebe naglašavati kombinirani pristup, niti posebno naglašavati neku metodu. Očekivano je i normalno da se sve metode primjenjuju simultano prema potrebi s ciljem provođenja što kvalitetnijeg istraživanja te dobivanja relevantnih informacija [8., 9., 10., 11., 12., 13., 14., 15.].

U ovom radu u svakoj je fazi korištena kombinacija obaju pristupa istraživanju. Namjera je bila da se u svakoj fazi istraživanja uspoređuje kvalitativni pristup s kvantitativnim te u što većoj mjeri evaluira [16.].

Podaci su prikupljeni na dva načina: razgovorom i upitnikom. Time su uključena oba pristupa u svojim osnovnim postavkama.

Razgovor je vodio ispitivač osobno te zapisivao bez bitnih promjena. Vođenjem tzv. tematski ciljanog ispitivanja kroz razgovor namjera je bila više približiti se razgovoru kakav očekuje medicinsko osoblje, osobito liječnik s pacijentom. Pitanja poput „Kako se osjećate?“, „Koje tegobe imate?“, „Što vas ometa u svakodnevnom životu?“, „Sprečava li vas vaše zdravstveno stanje u vašim svakodnevnim aktivnostima?“ i slična uobičajena su pitanja u komunikaciji svakog liječnika s njegovim pacijentom i trebala bi biti dio svakog razgovora [17., 18., 19.].

Upitnik je opće poznat i priznat Upitnik SF-36 čiji se rezultati izražavaju numerički i omogućavaju daljnju kvantitativnu obradu, primjenu i usporedbu [20.]. Ostali podaci o ispitanicima korišteni u ovom istraživanju prikupljeni su upitnikom koji su ispitanici ispunjavali uz SF-36.

Analiza podataka je provedena ovisno o vrsti podataka. Zapisani razgovori analizirani su na dva načina – interpretativnom metodom tipičnom samo za kvalitativni pristup istraživanjima te uz pomoć posebnog računalnog programa koji je na određeni način kvantitativna metoda analize podataka dobivenih kvalitativnim pristupom. Ostali podaci, Upitnik SF-36 i osobni podaci analizirani su standardnim kvantitativnim metodama.

Interpretativna metoda je način na koji i liječnik stvara dojam pri razgovoru s pacijentom.

Kvantitativna metode analize zapisanih razgovora uključila je posebni računalni program i statističke metode obrade tekstnih podataka [21.].

Interpretacija dobivenih rezultata u potpunosti je kombinacija obaju pristupa i pritom su istodobno primijenjene sve metode kombiniranja.

Kombinacija kvalitativnog i kvantitativnog pristupa osobito se čini dobro primjenjivom kod ispitivanja u epidemiološkim istraživanjima ponajprije s tekstnim podacima. Analiza tekstnih podataka predodređena je za interpretativnu metodu, a u novije vrijeme razvijene su metode tekstne statistike taj isti tekst obrađuju u skladu sa statističkim metodama obrade kvalitativnih podataka.

2. HIPOTEZA I CILJ RADA

2.1. HIPOTEZA

Kvalitativni i kvantitativni pristup u epidemiološkim istraživanjima su komplementarni postupci, a ne suprotstavljeni postupci.

U kvalitativnom pristupu epidemiološkim istraživanjima rezultati dobiveni računalnom tehnikom (statističkom analizom tekstnih podataka) i rezultati dobiveni obradom podataka interpretativnom tehnikom usporedivi su, nadopunjuju se i obogaćuju epidemiološko istraživanje.

2.2. CILJEVI ISTRAŽIVANJA

Ciljevi istraživanja:

- Usporediti rezultate kvalitativne analize podataka dobivenih otvorenim ispitivanjima u epidemiološkim istraživanjima s podacima dobivenima statističkom analizom podataka iz zatvorenih te prethodno patentiranih i validiranih upitnika.

- Kvalitativna analiza provest će se na dva načina
 - interpretativnom tehnikom – osobnim objektivnim dojmom ispitivača
 - računalnom tehnikom – statističkom analizom tekstnih podataka (slobodnog teksta).

- Usporediti primjenjivost statističke analize tekstnih podataka te njezin doprinos saznanjima o ispitanicima dobivenim interpretativnom tehnikom.

- Podaci će biti prikupljeni metodom *intensity sampling*. To je metoda izbora ili traženja uzoraka koji su osobito bogati obilježjem koje se istražuje. U principu to nisu ekstremni niti devijantni uzorci. Takva metoda ne temelji se na slučajnom uzorku nego pretpostavlja prethodno prikupljanje informacija o mogućem uzorku i dodatni rad na traženju takvog uzorka i prikupljanju podataka o njima.

3. ETIČKI OBZIRI

Svakom ispitaniku je bilo objašnjeno istraživanje, opisana su mu oba postupka, ispunjavanje Upitnika SF-36 te razgovor koji se snima diktafonom i poslije upisuje u računalo.

Zajamčena mu je privatnost tako da se iz prikupljenih i kasnije obrađenih podataka neće moći prepoznati niti jedan osobni podatak. Zatražena je njegova suglasnost da sudjeluje u tom istraživanju, a pisanu suglasnost svaki je ispitanik i potpisao.

Te su potpisane izjave dio ovog istraživanja.

4. ISPITANICI I METODE

4.1. ISPITANICI

Uzorak čini 117 ispitanika prikupljenih metodom *intensity sampling*. Dakle ispitanici su odabrani kao tipični slučajevi izbjegavajući ekstremne i devijantne kako bi se dobio uzorak vrlo bogat podacima i varijabilnošću.

Ispitanici s područja Grada Zagreba stanuju u različitim dijelovima grada i to im zbog specifičnosti tih pojedinih dijelova grada daje veliku različitost. Dio njih je s područja Novog Zagreba, stanuje u velikim stambenim zgradama, uglavnom su se prije više od 20 – 30 godina doselili u Zagreb. Dio ispitanika je s gradskog područja Kustošije. Oni su većinom druga generacija pridošlih u Zagreb. Ostali ispitanici iz Zagreba iz drugih su dijelova grada, većina je iz centra ili šireg centra grada.

Dvanaest ispitanika je s područja slavonske općine Velika i stanovnici su samo ruralnog područja.

Četrnaest ispitanika je s područja istarskog mjesta Tinjan, većina ih je s ruralnog područja, ali dio su stanovnici gradića tipičnog za unutrašnjost Istre.

U Tablici 4.1. je razdioba ispitanika prema mjestu stanovanja.

Tablica 4.1. Razdioba ispitanika prema mjestu stanovanja.

Mjesto stanovanja		Broj	Ukupno
Zagreb	Novi Zagreb	21	18,0
	Kustošija	17	14,5
	(širi) centar grada	49	41,9
Velika		12	10,3
Tinjan		14	12,0
Okolica Zagreba		4	3,3
Ukupno		117	100

Svi su ispitanici punoljetni. U uzorku je 52 muškarca i 65 žena (Tablica 4.2.).

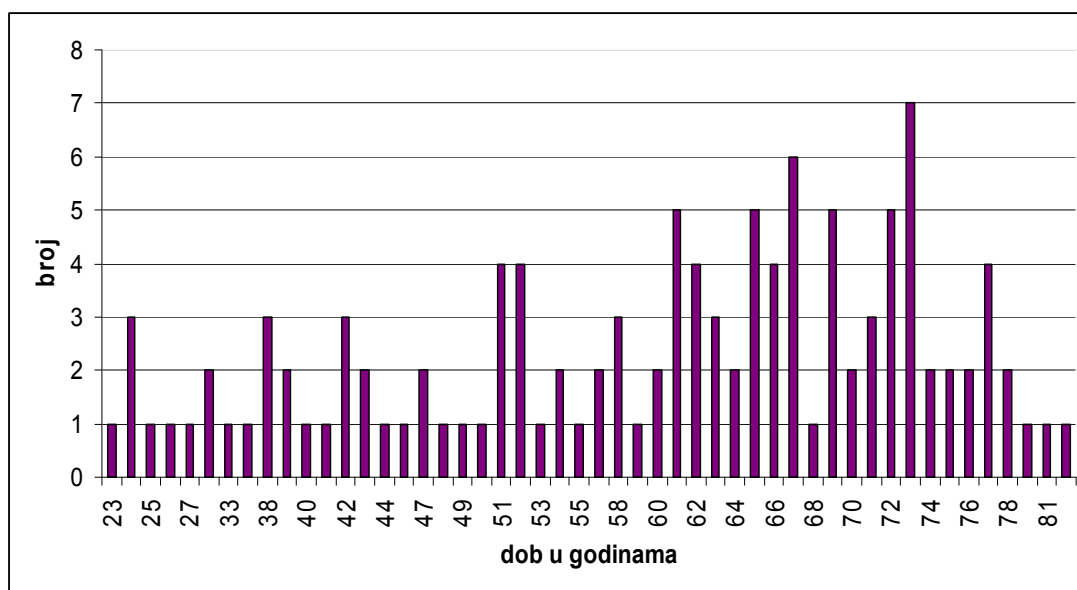
Tablica 4.2. Razdioba ispitanika prema spolu.

Spol	Broj	Postotak
Muškarci	52	44,4
Žene	65	55,6
Ukupno	117	100

Dob ispitanika je vrlo različita, najmlađi je imao u vrijeme ispitivanja 23 godine, a najstariji 82 godine.

Prosječna dob svih ispitanika je 58,7 godina sa standardnom devijacijom 15,1; medijan je 62,0.

Histogram ispitanika prema dobi je na Slici 4.1.



Slika 4.1. Histogram ispitanika prema dobi.

Ispitanici su različitih zvanja. Najviše, njih 13, bilo je ekonomskoga zvanja (diplomirani ekonomist 6, ekonomist 7), zatim slijede profesori (6), trgovci (5), krojači / krojačice (5), diplomirani inženjeri arhitekture (4), liječnici (4), diplomirani inženjeri

strojarstva (3), kuharice (3), radnici (3) te ispitanici ostalih zvanja, dok se 14 nije izjasnilo o svojem zvanju, odnosno završena škola im nije dala neko određeno zvanje.

Čak 46% ispitanika završilo je višu ili visoku školu, 40% srednju, a samo nešto manje od 14% osnovnu školu, tako da razdioba ispitanika prema završenoj školi ne odgovara razdiobi u Hrvatskoj. Naime, prema popisu stanovništva iz 2001. postotak stanovnika starijih od 15 godina sa završenom višom ili visokom školom uključujući i one koji su magistrirali i doktorirali je 12%, sa završenom srednjom školom 47%, a s osnovnom (uključujući i nepotpunu) je 38% [22.].

Razdioba ispitanika prema završenoj školi je u Tablici 4.3.

Tablica 4.3. Razdioba ispitanika prema završenoj školi.

Završena škola	Broj	Postotak
Završena osnovna škola	16	13,7
Završena srednja škola	47	40,2
Završena viša/visoka škola	54	46,1
Ukupno	117	100

Ispitanici su velikom većinom (73,5 %) u braku. Bračno stanje ispitanika prikazano je u Tablici 4.4.

Tablica 4.4. Razdioba ispitanika prema bračnom stanju.

Bračno stanje	Broj	Postotak
Neoženjen / neudana	18	15,4
Oženjen / udana	86	73,5
Rastavljen / rastavljena	6	5,1
Udovac / udovica	7	6,0
Ukupno	117	100

Dob ispitanika odredila je i njihov radni status, velikom su većinom, njih 69 ili 59,0% umirovljenici. Od svih ispitanika njih 39 ili 33,3% rade, a posebno su izdvojene ispitanice koje su se izjasnile da su kućanice, njih 9 ili 7,7%. Naime, činilo se ipak bitnim istaknuti da kućanice imaju malo drugačiji status: rade i to više od punog radnog vremena, a prema nekom uvriježenom shvaćanju nemaju radni status i čini se da ne privređuju.

Za potrebe pohrane, pretraživanja i temeljnih opisa podataka izrađena je vlastita relacijska baza podataka u programu MS Access 2000® (Microsoft Corporation, Redmond, WA, SAD).

4.2. METODE

4.2.2. UPITNIK SF-36

4.2.2.2. DEFINICIJA I STRUKTURA UPITNIKA SF-36

Upitnik SF-36 (Short Form 36) je višestruko primjenljiv upitnik zdravstvenog statusa s 36 pitanja. To su pitanja s odgovorima tipa višestrukog izbora: od 3 do 6 ponuđenih.

Rezultat se standardno izražava u osam dimenzija koje čine profil zdravstvenog statusa:

1. fizičko funkcioniranje (PF, engl. *physical functioning*): 10 pitanja
2. ograničenje zbog fizičkih teškoća (RP, engl. *role limitation due to physical problems*): 4 pitanja
3. tjelesni bolovi (BP, engl. *body pain*): 2 pitanja
4. percepcija općeg zdravlja (GH, engl. *general health perception*): 5 pitanja
5. vitalnost i energija (VT, engl. *vitality / energy*): 4 pitanja
6. socijalno funkcioniranje (SF, engl. *social functioning*): 2 pitanja
7. ograničenje zbog emocionalnih teškoća (RE, engl. *role limitation due to emotional problems*): 3 pitanja
8. psihičko zdravlje (MH, engl. *mental health*): 5 pitanja

U profil fizičkog zdravlja (PCS, engl. *Physical Component Summary Measure*) ulaze od prije spomenutih osam dimenzija sljedeće četiri: fizičko funkcioniranje, ograničenje zbog fizičkih teškoća, tjelesni bolovi i percepcija općeg zdravlja. U profil psihičkog zdravlja (MCS, engl. *Mental Component Summary Measure*) ubrajaju se: vitalnost i energija, socijalno funkcioniranje, ograničenje zbog emocionalnih teškoća i psihičko zdravlje.

Na Slici 4.2. (1/3 – 3/3) je hrvatska verzija upitnika SF-36.

Upitnik SF-36 predstavlja teorijski utemeljenu i empirijski provjerenu operacionalizaciju dvaju generalnih koncepta zdravlja – fizičko zdravlje i psihičko zdravlje te njihove dvije općenite manifestacije: funkcioniranje i dobrobit [20.].

SUBJEKTIVNA OCJENA ZDRAVLJA (SF - 36)

1. Općenito, biste li rekli da je Vaše zdravlje (zaokružite jedan odgovor):

Izvršno	1
Vrlo dobro	2
Dobro	3
Zadovoljavajuće	4
Loše	5

2. U usporedbi s prošlom godinom, kako biste sada ocijenili svoje zdravlje?

Puno bolje nego prije godinu dana	1
Malo bolje nego prije godinu dana	2
Otprilike isto kao prije godinu dana	3
Malo lošije nego prije godinu dana	4
Puno lošije nego prije godinu dana	5

Sljedeća pitanja se odnose na aktivnosti kojima se možda bavite tijekom jednog tipičnog dana. Je li Vas trenutno Vaše zdravlja ograničava u obavljanju tih aktivnosti? Ako DA, u kojoj mjeri?

	DA puno	DA malo	NE nimalo
3. Fizički naporne aktivnosti kao što su trčanje, podizanje teških predmeta	1	2	3
4. Umjereno naporne aktivnosti npr. pomicanje stola, boćanje, vožnja biciklom	1	2	3
5. Podizanje ili nošenje torbe s namirnicama	1	2	3
6. Uspinjanje uz stepenice (nekoliko katova)	1	2	3
7. Uspinjanje uz stepenice (jedan kat)	1	2	3
8. Saginjanje, klečanje, prigibanje	1	2	3
9. Hodanje više od 1 km	1	2	3
10. Hodanje oko pola km	1	2	3
11. Hodanje 100 m	1	2	3
12. Kupanje ili oblačenje	1	2	3

Jeste li u proteklih mjesec dana u svom radu ili drugim redovitim dnevnim aktivnostima imali neki od sljedećih problema zbog svojeg fizičkoga zdravlja? (zaokružite jedan broj u svakom redu)

	DA	NE
13. Skratili ste vrijeme provedeno u radu i drugim aktivnostima	1	2
14. Obavili ste manje nego što ste željeli	1	2
15. Niste mogli obavljati neke poslove ili druge aktivnosti	1	2
16. Imali ste poteškoća pri obavljanju posla ili nekih drugih aktivnosti (npr. morali ste uložiti dodatni trud)	1	2

Slika 4.2. (1/3) Upitnik SF-36 (hrvatska verzija).

Jeste li u proteklih mjesec dana imali neke od dolje navedenih problema na poslu ili pri obavljanju nekih drugih svakodnevnih aktivnosti zbog bilo kakvih emocionalnih problema? (npr. osjećaj depresije ili tjeskobe)

	DA	NE
17. Skratili ste vrijeme provedeno u radu i drugim aktivnostima	1	2
18. Obavili ste manje nego što ste željeli	1	2
19. Niste obavili posao ili neke druge aktivnosti onako pažljivo kao obično	1	2
20. U kojoj su mjeri u proteklih mjesec dana Vaše fizičko zdravlje ili Vaši emocionalni problemi utjecali na Vaše uobičajene društvene aktivnosti u obitelji, s prijateljima, susjedima ili drugim ljudima ?		
Uopće ne	1	
U manjoj mjeri	2	
Umjereno	3	
Prilično	4	
Izrazito	5	
21. Kakve ste tjelesne bolove imali u proteklih mjesec dana?		
Nikakve	1	
Vrlo blage	2	
Blage	3	
Umjerene	4	
Teške	5	
Vrlo teške	6	
22. U kojoj su Vas mjeri ti bolovi u proteklih mjesec dana ometali u vašem uobičajenom radu (uključujući rad izvan kuće i kućne poslove)?		
Uopće ne	1	
Malo	2	
Umjereno	3	
Prilično	4	
Izrazito	5	

Sljedeća pitanja govore o tome kako se osjećate i kako ste se osjećali u proteklih mjesec dana.

Molim Vas da za svako pitanje odaberete po jedan odgovor koji će najbliže odrediti kako ste se osjećali.

Koliko ste (se) vremena u proteklih mjesec dana: (zaokružite jedan odgovor u svakom redu)

Slika 4.2. (2/3) Upitnik SF-36 (hrvatska verzija).

	Stalno	Skoro uvijek	Dobar dio vremena	Povremeno	Rijetko	Nikada
23. Osjećali puni života	1	2	3	4	5	6
24. Bili vrlo nervozni	1	2	3	4	5	6
25. Osjećali tako potištenim da Vas ništa nije moglo razvedriti	1	2	3	4	5	6
26. Osjećali spokojnim i mirnim	1	2	3	4	5	6
27. Bili puni energije	1	2	3	4	5	6
28. Osjećali malodušnim i tužnim	1	2	3	4	5	6
29. Osjećali iscrpljenim	1	2	3	4	5	6
30. Bili sretni	1	2	3	4	5	6
31. Osjećali umornim	1	2	3	4	5	6
32. Koliko su Vas u proteklih mjesec dana vaše fizičko zdravlje ili emocionalni problemi ometali u društvenim aktivnostima (npr. posjete rodbini, prijateljima itd.)?						
Stalno	1					
Skoro uvijek	2					
Povremeno	3					
Rijetko	4					
Nikada	5					
Koliko je u Vašem slučaju TOČNA ili NETOČNA svaka od dolje navedenih tvrdnji						
	Potpuno točno	Uglavnom točno	Ne znam	Uglavnom netočno	Potpuno netočno	
33. Čini mi se da se razbolim lakše nego drugi ljudi	1	2	3	4	5	
34. Zdrav sam kao i bilo tko drugi koga poznajem	1	2	3	4	5	
35. Mislim da će mi se zdravlje pogoršati	1	2	3	4	5	
36. Zdravlje mi je odlično	1	2	3	4	5	

Slika 4.2. (3/3) Upitnik SF-36 (hrvatska verzija).

Na Slici 4.3. je grafički prikaz strukture upitnika SF-36 njegovih dimenzija i ukupnih profila.

pitanje	dimenzija	profil
3. Fizički naporne aktivnosti 4. Umjereno naporne aktivnosti 5. Podizanje ili nošenje torbe 6. Uspinjanje uz stepenice (više katova) 7. Uspinjanje uz stepenice (jedan kat) 8. Saginjanje, klečanje, prigibanje 9. Hodanje više od 1 km 10. Hodanje oko pola km 11. Hodanje 100 m 12. Kupanje ili oblačenje	fizičko funkcioniranje (PF)	fizičko zdravlje (PCS)
13. Skratili ste vrijeme 14. Obavili ste manje 15. Niste mogli obavljati 16. Imali ste teškoća	ograničenje zbog fizičkih teškoća (RP)	
21. Jačina bolova 22. Utjecaj bolova	tjelesni bolovi (BP)	
1. Subjektivni osjećaj zdravlja 33. Razbolim se lakše nego drugi 34. Zdrav sam kao i drugi 35. Pogoršanje zdravlja 36. Odlično zdravlje	percepcija općeg zdravlja (GH)	
23. Osjećali se puni života 27. Bili puni energije 29. Osjećali se iscrpljenim 31. Osjećali se umornim	vitalnost i energija (VT)	psihičko zdravlje (MCS)
20. Društvene aktivnosti 32. Ometanje društvenih aktivnosti	socijalno funkcioniranje (SF)	
17. Skratili ste vrijeme 18. Obavili ste manje 19. Niste obavili aktivnosti	ograničenje zbog emocionalnih teškoća (RE)	
24. Bili ste vrlo nervozni 25. Osjećali se potištenim 26. Osjećali se spokojnim i mirnim 28. Osjećali se malodušnim i tužnim 30. Bili ste sretni	psihičko zdravlje (MH)	

Slika 4.3. Struktura upitnika SF-36.

Pojedini odgovori na svako pitanje se različito boduju, odnosno diferencijalno ponderiraju prema unaprijed utvrđenim empirijskim normama, a s obzirom na dijagnostičku vrijednost određenog odgovora ispitanika. Pitanje koje se odnosi na promjenu u zdravlju prikazuje se odvojeno, distribucijom frekvencija. Osim toga, pojedine manifestacije zdravlja obuhvaćene su različitim brojem pitanja, kako je to vidljivo iz strukture upitnika na Slici 4.3. Taj je broj u skladu s psihometrijskim kriterijima pouzdanosti i valjanosti. Svaki broj bodova zabilježen u svakoj skali

upitnika transformira se u standardne vrijednosti i baždaren je na jedinstvenu skalu čiji je teorijski minimum jednak 0, a maksimum 100 bodova. Takav način vrednovanja omogućava kvantitativno uspoređivanje različitih manifestacija zdravlja koje se mjere upitnikom, interpretaciju ukupne razine i diferenciranost osam dimenzija [20.].

4.2.2.3. ZAŠTO UPITNIK SF-36 ?

Upitnik SF-36 je poznat i vrlo često primjenjivan upitnik zdravstvenog statusa širom svijeta tako da ima pozitivnu međunarodnu reputaciju. Pokazao se vrlo primjenljiv i kod istraživanja zdravstvenog statusa starijih osoba. Razumljiva i obična pitanja omogućavaju njegovu primjenu neovisno o stupnju obrazovanja ispitanika i njihovog sociokulturnog podrijetla. Odgovaranje na svih 36 pitanja u upitniku je kratkotrajno (10 – 20 minuta) tako da ga ispitanici ispune brzo i ne odustaju zbog nezainteresiranosti i opterećenosti [23., 24.].

Upitnik SF-36 primijenjen je u Hrvatskoj u nizu istraživanja. Na Školi narodnog zdravlja „Andrija Štampar“ korišten je u sklopu projekta „Analiza tranzicije zdravstvenog sustava u Hrvatskoj“ u razdoblju 1997-1999. na reprezentativnom uzorku hrvatske populacije koji je obuhvatio oko 1% stanovnika svih regija u Hrvatskoj. Za tu je potrebu preveden i pripremljeno je hrvatsko izdanje. Taj je upitnik korišten i u ovom istraživanju [25., 26., 27.].

4.2.2.4. KADA I ZAŠTO JE NASTAO UPITNIK SF-36

Krajem osamdesetih godina 20. stoljeća pacijentovo gledište u praćenju kakvoće medicinske njege postaje jedno od važnijih dostignuća na polju zdravstvene zaštite. Pacijent, odnosno ispitanik, najbolji je izvor informacija o postizanju osnovnog cilja zdravstvene zaštite: kakvoće medicinske njege ili, vrlo općenito, zdravstvene usluge. Naime, u većini kliničkih ispitivanja, pa ni u medicinskoj praksi nisu se rutinski prikupljale informacije kako pacijenti doživljavaju bolest i liječenje. Te informacije nisu bile dio zdravstvenih dokumenata, a većinom nisu ni danas dio

(elektroničkog) zapisa pacijenta. To je razlog što se takva informacija nije ni mogla analizirati, ni upotrijebiti za unaprjeđenje zdravstvene njege. Međutim u zadnjim desetljećima 20. stoljeća započinje razdoblje u kojem (politički) analitičari uspoređuju troškove i investicije organiziranja i financiranja službi za zaštitu zdravlja. Sve je, pa tako i zdravlje na otvorenom tržištu, tako da je važna svaka informacija o stanju funkcija, dobrobiti i drugim ishodima zdravlja. Te su informacije podjednako važne kliničkim istraživačima pri vrednovanju novih vrsta liječenja i tehnologija, ali i liječnicima te ostalim davateljima zdravstvenih usluga koji su u svakodnevnom doticaju s pacijentom i pokušavaju postići najbolje moguće ishode za pacijenta.

U tom se razdoblju stvaraju i ubrzo etabliraju razni standardizirani upitnici za pacijente koji tako postaju glavnim izvorom novih informacija o općim ishodima zdravlja. Razvijaju se različiti mjerni instrumenti za mjerenja zdravstvenog stanja pojedinaca. Sve se više prihvaća shvaćanje da su fizičko i psihičko zdravlje među glavnim komponentama kvalitete života povezanim sa zdravljem. Tek u zadnjem desetljeću 20. stoljeća možemo uočiti znatan napredak u stvaranju koncepcije zdravlja prema mnogo široj koncepciji kvalitete življenja, a povezano s time i napredak u razvijanju metoda za mjerenje zdravlja s pomoću standardiziranih upitnika. To je vrijeme kada istraživanja u svijetu ne vrednuju samo i u prvom redu troškove zdravstvene zaštite, nego se sve više usredotočuju na ukupne ishode zdravstvene zaštite. Tada nastaje niz upitnika koji ispituju zdravlje ispitanika, ali preko njegova osobnog stava prema svojem zdravlju, svojim tegobama, bolestima i zdravstvenoj zaštiti [28., 29., 30.].

Taj niz upitnika otvorio je novu problematiku: kako pristupiti razvoju i standardizaciji takvih upitnika s prvenstvenom namjerom da rezultati mogu biti ponovljivi u raznim zemljama radi mogućnosti usporedbe.

U to je vrijeme razvijan i Upitnik SF-36. Taj je upitnik uskoro postao dobro poznat i često primjenjivan za mjerenje zdravlja i dobrobiti. Testiran je u različitim sredinama i vrstama upravljanja, na ispitanicima različitog stupnja obrazovanja i dobi te vrlo različitog zdravstvenog statusa. Upitnik SF-36 je metodološki veoma utemeljen kroz svoj razvoj, koji nije prestao sve do današnjih dana i stalno se dograđuje [20.].

4.2.3. KVALITATIVAN PRISTUP U ZNANSTVENOM ISTRAŽIVANJU

Kvalitativan pristup u znanstvenom istraživanju primjenljiv je u mnogim strukama i različitim sadržajima. Takav se pristup primjenjuje osobito u cilju dubljeg prepoznavanja ljudskog ponašanja i razloga koji uvjetuju određena ponašanja. Karakteristična pitanja su *zašto* i *kako*, a ne samo *što*, *gdje*, *kada* i *koliko*. Karakterističan je i uzorak koji se ispituje – obično su to manje i ciljane skupine.

4.2.3.2. POVIJESNE NAPOMENE

Kvalitativan pristup u znanstvenom istraživanju bio je početkom 20. stoljeća prvi način osobito u socijalnim istraživanjima. Međutim pedesetih i šezdesetih godina prevladava kvantitativan pristup, tako da se čak govori o kvantitativnoj revoluciji. Ipak je sedamdesetih godina prošlog stoljeća ponovno „otkriven“ kvalitativan pristup, ponajprije kao dio istraživanja u antropologiji i sociologiji. Kasnih sedamdesetih i osamdesetih godina 20. stoljeća kvalitativan pristup u znanstvenim istraživanjima počinje se rabiti i u drugim znanostima. Postaje vodeći pristup u istraživanju obrazovanja, socijalnim istraživanjima, informacijskim istraživanjima, istraživanjima iz područja promidžbe, posloводства, komunikacije te psihologije, nešto kasnije primjenjuje se u istraživanjima iz područja sestrintstva, skrbi, psihologije i zdravstvene zaštite. Kvalitativan pristup istraživanjima doživljava veliku primjenu i nagli razvoj uporabom u analizama tržišta te proučavanjima mogućnosti i razvoja te oglašavanja i reklame. U zadnja dva desetljeća 20. stoljeća razvijene su nove metode kvalitativnog pristupa, osobito nakon niza vrlo kritičkih primjedbi sljedbenika (samo) kvantitativnog pristupa [31., 32., 33.].

Za potvrdu etabliranja kvalitativnog pristupa u znanstvenom istraživanju osobito je važna činjenica da unatrag dvadesetak godina mnogi vodeći znanstveni i stručni časopisi prihvaćaju radove s tematikom kvalitativnog pristupa istraživanjima, pa čak i u prirodnim znanostima. Dapače, zadnjih godina postoje časopisi koji objavljuju u prvom redu radove s tematikom kvalitativnog pristupa istraživanjima i njihov je broj u stalnom porastu. Neka istraživanja pokazuju da se svake godine pojavi barem jedan časopis takvog usmjerenja [34.].

4.2.2.2. KVALITATIVAN PRISTUP NASUPROT KVANTITATIVNOM

O razlikama i suprotnostima kvalitativnog i kvantitativnog pristupa znanstvenom istraživanju postoji golema količina radova, od onih koji u potpunosti negiraju kvalitativni pristup u korist kvantitativnog do onih u kojima se tvrdi suprotno [35].

Tri su činjenice u kojima se ta dva pristupa metodološki razlikuju:

I. Ispitanici.

U kvalitativnom pristupu nikad nije slučajni uzorak, nego su ispitanici izabrani prema tematici istraživanja.

II. Istraživač.

U kvalitativnom pristupu istraživač mora biti osobito oprezan kako osobnim stavovima i / ili pristupom ne bi poremetio objektivnost istraživanja. Uvriježeno je mišljenje da istraživač nikad nije posve objektivan u takvim istraživanjima.

III. Podaci i analiza podataka.

U kvalitativnom pristupu podaci su najčešće izraženi riječima. To znači da se kvalitativna analiza podataka s jedne strane treba temeljiti na postupcima vezanim s jezikom, znakovima i značenjem, ali istodobno i na postupcima koji uključuju kontekstni i holistički pristup.

Uvriježen je stav da se kvalitativni pristup primjenjuje kao predistraživanje ili uvod u velika istraživanja koja će se (ipak) provesti kvantitativnim pristupom, u potpunosti ili većim dijelom. Naime, kvantitativni pristup je etabliran kao tipičan, pouzdan i precizan u istraživanjima koja se onda potvrđuju mjernim instrumentima i matematičkim postupcima. Nasuprot tome kvalitativni podaci se obično teško prikazuju grafički ili kroz matematičke pojmove.

Kvalitativan pristup se često primjenjuje i za vrednovanje smjera i programa istraživanja. Naime, tim se pristupom lakše dolazi do saznanja kako i zašto su se neki (očekivani) rezultati pojavili ili nisu, ali i do saznanja o relevantnosti istraživanja, neočekivanim učincima i djelovanjima. Velika je prednost kvalitativnog pristupa

istraživanju ta što dopušta veliku različitost u odgovorima te se može prilagoditi novim spoznajama ili materijalima odmah još u tijeku samog istraživanja.

Kvalitativna istraživanja smatraju se skupljima od kvantitativnih, što je dosta teško objektivno procijeniti, ali činjenica oko koje se velika većina slaže jest da kvalitativna istraživanja traju dulje.

4.2.3.3. PRIKUPLJANJE I OBRADA PODATAKA

Istraživači se pri kvalitativnom pristupu znanstvenom istraživanju koriste vrlo različitim metodama prikupljanja podataka i njihovim kombinacijama. Te metode ovise o području istraživanja, sadržaju, ali i cilju istraživanja.

Obično se primjenjuju sljedeće metode prikupljanja podataka [36.]:

- sudioničko promatranje
- nesudioničko promatranje
- strukturirani razgovor
- nestrukturirani razgovor
- terenske bilješke
- analiza dokumenata, slika, predmeta, ostalih materijala.

Podaci prikupljeni nekom od prije navedenih metoda ili najčešće njihovom kombinacijom obrađuju se i interpretiraju za konačan izvještaj.

Metode obrade podataka dobivenih kvalitativnim pristupom su sljedeće:

- interpretativne metode
- kodiranje
- rekurzivno skraćivanje
- obrade podataka računalom.

Interpretativna metoda kao analiza podataka je vrlo uobičajena u kvalitativnom pristupu i ona je istraživačev / promatračev dojam. Promatrač istražuje podatke, interpretira ih stvarajući dojmove te prikazuje te dojmove u strukturiranom, čak ponekad i kvantitativnom obliku.

Kodiranje podataka je metoda obrade podataka dobivenih kvalitativnim pristupom kojom se takvi podaci istodobno organiziraju i pripremaju za neku od kvantitativnih obrada podataka. Zahtijeva da se podaci vrlo pažljivo čitaju, određuju segmenti koji se onda označavaju istim kodom (riječu, frazom, oznakom). Ta je metoda postala vrlo česta pri računalnoj obradi podataka. Postoje računalni programi koji takva kodiranja rade automatski i pripremaju za kasniju računalnu obradu. Nedostaci ove metode su da se gubi raznolikost, obilje i individualnost podataka.

Rekurzivno skraćivanje je niz postupaka sažimanja osnovnih podataka: podaci se ne kodiraju nego se iz pojedinih skupova podataka stvaraju sažeci, koji se ponovno sažimaju u nove sažetke itd. Krajnji rezultat je obično vrlo kratak sažetak u kojem se teško prepoznaju početni podaci i njihove međufaze. To je upravo i često navođeni nedostatak ove metode, gdje se gubi gotovo svaka individualnost i varijabilnost među podacima.

Računalne obrade podataka dobivenih kvalitativnim pristupom danas su uobičajene i česte. One se koriste mogućnostima suvremenih informacijskih i komunikacijskih tehnologija (ICT) za pretraživanje, sortiranje i spremanje velikih količina podataka koji su u ovom slučaju većinom riječi, rečenice ili odgovarajući znakovi. Te se metode osobito koriste u slučaju da je količina podataka prevelika za analizu i / ili trošak obrade nadmašuje korist dobivene obrade. Primjenjuju se i u slučaju da se traže točno određeni podaci – riječi, izrazi, dijelovi tekstova. Najčešće navođeni nedostatak ove metode je računalna analiza, odnosno izostanak istraživača u analizi podataka. Bez obzira na to što te sve sofisticiranije programe izrađuju vrhunski

stručnjaci za obradu podataka dobivenih kvalitativnim pristupom, sama je analiza automatizirana, što znači izostanak ljudskog elementa. Ipak takvi se programi primjenjuju sve više, dijelom kao pomoć u analizi, a dijelom se izostanak ljudskog elementa pri obradi podataka pokušava nadoknaditi vrsnom interpretacijom dobivenih rezultata od strane vrhunskih stručnjaka iz istraživnog područja.

Prosudba o ispravnosti istraživanja ovisi o mogućnosti stvaranja prosudbe o tijeku istraživanja. Istraživanja s nestandardnim kvalitativnim pristupom moraju biti detaljnije opisana u svakom koraku, a osobito o načinu prikupljanju podataka i mogućem utjecaju samog ispitivača na prikupljanje podataka [37., 38., 39., 40.].

4.2.3. OBRADA TEKSTNIH PODATAKA RAČUNALOM

4.2.3.1. ANALIZA TEKSTNIH PODATAKA

Tekstna statistika je primjena statističkih metoda u obradi tekstnih podataka. Zadnjih desetljeća 20. stoljeća tekstna je statistika doživjela niz promjena kako u samom području istraživanja tako i u svojim ciljevima i metodologiji koju koristi.

Tekstna statistika jedna je od više znanosti koja se bavi tekstom, odnosno tekstnim podacima. Neke od njih su lingvistika, analiza razgovora, analiza sadržaja, umjetna inteligencija. Sve se one u istraživanjima koriste kombinacijama lingvistike, matematike i računalnih znanosti, ali manje statističkim modelima i metodama. Uspješnom primjenom statističkih metoda u prirodnim i društvenim znanostima, slijedom toga i u područjima tih znanosti gdje se rabe tekstni podaci povećava se i zanimanje istraživača za statističku analizu tekstnih podataka, osobito pedesetih i šezdesetih godina 20. stoljeća. U svojim počecima tekstna je statistika bila uglavnom leksikometrija te prepoznavanje ključnih riječi i njihovih tekstnih okruženja.

Unaprjeđivanje informacijskih tehnologija (IT) omogućilo je stvaranje velikih baza podataka, njihovo pretraživanje i analiziranje. Tada se u statistici razvijaju multivarijatne metode statističke analize. One se sve više primjenjuju i na tekstne podatke. To je i vrijeme kada se pojavljuje i obrada prirodnog jezika (engl. *natural language processing*) koja postaje osobito zanimljiva za moguću analizu sadržaja odabranih tekstova. Objedinjavanjem svih tih dosega iz raznih područja u statističkoj analizi tekstnih podataka unapređuju se metode tekstne deskriptivne analize (unaprijeđena leksikometrija) i metode korespondencijske analize. U vrijeme prelaska iz drugog u treće tisućljeće IT su napravile poseban iskorak u prikazivanju grafike tako da su i statističke analize tekstnih podataka dobile kvalitetan grafički prikaz rezultata, lako razumljiv te blizak i istraživačima s manjim predznanjima u statistici [41., 42.].

4.2.3.2. STATISTIČKA OBRADA TEKSTA

Tekst se statistički analizira kao niz diskretnih kvalitativnih varijabli. Kolikogod se čini da takva primjena statističke analize teksta jednostavno rješava sve probleme u obradi teksta, noviji pristupi pokazuju mnoge nedostatke. Osobito je to vidljivo ako se želi u potpunosti automatizirati (statističku) analizu tekstnih podataka. Naime, postoji velika količina informacija koja se odnosi na matricu podataka koja se analizira, a te se informacije ne pojavljuju u samoj matrici. To su meta-informacije i njima obiluje svaki tekst. One mogu biti razne nejednakosti među podacima, odnosi među vrijednostima pojedinih varijabli, granične vrijednosti, ali i mnogo toga drugoga što se rabi gotovo rutinski ili intuitivno kad se radi o tekstu.

Svaka riječ je nedjeljivi dio teksta i nije ju moguće izostaviti bez posljedica za statističku tekstnu analizu. Posebnu pozornost pri analizi zahtijevaju gramatičke (neleksičke) riječi koje često zovemo prazne ili pomoćne riječi. One svojim velikim brojem i višeznačnošću opterećuju statističku analizu teksta, međutim njihovo izostavljanje mijenja smisao teksta, a time je i analiza danog teksta nepotpuna.

4.2.3.3. POSEBNI TEKSTOVI – ODGOVORI NA OTVORENA PITANJA

Statistička tekstna analiza uspješno se primjenjuje na sve vrste tekstova: književne (prozu, poeziju i dramu) i stručne iz raznih znanstvenih područja, gdje se posebno rabi u analizi odgovora na otvorena pitanja u raznim upitnicima.

Različiti upitnici za prikupljanje podataka primjenjuju se u mnogim istraživanjima u svim područjima znanosti. Zadnjih desetljeća vrlo se često koriste u socioekonomskim, poslovnim, marketinškim i političkim istraživanjima, ali i u epidemiologiji, procjeni sestrinske skrbi, zdravstvenoj skrbi općenito te javnozdravstvenim istraživanjima. U sve su većem broju to upitnici s tzv. otvorenim pitanjima u kombinaciji s pitanjima na koja su ponuđeni odgovori i / ili se prikupljaju osobni podaci ispitanika te bilo koji drugi podaci.

Otvorena pitanja su pitanja na koja ispitanik odgovara tekstem prema osobnom izboru. Odgovore na otvorena pitanja zovemo otvoreni odgovori.

Otvorene odgovore možemo dobiti kao rezultat triju vrsta upita, a mogu biti i kombinirani u istom upitniku. To su:

- neformalni upiti kroz razgovor
- tematski ciljani upiti kroz razgovor
- upiti u standardiziranim upitnicima sa slobodnim odgovorima.

Slobodni odgovori prema svojem nastanku kao samosvojno izražavanje svakog pojedinog ispitanika nose mnogo informacija. Često su podaci u tim odgovorima redundantni, osobito u slučaju velikog broja ispitanika pa su i veliki izazov, ali i veliko opterećenje za statističku analizu. To je razlog što se statistička analiza takvih podataka pokušava pojednostavniti naknadnim kodiranjem tih odgovora, što dovodi do niza pogrešaka koje su nepopravljive u daljnjoj statističkoj analizi.

Upitnici koji uključuju i otvorena pitanja imaju dodatnu vrijednost, jer zahvaljujući naprednim IT u statističkoj analizi tih upitnika možemo kombinirati sve vrste podataka. Otvoreni odgovori unose se u računalo u izvornom obliku i mogu se kombinirati s demografskim značajkama ispitanika te njihovim odgovorima na zatvorena pitanja. Takvim se postupcima otvoreni odgovori mogu kategorizirati i grupirati, a da se nimalo ne mijenjaju [43.].

4.2.3.4. POSTUPCI PRI ANALIZI TEKSTNIH PODATAKA

Analiza tekstnih podataka počinje njihovim unosom u računalo, nazvanim i kompjuterizacija teksta. Najčešće je to unos preko tipkovnice, međutim IT danas omogućuju relativno jednostavan unos velikih količina teksta.

Tekstni podaci spremljeni u računalo u sljedećem se koraku segmentiraju, što znači podjelu teksta u tekstne jedinice, tj. minimalne dijelove teksta koji se više ne mogu dijeliti. Tekstna jedinica je definirana kao skup znakova između dvaju delimitera. Takvu tekstnu jedinicu zovemo grafička forma ili, češće, riječ. Uobičajeno

je da se delimiterom za riječ smatra bjelina (praznina) kojom inače odjeljujemo riječi pri pisanju. Međutim kao delimiter može se definirati, i obično se definira, čitav skup znakova, najčešće interpunkcija.

Jedno pojavljivanje riječi je niz znakova izvan skupa delimitera koji je s obje strane ograničen delimiterom. Dvije su riječi jednake ako su to dva jednaka niza znakova izvan skupa delimitera koji su s obje strane ograničeni delimiterom.

Vokabular čine sve različite riječi u danom tekstu. Veličina ili duljina teksta je ukupni broj pojavljivanja svih riječi u tekstu.

U tekstnoj analizi nakon segmentacije teksta slijedi numeričko kodiranje teksta. To je postupak kojim se svakoj riječi pridjeljuje njezin brojčani kôd ili jednostavno broj koji se u statističkoj analizi rabi kao elementarni podatak. Ti su kodovi spremljeni u rječnik riječi koji je jedinstven za svaku pojedinu primjenu. Nakon svih provedenih statističkih analiza s pomoću tog rječnika rezultati se prikazuju s pomoću početnih riječi.

Broj pojavljivanja pojedine riječi je vrlo različit. One riječi koje se pojavljuju samo jedanput posebno su istaknute svojim imenom. To su hapaksi. Taj naziv dolazi iz starogrčkog jezika *hapax legomenon* (= jedanput izrečeno).

Uobičajeno je da su analizirani tekstovi različite duljine pa broj riječi u tekstu može biti od nekoliko stotina do više tisuća. Vokabular je strogo vezan uz osobu čije su to riječi. To znači da svaka riječ svojim pojavljivanjem ili nepojavlivanjem karakterizira autora teksta. Istraživanja tekstova vezano uz duljinu i vokabular pokazala su opću karakteristiku vezanu uz učestalost pojavljivanja riječi. Ta je karakteristika nazvana Zipfov zakon (1949.) u čast G. U. Zipfu i glasi:

Umnožak učestalosti pojavljivanja određene riječi u tekstu i njezina ranga u tablici učestalosti je konstantan.

To znači da će u svakom tekstu biti malo riječi velike učestalosti, a mnogo riječi male učestalosti. To je pravilo i uvjetovalo da se u tekstu posebno istaknu hapaksi, jer su oni najučestaliji.

Tekstna analiza nastavlja se leksikometrijom – prebrojavanjem riječi i stvaranjem indeksa cijelog teksta ili pojedinih njegovih dijelova. Indeksi mogu biti hijerarhijski (po učestalosti pojedine riječi) ili abecedni. U hijerarhijskom indeksu teksta riječi iste učestalosti poredane su prema abecedi.

Ponekad se kod proučavanja samih riječi u danom tekstu provodi i lematizacija (stvaranje lema). Lematizacija je stavljanje svih promjenljivih riječi u osnovni oblik, primjerice sve imenice u nominativ jednine, sve pridjeve u nominativ jednine muškog roda, sve glagole u infinitiv i slično. Takav postupak provodi se uglavnom kod izučavanja književnih tekstova, a u nekim jezicima je razvijena i automatska lematizacija na računalu.

Nakon leksikometrije tekstna analiza nastavlja se uglavnom u dva smjera prema izboru istraživača:

- kontekstne analize
- statističke analize kvalitativnih podataka.

Kontekstne analize su automatizirani postupci kojima se pokušava približiti sadržaju kompjuteriziranog teksta. U tu se svrhu provode postupci konkordancije s odabranom (ključnom riječi) te pronalaženje segmenata koji se ponavljaju i njihova leksikometrija.

Statističke analize najčešće korištene za obradu tekstnih podataka su:

- korespondencijska analiza
- klaster analiza.

Korespondencijska analiza jedna je od metoda analize glavnih komponenti prilagođena za analizu tablica kontingencije i binarnih tablica. Osniva se na linearnoj algebri. Korespondencijska analiza je primjenjiva i na kategoričke varijable i na

binarne varijable. Rezultat je, osim brojčanih vrijednosti, grafički prikaz (položaj točaka u ravninama projekcije) na kojemu geometrijska bliskost između elemenata retka i elemenata stupca prenosi statističku povezanost između redaka i stupaca, odnosno podataka u njima.

Klaster analiza je statistička metoda za utvrđivanje relativno homogenih grupa objekata klasificirajući pojedine jedinice analize s obzirom na njihovu sličnost, odnosno različitost prema nekim njihovim mjerenim obilježjima. U analizi tekstnih podataka ta se metoda rabi za određivanja bliskosti među elementima leksičke tablice (za retke ili za stupce) grupirajući ih, odnosno stvarajući klastere bliskosti.

Primjenjuju se dvije glavne metode klasteriranja:

- hijerarhijsko
- direktno sa zadanim brojem klastera.

Kod vrlo velikih skupova podataka koristi se kombinacija obiju metoda [43., 44., 45.].

4.2.3.5. RAČUNALNI PROGRAM ZA OBRADU TEKSTNIH PODATAKA

U ovom radu za računalnu obradu podataka korišten je program Dtm-Vic engleskog imena *Data and Text Mining – Visualization, Inference, Classification* verzija 4.2 iz 2008. godine. To je softver iz skupine tzv. *free software* programa. Namijenjen je u prvom redu studentima i znanstvenim istraživačima za statističku analizu složenih skupova podataka, koji sadržavaju i numeričke i tekstne podatke. Autor tog programa je Loudovic Lebart, a suautori su Monica Becue i André Salem.

Taj je program posljednji u nizu programa Loudovica Lebart koji se koriste metodama multivarijatne statističke analize u istraživanju uzoraka podataka. Prvi je bio 1987. godine SPAD, koji ubrzo dobiva svoju verziju SPAD.T koja je omogućila istodobnu statističku analizu i analizu tekstnih podataka. Taj je softver doživio niz verzija i prilagodbi s napretkom IT [46., 47.].

Program Dtm-Vic omogućava obradu podataka u skladu s najnovijim teorijskim spoznajama iz područja multivarijatne statističke analize te analize podataka dobivenih iz raznih upitnika koji uključuju kvantitativne i kvalitativne podatke, osobito odgovore na otvorena pitanja.

Tri su posebne prednosti:

- komplementarno korištenje metoda vizualizacije (analiza glavnih komponenata, korespondencijska analiza) i metoda klasteriranja,
- procjenjivanje metoda vizualizacije: metode ponovnog uzorkovanja
- analiza bliskosti (engl. *contiguity analysis*) i odgovarajuće metode.

Verzija 4.2 programa omogućava obradu 22500 ispitanika (redaka), 1000 varijabli (numeričkih ili kategoričkih), 100000 znakova po ispitaniku u skupu otvorenih pitanja. Unos podataka moguć je direktno ali i iz programa Microsoft Excel[®], a tekstni podaci unose se u txt formatu. Podaci iz Dtm-Vic programa mogu se prenijeti u Microsoft Excel[®], a rezultati su mogući i u formatu txt. Korištenje programom nije jednostavno, zahtijeva dosta predznanja iz teorije statističke analize, posebno analize tekstnih podataka. Međutim taj nedostatak se kompenzira velikim mogućnostima statističkih analiza svih vrsta podataka i njihovim kombinacijama, te vrlo transparentnim rezultatima, osobito grafičkim [48.].

5. REZULTATI

5.1. OSNOVNE ZNAČAJKE ISPITANIKA

Za potrebe ovog istraživanja ispitanici su razvrstani u tri dobne skupine i to: do 50 godina, od 50 do 64 godine te 65 i više godina (Tablica 5.1.). Najveći broj ispitanika starije je životne dobi, 65 i više godina.

Tablica 5.1. Razdioba ispitanika prema dobnim skupinama.

<i>Kategorija</i>	<i>Obuhvat godina</i>	<i>Broj</i>	<i>Postotak</i>
Mlađi	mlađi od 50 godina	31	26,5
Srednji	od 50 do 64 godine	33	28,2
Stariji	65 godina i stariji	53	45,3
	Ukupno	117	100

Jedno od dodatnih pitanja postavljenih u istraživanju bilo je vezano uz samoprocjenu svojeg zdravlje u usporedbi sa zdravljem svojih vršnjaka. Pitanje je bilo:

Biste li svoje zdravlje u usporedbi sa zdravljem svojih vršnjaka ocijenili kao:
ponuđeni odgovori:

znatno bolje, bolje, podjednako, lošije, znatno lošije.

U istraživanju su odgovori *znatno bolje* i *bolje* spojeni u odgovor *bolje* te odgovori *lošije* i *znatno lošije* u odgovor *lošije*.

Najveći broj ispitanika, njih 61 (52,2 %) procijenio je svoje zdravlje jednakim u usporedbi sa zdravljem svojih vršnjaka, a jednaki ih je broj (28 ili 23,9%) procijenio boljim, odnosno lošijim (Tablica 5.2.).

Tablica 5.2. Samoprocjena zdravlja ispitanika u usporedbi sa zdravljem svojih vršnjaka.

Samoprocjena svojeg zdravlja u usporedbi sa zdravljem vršnjaka	Broj	Postotak
Bolje	28	23,9
Jednako	61	52,2
Lošije	28	23,9
Ukupno	117	100

Kako bi ispitanici bili detaljnije opisani, analizirane su međusobno sljedeće njihove karakteristike: spol, dob, bračno stanje, završena škola i samoprocjena svojeg zdravlja u usporedbi sa zdravljem svojih vršnjaka. Za ocjenu značajnosti dobivenih rezultata odabrana je razina značajnosti $p = 0,05$.

Razdioba ispitanika prema dobnim skupinama ne razlikuje se statistički značajno u skupinama muškaraca i žena ($\chi^2 = 4,505$; 2 stupnja slobode, $p = 0,105$) (Tablica 5.3.).

Tablica 5.3. Razdioba ispitanika prema spolu i dobnim skupinama.

Spol	Dobna skupina			Ukupno
	mlađi	srednji	stariji	
Muškarci	9	18	25	52
Žene	22	15	28	65
Ukupno	31	33	53	117

Muškarci i žene statistički se značajno razlikuju s obzirom na bračno stanje (Fisherov egzaktni $p = 0,027$). Udio oženjenih muškaraca je veći (82,7%) od udjela udanih žena (66,2%), dok je udio rastavljenih žena (7,7%) veći od rastavljenih

muškaraca (1,9%). Zanimljivo je da među ispitanicima nema niti jednog udovca, dok je čak 7 ili 10,8% udovica (Tablica 5.4.).

Tablica 5.4. Razdioba ispitanika prema spolu i bračnom stanju.

Spol	Bračno stanje				Ukupno
	neoženjen/ neudana	oženjen/ udana	rastavljen/ rastavljena	udovac/ udovica	
Muškarci	8	43	1	0	52
Žene	10	43	5	7	65
Ukupno	18	86	6	7	117

Muškarci i žene statistički su se značajno razlikovali s obzirom na završenu školu ($\chi^2 = 7,43$; 2 stupnja slobode, $p = 0,024$). Udio muškaraca sa završenom višom / visokom školom veći je (59,6%) od udjela žena (35,4%) iste završene škole. Kod ispitanika sa završenom srednjom školom udio žena je veći (46,2%) od udjela muškaraca (32,7%). Osobito je ta razlika u udjelu statistički značajno različita kod završenog samo osnovnog obrazovanja: udio žena sa završenom samo osnovnom školom je 18,5%, dok je udio muškaraca istog završenog obrazovanja 7,7% (Tablica 5.5.).

Tablica 5.5. Razdioba ispitanika prema spolu i završenoj školi.

Spol	Završena škola			Ukupno
	viša/visoka škola	srednja škola	osnovna škola	
Muškarci	31	17	4	52
Žene	23	30	12	65
Ukupno	54	47	16	117

Muškarci i žene nisu se statistički značajno razlikovali u samoprocjeni svojeg zdravlja s obzirom na zdravlje svojih vršnjaka ($\chi^2 = 2,713$; 2 stupnja slobode, $p = 0,258$). U obje je skupine najviše onih ispitanika koji su svoje zdravlje procijenili jednakim sa zdravljem svojih vršnjaka (Tablica 5.6.).

Tablica 5.6. Razdioba samoprocjene svojeg zdravlja prema zdravlju svojih vršnjaka s obzirom na spol.

Spol	Samoprocjena svojeg zdravlja s obzirom na zdravlje vršnjaka			Ukupno
	bolje	jednako	lošije	
Muškarci	15	28	9	52
Žene	13	33	19	65
Ukupno	28	61	28	117

Ispitanici podjednako samoprocjenjuju svoje zdravlje sa zdravljem svojih vršnjaka (Fisherov egzaktni $p = 0,932$). U svim skupinama najviše je onih koji su svoje zdravlje samoprocijenili jednakim s obzirom na zdravlje svojih vršnjaka. (Tablica 5.7.).

Tablica 5.7. Razdioba samoprocjene svojeg zdravlja prema zdravlju svojih vršnjaka s obzirom na završenu školu.

Završena škola	Samoprocjena svojeg zdravlja s obzirom na zdravlje vršnjaka			Ukupno
	bolje	jednako	lošije	
Viša/visoka škola	14	26	14	54
Srednja škola	11	25	11	47
Osnovna škola	3	10	3	16
Ukupno	28	61	28	117

5.2. REZULTATI ANALIZE UPITNIKA SF-36

Svi su ispitanici ispunili upitnik SF-36. Rezultati standardizirane obrade tog upitnika daju osam dimenzija fizičkog i psihičkog zdravlja ispitanika. Nazivi pojedinih dimenzija te definicija obaju profila fizičkog i psihičkog zdravlja su u Tablici 5.8.

Tablica 5.8. Upitnik SF-36: dimenzije i profili fizičkog i psihičkog zdravlja.

<i>dimenzija</i>	<i>profil</i>
fizičko funkcioniranje (PF) ograničenje zbog fizičkih teškoća (RP) tjelesni bolovi (BP) percepcija općeg zdravlja (GH)	fizičko zdravlje (PCS)
vitalnost i energija (VT) socijalno funkcioniranje (SF) ograničenje zbog emocionalnih teškoća (RE) psihičko zdravlje (MH)	psihičko zdravlje (MCS)

Osnovni statistički podaci pojedinih dimenzija pokazuju nešto više vrijednosti dimenzija psihičkog zdravlja s obzirom na vrijednosti dimenzija fizičkog zdravlja (Tablica 5.9.).

Više vrijednosti u psihičkom zdravlju još su uočljivije u izračunu osnovnih statističkih podataka za profil fizičkog zdravlja i profil psihičkog zdravlja (Tablica 5.10.).

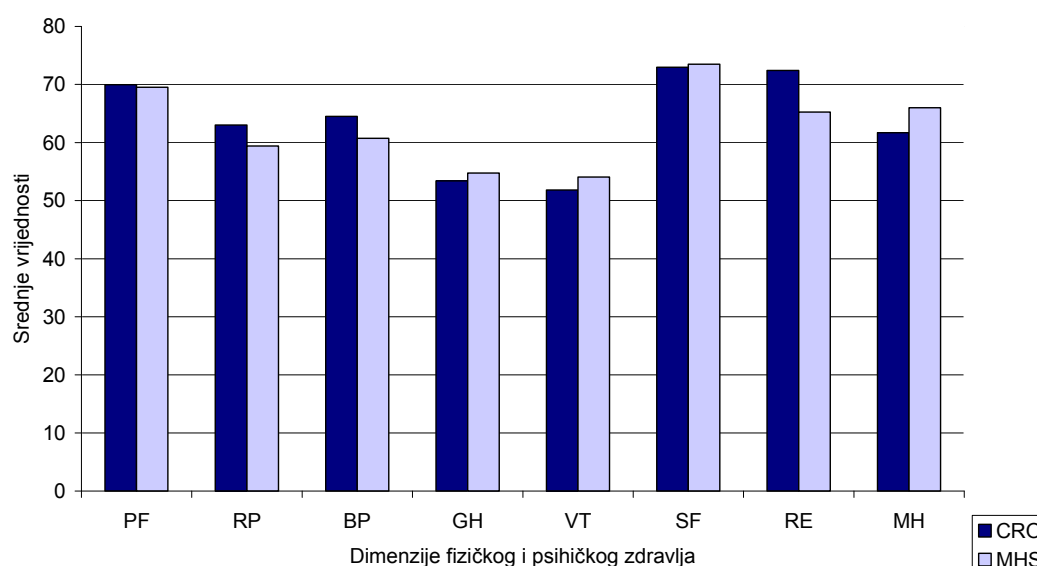
Tablica 5.9. Osnovni statistički podaci za osam dimenzija fizičkog i psihičkog zdravlja.

	<i>dimenzije fizičkog zdravlja</i>				<i>dimenzije psihičkog zdravlja</i>			
	PF	RP	BP	GH	VT	SF	RE	MH
Srednja vrijednost	69,53	59,40	60,73	54,74	54,06	73,50	65,24	65,98
Medijan	75,00	75,00	61,00	52,00	55,00	75,00	100,00	68,00
Standardna devijacija	26,06	40,57	29,16	21,72	21,39	25,22	40,68	20,18
Varijanca	678,87	1645,76	849,99	471,83	457,30	636,21	1655,07	407,07
Minimum	0	0	0	0	0	0	0	0
Maksimum	100	100	100	100	95	100	100	100

Tablica 5.10. Osnovni statistički podaci za profil fizičkog zdravlja i za profil psihičkog zdravlja.

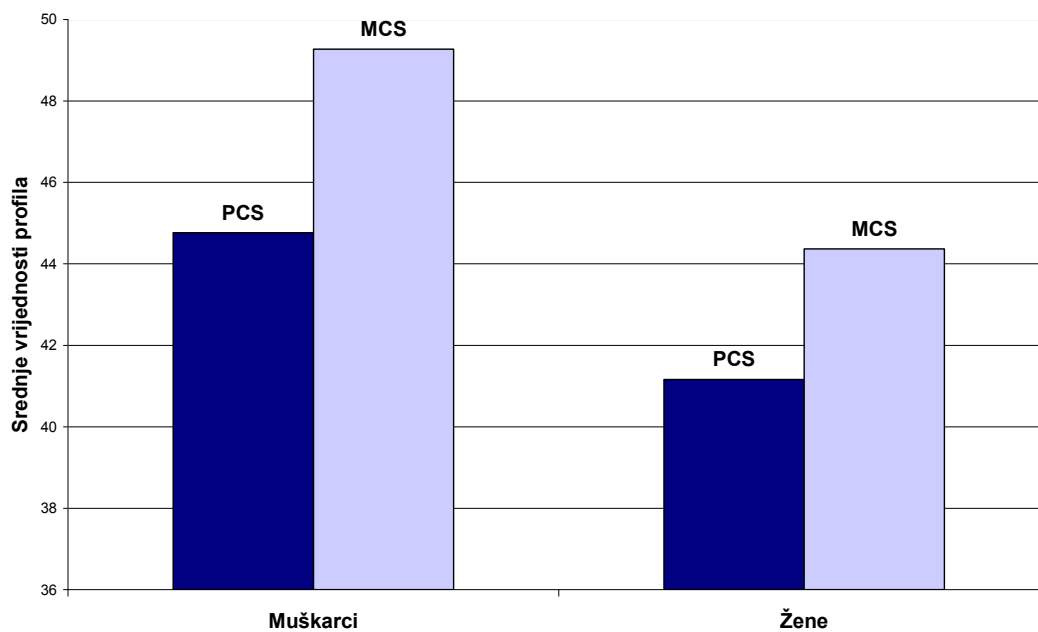
	PCS	MCS
Srednja vrijednost	42,76	46,55
Standardna devijacija	10,84	11,03
Varijanca	117,47	121,75
Raspon	47,13	47,90
Minimum	15,18	17,18
Maksimum	62,31	65,08
I. kvartil	34,71	37,10
II. kvartil	42,59	49,35
III. kvartil	51,34	55,66

Srednje vrijednosti samoprocjene zdravlja prema SF-36 za 117 ispitanika iz ovog istraživanja (MHS) uspoređene su sa srednjim vrijednostima dobivenima u istraživanju na Školi narodnog zdravlja „Andrija Štampar“ u sklopu projekta „Analiza tranzicije zdravstvenog sustava u Hrvatskoj“ u razdoblju godina 1997-1999. na reprezentativnom uzorku hrvatske populacije koji je obuhvatio oko 1% stanovnika svih regija u Hrvatskoj (CRO). Pokazuje se da srednje vrijednosti ovog istraživanja ne odstupaju od rezultata za Hrvatsku. Ispitivanje χ^2 testom pokazalo je da nema statistički značajne razlike, za sve dimenzije je $p > 0,05$ (Slika 5.1.).

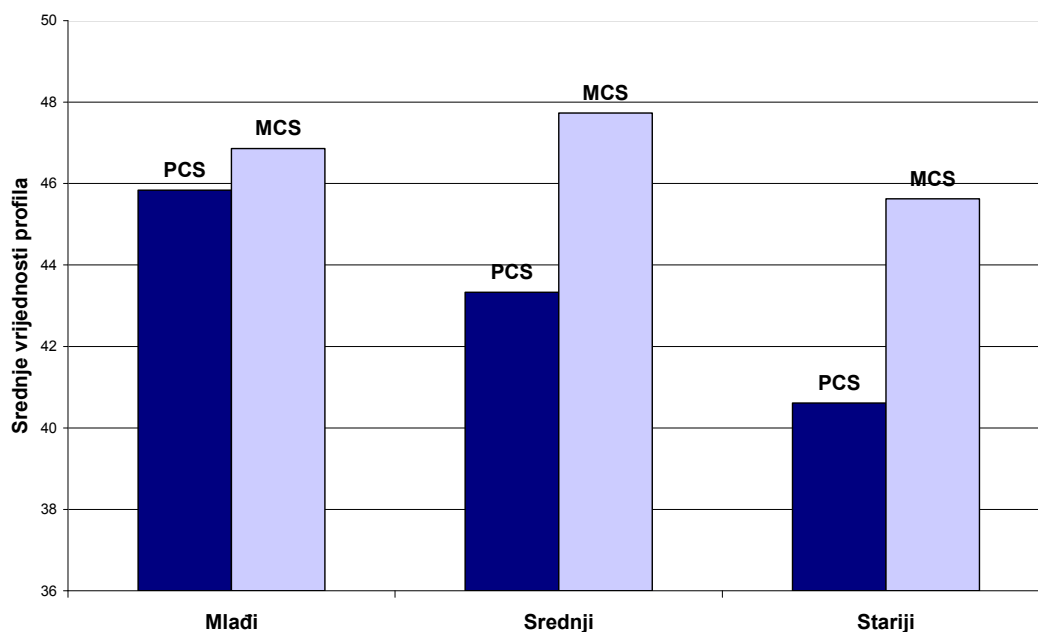


Slika 5.1. Grafički prikaz srednjih vrijednosti samoprocjene zdravlja u osam dimenzija za Hrvatsku (CRO) i za ispitanike iz ovog istraživanja (MHS).

Muškarci i žene u ovom istraživanju statistički se značajno razlikuju po profilima fizičkoga i psihičkoga zdravlja (PCS i MCS). Srednje vrijednosti oba profila više su kod muškaraca nego kod žena. Razlika je veća kod srednje vrijednosti profila psihičkog zdravlja (Slika 5.2.).

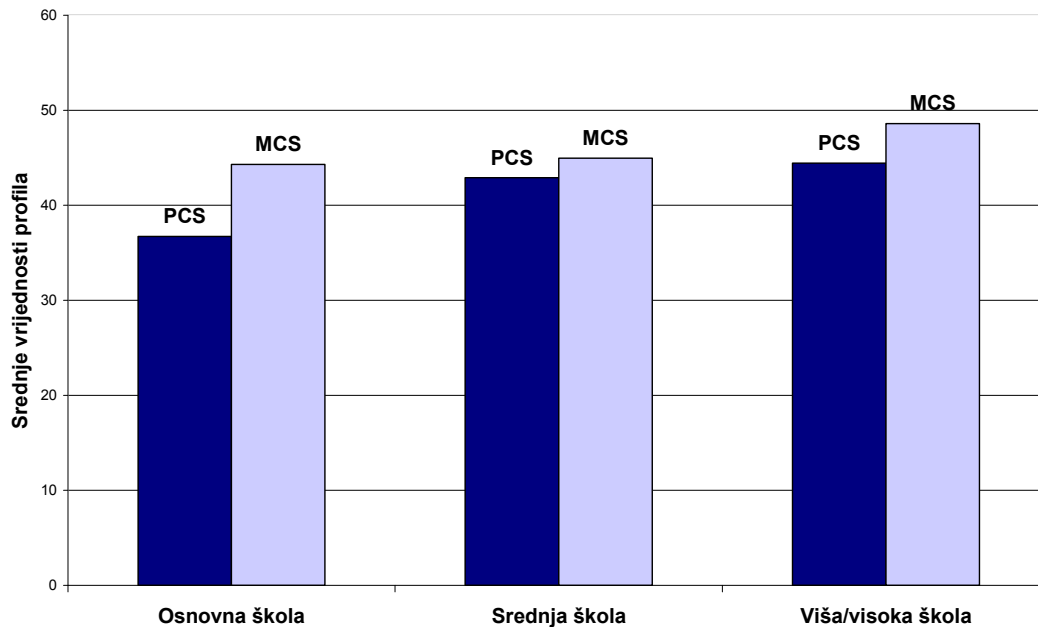


Slika 5.2. Grafički prikaz srednjih vrijednosti profila fizičkog i psihičkog zdravlja kod muškaraca i žena iz istraživanja.



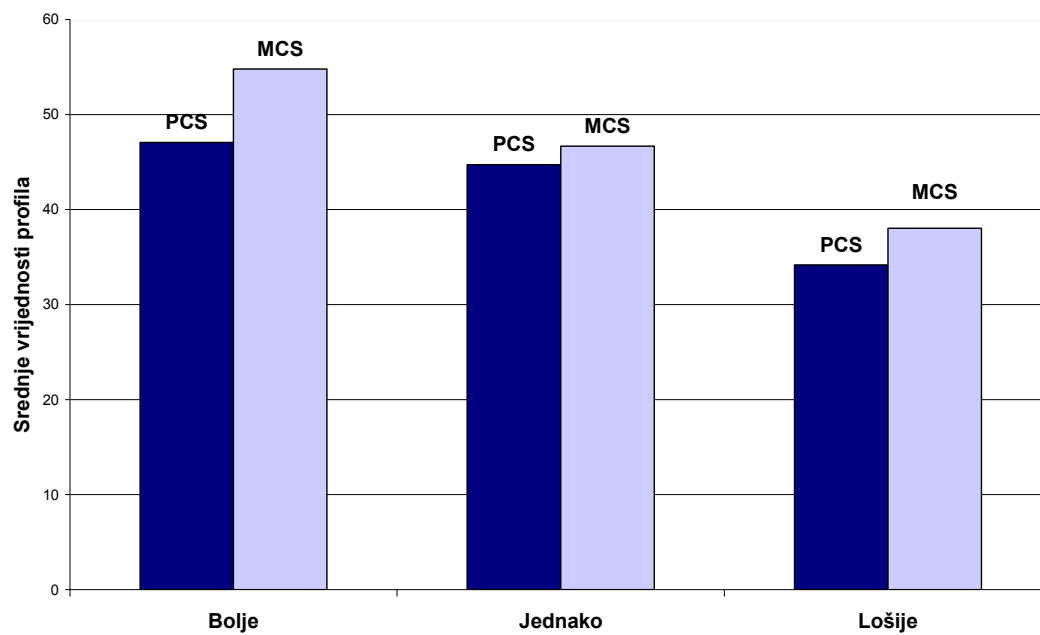
Slika 5.3. Grafički prikaz srednjih vrijednosti profila fizičkog i psihičkog zdravlja kod ispitanika različitih dobnih skupina.

Ispitanici iz ovog istraživanja sa završenom višom ili visokom školom imaju nešto više vrijednosti i profila fizičkog i profila psihičkog zdravlja (Slika 5.4.).



Slika 5.4. Grafički prikaz srednjih vrijednosti profila fizičkog i psihičkog zdravlja ispitanika s obzirom na završenu školu.

Samoprocjena svojeg zdravlja s obzirom na zdravlje svojih vršnjaka pokazala se vjerodostojnom u odnosu na profile fizičkog i psihičkog zdravlja iz Upitnika SF-36. Ispitanici koji su svoje zdravlje procijenili boljim od zdravlja svojih vršnjaka imaju i najviše vrijednosti u oba profila. Nasuprot tomu, ispitanici koji su svoje zdravlje procijenili lošijim od zdravlja svojih vršnjaka imaju vrlo niske vrijednosti i profila fizičkog zdravlja i profila psihičkog zdravlja (Slika 5.5.).



Slika 5.5. Grafički prikaz srednjih vrijednosti profila fizičkog i psihičkog zdravlja skupina ispitanika prema samoprocjeni svojeg zdravlja s obzirom na zdravlje svojih vršnjaka.

5.3. REZULTATI KVALITATIVNE ANALIZE

5.3.1. PRIKUPLJANJE PODATAKA

Tekst, odnosno tekstni podaci su zapis razgovora s ispitanicima koji je vođen u obliku tematski ciljanog ispitivanja kroz razgovor (engl. *interview guide approach*).

Postupak pri razgovoru bio je sljedeći:

- Svim je ispitanicima objašnjeno da se razgovor snima i poslije zapisuje.
- Od svih ispitanika je nakon kraćeg objašnjenja zatražen i usmeni i potpisani pristanak za sudjelovanje u istraživanju.
- Za vrijeme razgovora ispitivač nije držao diktafon u ruci nego ga je odložio, kako se ispitanik ne bi nelagodno osjećao zbog snimanja.
- Zatim bi ispitivač zamolio svojeg sugovornika da nešto kaže o svojem zdravlju, svojim zdravstvenim tegobama te ograničenjima ili neograničenjima koje ima vezano uz svoje zdravlje.
- Ako je tijekom razgovora uočena trema ili zbunjenost zbog razgovora, ispitivač je poticao razgovor dodatnim pitanjima, ali uvijek samo povezanim s danom temom – ispitanikovo zdravstveno stanje. Ako je ispitivač primijetio da se ispitanik nelagodno osjeća tijekom razgovora i da potpitanja ne pomažu, ispitivanje je prekidao.
- Razgovor je obično trajao 10-15 minuta, ali bilo je ispitanika koji su svojevóljno govorili o svojem zdravlju i više od 20 minuta.

Razgovor je sniman diktafonom s mini kasetama, zatim je „skidan“ s kasete i zapisan u računalo. Tekst je upisan u računalo putem tipkovnice jer broj ispitanika nije dopuštao upotrebu digitalnog diktafona zbog nepouzdanosti prijenosa u računalo.

Valja naglasiti da sam za vrijeme razgovora kao ispitivačica imala popriličnu empatiju za sugovornike. Nastojala sam je ne pokazati previše, iako mi je ta empatija pomogla da ispitanici prevladaju nelagodu zbog razgovora o svojem zdravstvenom stanju.

5.3.2. REZULTATI KVALITATIVNE ANALIZE:

INTERPRETATIVNA TEHNIKA

Razgovore s ispitanicima pomno sam pročitala i prema njima rasporedila 117 ispitanika u klasterne. Klasteri su nazvani prema osobnom dojmu koji sam imala nakon čitanja tih tekstova. Koji će ispitanik biti obuhvaćen kojim klasterom odredila sam prema dojmu njegove samoprocjene svojeg zdravlja. Nisam bila apsolutno objektivna jer je svako čitanje razgovora vratilo sjećanje na pojedinog ispitanika i ponovno izazvalo empatiju. Odlučila sam se za 12 klastera. Imena klastera i broj ispitanika u njima navedeni su u Tablici 5.11.

Tablica 5.11. Klasteri definirani interpretativnom tehnikom.

	<i>Ime klastera</i>	<i>Broj ispitanika</i>
M1	„ZDRAV“	19
M2	„ZADOVOLJAN“	25
M3	„UMORAN“	4
M4	„UZNEMIREN“	7
M5	„FILOZOF“	7
M6	„BOLESNI SAMAC“	7
M7	„BOLESTAN, AKTIVAN“	6
M8	„OPTEREĆEN BOLEŠĆU“	5
M9	„POMIREN S BOLEŠĆU I / ILI ŽIVOTOM“	18
M10	„BOLESTAN, OPTIMIST“	10
M11	„BOLESTAN, PESIMIST“	3
M12	„NEDEFINIRAN“	6

Imena klastera jasno definiraju i moj dojam o pojedinom ispitaniku iz tog klastera. Napomena uz klaster M5: ‘Filozof’. U njemu su obuhvaćeni ispitanici koji su najmanje

govorili o svojem zdravstvenom statusu, a mnogo su više govorili općenito o zdravlju, bolestima, mogućnostima raznih vrsta liječenja i slično.

Ispitanici u svakom klasteru opisani su svojim osnovnim značajkama obuhvaćenim ovim istraživanjem (dob, spol, bračno stanje, završena škola i samoprocjena svojeg zdravlja s obzirom na zdravlje svojih vršnjaka) te dobivenim vrijednostima samoprocjene zdravlja iz Upitnika SF-36.

Svaki je klaster karakteriziran brojem ispitanika s istom prije navedenom značajkom (Tablica 5.12.). Na taj način veći broj ispitanika s istom značajkom opisuju pojedini klaster. Tako primjerice klaster M1 prema dobi karakteriziraju mlađi ispitanici, prema spolu žene, prema bračnom stanju ispitanici koji su u braku, prema završenoj školi oni ispitanici koji su završili višu/visoku školu te oni koji su svoje zdravlje procijenili jednakim s obzirom na zdravlje svojih vršnjaka.

Tablica 5.12. Opis klastera s pomoću broja ispitanika istih značajki.

Klaster	broj isp.	dob			spol		bračno stanje				završena škola			samoprocjena zdravlja		
		mlad	sred	star	m	z	neoz	brak	rast	udov	osn	sred.šk	vis.šk	bolje	jedn	lošije
M1	19	12	5	2	8	11	7	12	0	0	1	8	10	4	13	2
M2	25	4	8	13	9	16	2	17	2	4	1	11	13	8	16	1
M3	4	3	1	0	1	3	0	4	0	0	0	4	0	1	1	2
M4	7	3	2	2	3	4	1	4	1	1	1	5	1	3	3	1
M5	7	3	1	3	5	2	1	6	0	0	1	1	5	3	2	2
M6	7	0	1	6	1	6	5	0	1	1	3	1	3	1	3	3
M7	6	0	1	5	5	1	1	5	0	0	2	1	3	1	3	2
M8	5	2	0	3	0	5	0	3	2	0	1	1	3	0	0	5
M9	18	0	7	11	11	7	1	17	0	0	4	6	8	4	10	4
M10	10	2	2	6	8	2	1	8	0	1	1	1	8	5	2	3
M11	3	0	2	1	1	2	0	3	0	0	1	1	1	0	0	3
M12	6	1	3	2	0	6	0	6	0	0	0	5	1	0	5	1
	117	30	33	54	52	65	19	85	6	7	16	45	56	30	58	29

Od vrijednosti dobivenih iz Upitnika SF-36 u opisu klastera korišteni su profil fizičkog zdravlja (PCS) i profil psihičkog zdravlja (MCS). Za svaki je klaster izračunana srednja vrijednost obaju profila svih ispitanika. Srednje vrijednosti tih profila označene su PCS_{kv} i MCS_{kv} (Tablica 5.13.).

Tablica 5.13. Srednje vrijednosti profila fizičkog i psihičkog zdravlja ispitanika u pojedinim klasterima.

<i>Klaster</i>	<i>Broj isp.</i>	<i>PCS_{kv}</i>	<i>MCS_{kv}</i>
M1	19	52,40	50,43
M2	25	46,75	51,54
M3	4	37,85	36,91
M4	7	45,84	41,88
M5	7	48,52	51,32
M6	7	36,89	36,66
M7	6	37,23	44,90
M8	5	34,86	33,14
M9	18	40,59	49,27
M10	10	39,09	54,11
M11	3	21,68	22,14
M12	6	34,15	41,94
	117		

Sljedeći je postupak bio opisati što detaljnije klasterne definirane interpretativnom tehnikom kao jednom od metoda kvalitativne analize. Iz Tablice 5.12. odabrane su za svaki klaster tri (3) kategorije s najvećom učestalošću među svim kategorijama svih osnovnih značajki. Tim su karakteristikama dodane i srednje vrijednosti obaju profila zdravlja svih ispitanika u pojedinom klasteru. Time je svaki klaster opisan brojem ispitanika, trima kategorijama osnovnih značajki obuhvaćenih ovim istraživanjima, profilima fizičkog i psihičkog zdravlja te osobnim dojmom ispitivača (Tablica 5.14.).

Tablica 5.14. Opis klastera s pomoću broja ispitanika, najučestalijih kategorija, srednjih vrijednosti oba profila te dojma ispitivača.

	<i>Klaster</i>	<i>Broj isp.</i>	<i>Kat I.</i>	<i>Kat II.</i>	<i>Kat III.</i>	<i>PCS_{kv}</i>	<i>MCS_{kv}</i>
M1	„ZDRAV“	19	jedn	mlad	brak	52,40	50,43
M2	„ZADOVOLJAN“	25	brak	z	jedn	46,75	51,54
M3	„UMORAN“	4	brak	sred.šk.	mlad	37,85	36,91
M4	„UZNEMIREN“	7	sred.šk.	z	brak	45,84	41,88
M5	„FILOZOF“	7	brak	m	vis.šk.	48,52	51,32
M6	„BOLESNI SAMAC	7	star	z	neoz	36,89	36,66
M7	„BOLESTAN, AKTIVAN“	6	star	m	brak	37,23	44,90
M8	„OPTEREĆEN BOLEŠĆU“	5	z	lošije		34,86	33,14
M9	„POMIREN S BOLEŠĆU I / ILI ŽIVOTOM“	18	brak	star	m	40,59	49,27
M10	„BOLESTAN, OPTIMIST“	10	m	brak	vis.šk.	39,09	54,11
M11	„BOLESTAN, PESIMIST“	3	brak	lošije		21,68	22,14
M12	„NEDEFINIRAN“	6	z	brak	sred.šk.	34,15	41,94

5.3.3. REZULTATI KVALITATIVNE ANALIZE:

RAČUNALNA TEHNIKA – ANALIZA TEKSTNIH PODATAKA

5.3.3.1. KOMPJUTERIZACIJA TEKSTA

Cjelokupni razgovori s ispitanicima uneseni su u računalo putem tipkovnice u txt formatu u obliku u kojem ga program Dtm-Vic prepoznaje i može dalje obrađivati. Valja napomenuti da je tekst zapisan u gotovo izvornom obliku.

Minimalni zahvati u izvorni tekst bili su sljedeći:

- Riječi su napisane bez dijakritičkih znakova.
- Označeni su glagoli u negaciji. Razlog je tomu što se smisao izgovorenog teksta u analizi mijenja ako se bilježi kao afirmativni oblik glagola i riječ ne.
- Prijedlog *s* ili *sa* zapisan je uvijek kao *sa*. Stajalište je da taj zahtjev ne mijenja smisao izgovorenog teksta za statističku analizu.
- Riječ *kad* ili *kada* zapisana je kao *kada*. Razlog je opet da to ne mijenja smisao teksta za statističku analizu.
- Isto tako i riječ *sad* ili *sada*.
- Kondicional glagola zapisan je uvijek u gramatički pravilnom obliku (bih, bismo, biste, ...) kako bi se zadržala različitost u glagolskim oblicima.
- Dosljedno je zapisivano *jedanput*.
- Brojevi do 10 zapisivani su riječima, a veći brojkom. Brojevi do 10 bili su češći u razgovorima. Za kalendarske godine su uvijek pisane samo zadnje dvije znamenke, a da bi se razlikovale od nekih drugih brojeva neposredno prije njih stavljen je znak &.
- Neke su riječi pisane uvijek istom kraticom, bez obzira na padež, npr. tv je i televizija i televizor te svi padeži tih riječi.
- Napravljene su još neke minorne promjene pri unosu, ali one nisu bitne za daljnju tekstnu analizu.
- Drugi zahvati u tekst nisu napravljeni kako bi se što je moguće više zadržao izvorni tekst odgovora ispitanika.

Razgovor s ispitanikom vođen je kao tematski ciljani razgovor. U obradi tekstnih podataka razgovor s pojedinim ispitanikom obrađen je kao cjelina, što znači kao jedan odgovor. Nisu zapisane nikakve intervencije ispitivača.

5.3.3.2. REZULTATI LEKSIKOMETRIJSKE ANALIZE

Rezultati leksikometrijske analize u prvom su redu rezultati koji se odnose na broj riječi te njihovu učestalost, što znači na ukupan broj odgovora, ukupan broj riječi, broj različitih riječi i njihov postotak u odnosu na ukupan broj riječi (Tablica 5.15.).

Tablica 5.15. Rezultati leksikometrijske analize vezani uz broj riječi.

Ukupan broj odgovora	117
Ukupan broj riječi	32356
Broj različitih riječi	5237
Postotak različitih riječi	16,2 %

Rezultat leksikometrije su i indeksi korištenih riječi. Za potrebe ove analize zanimljiv je bio hijerarhijski indeks, indeks dobiven prema učestalosti pojavljivanja korištenih riječi. U Tablici 5.16. dio je tog indeksa i to samo riječi čija je učestalost 80 i veća. I iz tog malog dijela hijerarhijskog indeksa vidljivo je da vrijedi tzv. Zipfov zakon: Produkt ranga riječi u hijerarhijskom indeksu i njihove učestalosti u nekom tekstu približno je konstantan, tj. imamo mali broj riječi velike učestalosti, a veliki broj riječi male učestalosti.

Tablica 5.16. Najučestalije riječi (učestalost ≥ 80).

<i>rbroj</i>	<i>riječ</i>	<i>učestalost</i>
1	sam	1129
2	i	929
3	se	877
4	je	804
5	da	693
6	ja	654
7	to	648
8	u	596
9	na	442
10	mi	429
11	me	336
12	sa	326
13	imam	302
14	ali	277
15	sto	273
16	onda	252
17	kada	230
18	sve	220
19	sada	206
20	tako	206
21	pa	202
22	su	191
23	malo	176
24	od	173
25	nemam	172
26	nisam	165
27	za	163
28	jer	147
29	vise	147
30	ne_mogu	128
31	dobro	127
32	nista	121
33	puno	116
34	jako	111
35	mogu	108
36	a	107
37	tu	107
38	prije	106
39	nije	105
40	bila	101
41	bio	98
42	problema	97
43	idem	96
44	vrijeme	95
45	mislim	94
46	toga	94
47	nesto	93
48	jos	89
49	bilo	88
50	dosta	88
51	osjecam	87
52	po	85
53	volim	85
54	kao	84
55	ili	83
56	godine	82
57	radim	80

Tekstovi, odnosno zapisi razgovora, koji nisu mijenjani tijekom upisa u računalo vrlo su zanimljivi i za analizu duljine korištenih riječi. U ovom istraživanju ispitanici su najviše rabili riječi od 6 i 7 slova. Općenito, najveće su učestalosti riječi od 4 do 10 slova (Tablica 5.17.).

Tablica 5.17. Razdioba učestalost riječi obzirom na broj slova.

Broj slova	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Učestalost	5	74	164	448	769	916	913	749	527	334	179	127	56	30	20	7

5.3.3.3. REZULTATI KORESPONDENCIJSKE I KLASTER ANALIZE

Program Dtm-Vic automatski odbacuje hapakse (samo jednom izrečene riječi) iz korespondencijske i klaster analize. U ovim tekstnim podacima bilo je 2966 hapaksa, tako da je ukupan broj riječi s kojima je nastavljena analiza bio 29390.

Korespondencijska analiza obuhvatila je 2271 različitih riječi koje je izgovorilo 117 ispitanika, dakle matricu s 2271 redaka i 117 stupaca.

Program Dtm-Vic automatski je odabrao 12 prvih svojstvenih vrijednosti (Tablica 5.18.). U tablici je i postotak i kumulativni postotak varijance.

Trag (zbroy svih svojstvenih vrijednosti) je 10,43.

Tablica 5.18. Svojstvene vrijednosti.

<i>svojstvena vrijednost</i>	<i>postotak</i>	<i>kumulativni postotak</i>
1	0,1903	1,83
2	0,1701	3,46
3	0,1507	4,90
4	0,1468	6,31
5	0,1429	7,68
6	0,1384	9,01
7	0,1358	10,31
8	0,1342	11,60
9	0,1329	12,87
10	0,1306	14,12
11	0,1304	15,37
12	0,1285	16,61

Kao rezultati korespondencijske analize dobivene su i koordinatne osi prostora definiranog riječima (2271 riječi) i ispitanicima (117 ispitanika). Svaka je os opisana riječima i ispitanicima koji se nalaze što dalje od ishodišta (0) na pojedinoj osi u obje orijentacije, pozitivnoj (predznaka +) i negativnoj (predznaka -).

U svim grafičkim prikazima u vezi s tekstnom analizom prikazani su rezultati samo prve koordinatne ravnine (opisane prvom i drugom osi). Tako su tu i rezultati samo u vezi s prvima dvjema osima. Prikazano je po 20 krajnjih točaka riječi i njihovih koordinata te ispitanika i njihovih koordinata na objema orijentacijama svake osi (Tablica 5.19. i Tablica 5.20.).

Tablica 5.19. Krajnje točke prve osi.

	<i>riječ i koordinata</i>		<i>ispitanik i koordinata</i>	
P R V A O S	ja	- 8,925	T034	- 0,828
	onda	- 4,956	T009	- 0,781
	mene	- 4,393	P006	- 0,726
	bila	- 4,165	T006	- 0,719
	do	- 3,031	S014	- 0,702
	tu	- 3,016	P002	- 0,696
	živac	- 2,909	T025	- 0,688
	su	- 2,864	T005	- 0,686
	kicma	- 2,855	J002	- 0,657
	dosla	- 2,852	T010	- 0,642
	radila	- 2,851	T004	- 0,618
	kada	- 2,762	S025	- 0,608
	može	- 2,759	T008	- 0,587
	dodje	- 2,720	S015	- 0,585
	znala	- 2,710	P001	- 0,580
	opet	- 2,649	P005	- 0,558
	cu	- 2,646	S024	- 0,547
	samo	- 2,637	E007	- 0,521
	da	- 2,614	T012	- 0,463
	mogla	- 2,594	P003	-0,463

	ne_pase	6,039	J017	0,692
	nekih	6,146	J015	0,821
	skladu	6,193	M007	0,856
	vino	6,252	J009	0,868
	vrijeme	6,383	M001	0,894
	zestoka	6,414	J018	1,038
	alkohola	6,415	J014	1,055
	racunalom	6,592	J012	1,059
	sjedim	7,355	M005	1,071
	ne_pusim	7,558	J013	1,089
	nemam	7,582	M010	1,148
ledja	7,856	M002	1,152	
pazim	7,893	M012	1,284	
cesto	7,904	M003	1,343	
virtu	8,027	M008	1,452	
kolicinama	8,316	J019	1,456	
obaveza	9,474	E013	1,495	
ponekad	10,612	M006	1,779	
zbog	10,612	M009	2,038	
problema	11,086	M011	2,096	

Tablica 5.20. Krajnje točke druge osi.

	<i>riječ i koordinata</i>		<i>ispitanik i koordinata</i>	
	me	- 8,402	E007	- 1,145
	kicma	- 5,846	M008	- 0,924
	boli	- 5,700	P005	- 0,903
	biti	- 5,563	M009	- 0,755
	daj	- 5,395	P006	- 0,616
	pa	- 5,205	T013	- 0,611
	uhvati	- 5,061	E013	- 0,597
	kada	- 4,958	T001	- 0,581
	ce	- 4,864	J005	- 0,574
	pati	- 4,809	J007	- 0,549
	zdravlje	- 4,803	T035	- 0,546
	ako	- 4,744	P001	- 0,473
	ponekad	- 4,575	P003	- 0,454
	ni	- 4,512	T012	- 0,443
	ne_znam	- 4,414	T015	- 0,421
D	ne_razmislijam	- 4,403	T025	- 0,416
R	delati	- 4,230	T010	- 0,412
U	vrata	- 4,230	T009	- 0,380
G	volje	- 4,099	J013	- 0,365
A	nemam	- 4,003	M002	-0,346

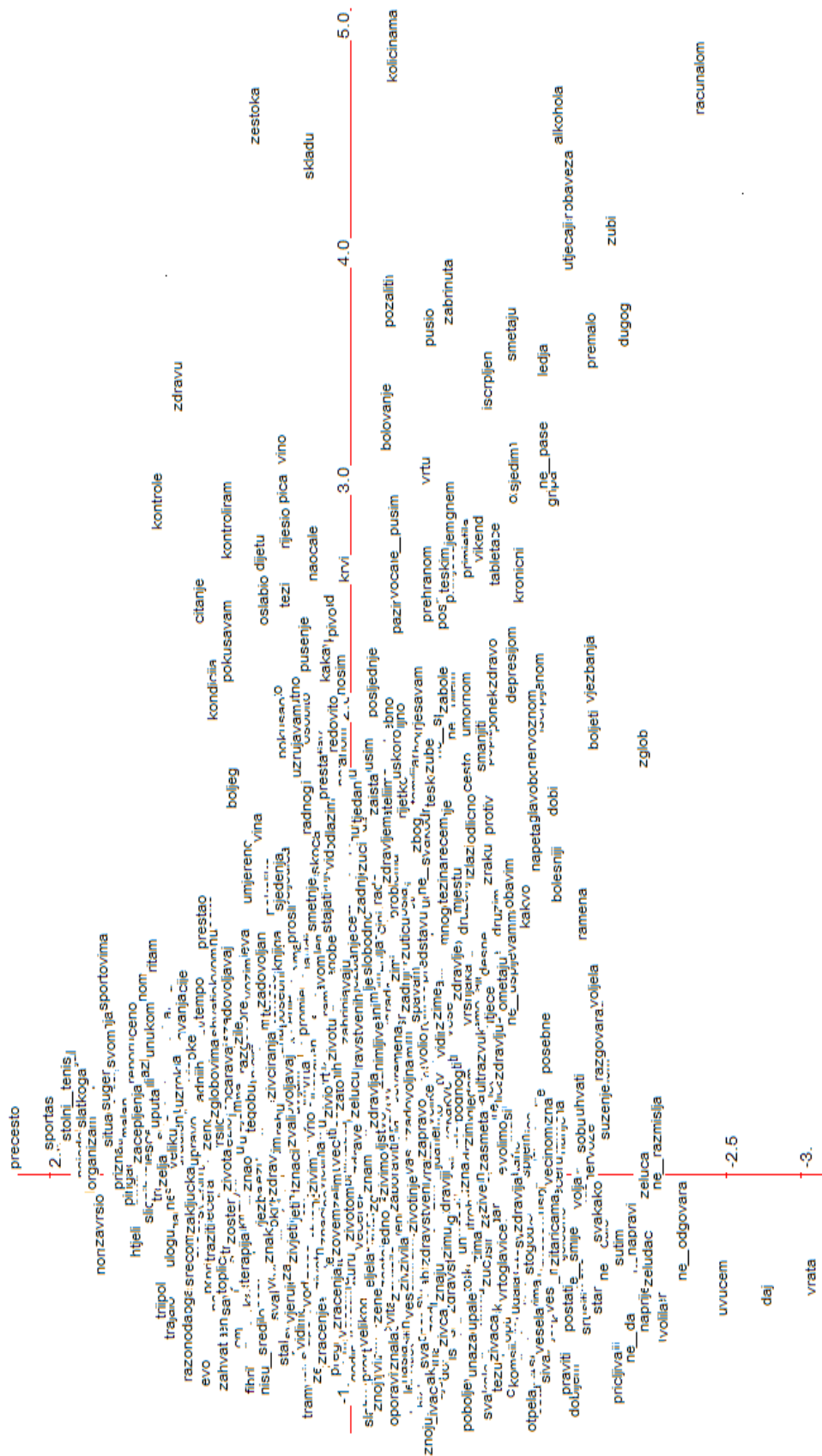
O	normalan	3,828	S011	0,414
S	bito	3,854	M011	0,439
	bavio	3,859	S027	0,439
	godina	3,905	T002	0,445
	stolni_tenis	3,918	S007	0,451
	opsezan	3,993	S038	0,452
	potpuno	4,088	S014	0,467
	situacija	4,117	S039	0,484
	u	4,185	S020	0,504
	promijenio	4,198	S033	0,522
	prije	4,208	S023	0,551
	infarkt	4,305	T005	0,562
	zivot	4,630	S037	0,577
	nacin	4,714	S021	0,670
	mogao	4,769	S034	0,723
	imao	4,795	S004	0,741
	je	5,157	S031	0,806
	igrati	5,401	S009	0,849
	sam	7,020	S022	0,856
	bio	7,353	S003	1,040

Program Dtm-Vic kao rezultat korespondencijske analize daje i grafički prikaz (položaj točaka u ravninama projekcije) na kojemu geometrijska bliskost između elemenata retka i elemenata stupca prenosi statističku povezanost između redaka i stupaca, odnosno podataka u njima.

Zbog vrlo velikog broja varijabli u recima polazne matrice, tj. 2271 riječi, grafički prikaz rezultata korespondencijske analize po recima ograničen je na samo središnji dio prve koordinatne ravnine. Tako je prva os prikazana u rasponu od -1 do +5, a druga os od -3 do +2. Riječi koje imaju najveću frekvenciju distribuirane su upravo oko sjecišta prve i druge osi te je na tom mjestu najveća gustoća, a time i preklapanje mnogih riječi (Slika 5.6.).

Rezultat korespondencijske analize po stupcima, prva koordinatna ravnina, u potpunosti je grafički prikazan. Razdioba ispitanika u prvoj koordinatnoj ravnini pokazuje da postoji skupina ispitanika bliska prema svojim odgovorima i koja se nalazi oko ishodišta prve koordinatne ravnine. Međutim, neki ispitanici se svojim vokabularom dosta razlikuju od te skupine pa su na grafičkom prikazu udaljeni od ishodišta. Neki među njima vrlo su bliski s obzirom na distribuciju na prvoj osi, ali su vokabularom različiti s obzirom na distribuciju po drugoj osi (Slika 5.7.).

Slika 5.6. Vizualizacija rezultata korepondencijske analize (prva i druga os) po recima (riječima).

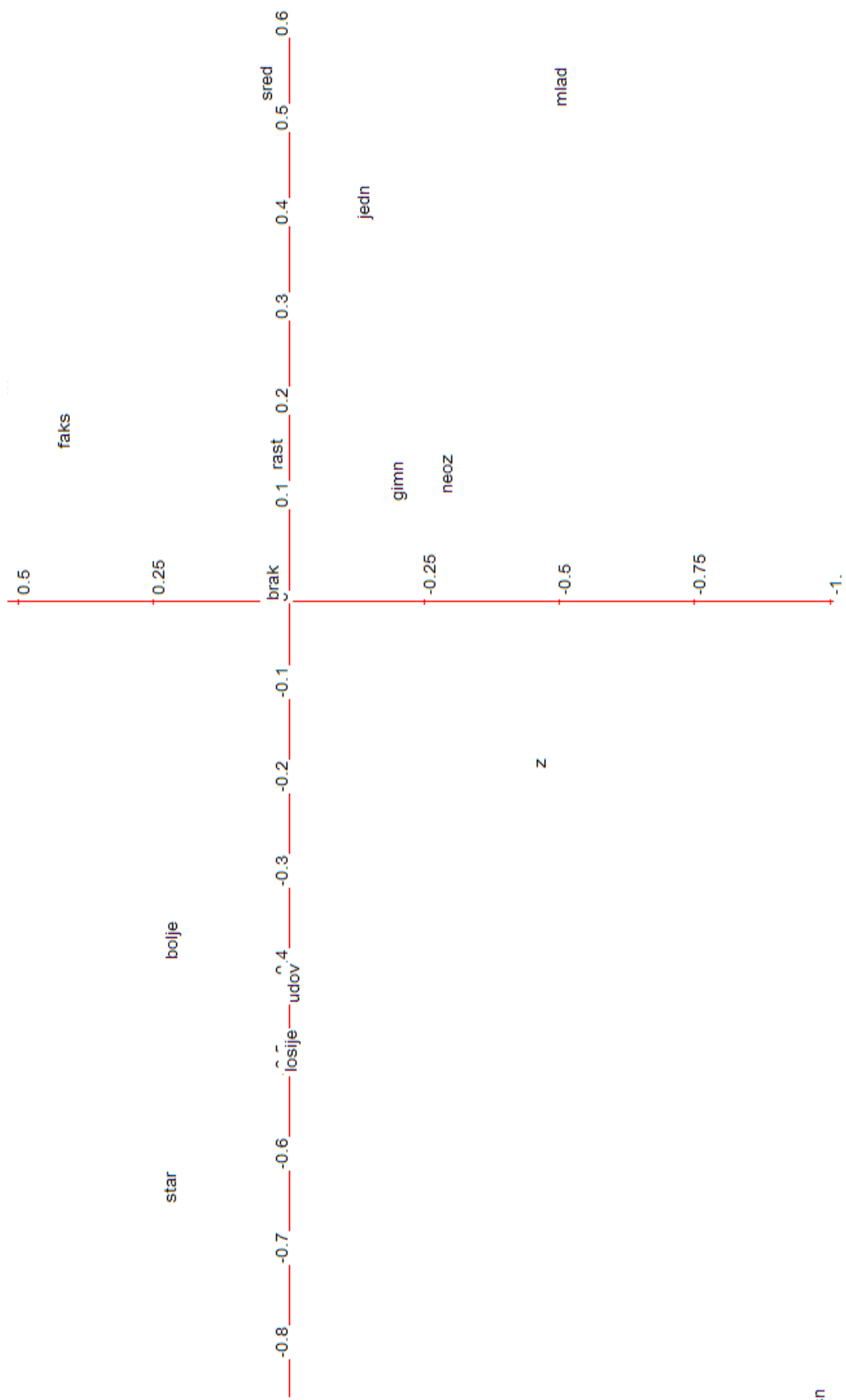


Vizualizacija rezultata korespondencijske analize daje i razdiobu osnovnih značajki ispitanika analiziranih u ovom istraživanju u prvoj koordinatnoj ravnini prema korespondencijskoj analizi. Razdioba pokazuje različitost ispitanika s obzirom na značajke koja proizlazi iz njihova vokabulara. Osobito je vidljiva različitost muškaraca i žena, ispitanika sa završenih osnovnom školom i ispitanika sa završenom višom/visokom školom, starijih i mlađih ispitanika te ispitanika koji samoprocjenjuju svoje zdravlje boljim ili lošijim u usporedbi sa zdravljem svojih vršnjaka od onih koji ga procjenjuju boljim. Oni se razlikuju u razdiobi u prvoj koordinatnoj ravnini i po prvoj i po drugoj osi. Ispitanici sa završenom srednjom školom i višom/visokom školom ne razlikuju se bitno s obzirom na prvu os, ali su na suprotnim orijentacijama druge osi. Ispitanici mlađe i srednje dobi po svom su vokabularu bliski s obzirom na prvu os. Zanimljiva je raspodjela s obzirom na bračno stanje: udovci/udovice, oženjeni/udane i rastavljeni/rastavljene razlikuju se međusobno svojim vokabularom samo s obzirom na prvu os. Neoženjeni/neudane svojim se vokabularom razlikuju od njih prema drugoj osi, a prema prvoj osi su blizu rastavljenima (Slika 5.8.). Oznake značajki ispitanika na Slici 5.8. nalaze se u Tablici 5.21.

Tablica 5.21. Oznake značajki ispitanika korištene na Slici 5.8.

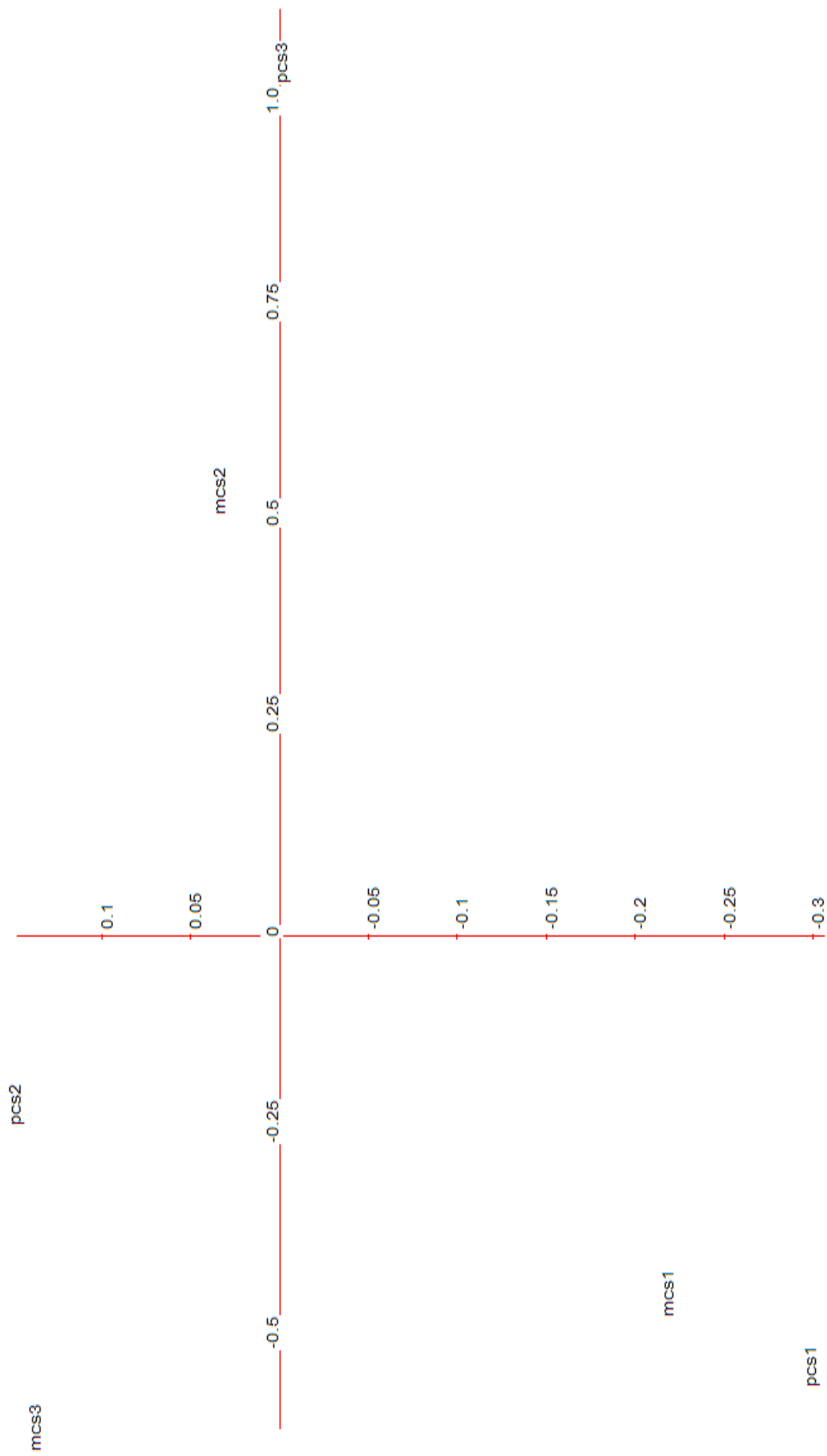
Značajka ispitanika na Slici 5.8	Opis
m	Muškarci
z	Žene
mlad	Ispitanici mlađi od 50 godina
sred	Ispitanici od 50 do 64 godine
star	Ispitanici stariji od 65 godina
neoz	Neoženjeni/neudane
brak	Oženjeni/udane
rast	Rastavljeni/rastavljane
udov	Udovci/udovice
osn	Završena osnovna škola
gimn	Završena srednja škola
faks	Završena viša/visoka škola
bolje	Samoprocjena zdravlja boljim
jedn	Samoprocjena zdravlja jednakim
losije	Samoprocjena zdravlja lošijim

Slika 5.8. Vizualizacija korespondencijske analize (prva i druga os): značajke ispitanika ispitanika.



Vizualizacija rezultata korespondencijske analize daje i razdiobu vrijednosti profila fizičkog zdravlja (PCS) i profila psihičkog zdravlja (MCS) dobivenih analizom Upitnika SF-36 u prvoj koordinatnoj ravnini iz korespondencijske analize. U korespondencijsku analizu uvršteni su kvartili pojedinih profila. Razdioba pokazuje različitost ispitanika s obzirom na kvartile pojedinih profila koja proizlazi iz njihovog vokabulara. Prvi i drugi kvartil fizičkog profila zdravlja ispitanika iz ovog istraživanja su na istoj orijentaciji prve osi, dok je treći kvartil na suprotnoj orijentaciji. S obzirom na drugu os prvi i drugi kvartil fizičkog profila su na suprotnim orijentacijama te osi, a treći je između njih. Kod psihičkog profila zdravlja prvi i treći kvartil su na istoj orijentaciji prve osi, a drugi je na suprotnoj. Prvi i treći kvartil psihičkog profila zdravlja su na suprotnim orijentacijama druge osi, a drugi je između njih s obzirom na drugu os (Slika 5.9.). Na slici su sljedeće oznake: pcs1, pcs2, pcs3 su prvi, drugi, treći kvartil profila fizičkog zdravlja, mcs1, mcs2, mcs3 su prvi, drugi, treći kvartil profila psihičkog zdravlja.

Slika 5.9. Vizualizacija korespondencijske analize (prva i druga os): kvartili profila fizičkog i psihičkog zdravlja



Sljedeći postupci u statističkoj analizi tekstnih podataka bili su postupci klasteriranja. Program Dtm-Vic omogućava kombinaciju dviju metoda klasteriranja: hijerarhijskog i direktnog s unaprijed zadanim brojem klastera. Program prema veličini matrice sam određuje parametre klasteriranja, ali dopušta i neke promjene. U ovoj analizi jedina je promjena bila izbor broja klastera, odabrano je 12 klastera.

Prema vokabularu ispitanika program je rasporedio ispitanike u pojedine klasterne, K1 – K12. Klasteri su prema broju ispitanika različitih veličina, od klastera K10 sa samo jednim ispitanikom do klastera K5 koji je najbrojniji s čak 28 ispitanika (Tablica 5.22.).

Tablica 5.22. Klasteri dobiveni statističkom analizom tekstnih podataka.

<i>Ime klastera</i>	<i>Broj ispitanika</i>
K1	19
K2	6
K3	2
K4	8
K5	28
K6	11
K7	11
K8	4
K9	10
K10	1
K11	13
K12	4

Kategorije značajki ispitanika koje obilježavaju pojedine klasterne prikazane su u Tablici 5.23. Njih je program Dtm-Vic izračunao prema učestalosti riječi u vokabularu ispitanika u pojedinom klasteru.

Tablica 5.23. Karakteristične kategorije značajki ispitanika za pojedine klustere.

Klaster	broj isp.	značajka	kategorija
K1	19	spol završena škola	žene osnovna
K2	6	završena škola	osnovna
K3	2	završena škola	osnovna
K4	8	-	-
K5	28	-	-
K6	11	spol brač. stanje	muškarci oženjen/udata
K7	11	završena škola spol	viša/visoka. muškarci
K8	4	-	-
K9	10	dob	mlađa
K10	1	-	-
K11	13	dob	srednja
K12	4	-	-
	117		

Program Dtm-Vic odredio je 10 relativno najučestalijih riječi ispitanika u svakom pojedinom klasteru (Tablica 5.24.).

Tablica 5.24. Deset relativno najučestalijih riječi u svakom klasteru.

	10 najučestalijih riječi									
K1	tesko	skroz	rucak	zadovoljstvo	imam	boli	me	biti	obaveze	si
K2	nekad	ducana	moze	kada	djeca	onda	ovako	nikuda	noc	pa
K3	me	pati	daj	ni	kicmu	poslu	uhvati	ne_razmisljam	smeta	uvucem
K4	ja	bila	su	zivac	meni	znala	mene	ca	sam	dobila
K5	more	jako	odrzavam	stigla	stres	ok	preglede	povisen	koljenima	unuka
K6	imao	poteskoce	svjestan	ne_smeta	dob	bavio	prije	fizicke	tjedno	ne_osjecam
K7	igrati	mogu	bio	situacija	mogao	u	priznati	igramo	doslo	zivot
K8	zdravlje	kakvo	biti	ne_druzim	buducnosti	ce	lose	ponekad	ne_izlazim	tv
K9	imao	poteskoce	svjestan	ne_smeta	dob	bavio	prije	fizicke	tjedno	ne_osjecam
K10	redovito	kontroliram	kontrola	zestoka	dijetu	voca	skladu	krvi	zdravu	zube
K11	zbog	problema	ne_pusim	cesto	pazim	dobro	ponekad	bolovanje	vrtu	ne_pijem
K12	problema	obveza	racunalom	zubi	ponekad	zbog	gripa	prehranom	ne_pase	ne_stignem

Klasteri dobiveni statističkom analizom opisani su brojem ispitanika u kategorijama značajki korištenih u ovom istraživanju: dob, spol, bračno stanje, završena škola i samoprocjena svojeg zdravlja s obzirom na zdravlje svojih vršnjaka (Tablica 5.25.).

Tablica 5.25. Opis klastera programa Dtm-Vic s pomoću broja ispitanika istih značajki.

Klaster	broj isp.	dob			spol		bračno stanje				završena škola			samoprocjena zdravlja		
		mlad	sred	star	m	z	neoz	brak	rast	udov	osn	sred.šk	vis.šk	bolje	jedn	lošije
K1	19	8	5	6	3	16	3	14	1	1	6	10	3	4	11	4
K2	6	1	1	4	1	5	1	5	0	0	4	2	0	1	3	2
K3	2	1	0	1	0	2	0	2	0	0	2	0	0	0	1	1
K4	8	2	1	5	2	6	2	3	1	2	2	2	4	3	1	4
K5	28	4	9	15	13	15	6	17	2	3	0	13	15	9	12	7
K6	11	1	2	8	11	0	0	11	0	0	0	2	9	3	6	2
K7	11	0	3	8	9	2	1	10	0	0	0	1	10	4	5	2
K8	4	1	0	3	0	4	0	4	0	0	0	2	2	0	2	2
K9	10	7	2	1	3	7	2	7	1	0	2	5	3	1	6	3
K10	1	0	1	0	1	0	0	1	0	0	0	0	1	0	1	0
K11	13	4	8	1	7	6	1	10	1	1	0	6	7	3	10	0
K12	4	3	1	0	2	2	2	2	0	0	0	2	2	0	4	0
	117	32	33	52	52	65	18	86	6	7	16	45	56	28	62	27

Iz Upitnika SF-36 za opis ovih klastera dobivenih računalnom obradom tekstnih podataka te značajki pojedinih ispitanika korišteni su ponovno profili fizičkog (PCS) i psihičkog (MCS) zdravlja. Za svaki je klaster izračunana srednja vrijednost obaju profila svih ispitanika. Srednje vrijednosti tih profila označene su \underline{PCS}_{kvan} i \underline{MCS}_{kvan} (Tablica 5.26.).

Tablica 5.26. Srednje vrijednosti profila fizičkog i psihičkog zdravlja u klasterima određenih programom Dtm-Vic.

Klaster	broj isp.	<u>PCS</u>_{kvan}	<u>MCS</u>_{kvan}
K1	19	41,23	49,27
K2	6	29,25	44,55
K3	2	36,29	45,15
K4	8	34,70	39,61
K5	28	44,42	45,49
K6	11	43,14	49,08
K7	11	44,67	50,99
K8	4	27,61	35,62
K9	10	41,79	46,71
K10	1	58,58	50,69
K11	13	51,27	48,27
K12	4	57,76	43,30
	117		

Klasteri dobiveni programom Dtm-Vic statističkom obradama tekstova koji su izrekli ispitanici i potom digitaliziranih i prilagođenih za obradu računalnim programom opisani su zatim s pomoću kategorija značajki ispitanika u tim klasterima i profila zdravlja iz Upitnika SF-36. Iz Tablice 5.25. odabrane su za svaki klaster tri (3) najučestalije kategorije među svim kategorijama svih osnovnih značajki. Tim su karakteristikama dodane i srednje vrijednosti obaju profila zdravlja svih ispitanika u pojedinom klasteru. Time je svaki klaster dobiven računalnom metodom (programom Dtm-Vic) opisan brojem ispitanika, trima kategorijama osnovnih značajki obuhvaćenih ovim istraživanjima, profilima fizičkog i psihičkog zdravlja iz Upitnika SF-36 (Tablica 5.27.).

Tablica 5.27. Opis klastera određenih programom Dtm-Vic s pomoću kategorija i srednjih vrijednosti obaju profila.

<i>Klaster</i>	<i>broj isp.</i>	<i>Kat I.</i>	<i>Kat II.</i>	<i>Kat III.</i>	<i>PCS_{kvan}</i>	<i>MCS_{kvan}</i>
K1	19	z	brak	jedn	41,23	49,27
K2	6	z	brak	star	29,25	44,55
K3	2	z	brak	osn	36,29	45,15
K4	8	z	star		34,70	39,61
K5	28	brak	star	vis.šk.	44,42	45,49
K6	11	m	brak	vis.šk.	43,14	49,08
K7	11	brak	vis.šk.	m	44,67	50,99
K8	4	z	brak	star	27,61	35,62
K9	10	mlad	z	brak	41,79	46,71
K10	1				58,58	50,69
K11	13	brak	jedn	sred	51,27	48,27
K12	4	jedn	mlad		57,76	43,30

Program Dm-Vic kombinira rezultate korespondencijske analize i klaster analize. Klasteri su dobiveni kombiniranom metodom hijerarhijskog i direktnog klasteriranja, a zatim je izračunata udaljenost pojedinih klastera od faktorskih osi i ishodišta iz pojedinih ravnina definiranih tim osima. Numerički rezultati vezani uz prvu ravninu i dobivene klastere dani su u Tablici 5.28.

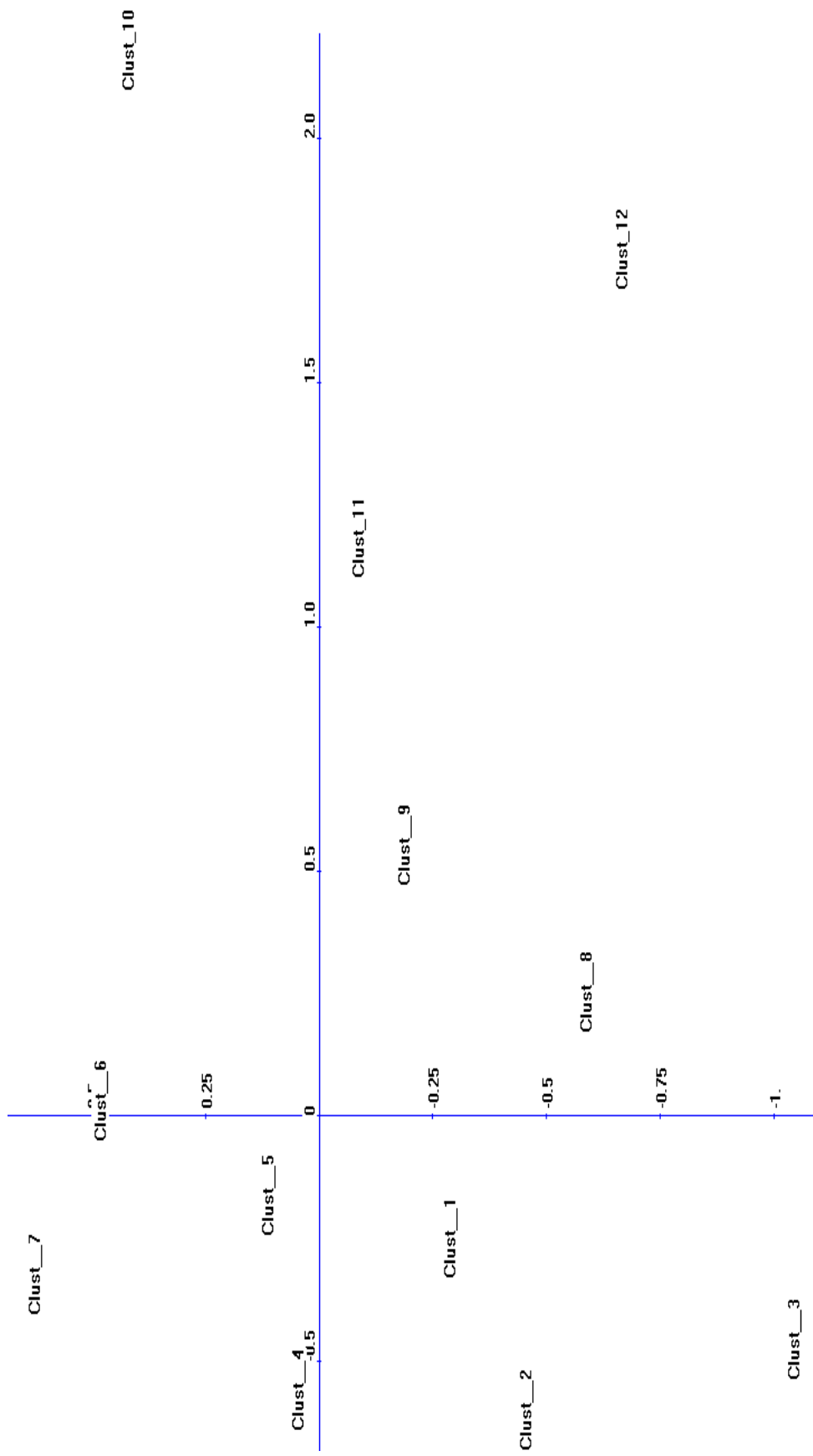
Tablica 5.28. Udaljenosti klastera dobivenih programom Dtm-Vic od ishodišta prve ravnine te prve i druge osi.

<i>Klaster</i>	<i>broj isp.</i>	<i>udaljenost od ishodišta</i>	<i>prva os</i>	<i>druga os</i>
K1	19	0,18	- 0,33	- 0,27
K2	6	0,66	- 0,69	- 0,44
K3	2	1,34	- 0,54	- 1,02
K4	8	0,42	- 0,64	0,06
K5	28	0,08	- 0,25	0,13
K6	11	0,25	- 0,05	0,50
K7	11	0,58	- 0,41	0,65
K8	4	0,35	0,17	- 0,57
K9	10	0,25	0,47	- 0,17
K10	1	4,59	2,10	0,44
K11	13	1,22	1,10	- 0,07
K12	4	3,28	1,69	- 0,65

Vizualizacija rezultata korespondencijske i klaster analize je grafički prikaz klastera u pojedinim ravninama definiranim osima dobivenim korespondencijskom analizom (Slika 5.10.).

Na slici su prva i druga faktorska os iz korespondencijske analize redaka matrice (riječi) i u tu prvu ravninu smješteno je 12 klastera definiranih računalnom metodom s pomoću programa Dtm-Vic i opisanih ranije.

Slika 5.10. Vizualizacija korespondencijske i klaster analize.



6. RASPRAVA

Ovo se istraživanje temeljilo na tekstnim podacima, kakve i koliko informacija donose izgovorene riječi u opisu nekog zdravstvenog događaja, a koje informacije standardizirani upitnik kojim također možemo saznati podatke o istom događaju. U ovom istraživanju zdravstveni događaj je ispitanikovo sadašnje zdravstveno stanje, i još malo više: ima li to zdravstveno stanje utjecaja na njegov svakodnevni život, aktivnosti, navike i druženje s bliskim i dragim osobama.

Taj je zdravstveni događaj dio uobičajene komunikacije između liječnika i pacijenta, ili bi, možda, trebao biti. Pacijenti smatraju da liječnici premalo s njima razgovaraju, da se sada i liječnikovo računalo ispriječilo između njihova otvorena razgovora. Liječnici, pak, misle da nemaju vremena za razgovore kakve bi pacijenti željeli. Američka udruga American Healthways organizirala je 2003. godine veliki skup posvećen promijenjenim odnosima pacijenata i liječnika na kojemu su dane smjernice tih odnosa u 21. stoljeću [49.]. Ta se komunikacija osobito mora njegovati i poboljšati kod nekih posebnih skupina pacijenata [50.]. Tange [51.] je još 1999. godine zagovarao da razgovori s pacijentima uđu u elektronički zdravstveni zapis (EHR). Međutim, neka istraživanja pokazuju da svi liječnici još nisu spremni za izazove informatizacije svojih ambulanti [52.].

Razgovor s ispitanikom koji bi trebao rezultirati nizom važnih podataka o tako osjetljivoj temi kao što je zdravstveno stanje nije nimalo jednostavno provesti i pripada skupu naturalističkih raspitivanja. Ta su raspitivanja danas mnogo češća i zahtjevnija nego prije. Godine 1985. S.Y. Lyncoln i E. Guba u svojoj knjizi *Naturalistic Inquiry* [53.] detaljno su opisali kako takvo raspitivanje treba pripremiti, provesti i analizirati. Ta se knjiga s detaljno opisanim postupcima i uputama te ostali njihovi radovi citiraju i smatraju nezaobilaznim uputama u svakom istraživanju koje uključuje naturalističko raspitivanje [54., 55.]. Zadnjih desetak godina ta su raspitivanja unaprijeđena pa je A. Bowling istraživao kako ih učiniti još kvalitetnijima [56., 57.]. Neki radovi donose iskustva u provođenju raspitivanja na velikom broju ispitanika [58.]. Naturalistička raspitivanja primijenjena su i u istraživanjima u Hrvatskoj [59., 60., 61., 62., 63.].

Uporaba interneta pokazuje neka nova iskustva u provođenju naturalističkih raspitivanja [64.].

Iskustva i upute iz literature te osobnih ranijih radova primijenila sam u ovom istraživanju. No pokazalo se da ispitanicima nije jednostavno govoriti o svojem zdravlju.

Ako problema i teškoća sa zdravljem nema, onda o tome nemamo što reći, zdravlje prihvaćamo kao činjenicu kojom se ne bavimo sve dok se ne poremeti. Razgovori sa zdravim ispitanicima bili su kratki, često neinformativni, a potpitanja koja su trebala pomoći jednostavno nisu poticala daljnji razgovor jer su ispitanici smatrali da nemaju što reći. Neki pričljiviji, u dobroj namjeri da ipak nešto kažu kad su već upitani, govorili su općenito o zdravlju, o različitim pristupima zdravlju i liječenju. To su oni „filozofi“ iz klastera definiranih interpretativnom metodom.

Postoje li zdravstveni problemi i teškoće, treba se htjeti i moći otvoriti nepoznatoj osobi koju prvi put susrećete i razgovarati o tome. Svakom se čini da će lakše početi svoju priču ako počne „od početka“. Ispitanici su većinom govorili o počecima svojih teškoća, problema, bolesti. Znali su u svojim pričama o zdravlju otići daleko u svoje djetinjstvo i povezati mnogo godina svojeg života. Često se događalo da zdravstveni problem koji su mi ispričali u tančine uopće više nije ni aktualan, ali on je ostao zauvijek u njihovu pamćenju kao problem i uvijek se na njega vraćaju. Dakle, povijesno prikazivanje svojih zdravstvenih teškoća odvlači pozornost od sadašnjeg statusa, ali može biti informativno zbog nekih zdravstvenih podataka koji se mogu preko njih doznati. Mnogo je veći problem niz informacija o obitelji, užoj i široj, prijateljima, susjedima, događajima, zatim raznorazni komentari koji nemaju mnogo veze s objektivnim zdravstvenim statusom.

Kriteriji za izbor ispitanika za istraživanje bili su njihov pristanak za sudjelovanje u ovom istraživanju i volja za razgovor o svojem zdravlju. Ovo je primjer metode *intensity sampling*, a ekstremni slučajevi ispitanika koji nisu znali što reći o svojem zdravlju izostavljeni su, iako su potpisali svoj pristanak za sudjelovanje u istraživanju i ispunili upitnik SF-36 [37.].

Dodatne značajke 117 ispitanika u ovom istraživanju statistički su obrađene standardnim metodama [65., 66., 67., 68.].

Specifičnosti s obzirom na dodatne značajke:

- **DOB.**

Uzorak je izrazito starije dobi – više od 45% ih je starijih od 65 godina. Oni su imali vremena za razgovor, nekima je to čak bilo zanimljivo iskustvo, oni su i bolesniji od mlađe populacije pa imaju što reći u takvom razgovoru. Međutim ta starija dob definirala je i zanimanje ispitanika – oni su većinom (59%) umirovljenici.

- **SPOL.**

Kolikogod su ispitanici birani prema kriteriju „pričljivosti“, struktura prema spolu s obzirom na prevladavajuću stariju dob je prilično realna: 44,4% ispitanika su muškarci, a 55,6% žene.

Međutim zanimljivija je razlika prema spolu u sadržaju njihovih razgovora. Muškarci su većinom vrlo sustavno izlagali o svojem zdravlju, točnije bolestima s čestim dodatnim objašnjenjima i detaljima. Vrlo često su svoje zdravstveno stanje povezivali sa svojim poslom i teškoćama vezanim uz stvaranje karijere, zatim s imovinom, obično vikendicom i autom. Svoje bračne družice i djecu spominjali su uglavnom usputno, no unuci su češće bili u „priči“. Žene su često naizmjenično „skakale“ s teme na temu. Obavezno bi, ako su udane, ili bile, te imale djecu, u svoj odgovor „uplele“ djecu i / ili unuke, bračne drugove i to kao paralelnu „priču“. Ta razlika u muškom i ženskom razgovoru pokazala se kasnije kod klasteriranja, i interpretativnom metodom i računalnom metodom, presudnom u opisu nekih klastera.

- **BRAČNO STANJE.**

Velika većina ispitanika je u braku, čak njih 73,5%. Međutim zanimljivo je da su svih 7 ispitanika iz kategorije udovac / udovica žene, a od 6 njih iz

kategorije rastavljen / rastavljena 5 je žena. Pretpostavka je da muškarci iz tih dviju kategorija nisu zadovoljili kriterije za *intensity sampling* ili nisu pristali sudjelovati u ovom istraživanju.

- ZAVRŠENA ŠKOLA.

Ispitanici u ovom istraživanju bili su višeg obrazovanja nego što je prosjek za Hrvatsku [22.]. Čak njih 46% ima završenu višu ili visoku školu. Pretpostavka je te činjenice da je velika većina ispitanika iz Zagreba i da su obrazovaniji bili skloniji sudjelovati u istraživanju. Ta se činjenica odrazila i na vokabular u tekstnoj analizi.

Statistički je značajna razlika u završenoj školi ispitanika prema spolu, više muškaraca ima završenu višu ili visoku školu, a više žena srednju ili osnovnu školu.

- SAMOPROCJENA SVOJEG ZDRAVLJA S OBZIROM NA ZDRAVLJE SVOJIH VRŠNJAKA.

Malo više od polovine (52,2%) ispitanika samoprocijenilo je svoje zdravlje jednakim u odnosu na zdravlje svojih vršnjaka, a jednaki broj (23,9%) smatra svoje zdravlje boljim, odnosno lošijim od zdravlja svojih vršnjaka. Ispitivanjem statističke značajnosti samoprocjene zdravlja s obzirom na zdravlje svojih vršnjaka pokazalo se da nema razlike prema spolu niti prema završenoj školi. Zanimljivo je da su se neki ispitanici u razgovoru tužili na svoje zdravstvene probleme i teškoće, a ipak su svoje zdravlje procijenili podjednakim sa zdravljem svojih vršnjaka.

Ispitanici su bili iz različitih mjesta stanovanja. Većina (75%) ih je bila iz Zagreba, iz triju dijelova grada: Novog Zagreba, Kustošije i (šireg) središta grada, ostali su bili iz različitih ruralnih područja: Slavonije, Istre, Prigorja. Htjela sam izbjeći da samo stanovnici Zagreba budu obuhvaćeni ovim istraživanjem. Izbor ispitanika iz raznih dijelova Zagreba prvenstveno je utjecao na interpretativnu metodu, što je dijelom bilo i očekivano. Stanovnici ruralnih područja teže su davali pristanak za

sudjelovanje u istraživanju. Razgovor je s onima koji su pristali bio ugodan, uspješan i informativan. Ipak za istraživanje je bilo presudna metoda *intensity sampling*, a ne mjesto stanovanja.

Izbor Upitnika SF-36 kao jednog od najčešće korištenog za evaluaciju zdravstvenog stanja potaknut je i nizom rezultata istraživanja širom svijeta, ali i u Hrvatskoj [23., 26., 27., 69., 70.].

Iako je ovih 117 ispitanika izabrano metodom *intensity sampling*, ipak prema standardiziranom Upitniku SF-36 oni ne odstupaju od prosjeka Hrvatske u 8 dimenzija. Srednje vrijednosti dimenzija fizičkog i socijalnog funkcioniranja podudaraju se u oba uzorka, a srednje vrijednosti dimenzija percepcije općeg zdravlja i vitalnosti su nešto malo više u ovom uzorku. Nešto je veća razlika u srednjim vrijednostima dimenzija fizičkih ograničenja i tjelesnih bolova, u uzorku su niže od prosječnih hrvatskih. U dimenziji psihičkog zdravlja srednja vrijednost uzorka veća je od prosječnih hrvatskih. Spomenute razlike nisu statistički značajne za $p = 0,05$. Za hrvatski prosjek su uzeti rezultati dobiveni u istraživanju provedenom 1997-1999. na Školi narodnog zdravlja „Andrija Štampar“ na reprezentativnom uzorku hrvatske populacije koji je obuhvatio oko 1% stanovnika svih regija u Hrvatskoj [25.].

Rezultati dobiveni upitnikom SF-36 u ovom istraživanju dali su još neke zanimljive informacije o tom uzorku. Vrijednosti profila psihičkog zdravlja više su od vrijednosti profila fizičkog zdravlja. Među ispitanicima starije životne dobi prevladavaju ispitanici koji imaju niske vrijednosti profila fizičkog zdravlja, ali više vrijednosti profila psihičkog zdravlja. Međutim postoje ispitanici kod kojih su te vrijednosti potpuno suprotne. Naime, neki, osobito mlađi ispitanici, imaju više vrijednosti profila fizičkog zdravlja s obzirom na ostale ispitanike, ali su im vrijednosti profila psihičkog zdravlja niže od ostalih.

Odabir tekstnih podataka kao podataka na kojima se temelji ovo istraživanje definirao je i kvalitativni pristup istraživanjima kao metodu.

Kvalitativni pristup istraživanju kao metoda istraživanja u epidemiologiji te zdravstvu općenito tek je posljednjih 20tak godina prihvaćen kao ravnopravna metoda kvantitativnom pristupu i primjenjuje se sve češće. U početku su se jednostavno primjenjivali postupci uobičajeni u sociologiji i etnologiji gdje je kvalitativni pristup ranije korišten. Početkom 2000-tih godina kvalitativan pristup u medicini i zdravstvu već je prihvaćena metoda, tako Giacommini daje upute kako se njome koristiti [71.], dok se neki autori više zanimaju analizom takvih podataka [72., 73.] ili primjenom [74., 75.]. Neki autori, poput Maysa i McCavitta u svojim radovima [76., 77.] rezultate istraživanja dobivene kvalitativnim pristupom vide primijenjene na istraživanja kvalitete života ispitanika. Istodobno se, a osobito kasnije, ti rezultati primjenjuju i za istraživanja i poboljšanja kvalitete procesa u zdravstvu te zdravstvenog menadžmenta [78., 79.].

Ciljevi istraživanja definirali su metode prikupljanja i analize podataka. Tekstni podaci prikupljeni su metodama sudioničkog promatranja i strukturiranog razgovora, a analizirani su interpretativnom metodom i računalnom obradom [37.].

Kako je jedan od ciljeva istraživanja upravo bio usporediti dvije metode analize podataka dobivenih kvalitativnim pristupom istraživanju, interpretativnu metodu i metodu obrade računalom, osnovni skup riječi morao je ostati isti. Prema tome, nikakvi zahvati u tekst, koji inače daju jasniju vizualizaciju nekih multivarijantnih analiza u tekstnoj statistici, u ovom slučaju nisu dolazili u obzir; tekst je morao biti jedan jedini za obje metode.

Duljina razgovora, ponavljanje riječi, hapaksi i slično nisu bili veliki problem u interpretativnoj tehnici. Vjerojatno je i relativno malen broj ispitanika, njih 117, tomu doprinio. Vjerujem da bi velik ili čak samo nešto veći broj ispitanika ipak bio zapreka provođenju kvalitetne analize interpretativnom metodom. Naime, kod interpretativne metode glavni je zahtjev na istraživača da zadrži ista mjerila za svakog ispitanika, da jednakom pozornošću obradi podatke svakog ispitanika i, naravno, da podatke obradi u „jednom dahu“ kako bi ti zahtjevi bili poštovani.

Interpretativna metoda u analizi prikupljenih tekstnih podataka u ovom radu praktički je bila sljedeća: pročitala sam svih 117 kompletnih razgovora odjednom, napravila bilješku o svojem dojmu s obzirom na tekst svakog ispitanika, zatim sam za slične dojmove našla zajednički naziv i iz takve je obrade nastalo 12 grupa ispitanika, odnosno klastera. Klastera sam imenovala prema prevladavajućoj osobini koja mi se nametnula pri čitanju tih razgovora.

Tako su definirani klasteri:

- „ZDRAV“
ispitanici koji su mi se činili zdravima, bilo da su to sami izjavili ili je to proizlazilo iz razgovora.
- „ZADOVOLJAN“
ispitanici koji su možda imali trenutačne zdravstvene tegobe, ali su u biti zadovoljni svojim zdravstvenim stanjem.
- „UMORAN“
ispitanici koji se nisu tužili na neke posebne zdravstvene teškoće, ali su više puta u razgovoru istaknuli svoj umor i preopterećenost.
- „UZNEMIREN“
ispitanici koji su nervozni, ne uvijek i ne samo, zbog svojih zdravstvenih teškoća. Smeta im mnogo toga iz njihove neposredne okoline i zdravstvene teškoće su im dodatni razlog za nerviranje.
- „FILOZOF“
ispitanici koji ili uopće nisu imali zdravstvenih teškoća ili imaju ozbiljne teškoće koje ne žele priznati pa ih relativiziraju. Oni u razgovoru ne govore o svojem zdravstvenom stanju nego o raznim mogućnostima liječenja i izlječenja, što uopće znači imati zdravstvene teškoće, kako je teško biti bolestan, i slično.
- „BOLESNI SAMAC“
među ispitanicima se prema izrečenom u razgovoru izdvojilo 7 ispitanika koje sam odmah prepoznala da žive sami, ali i da su bolesni.

- „BOLESTAN, AKTIVAN“
klaster je ispitanika koji svoje, često ozbiljne, zdravstvene teškoće nadoknađuju velikom aktivnošću. Ponekad mi se kod nekih ta velika aktivnost, gotovo hiperaktivnost, činila i nerealnom s obzirom na zdravstvene teškoće o kojima su govorili.
- „OPTEREĆEN BOLEŠĆU“
razgovor s tim ispitanicima bio je razgovor samo o bolesti od koje boluju ili su болоvali, oni su mi ispričali u detalje svaku fazu svoje bolesti, svaki odlazak liječniku, u bolnicu, svaki postupak, lijek. Dojam je da oni i nemaju druge interese nego se sve „vrti“ oko bolesti.
- „POMIREN S BOLEŠĆU I / ILI ŽIVOTOM“
neki su ispitanici, čak neki među njima i ozbiljnije bolesni prihvatili svoju bolest kao činjenicu i uskladili svoje životne navike s novom situacijom. Čini se da im je takvo stajalište pomoglo da se nose sa svim teškoćama i dojam je da su na takav način izbjegli stres koji prati zdravstvene teškoće.
- „BOLESTAN, OPTIMIST“
ispitanici koji su svjesni svojeg lošeg zdravlja i ne baš dobre budućnosti, ali oni ipak vjeruju u neko rješenje. Oni su vrlo slični ispitanicima u prethodnom klasteru, ali dojam je bio da ipak imaju ili samo iskazuju neko veće, pozitivno očekivanje.
- „BOLESTAN, PESIMIST“
nekoliko ispitanika ne vidi nikakvo rješenje svojih zdravstvenih teškoća, oni ne skrivaju čak ni suicidalne ideje.
- „NEDEFINIRAN“
u tom su klasteru ispitanici koji u svojem razgovoru nisu na mene ostavili nikakav dojam.

Ovi klasteri pokazali su se dobro definiranima u usporedbi s vrijednostima profila fizičkog i psihičkog zdravlja za svakog ispitanika iz Upitnika SF-36.

- Klaster „ZDRAV“ ima srednje vrijednosti obaju profila vrlo visoke, iznad trećeg kvartila za profil fizičkog zdravlja cijelog uzorka, a iznad drugog za profil psihičkog zdravlja cijelog uzorka.

- Klaster „ZADOVOLJAN“ ima srednju vrijednost profila fizičkog zdravlja nižu od srednje vrijednosti za klaster „ZDRAV“, ali višu srednju vrijednost profila psihičkog zdravlja.
- U klasteru „UZNEMIREN“ srednja vrijednost profila fizičkog zdravlja je viša od drugog kvartila istog profila cijelog uzorka, ali je srednja vrijednost profila psihičkog zdravlja tog klastera niža od srednje vrijednosti tog profila za cijeli uzorak.
- Zanimljive su vrijednosti za klaster „BOLESTAN, OPTIMIST“. Srednja vrijednost profila fizičkog zdravlja je niža od srednje vrijednosti za cijeli uzorak i nalazi se između prvog i drugog kvartila, ali srednja vrijednost profila psihičkog zdravlja najveća je među svim ovako definiranim klasterima.
- Slično je sa srednjim vrijednostima profila za klaster „POMIREN S BOLEŠĆU I / ILI ŽIVOTOM“. One su između prvog i drugog kvartila za profil fizičkog zdravlja cijelog uzorka, ali srednja vrijednost profila psihičkog zdravlja odgovara drugom kvartilu toga profila za cijeli uzorak.
- Srednje vrijednosti obaju profila za klaster „BOLESTAN, PESIMIST“ najniže su među svim klasterima, mnogo niže od prvih kvartila cijelog uzorka.

Analizirajući klustere dobivene interpretativnom metodom s obzirom na profile fizičkog i psihičkog zdravlja dobivene iz upitnika SF-36 može se uočiti sljedeće:

- Srednje vrijednosti profila fizičkog zdravlja gotovo svih klastera nalaze se u interkvartilu (između prvog i trećeg kvartila) cijelog uzorka. Srednje vrijednosti profila fizičkog zdravlja za klustere „BOLESTAN, PESIMIST“ i „NEDEFINIRAN“ manje su od vrijednosti prvog kvartila cijelog uzorka. Srednje vrijednosti profila fizičkog zdravlja za klaster „ZDRAV“ veća je od vrijednosti trećeg kvartila cijelog uzorka.
- Za profil psihičkog zdravlja srednje vrijednosti za četiri klastera manje su od vrijednosti prvog kvartila za cijeli uzorak, to su klasteri „UMORAN“, „BOLESNI SAMAC“, „OPTEREĆEN BOLEŠĆU“ i „BOLESTAN, PESIMIST“. Za ostalih osam klastera srednje vrijednosti se nalaze u interkvartilu. Srednja vrijednost profila psihičkog zdravlja niti u jednom klasteru nije veća od vrijednosti trećeg kvartila za cijeli uzorak.

U opis pojedinih klastera uvrštene su i značajke koje dodatno opisuju ispitanike: spol, dob, bračno stanje, završena škola i samoprocjena svojeg zdravlja s obzirom na zdravlje svojih vršnjaka. Analizirajući svaki klaster prema kategorijama pojedinih varijabli i učestalosti ispitanika dobila sam dodatni opis pojedinog klastera.

- Većina ispitanika u klasteru „ZDRAV“ smatra svoje zdravlje jednakim sa zdravljem svojih vršnjaka, prema dobi su mlađi od 50 godina, u braku su.
- U klasteru „ZADOVOLJAN“ većina ispitanika je u braku, ženskoga je spola, svoje zdravlje samoprocjenjuju jednakim s obzirom na zdravlje svojih vršnjaka.
- U klasteru „POMIREN S BOLEŠĆU I / ILI ŽIVOTOM“ većina ispitanika je u braku, starija je od 65 godina, muškog spola.
- Ispitanici u klasteru „BOLESTAN, OPTIMIST“ većinom su muškog spola, u braku su, imaju završenu višu ili visoku školu.
- U klasteru „UMORAN“ svi su ispitanici u braku i imaju završenu srednju školu.
- Svi ispitanici u klasteru „BOLESNI SAMAC“ stariji su od 65 godina i, naravno, nijedan nije u braku.
- U klasteru „OPTEREĆEN BOLEŠĆU“ svi su ispitanici ženskoga spola i svi samoprocjenjuju svoje zdravlje lošijim s obzirom na zdravlje svojih vršnjaka.
- Svi ispitanici u klasteru „BOLESTAN, PESIMIST“ u braku su i svoje zdravlje samoprocjenjuju lošijim u odnosu na zdravlje svojih vršnjaka.

Opis klastera prema učestalosti ispitanika u pojedinim kategorijama značajki potvrđuje dobru procjenu ispitanika interpretativnom metodom analize tekstnih podataka – razgovora.

Druga metoda analize podataka dobivenih kvalitativnim pristupom istraživanju je obrada podataka, u ovom istraživanju razgovora s ispitanicima, računalom, odnosno računalnim programom.

Analiza tekstova računalnim programom provodi se većinom s dvama različitim pristupima. Jedan je, već od prije poznat i korišten, lingvistički, a drugi statistički. Lingvistički je prihvaćen i primjenjivan od prvih lingvističkih analiza

tekstova najrazličitijih veličina, sadržaja, namjene i primjene, a osobito je unaprijeđen razvojem IT [80.]. Mitkov i Taylor u svojim priručnicima za kompjuteriziranu lingvistiku [81., 82.] osim uputa za analizu daju i pregled prikaza rezultata u lingvističkim istraživanjima. Jurafski i Martin [83.] prenose iskustva iz lingvističkih istraživanja u obradu prirodnog jezika (engl. *natural language processing*). U ovom je istraživanju lingvistički pristup analizi teksta sveden na leksikometriju te korišten u organizaciji i sistematizaciji tekstnih podataka.

Drugi pristup analizi tekstova je statistički i temelji se na metodama tekstne statistike koje su krajem 70-tih godina ustanovili L. Lebart i A. Morineau [84.], zatim J.-P. Benzécri, M. Becue i A. Salem [43., 46.]. Riječi se tretiraju kao kvalitativne varijable i na njih se primjenjuju odgovarajući statistički postupci [43., 85., 86.]. M. Greenacre je opisao praktične primjene metode korespondencijske analize [87., 88.] koja je ovdje prilagođena tekstnim podacima [43.]. Vizualizacija korespondencijske analize osim kod Lebartovih računalnih programa koristi se i u drugim statističkim programima [89.]. Ove metode tekstne statistike osobitu vrijednost imaju u činjenici što se tekst analizira nepromijenjen, dakle upravo onako kako je izrečen ili napisan, bez ikakvih zahvata u njegov korpus [90.].

U ovom sam se istraživanju za računalnu obradu podataka dobivenih kvalitativnim pristupom koristila programom za statističku analizu Dtm-Vic [48.] koji uz statističku analizu tekstnih podataka omogućava njihovu obradu zajedno s ostalim podacima vezanim uz ispitanike.

Ukupan broj riječi (tekst) koji su izrekli ispitanici vrlo je velik 32365, što znači da je svaki ispitanik izrekao prosječno 277 riječi. Kako je određeni broj ispitanika ipak, bez obzira na *intensity sampling*, bio dosta suzdržan u razgovoru, jasno je da su neki bili vrlo „pričljivi“ u opisu svojeg zdravstvenog stanja rabeći i više od 1000 riječi. Broj različitih riječi je 5237 ili 45 po ispitaniku, što znači da je vokabular ovog uzorka ispitanika 16,2% od ukupnog broja izrečenih riječi.

Problem ovog teksta bio je veliki broj hapaksa (samo jednom izrečenih riječi), čak 2966. Praktički to znači da je broj hapaksa 9% od ukupnog broja riječi ili čak 56,6% od broja različitih riječi.

Veliki broj hapaksa u tekstovima na hrvatskom jeziku je očekivan. Proizlazi iz gramatike našeg jezika: 7 padeža u jednini i množini, 3 roda i složeni glagolski oblici glavni su problem. Takvi se problemi obično izbjegavaju nekim zahvatima u tekst, primjerice djelomičnom lematizacijom, spajanjem sinonima, izostavljanjem pomoćnih riječi i slično. U ovom istraživanju to nije bilo moguće učiniti zbog postavljenih osnovnih ciljeva istraživanja.

Osim ovih rezultata vezanih uz broj riječi, jedan od rezultata leksikometrije analiziranog teksta je indeks učestalosti. U ovom tekstu među prvih pedesetak najučestalijih riječi velik je broj pomoćnih riječi. Međutim riječi „ja“, „mi“, „me“, „imam“ nalaze se među 15 najučestalijih i pokazuju osobni pristup i stajalište u razgovoru.

Programom Dtm-Vic uspjela sam provesti korespondencijsku i klaster analizu tako velikog broja riječi i kombinirati s ostalim podacima ispitanika u ovom uzorku. Ipak, u vizualizaciji ima nekih manjkavosti jer toliki broj riječi nije dopustio uporabu dodatnih funkcija u grafičkom prikazu, tako da je ostao standardni prikaz. Taj veliki broj riječi nije baš uobičajen za statističku analizu otvorenih odgovora pa i nema uputa kako provesti analize. Savjet je i preporuka smanjiti broj riječi, što bi ovdje narušilo postupak istraživanja.

Program Dtm-Vic u statističkim analizama automatski odbacuje hapakse, jer se pretpostavlja da su oni „riješeni“ nakon leksikometrije nekim od uobičajenih načina, odnosno da oni koji su ostali nisu bitni za analize. Tako je nakon odbacivanja hapaksa za analizu ostalo 2271 riječi koje su se pojavljivale od najmanje dvaput do riječi „sam“ koja se pojavljuje 1129 puta.

Korespondencijska analiza tako velike matrice (2271 x 117) dala je vrlo skromne rezultate. Prve dvije svojstvene vrijednosti određuju samo 3,46% varijance, a čak 12 svojstvenih vrijednosti odredilo je samo 16,61% varijance.

Program je izračunao i koji elementi redaka i stupaca imaju najbolju distribuciju na prve dvije osi. Kako je vidljivo iz tablica (Tablice 5.19. i 5.20.), distribucija na prvu i drugu os dovoljno je jasna za grafički prikaz. Tako da je za grafičke prikaze u ovom radu korištena distribucija samo na prvu i drugu os, odnosno u prvoj korespondencijskoj ravnini.

Zbog velikog broja elemenata redaka matrice (riječi) nije bilo jednostavno definirati prvu i drugu os s pomoću njih. Riječi s najvećim doprinosom na jedan smjer svake osi međusobno su nepovezane i teško je naći čvrsti opis. Tako da u ovom slučaju osi ostaju nedefinirane s obzirom na značenje u skupu korištenih riječi.

Za definiciju prve dvije osi elementima stupaca, dakle ispitanicima, koristila sam se ostalim značajkama koje opisuju ispitanike. To su ponajprije profili fizičkog i psihičkog zdravlja iz Upitnika SF-36, te značajke dob, spol, bračno stanje, završena škola i samoprocjena procjena svojeg zdravlja s obzirom na zdravlje svojih vršnjaka.

Rezultati dobiveni analizom svih tih podataka daju sljedeći opis prve i druge osi:

- Prva os definira ispitanike prema fizičkom zdravlju, na negativnom smjeru osi su ispitanici čija je vrijednost profila PCS niska, a na pozitivnom smjeru ispitanici čija je vrijednost profila PCS visoka.
- Druga os definira ispitanike prema spolu, na negativnom smjeru druge osi su ispitanici ženskog spola, a na pozitivnom smjeru druge osi su ispitanici muškog spola.

Ti opisi su vidljivi i iz vizualizacija korespondencijske analize za značajke ispitanika i za kvartile profila njihova fizičkog i psihičkog zdravlja (Slike 5.8. i 5.9.).

Programom Dtm-Vic napravljena je procedura klasteriranja i to kombinacijom metoda hijerarhijskog i direktnog s unaprijed zadanim brojem klastera. Program je odredio 15 klastera kao odgovarajući broj klastera s obzirom na broj redaka i stupaca matrice. Izračunom su dobivena dva klastera s po jednim ispitanikom te nekoliko s 2 do 4 ispitanika. Kako bih smanjila broj klastera s malim brojem ispitanika, odabrala

sam da unaprijed zadani broj klastera bude 12. Taj je broj klastera dao zadovoljavajuću distribuciju ispitanika po klasterima. Međutim ostala su četiri klastera s malim brojem ispitanika: po jedan je klaster s jednim ispitanikom i s dvama ispitanicima, a dva su s četiri ispitanika. Broj od 12 klastera čini se prikladnim jer je i interpretativnom metodom dobiveno 12 klastera pa se na taj način pokušalo usporediti klasterne dobivene s pomoću različitih metoda.

Program je vrlo precizno izračunao koje kategorije kojih značajki najbolje karakteriziraju pojedine klasterne računalnom analizom teksta:

- Klaster K1 karakteriziraju ženski spol i završena osnovna škola.
- Klasterne K2 i K3 karakterizira završena osnovna škola.
- Klaster K6 karakterizira muški spol i bračno stanje oženjen/udana.
- Klaster K7 karakterizira završena viša ili visoka škola i muški spol.
- Klaster K9 karakterizira mlađa dob.
- Klaster K11 karakterizira srednja dob.

Program je izračunao i 10 relativno najučestalijih riječi koje su izgovorili ispitanici u pojedinom klasteru. S obzirom na veliki broj različitih riječi taj je indeks učestalosti prema klasterima slabo informativan. Pokazuje da su svi ispitanici imali osobni pristup u svojim odgovorima, govorili su o sebi i pri tom su upotrijebljeni glagoli u prvom licu jednine.

Sljedeći je postupak bio tih 12 klastera (K1 – K12) dobivenih s pomoću računalne obrade teksta opisati na isti način kao i klasterne dobivene interpretativnom tehnikom (M1 – M12).

Analizirajući klasterne dobivene računalnom obradom podataka u odnosu na profile fizičkog i psihičkog zdravlja dobivene iz upitnika SF-36 može se uočiti sljedeće:

- Srednje vrijednosti profila fizičkog zdravlja gotovo svih klastera nalaze se u interkvartilu (između prvog i trećeg kvartila) za cijeli uzorak osim srednjih vrijednosti za klasterne K2 i K8 koje su manje od vrijednosti prvog kvartila cijelog

uzorka te srednje vrijednosti za klaster K10 (samo jedan ispitanik) i K12 koji su veći od vrijednosti trećeg kvartila cijelog uzorka.

- Za profil psihičkog zdravlja srednje vrijednosti svih klastera nalaze se u interkvartilu, osim za klaster K8 čija je srednja vrijednost manja od vrijednosti prvog kvartila za cijeli uzorak.

Zanimljivo je da su srednje vrijednosti obaju profila zdravlja ujednačene i teško je prepoznati karakteristike pojedinog klastera kroz profile PCS i MCS. Jedino se u klasterima K11 i K12 može uočiti da su srednje vrijednosti profila fizičkog zdravlja više od srednjih vrijednosti profila psihičkog zdravlja. Srednja vrijednost profila PCS u klasteru K11 približno je ista trećem kvartilu cijelog uzorka, a nešto viša u K12, dok su srednje vrijednosti profila psihičkog zdravlja između drugog i trećeg kvartila cijelog uzorka.

Analiziraju li se klasteri dobiveni računalnom obradom podataka u odnosu na kategorije pojedinih značajki i učestalost ispitanika, klasteri se mogu opisati na sljedeći način:

- U klasteru K1 većina je ispitanika ženskoga spola, u braku je, samoprocjenjuje svoje zdravlje podjednako sa zdravljem svojih vršnjaka.
- U klasteru K2 većina ispitanika je ženskoga spola, u braku.
- U klasteru K3 svi su ispitanici ženskoga spola, u braku su i imaju završenu osnovnu školu.
- U klasteru K4 svi su ispitanici ženskoga spola, većina je starije dobi (iznad 65 godina).
- U klasteru K5 većina je ispitanika u braku.
- U klasteru K6 svi ispitanici su muškoga spola i u braku, većina ima završenu višu ili visoku školu.
- U klasteru K7 većina ispitanika je u braku, ima završenu višu ili visoku školu, muškoga je spola.
- U klasteru K8 svi su ispitanici ženskoga spola i u braku.
- U klasteru K9 većina je ispitanika mlađe dobi (do 50 godina), ženskoga je spola, u braku.

- U klasteru K11 većina je ispitanika u braku, većina samoprocjenjuje svoje zdravlje podjednako sa zdravljem svojih vršnjaka, većina je srednje dobi (50 do 65 godina).
- U klasteru K12 svi ispitanici samoprocjenjuju svoje zdravlje podjednako sa zdravljem svojih vršnjaka, većina je mlađe dobi (do 50 godina).

Takav opis klastera podudara se s opisom tih klastera dobivenim računalnom obradom u gotovo svim klasterima (ne podudara se u K2), a u opisu s pomoću učestalosti ispitanika u K1 nema kategorije završena osnovna škola.

U ovom istraživanju klasteri dobiveni interpretativnom metodom (M1 – M12) i klasteri dobiveni računalnom metodom (K1 – K12) ne mogu se međusobno uspoređivati jer ne obuhvaćaju iste ispitanike.

Klasteri dobiveni računalnom metodom rezultat su samo učestalosti pojedinih riječi i povezani su samo na taj način sa svojim autorima – ispitanicima koji su ih izrekli. Vidljiva je činjenica da su se ispitanici s istim značajkama koristili istim riječima za opis svojeg zdravstvenog stanja, npr. u klasteru K1 su većinom žene, u klasteru K3 većina ispitanika ima završenu samo osnovnu školu, u klasteru K7 su većinom muškarci i ispitanici sa završenom višom ili visokom školom. Slično je i s profilima fizičkog i psihičkog zdravlja, u klasterima od K1 do K9 srednja vrijednost profila fizičkog zdravlja niža je od srednje vrijednosti profila psihičkog zdravlja, što vrijedi i za cijeli uzorak. Klasteri dobiveni računalnom metodom mogu se opisati s pomoću kategorija značajki ispitanika u njima te profilima fizičkog i psihičkog zdravlja, ali ne mogu se imenovati.

Klasteri M1 – M12 definirani interpretativnom metodom rezultat su razumskog pristupa izrečenom tekstu i empatičkog pristupa propitivanju i ispitanicima. Nije bilo važno kojim će riječima neki ispitanik opisati svoje zdravstveno stanje, nego kakav je dojam svojim riječima ostavio na ispitivača. Tako se čini da su klasteri M1 – M12 mnogo razumljivije i logičnije opisani profilima fizičkog i psihičkog zdravlja (PCS i MCS) te korištenim značajkama ispitanika, a svaki je klaster imenovan prema dojmu ispitivača.

7. ZAKLJUČAK

1. Kvalitativni i kvantitativni pristup epidemiološkim istraživanjima nisu suprotstavljani nego se nadopunjuju i obogaćuju interpretaciju rezultata.
2. Tekstni podaci dobiveni tematski ciljanim upitima kroz razgovor pogodni su za interpretativnu metodu analize podataka.
3. Rezultati dobiveni interpretativnom metodom analize usporedivi su s rezultatima analiza podataka dobivenih kvantitativnim pristupom istraživanjima.
4. Tekstni podaci koji su rezultat kvalitativnog pristupa epidemiološkim istraživanjima mogu se obrađivati posebnim računalnim programima za analizu tekstnih podataka. Definiraju se kao kvalitativni podaci i na njih se mogu primjenjuju svi postupci multivarijatne analize kvalitativnih podataka.
5. Rezultati dobiveni računalnom obradom tekstnih podataka usporedivi su s rezultatima analiza podataka dobivenih kvantitativnim pristupom istraživanjima.
6. Rezultati analize tekstnih podataka dobiveni interpretativnim metodama i rezultati analize tekstnih podataka dobiveni računalnim programima međusobno su usporedivi.
7. Rezultati analize tekstnih podataka dobiveni interpretativnim metodama i rezultati analize tekstnih podataka dobiveni računalnim programima međusobno su nadopunjuju. Prethodno provedena interpretativna analiza omogućuje prilagodbu istraživanja drugim metodama analize podataka.

8. Tekstni podaci dobiveni tematski ciljanim upitima kroz razgovor nisu pogodni za metodu računalne analize podataka, osobito ako su to dulji razgovori s puno izrečenog teksta ispitanika.

9. Za metodu računalne obrade podataka mnogo su pogodniji podaci koji se dobivaju upitima u standardiziranim upitnicima sa slobodnim odgovorima. Takvi tekstovi su obično kraći, bez pomoćnih riječi i s manjim varijacijama u padežima te s manje raznih složenih glagolskih oblika

8. SAŽETAK

U epidemiološkim istraživanjima danas su se ustalila dva metodološka pristupa, kvantitativni i kvalitativni. U ovom su radu primjenjena oba postupka upravo zbog svoje komplementarnosti. Polustrukturirani intervjui uz primjenu rafiniranog analitičkog postupka statističke analize teksta otvaraju mogućnost ulaska u dubinu i bit ispitivanog problema (doživljaja kvalitete života).

Ciljevi istraživanja bili su usporediti rezultate kvalitativne analize podataka dobivenih otvorenim ispitivanjima u epidemiološkim istraživanjima s rezultatima dobivenima analizom podataka iz Upitnika SF-36 te usporediti primjenjivost računalne analize tekstnih podataka i njezin doprinos saznanjima o ispitanicima s interpretativnom tehnikom.

Uzorak čini 117 ispitanika u dobi od 23 do 82 godine, žena je nešto više (55,6%), različitih su zvanja, većinom su (73,5%) u braku, više od 85% ima završenu barem srednju školu, većinom su umirovljenici (59%), više od polovice (52,2%) samoprocjenjuje svoje zdravlje jednakim u odnosu na zdravlje svojih vršnjaka, većina (75%) ih je iz raznih dijelova Zagreba.

Kao patentiran i validiran upitnik primijenjen je SF-36, hrvatsko izdanje. Rezultati analize upitnika SF-36 pokazuju da se rezultati statistički ne razlikuju od rezultata dobivenih ispitivanjem istim upitnikom oko 1% stanovnika svih regija Hrvatske.

Analizom otvorenih ispitivanja interpretativnom tehnikom dobiveno je 12 klastera ispitanika različite veličine. Svaki je klaster analiziran i opisan značajkama ispitanika te profilima fizičkog i psihičkog zdravlja dobivenim Upitnikom SF-36.

Računalna obrada tekstnih podataka programom Dtm-Vic kao rezultat leksikometrijske analize daje ukupni broj riječi 32365, od toga je 5381 ili 16,4% različitih riječi što pokazuje da se ispitanici uglavnom služe istim riječima u opisu svojeg zdravstvenog stanja, a učestalost pojedinih riječi pokazuje osobni pristup svojem zdravlju. Rezultat korespondencijske i klaster analize je 12 klastera ispitanika

različite veličine. Ti su klasteri također opisani značajkama ispitanika te profilima fizičkog i psihičkog zdravlja dobivenim Upitnikom SF-36.

Skupine (klasteri) dobivene različitim tehnikama (interpretativnom i računalnom) nisu međusobno usporedive zbog različitih ispitanika u njima.

Usporedba analiza upitnika SF-36 i tekstnih podataka dobivenih razgovorom pokazuje veliku sličnost u krajnjim rezultatima. Međutim tekstni podaci, odnosno rezultati njihovih analiza donose mnogo više informacija, mnogo su bolje nijansirani, čuvaju individualnost ispitanika.

Rad pokazuje da su tematski ciljane ankete kroz razgovor osobito primjenljive i korisne u epidemiološkim istraživanjima zbog bogatstva informacija na uzorku populacije koji je u ovom istraživanju bio izabran metodom *intensity sampling*.

9. SUMMARY

VALUE OF TEXTUAL DATA IN OBSERVATIONAL EPIDEMIOLOGICAL RESEARCH

There are two established methodological approaches in present-day epidemiological research: quantitative and qualitative. Both approaches were used in this paper due to their complementary quality. Semi-structured interviews used with a refined analytical approach of statistical text analysis provide an opportunity to reach to the core and essence of the subject (personal estimate of their quality of life).

The goal of this research was to compare the results of qualitative analysis of textual data acquired through general interview guided approach in epidemiological research with results acquired through analysis of data from SF-36 Health Survey. Additional goal was to compare the applicability of textual analysis done by computer programs and its contribution in gaining information on examinees in comparison with interpretative methods of textual analysis.

Sample consist of 117 examinees aged from 23 to 82, more women (55,6%) than men, of various professions, most of them married (73,5%), more than 85% have secondary-school education, most of them are retired (59 %), more than half (52,2%) of them self-evaluated their health is the same as health of their peers, most of them (75%) live in Zagreb.

This research used a Croatian edition of SF-36 as a validated and standardized questionnaire. Analysis of questionnaire SF-36 shows that the sample correlated to results gained through research based on the same questionnaire used on 1% of inhabitants in Croatian regions.

The analysis of interview guided approach through interpretative techniques resulted in 12 clusters with different numbers of examinees. Every cluster was analyzed and described by examinees' characteristics, as well as by the Physical Component Summary (PCS) measure and the Mental Component Summary (MCS) measure from SF-36 Health Survey.

The analysis of interview guided approach through computer program Dtm-Vic provided two kinds of results: results of lexicometrical analysis and results of correspondence and cluster analyses. Lexicometrical analysis of textual data provided 32365 words all together, 5381 or 16.4% are different words, which shows that examinees mostly use the same language in describing their health condition. 35 most frequently used words are: *sam (od jesam)* (= I am (from "to be")), *ja* (= I), *mi* (= to me), *me* (= me), *imam* (= I have), *nemam* (I don't have), *ne_mogu* (I can't), *mogu* (I can). That shows a high level of individual approach in health evaluation which was insisted upon. Final results of correspondence and cluster analyses were 12 clusters of examinees with different numbers of examinees. They were analyzed and described by examinees' characteristics, as well as by the Physical Component Summary (PCS) measure and the Mental Component Summary (MCS) measure from SF-36 Health Survey.

Clusters obtained by different techniques of analyses (interpretative and computer) cannot be mutually compared as they included different examinees.

Comparison of questionnaire SF-36 analysis and textual data acquired through general interview guided approach shows great similarity in final results. However, textual data, i.e. their analysis, provide much more information, they are more detailed and show examinees' individuality.

The paper shows that general interview guided approach is particularly suitable and useful in epidemiological researches because it provides plentiful information on population sample chosen for this research using the method of intensity sampling.

10. POPIS LITERATURE

- [1.] Gubrium JF, Holstein JA. ur. Handbook of interview research: context & method. Vol. xiii. 2003, Sage: Thousand Oaks, CA. str 981.
- [2.] Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualitative research methods in health technology assessment: a review of the literature. Health Technology Assessment, 1998. 2(16).
- [3.] Reichardt CS, Cook TD. Beyond qualitative versus quantitative methods, U Quantitative methods in evaluation research, Cook TD, Reichardt CS. ur. 1979, Sage: Beverly Hills, CA. str. 7-32.
- [4.] Reichardt CS. Rallis SF. ur. The qualitative-quantitative debate: New perspectives. 1994, Jossey-Bass: San Francisco.
- [5.] de Jong JTVM. van Ommeren M. Toward a Culture-Informed Epidemiology: Combining Qualitative and Quantitative Research in Transcultural Contexts. Transcultural Psychiatry, 2002. 39(4): str. 422-33.
- [6.] Popay J. Qualitative research and the epidemiological imagination: a vital relationship. Gac Sanit, 2003. 3: str. 56-63.
- [7.] Goldsmith MR, Bankhead CR, Austoker J. Synthesising quantitative and qualitative research in evidence-based patient information. Journal of Epidemiology and Community Health, 2007. 61: str. 262-70.
- [8.] Dixon-Woods M, Fitzpatrick R. Qualitative research in systematic reviews. Has established a place for itself. BMJ, 2001. 323(7316): str. 765-6.
- [9.] Giacomini MK. The rocky road: qualitative research as evidence. Evid Based Med, 2001. 6: str. 4-6.
- [10.] Sale JEM, Lohfeld LH, Brazil K. Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research. Quality & Quantity, 2002. 36: str. 43-53.
- [11.] Patton MQ. Qualitative Research and Evaluation Methods. 2002, London: Sage publications.

- [12.] Dharamsi S, Scott I. Quantitative and qualitative research: Received and interpretivist views of science. *Can Fam Physician*, 2009. 55(8): str. 843-4.
- [13.] Newman M, Thompson C, Roberts AP. Helping practitioners understand the contribution of qualitative research to evidence-based practice. *Evid Based Nurs*, 2006. 9: str. 4-7.
- [14.] Curry LA, Nembhard IM, Bradley EH. Qualitative and Mixed Methods Provide Unique Contributions to Outcomes Research. *Circulation*, 2009. 119: str. 1442-52.
- [15.] Saxena H, Wiedermann BL. Is Qualitative Research High Quality Research? *AAP Grand Rounds*, 2009. 21: str. 64.
- [16.] Thomas RM. *Blending Qualitative and Quantitative Research Methods in Theses and Dissertations*. 2003, Thousand Oaks, CA: Corwin Press. str. 256.
- [17.] Kocijan-Hercigonja D, Despot-Lučanin J. Komunikacijske vještine u radu zdravstvenih djelatnika, ur. Hebrang A. 2004, Zagreb: Ministarstvo zdravstva i socijalne skrbi Republike Hrvatske.
- [18.] van den Brink-Muinen A, van Dulmen A, Jung HP, Bensing JM. Do our talks with patients meet their expectations? *The Journal of Family Practice*, 2007. 56(7).
- [19.] Park L, Watson B. *The Influence of Patients' and Health Professionals' Communication on Quality of Care, Health Outcomes, and Treatment Compliance*. 2008, All Academic, Inc.
- [20.] Ware jr JE. SF-36® Health Survey Update. 2009, SF-36.org: Lincoln, RI. <http://www.sf-36.org>
- [21.] Hercigonja-Szekeres M. Analiza tekstnih podataka, U *Medicinska informatika*, Kern J, Petrovečki M. ur. 2009, Medicinska naklada: Zagreb.
- [22.] Republika Hrvatska Državni zavod za statistiku. *Popis stanovništva 2001. 2009*, Državni zavod za statistiku RH. <http://www.dzs.hr/hrv/censuses/census2001/census.htm>
- [23.] Hayes V, Morris J, Wolfe C, Morgan M. The SF-36 health survey questionnaire: is it suitable for use with older adults? *Age Ageing*, 1995. 24(2).

- [24.] Ball AE, Russell EM, Gwyn Seymour D, Primrose WR, Garratt AM. Problems in Using Health Survey Questionnaires in Older Patients with Physical Disabilities. Can Proxies Be Used to Complete the SF-36? *Gerontology*, 2001. 47: str. 334-40.
- [25.] Jureša V, Ivanković D, Vuletić G, Babić-Banaszak A, Srček I, Mastilica M, Budak A. The Croatian Health Survey - SF-36: I. General quality of life assessment. *Collegium Anthropologicum*, 2000. 24: str. 69-78.
- [26.] Vuletić G, Mujkić A. Što čini osobnu kvalitetu života: studija na uzorku hrvatske gradske populacije. *Liječnički vjesnik*, 2002. 124(2): str. 64-70.
- [27.] Maslić-Seršić D, Vuletić G. Psychometric evaluation and establishing norms of Croatian SF-36 health survey: Framework for subjective health research. *CMJ*, 2006. 47(1): str. 95-102.
- [28.] Ruta AD, Abdalla MI, Garratt AM, Coutts A, Russell IT. SF 36 health survey questionnaire: I. Reliability in two patient based studies. *Quality and Safety in Health Care*, 1994. 3: p. 180-5.
- [29.] de Bruin A, Picavet HSJ, Nossikov A. ur. Health interview surveys. Towards international harmonization of methods and instruments. WHO Regional Publications. European Series. Vol. 58. 1996, World Health Organization, Regional Office for Europe: Copenhagen.
- [30.] Bowling A, Bond M, Jenkinson C, Lamping DL. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health Medicine*, 1999. 21: str. 255-70.
- [31.] Schwandt TD, Denzin NK, Lincoln YS. Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics, and social constructionism, U *The SAGE handbook of qualitative research*. 2000, Sage Publications, Inc: Thousand Oaks, CA. str. 189-213.
- [32.] Savage J. Ethnography and health care. *BMJ*, 2000. 31: str. 1400-2.
- [33.] Davey Smith G, Ebrahim S. Epidemiology – is it time to call it a day? *International J Epidemiol*, 2001. 30: str. 1-11.

- [34.] Hands DW, Reconsidering the received view of the "received view": Kant, Kuhn, and the demise of positivist philosophy of science. *Soc Epistemol*, 2003. 17(2-3): str. 169-73.
- [35.] Denzin NK, Lincoln YS, ur. *The Sage Handbook of Qualitative Research*. 2005, Sage: Thousand Oaks, CA.
- [36.] Kinn S, Curzio J, Integrating qualitative and quantitative research methods. *Journal of Research in Nursing*, 2005. 10(3): str. 317-36.
- [37.] Constable R et al. *Ethnography, Observational Research, and Narrative Inquiry*. 2005 <http://writing.colostate.edu/guides/research/observe/>.
- [38.] Lindlof TR, Taylor BC. *Qualitative communication research methods: Second edition*. 2002, Thousand Oaks, CA: Sage Publications, Inc.
- [39.] Patton MQ. *Utilization-Focused Evaluation (U-FE) Checklist*. 2002. <http://www.wmich.edu/evalctr/checklists/ufe.pdf>.
- [40.] Patton MQ, Patrizi P. *Teaching Evaluation Using the Case Method*. *New Directions for Evaluation*, 2005. 105.
- [41.] Manning CD, Schutze H. *Foundations of Statistical Natural Language Processing*. 1999, Cambridge, MA: The MIT Press. str. 680.
- [42.] Carroll J. Parsing, U *The Oxford Handbook of Computational Linguistics*, Mitkov R. ur. 2003, Oxford University Press: Oxford. str. 233-48.
- [43.] Lebart L, Salem A, Berry L. *Exploring Textual Data*. 1998, Dordrecht / Boston / London: Kluwer Academic Publishers. str. 245.
- [44.] Woods A, Fletcher P, Hughes A. *Statistics in language studies*. 1996, Cambridge / New York / Melbourne: Cambridge University Press. str. 326.
- [45.] Sirmakessis S. ur. *Text Mining and its Applications*. *Studies in Fuzziness and Soft Computing*. Vol. 138. 2004, Springer. str. 204.
- [46.] Becue-Bertaut M. SPAD.T: Package for Statistical Analysis of Textual Data. U 13th International Conference on Information Technology Interfaces. 1991. Pula: University Computing Centre Zagreb.
- [47.] Hercigonja-Szekeres M, Kern J, Sonicki Z, Ivanković D, Vuletić S. *Data Organisation and Analysis of Informal Conversation Interview by SPAD.T*. U 16th Conference on Information Technology Interfaces. 1994. Pula: University Computing Centre.

- [48.] Lebart L. Software DtmVic: Exploratory statistical processing of complex data sets comprising both numerical and textual data. 2009: Paris. <http://www.dtm-vic.com>
- [49.] Report, Defining the Patient-Physician Relationship for the 21st Century, in 3rd Annual Disease Management Outcomes Summit. 2003, Johns Hopkins and American Healthways: Phoenix, Arizona.
- [50.] Scott JT, Entwistle VA., Sowden AJ, Watt I., Communicating with children and adolescents about their cancer (Cochrane Review). The Cochrane Library 2, 2001.
- [51.] Tange H., Consultation of medical narratives in the electronic medical record. *Methods Inf Med*, 1999. 38(4-5): str. 289-93.
- [52.] Kern J., Hercigonja-Szekeres M. Medical Informatics - a Challenge for Physician. U EFMI Special Topics Conference Timisoara 2006. 2006. Timisoara.
- [53.] Lyncoln SY, Guba EG. *Naturalistic Inquiry*. 1985: SAGE Publication.
- [54.] Guba EG, Lyncoln YS. Paradigmatic controversis, contradictions and emerging influences. U Denzin NK, Lyncoln YS ur. *The Sage Handbook of Qualitative Research* 3rd ed, 2005: Thousand Oaks, CA, Sage, str. 191-205
- [55.] Patton MQ. R-E-S-P-E-C-T: Remembering Egon Guba. *American Journal of Evaluation*, 2008. 29(3): str. 318-9.
- [56.] Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J. Public Health Med*, 2005. 27(3): str. 281-91.
- [57.] Bowling A, Windsor J. The effects of question order and response-choice on self-rated health status in the English Longitudinal Study of Ageing (ELSA). *J Epidemiol Community Health*, 2008. 62(1): str. 81-5.
- [58.] Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. Patterns of Consent in Epidemiologic Research: Evidence from Over 25,000 Responders. *Am. J. Epidemiol*, 2004. 159(11): str. 1087-94.
- [59.] Hercigonja-Szekeres M, Kern J, Ivanković D, Vuletić S. Textual Analysis in Elderly People. U ISCB-GMDS-99 20th Annual Conference of the International Society for Clinical Biostatistics & 44th Jahrestagung der

Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie. 1999. Heidelberg.

- [60.] Hercigonja-Szekeres M, Ivanković D. Open or closed interviews in life quality evaluation. U Medicon 2001 IX mediterranean Conference on Medical and Biological Engineering and Computing. 2001. Pula: IFMBE.
- [61.] Heim I, Vuletić S, Hercigonja-Szekeres M. The diet of elderly people in Croatia - textual analyses. U Medicon 2001 IX Mediterranean Conference on Medical and Biological Engineering and Computing. 2001. Pula: IFMBE.
- [62.] Stevanović R. Zdravstvene, sociopsihološke i genetske osobitosti autohtonog stanovništva otoka Suska, doktorska disertacija, Medicinski fakultet. 2004, Sveučilište u Zagrebu: Zagreb. str. 118+13.
- [63.] Šogorić S, Middleton J, Lang S, Ivankovic D, Kern J. A naturalistic inquiry on the impact of interventions aiming to improve health and the quality of life in the community. *Social science & medicine*, 2005. 60(1): str. 153-64.
- [64.] Taha J, Sharit J, Czaja S. Use of and Satisfaction With Sources of Health Information Among Older Internet Users and Nonusers. *The Gerontologist*, 2009. 49: str. 663-73.
- [65.] Altman DG, Machin D, Bryant TN, Gardner MJ. *Statistics with Confidence* 2nd ed. Bristol: BMJ Books, 2005.
- [66.] Bland M. *An Introduction to Medical Statistics*. 3rd ed. 2005, Oxford: Oxford University Press.
- [67.] Petrie A, Sabin C. *Medical Statistics at a Glance*. 2nd ed. 2005, Oxford: Blackwell Science Ltd.
- [68.] Jazbec A. *Osnove statistike*. 2008, Zagreb: Šumarski fakultet Sveučilišta u Zagrebu. str. 141.
- [69.] Ziebland S. The short form 36 health status questionnaire: clues from the Oxford region's normative data about its usefulness in measuring health gain in population surveys. *J Epidemiol Community Health*, 1995. 49: str. 102-5.
- [70.] Bowling A, Bond M, Jenkinson C, Lamping DL. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used. *J Public Health Med*, 1999. 21(3): str. 255-70.

- [71.] Giacomini MK, Cook DJ., Users' guides to the medical literature: XXIII. Qualitative research in health care B. What are the results and how do they help me care for my patients? Evidence-Based Medicine Working Group. JAMA, 2000. 284: str. 478-82.
- [72.] Pope C, Ziebland S, Mays N. Analysing qualitative data. BMJ, 2000. 320: str. 114-6.
- [73.] Thorne S. Data analysis in qualitative research. Evidence-Based Nursing, 2000. 3: str. 68-70.
- [74.] Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ, 2000. 321: str. 694-6.
- [75.] Dixon-Woods M, Fitzpatrick R, Roberts K. Including qualitative research in systematic reviews: problems and opportunities. J Eval Clin Pract, 2001. 7: str. 125-33.
- [76.] Mays N, Pope C. Assessment quality in qualitative research. BMJ, 2000. 320: str. 50-2.
- [77.] McKeivitt C, Wolfe C. Quality of life: what, how, why? The views of health care professionals. Qual Ageing, 2002. 3: str. 12-9.
- [78.] Bossuyt P. Better standards for better reporting of RCTS. BMJ, 2001. 322: str. 1317-8.
- [79.] Patton, M.Q., State of the Art in Measuring Development Assistance, in Utilization-Focused Evaluation. 2008: Saint Paul, Minnesota.
- [80.] Biber D, Conrad S, Reppen R. Corpus Linguisticus - Investigating Language Structure and Use. 1998: Cambridge University Press.
- [81.] Mitkov R. ur. The Oxford Handbook of Computational Linguistics. 2004, Oxford University Press: Oxford. str. 763.
- [82.] Taylor AG. The organization of information. 2nd ed. Library and information science text series. 2004, London: Libraries Unlimited. str. 417.
- [83.] Jurafsky D, Martin JH. Speech and Language processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall series in artificial intelligence. 2009, New Jersey: Pearson education. str. 988.

- [84.] (35) Lebart L, Piron M, Morineau A. Statistique Exploratoire Multidimensionnelle (Visualisation et Inference en Fouille de Donnees) 4eme edition. 2006: Dunod. str. 480.
- [85.] Stokes ME, Davis CS, Koch GG. Categorical Data Analysis using the SAS System. 2001, Cary, NC: SAS Institute. str. 629.
- [86.] Morin A, Bosc P, Hebrail G, Lebart L. Bases de données et statistique. 2002: Dunod. str. 368.
- [87.] Greenacre MJ, Blasius J. ur. Multiple Correspondence Analysis and Related Methods. 2006, Chapman-HallCRC.: Boca-Raton, FL. str. 581.
- [88.] Greenacre, MJ. Correspondence Analysis in Practice, 2nd edition. 2007: Chapman & Hall\CRC. str. 280.
- [89.] Friendly M. Visualizing Categorical Data. 2001, Cary, NC: SAS Institute. str. 436.
- [90.] Stein HD, Nadkarni P, Erdos J, Miller PL, Exploring the Degree of Concordance of Coded and Textual Data in Answering Clinical Queries from Clinical Data Repository. JAMIA, 2000. 7: str. 42-54.

11. ŽIVOTOPIS

Rođena sam 4. veljače 1951. Osnovnu školu i Klasičnu gimnaziju završila sam u Zagrebu. Diplomirala sam na zagrebačkom Prirodoslovno-matematičkom fakultetu studij Matematike, stručni smjer Praktična matematika i informatika.

Radila sam kao nastavnik matematike u Školi za osnovno obrazovanje odraslih, a od 1980. kao profesor matematike u Građevinskom obrazovnom centru. Od 1984. u RO Laboratorij mjerne elektronike bila sam direktorica, zatim radila marketing, prodaju, sudjelovala u izradi i održavanju softvera za medicinske aparate te na povezivanju medicinske opreme s računalom. Od 1987. do 1991. bila sam samostalni obrtnik na održavanju medicinskih aparata i obradi podataka računalom. Od 1991. do 2009. radila sam u tvrtki POLIMEDIKA d.o.o. čiji sam i suvlasnik. Od 2009. radim na Visokoj školi Hrvatsko zagorje Krapina.

Dodatno sam se obrazovala u elektronici: godine 1985. izvanrednim školovanjem stekla sam stručnu spremu automatičara u struci automatika.

Na Medicinskom fakultetu Sveučilišta u Zagrebu godine 1988. upisala sam poslijediplomski studij Zdravstveni informacijski sistemi i završila ga 1996. obranom magistarskog rada pod naslovom “Informatički pristup analizi promjena reoencefalografskog signala tokom snimanja”.

Na zagrebačkom Medicinskom fakultetu sudjelujem u poslijediplomskoj nastavi. Na Zdravstvenom veleučilištu u Zagrebu od 2000. do 2007. predavala sam kolegij Informatika. U razdoblju 2003–2006. bila sam u statusu naslovnog nastavnog zvanja predavač visoke škole, te pročelnik Katedre za informatiku. Na Visokoj školi Hrvatsko zagorje Krapina od ak. godine 2007/08. predajem kolegije iz područja informatike.

Autor sam ili suautor četrdesetak stručnih i znanstvenih radova iz područja Informatike, Medicinske informatike i Tekstne analize. Od 1990. godine bila sam aktivni sudionik više od 40 stručnih i znanstvenih domaćih i međunarodnih skupova iz područja informatike, statistike, medicinske informatike te statističke analize tekstnih podataka.

Član sam Hrvatskog društva za medicinsku informatiku, Hrvatskog biometrijskog društva, Hrvatskog društva za medicinsku i biološku tehniku te Udruge HL7 Hrvatska.