

Gordana Budimir, Philip Rachimis, Luka Juršnik

## Analiza odmevnosti slovenskih raziskovalcev na področju medicine v podatkovnih zbirkah WoS in Scopus v obdobju 1970-2016

**Povzetek.** Za evalvacijo slovenskih raziskovalcev na razpisih Javne agencije za raziskovalno dejavnost RS (ARRS) za (so)financiranje raziskovalnih projektov in programov je pomembna odmevnost njihovih znanstvenih člankov v citatnih podatkovnih zbirkah WoS in Scopus. Raziskava člankov slovenskih raziskovalcev na področju medicine v obdobju 1970-2016 je pokazala njihovo veliko odmevnost ter podoben vzorec citiranja in samocitiranja v obeh zbirkah. Ker pa se po metodologiji ARRS samocititi člankov ugotavljajo na osnovi podatkov iz sistema COBISS, ki za razliko od podatkov v zbirkah WoS in Scopus ne vsebujejo vseh avtorjev člankov in vseh njihovih citatov, prihaja do anomalij pri vrednotenju člankov, zaradi katerih so določeni raziskovalci lahko prikrajšani pri razpisih ARRS. Tako je za področje medicine ugotovljeno, da več kot polovica slovenskih raziskovalcev, ki so citirani v zbirkah WoS in Scopus, izkazuje skoraj 8% več citatov po metodologiji ARRS kot v zbirkah WoS in Scopus. Dodatna analiza člankov je pokazala, da do največjih odstopanj prihaja pri člankih v sistemu COBISS v obdobju 2003-2014, ki imajo do 4 avtorje. Na osnovi ugotovitev raziskave bi bilo možno dopolniti bibliografije raziskovalcev v sistemu COBISS, da bi se v največji meri zmanjšala odstopanja pri vrednotenju odmevnosti po metodologiji ARRS.

## Recognition Analysis of Slovene Researchers in the Field of Medicine in WoS and Scopus Databases from 1970 to 2016

**Abstract.** In order to evaluate Slovenian researchers in tenders of Slovenian Research Agency (SRA) for (co)financing research projects and programmes, the recognition of their scientific articles in WoS and Scopus citation databases is required. A study of articles by Slovenian researchers in the field of medicine in the years 1970-2016 has shown their great recognition and a similar pattern of citation and self-citation in both databases. However, according to the SRA methodology self-citations are determined from the data in the COBISS system, which (unlike the data in the WoS and Scopus databases) does not include all authors of all articles and all their citations; anomalies therefore can occur when evaluating the articles, causing some researchers to be deprived in SRA tenders. It was discovered that in the field of medicine more than half of Slovene researchers cited in the WoS and Scopus databases have almost 8% more citations according to SRA methodology than in the WoS and Scopus databases. An additional analysis of articles revealed that the largest anomalies occur in articles in the COBISS system for the years 2003-2014 with up to 4 authors. Based on these findings, the researchers' bibliographies could be supplemented in the COBISS system in order to minimise the anomalies in evaluating the recognition according to the SRA methodology.

■ **Infor Med Slov** 2017; 22(1-2): 1-8

*Institucije avtorjev / Authors' institutions: Institut informacijskih znanosti, Maribor (GB, LJ), Medicinska fakulteta, Maribor (PR).*

*Kontaktna oseba / Contact person: mag. Gordana Budimir, IZUM, Prešernova 17, 2000 Maribor, Slovenija. E pošta / E mail: gordana.budimir@izum.si.*

*Prispelo / Received: 2. 7. 2017. Sprejeto / Accepted: 18. 1. 2018.*

## Uvod

V Sloveniji evalvacijo znanstvene uspešnosti raziskovalcev, registriranih v sistemu SICRIS (<http://www.sicris.si>), izvaja ARRS, ki v svoji metodologiji ocenjevanja prijav za razpise<sup>1</sup> upošteva tudi odmevnost raziskovalcev v zbirkah Web of Science (WoS, <http://isiknowledge.com>) in Scopus (<http://www.scopus.com>), in sicer tako, da pri določanju citatov njihovih člankov izloča njihove samocitate. Medtem ko je citiranost eden od najbolj enostavnih bibliografskih indikatorjev, definicija samocitata ni tako preprosta<sup>2-13</sup>. Sam način definiranja samocitata je zelo pomemben za razlago in primerjavo rezultatov raziskav, v katerih se samocitate analizira z različnih vidikov (avtorski ali dokumentni) in za različne ravni združevanja publikacij ter njihovih citatov, npr. na nacionalnem makro nivoju<sup>7,13</sup> ali na mezo nivoju<sup>8,11,12</sup> za določeno revijo, znanstveno področje, institucijo itn. Na področju medicine je večina raziskav samocitiranosti narejena za določeno (pod)področje ali revijo,<sup>14,15</sup> manj pa je širših raziskav, npr. na nacionalni ravni.<sup>16</sup>

V novejših raziskavah se za samocitate večinoma uporablja definicija, ki sta jo objavila Snyder in Bonzi leta 2009, po kateri samocitat dokumenta nastane, ko množici soavtorjev citiranega članka in soavtorjev člankov, ki citirajo ta članek, nista disjunktni, tj. kadar imata skupnega vsaj enega avtorja.<sup>6</sup> Ta definicija se uporablja tudi za ugotavljanje samocitativ v sistemu Scopus<sup>10</sup>, sistem WoS pa za samocitate nekega članka določa članke, ki citirajo ta članek in so v seznamu zadetkov vhodne iskalne zahteve,<sup>9</sup> zaradi česar v zbirki WoS ni vedno mogoče ugotoviti vseh samocitativ. Poleg tega v teh sistemih velikokrat prihaja do napak pri primerjanju avtorjev člankov zaradi nepopolnih ali neobstoječih podatkov o avtorjih in njihovih ustanovah<sup>17,18</sup> ali zaradi različnih oblik njihovih imen<sup>11-13</sup> (homonimi, spremembe imen).

ARRS pridobiva podatke o citiranosti s portala COBISS/SciMet (<http://scimet.izum.si>), v okviru katerega se agregirajo podatki iz vseh informacijskih virov za potrebe vrednotenja del slovenskih raziskovalcev.<sup>19</sup> Pri tem se njihovi članki v zbirkah WoS in Scopus samodejno povezujejo z ustreznimi članki v sistemu COBISS (<http://www.cobiss.si>), v katerem se od leta 1997 vodijo bibliografije slovenskih raziskovalcev, ki imajo šifro raziskovalca v sistemu SICRIS.<sup>20</sup> Na ta način je v okviru portala COBISS/SciMet omogočeno natančneje ugotavljanje samocitativ, ker se ne primerjajo imena avtorjev pač pa enotni identifikatorji, ki so avtorjem dodeljeni v sistemih SICRIS, COBISS, WoS in Scopus.<sup>21</sup> Po

metodologiji ARRS pa se pri tem primerjajo samo šifre raziskovalca v sistemu SICRIS ali njihova identifikacijska številka v sistemu COBISS iz normativne zbirke imen avtorjev CONOR.SI. Ker pa se v sistemu COBISS pri člankih ne vnašajo vsi avtorji, prihaja do večjih odstopanj glede na dejansko število samocitativ v zbirkah WoS in Scopus, kot je že ugotovljeno v raziskavi samocitiranosti objav slovenskih raziskovalcev v zbirkah WoS in Scopus za obdobje 1996-2013.<sup>21</sup> V naši raziskavi so ta odstopanja podrobneje analizirana za članke na področju medicine, da bi ugotovili morebitne vzorce razlik v samocitiranosti in vpliv teh razlik na evalvacijo znanstvene uspešnosti raziskovalcev na tem področju. Cilji raziskave so torej ugotoviti:

- značilnosti citiranja in samocitiranja člankov slovenskih raziskovalcev na področju medicine v zbirkah WoS in Scopus v obdobju 1970-2016;
- razlike med samocitiranostjo teh člankov po metodologiji ARRS in številom samocitativ v zbirkah WoS in Scopus;
- možnosti dopolnitve bibliografij raziskovalcev v sistemu COBISS, s katerimi bi se zmanjšale ugotovljene razlike.

## Metode

V raziskavo so vključeni članki slovenskih raziskovalcev, registriranih v sistemu SICRIS, na področju medicine, ki jih indeksirata WoS in Scopus v obdobju 1970-2016 in so hkrati obdelani v sistemu COBISS za potrebe vrednotenja raziskovalcev po metodologiji ARRS. Podatki so v surovi obliki pridobljeni s portala COBISS/SciMet in preneseni v relacijsko podatkovno zbirko ter obdelani s posebej razvito skripto SQL in programom R.

Število vseh slovenskih raziskovalcev, registriranih v sistemu SICRIS na področju medicine, je 3.244, od tega je 2.198 (so)avtorjev člankov, indeksiranih v WoS in Scopus. Skupno število njihovih člankov, ki so hkrati obdelani v sistemu COBISS, je 13.522, od tega je 11.275 člankov citiranih v zbirki WoS (za 2.063 raziskovalcev), v zbirki Scopus pa 11.636 člankov (za 2.084 raziskovalcev). Ker je glavni namen raziskave analiza samocitiranosti, so iz raziskave izločeni članki brez citatov, zaradi razlik pri evidentiranju avtorjev člankov v zbirkah WoS in Scopus pa so v analizo zajeti samo članki, ki imajo enako število avtorjev v obeh zbirkah (10.317 člankov za 2.031 raziskovalcev). Za te članke smo naredili analizo dveh vrst samocitativ, definiranih po Snyderju in Bonzi<sup>6</sup> za različne nabore avtorjev, ki se upoštevajo pri primerjanju članka in njegovega citata:

- $AC_W$  oz.  $AC_S$  – samocitati v zbirki WoS oz. Scopus, ki se ugotavljajo s primerjanjem identifikatorjev vseh soavtorjev članka v zbirki WoS oz. Scopus z identifikatorji vseh soavtorjev njegovega citata v zbirki WoS oz. Scopus;
- $AC_{WC}$  oz.  $AC_{SC}$  – samocitati v sistemu COBISS, ki se upoštevajo po metodologiji ARRS in se ugotavljajo s primerjanjem identifikatorjev vseh soavtorjev članka v sistemu COBISS z identifikatorji vseh soavtorjev njegovega citata v sistemu COBISS; če določen citat ni obdelan v sistemu COBISS, se ta citat ne šteje za samocitat članka.

Dodatno so za posamezni članek analizirane še naslednje spremenljivke:

- AU oz.  $AU_C$  – število avtorjev članka v zbirki WoS ali Scopus oz. število avtorjev članka v sistemu COBISS, v katerem je bilo do leta 2014 možno evidentirati največ 11 soavtorjev (prvenstveno slovenskih), od leta 2014 pa je možno evidentirati vse avtorje, vendar se v praksi pri člankih z večjim številom avtorjev evidentirajo samo določeni po izboru katalogizatorjev;
- $TC_W$  oz.  $TC_S$  – število vseh citatov članka v zbirki WoS oz. Scopus.

Za primerjavo citiranosti in samocitiranosti člankov v zbirki WoS in zbirki Scopus je preverjena korelacija opazovanih spremenljivk v teh zbirkah, za ugotavljanje razlik med samocitati v zbirkah WoS oz. Scopus ter samocitati v sistemu COBISS pa je narejena dodatna analiza člankov po številu njihovih avtorjev, viru člankov in letu objave člankov.

## Rezultati

### Značilnosti citiranosti in samocitiranosti člankov

Skupno število citatov v zbirki WoS oz. Scopus za vseh  $N = 10.317$  člankov je  $\Sigma(TC_W) = 196.065$  oz.  $\Sigma(TC_S) = 228.355$ , samocitativ v teh zbirkah pa je  $\Sigma(AC_W) = 37.232$  (18,99%) oz.  $\Sigma(AC_S) = 41.641$  (18,24%). Samocitativ v sistemu COBISS je za članke iz zbirke WoS  $\Sigma(AC_{WC}) = 22.003$  (11,22%), za članke iz zbirke Scopus pa  $\Sigma(AC_{SC}) = 23.501$  (10,29%). Opisne statistike posameznih spremenljivk so podane v tabeli 1, iz katere je razvidna njihova pozitivna asimetrična porazdelitev, kar kaže na to, da je veliko člankov z majhnim številom avtorjev, citatov ali samocitativ, zelo malo člankov pa ima veliko število avtorjev, citatov ali samocitativ. To je razvidno tudi iz tabele 2, v kateri so prikazani podrobnejši percentili za opazovane spremenljivke.

**Tabela 1** Opisne statistike za opazovane spremenljivke ( $N = 10.317$ ).

Sprem.	AVG	SD	Med	Mod	IQR	VR
AU	5,82	8,03	4	3	[3-6]	[1-366]
$AU_C$	4,32	2,75	4	3	[2-6]	[1-36]
$TC_W$	19,00	39,92	8	1	[3-20]	[1-1.074]
$TC_S$	22,13	45,84	10	1	[4-24]	[1-1.264]
$AC_W$	3,61	7,44	1	0	[0-4]	[0-151]
$AC_S$	4,04	8,13	1	0	[0-4]	[0-168]
$AC_{WC}$	2,28	4,29	1	0	[0-2]	[0-83]
$AC_{SC}$	2,13	4,53	1	0	[0-3]	[0-91]

Oznake: *AVG* – povprečje; *SD* – standardni odklon; *Med* – mediana; *Mod* – modus; *IQR* [Q1-Q3] – medčetrtinski razmik; *VR* [Min-Max] – variacijski razmik.

**Tabela 2** Percentili za opazovane spremenljivke ( $N = 10.317$ ).

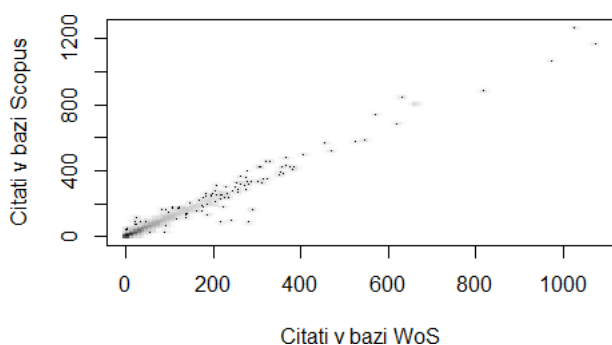
Per.	AU	$AU_C$	$TC_W$	$TC_S$	$AC_W$	$AC_S$	$AC_{WC}$	$AC_{SC}$
5	1	1	1	1	0	0	0	0
10	2	1	1	2	0	0	0	0
15	2	2	2	3	0	0	0	0
20	3	2	3	3	0	0	0	0
<b>25</b>	<b>3</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
30	3	3	4	5	0	0	0	0
35	3	3	5	6	1	1	0	0
40	4	3	6	7	1	1	0	0
45	4	4	7	9	1	1	0	1
<b>50</b>	<b>4</b>	<b>4</b>	<b>8</b>	<b>10</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
55	5	4	10	12	2	2	1	1
60	5	4	12	14	2	2	1	1
65	5	5	14	16	2	3	1	2
70	6	5	16	19	3	4	2	2
<b>75</b>	<b>6</b>	<b>6</b>	<b>20</b>	<b>24</b>	<b>4</b>	<b>4</b>	<b>2</b>	<b>3</b>
80	7	6	25	29	5	6	3	3
85	8	7	31	37	7	7	4	4
90	10	8	42	50	9	10	6	6
95	14	9	67	77	14	16	9	10

Oznake: Per. – percentil.

Glede na omejitve sistema COBISS pri evidentiranju avtorjev člankov je število avtorjev pri člankih v sistemu COBISS ( $AU_C$ ) pričakovano manjše kot pa v zbirkah WoS in Scopus (AU). Povprečno število avtorjev na članek v sistemu COBISS je 4,32, v zbirkah WoS in Scopus pa je nekoliko večje in znaša 5,82. Ne glede na zbirko ima največje število člankov (približno četrtnina) 3 avtorje ( $Mod = 3$ ), polovica jih ima do 4 avtorje ( $Med = 4$ ), tri četrtine člankov pa ima do 6 avtorjev ( $Q3 = 6$ ). Približno 20% člankov ima 4 avtorje, četrtnina člankov ima 5 ali 6 avtorjev, samo 5% člankov pa ima več kot 14 in največ 366 avtorjev v zbirkah WoS in Scopus oz. več kot 9 in največ 36 avtorjev v sistemu COBISS. Razlike v številu avtorjev se pojavijo pri 1.887 člankih; 1.867 jih ima več avtorjev v zbirkah WoS in Scopus, 20 pa več avtorjev v sistemu COBISS, kar je posledica navajanja članov

raziskovalnih skupin kot avtorjev člankov v sistemu COBISS, čeprav niso navedeni na samih publikacijah.

V povprečju je vsak članek v zbirki WoS citiran 19-krat ( $TC_w$ ), v zbirki Scopus pa 22-krat ( $TC_s$ ). V obeh zbirkah ima največ člankov samo en citat, ( $Mod = 1$ ). Iz tabele 2 in glede na majhne razlike med mediano in medčetrtnskim razmikom ( $Med = 8$  oz. 10,  $IQR$  med 3 in 20 oz. med 4 in 24 za zbirko WoS oz. Scopus) pa je razvidno, da je porazdelitev citatov v teh zbirkah zelo podobna; vsaj četrtnina člankov ima do 4 citate, vsaj polovica člankov ima do 10 citatov, vsaj tri četrtine člankov pa ima do 24 citatov. Približno četrtnina člankov ima od 5 do 10 citatov, četrtnina od 12 do 24 citatov, 5% člankov pa jih ima več kot 67 oz. 77 v zbirki WoS oz. Scopus. Opazna je tudi močna pozitivna korelacija ( $r = 0,987$ ) med številom citatov v zbirki WoS in v zbirki Scopus (slika 1).

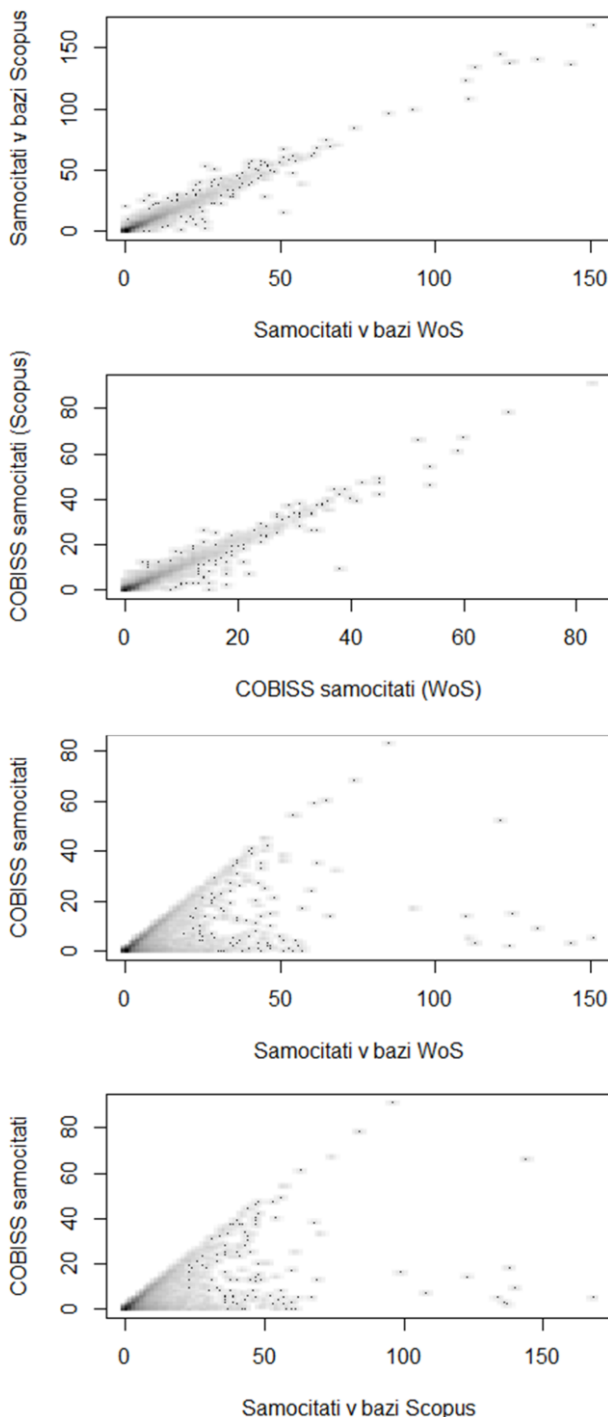


**Slika 1** Korelacija med številom citatov v zbirki WoS in zbirki Scopus ( $N = 10.317$ ).

Povprečno število samocitatov na članek v zbirki WoS oz. Scopus je 3,61 ( $AC_w$ ) oz. 4,04 ( $AC_s$ ), povprečno število samocitatov v sistemu COBISS pa je 2,28 za članke iz zbirke WoS ( $AC_{wc}$ ) oz. 2,13 za članke iz zbirke Scopus ( $AC_{sc}$ ). Ne glede na vrsto samocitatov največje število člankov nima samocitatov ( $Mod = 0$ ) in to več kot četrtnina v zbirkah WoS in Scopus in skoraj polovica v sistemu COBISS. Približno četrtnina člankov ima en samocitat ( $Med = 1$ ), četrtnina člankov pa ima od 2 do 4 samocitate ( $IQR \leq 4$ ). Več kot 5 samocitatov v zbirkah WoS in Scopus ima približno četrtnina člankov, v sistemu COBISS pa približno desetina člankov; več kot 10 samocitatov v zbirkah WoS in Scopus ima manj kot 10% procentov člankov, v sistemu COBISS pa manj kot 5% člankov.

Za samocitate iste vrste v zbirki WoS in Scopus obstaja tudi zelo visoka pozitivna korelacija, kot je prikazano na sliki 2 (za  $AC_w$  in  $AC_s$  je  $r = 0,977$ , za  $AC_{wc}$  in  $AC_{sc}$  je  $r = 0,972$ ). Za različne vrste samocitatov v isti zbirki podatkov pa ta korelacija ni tako visoka (za članke iz zbirke WoS med  $AC_w$  in

$AC_{wc}$  je  $r = 0,670$ , za članke iz zbirke Scopus med  $AC_s$  in  $AC_{sc}$  je  $r = 0,664$ ), kar kaže na to, da med samocitami v zbirki WoS oz. Scopus ter samocitami v sistemu COBISS za članke iz zbirke WoS oz. Scopus prihaja do nezanemarljivih razlik.



**Slika 2** Korelacija med številom samocitatov v zbirkah WoS in Scopus, v zbirki COBISS za zapise iz zbirk WoS in Scopus, v sistemu COBISS in zbirki WoS, ter v sistemu COBISS in zbirki Scopus ( $N = 10.317$ ).

### Razlike v samocitiranosti člankov v sistemu COBISS in zbirkah WoS in Scopus

Glede na definicijo uporabljenih vrst samocitativ se razlike med številom samocitativ v zbirki WoS in Scopus ter številom samocitativ v sistemu COBISS (med  $AC_W$  in  $AC_{WC}$  oz. med  $AC_S$  in  $AC_{SC}$ ) pojavijo pri naslednjih skupinah člankov:

1. pri člankih z različnim številom avtorjev v sistemu COBISS ter v zbirkah WoS in Scopus ( $AUC \neq AU$ ): 1.274 (12,35%) člankov, ki imajo v sistemu COBISS skupaj 9.120 (24,50%) oz. 10.473 (25,15%) samocitativ manj kot v zbirki WoS oz. Scopus;
2. pri člankih z enakim številom avtorjev ( $AUC = AU$ ) in s citati v zbirkah WoS ali Scopus, ki pa niso zavedeni v sistemu COBISS: 1.696 (16,44%) oz. 2.153 (20,87%) člankov iz zbirke WoS oz. Scopus; skupaj 2.365 (22,92%) različnih člankov, ki imajo v sistemu COBISS 5.969 (16,04%) oz. 7.667 (18,41%) samocitativ manj kot v zbirki WoS oz. Scopus.

Člankov, pri katerih prihaja do odstopanj v številu samocitativ različnih vrst, je skupaj 3.639 (35,27%), število raziskovalcev, ki so soavtorji teh člankov, pa je 1.176 (56,43%). Za te članke je število samocitativ v sistemu COBISS manjše od števila samocitativ v zbirkah WoS oz. Scopus za 15.229 (40,90%) oz. 18.140 (43,56%), za toliko pa je število citatov po metodologiji ARRS večje kot v zbirkah WoS oz. Scopus, kar znaša 7,77% oz. 7,95% od skupnega števila citatov v teh zbirkah.

Za podrobnejšo analizo razlik v številu samocitativ so članki iz prve skupine ( $AUC \neq AU$ ) razdeljeni v podskupine po številu avtorjev v sistemu COBISS. Pri tem so izločeni članki, ki imajo v sistemu COBISS več avtorjev kot v zbirkah WoS in Scopus (17 člankov z  $AUC > AU$ ), ker pa je število samocitativ za te članke še vedno večje v zbirkah WoS in Scopus kot v sistemu COBISS ( $AC_W > AC_{WC}$  in  $AC_S > AC_{SC}$ ), ti dodatni avtorji v sistemu COBISS nimajo pomembnejšega vpliva na njihovo samocitiranost. Tabela 3 za podskupine člankov z 1 do 11 avtorji v sistemu COBISS ( $AUC$ ) prikazuje število člankov ( $n$ ), povprečno število avtorjev v zbirkah WoS in Scopus ( $AVG_{AU}$ ), povprečje razlik v številu samocitativ na članek za različne vrste samocitativ v zbirki WoS ( $AVG_W = AVG(AC_W - AC_{WC})$ ) in v zbirki Scopus ( $AVG_S = AVG(AC_S - AC_{SC})$ ) ter povprečje deležev teh razlik ( $AVG_{W\%}$  in  $AVG_{S\%}$ ) glede na skupno število citatov člankov v zbirki WoS ( $TC_W$ ) in v zbirki Scopus ( $TC_S$ ).

**Tabela 3** Razlike v samocitiranosti v zbirkah WoS oz. Scopus in v sistemu COBISS za različne vrste samocitativ glede na število avtorjev v sistemu COBISS ( $N = 1.257$ ).

$AUC$	$n$	$AVG_{AU}$	$AVG_W$	$AVG_S$	$AVG_{W\%}$	$AVG_{S\%}$
1	343	7,36	7,81	8,26	18,69	17,91
2	339	14,04	6,29	7,29	18,19	19,87
3	132	14,83	5,27	6,01	20,22	20,64
4	234	14,26	8,19	9,81	21,27	22,98
5	57	19,30	7,46	8,54	16,95	18,42
6	28	12,14	4,61	5,07	14,44	15,20
7	20	12,90	6,95	6,01	18,75	20,17
8	8	16,38	12,00	14,86	13,42	18,77
9	11	26,09	11,45	12,18	20,32	19,80
10	65	24,58	8,22	9,42	21,87	27,90
11	20	43,55	11,50	13,60	24,64	31,63

Oznake:  $AUC$  – število avtorjev v sistemu COBISS;  $n$  – število člankov;  $AVG_{AU}$  – povprečno število avtorjev na članek v zbirkah WoS in Scopus;  $AVG_W$  oz.  $AVG_S$  – povprečno število dodatnih samocitativ na članek v zbirki WoS oz. Scopus;  $AVG_{W\%}$  oz.  $AVG_{S\%}$  – povprečje deležev števila dodatnih samocitativ v zbirki WoS oz. Scopus glede na skupno število citatov članka v zbirki WoS oz. Scopus.

Podskupine člankov z več kot 11 avtorji (od 12 do 18 avtorjev) v tabeli 3 niso prikazane, skupaj pa zajemajo 17 člankov, ki imajo v povprečju večje število dodatnih samocitativ v zbirkah WoS oz. Scopus (156 oz. 192), kar je posledica veliko večjega števila avtorjev teh člankov v zbirkah WoS in Scopus (do 74) kot v sistemu COBISS. Velika večina člankov (1.048) z razlikami v številu samocitativ ima v sistemu COBISS do 4 avtorje, povprečno pa imajo ti članki v zbirkah WoS in Scopus približno od 6 do 12 avtorjev več, kar se odraža tudi na povečanem številu dodatnih samocitativ na članek v teh zbirkah (približno od 7 do 10 dodatnih samocitativ ali povprečno od 18% do 23% od skupnega števila citatov članka v teh zbirkah). Podobno je pri ostalih 209 člankih z več kot štirimi avtorji v sistemu COBISS, pri katerih pa je povprečen delež dodatnih samocitativ na članek v zbirkah WoS oz. Scopus glede na skupno število citatov članka v teh zbirkah nekoliko večji, kar je posledica dosti večjega števila avtorjev v zbirkah WoS in Scopus. Na primer članki z 11 avtorji v sistemu COBISS, ki imajo v zbirkah WoS in Scopus v povprečju preko 43 avtorjev več, imajo v zbirki Scopus povprečno 13,6 samocitativa več kot v zbirki COBISS, kar je povprečno skoraj 32% od skupnega števila njihovih citatov v zbirki Scopus. Med številom avtorjev članka v sistemu COBISS in številom dodatnih samocitativ članka v zbirkah WoS in Scopus pa nismo ugotovili korelacije.

V tabeli 4 je seznam revij ( $N = 17$ ), pri katerih prihaja do največjih razlik v številu samocitativ za članke iz druge skupine ( $AUC = AU$ ). Iz faktorja vpliva teh revij je razvidno, da odstopanje v številu samocitativ ni povezano s citiranostjo revije. Na primer, revija

*Critical care* ima dosti manjši faktor vpliva ( $IF = 5,053$ ) od revije *The American journal of psychiatry* ( $IF = 14,721$ ), ki pa ima dosti manjše število dodatnih samocitativ v zbirki WoS ( $dAC_W = 39$ ) in zbirki Scopus ( $dAC_S = 52$ ). Po drugi strani pa ima revija *Disability and rehabilitation* ( $IF = 1,985$ ), ki ima tudi manjše število dodatnih samocitativ ( $dAC_W = 47$ ,  $dAC_S = 63$ ) od revije *Critical care* ( $dAC_W = 100$ ,  $dAC_S = 134$ ), veliko manjši faktor vpliva.

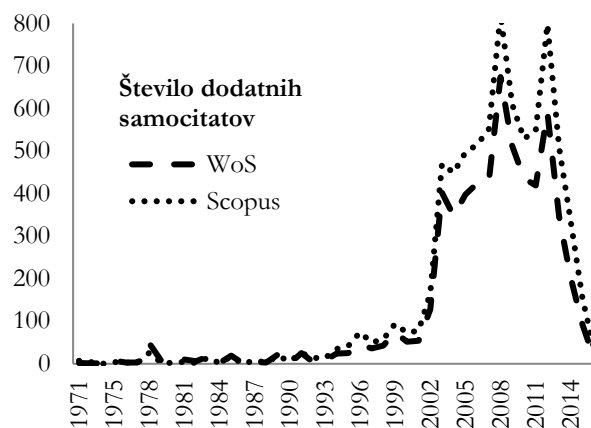
**Tabela 4** Seznam revij ( $N = 17$ ) z največjimi razlikami v številu samocitativ v sistemu COBISS in v zbirkah WoS in Scopus.

ISSN	Naslov revije	IF	$dAC_W$	$dAC_S$
1364-8535	Critical care	5,035	100	134
0098-7484	JAMA	37,684	96	109
1932-6203	PloS one	4,411	83	103
0021-9258	The Journal of biological chemistry	7,716	74	92
0167-8140	Radiotherapy and oncology	5,58	67	72
0270-6474	The journal of neuroscience	8,955	50	67
1439-6319	European journal of applied physiology	2,66	52	66
0963-8288	Disability and rehabilitation	1,985	47	63
0027-8424	Proceedings of the National Academy of Sciences of the United States of America	10,896	48	59
1741-7015	BMC medicine	8,005	49	56
2190-5991	Journal of cachexia, sarcopenia and muscle	7,883	46	56
0013-7227	Endocrinology	5,365	48	55
0002-953X	The American journal of psychiatry	14,721	39	52
0894-1491	GLIA	6,031	44	52
0143-4160	Cell calcium	5,186	26	51
0033-2917	Psychological medicine	6,159	52	51
0959-8049	European journal of cancer	6,163	50	50

Oznake: ISSN – ISSN številka revije; IF – faktor vpliva revije;  $dAC_W$  oz.  $dAC_S$  – dodatni samocitativ v zbirki WoS oz. Scopus.

Na sliki 3 so za članke iz druge skupine ( $AUC = AU$ ) prikazane razlike med številom samocitativ v zbirkah WoS oz. Scopus in številom samocitativ v sistemu COBISS po letih izdaje člankov. Iz slike je razvidno, da število dodatnih samocitativ v zbirki WoS in v zbirki Scopus naglo raste od leta 2003 in nato naglo upade v letu 2015 (kar je pričakovano, ker tudi število citatov v teh zbirkah upada od leta 2015 dalje, od leta 2014 pa je možen tudi vnos vseh avtorjev v sistem COBISS), največ dodatnih samocitativ (preko 500 na leto) pa imajo članki iz leta 2008 in 2012. Obstaja

skoraj popolna linearna korelacija med številom dodatnih samocitativ v zbirki WoS in v zbirki Scopus ( $r = 0,996$ ), ni pa omembe vredne korelacije med letom izdaje članka in dodatnimi samocitativ v zbirkah WoS oz. Scopus.



**Slika 3** Število dodatnih samocitativ v zbirki WoS in zbirki Scopus po letih izdaje člankov ( $N = 2.365$ ).

## Razprava

Več kot dve tretjini (67,8%) vseh raziskovalcev, registriranih v sistemu SICRIS na področju medicine, je (so)avtorjev člankov, indeksiranih v zbirkah WoS ali Scopus, velika večina teh člankov pa je citiranih v zbirki WoS (83,38%) in zbirki Scopus (86,04%), kar kaže na veliko odmevnost objav slovenskih raziskovalcev na tem področju. Praktično vsak od raziskovalcev ima vsaj en citat v zbirki WoS (93,9%) ali Scopus (94,8%). Ugotovili smo zelo visoko korelacijo med citiranostjo člankov v teh zbirkah in podoben vzorec citiranja (največ člankov ima en citat, približno četrtina člankov ima do 4 citate, četrtina od 5 do 10 citatov, četrtina od 12 do 24 citatov in manj kot 5% člankov ima več kot 77 citatov). Majhne razlike v citiranosti v zbirkah WoS in Scopus izhajajo iz razlik v pokritosti revij, v katerih objavljajo naši raziskovalci na področju medicine, in različnega časovnega intervala evidentiranja citatov v teh sistemih<sup>22</sup>.

Ne glede na vrsto samocitativ večina citiranih člankov slovenskih raziskovalcev na področju medicine v zbirkah WoS in Scopus nima samocitativ ali jih ima zelo malo (do 4 samocitate), več kot 10 samocitativ pa ima zelo malo člankov, kar kaže na pomembno odmevnost teh člankov izven slovenskih znanstvenih krogov. Delež samocitativ v zbirkah WoS in Scopus glede na skupno število citatov v teh zbirkah se praktično ne razlikujeta (manj kot 1%), podobno pa velja tudi za deleže samocitativ v sistemu

COBISS za članke v teh zbirkah. Močna korelacija samocitiranosti člankov v zbirkah WoS in Scopus za isto vrsto samocitativ kaže na podoben vzorec samocitiranosti v teh zbirkah. Raziskava pa je pokazala, da za različne vrste samocitativ do razlik v samocitiranosti člankov iz določene zbirke prihaja pri približno tretjini člankov, katerih soavtorji predstavljajo več kot polovico slovenskih raziskovalcev na področju medicine s citati v zbirkah WoS ali Scopus. Ker ARRS upošteva samocitate iz sistema COBISS in ne iz zbirk WoS in Scopus, za te raziskovalce prihaja do anomalij pri evalvaciji njihove znanstvene uspešnosti. Tem raziskovalcem je upoštevano za več kot 40% samocitativ manj, posledično pa jim je zato izkazana višja citiranost za približno 8% glede na dejansko citiranost v zbirkah WoS in Scopus. Tako je zaradi specifičnosti obdelave člankov v sistemu COBISS pri evalvaciji raziskovalne uspešnosti po metodologiji ARRS v neenakopravnem položaju skoraj polovica slovenskih raziskovalcev na področju medicine.

Da bi ugotovili, pri katerih člankih prihaja do največjih odstopanj, smo naredili dodatno analizo razlik v samocitiranosti člankov, ki imajo enako število avtorjev v sistemu COBISS in v zbirkah WoS in Scopus, ter samocitiranosti člankov, ki imajo različno število avtorjev, nimajo pa vseh citatov obdelanih v sistemu COBISS. Člankov v drugi skupini je več kot v prvi, do večjih razlik pa prihaja pri člankih iz prve skupine, kar pomeni, da ima neevidentiranje vseh avtorjev člankov v sistemu COBISS večji vpliv na povečano citiranost po metodologiji ARRS kot pa manjkajoči citati v sistemu COBISS. Največja odstopanja so bila ugotovljena pri člankih, ki imajo v sistemu COBISS do 4 avtorje, pri katerih se v povprečju upošteva od 7 do 10 citatov več po metodologiji ARRS. Pri člankih z manjkajočimi citati v sistemu COBISS pa so bila ugotovljena največja odstopanja pri člankih iz obdobja 2003-2014, ki se jim po metodologiji ARRS upošteva skupaj preko 300 dodatnih citatov letno (največ člankom iz leta 2008 in 2012; glede na faktor vpliva revije, ki je zelo pomemben pri evalvaciji raziskovalcev, pa so pomembna odstopanja ugotovljena pri člankih iz revij *JAMA*, *Proceedings of the National Academy of Sciences of the United States of America* in *The American journal of psychiatry*).

Linearne korelacije med zmanjšanjem števila samocitativ člankov v sistemu COBISS in številom njihovih avtorjev, letom izida članka ali faktorjem vpliva revije, iz katere so članki, nismo ugotovili, kar kaže na to, da do razlik v citiranosti po metodologiji ARRS ne prihaja sistematično, pač pa glede na prakso katalogizatorjev v sistemu COBISS pri izbiri avtorjev

in citatov člankov, ki bodo evidentirani v sistemu COBISS. To pomeni, da bi z ureditvijo bibliografij raziskovalcev na področju medicine v sistemu COBISS z dodatnim vnosom avtorjev ali evidenco manjkajočih citatov v sistemu COBISS predvsem pri člankih, pri katerih so ugotovljena največja oz. pomembnejša odstopanja pri številu samocitativ, znatno zmanjšali anomalije pri evalvaciji raziskovalcev po metodologiji ARRS na tem področju. Analiza v raziskavi je bila omejena na število avtorjev, leto izida in faktor vpliva revije, za natančnejšo določitev člankov, pri katerih te razlike najbolj vplivajo na evalvacijo po metodologiji ARRS, pa bi bilo treba analizirati dodatne spremenljivke (npr. tipologije del ali založnike).

Odstopanja samocitiranosti, do katerih prihaja pri evalvaciji raziskovalne uspešnosti, zaradi česar so določeni raziskovalci na področju medicine lahko prikrajšani pri razpisih ARRS, smo ugotovili na podlagi dveh vrst samocitativ, ki se določajo glede na vse avtorje člankov in njihovih citatov v zbirkah WoS in Scopus ali pa v sistemu COBISS. Za podrobnejšo analizo vpliva samocitiranosti na evalvacijo raziskovalne uspešnosti bi raziskavo morali razširiti na dodatne vrste samocitativ, npr. normirane glede na število avtorjev, kot je predlagano v raziskavi medicinskih člankov slovenskih raziskovalcev v obdobju 1986-2007,<sup>16</sup> ter na dodatne bibliometrične indikatorje, kot so npr. različne vrste h-indeksov.

## Zaključek

Citiranost znanstvenih objav je še vedno eden od najenostavnejših in najbolj uporabljenih bibliografskih kazalnikov za spremljanje znanstvene uspešnosti raziskovalcev in njihovo vrednotenje. Raziskava člankov slovenskih raziskovalcev na področju medicine v zbirkah WoS in Scopus v obdobju 1970-2016 je pokazala njihovo veliko odmevnost v znanstveni literaturi, saj je citiranih več kot 80% objav, citatov pa so deležni praktično vsi raziskovalci in ne samo posamezniki ali skupine raziskovalcev. Vzorec citiranosti in samocitiranosti teh raziskovalcev je zelo podoben v zbirkah WoS in Scopus, večina člankov ima tako do 24 citatov in do 4 samocitate, zelo malo člankov pa ima več kot 77 citatov in več kot 10 samocitativ.

Toda po metodologiji ARRS za evalvacijo raziskovalne uspešnosti se samocitati ne ugotavljajo v zbirkah WoS in Scopus, pač pa v sistemu COBISS, v katerem niso evidentirani vsi avtorji člankov in vsi njihovi citati. Raziskava je pokazala, da zato več kot polovica raziskovalcev izkazuje za skoraj 8% večjo citiranost pri evalvaciji raziskovalne uspešnosti po

metodologiji ARRS kot je ta dejansko v zbirkah WoS in Scopus. Do odstopanj prihaja predvsem pri člankih, ki imajo v sistemu COBISS manj kot 4 avtorje (v povprečju prejmejo 7-10 citatov več) in članke v obdobju 2003-2014, za katere citati v sistemu COBISS niso obdelani. Pri tem je bilo ugotovljeno, da ima na odstopanja največ vpliva katalogizacijska praksa pri obdelavi člankov v sistemu COBISS (izbira avtorjev in njihovih citatov). Da bi se izognili negativnim vplivom obdelave člankov na vrednotenje raziskovalcev po metodologiji ARRS, bi bilo treba ustrezno urediti njihove bibliografije v sistemu COBISS (predvsem članke, pri katerih so ugotovljene največje razlike), za ugotovljane najprimernejše metodologije merjenja odmevnosti člankov, ki bi v največji meri zagotovila enakovredne pogoje za vse raziskovalce pri njihovem vrednotenju, pa bi raziskavo morali razširiti na analizo dodatnih vrst samocitatov in bibliometričnih kazalnikov.

## Reference

- ARRS: Metodologija ocenjevanja prijav za razpise. <http://www.arrs.gov.si/sl/akti/metod-skupna-15-8.asp> (1. 12. 2016).
- Garfield E: Citation analysis as a tool in journal evaluation. *Science* 1972; 178(4060): 471-479. doi: 10.1126/science.178.4060.471
- Tagliacozzo R: Self-citation in scientific literature. *J Doc* 1977; 33(4): 251-265. doi: 10.1108/eb026644.
- Garfield E: Is citation analysis a legitimate evaluation tool? *Scientometrics* 1979; 1(4): 359-375. doi: 10.1007/BF02019306
- Lawani SM: On the heterogeneity and classification of author self-citations. *J Am Soc Inf Sci* 1982; 33(5): 281-284. doi: 10.1002/asi.4630330506
- Snyder H, Bonzi S: Patterns of self-citation across disciplines. *J Inf Sci* 1998; 24(6): 431-435. doi: 10.1177/016555159802400606
- Aksnes DW: A macro study of self-citation. *Scientometrics* 2003; 56(2): 235-246. doi: 10.1023/A:1021919228368
- Costas R, van Leeuwen TN, Bordons M: Self-citations at the meso and individual levels: effects of different calculation methods. *Scientometrics* 2010; 82(3): 517-537. doi: 10.1007/s11192-010-0187-7
- Carley S, Porter AL, Youtie J: Toward a more precise definition of self-citation. *Scientometrics* 2013; 94(2): 777-780. doi: 10.1007/s11192-012-0745-2
- Ioannidis JPA: A generalized view of self-citation: direct, co-author, collaborative, and coercive induced self-citation. *J Psychosom Res* 2015; 78(1): 7-11. doi: 10.1016/j.jpsychores.2014.11.008
- Glänzel W, Thijs B, Schlemmer B: A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics* 2004; 59(1): 63-77. doi: 10.1023/B:SCIE.0000013299.38210.74
- Shah TA, Gul S, Gaur R: Authors self-citation behaviour in the field of library and information science. *AJIM* 2015; 67(4): 458-468. doi: 10.1108/AJIM-10-2014-0134
- Fowler JH, Aksnes DW: Does self-citations pay?. *Scientometrics* 2007; 72(3): 427-437. doi: 10.1007/s11192-007-1777-2
- Tolisano AM, Song SA, Cable BB: Author self-citation in the otolaryngology literature. *Otolaryngol-Head Neck Surg* 2016; 154(2): 282-286. doi: 10.1177/0194599815616111
- Sewell JM, Adejoro OO, Fleck JR et al.: Factors associated with the Journal Impact Factor (JIF) for Urology and Nephrology Journals. *Int Braz J Urol* 2016; 41(6): 1058-1066. doi: 10.1590/S1677-5538.IBJU.2014.0497
- Blagus R, Leskošek BL, Stare J: Comparison of bibliometric measures for assessing relative importance of researchers. *Scientometrics* 2015; 105(3): 1743-1762. doi: 10.1007/s11192-015-1622-6
- Vieira ES, Gomes JANF: A comparison of Scopus and Web of Science for a typical university. *Scientometrics* 2009; 81(2): 587-600. doi: 10.1007/s11192-009-2178-0
- Jacsó P: Errors of omission and their implications for computing scientometric measures in evaluating the publishing productivity and impact of countries. *Online information review* 2009; 33(2): 376-385. doi: 10.1108/14684520910951276
- Budimir G, Juršnik J: COBISS/SciMet. *Organizacija znanja* 2015; 20(2): 68-74. doi: 10.3359/oz150268
- Demšar F, Južnič P: Transparency of research policy and the role of librarian. *Journal of librarianship and information science* 2014; 46(2): 139-147. doi: 10.1177/0961000613503002
- Budimir G, Juršnik L, Rachimis P: Samocitiranost objav slovenskih raziskovalcev v podatkovnih zbirkah Web of Science in Scopus za obdobje 1996-2013. *Knjižnica* 2016; 60(1): 45-60.
- Bartol T, Budimir G, Dekleva-Smrekar D et al.: Assessment of research fields in Scopus and Web of Science in the view of national research evaluation in Slovenia. *Scientometrics* 2014; 98(2): 1491-1504. doi: 10.1007/s11192-013-1148-8