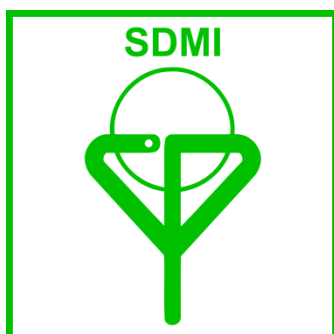


INFORMATICA **M**EDICA **S**LOVENICA



Časopis Slovenskega društva za medicinsko informatiko
Journal of the Slovenian Medical Informatics Association

LETNIK / VOLUME **22 (2017)**, ŠTEVILKA / NO. **1-2**

ISSN 1318-2129 (tiskana izdaja / printed edition)

ISSN 1318-2145 (spletna izdaja / online edition)

<http://ims.mf.uni-lj.si/>

Editor in Chief / Glavni urednik

Gaj Vidmar

Associate Editors / Souredniki

Riccardo Bellazzi
Bjoern Bergh
Jure Dimec
Peter Juvan
Brane Leskošek
Blaž Zupan

Technical and Web Editor / Tehnični in spletni urednik

Peter Juvan

Editorial Board Members / Člani uredniškega odbora

Gregor Anderluh
Janez Demšar
Emil Hudomalj
Izet Mašič
Marjan Mihelin
Mojca Paulin
Uroš Petrovič
Primož Zihlerl

Former Editors in Chief / Bivši glavni uredniki

Martin Bigec
Peter Kokol
Janez Stare

About the Journal

Informatica Medica Slovenica (IMS) is an interdisciplinary professional journal that publishes contributions from the field of medical informatics, health informatics, nursing informatics and bioinformatics. Journal publishes scientific and technical papers and various reports and news. Especially welcome are the papers introducing new applications or achievements.

IMS is the official journal of the Slovenian Medical Informatics Association (SIMIA). It is published two times a year in print (ISSN 1318-2129) and electronic editions (ISSN 1318-2145, available at <http://ims.mf.uni-lj.si>). Prospective authors should send their contributions in Slovenian, English or other acceptable language electronically to the Editor in Chief Assoc.Prof. Gaj Vidmar, PhD. Detailed instructions for authors are available online.

The journal subscription is a part of the membership in the SIMIA. Information about the membership or subscription to the journal is available from the secretary of the SIMIA (Mrs. Mojca Paulin, mojca.paulin@gmail.com).

O reviji

Informatica Medica Slovenica (IMS) je interdisciplinarna strokovna revija, ki objavlja prispevke s področja medicinske informatike, informatike v zdravstvu in zdravstveni negi, ter bioinformatike. Revija objavlja strokovne prispevke, znanstvene razprave, poročila o aplikacijah ter uvajanju informatike na področjih medicine in zdravstva, pregledne članke in poročila. Še posebej so dobrodošli prispevki, ki obravnavajo nove in aktualne teme iz naštetih področij.

IMS je revija Slovenskega društva za medicinsko informatiko (SDMI). Izhaja dvakrat letno v tiskani (ISSN 1318-2129) in elektronski obliki (ISSN 1318-2145, dostopna na naslovu <http://ims.mf.uni-lj.si>). Avtorji člankov naj svoje prispevke pošljejo v elektronski obliki glavnemu uredniku izr.prof.dr. Gaju Vidmarju. Podrobnejša navodila so dosegljiva na spletni strani revije.

Revijo prejemaajo vsi člani SDMI. Informacije o članstvu v društvu oziroma o naročanju na revijo so dostopne na tajništvu SDMI (Mojca Paulin, mojca.paulin@gmail.com).

Contents

Research Papers

- 1 Gordana Budimir, Philip Rachimis, Luka Juršnik**
Recognition Analysis of Slovene Researchers in the Field of Medicine in WoS and Scopus Databases from 1970 to 2016
- 9 Dalibor Stanimirović, Eva Murko**
Rare Disease Registry: Systemic Starting Points and Development Guidelines

Research Review Paper

- 22 Matevž Kastrin, Dimitar Hristovski, Andrej Kastrin**
Literature-Based Discovery in the Field of Life Sciences

SIMLA Bulletin

- 32 Drago Rudel**
Report on Active Participation at the Med-e-Tel 2017 International Congress in Luxembourg
- 33 Ema Dornik**
Digital Health – Opportunities and Challenges: Report from the Meeting of the Nursing Informatics Section 2017

Vsebina

Izvirna znanstvena članka

- 1 Gordana Budimir, Philip Rachimis, Luka Juršnik**
Analiza odmevnosti slovenskih raziskovalcev na področju medicine v podatkovnih zbirkah WoS in Scopus v obdobju 1970-2016
- 9 Dalibor Stanimirović, Eva Murko**
Register redkih bolezni: sistemska izhodišča in razvojne usmeritve

Pregledni znanstveni članek

- 22 Matevž Kastrin, Dimitar Hristovski, Andrej Kastrin**
Odkrivanje zakonitosti iz literature na področju znanosti o življenju

Bilten SDMI

- 32 Drago Rudel**
Poročilo o aktivni udeležbi na mednarodnem kongresu Med-e-Tel 2017 v Luksemburgu
- 33 Ema Dornik**
Digitalno zdravje – priložnosti in izzivi: poročilo s srečanja Sekcije za informatiko v zdravstveni negi 2017

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 samocitati 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@izj.um.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 razpise¹ 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 preprosta²⁻¹³. 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^{7,13} ali na mezo nivoju^{8,11,12} za določeno revijo, znanstveno področje, institucijo itn. Na področju medicine je večina raziskav samocitiranosti narejena za določeno (pod)področje ali revijo,^{14,15} manj pa je širših raziskav, npr. na nacionalni ravni.¹⁶

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.⁶ Ta definicija se uporablja tudi za ugotavljanje samocitativ v sistemu Scopus¹⁰, sistem WoS pa za samocitate nekega članka določa članke, ki citirajo ta članek in so v seznamu zadetkov vhodne iskalne zahteve,⁹ 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^{17,18} ali zaradi različnih oblik njihovih imen¹¹⁻¹³ (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.¹⁹ 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.²⁰ 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.²¹ 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 1996-2013.²¹ 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 Bonzi⁶ 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. $\Sigma(TC_S) = 228.355$, samocitatov v teh zbirkah pa je $\Sigma(AC_W) = 37.232$ (18,99%) oz. $\Sigma(AC_S) = 41.641$ (18,24%). Samocitatov 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 samocitatov, zelo malo člankov pa ima veliko število avtorjev, citatov ali samocitatov. 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] – medčetrtnski razmik; VR [Min-Max] – variacijski razmik.

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

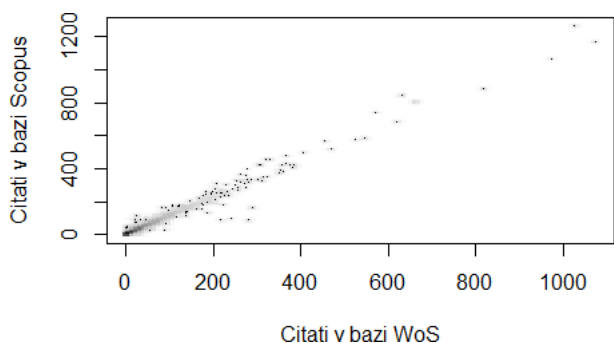
Per.	AU	AU_C	TC_W	TC_S	AC_W	AC_S	AC_{WC}	AC_{SC}
5	1	1	1	1	0	0	0	0
10	2	1	1	2	0	0	0	0
15	2	2	2	3	0	0	0	0
20	3	2	3	3	0	0	0	0
25	3	2	3	4	0	0	0	0
30	3	3	4	5	0	0	0	0
35	3	3	5	6	1	1	0	0
40	4	3	6	7	1	1	0	0
45	4	4	7	9	1	1	0	1
50	4	4	8	10	1	1	1	1
55	5	4	10	12	2	2	1	1
60	5	4	12	14	2	2	1	1
65	5	5	14	16	2	3	1	2
70	6	5	16	19	3	4	2	2
75	6	6	20	24	4	4	2	3
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 četrtina) 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, četrtina 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 četrtina člankov ima do 4 citate, vsaj polovica člankov ima do 10 citatov, vsaj tri četrtine člankov pa ima do 24 citatov. Približno četrtina člankov ima od 5 do 10 citatov, četrtina 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).

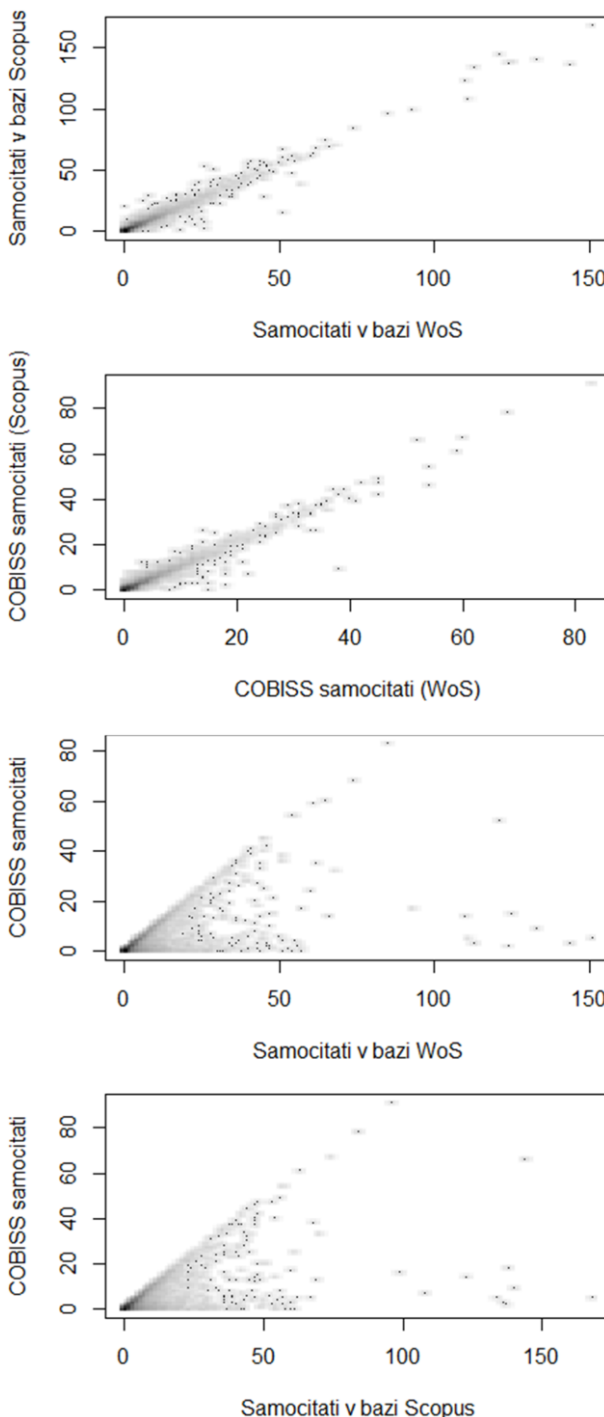


Slika 1 Korelacija med številom citatov v zbirki WoS in zbirki Scopus ($N = 10.317$).

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 četrtina v zbirkah WoS in Scopus in skoraj polovica v sistemu COBISS. Približno četrtina člankov ima en samocitat ($Med = 1$), četrtina člankov pa ima od 2 do 4 samocitate ($IQR \leq 4$). Več kot 5 samocitatov v zbirkah WoS in Scopus ima približno četrtina 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 samocitativ 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.

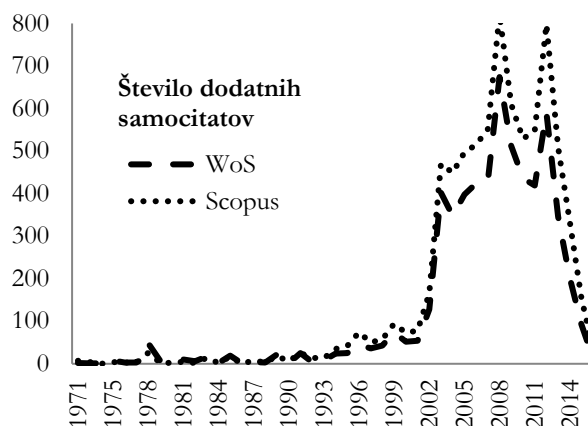
Tabela 4 Seznam revij ($N = 17$) z največjimi razlikami v številu samocitativ v sistemu COBISS in v zbirkah WoS in Scopus.

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 sistemih²².

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 zbirke 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 citativ 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 citativ 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 citativ 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 citativ 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 citativ 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 citativ 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 1986-2007,¹⁶ 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, citativ 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 citativ in do 4 samocitate, zelo malo člankov pa ima več kot 77 citativ 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 avtorjev (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 ugotavljane 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

Dalibor Stanimirović, Eva Murko

Register redkih bolezni: sistemska izhodišča in razvojne usmeritve

Povzetek. Redke bolezni so posamič redke – v Evropi jih definiramo s pojavnostjo manj kot 1 pacient na 2.000 (oz. 5 na 10.000) oseb, vendar je njihovo skupno število ocenjeno na preko 7.000 bolezni, torej so kumulativno pogoste. V Sloveniji je po grobih ocenah okoli 150.000 pacientov z redko boleznijo. Za ustrezno oblikovanje zdravstvenih politik in vrednotenje učinkov zdravstvenih politik ter spremljanje pacientov z redko boleznijo je nujno potrebno poznati epidemiološke podatke, vključno s podatki o zdravljenju in uporabi zdravil sirot. Registri predstavljajo učinkovit način za nepristransko zbiranje podatkov, spremljanje področja redkih bolezni ter epidemiološke ali klinične raziskave in lahko v veliki meri pripomorejo k izboljšanju zdravstvenega varstva ter načrtovanja zdravstvene oskrbe pacientov. Zaradi teh razlogov je razvoj registrov redkih bolezni ena izmed prioritiet na področju spremljanja in obvladovanja redkih bolezni v Evropski uniji. Nacionalni register redkih bolezni v Sloveniji še ni vzpostavljen, čeprav je njegova vzpostavitev navedena med ključnimi aktivnostmi Načrta dela na področju redkih bolezni v Sloveniji. Prispevek analizira trenutno stanje na področju redkih bolezni v Sloveniji, identificira izzive za vzpostavitev registra redkih bolezni ter ob koncu podaja praktično uporabne usmeritve za razvoj registra redkih bolezni.

Rare Disease Registry: Systemic Starting Points and Development Guidelines

Abstract. Rare diseases are individually rare – in Europe, they are defined with the incidence of less than 1 patient per 2,000 (or 5 per 10,000) people, but their total number is estimated to over 7,000 diseases, so they are cumulatively frequent. According to rough estimates, there are around 150,000 patients with rare diseases in Slovenia. In order to conduct a proper policy-making process, evaluate the effects of health policies and to monitor patients with rare diseases, it is essential to have epidemiological data, including information on the treatment and use of orphan medicinal products. Registries represent an effective way of impartial data collection, monitoring of rare diseases and epidemiological or clinical research, and can greatly contribute to improving healthcare and planning healthcare process for patients. For these reasons, the development of rare disease registries is one of the priorities in the field of monitoring and management of rare diseases in the European Union. The national rare disease registry in Slovenia has not been established yet, although its establishment is cited as one of the key activities of the Work Plan in the Field of Rare Diseases in Slovenia. The article analyses the current situation in the field of rare diseases in Slovenia, identifies the challenges for establishing a rare disease registry, and ultimately provides practically applicable guidelines for the development of the rare disease registry.

■ **Infor Med Slov** 2017; 22(1-2): 9-21

Institucije avtorjev / Authors' institutions: Nacionalni inštitut za javno zdravje, Ljubljana (DS, EM).

*Kontaktna oseba / Contact person: doc. dr. Dalibor Stanimirović, Nacionalni inštitut za javno zdravje, Trubarjeva 2, 1000 Ljubljana, Slovenija.
E-pošta / E-mail: dalibor.stanimirovic@nijz.si.*

Prispelo / Received: 11. 10. 2017. Sprejeto / Accepted: 18. 1. 2018.

Uvod

Izraz »redke bolezni« izhaja iz ugotovitve, da imajo številne bolezni, zlasti presnovne, nekaj skupnih značilnosti pri diagnosticiranju in zdravstveni oskrbi, čeprav se zelo razlikujejo po klinični sliki in načinu zdravljenja. Te bolezni niso le redke, ampak pogosto tudi kronične, napredujoče in degenerativne. V 80 odstotkih primerov so genetskega izvora, v 75 odstotkih primerov prizadenejo otroke in so pogosto neozdravljive. Večina redkih bolezni močno vpliva na kakovost življenja prizadete osebe in njenih bližnjih. V svetu ni skupne opredelitve redke bolezni. Akcijski program Evropske unije (EU) glede redkih bolezni, vključno z genetskimi boleznimi, šteje bolezni za redke, kadar prizadenejo največ 5 na 10.000 oseb,¹ v ZDA pa kot stanje ali bolezen, ki prizadene manj kot 200.000 posameznikov.² Takih bolezni je med 5.000 in 8.000 in prizadenejo od 6 do 8 odstotkov prebivalstva.¹ Skupaj te bolezni prizadenejo 30 milijonov ljudi v Evropi, 25 milijonov v ZDA, in okoli 400 milijonov po celem svetu. V Sloveniji je po grobih ocenah okoli 150.000 pacientov z redko boleznijo.³ Kljub odsotnosti natančnih epidemioloških podatkov o stanju v Republiki Sloveniji je na podlagi zgoraj navedenih števil jasno, da gre za pomembno področje zdravstvene oskrbe tudi pri nas. Redkost teh bolezni pogosto povzroča težave pri njihovem prepoznavanju. Ker so kronične, pogosto zahtevajo celovito medicinsko obravnavo. Spremljanje pacientov z redkimi boleznimi omogoča boljše oskrbo zlasti z vidika usklajenosti vpletenih izvajalcev zdravstvene dejavnosti in z vidika dostopnosti epidemioloških podatkov.¹

Številne redke bolezni imajo dolg seznam skupnih značilnosti, ki predstavljajo resne izzive zdravstvenim delavcem:

- težavna in dolgotrajna diagnostika;
- diagnoze Mednarodne klasifikacije bolezni (MKB, ang. International Classification of Diseases, ICD) so neustrezno opredeljene;
- vzročni molekularni ali fiziološki mehanizmi so neznani;
- primanjkuje specializirane in koordinirane oskrbe;
- zdravljenje je pogosto kompleksno;
- malo je longitudinalnih zbirk podatkov;
- razvoj novih zdravil in zdravljenj je fragmentiran in počasen;
- strategijam presejanja manjka učinkovitosti;
- obseg in zmogljivost večine registrov in podatkovnih zbirk sta omejena;
- standardi oskrbe za zdravljenje in rehabilitacijo niso podprti s trdnimi dokazi, ker so raziskave izvedene na majhnih vzorcih.

Zaradi navedenih izzivov redke bolezni predstavljajo velike težave na področju javnega zdravja in terjajo celovit ter ustrezno organiziran pristop. Problematika sedanje obravnave nekaterih pacientov z redkimi boleznimi sega na področja zahtevne, drage in pozne diagnostike, visokih stroškov zdravljenja, dolgotrajne rehabilitacije in še ne organizirane pomoči svojcem pacientov z redkimi boleznimi. Večina redkih bolezni se začne v otroštvu in imajo lahko uničujoče posledice za zdravje, vključno s prezgodnjo smrtjo. Resno lahko prizadenejo življenja skrbnikov, njihov ekonomski vpliv je precejšen za starše in družbo na splošno. Čeprav za številne redke bolezni ni zdravila, se lahko nekatere bolezni z medicinsko intervencijo prepreči ali nadzoruje, življenjska doba pacientov pa se lahko podaljša. Pomembna podpora javnozdravstvenim pristopom k redkim boleznim je prišla preko organizacij pacientov in skozi napore politike, ki vključujejo sprejem Zakona o zdravilih sirotah leta 1983 in zakona o redkih boleznih v ZDA leta 2002, ki sta spodbudila raziskovanje in razvoj ustreznih zdravil.⁴

Eden izmed temeljnih ukrepov za vzpostavitev sistemskega nadzora nad celotnim področjem redkih bolezni in obvladovanja navedenih izzivov je vzpostavitev registra redkih bolezni. Strateško pomembnost registrov redkih bolezni je v dokumentu »Redke bolezni: Evropski izziv« iz leta 2008 poudarila tudi Evropska Komisija,⁵ pomembnost tovrstnih registrov je poudarjena tudi v različnih razvojnih in strateških dokumentih nekdanjega EUCERD-a (ang. European Union Committee of Experts on Rare Diseases), oz. zdajšnje Strokovne skupine za redke bolezni pri Evropski komisiji (ang. EC Expert Group on Rare Diseases). Nacionalni registri redkih bolezni že obstajajo v Franciji,⁶ Italiji,⁷ Španiji,⁸ na Slovaškem⁹ in v Belgiji, v pripravi sta bolgarski in švedski nacionalni register redkih bolezni.¹⁰ V Republiki Sloveniji nimamo natančnih epidemioloških ocen o pacientih z redko boleznijo. Posamezne klinike univerzitetnih kliničnih centrov ali specializirani centri splošnih bolnišnic sicer vodijo zbirke podatkov o pacientih za klinične namene, kar pa ne zadošča za celovito longitudinalno spremljanje pacientov, pregled stanja in izdelavo ocen na nacionalni ravni³ ter sprejem ustreznih zdravstvenih programov in politik na področju redkih bolezni. Zaradi tega je vzpostavitev nacionalnega registra za redke bolezni ena izmed ključnih aktivnosti Načrta dela na področju redkih bolezni v Republiki Sloveniji.¹ V skladu s tem so tudi cilji prispevka, ki se osredotočajo na analizo trenutnega stanja na področju redkih bolezni, identifikacijo izzivov za vzpostavitev registra redkih

bolezni ter pripravo praktično uporabnih usmeritev za razvoj registra redkih bolezni v Sloveniji.

Metode

Metodološko gledano je raziskava zasnovana kot študija primera, pri izvedbi celotne raziskave pa so bile uporabljene že uveljavljene metode družboslovnega raziskovanja. Raziskava se je najprej oprla na analizo primarnih in sekundarnih virov iz domače in tuje literature, mednarodnih projektov ter evropskih strateških dokumentov in usmeritev. Pregled smo izvedli s pomočjo spletnega brskalnika decembra 2016. Pregledali smo vso dostopno literaturo v dveh elektronskih podatkovnih zbirkah – PubMed in PubMed Central. Pri iskanju so bile uporabljene ključne besede oz. kombinacije ključnih besed »rare disease / registry / development / registration«. V nadaljevanju raziskave je bila izvedena sinteza raziskovalnih ugotovitev, v kateri so bila celovito povezana teoretična in praktična spoznanja z obravnavanega področja. Metodološki okvir je prilagojen vsebini in kompleksnosti raziskovalne problematike¹¹ in omogoča celovito analizo dosedanjih izsledkov na področju redkih bolezni, sintezo raziskovalnih ugotovitev¹² ter oblikovanje usmeritev za razvoj registra redkih bolezni v Sloveniji.

Rezultati

Trenutno stanje na področju redkih bolezni v Sloveniji

Za kodiranje bolezni se v Sloveniji trenutno uporablja posodobljena avstralska modifikacija desete revizije MKB in sorodnih zdravstvenih problemov za statistične namene (MKB-10-AM, verzija 6).¹³ Skladno s to klasifikacijo se kodira tudi redke bolezni, pri čemer je na voljo le majhen nabor specifičnih kod za približno 6.000 obstoječih redkih bolezni. S seznamom redkih bolezni, ki so vključene v omenjeno klasifikacijo, v Sloveniji ne razpolagamo. Nacionalni inštitut za javno zdravje (NIJZ) skladno z zakonom vodi evidenco o bolnišničnem zdravljenju pacientov, v kateri se beleži glavna diagnoze ter dodatne diagnoze.¹⁴ Incidenca redkih rakov se v Sloveniji spremlja v okviru Onkološkega inštituta preko njegove vključenosti v panevropski projekt RARECARE.¹⁵

Zakon o zbirkah podatkov s področja zdravstvenega varstva (ZZPPZ)¹⁶ določa obdelavo podatkov in zbirke podatkov s področja zdravstvenega varstva ter eZdravja, njihove upravljavce in upravičence do podatkov. V Prilogi 1 zakona so opredeljene tudi vrste in vsebina posameznih zbirk podatkov s področja

zdravstvenega varstva, njihov namen, obdobja poročila, pošiljatelji in časovni roki, upravljavec zbirke, način poročanja in čas hranjenja podatkov.¹⁶ Zbirka oziroma register redkih bolezni v zakonu ni posebej omenjen, opredeljenih pa je 5 registrov bolezni, ki so bodisi redke bodisi lahko vsebujejo redke bolezni. Od registrov, ki jih našteva Priloga 1 omenjenega zakona, obstajajo v praksi samo trije registri, ki vključujejo redke bolezni, in sicer register prirojenih anomalij, register rizičnih otrok in register oseb z motnjo strjevanja krvi. Na drugi strani obstaja v Sloveniji več kliničnih zbirk podatkov, ki še nimajo urejene pravne podlage v zakonu, čeprav se iz nekaterih izmed njih že pošiljajo podatki v evropske registre (npr. register Fabryjeve bolezni, register cistične fibroze itd.):

- Register pacientov s cistično fibrozo
- Register pacientov s prirojenimi presnovnimi boleznimi
- Nacionalni register pacientov s Fabryjevo boleznijo
- Register oseb z motnjo strjevanja krvi
- Register pacientov z odpovedjo ledvic
- Register pacientov z živčnomišičnimi boleznimi
- Slovenski nacionalni register pacientov s primarno imunsko pomanjkljivostjo
- Register prirojenih anomalij

V Sloveniji se pacienti z redkimi boleznimi zdravijo samo v štirih ustanovah, in sicer v Univerzitetnem kliničnem centru Ljubljana, KOPA Golnik, Splošni bolnišnici Slovenj Gradec in v Ortopedski bolnišnici Valdoltra, rehabilitacija pacientov pa se izvaja na URI – Soča. Center za nediagnosticirane redke bolezni, ki deluje znotraj Kliničnega inštituta za medicinsko genetiko UKC Ljubljana, predstavlja prvi specializirani oddelek za tovrstne bolezni v regiji.¹⁷

Čeprav register ni neposredno omenjen v ZZPPZ,¹⁶ je vzpostavitev registra eden izmed temeljnih predpogojev za obvladovanje in celovito upravljanje z redkimi boleznimi v Sloveniji in vsekakor sodi med prioritete naloge na celotnem področju redkih bolezni. Glede na trenutno stanje in dosedanje potrebe deležnikov na področju redkih bolezni v Sloveniji, pa tudi na podlagi mednarodnih izkušenj,^{10,18} ustrezno zasnovan in učinkovito upravljan register redkih bolezni ponuja številne priložnosti za izboljšanje stanja na področju redkih bolezni, in sicer:

- nadzor nad prevalenco in incidenco ter signaliziranje zgodnjih opozoril;
- zagotovitev informacij za razvoj ustreznih storitev na nacionalni ravni oziroma razvoj ustreznih

kliničnih poti z uporabo specialističnih storitev v tujini;

- razkritje naravnega poteka bolezni – značilnosti bolezni, vodenje in izidi z ali brez zdravljenja;
- spremljanje varnosti po vpeljavi novega oziroma eksperimentalnega načina zdravljenja;
- vrednotenje klinične učinkovitosti novih intervencij;
- spremljanje izidov oskrbe in omogočanje primerjave z evropskimi ali mednarodnimi standardi;
- zagotavljanje seznama pacientov, h katerim se lahko pristopi za klinične raziskave in sodelovanje multicentričnih študijah;
- zagotavljanje podatkov o ekonomskih ocenah zdravja, kot so stroški bolezni in študije stroškovne učinkovitosti.

V skladu s temi ugotovitvami in strateškimi usmeritvami Načrta dela na področju redkih bolezni v Sloveniji¹ se je oktobra 2015 pričel izvajati dvoletni ciljni raziskovalni projekt »Analiza in razvoj področja redkih bolezni v Sloveniji«, v okviru katerega se načrtuje zasnova nacionalnega sistema za spremljanje redkih bolezni vključno z razvojem pilotnega registra za redke bolezni.³ Pilotni register redkih bolezni je zasnovan kot spletna aplikacija, s pomočjo katere bodo pristojni subspecializirani izvajalci zdravstvene dejavnosti poročali podatke povezane z diagnosticirano redko boleznijo. V sklopu razvoja pilotnega registra redkih bolezni v projektu sodeluje pet poročevalskih institucij, ki sodelujejo tako pri zasnovi samega pilotnega registra redkih bolezni kot tudi pri njegovem testiranju in končni optimizaciji. Izgradnja pilotnega registra redkih bolezni temelji na metodologiji OpenEHR, ki se v zadnjih letih vse bolj uveljavlja za doseganje semantične interoperabilnosti na področju zdravstva in je bila uporabljena tudi v projektu PARENT JA za oblikovanje Registra endoprotetike Slovenije. Metodologija razvoja registrov s pomočjo OpenEHR temelji na ločenem obravnavanju klinične vsebine od preostale informacijsko-tehnološke (IT) podpore, ki v končni fazi upravlja z vsebinami. Klinične vsebine se načrtujejo in strukturirajo ločeno od preostalega načrtovanja IT rešitve. Celoten koncept temelji na tem, da načrtovanje klinične vsebine izvajajo zdravniki oziroma strokovnjaki za klinične vsebine, kar omogoča razvijalcem IT rešitve večjo osredotočenost na tehnični vidik rešitve. V projektu izdelave pilotnega registra redkih bolezni tako sodelujejo priznani slovenski strokovnjaki s področja redkih bolezni.

V letu 2015 je Ministrstvo za zdravje pričelo tudi s projektom razvoja spletne strani Nacionalne

kontaktne točke (NKT) za redke bolezni, ki ga vodi Pediatrična klinika Ljubljana. NKT je bila uspešno vzpostavljena leta 2016. Cilj te točke je povezati med seboj ustanove, strokovnjake in paciente z redkimi boleznimi ter njihove svojce, obenem pa pacientom in strokovnjakom omogočiti dostop do kakovostnih informacij o obravnavi redkih bolezni v Sloveniji.¹⁹ Omenjeni deležniki bodo, poleg drugih, predvidoma tudi glavni viri podatkov za nacionalni register redkih bolezni, ko bo le-ta vzpostavljen (slika 1). Ob nadaljnjem informiranju splošne javnosti o redkih boleznih in obstoju NKT ter z naprednimi IT rešitvami bi lahko, dolgoročno gledano, stremeli k možnosti samo-registracije pacientov tudi preko NKT.



Slika 1 Potencialni viri podatkov v slovenskem nacionalnem registru redkih bolezni.

Izzivi za vzpostavitev registra redkih bolezni v Sloveniji

V pripravi je 11. verzija MKB, ki za enkrat vsebuje okoli 5.400 kod za redke bolezni.²⁰ Pričakovati je, da bo 11. verzijo s prilagoditvami sprejela tudi Slovenija, vendar pa proces popolnega prevzema omenjene nove klasifikacije lahko traja več let. Podatki o bolnišničnih obravnavah so pomemben vir informacij za spremljanje, načrtovanje, vodenje in razvijanje zdravstvene dejavnosti, ocenjevanje zdravstvenega stanja prebivalstva, so pa tudi osnova za poročanje in mednarodne primerjave v okviru obveznosti do EU, Svetovne zdravstvene organizacije (SZO) in drugih. Leta 2013 je bila uvedena enotna metodologija poročanja o celotni bolnišnični obravnavi, tako

obravnavi istega tipa kakor tudi o posameznih epizodah bolnišnične obravnave po enotni metodologiji Spremljanje bolnišničnih obravnjav, preko informacijskega sistema ePrenosi.¹⁴ Glede na načrtovan prevzem MKB-11 v Sloveniji bi bilo smiselno razmisliti, ali bi v podatke o bolnišničnih obravnava vpeljali tudi mednarodne standarde kodiranja za redke bolezni (uporaba kod iz Orphanet nomenklature za redke bolezni). Po drugi strani pa bi bilo – tudi v luči potreb registra redkih bolezni – v bolnišničnih informacijskih sistemih nujno zagotoviti možnosti za ustrezno in standardizirano kodiranje redkih bolezni ter prenos podatkov v register redkih bolezni.

Pri snovanju dolgoročnih IT projektov na področju zdravstva (in tudi drugje) je potrebno že v zgodnji fazi načrtovanja zasnovati ustrezen model upravljanja in trajnostnega financiranja razvite IT rešitve. Pogosto se namreč dogaja, da se pri načrtovanju zagotovijo samo zagonska sredstva za razvoj in eventualno izvedbo pilotnega projekta. Kasneje se pogosto pojavijo vprašanja o dolgoročnem upravljanju, v smislu funkcij posameznih deležnikov, organizacije njihovega dela ter oblikovanja poslovnih procesov, ki se morajo ustrezno umestiti v sistem in zagotoviti učinkovito uporabo IT rešitve. S tem je povezana tudi zagotovitev sredstev za finančno vzdržnost projekta ter dolgoročno upravljanje, vzdrževanje in razvoj IT rešitve v prihodnjem obdobju. Kakovost in uporabnost končne IT rešitve sta namreč v veliki meri odvisna od razpoložljivih virov (finančnih, kadrovskih, informacijskih, organizacijskih), ki so na voljo za tekoče upravljanje registra in prihodnji razvoj. Zagotavljanje ustreznih virov je ena izmed ključnih predpostavk za učinkovit razvoj IT rešitve, njeno uspešno implementacijo ter kasnejše kakovostno upravljanje, ki vključuje vzdrževanje in razvoj. V skladu s težnjami v strateških dokumentih po ureditvi področja redkih bolezni v Sloveniji in vzpostavitvi registra redkih bolezni je bilo v zadnjih letih narejenih nekaj pomembnih premikov, predvsem v smislu resnejšega pristopa k celotni problematiki in vzpostavitvi NKT za redke bolezni. Kljub vsemu vprašanja institucionalnega upravljanja in dolgoročnega financiranja registra redkih bolezni še danes v veliki meri ostajajo odprta. To pomeni, da področje redkih bolezni še ni sistemsko urejeno, saj predstavlja register redkih bolezni enega izmed ključnih predpogojev za izboljšavo stanja ne celotnem področju.

Trenutno veljavni ZZPPZ¹⁶ ne nudi zakonske podlage za zbiranje podatkov o pacientih z redkimi boleznimi. Nov ali prenovljen Zakon o zbirkah podatkov s področja zdravstvenega varstva je bil

vključen v normativni delovni program Vlade Republike Slovenije za leto 2016, vendar pa sprejem tega zakona ni bil obvezen.²¹ Za vzpostavitev nacionalnega registra redkih bolezni je potrebno pripraviti ustrezen zakonski predlog in zagotoviti pravno podlago za zbiranje tovrstnih občutljivih osebnih podatkov.

Področje zdravstvene informatike v Sloveniji je izredno heterogeno tako v smislu različne stopnje informatiziranosti posameznih zdravstvenih dejavnosti (in različnih zdravstvenih ustanov) kot tudi številnih različnih (pogosto neusklajenih in neinteroperabilnih) IT rešitev, ki se v zadnjih letih uporabljajo v zdravstvenem sistemu. Zato so se na področju interoperabilnosti v predhodnem obdobju pojavljale številne težave, ki jih je ponekod in v manjši meri moč zaslediti še danes. Z uvedbo rešitev eZdravja in uporabo enotnih standardov na področju izmenjave podatkov so bile težave v veliki meri odpravljene. Potrebno pa je poudariti, da vsi izvajalci zdravstvene dejavnosti še ne uporabljajo rešitev eZdravja oziroma jih določeni izvajalci zdravstvene dejavnosti ne uporabljajo v zadostni meri. V kontekstu razvoja in uporabe registra za redke bolezni je pomemben predvsem Centralni register podatkov o pacientih (CRPP), ki je zbirka podatkov eZdravja o pacientih s stalnim ali začasnim prebivališčem v Republiki Sloveniji. CRPP vsebuje povzetek podatkov o pacientu in pacientovo zdravstveno dokumentacijo. CRPP omogoča izvajalcem zdravstvene dejavnosti v Sloveniji dostop in izmenjavo podatkov za izvajanje ustrezne in kakovostne zdravstvene oskrbe pacientov.¹⁶ Podatkov v CRPP trenutno še ne pošiljajo vsi javni zdravstveni zavodi v Sloveniji, bodisi zaradi tehničnih bodisi zaradi drugih razlogov, čeprav je centralna infrastruktura funkcionalna že dlje časa.

Usmeritve za razvoj registra redkih bolezni

Registracija je proces stalnega sistematičnega zbiranja podatkov o pojavljanju in značilnosti pomembnih pojavov z namenom čim bolj natančne ocene razsežnosti pojava in njegovega vpliva na prebivalstvo ter čim bolj kvalitetnega načrtovanja ukrepov za njegovo obvladovanje. Register je zbirka/arhiv, v katerem se shranjujejo podatki o pojavu, ki ga registriramo. Zapis je vezan na posamezno osebo. Teoretično naj bi se registrirali vsi primeri pojava, ki ga registriramo, če pa to ni možno, je potrebno doseči čim večjo pokritost.²² Registri so posebno primerni za situacije, kjer eksperimentalno raziskovanje ni izvedljivo ali praktično, kot je npr. nadzor nad redkimi dogodki ali redkimi boleznimi.²³

Registri pacientov so v rabi že več desetletij na področjih kot so rak, prirojeni defekti in kardiovaskularne bolezni.⁵ Register raka v Sloveniji je bil vzpostavljen že leta 1950.²⁴ Ta dolga in obširna zgodovina zbiranja podatkov je osnova, na kateri se gradijo tudi usmeritve za registracijo pacientov z redkimi boleznimi, čeprav imajo registri redkih bolezni tudi nekaj svojstvenih lastnosti. Razmeroma majhno število primerov in kompleksnost bolezni nalagata veliko geografsko pokritost in obsežno izmenjavo podatkov (običajno mednarodno), poleg tega je večina redkih bolezni genetskega izvora, kar zahteva, da so posamični primeri sledljivi v okviru družinskih skupnosti. Zato so stroški vzpostavitve in vzdrževanja registrov redkih bolezni skoraj enaki kot za registre, ki spremljajo običajne bolezni.⁵

Končni cilj registra redkih bolezni je vzpostavitev celovitega sistema na področju redkih bolezni, ki omogoča učinkovitejšo zdravstveno oskrbo za paciente in družine preko izboljšanja znanja o naravnem poteku bolezni ter zagotavlja informacije deležnikom; zdravnikom, organizacijam pacientov in oblikovalcem politik.²⁵ Ustrezno zasnovan, strukturiran in metodičen pristop je ključnega pomena za učinkovit razvoj registra redkih bolezni.

Načrtovanje registra redkih bolezni

Registri pacientov, ki so bili zasnovani za paciente s skupnimi karakteristikami (diagnoza, vgrajena naprava, terapija, tveganje za razvoj bolezni ipd.), predstavljajo enega izmed pomembnih sistemskih ukrepov na področju javnega zdravja. Tovrstni registri so desetletja služili kot važen vir podatkov za oceno klinične učinkovitosti, oceno zdravstvene tehnologije ali oceno posledic zdravstvenih politik/programov na lokalni, regionalni, nacionalni, pa tudi mednarodni ravni.⁴

Splošni principi načrtovanja tovrstnih IT rešitev in specifične smernice za vzpostavitev registrov na področju zdravstvene oskrbe pacientov poudarjajo, da je pri načrtovanju registrov potrebno slediti zaporedju korakov²⁶: 1) artikulacija namena registra, 2) presoja, ali je register ustrezen način, da se doseže namen, 3) identifikacija ključnih deležnikov in 4) ocena izvedljivosti registra. Po ustrezni oceni izvedljivosti oz. odločitvi o nadaljevanju projekta je potrebno izvesti še naslednje korake²⁶: 5) oblikovanje skupine za razvoj registra, 6) vzpostavitev načrta za upravljanje in nadzor registra, 7) opredelitev potrebnega obsega in zahtevnosti registra, 8) opredelitev podatkovnega nabora, izidov zdravljenja in tarčne populacije, 9) razvoj raziskovalnega načrta ali protokola in 10) razvoj projektnega načrta. Posamezni

koraki v procesu načrtovanja se lahko izvajajo vzporedno. Nekateri od naštetih korakov so podrobneje predstavljeni v nadaljevanju.

Artikulacija namena registra

Eden izmed prvih korakov pri načrtovanju registra je artikulacija njegovega namena.²⁶ Cilj oziroma namen je potrebno jasno definirati, saj bo to pomagalo pri oceni, ali predstavlja register ustrezen pristop za zajem informacij, ki nas zanimajo. Jasno definiran namen bo hkrati pomagal osvetliti potrebo po določenih podatkih. Register ima lahko enega ali več namenov. V vsakem primeru mora biti celoten namen preveden v specifične cilje ali vprašanja, ki se jih bo naslavljalo z registrom.²³ Jasni cilji so nujni za definicijo strukture in procesov zbiranja podatkov ter zagotavljajo, da bo register učinkovito naslovil pomembna vprašanja skozi ustrezne analize rezultatov. Načrtovalci registra morajo vzpostaviti specifične cilje upoštevaje ključna vprašanja, na katera mora register odgovoriti.

Identifikacija ključnih deležnikov

Tipično ima vsak register primarne in sekundarne deležnike. Primarni deležniki so odgovorni za vzpostavitev in financiranje registra. Sekundarni deležnik je nekdo, ki bi imel koristi od poznavanja podatkov ali na katerega bi imeli rezultati vpliv, vendar pa ni ključen za vzpostavitev registra. V širši nabor možnih deležnikov, tako primarnih kot sekundarnih, sodijo pacienti (društva, združenja, zagovorniške skupine), izvajalci zdravstvene oskrbe, plačniki zdravstvenih storitev (zavodi za zdravstveno zavarovanje, zavarovalnice), proizvajalci zdravil, javnozdravstveni in zakonodajni organi, akademske in raziskovalne inštitucije ter strokovna združenja. Koordinirane aktivnosti in interakcije med deležniki v času uporabe registra lahko zagotovijo, da je register povezan z nameni in cilji, ki so bili določeni v fazah načrtovanja, in da je delovanje registra skladno s smernicami, pravili in zakoni.²⁶

Ocena izvedljivosti registra

Ključni element ugotavljanja izvedljivosti razvoja novega registra se nanaša na financiranje (kar vključuje tako zagonske stroške za razvoj in vzpostavitev registra kot tudi njegovo dolgoročno vzdrževanje in nadgrajevanje). Višina stroškov je odvisna od obsega registra, obsežnosti in natančnosti zbirke podatkov ter potrebnega testiranja in nadzora. K stroškom prispeva tudi metoda zbiranja podatkov, stroški prenove poslovnih procesov, upravljavski stroški idr. Potrebe po zagotavljanju finančnih sredstev je potrebno analizirati v luči projekcij o življenjski dobi registra in njegovi dolgoročni finančni

vzdržnosti. Skladno z nameni, cilji in potrebami po registru se lahko razlikujejo tudi možni viri financiranja razvoja in vzdrževanja registrov.

Oblikovanje skupine za razvoj registra

Za načrtovanje in implementacijo registra je potrebno raznoliko znanje, izkušnje in sposobnosti. Člani razvojne ekipe morajo razumeti cilje in namen registra, dobro poznati vire in značilnosti podatkov ter imeti dober sistemski in vsebinski vpogled v področje, ki ga register pokriva. Za uspešno in učinkovito delo morajo člani skupine za razvoj registra dobro poznati vsebinsko ozadje problematike ter imeti znanja in izkušnje s področja registrov, podatkovnih zbirk, projektnega menedžmenta, zakonodaje ter zagotavljanja kakovosti.

Vzpostavitev načrta za upravljanje in nadzor registra

Upravljanje se nanaša na vodstvo in sprejemanje odločitev na visokem nivoju, vključno z idejnimi rešitvami, financiranjem, izvedbo in razširjanjem informacij. Odvisno od velikosti registra, namena in vsebine registra, upravljanje običajno izvajajo različna telesa (nadzorni, usmerjevalni odbor), ki jih sestavljajo strokovnjaki, ki so del razvojne skupine (interno upravljanje), ali strokovnjaki, ki ne sodelujejo pri izvajanju dnevnih operativnih aktivnosti na področju registra (zunanje upravljanje). V skladu s tem je tudi izbor pristopa (in organizacijske oblike) za izvajanje in stopnjo nadzora nad delovanjem registra, ki je odvisen od pomembnosti vsebinskega področja, ki ga pokriva register, višine sredstev, ki jih ima register na voljo za svoje delovanje, zanesljivosti virov podatkov, stabilnosti delovanja registra, zadovoljstva uporabnikov ipd. Ključne funkcije na področju upravljanja in nadzora registrov se osredotočajo na izvajalske ali usmerjevalne aktivnosti, ki preko sprejemanja odločitev na finančnem, organizacijskem, zakonodajnem, etičnem in znanstvenem področju uokvirjajo delovanje registra in določajo nadaljnje razvojne smernice registra. Odvisno od zmožnosti, virov in ciljev, lahko upravljavska in nadzorna telesa prevzamejo določene funkcije tudi na znanstvenem, promocijskem in drugih področjih. Med pomembne funkcije spadajo tudi dolgoročne projekcije na področju delovanja registra. Že v fazi načrtovanja registra je namreč potrebno razmišljati o življenjski dobi registra in prenehanju njegovega delovanja, predvsem v smislu lastništva in prenosa podatkov, lastnine nad strojno in programsko opremo ter zaupnosti in varovanju osebnih podatkov pacientov. Sodelujoče paciente je potrebno obvestiti o morebitni

ukinitvi registra in usodi podatkov, ki so jih prispevali v register.²³

Opredelitev podatkovnega nabora, izidov zdravljenja in tarčne populacije

Podatkovni elementi, ki naj bodo vključeni v register, morajo imeti uporabno vrednost za znanstvene raziskave in klinično prakso ter morajo biti potrjeni s strani strokovnjakov, vključno z mnenji biostatistikov, informatikov in epidemiologov. Vsak podatkovni element se mora nanašati na namen in specifične cilje registra. Spremenljivke jedrnega podatkovnega nabora definirajo informacije, ki so potrebne za naslavljanje kritičnih vprašanj, zaradi katerih je bil register zgrajen. Če se doda druge spremenljivke, ki so bolj opisne narave, je pomembno, da taki podatkovni elementi sovpadajo s cilji registra in njihovo zbiranje in vnos ne predstavljata prevelikega bremena. Tudi podatke, ki jih je nujno poznati, je včasih težko zanesljivo zbrati ali težko zbrati brez precejšnjega bremena. Hkrati z razvojem jedrnega podatkovnega nabora je potrebno razmišljati tudi o načrtu za zagotavljanje kakovosti zbranih podatkov. Register bi se moral izogibati doseganju preštevilnih ciljev, sicer bo njegovo breme pretehtalo njegovo uporabnost za klinične strokovnjake in raziskovalce. V fazi načrtovanja je nujno ugotoviti, kateri podatki so že na voljo in katere je potrebno še dodatno zbrati.²⁶ Pri zbiranju podatkov je potrebno upoštevati načelo sorazmernosti, zakonodajo s področja varovanja osebnih podatkov, nacionalna in mednarodna etična merila ter druga pravila, ki naslavlajo preprečevanje tveganj pri zbiranju podatkov in upravljanju registra.²⁶

V zgodnji fazi zasnove registra je potrebno identificirati najpomembnejše izide zdravljenja. Opredelitev izidov zdravljenja bo za snovalce registra predstavljala izhodišča za vzpostavitev prioritet pri načrtovanju in funkcionalnostih registra ter določanje podatkovnega nabora. Potrebno je jasno vzpostaviti načine in metodologijo preverjanja glavnih izidov zdravljenja. Pri vzpostavitvi ustrezne strukture registra za spremljanje izidov zdravljenja je potrebno upoštevati tudi diagnostične zahteve, raven podrobnosti podatkov ter stopnjo potrditve in vrednotenja podatkov.²⁶

Tarčna populacija je populacija, na katero se bodo nanašali podatki iz registra. Definirana mora biti iz dveh osnovnih razlogov: ker služi kot osnova za načrtovanje registra in ker predstavlja skupnost, na katero bodo vplivali rezultati registra. Opredelitev tarčne populacije je odvisna od številnih dejavnikov, ključna determinanta pri opredelitvi pa je namen

registra. Potrebno je natančno preučiti vse dejavnike in sprejeti metodologijo za določanje članov tarčne populacije, kakor tudi potencialnih virov za dodatne člane tarčne populacije izven zastavljenih okvirjev registra (društva pacientov). Večja dostopnost do tarčne populacije omogoča večjo reprezentativnost in statistično moč raziskav,² ki se lahko uporabijo za oblikovanje ustreznih zdravstvenih politik oziroma programov na področju redkih bolezni.

Razvoj raziskovalnega načrta ali protokola

Raziskovalni načrt opredeljuje in dokumentira cilje registra ter opisuje, kako se bodo ti cilji dosegli. Celovit raziskovalni načrt mora podati poglobljen vpogled v razmere na obravnavanem področju ter opisati, kako bo register pripomogel k reševanju prepoznanih težav. V vsakem primeru mora vključevati cilje registra, vključitvene kriterije za udeležence in postopke zbiranja podatkov. Idealno pripravljen raziskovalni načrt bi moral ustrezno opredeliti cilje, podatkovno strukturo registra, arhitekturno zasnovo, infrastrukturo, vključitvene kriterije, izide zdravljenja, postopke zbiranja podatkov, postopke upravljanja ter načrte za zagotavljanje etičnih meril in varstvo osebnih podatkov pacientov.²³

Razvoj projektnega načrta

V začetni fazi načrtovanja registra je potrebno pripraviti celovit projektni načrt. Tovrsten dokument predstavlja osnovo za vse projektne aktivnosti, ki jih izvaja skupina za razvoj registra in je kot takšen kritičnega pomena za učinkovito vodenje in uspešno realizacijo projekta. Odvisno od kompleksnosti projekta registra, lahko projektni načrt vključuje nekatere ali vse naslednje elemente²⁶:

- načrt vodenja projekta;
- podrobno časovnico za vse predvidene aktivnosti;
- načrt upravljanja stroškov;
- načrt upravljanja s kakovostjo;
- načrt upravljanja s kadrovskimi viri;
- načrt komunikacije;
- načrt nabave strojne opreme in razvoja programske opreme za register;
- načrt upravljanja s tveganji, ki zagotavlja identifikacijo tveganj in ukrepe za njihovo zmanjšanje.

Vsebinska zasnova registra redkih bolezni

Ko so nameni in glavni cilji registra jasno definirani, je v naslednjem koraku potrebno prevesti namen v temeljno (raziskovalno) vprašanje. Ne glede na naravo raziskovalnega vprašanja je definicija le-tega ključna pomena za načrtovalce registra, saj bodo ta

raziskovalna vprašanja usmerjala nadaljnje odločitve (kateri podatki se bodo v registru zbirali, populacija registra itd.) in delo pri razvoju registra.²⁷ Pri definiranju raziskovalnih vprašanj ali bolj specifičnih raziskovalnih hipotez je pomembno, da so le-ta natančna, razumljiva ter vsebinsko osredotočena na specifičen register.²⁸ Raziskovalna vprašanja pri razvoju registra redkih bolezni se običajno nanašajo na stopnjo incidence in prevalence redkih bolezni v določeni populaciji.

Čeprav so študije, izvedene na podlagi podatkov iz registrov, po definiciji opazovalne študije, je analitični okvir tisti dejavnik, ki sproži zbiranje podatkov, določa način zbiranja podatkov ter izbiro pacientov za vključitev v študijo.²⁶ V fazi zasnove registra je potrebno eksplicitno določiti, kako velik bo register, koliko časa se bodo spremljali pacienti in na čem morajo sloneti utemeljitve za sprejem tovrstnih odločitev. Te odločitve morajo biti ustrezno utemeljene v splošnem namenu registra.²³

Podatkovni elementi za register

Podatkovni elementi morajo biti skrbno izbrani glede na namen registra; vsak podatkovni element mora podpirati namen in cilje registra. Izbor podatkovnih elementov za register se začne z identifikacijo domen oziroma zbirk podatkovnih elementov, ki se vežejo na skupno temo. V registrih se običajno beležijo zbirke podatkovnih elementov, ki so vsebinsko nanašajo na pacienta, izvajalca zdravstvene dejavnosti, trenutni status pacienta in bolezni, izide zdravljenja, mejnike v procesu zdravstvene oskrbe ter administrativne zadeve.

Po identifikaciji podatkovnih domen za register sledijo odločitve o tem, katere specifične podatke bo zbiral register. Če register postavi preveč kompleksen proces zbiranja podatkov, kar poveča stroške in breme, lahko ogrozi delovanje ter vzdržnost registra.²⁸ Definicije podatkovnih elementov morajo biti eksplicitne in bi morale zagotavljati, da ni nobenih razlik v pojmovanju, zbiranju ali formatu med ustanovami in posamezniki, ki zbirajo in poročajo podatke.²⁸ Zato je uporaba podatkovnih standardov zelo pomembna pri izgradnji registra. Standardni podatkovni elementi in definicije bi morali biti uporabljani, če obstajajo in če je le mogoče. Standardi izboljšujejo konsistentnost, primerljivost in skupno razumevanje podatkovnih elementov. Uporaba obstoječih podatkovnih standardov, kot so klasifikacije, klinične terminologije in skupni podatkovni nabori, omogoča primerjavo rezultatov, interoperabilnost ter izmenjavo in ponovno uporabo podatkov.²⁸ V tej fazi je potrebno opredeliti tudi

politike dostopnosti do podatkov in načrt za objavlanje podatkov in ugotovitev.

Metodološka navodila

Vsak podatkovni nabor zahteva metodološka navodila, ki vključujejo podrobne informacije o tem, kaj se zbira in kako. Navodila lahko vključujejo a) interpretacije definicij podatkovnih elementov, b) pojasnila, kateri podatki se zbirajo oz. kateri pojavi so vključeni v opazovanje in kateri ne, ter obrazložitev potencialno nejasnih primerov/situacij, c) pravila in omejitve za specifične podatkovne elemente, vključno z informacijo o formatu podatkovnih elementov in tipu obveznosti vnosa in d) informacije o zbiranju in poročanju podatkov (kdo mora zbirati podatke o posameznem podatkovnem elementu, kdaj in s kakšno metodo oziroma instrumentom mora zbirati podatke, kdo je dolžan poročati o zbranih podatkih in kakšna je dinamika poročanja). Podatkovni elementi morajo biti dokumentirani in dostopni vsem zainteresiranim deležnikom. Pomembno je, da register vsebuje vso potrebno spremljajočo dokumentacijo, vključno s podatkovnim slovarjem – popisom in opisom vseh podatkovnih elementov, vključenih v register.²⁸

Viri podatkov za register

Viri podatkov se delijo na primarne in sekundarne, odvisno od pomembnosti za namen registra. Primarni viri podatkov zajemajo podatke, ki so zbrani za neposredne namene registra. Uporabljamo jih, ko želenih podatkov ni mogoče pridobiti nikjer drugje. Zbiranje podatkov iz primarnih virov povečuje verjetnost popolnosti, veljavnosti in zanesljivosti podatkov, kar je zelo pomembno za vrednotenje kakovosti podatkov in verodostojnosti registra nasploh.²⁹ Ti podatki se prospektivno načrtujejo in zbirajo v skladu s protokolom ali raziskovalnim načrtom z uporabo skupnih postopkov in enakega formata pri vseh pacientih. Vnesene podatke je možno slediti do posameznika, ki jih je zbral, zato je podatke iz primarnih virov preprosto preverjati z avtomatiziranimi kontrolami.

Sekundarni viri podatkov vsebujejo podatke, ki so bili izvorno zbrani za druge namene (standardna medicinska oskrba, krvne banke, zavarovalniški zahtevki, administrativne podatkovne baze itd.). Ti podatki so pogosto shranjeni v elektronskem formatu in razpoložljivi za rabo z ustreznimi dovoljenji. Za natančno ujemanje podatkov iz sekundarnih virov in pacientov so potrebni ustrezni in zanesljivi identifikatorji. Pri procesu uparjanja je potrebno biti pozoren na morebitne navzkrižne napake in podvajanje podatkov. Preden se podatki iz

sekundarnih virov vključijo v register, je potrebno kritično oceniti potencialen vpliv kakovosti teh podatkov na celotno kakovost podatkov v registru.

Zagotavljanje interoperabilnosti

Za zagotavljanje interoperabilnosti je bistvena uporaba ustreznih standardov za strukturiranje, berljivost in prenos zdravstvenih podatkov preko zdravstvenih informacijskih sistemov. Med najbolj uveljavljenimi standardi na področju zdravstvenih podatkov so Health Level Seven (HL7), OpenEHR in terminologije Logical Observation Identifiers Names and Codes (LOINC) za laboratorijske izvide, Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) za klinične izraze, Orphanet za zdravila sirote in diagnoze redkih bolezni ter Online Mendelian Inheritance in Man (OMIM) za gene in genetske fenotipe. Standardi omogočajo interoperabilnost in izboljšujejo splošno kakovost podatkov v registrih. Načrtovanje uporabe podatkovnih standardov je nujno že v fazi zasnove ter kasnejše izgradnje in uporabe registra. Podatkovni standardi se morajo upoštevati tudi pri oblikovanju skupnih podatkovnih elementov za različne bolezni ali zasnovi minimalnih podatkovnih naborov. Le tako strukturirane podatke je namreč mogoče integrirati v elektronske zdravstvene kartoteke pacientov in kasneje uporabiti za različne namene (spremljanje ter načrtovanje zdravstvene oskrbe, epidemiološke in druge raziskave na področju javnega zdravja).³⁰ Kodiranje bolezni pri izvajalcih zdravstvene dejavnosti je prilagojeno obstoječim nacionalnim praksam. Diagnoze redkih bolezni so lahko kodirane v Orphanetu, OMIM ali SNOMED CT, odvisno od konteksta in od zahtevane granularnosti kodiranja.³⁰

Kodiranje in razvrščanje redkih bolezni

V mednarodnih nomenklaturah ima trenutno le majhen delež redkih bolezni ustrezne kode, zaradi česar ni mogoče slediti vseh pacientov z redko boleznijo v zdravstvenih informacijskih sistemih na nacionalni ali mednarodni ravni.³¹ Klasifikacija bolezni Svetovne zdravstvene organizacije v sistemu MKB-10 ali 9 se uporablja v večini držav članic, samo Velika Britanija uporablja SNOMED CT. V omenjenih sistemih ima le majhen delež redkih bolezni svojo posebno kodo: 250 oziroma 466 bolezni v MKB-10 in 2.883 v SNOMED CT od skupno skoraj 7.000 redkih bolezni, ki potrebujejo kodiranje.³² Izboljšanje kodiranja redkih bolezni je navedeno kot prioriteta v Priporočilih Sveta o aktivnostih na področju redkih bolezni iz leta 2009.³³ V obravnavi je nova verzija MKB-11, ki bo vsebovala tudi kode za večino redkih bolezni, vendar bo uvedba nove verzije

(vsebinsko in tehnično) dolgotrajna, povzročila pa bo tudi pomembne dodatne finančne obremenitve, kar ne obeta hitrih in enostavnih rešitev.³² Trenutno je za 5.400 redkih bolezni, ki so navedene v podatkovni zbirki Orphanet, potrjena vključitev v MKB-11 in dodelitev unikatnega identifikatorja, kar je približno 10-krat več kot v MKB-10.³³ Orphanet daje prost dostop do nomenklature redkih bolezni, ki je usklajena z MKB-10 in MKB-11 (beta verzijo) ter drugimi nomenklaturami, vključno z SNOMED CT, OMIM ipd. Nomenklatura Orphanet je dostopna v različnih formatih, kar omogoča poenostavljen prenos in integracijo z različnimi informacijskimi sistemi. Po mnenju Commission Expert Group on Rare Diseases (CEGRD) bi morale države članice EU razmisliti o dodajanju Orphanet državnim zdravstvenim informacijskim sistemom in raziskati možnosti in potrebne vire za izvedbo.³⁴ Nekatere države so že same prilagodile MKB-10 z razširitvami sistema za posebne potrebe redkih bolezni, npr. Nemčija in Francija.³⁵ V Franciji so vzpostavili eno od ključnih aktivnosti drugega francoskega načrta za redke bolezni: ORPHA številke so stopile v rabo v bolnišničnih informacijskih sistemih, kar je omogočilo boljšo sledljivost pacientov z redko boleznijo v zdravstvenem sistemu.³⁶ Izboljšave in prilagoditve kodiranja redkih bolezni so eden glavnih ciljev v mnogih nacionalnih akcijskih načrtih na področju redkih bolezni. Nacionalni zdravstveni sistemi pri tem uporabljajo različne pristope; v mnogih državah kodiranje predstavlja glavni vir podatkov za vse nadaljnje raziskave in statistične analize. Ponekod imajo postopki kodiranja tudi pomembne finančne posledice ter pomembno vlogo pri načrtovanju zdravstvenih programov in politik.³⁵

Zakonodajni vidiki ter varstvo osebnih podatkov

Direktiva 95/46/EC Evropskega parlamenta in Sveta z dne 24. 10. 1995 o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov predstavlja zakonodajni okvir na področju varovanja osebnih podatkov v EU. Direktiva navaja pravila glede zakonite obdelave in varstva osebnih podatkov ter s tem povezanih pravic državljanov, kot tudi nujnost zagotavljanja ustreznih nadzornih mehanizmov na področju varstva osebnih podatkov. Pravila iz direktive vpeljujejo privolitev posameznika in vnaprejšnjo obveščeno o obdelavi njegovih osebnih podatkov. V Sloveniji je to področje urejeno z Zakonom o varstvu osebnih podatkov,³⁷ ki uvaja določbe iz omenjene direktive v slovenski pravni red in v 6. poglavju jasno navaja pogoje, v katerih je dovoljeno povezovanje zbirk osebnih podatkov.

Evropska komisija je 25. 1. 2012 objavila obsežen zakonodajni sveženj za reformo zakonodaje EU na področju varstva osebnih podatkov. Reforma naj bi zaščitila osebne podatke v EU, obenem pa povečala nadzor državljanov nad svojimi podatki ter zmanjšala stroške za podjetja. Nova Splošna uredba o varstvu podatkov (ang. General Data Protection Regulation, GDPR) (EU) 2.016/679 Evropskega parlamenta in Sveta z dne 27. 4. 2016 o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov ter o razveljavitvi Direktive 95/46/ES je bila objavljena aprila 2016 in se bo začela uporabljati maja 2018.³⁸

Razprava

Čeprav raziskave kažejo, da redke bolezni predstavljajo resen problem na področju javnega zdravja, je zaradi splošnega pomanjkanja epidemioloških podatkov o številnih redkih boleznih nemogoče natančno oceniti njihovo dejansko breme. Register na področju redkih bolezni pomeni več kot običajno epidemiološko orodje za zbiranje sekundarnih podatkov, povezanih s pacienti z določeno diagnozo redke bolezni. Registri redkih bolezni predstavljajo učinkovit način za doseganje dovolj velikega vzorca za epidemiološke in klinične raziskave in so ključni instrumenti za nadzor teh bolezni, namenjeni izboljšanju oskrbe pacientov in načrtovanju zdravstvene oskrbe.³⁹ Obenem je dokazano, da kjer obstajajo dobro implementirani registri redkih bolezni in aktivne organizacije pacientov, se verjetnost razvoja zdravila za obravnavano redko bolezen poveča. Hkrati dosledno longitudinalno zbiranje pacientovih podatkov olajša izdelavo standardov oskrbe ter znatno izboljša rezultate zdravljenja in pričakovano preživetje tudi v odsotnosti novih terapij.⁴⁰ Zaradi teh razlogov je razvoj ustreznih registrov ena od prioritet na področju redkih bolezni v EU. Slednje potrjujejo tudi različni projekti na področju redkih bolezni, ki jih financira EU in katerih strateški cilj je oblikovanje evropske platforme za registracijo in spremljanje pacientov z redko boleznijo³⁹ ter zagotavljanje podpore za obstoječe (in bodoče) registre redkih bolezni.

Na drugi strani spremljajo razvoj in upravljanje registrov redkih bolezni razmeroma visoki stroški in zahtevna organizacija. Vzpostavitev registra je z javnofinančnega in javnozdravstvenega vidika opravičljiva torej le v povezavi z reševanjem hudih oziroma specifičnih problemov na področju javnega zdravja.⁴¹ Epidemiološka uporabnost registra narašča z dolžino obstoja registra, kar pogosto pomeni, da se zbiranje podatkov, dokumentiranje in kontrole

kakovosti izvajajo veliko let preden postane register redkih bolezni polno uporaben za epidemiološke namene. Splošno pravilo je, da registri pacientov zahtevajo kontinuirano financiranje ter dolgoročno zavezanost vseh deležnikov in bi morali biti vzpostavljeni le z močnim zagotovitvom, da bo register ustrezno tehnološko in materialno podprt še leta v prihodnosti.⁴²

Nekatere države v EU so prepoznale koristi in priložnosti, ki jih ponujajo registri redkih bolezni, in so že pred časom pristopile k njihovi vzpostavitvi preko različnih nacionalnih ali mednarodnih projektov. Izkušnje iz teh držav in primeri dobrih praks potrjujejo, da je za uspešno vzpostavitev in kasnejšo učinkovito uporabo registra redkih bolezni potrebno zagotoviti stabilno sistemsko infrastrukturo ter izbrati ustrezen in metodičen razvojni pristop, ki vključuje podporo zdravstvene politike ter dobro koordinacijo in sodelovanje vseh deležnikov. Pred pričetkom projekta ali pa najkasneje vzporedno z aktivnostmi za vzpostavitev registra redkih bolezni je potrebno zagotoviti zakonske podlage za delovanje registra, ki opredeljujejo obveznosti in organizacijo na področju zbiranja, prenosa in analize podatkov ter materialne pogoje za upravljanje in dolgoročno vzdržnost registra. Te pogoje se velikokrat spregleda, saj se realizacija tovrstnih projektov v zdravstvu pogosto tolmači kot enkratna naložba.

Število pacientov z redko boleznijo v Sloveniji po nekaterih ocenah presega število pacientov s sladkorno boleznijo tipa 2. Tudi pri nas se pacienti z redko boleznijo soočajo s pozno diagnozo, dragim zdravljenjem (če sploh obstajajo zdravila) in socialnimi posledicami. Za posameznike, ki jih prizadene omejujoča, življenje ogrožajoča bolezen, bi bil celovit javnozdravstveni pristop z vzpostavitvijo nacionalnega registra redkih bolezni pomemben korak k izboljšanju stanja na celotnem področju.⁴ Pričakovati je namreč, da bi ustrezen register redkih bolezni tudi v Sloveniji prispeval k bolj učinkovitemu nadzoru in spremljanju teh bolezni, izboljšanju zdravstvene oskrbe pacientov in zmanjšanju neenakosti. Kratkoročno bi pomemben korak naprej pomenila že uvedba Orphakod v procese kodiranja, ki bi omogočila bolj ustrezno kodiranje redkih bolezni in evidentiranje pacientov. Toda v zvezi z vzpostavitvijo slovenskega nacionalnega registra redkih bolezni se poleg težav na področju kodiranja pojavljajo še številni drugi izzivi, kot so dolgoročno financiranje in upravljanje registra ter neinteroperabilnost zdravstvenih informacijskih sistemov. Poleg tega tovrsten register pri nas še nima ustrezne zakonske podlage v ZZPPZ.¹⁶

Kljub naraščajočemu zanimanju raziskovalne in akademske sfere za različne vidike registrov redkih bolezni in njihove implikacije⁴³ so empirične raziskave, ki bi celovito osvetljevale to problematiko, še vedno maloštevilne. Večina raziskav potrjuje kompleksen in večrazsežen vpliv registrov redkih bolezni na celotno področje in soglašajo, da je izgradnja celovitih registrov eden od temeljnih korakov za sistemsko ureditev področja. Glavna metodološka omejitev študije primera, uporabljene v tem prispevku, je nezmožnost empiričnega vrednotenja in preverjanja dejanskih vplivov registra redkih bolezni v realnem okolju slovenskega zdravstvenega sistema oziroma konkretno na področju redkih bolezni, saj Slovenija še nima nacionalnega registra redkih bolezni. Prispevek naslavlja aktualno tematiko in odpira vprašanja, ki bodo v prihodnjih raziskavah zahtevala poglobljeno obravnavo. V prihodnjih raziskovalnih poskusih bo potrebno preučiti dejanski vpliv registra redkih bolezni preko empiričnih analiz ter v nadaljevanju s pomočjo sinteze teoretičnih in empiričnih spoznanj z obravnavanega področja oblikovati ustrezne sklepe.

Zaključek

Vzpostavitev registra redkih bolezni predstavlja pomembno pridobitev z vidika pacientov, saj lahko v veliki meri prispeva k načrtovanju zdravljenja in izboljšanju oskrbe pacientov ter omogoči celovito spremljanje izidov in kakovosti zdravljenja. Zlasti pomembne so evidence pacientov, ki se zdravijo z zdravili sirotami, ker omogočajo zbiranje dokazov o učinkovitosti zdravljenja in možnih stranskih učinkih. Na drugi strani lahko ustrezno zasnovan register redkih bolezni pomeni znatne sistemske koristi za celotno področje redkih bolezni, saj lahko zagotovi uporabna izhodišča za razvoj epidemioloških in kliničnih raziskav, programske oziroma politične analize, načrtovanje poslovanja zdravstvenih institucij, tehnološke inovacije ter oblikovanje ustrežnejše organizacijske strukture in optimalnejših delovnih procesov. Izgradnja dolgoročno učinkovitega, večstransko uporabnega ter ustrezno informatično podprtega registra redkih bolezni ni odvisna samo od tehnoloških rešitev, ampak zahteva spremenjeno paradigmo delovanja vseh deležnikov, vključno z zagotovitvijo ustreznih materialnih in nematerialnih virov za izvedbo potrebnih razvojnih in kasnejših upravljaljskih procesov.

Dosedanji naporji za ureditev razmer na področju redkih bolezni kažejo, da je za učinkovito sistemsko rešitev tako kompleksne problematike potrebno uskladiti zdravstveno politiko, organizacijske,

informacijske, zakonodajne in druge dejavnike na področju redkih bolezni. Vzpostavitev registra redkih bolezni ima velik potencial in bo v prihodnosti pomembno določala razvoj področja redkih bolezni, kljub temu pa informatizacija kot izoliran dejavnik ne more spodbuditi uspešne spremembe trenutnih razmer in izboljšati rezultatov zdravstvene oskrbe pacientov na področju redkih bolezni. Uspešno poslovanje in učinkovitost zdravstvene oskrbe na določenem področju zdravstvenega sistema sta namreč kritično odvisna od usklajenosti IT rešitev s številnimi drugimi sistemskimi dejavniki in prilagoditve modela zdravstvene oskrbe potrebam pacientov in specifikam zdravstvenega okolja. Uspešna vzpostavitev registra redkih bolezni terja mobilizacijo različnih deležnikov ter koordinacijo nasprotujočih si interesov znotraj zdravstvenega sistema, kar otežuje njeno realizacijo in predstavlja velik izziv za dolgoročno ureditev razmer na področju redkih bolezni v Sloveniji. Kljub omenjenim težavam vzpostavitev registra redkih bolezni predstavlja pomembno razvojno priložnost, ki lahko ob ustreznih odpravi ovir učinkovito poveže različne deležnike in poveča izkoristek institucionalnih potencialov, ki nedvomno obstajajo, ter pripomore k izboljšanju zdravstvene oskrbe vseh pacientov z redko boleznijo.

Reference

1. Ministrstvo za zdravje: *Načrt dela na področju redkih bolezni v Republiki Sloveniji*. Ljubljana 2011: Ministrstvo za zdravje Republike Slovenije.
2. Gliklich RE, Dreyer NA, Leavy MB: *Rare disease registries*. Rockville 2014: AHRQ.
3. Grošelj U: Analiza in razvoj področja redkih bolezni v Sloveniji - predstavitev Ciljnega raziskovalnega projekta 2015 - 2017. V: *Redke bolezni 2016* (6. izd.). Ljubljana: Društvo bolnikov s krvnimi boleznimi Slovenije.
4. Valdez R, Ouyang L, Bolen J: Public health and rare diseases: oxymoron no more. *Prev Chronic Dis* 2016, 13: E05.
5. Rare diseases task force: *Disease registries in Europe*. Orphanet 2011: Rare Diseases collection.
6. Choquet R, Landais P: The French national registry for rare diseases: an integrated model from care to epidemiology and research. *Orphanet J Rare Dis* 2014, 9(1): O7.
7. Taruscio D, Kodra Y, Ferrari G *et al.*: The Italian National Rare Diseases Registry. *Blood Transfus* 2014, 12(3): 606-613.
8. Alonso V, Abaitua I, Zurriaga Ó *et al.*: National rare diseases registry in Spain: pilot study of the Spanish Rare Diseases Registries Research Network (SpainRDR). *Orphanet J Rare Dis* 2014, 9(Suppl 1): P5.
9. Cisarik F, Barakova A: The creation of the national registry of rare diseases in the Slovak Republic. *Acta Fac Phar Univ Comen* 2014, 61(1): 9-15.
10. Taruscio D, Gainotti S, Mollo E *et al.*: The current situation and needs of rare disease registries in Europe. *Public Health Genomics* 2013, 16(6), 288-298.
11. Patton M: *Qualitative evaluation and research methods* (2nd ed.). Thousand Oaks 1990: Sage Publications.
12. Yin R: *Case study research: design and methods* (3rd ed.). Thousand Oaks 2003: Sage Publications.
13. Nacionalni inštitut za javno zdravje – NIJZ: *Mednarodna klasifikacija bolezni*. Ljubljana 2016: NIJZ.
14. Nacionalni inštitut za javno zdravje – NIJZ: *Spremljanje bolnišničnih obravnav*. Ljubljana 2015: NIJZ.
15. Rarecare: *Projekt Rarecare. Pobuda za opredelitev in širjenje informacij o bremenu redkih rakov v Evropi*. Ljubljana 2011: Onkološki inštitut.
16. Zakon o zbirkah podatkov s področja zdravstvenega varstva (ZZPPZ). *Uradni list RS*, št. 65/00 in 47/15.
17. Klinični inštitut za medicinsko genetiko UKC Ljubljana – KIMG: *Center za nediagnostičirane redke bolezni*. Ljubljana 2016: UKCL.
18. *National rare disease plan for Ireland*. Dublin 2016: Department of health of the Republic of Ireland.
19. Nacionalna kontaktna točka za redke bolezni – NKT: *Evropska priporočila za razvoj kontaktnih točk*. Ljubljana 2016: UKCL.
20. Aymé S, Bellet B, Rath A: Rare diseases in ICD11: making rare diseases visible in health information systems through appropriate coding. *Orphanet J Rare Dis* 2015, 10(1): 35.
21. Vlada Republike Slovenije: *Program dela Vlade RS za leto 2016*. Ljubljana 2016: Vlada republike Slovenije.
22. Zaletel-Kragelj L, Eržen I, Premik M: *Uvod v javno zdravje*. Ljubljana 2011: Medicinska fakulteta.
23. Gliklich RE, Dreyer NA: *Registries for evaluating patient outcomes: a user's guide*. Rockville 2010: AHRQ.
24. Register raka: *Epidemiologija in register raka*. Ljubljana, 2017: Onkološki inštitut.
25. The Epirare project: *EPIRARE (European platform for rare disease registries)*. Brussels 2011: European Commission.
26. Gliklich RE, Dreyer NA, Leavy MB: *Registries for evaluating patient outcomes: A User's Guide* (3rd ed.). Rockville 2014: AHRQ.
27. Richesson RL, Shreff D, Andrews JE: Standardization of Questions in Rare Disease Registries: The PRISM Library Project. *Interact J Med Res* 2012, 1(2): e10.
28. Cross Border PATient REGistries iNiTiative – PARENT: *Methodological guidelines and recommendations for efficient and rational governance of patient registries*. Brussels 2015: European Commission.
29. Zoni AC, Berjón MD, Barceló E *et al.*: Identifying data sources for a national population-based registry: the experience of the Spanish Rare Diseases Registry. *Public Health* 2015, 129(3): 271-275.
30. Choquet R, Maaroufi M, de Carrara A *et al.*: A methodology for a minimum data set for rare diseases to support national centers of excellence for healthcare and research. *J Am Med Inform Assoc* 2015, 22(1): 76-85.
31. European Union committee of experts on rare diseases – EUCERD: *Patient registries and databases in the*

- field of rare diseases*. Brussels 2016: European Commission.
32. Žerjav Tanšek M: *Urejanje področja redkih bolezni v Evropi in Sloveniji*. Ljubljana 2015: Glasilo Zdravniške zbornice Slovenije.
 33. European Union committee of experts on rare diseases – EUCERD: *Relevant coding and classification of RD in international nomenclatures*. Brussels 2015: European Commission.
 34. Commission expert group on rare diseases – CEGRD: *Recommendation on ways to improve codification for rare disease in health information systems*. Brussels 2014: European Commission.
 35. Žerjav Tanšek M: Evropske pobude na področju redkih bolezni. *Redke bolezni* 2015 (5. izd.). Ljubljana: Društvo bolnikov s krvnimi boleznimi Slovenije.
 36. Mazzucato M, Visonà Dalla Pozza L, Manea S *et al.*: A population-based registry as a source of health indicators for rare diseases: the ten-year experience of the Veneto Region's rare diseases registry. *Orphanet J Rare Dis* 2014, 9(1), 37.
 37. Zakon o varstvu osebnih podatkov (ZVOP-1). *Uradni list RS*, št. 94/07.
 38. Evropska komisija: *Varstvo osebnih podatkov*. Bruselj 2016: Evropska komisija.
 39. Taruscio D, Mollo E, Gainotti S *et al.*: The EPIRARE proposal of a set of indicators and common data elements for the European platform for rare disease registration. *Archives of Public Health* 2014, 72(1): 35.
 40. Eurordis: *Rare disease patient registries. Policy fact sheet*. Brussels 2013: European Commission.
 41. Premik M: *Uvod v epidemiologijo*. Ljubljana 1998: Medicinska fakulteta.
 42. Posada de la Paz M, Groft SC: *Rare diseases epidemiology* 2010 (Vol. 686). New York: Springer Science & Business Media.
 43. Groft SC: Rare diseases research: expanding collaborative translational research opportunities. *Chest* 2013, 144(1): 16-23.

■ *Pregledni znanstveni članek*

Matevž Kastrin, Dimitar Hristovski, Andrej Kastrin

Odkrivanje zakonitosti iz literature na področju znanosti o življenju

Povzetek. Odkrivanje zakonitosti iz literature (OZL) je razvijajoče se znanstveno področje, ki se ukvarja s samodejnim odkrivanjem novega znanja iz bibliografskih podatkov. V prispevku najprej naredimo kratek pregled problematike znanstvene ustvarjalnosti, ki je neločljivo povezana s področjem OZL. OZL naslonimo na Mednickovo teorijo ustvarjalnosti in Koestlerjev koncept bisociacij. V nadaljevanju predstavimo glavne teoretične modele OZL. Največji del prispevka posvetimo pregledu aplikacij s področja OZL, ki jih razdelimo glede na osnovni mehanizem odkrivanja novega znanja: aplikacije, ki temeljijo na načelu sopoavnosti, in aplikacije, ki temeljijo na semantičnih povezavah med biomedicinskimi koncepti.

Literature-Based Discovery in the Field of Life Sciences

Abstract. Literature-based discovery (LBD) is a growing scientific field that deals with automatic discovery of new knowledge from bibliographic data. We first make a brief overview of the problem of scientific creativity, which is inextricably associated with the field of LBD. We base LBD on Mednick's theory of creativity and Koestler's concept of bisociations. Next, we present the main theoretical models of LBD. The largest part of the paper is devoted to a comprehensive review of LBD applications. We categorise LBD applications according to the main mechanism that provides knowledge discovery: applications based on the co-occurrence principle, and applications based on semantic relations among biomedical concepts.

■ **Infor Med Slov** 2017; 22(1-2): 22-31

Institucije avtorjev / Authors' institutions: Klinični oddelek za travmatologijo, Kirurška klinika, Univerzitetni klinični center Ljubljana (MK); Inštitut za biostatistiko in medicinsko informatiko, Medicinska fakulteta, Univerza v Ljubljani (DH, AK).

Kontaktna oseba / Contact person: dr. Andrej Kastrin, Inštitut za biostatistiko in medicinsko informatiko, Medicinska fakulteta, Univerza v Ljubljani, Vrazov trg 2, 1000 Ljubljana, Slovenija. E-pošta / E-mail: andrej.kastrin@mf.uni-lj.si.

Prispelo / Received: 18. 4. 2017. Sprejeto / Accepted: 18. 1. 2018.

Uvod

Pozornost raziskovalcev na področju znanosti o življenju se je pred dobrima dvema desetletjema tudi na eksperimentalni ravni preselila iz zanimanja za posamezne biološke koncepte (npr. gene in proteine) k poskusu razumevanja celih bioloških sistemov. Rezultat tega preskoka je med drugim vedno večja količina podatkov, ki jo je potrebno ne le analizirati, pač pa tudi smiselno interpretirati. To zahteva sposobnost integracije podatkovnih virov z obstoječim znanjem, ki je dostopno v literaturi.

Informacijski viri na področju biomedicine so zelo obsežni in v veliki večini primerov tudi prosto dostopni. MEDLINE, najobsežnejša bibliografska zbirka na področju biomedicine, obsega že več kot 25 milijonov bibliografskih zapisov, s prirastkom od 2.000 do 4.000 zapisov dnevno. Pred raziskovalci je zato zahtevna naloga. Integracija podatkovnih virov z eksperimentalnimi podatki zahteva dobro poznavanje in redno spremljanje različnih raziskovalnih smeri, kar je za človekov spoznavni aparat prehud zalogaj. Orodja, ki omogočajo (pol)samodejno rudarjenje po besedilih znanstvenih člankov so zato primarnega pomena za nadaljnji razvoj znanosti o življenju.

Orodja za rudarjenje besedilnih podatkov nam v grobem pomagajo pri štirih, zaporednih in medsebojno povezanih procesih.^{1,2} Prvič, raziskovalcu omogočajo identifikacijo ustreznih virov znanja oz. priklic informacij (angl. *information retrieval*). Drugič, olajšajo identifikacijo (bioloških) entitet oz. konceptov (npr. genov in proteinov) v besedilu (angl. *entity recognition*). Tretjič, pomagajo pri ekstrakciji predoločeni entitet in relacij (npr. eksplicitnih relacij med geni) iz besedila (angl. *information extraction*). Četrto sklop procesov se nanaša na sisteme za odkrivanje zakonitosti iz literature (OZL), ki temeljijo na prepoznavi posrednih relacij med posameznimi biološkimi koncepti (angl. *literature-based discovery*). Za razliko od prvih treh, je področje OZL najmanj razvito, polno raziskovalnih izzivov in zato glavni predmet tega pregleda.

Znanstvena ustvarjalnost

Ustvarjalnost je pomembna intelektualna aktivnost, pri kateri tvorimo nove ideje in povezave med njimi, ter obenem temeljno gonilo človeškega razvoja. Stari Grki so sicer zasejali seme dvoma, da je ustvarjalnost povezana z norostjo, vendar take popreproščene definicije danes nihče več ne jemlje resno. Brez ustvarjalnosti bi človek ostal lačen napredka in bi pri reševanju problemskih nalog ponavljal vedno enake vzorce vedenja.

Znanost običajno opredelimo v smislu sistematičnega preučevanja naravnih in družbenih pojavov. Hipotetično-deduktivni pristop, ki se dandanes večinoma uporablja v znanosti, je sestavljen iz več medsebojno povezanih korakov:³ (i) opredelitev problema, (ii) identifikacija raziskovalnega vprašanja, (iii) formulacija raziskovalne domneve, (iv) oblikovanje pričakovanega rezultata, (v) izvedba poskusa, (vi) primerjava pričakovanega in dejanskega rezultata ter (vii) oblikovanje sklepa. Odkritje je pogoj *sine qua non* v znanosti in predstavlja vrh človekove ustvarjalne misli. Odkritja omogočajo raziskovalcem spoznavanje sveta okoli sebe, osmišljajo realnost in predstavljajo meje mogočega. Odkritje se lahko nanaša na rešitev znanega problema ali pa se nanaša na problem, ki se šele kasneje izkaže za pomembnega. Pomembna odkritja se običajno dogajajo na mejah znanega, v prostoru brez domenskih ekspertov in ustreznega raziskovalnega instrumentarija. Možgani raziskovalca v tem prostoru delujejo po principu selektivnih poskusov in napak ter na osnovi hevrstičnih pravil, ki so osnovana na preteklih izkušnjah.⁴ Generiranje novih idej je ključni element ustvarjalnega procesa znanstvenega spoznavanja realnosti.^{5,6}

Sodobno znanstveno raziskovanje je ena od najbolj kompleksnih človeških aktivnosti, ki zahteva različne tipe sposobnosti. Znanstveno ustvarjalnost tako lahko vidimo kot skupek kognitivnih in računskih konceptov, med katere uvrščamo (i) motivacijo za raziskovalno delo, (ii) sposobnost pravilne formulacije raziskovalnega problema znotraj obstoječe zakladnice znanja, (iii) sposobnost konstrukcije izčrpnega iskalnega prostora za rešitev raziskovalnega problema, (iv) sposobnost implementacije hevrstik za redukcijo prostora iskanja ter (v) potrpljenje in vzdržljivost.⁷ Kljub neustavljivi želji, da bi bili sposobni razumeti dinamiko znanstvenih odkritij ali bili celo sposobni njihovih napovedi, je naše znanje o tem še zelo skopo.⁸ Nekatera odkritja so zaradi vidne akumulacije teoretičnih spoznanj in empiričnih dokazov kot na dlani. Dober primer je npr. nedavna potrditev obstoja gravitacijskih valov. Po drugi strani se zdijo nekatera odkritja popolnoma nepredvidljiva; tak primer je npr. odkritje penicilina.

Element novosti je glavni sestavni del znanstvene ustvarjalne misli, vendar je za nas ključnega pomena spoznanje, da so sestavni bloki novih idej pogosto že vključeni v obstoječem znanju. Uzzi⁹ namreč na osnovi obširne bibliografske analize dobrih 18 milijonov znanstvenih člankov ugotavlja, da sodobna znanost temelji na novih, a močno konvencionalnih kombinacijah obstoječega znanja. Posledično

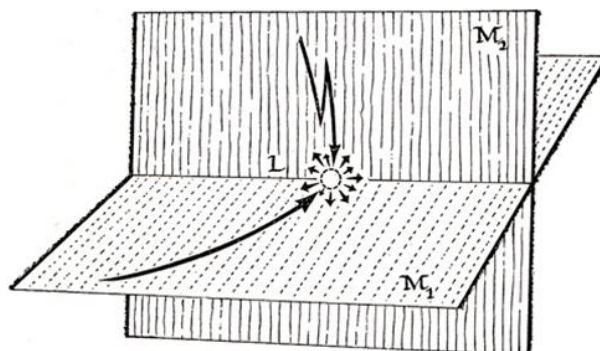
ugotavlja, da bo znanstveni članek, ki je zmes novega in konvencionalnega, imel zelo visoko verjetnost, da bo visoko citiran. Znanstveno ustvarjalnost si zato lahko predstavljamo kot kontinuum, kjer na enem koncu domuje novost, na drugem pa konvencionalnost.

Teoretičnih modelov, ki pojasnjujejo človekovo ustvarjalnost, je ogromno, njihov pregled pa bi presegal okvire tega prispevka. Za naše potrebe je dovolj, da iz zgodovinskega spomina obudimo Mednickovo¹⁰ teorijo ustvarjalnosti in Koestlerjev¹¹ model bisociacij. Mednick je zasnoval teoretičen model ustvarjalnosti, v katerem je izpostavil pomen asociacij pri porajanju novih idej. Proces ustvarjalnega mišljenja definira kot povezovanje asociacijskih elementov v nove kombinacije, ki so na svojstven način uporabne. Bolj kot sta si elementa v asociacijskem prostoru oddaljena, bolj ustvarjalen je proces oziroma njegova rešitev. Do novega odkritja po Mednicku lahko pridemo po slučaju ali pa zaradi podobnosti asociacijskih elementov.

Koestlerjev osnovni koncept bisociacij je predstavljen na sliki 1. Diagram predstavlja dve domeni znanja M_1 in M_2 , kot ravnini, ki sta pravokotni druga na drugo. Vsaka od ravnin se nanaša na svoj asociativni kontekst. Koestler definira ustvarjalnost kot bisociacijo, povezavo med elementi iz dveh različnih asociativnih kontekstov. Z drugimi besedami je bisociacija nov kos znanja, ki povezuje dve domeni znanja na doslej še nepovezan način. Formalno, med dvema konceptoma obstaja bisociacija, če: (i) nista koncepta evidentno neposredno povezana, (ii) vsak koncept prihaja iz svoje domene in (iii) nova povezava odkriva nov, še nepoznan, pogled na problemsko domeno. Iz zgodovine je najbolj znan primer Arhimeda, ki je poskušal ugotoviti, ali je zlatu v kroni primešano srebro, ne da bi krono kakorkoli poškodoval. Prvi asociacijski kontekst je predstavljal problem, kako izmeriti prostornino nepravilno oblikovanega predmeta. Drugi asociacijski kontekst je bila Arhimedova ugotovitev, da je prostornina vode, ki se prelije iz kadi, enaka prostornini kopalca. S tem spoznanjem Arhimedu ni bilo težko rešiti naloge s krono, tako da je vzpostavil bisociacijo med obema asociacijskima kontekstoma in tako razkril zlatarjevo goljufijo. Do podobnega »heureka« momenta je prišlo ob iznajdbi metode verižne reakcije s polimerazo, katere odkritju je botrovala premetena kombinacija dotedanjih dobro poznanih tehnik v biokemiji.

Dolgo časa je veljalo, da je racionalen model odkrivanja znanja nemogoč. Raziskovalci na področju umetne inteligentnosti so si prizadevali proces odkrivanja znanja natančno opisati s pomočjo

algoritmov in ga čim bolj avtomatizirati. Na nekaterih področjih jim je to celo uspelo. V okviru Microsoftovega projekta Adam so npr. raziskovalci razvili avtomatiziran sistem za prepoznavo slik, ki temelji na metodi globokega učenja (angl. *deep learning*).¹² Pasma psa je Adam sposoben prepoznati na podlagi enega samega posnetka. Na področju bibliometrije so raziskovalci s pomočjo računalnika poskušali simulirati pomembna zgodovinska odkritja, drugi pa so se usmerili v računsko odkrivanje novega znanja.¹³ Čeprav že sam termin »računsko odkrivanje znanja« implicira avtomatiziran proces, bolj podroben pregled literature razkriva, da pri praktično vsakem uspešnem poskusu človek igra pomembno vlogo. Proces računskega odkrivanja znanja lahko podobno kot pri človeku opišemo z naslednjimi koraki: (i) formulacija problema, (ii) reprezentacija problema, (iii) manipulacija podatkov, (iv) manipulacija algoritmov ter (v) filtriranje rezultatov in njihova interpretacija.



Slika 1 Ravnini predstavljata nepovezana asociacijska konteksta (domeni znanja) M_1 in M_2 . Koestler definira ustvarjalnost kot bisociacijo, povezavo med dvema elementoma iz različnih asociativnih kontekstov.

Znanstvena skupnost je danes kompleksen ekosistem, s stotinami bolj ali manj prepletenih raziskovalnih področij, desettisoči raziskovalcev in vrtoglavim številom znanstvenih publikacij. Znanstveni napredek bi bil seveda idealen, če bi imel raziskovalec pregled nad vso relevantno literaturo s svojega znanstvenega področja. Wyatt¹⁴ je že v devetdesetih letih prejšnjega stoletja ugotovil, da se število biomedicinskih revij podvoji približno na vsakih 20 let. V pogled v PubMed npr. razkriva, da je danes o diabetesu napisanih dobrih 575.000 znanstvenih člankov. Predstavljajte si diabetologa, ki bi vsak dan prebral 20 člankov. Da bi se prebil čez celo to skladovnico člankov, bi potreboval kar 80 let! Danes velja, da je praktično nemogoče biti odlični ekspert na več področjih hkrati. Znanosti o življenju čedalje bolj stremijo k podatkovno intenzivnemu raziskovanju. Kljub temu se še vedno zanašajo na tradicionalne metode

raziskovanja, ki med drugim slonijo na ročnih interpretacijah domenskega eksperta, v upanju, da bo našel uporabne vzorce v podatkih. Avtomatizacija znanstvenega procesa je zato ne samo teoretično, pač pa tudi povsem praktično vprašanje.¹⁵ Glavna predpostavka avtomatskega odkrivanja zakonitosti iz literature je uporabniku omogočiti samodejno iskanje novega, potencialno uporabnega in razumljivega znanja.

Swanson in odkrivanje zakonitosti iz literature

Zelo površno rečeno je znanstvena literatura na področju znanosti o življenju nestrukturiran skupek konceptov, kot so npr. ključne besede, imena genov, zdravil, terapevtskih postopkov in relacij med njimi. Akumulacija znanja temelji na njihovi učinkoviti prepoznavi, ekstrakciji in transformaciji v dobro strukturirano in semantično stabilno obliko.

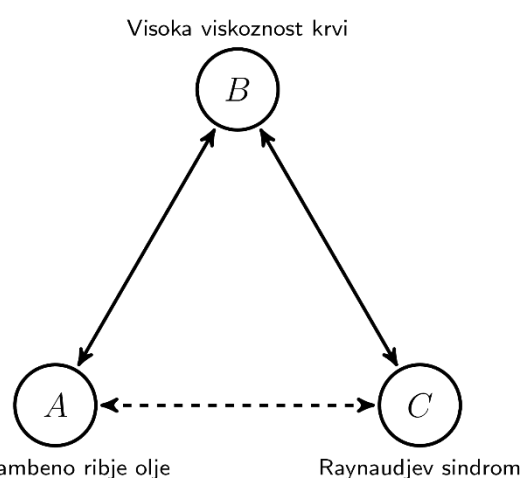
V sredini 80. let prejšnjega stoletja je Swanson¹⁶ sprožil novo vejo interdisciplinarnega raziskovanja. Pri OZL je šlo za slučajno odkritje, ki je prineslo nov veter v jedra tehnološko osveščenih raziskovalcev na področju znanosti o življenju, po drugi strani pa priložnost za razvoj novih računalniških aplikacij. Podrobnejši opis Swansonovega odkritja sledi v naslednjem razdelku. Formalno OZL opredelimo kot avtomatiziran proces iskanja komplementarnih (vsebinsko dopolnjujočih) struktur med disjunktными množicami znanstvene literature.¹⁷

Teoretično ozadje Swansonovega pristopa k OZL lahko navežemo na Mednickovo¹⁰ teorijo ustvarjalnosti, s posebnim poudarkom na medkontekstnih asociacijah, ki jim Koestler¹¹ pravi bisociacije. Če bi iz množice člankov po slučaju potegnili dva, bi z veliko verjetnostjo ne bila komplementarna. Za ta namen zato potrebujemo avtomatiziran iskalni proces, ki bo karseda dobro izkoristil človekovo znanje in sposobnost presoje ter ju kombiniral s hitrostjo računalnika in njegovo učinkovitostjo pri obdelavi velikih količin podatkov. Koncept OZL je osnovan na treh medsebojno povezanih predpostavkah:¹⁶ (i) količina znanstvene literature je prevelika, da bi jo en sam raziskovalec lahko pregledal, (ii) raziskovalci so običajno specializirani za eno področje, brez dobrega vpogleda v druga raziskovalna področja in (iii) znanost je fragmentirana, sestavlja jo večje število bolj ali manj tesno povezanih skupnosti.

OZL je bilo doslej, z redkimi izjemami,¹⁸ uporabljeno zgolj na področju znanosti o življenju; Swanson¹⁶ pravi, da bržkone zaradi goste povezanosti konceptov

v njem, sami pa pristavljamo, da je glavni vzrok temu predvsem prosta dostopnost in enostavnost strojne manipulacije bibliografske zbirke MEDLINE. Swanson¹⁶ pojmuje OZL primarno kot človeško aktivnost, ki pa nujno potrebuje računalniško podporo. OZL zato vidi predvsem kot orodje za spodbujanje raziskovalčeve ustvarjalnosti pri tvorjenju plavzibilnih in preverljivih domnev.

Najbolj znan in v aplikacijah najpogosteje uporabljen je Swansonov¹⁷ *ABC* model odkrivanja znanja. Model je zasnovan na preprosti ideji, da lahko novo znanje odkrijemo med neposredno nepovezanima konceptoma (*A*) in (*C*) preko vmesnega koncepta (*B*). Shematsko je model prikazan na sliki 2.

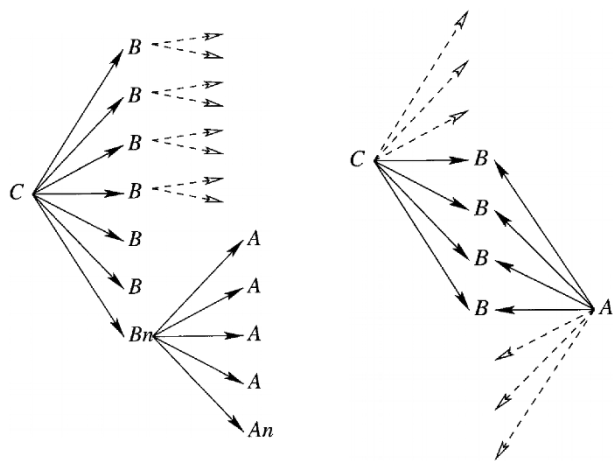


Slika 2 Swansonov *ABC* model odkrivanja znanja. Novo znanje predstavlja povezava med neposredno nepovezanima konceptoma (*A*) in (*C*) preko vmesnega koncepta (*B*).

Model *ABC* se uporablja v dveh izpeljankah:¹⁹ odprti in zaprti. Pri odprtem iskanju je glavni namen generiranje raziskovalne domneve. Običajno imamo na voljo raziskovalni problem. Proces iskanja začnemo z enim konceptom (*A*), kateremu priredimo množico konceptov (*B*). Postopek iskanja ponovimo za vsakega ob (*B*) konceptov. Dobljeni zadetki predstavljajo koncepte (*C*). Te filtriramo tako, da izločimo vse neposredne povezave s konceptom (*A*). Postopek odprtega iskanja je shematsko prikazan na sliki 3. Pri drugi obliki, zaprtem iskanju, raziskovalno domnevo le preverjamo. Začetni (*A*) in ciljni (*C*) koncept poznamo, iščemo pa vmesne koncepte (*B*). Tudi model zaprtega iskanja je shematsko prikazan na sliki 3.

Čeprav je model *ABC* največkrat uporabljen, pa ni edini. Wilkowski²⁰ je predlagal razširitev modela *ABC* tako, da je enovozliščne (*B*) koncepte nadomestil z večvzliščnimi vmesnimi koncepti. Model ima

oznako AnC , kjer je $n = (B_1, B_2, \dots, B_m)$. Van der Eijk²¹ je za OZL prvi predlagal uporabo modela omrežja ob predpostavki, da koncept sopojavnosti (angl. *co-occurrence*) zagotavlja celosten pogled na kompleksnost povezav med koncepti.



Slika 3 Odprt (levo) in zaprt (desno) proces odkrivanja novega znanja. Polne povezave vodijo do potencialno zanimivih konceptov, črtkane povezave pa do konceptov, ki so se ob naknadnem preverjanju izkazali za nezanimive.

Pregled sistemov za računalniško odkrivanje zakonitosti iz literature

Pristopi, osnovani na načelu sopojavnosti

Swanson²² je do svojega prvega odkritja prišel po naključju ob študiju literature o inuitski prehrani. Inuiti so staroselsko ljudstvo, ki prebiva na Grenlandiji, severnih obalah Kanade, Aljaski ter skrajnovzhodnem koncu Sibirije. Ob branju več kot 4.000 naslovov člankov v zbirki MEDLINE je prišel do spoznanja, da prehrabeno ribje olje znižuje viskoznost krvi, zmanjšuje strjevanje krvnih ploščic in inhibira odziv žilne stene. Hkrati je ugotovil, da zmanjšana viskoznost krvi in strjevanja krvnih ploščic ob hkratni inhibiciji odziva žilne stene preprečuje nastanek Raynaudovega sindroma. Pri Raynaudovem sindromu gre za občasna skrčenja manjših žilnih odvodnic, najpogosteje v prstih rok, lahko pa tudi na prstih nog, jeziku in nosu. Motnja v prekrvavitvi traja navadno nekaj minut do nekaj ur. Swanson je tako postavil domnevo, da prehrabeno ribje olje preprečuje nastanek Raynaudovega sindroma. Kasnejši klinični eksperiment je povezavo dejansko potrdil.²³ Kasneje je sam oz. s sodelavci opisal še več primerov odkritij. Večino od njih si oglejmo v nadaljevanju prispevka.

Pri iskanju citatov si je Swanson²² pomagal z iskalnikom Dialog SciSearch, s katerim je preiskal naslove in povzetke bibliografskih zbirk MEDLINE in Embase (Excepta Medica). Okoli 1.000 bibliografskih zapisov se je nanašalo na Raynaudov sindrom, 3.000 zapisov pa na prehrabeno ribje olje. Množico člankov je prečistil in izdvojil 489 citatov, od katerih so le štirje povezovali obe iskalni domeni. Še več, le dva citata sta govorila o Raynaudovem sindromu in ribjem olju, vendar ne v kontekstu Swansonovega odkritja. Swanson je zato skoval termin »neodkrito javno znanje« (angl. *undiscovered public knowledge*), s katerim je opisal fenomen nepovezanih, a logično komplementarnih delov znanja, ki se skrivajo v literaturi.

Dve leti kasneje je Swanson²⁴ opisal povezavo med migreno in magnezijem, temu pa je dodal še povezavo med somatomedinom C in argininom.²⁵ S Smalheiserjem²⁶ je opisal povezave med pomanjkanjem magnezija in različnimi možganskimi funkcijami, med indometacinom in Alzheimerjevo boleznijo,²⁷ estrogenom in Alzheimerjevo boleznijo²⁸ ter med od kalcija neodvisno fosfolipazo A2 in shizofrenijo.²⁹ Swanson in Smalheiser³⁰ sta bila prva, ki sta strokovni javnosti ponudila orodje Arrowsmith, prostodostopno aplikacijo za interaktivno OZL.

Gordon in Lindsay³¹ sta poskušala sistematično ponoviti Swansonovo domnevo "Raynaudov sindrom ↔ prehrabeno ribje olje" s pomočjo avtomatiziranega postopka odprtega odkrivanja. Pomagala sta si z metodami leksikalne statistike, tako da sta uporabila različne mere, kot so frekvenca posameznih besed, frekvenca dokumentov, v katerih nastopa posamezna beseda, in mera $TF \times IDF$.³² Za razliko od Swansona sta uporabila celoten zapis, vključno s povzetkom, in ne samo naslova zapisa v MEDLINE. Potrdila sta, da je avtomatizirano odkrivanje znanja na primeru Swansonove domneve možno. Za razliko od sistema Arrowsmith daje njun pristop večjo težo posameznim besedam in ne le dokumentom. Kasneje sta poskušala enako metodologijo uporabiti tudi za potrditev domneve "migrena ↔ magnezij", vendar jima je spodletelo.³³ Uspešna sta bila šele z novim pristopom, kjer sta uporabila model n -gramov.

Gordon in Dumais³⁴ izhajata iz predpostavke, da so v procesu OZL dobri kandidati tisti koncepti, ki so karseda podobni (semantično in statistično) z izhodiščnim konceptom. Kot metodo za odkrivanje podobnih konceptov predlagata latentno semantično indeksiranje (angl. *latent semantic indexing*, LSI). LSI definira vektorski prostor, v katerem semantično sorodni termini ležijo bolj skupaj. Proces iskanja

začnemo tako, da izhodiščnemu konceptu (*A*) s pomočjo LSI poiščemo karseda podobne koncepte (*B*). Urejeni seznam konceptov (*B*) ovrednoti domenski ekspert in izbere ožjo množico konceptov. Proces iskanja nato ponovimo še za iskanje konceptov (*C*). Z izjemo metodologije avtorja ne ponujata ničesar novega. Odkrivanje novega znanja na problemski nalogi odprtega tipa ovrednotita tako, da svoje rezultate primerjata z eksperimentom, ki sta ga izvedla Gordon in Lindsay,³¹ in potrđita komplementarnost obeh pristopov.

Weeber, Klein, denBerg in Vos¹⁹ so kot pomoč uporabniku pri odkrivanju znanja razvili dvostopenjski sistem DAD (angl. *Disease – Adverse drug reaction – Drug*), ki omogoča odprti in zaprti način odkrivanja znanja. DAD po mnenju njegovih avtorjev pomaga raziskovalcu na treh področjih: (i) močno zmanjša prostor iskanja novega znanja, (ii) pomaga pri interpretaciji rezultatov iskanja s pomočjo semantične analize in (iii) omogoča vpogled v tekstovni kontekst dane domneve. Eksperimentalno okolje predstavlja celotna zbirka MEDLINE zapisov. DAD uporablja sistem MetaMap,³⁵ ki prosto besedilo naslova in povzetka bibliografskega zapisa preslika v biomedicinske koncepte iz metatezavra UMLS. Glavna prednost uporabe konceptov je, da se različne jezikovne različice, sinonimi in izpeljanke preslikajo na eno entiteto; npr. simboli in termini IL12, IL-12 in interleukin 12 se vsi preslikajo v koncept Interleukin-12. DAD prav tako omogoča filtriranje konceptov po semantičnih tipih, kot jih najdemo v orodju UMLS Semantic Network. Mero povezanosti med koncepti (*AB*) in (*BC*) predstavlja število sopojavnosti danih konceptov v posameznih stavkih besedila. Sistem so ovrednotili s pomočjo Swansonovih domnev "Raynaudov sindrom ↔ ribje olje" in "migrena ↔ magnezij", vendar niso predstavili statističnih mer ustreznosti sistema. Glavna slabost sistema DAD je problem razreševanja dvoumnih konceptov; MetaMap npr. zmotno preslika termina »mg« (miligram) in »Mg« (magnezij) v koncept »magnezij«.

Stegmann in Grohmann³⁶ sta bazo znanja zgradila na podlagi sopojavnosti MeSH terminov. MeSH je kontroliran geslovník, ki vsebuje biomedicinske izraze (deskriptorje) na različnih nivojih specifičnosti. Za odkrivanje posrednih povezav med koncepti sta uporabila metodologijo strateškega diagrama, ki jo je prvi predlagal Callon.³⁷ Dvojice MeSH terminov najprej po podobnosti razvrstita v skupine, nato pa za vsako skupino izračunata njeno gostoto (moč povezav med MeSH termini znotraj ene skupine) in središčnost (moč povezav med MeSH termini med različnimi skupinami). Na primerih domnev

"Raynaudov sindrom ↔ ribje olje" in "migrena ↔ magnezij" sta odkrila, da vmesni koncepti (*B*) nastopajo v območju pod mediano gostote in središčnosti. Na novo sta tako odkrila povezavo med prioni in manganom, pri kateri kot vmesni koncept nastopijo fuzijski proteini.

Pratt in Yetisgen-Yildiz³⁸ sta predstavila sistem LitLinker, ki za odkrivanje novega znanja uporablja kombinacijo tehnik procesiranja naravnega jezika in algoritmov rudarjenja podatkov. Njun cilj je bil ponuditi sistem za odprto odkrivanje znanja, ki bo čim bolj avtomatiziran in neodvisen od uporabnika. Za ekstrakcijo znanja sistem uporablja le naslove citatov, katerih besede s pomočjo sistema MetaMap preslika v UMLS koncepte. Sistem množico konceptov nato prečisti, tako da odstrani (i) neinformativne koncepte, (ii) pomensko preveč sorodne koncepte ter (iii) koncepte, ki z vidika preučevanega problema niso zanimivi. Slednje naredi s pomočjo orodja UMLS Semantic Network, tako da za nadaljnjo analizo uporabi le tiste biomedicinske koncepte, ki ustrezajo podanemu semantičnemu tipu. Za razliko od prejšnjih sistemov, LitLinker za identifikacijo povezav med koncepti uporablja asociacijska pravila, natančneje algoritem Apriori. Ciljne koncepte grupira glede na število vmesnih konceptov. Dobljena skupina konceptov prevzame ime koncepta, katerega ime je najkrajše. Namen tega koraka je grupirati podobne koncepte tako, da lažje poiščemo ciljne koncepte. Za ovrednotenje sistema sta avtorja uporabila Swansonovo domnevo "migrena ↔ magnezij", vendar brez statističnega ovrednotenja. Glavna prednost sistema je iskanje vmesnih (*B*) konceptov, kjer uporabnikova pomoč ni potrebna.

Srinivasan³⁹ namesto surovih UMLS konceptov za odkrivanje znanja uporabi MeSH termine. Glavni namen njenega pristopa je čim bolj avtomatizirati proces odkrivanja novega znanja, brez pomoči uporabnika. Za izbran izhodiščni koncept (*A*) algoritem sestavi MeSH profil za izbrane semantične tipe. Profil predstavlja relativno pomembnost različnih MeSH terminov za množico izbranih semantičnih tipov. Za vsak semantični tip nato izberemo *n* najbolj zanimivih MeSH terminov; ti termini predstavljajo koncepte (*B*). Termine (*C*) dobimo tako, da enak postopek neodvisno ponovimo nad vsakim od terminov (*B*). Avtorica je uspešno ponovila iskanje nad dvema odprtima in petimi zaprtimi problemi. Glavna očitka temu pristopu sta: (i) nekateri, zlasti pa novi, bibliografski zapisi (še) nimajo pripisanih MeSH terminov in (ii) MeSH termini so pogosto bolj splošni od prostega besedila v naslovih in povzetkih bibliografskih zapisov ter od

UMLS konceptov, kar lahko močno ovira odkrivanje znanja.

Van der Eijk in sodelavci²¹ so uporabili preslikavo iz omrežja sopojavnosti v asociativni prostor konceptov (angl. *associative concept space*). Gre za večrazsežni evklidski prostor, v katerem so koncepti s pomočjo Hebbovega učnega algoritma razvrščeni tako, da so pri visoki sopojavnosti koncepti narisani skupaj in obratno, koncepti, ki redko nastopajo skupaj, pa so narisani daleč vsaksebi. Kombinacija omrežja sopojavnosti in prostorske predstavitve omogoča hitro iskanje poti v omrežju, saj lahko dolžino poti med dvema vozliščema ocenimo na podlagi evklidske razdalje v prostoru. Daljše poti (več kot en skok med konceptoma) kažejo na posredne relacije med koncepti in predstavljajo kandidatke za nove raziskovalne domneve. Sestavni del pristopa je vizualizacija rezultatov, kar predstavlja pomembno prednost pred ostalimi predhodno razvitimi pristopi. Avtorji so uspešnost algoritma pri odkrivanju novega znanja preverili na simuliranih in realnih podatkih iz bibliografske zbirke MEDLINE. Podobno kot pri ostalih, tudi tukaj pogrešamo poglobljeno statistično evalvacijo sistema.

Wren in sodelavci⁴⁰ izhajajo iz predpostavke, da je število konceptov (C) v klasičnem odprtem modelu odkrivanja znanja običajno zelo veliko. Relacije (A) \leftrightarrow (C) so zato rangirali s pomočjo modela slučajnega omrežja, tako da so izračunali razmerje med številom dejanskih povezav in pričakovanim številom povezav. Relacije, katerih količnik je presegel prazno vrednost, so predstavljale nove raziskovalne domneve, potencialno novo znanje. Omrežje povezav so zgradili tako, da so izdvojili koncepte, ki se nanašajo na gene, bolezni, fenotipe in kemične učinkovine. Z uporabo opisanega pristopa so avtorji predpostavili domnevo o obstoju vzročne povezave med srčno hipertrofijo in klorpromazinom. Klinični eksperiment na miših je domnevo potrdil. Statistične evalvacije sistema avtorji ne podajo.

Hristovski s sodelavci⁴¹ namesto besed iz naslovov in povzetkov bibliografskih zapisov za odkrivanje novih relacij uporabi MeSH deskriptorje. Za opisovanje znanih in napovedovanje potencialnih novih relacij avtorji uporabijo asociacijska pravila. Kot prvi v vrsti sistemov za OZL predstavijo tudi izčrpno statistično evalvacijo nad 10 izbranimi boleznimi. Kasneje Hristovski s sodelavci⁴² predstavi spletno aplikacijo BITOLA za interaktivno odkrivanje zakonitosti iz literature. Sistem je posebej prilagojen odkrivanju kandidatnih genov za bolezni, vendar se ga lahko uporablja tudi kot splošni generator novih raziskovalnih domnev na področju biomedicine. Baza

znanja je zgrajena na osnovi naslovov in povzetkov člankov, pripadajočih MeSH terminov ter imen genov. Število relacij (A) \leftrightarrow (C) je omejeno s filtriranjem po semantičnih tipih in kromosomski lokaciji gena. Statistične evalvacije sistema avtorji ne podajo, predstavijo pa praktični primer odkrivanja genov, kandidatov za polimikrogirijo.

Sistem Telemakus bazo znanja zgradi na osnovi ekstrakcije UMLS konceptov iz biomedicinske literature, tako da uporabi podatke iz tabel in slik polnih dokumentov.⁴³ Prednost sistema je velika natančnost, saj je del konceptov ekstrahirano ročno. V tem pa se skriva tudi slabost pristopa, saj je ročna ekstrakcija počasna in draga. Poleg tega je sistem omejen le na dve domeni: na omejevanje kalorij in biologijo staranja, vendar je po zagotovilih avtorjev razširljiv tudi na druge domene.

Na koncu tega razdelka ne smemo pozabiti na metodo RaJoLink, ki je plod domačega znanja.⁴⁴ Metoda implementira Swansonov model ABC tako, da samodejno predlaga kandidatne koncepte (A), ki so logično povezani s proučevanim fenomenom (C). Kandidatne koncepte izbira po načelu redkosti pojavljanja. Avtorji žal ne podajo statistične evalvacije sistema, prav tako ni dostopna računalniška aplikacija.

Pristopi osnovani na semantičnih povezavah

Doslej so vsi predstavljeni sistemi kot glavni mehanizem odkrivanja novih relacij uporabljali preprosto mero sopojavnosti. Sopojavnost ima vrsto pomanjkljivosti, med drugim (i) generira veliko število napačno pozitivnih relacij (A) \leftrightarrow (C) ter hkrati (ii) ničesar ne pove o vsebini relacije. Raziskovalce namreč poleg golega obstoja interakcije med koncepti zanimajo tudi mehanizmi, ki ležijo v ozadju ter vzročno-posledični odnosi med koncepti.⁴⁵

Hu s sodelavci⁴⁶ je bil prvi, ki je pri rudarjenju povezav med koncepti upošteval tudi semantiko. Razvili so sistem Bio-SbKDS, ki z uporabo semantičnega znanja močno zmanjša prostor preiskovanja plavzibilnih domnev v procesu odkrivanja novega znanja. V primerjavi s prejšnjimi pristopi Bio-SbKDS uporabi semantično omrežje iz paketa UMLS, s katerim poišče ustrezne relacije med koncepti, le te pa filtrira s pomočjo semantičnih tipov.

Hristovski s sodelavci⁴⁷ nato uvede inovativen koncept "vzorec odkrivanja", ki vsebuje množico pogojev, katerim mora zadoščati novo odkrita raziskovalna domneva. Pogoji so sestavljeni iz kombinacij semantičnih relacij med koncepti, ekstrahiranimi iz bibliografske zbirke MEDLINE. Semantične relacije med koncepti so pridobljene s

pomočjo orodij BioMedLee⁴⁸ in SemRep.⁴⁹ BioMedMee je zbirka modulov za ekstrakcijo genotipskih in fenotipskih entitet iz besedila. S kombinacijo slovničnih pravil ter slovarja posameznim besedam in besednim zvezam pripiše ustrezno semantično in sintaktično kategorijo. SemRep je sistem za procesiranje naravnega jezika, ki je namenjen identifikaciji semantičnih relacij iz biomedicinskih besedil. Semantična relacija ima v splošnem obliko (*subjekt, predikat, objekt*), kjer predikat izraža vsebino odnosa med subjektom in objektom. Uporabo vzorcev iskanja ilustrirajo na primeru Swansonove domneve "Raynaudov sindrom ↔ ribje olje", hkrati pa najdejo novo povezavo med Huntingtonovo boleznijo in inzulinom. Kasneje so iskanje ponovili še na primeru Parkinsonove bolezni in predlagali izbor antiepileptikov, ki vplivajo na povišanje funkcije GABA, za katero je znano, da je pri teh bolnikih močno znižana.⁵⁰ Ahlers in sodelavci⁵¹ so enako metodologijo uporabili v zaprtem modelu iskanja, kjer so poskušali ugotoviti vmesne koncepte med antipsihotiki (angl. *antipsychotic agents*) in zdravljenjem raka. Odkrili so pet biomolekul, opisali možne interakcije z obema začetnima konceptoma, evalvacije pa ne podajo. Možnost odkrivanja novega znanja z odprtim in zaprtim modelom iskanja in uporabo semantičnih relacij je v sistemu EpiphaNet predstavil Cohen (2010) s sodelavci.⁵² Gre za interaktivno aplikacijo, ki rezultate predstavi grafično v obliki omrežja konceptov. Zbirko znanja EpiphaNet zgradi na osnovi semantičnih predikatov iz baze SemMedDB.⁵³ Sistem lahko napoveduje sopojavnost parov konceptov ter s tem simulira proces odkrivanja novega znanja. EpiphaNet temelji na metodah distributivne semantike: povezave med pari konceptov napoveduje na osnovi metodologije reflektivnega slučajnega indeksiranja (angl. *reflective random indexing*). Wilkowski s sodelavci²⁰ je predstavil inovativen koncept pregledovalnega odkrivanja znanja, katerega namen je interaktivno odkrivanje znanja. Metoda uporabnika vodi skozi serijo vmesnih konceptov (*B*), kjer je ustreznost koncepta določena z njegovo stopnjo v omrežju. Znanje je opisano v obliki omrežja, v katerem vozlišča predstavljajo posamezne biomedicinske koncepte, povezave med vozlišči pa so opisane s semantičnimi relacijami iz sistema SemRep.

V zlati dobi funkcijske genomike se je več raziskovalcev ukvarjalo z izdelavo sistemov za pomoč pri interpretaciji eksperimentov DNA mikromrež.² V tem kontekstu zato omenjamo sistem SemBT, ki integrira rezultate poskusa z DNA mikromrežami s semantičnimi relacijami iz literature kot jih predlaga SemRep.⁵⁴ Mikromreže nam dajo vpogled v statistično značilno različno izražene gene pri danem

eksperimentalnem pogoju, iz semantičnih relacij pa lahko razberemo interakcije genov z ostalimi biomedicinskimi koncepti. V tem okviru nam SemBT lahko pomaga pri interpretaciji rezultatov mikromrež ali pa pri načrtovanju novih terapevtskih postopkov za proučevano bolezen. Generiranje domnev v zadnjem primeru temelji na dveh vzorcih iskanja: "inhibiraj prekomerno izražen gen" ter "stimuliraj premalo izražen gen". Ideja v ozadju prvega vzorca iskanja je najti tak koncept (npr. učinkovino), ki bo inhibiral gen, ki je prekomerno izražen. Avtorji v nadaljevanju ponujajo več raziskovalnih domnev, ki se navezujejo na zdravljenje Parkinsonove bolezni. Kot možni učinkovini za zdravljenje Parkinsonove bolezni se ponujata laclitaxel in quecetin, saj se je ob pregledu literature izkazalo, da inhibirata gen HSP27 (HSPB1), ki se je v eksperimentalni študiji izkazal za prekomerno izraženega v testni skupini.

Semantični MEDLINE je spletna aplikacija, ki omogoča iskanje po predikatih, ekstrahiranih s pomočjo sistema SemRep.⁵³ Avtorji aplikacije zagotavljajo bolj ali manj svežo posodobitev ekstrahiranih predikatov. Za dani koncept sistem ekstrahira ustrezne predikate in izloči neinformativne koncepte, kot je npr. "Human". Nato izbere najpogostejše predikate in jih uporabniku predstavi v obliki neusmerjenega omrežja, v katerem vozlišča predstavljajo argumente (objekte oz. subjekte), povezave pa se navezujejo na semantične relacije med njimi. Semantični MEDLINE je za potrebe OZL prvi uporabil Miller s sodelavci.⁵⁵ Uporabili so zaprt model odkrivanja znanja iz literature, v katerem so vnaprej predpostavili povezavo (*A*) ↔ (*C*), da je raven testosterona povezana s spanjem, tako da vpliva na etiologijo slabega spanca pri odraslih moških. S pomočjo pregledovanja semantičnih relacij so nato poskušali najti vzročni mehanizem, ki bi to povezavo pojasnil. Kot možna rešitev se je izkazal kortizol, s katerim lahko pojasnimo manjkajočo korelacijo med nizko ravniyo testosterona pri moških in slabo kvaliteto njihovega spanja.

SemPathFinder temelji na semantični analizi poti v omrežju, ki omogoča »pripovedovanje zgodb«. ⁵⁶ Pripovedovanje zgodb je potrebno razumeti v smislu sodobnega pristopa k predstavitvi podatkov, kot sprehajanje po poteh v omrežju, tako da skupaj povezujemo tiste pare nepovezanih konceptov, ki so si semantično blizu. SemPathFinder model *ABC* iskanja razširi, tako da namesto enega vmesnega koncepta vpelje serijo (*B*) konceptov. Izbira najustrenejše poti je še vedno prepuščena domenskemu ekspertu. Sistem je namenjen za odprto in zaprto odkrivanje znanja. Sistem so avtorji preizkusili na zaprtem modelu na domnevah

"Raynaudov sindrom ↔ ribje olje" in "migrena ↔ magnezij" ter potrdili uspešnost iskanja.

Zaključek

OZL je danes pomembna raziskovalna disciplina na področju rudarjenja tekstovnih podatkov. Kljub precejšnjemu napredku je glavni očitek vsem predstavljenim sistemom ta, da ne omogočajo popolnoma samodejnega rudarjenja in potrebujejo posredovanje človeka. Druga pomembna pomanjkljivost je uporaba metodologije za ustrezno ovrednotenje procesa OZL; večina sistemov ovrednotenje opravi kar na že preverjenih domnevah in zanemari statistično preverjanje veljavnosti.

Pregled, ki je za nami, poskuša biti celovit, vendar nikakor ni izčrpen. Zajeti poskuša tiste segmente OZL, ki so po mnenju avtorjev nujni za temeljit pregled vsebine. Menimo, da smo vanj vključili vse pomembne sklice na literaturo, ki bodo bralcu lahko v pomoč pri morebitnem nadaljnjem raziskovanju področja OZL.

Reference

- Jensen LJ, Saric J, Bork P: Literature mining for the biologist: From information retrieval to biological discovery. *Nat Rev Genet* 2006; 7: 119-29.
- Faro A, Giordano D, Spampinato C: Combining literature text mining with microarray data: Advances for system biology modeling. *Brief Bioinform* 2012; 13: 61-82.
- Lawson AE: What does Galileo's discovery of Jupiter's moons tell us about the process of scientific discovery? *Sci Educ* 2002; 11: 1-24.
- Simon HA, Newell A: Human problem solving: The state of the theory in 1970. *Am Psychol* 1971; 26: 145-159.
- Clement JJ: *Creative model construction in scientists and students: The role of imagery, analogy, and mental simulation*. New York, NY 2008: Springer.
- Nersessian N: *Creating scientific concepts*. Cambridge, MA 2008: MIT Press.
- Kocabas S: *Elements of scientific creativity*. AAAI Technical Report SS-93-01 1993.
- Clauset A, Larremore DB, Sinatra R: Data-driven predictions in the science of science. *Science* 2017; 355: 477-80
- Uzzi B, Mukherjee S, Stringer M *et al.*: Atypical combinations and scientific impact. *Science* 2013; 342: 468-72.
- Mednick AM: The associative basis of the creative process. *Psychol Rev* 1962; 69: 220-232.
- Koestler A: *The art of creation*. London 1964: Penguin Books.
- Microsoft research blog: *On Welsh Corgis, computer vision, and the power of deep learning*. <http://bit.ly/2pp9ZQB> (11.4.2017).
- Langley P: Computational support of scientific discovery. *Int J Hum Comput Stud* 2000; 53: 393-410.
- Wyatt J: Use and sources of medical knowledge. *Lancet* 1991; 338: 1368-73.
- King RD, Whelan KE, Jones FM *et al.*: Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* 2004; 427: 247-2252.
- Swanson DR: Literature-based knowledge discovery? The very idea. V: Bruza P, Weeber M (ur.), *Literature-based discovery*. Berlin 2008: Springer; 3-11.
- Swanson DR: Undiscovered public knowledge. *Libr Q* 1986; 56: 103-118.
- Gordon M, Lindsay RK, Fan W: Literature-Based Discovery on the World Wide Web. *ACM Trans Internet Tech* 2002; 2: 261-275.
- Weeber M, Klein H, De Jong-Van Den Berg LTW *et al.*: Using concepts in literature-based discovery: Simulating Swanson's Raynaud-fish oil and migraine-magnesium discoveries. *J Am Soc Inform Sci Tech* 2001; 52: 548-57.
- Wilkowski B, Fiszman M, Miller CM *et al.*: Graph-based methods for discovery browsing with semantic predications. *AMLA Annual Symposium Proceedings* 2011, 1514-23.
- van der Eijk CC, van Mulligen EM, Kors JA *et al.*: Constructing an associative concept space for literature-based discovery. *J Am Soc Inform Sci Tech* 2004; 55: 436-44.
- Swanson DR: Fish Oil, Raynaud's syndrom, and undiscovered public knowledge. *Perspect Biol Med* 1986; 30: 7-18.
- Digiaco RA, Kremer JM, Shah DM: Fish-oil dietary supplementation in patients with Raynaud's phenomenon: A double-blind, controlled, prospective study. *Am J Med* 1989; 86: 158-64.
- Swanson DR: Migraine and Magnesium: Eleven Neglected Connections. *Perspect Biol Med* 1988; 31: 526-557.
- Swanson DR: Somatomedin C and arginine: Implicit connections between mutually isolated literatures. *Perspect Biol Med* 1990; 33: 157-86.
- Smalheiser NR, Swanson DR: Assessing a gap in the biomedical literature: Magnesium deficiency and neurologic disease. *Neurosci Res Commun* 1994; 15: 1-9.
- Smalheiser NR, Swanson DR: Indomethacin and Alzheimer's disease. *Neurology* 1996; 46: 583.
- Smalheiser NR, Swanson DR: Linking estrogen to Alzheimer's disease: An informatics approach. *Neurology* 1996; 47: 809-10.
- Smalheiser NR, Swanson DR: Calcium-independent phospholipase A2 and schizophrania. *Arch Gen Psychiatry* 1998; 55: 752-3.
- Swanson DR, Smalheiser NR: An interactive system for finding complementary literatures: A stimulus to scientific discovery. *Artif Intell* 1997; 91: 183-203.
- Gordon MD, Lindsay RK: Toward discovery support systems: A replication, re-examination, and extension of Swanson's work on literature-based discovery of a connection between Raynaud's and fish oil. *J Am Soc Inform Sci* 1996; 47: 116-28.

32. Baeza-Yates R, Ribeiro-Neto B: *Modern information retrieval: The Concepts and Technology behind Search*. New York, NY 1999: ACM Press.
33. Lindsay RK, Gordon MD: Literature-based discovery by lexical statistics. *J Am Soc Inf Sci* 1999; 50: 575-87.
34. Gordon MD, Dumais S: Using latent semantic indexing for literature based discovery. *J Am Soc Inf Sci* 1998; 49: 674-85.
35. Aronson AR: Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. *AMLA Annual Symposium Proceedings* 2001, 17-21.
36. Stegmann J, Grohmann G: Hypothesis generation guided by co-word clustering. *Scientometrics* 2003; 56: 111-35.
37. Callon M, Courtial JP, Laville F: Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics* 1991; 22: 155-205.
38. Pratt W, Yetisgen-Yildiz M: (2003). LitLinker: Capturing connections across the biomedical literature. In: Gennari JH, Porter BW, Gil Y (ur.), *Proceedings of the 2nd International Conference on Knowledge Capture*, Sanibel Island, FL, oktober 23-25. New York, NY, 2003: ACM; 105-112.
39. Srinivasan P: Text mining: Generating hypotheses from MEDLINE. *J Am Soc Inform Sci Tech* 2004; 55: 396-413.
40. Wren JD, Bekeredjian R, Stewart JA *et al.*: Knowledge discovery by automated identification and ranking of implicit relationships. *Bioinformatics* 2004; 20: 389-98.
41. Hristovski D, Stare J, Peterlin B *et al.*: Supporting discovery in medicine by association rule mining in Medline and UMLS. In: Patel V, Rogers R, Haux R (eds.), *MEDINFO 2001 Proceeding of the 10th World Congress on Medical Informatics*, London, UK, September 2-5. Amsterdam, NL, 2001: IOS Press; 1344-1348.
42. Hristovski D, Peterlin B, Mitchell JA *et al.*: Using literature-based discovery to identify disease candidate genes. *Int J Med Informat* 2005; 74: 289-98.
43. Fuller SS, Revere D, Bugni PF *et al.*: A knowledgebase system to enhance scientific discovery: Telemakus. *Biomed Digit Libr* 2004; 1: 2.
44. Petrič I, Urbaničič T, Cestnik B *et al.*: Literature mining method RaJoLink for uncovering relations between biomedical concepts. *J Biomed Inform* 2008; 42: 219-27.
45. Cameron D, Bodenreider O, Yalamanchili H *et al.*: A graph-based recovery and decomposition of Swanson's hypothesis using semantic predications. *J Biomed Inform* 2013; 46: 238-51.
46. Hu X, Li G, Yoo I *et al.*: A semantic-based approach for mining undiscovered public knowledge from biomedical literature. In: Hu X, Liu Q, Skowron A, Lin TY, Yager RR, Zhang B (eds.), *Proceedings of the 2005 IEEE International Conference on Granular Computing*, Beijing, China, July 25-27. Piscataway, NJ, 2005: IEEE; 22-27.
47. Hristovski D, Friedman C, Rindflesch TC *et al.*: Exploiting semantic relations for literature-based discovery. *AMLA Annual Symposium Proceedings* 2006, 349-53.
48. Lussier Y, Borlawsky T, Rappaport D *et al.*: PhenoGO: Assigning phenotypic context to gene ontology annotations with natural language processing. *Pac Symp Biocomput* 2006; 64-75.
49. Rindflesch TC, Fiszman M: The interaction of domain knowledge and linguistic structure in natural language processing: Interpreting hypernymic propositions in biomedical text. *J Biomed Inform* 2003; 36: 462-477.
50. Hristovski D, Friedman C, Rindflesch TC *et al.*: Literature-based knowledge discovery using natural language processing. In: Bruza P, Weeber M (eds.), *Literature-based discovery*. Berlin 2008: Springer; 133-52.
51. Ahlers CB, Hristovski D, Kilicoglu H *et al.*: Using the literature-based discovery paradigm to investigate drug mechanisms. *AMLA Annual Symposium Proceedings* 2007, 6-10.
52. Cohen T, Whitfield GK, Schvaneveldt RW *et al.*: EpiphaNet: An interactive tool to support biomedical discoveries. *J Biomed Discov Collab* 2010; 5: 21-49.
53. Kilicoglu H, Shin D, Fiszman M *et al.*: SemMedDB: A PubMed-scale repository of biomedical semantic predications. *Bioinformatics* 2012; 28: 3158-60.
54. Hristovski D, Kastrin A, Peterlin B *et al.*: Combining semantic relations and DNA microarray data for novel hypotheses generation. V: Blaschke C, Shatkay H (ur.), *Linking Literature, Information, and Knowledge for Biology*. Berlin 2010: Springer; 53-61.
55. Miller CM, Rindflesch TC, Fiszman M *et al.*: A closed literature-based discovery technique finds a mechanistic link between hypogonadism and diminished sleep quality in aging men. *Sleep* 2012; 35: 279-285.
56. Song M, Heo GE, Ding Y: SemPathFinder: Semantic path analysis for discovering publicly unknown knowledge. *J Informetr* 2015; 9: 686-703.

Drago Rudel

Poročilo o aktivni udeležbi na mednarodnem kongresu Med-e-Tel 2017 v Luksemburgu

Kot že nekaj let zapored sem se v dneh od 5. 4. do 7. 4. 2017 udeležil mednarodnega kongresa **International Telemedicine and eHealth Forum, Med-e-Tel 2017** v Luksemburgu. Kongres z razstavo je organiziralo mednarodno združenje *International Society for Telemedicine and eHealth (ISfTeH)*, katerega član je tudi Slovensko društvo za medicinsko informatiko (SDMI).

Na kongresu sem sodeloval z referatom Rudel D, Epšek-Lenart M, Balorda Z *et al.*: *Use of home telemedicine to optimise insulin therapy in patients at their self-management of diabetes type 2*. V referatu sem prikazal razvojno-raziskovalna prizadevanja za nadgradnjo telemedicinskih storitev v SB Slovenj Gradec in v Sloveniji, kar sovpada s področjem mojega dela v okviru SDMI – storitve zdravja na daljavo.

Na konferenci, ki se je jo udeležilo okoli 450 delegatov, je svoje razvojno-raziskovalne in tržne dosežke predstavilo več kot 250 udeležencev. Prispevki so pokrivali področji telemedicine in e-zdravja na primarnem in sekundarnem področju zdravstvene oskrbe, obvladovanje bolezni, pametnih tehnologij in podpore iz okolij za samostojno življenje (AAL), telerehabilitacije, e-učenja ter kliničnih disciplin, kot so telekardiologija, obvladovanje sladkorne bolezni na daljavo, telekirurgija, urgentna medicina, dolgotrajna oskrba na domu, telemedicinsko spremljanje bolnikov na daljavo, e-farmacija, mobilno zdravje in oskrba na daljavo. Nekateri smo predstavili delo v okviru evropskih oz. nacionalnih projektov, drugi nacionalne iniciative ali prispevke s področja večje uveljavitve žensk v e-zdravju. Nekaj prispevkov je bilo namenjenih

ekonomiki zdravstvene oskrbe (na daljavo), zakonski ureditvi področja e-zdravja, zdravja na daljavo in telemedicine, standardom, informacijski varnosti in etičnim dilemam. Poudarek je bil na vzpostavitvah telemedicinskih centrov za podporo izvajanju storitev. Pojavlja se vse več centrov, ki nudijo storitve komercialno. Niti eden od teh ni "nacionalni", za čemer težimo v Sloveniji. Kar nekaj udeležencev iz držav v razvoju z nizkim bruto družbenim proizvodom (Bangladeš, Indija, države centralne Afrike) je poročalo o uporabi novih zdravstvenih storitev, ki temeljijo na mobilni tehnologiji (mHealth).

Kot že v letu 2016, sem se sestal sem se s tajnikom in hkrati blagajnikom ISfTeH, g. Frankom Lievensom (<http://www.isfteh.org>). Ponovno se je zahvalil za zgledno sodelovanje SDMI kot nacionalnega združenja v ISfTeH in ponovno izrazil željo, da svoje pionirsko delo na področju uvajanja telemedicinskih storitev promoviramo v mesečnem e-glasilu ISfTeH Newsletter (<http://www.isfteh.org/media/category/newsletters>) in e-reviji *Journal of the International Society for Telemedicine and eHealth* (<http://journals.ukzn.ac.za/index.php/JISfTeH>), ki je razpošljejo na 60.000 e-naslovov. Povabilo g. Lievensa zato ponovno prenašam bralcem IMS!

Ocenjujem, da je bila moja udeležba na ISfTeH 2017 koristna tako kot promocija dela in rezultatov Slovenije kot zaradi mreženja za morebitno razvojno sodelovanje. UO SDMI se zahvaljujem za kritje dela stroškov moje udeležbe na Med-e-Tel 2017.

■ **Infor Med Slov** 2017; 22(1-2): 32

Emma Dornik

Digitalno zdravje – priložnosti in izzivi: poročilo s srečanja Sekcije za informatiko v zdravstveni negi 2017

Tradicionalno srečanje članov Sekcije za informatiko v zdravstveni negi (SIZN), ki deluje pri Slovenskem društvu za medicinsko informatiko (SDMI), je potekalo v Zrečah 20. oktobra 2017. Tudi tokrat je bil programski sklop Sekcije za informatiko v zdravstveni negi (SIZN) namenjen strokovni rasti, izobraževanju in druženju članov.

V nadaljevanju so predstavljeni povzetki predstavitev v zaporedju, kot so si sledili po programu.

Matija Ravnikar: Informacijska podpora zdravstveni negi žilnih katetrov in kirurških ran

Jedro informatizacije zdravstva ostaja podpora poslovnim procesom, ki so časovno zahtevni in vsebinsko izjemno kompleksni. V luči tega sta razvoj in vpeljava novosti, ki naj bi nadomestila sedanje, na papirju vodene postopke, polna izzivov tako na strani vsebine kot tudi pri implementaciji rešitve na terenu.

Pri načrtovanju informacijske podpore spremljanja stanja žilnih katetrov in kirurških ran smo kmalu ugotovili, da bo za učinkovito in uporabnikom prijazno rešitev potrebno predvideti širšo sliko, kot jo uporabljamo v programu Birpis. Ključni razlog je, da se v osnovni aplikaciji vsi dogodki vodijo le znotraj enega ambulantnega obiska ali hospitalizacije pacienta. Vstavljen centralni venski kateter pa je nekaj, kar pacienta lahko spremlja tako pri premestitvah znotraj in zunaj ene ustanove kot tudi pri odpustu v domačo oskrbo. Ustvarili smo povsem modul znotraj Birpis-a in ga poimenovali "Zdravstvena nega". Znotraj modula lahko spremljamo stanje arterijskih in venskih katetrov, kirurških ran, stalno pa dodajamo nove podtipе nege (dreni, razjede zaradi pritiska, oviranost ipd.). Pri vsakem podtipu spremljamo specifičen nabor podatkov, ki je odvisen tako od lastnosti podtipa kot tudi od izvajalca storitev, saj se postopki med njimi lahko razlikujejo.

Spremljamo proces nastanka podtipa, s premišljeno zasnovo pa smo vsakodnevno izpolnjevanje obrazcev za spremljanje stanja podtipa zmanjšali na hiter in

enostaven vnos ključnih parametrov zdravstvene nege posameznega podtipa. Ob tem beležimo tudi material, porabljen za nego katetrov in ran, ter omogočamo vnos podatkov o morebitni pojavitvi okužbe na mestu katetra ali rane. Z uporabo naprednih komponent in premišljeno programsko logiko smo vnos podatkov poenostavili v največji možni meri, hkrati pa obdržali ključno lastnost katetrov in ran: pacienta lahko premestijo, lahko ga odpustijo v drugo ustanovo ali pa tudi v domačo oskrbo, pa bo program ob sprejemu na drugi oddelek ali pri ponovnem pregledu v ambulanti uporabnika opozoril, da ima pacient npr. aktivno kirurško rano, ki jo je potrebno pregledati in po potrebi zaključiti z vnosom vseh zahtevanih podatkov. Namen vsega tega je, da zagotovimo pravočasen vnos kvalitetnih podatkov, ki so osnova za kakovostno spremljanje obravnave pacientovega stanja. Tako vnesene podatke lahko uporabimo za izdelavo naprednih analiz in statistično analizo podatkov, hkrati pa omogočimo enostaven in pregleden nadzor nad vsemi pacienti, ki jih spremljamo v okviru zdravstvene nege.

Sebastijan Forstnerič, Nino Fijačko, Gregor Štiglic: Vzpostavitev elektronskega naročanja v specialistični ambulanti

Uporaba interneta za zdravstvene namene vztrajno raste, saj ljudje vse pogosteje dostopajo do spleta in vse pogosteje delijo svoje izkušnje ter iščejo informacije o zdravstvenih storitvah. Sočasno v Sloveniji poteka uvajanje rešitev projekta eZdravje v zdravstveni sistem. V raziskavi smo se osredotočili na rešitev eNaročanje in raziskali časovno obremenjenost medicinske sestre pri posameznih metodah naročanja pacienta ter raziskali težave in pomanjkljivosti, ki nastajajo pri sami vzpostavitvi in uporabi elektronskega naročanja pacientov.

Uporabili smo kvantitativno metodo dela. Zajeli smo vzorec vseh naročenih pacientov v mesecu maju 2017 v specialistični ambulanti na območju severovzhodne Slovenije. Zbirali smo podatke o času, ki je potreben

za izvedbo različnih metod naročanja pacienta, in jih interpretirali s pomočjo opisne statistike.

Ugotovili smo, da je med vsemi metodami naročanja v specialistični ambulanti za medicinsko sestro najmanj obremenjujoča metoda naročanje preko sistema eNaročanje. Za naročanje po telefonu in osebno medicinska sestra porabi največ časa. Med vsemi naročenimi (n = 298) je eNaročanje izbralo 60 (20 %) pacientov, osebno se je naročilo 62 (21 %) pacientov in preko telefona 176 (59 %) pacientov. Pri uporabi sistema eNaročanje so prisotne tudi določene pomanjkljivosti, kot so npr. nemožnost izbire vseh obstoječih zdravstvenih storitev, možnost napačne izbire zdravstvene storitve zaradi različnih vrst naročanja in nezmožnost naročanja na kontrolne preglede, kar lahko vpliva na izbiro naročanja.

Z raziskavo smo prikazali, da je eNaročanje pacienta za medicinsko sestro najmanj obremenjujoče, učinkovitejše in časovno manj potratno. Kljub manjšim pomanjkljivostim eNaročanje predstavlja uporabno rešitev ter prinaša veliko prednosti za zdravstvene delavce in za paciente.

Nino Fijačko, Gregor Štiglic, Dejan Dinevski: Prikaz protokola učenja oralne higijene z igrifikacijo na daljavo

Slaba oralna higiena predstavlja velik javnozdravstveni problem tako v Evropi kot v Sloveniji. Osvajanje pravih veščin ščetkanja zobovja kot najpogostejše tehnike oralne higijene predstavlja tako za otroka kot tudi za starše svojevrsten izziv. Medicinske sestre imajo tukaj zelo pomembno vlogo pri promociji pravilne oralne higijene. Ohranjanje dobre oralne higijene pri otroku je zahtevna naloga, zato predstavlja iskanje novih motivacijskih strategij v obliki učenja z elementi igre oz. igrifikacije potencialno rešitev na področju oralne higijene. V tem prispevku želimo prikazati protokol učenja oralne higijene z igrifikacijo na daljavo.

Z uporabo igrifikacije želimo predstaviti naprednejši način učenja oralne higijene v domačem okolju in predstaviti možnost preverjanja učinkovitosti oralne higijene na daljavo.

V raziskavo je bil vključen otrok s sekundarno denticijo, ki uporablja ročno zobno ščetko in si sam ščetka zobovje. Protokol je sestavljen iz šest faz. V prvi fazi se zjutraj izvede oralna higiena z zobno ščetko in zobno pasto brez igrifikacije. Po oralni higieni otrok zaužije tableto za razkrivanje zobnih oblog in jo žveči približno 30 sekund ter nato ostanek in slino izpljune. V drugi fazi starši fotografirajo otrokovo zobovje in fotografije preko storitve za

pošiljanje datotek pošljejo medicinski sestri. Slednja s pomočjo standardiziranih indeksov oceni oralno higieno in staršem posreduje povratne informacije o izboljšanju le-te. V tretji fazi se vključi igrifikacija z nastavkom za zobno ščetko. Na pametni telefon se predhodno naložili mobilna aplikacija in se ustvaril osebni profil. Po dveh tednih oralne higijene z igrifikacijo se ponovita prvi dve fazi protokola.

Elementi igre v obliki značk, pokalov in lestvic predstavljajo potencialno rešitev za izboljšanje oralne higijene otroka, povečanje motivacije za oralno higieno, utrjevanje higienske navade in vpliv na avtomatizem pri oralni higieni. Tekom raziskave se srečujemo tudi z različnimi ovirami, ki nastanejo pri uporabi igrifikacije (npr. težave pri povezavi dodatka s pametnim telefonom, angleška verzija mobilne aplikacije itd.) ali pri oceni fotografije (kakovost fotografij, narejenih s pametnim telefonom, problem zajeti celotno zobovje s fotografijami itd.). Kljub temu lahko medicinska sestra na daljavo oceni učinkovitosti oralne higijene in posreduje takojšnje informacije o kvaliteti oralne higijene.

Leona Cilar, Leon Kopitar, Nino Fijačko, Primož Kocbek, Petra Povalej-Bržan, Majda Pajnikihar, Gregor Štiglic: Spletna aplikacija za podporo preventivni dejavnosti na področju sladkorne bolezni tipa 2 v referenčnih ambulantah

Sladkorna bolezen tipa 2 (SB2) je najpogostejši tip sladkorne bolezni, ki se pogosto pojavi postopoma in brez izrazitih simptomov. Raziskave kažejo, da je pojavnost SB2 v zadnjih letih v porastu. Zato je bistvenega pomena zgodnje odkrivanje bolezni in takojšnje ukrepanje. Podatke za 11.391 pacientov smo pridobili s pomočjo elektronskih zdravstvenih zapisov v petih slovenskih zdravstvenih ustanovah. Za izdelavo napovednih modelov smo zbrali podatke o plazemski glukozi in podatke, ki se nanašajo na vprašalnik Finnish Diabetes Risk Score (FINDRISC) za 2.073 posameznikov. Zbrani podatki so nam omogočili razvoj napovednih modelov za oceno tveganja prisotnosti preddiabetesa in nediagnosticirane SB2 pri zdravi populaciji. Napovedna modela smo uporabili v okviru razvoja spletne aplikacije za podporo pri odločanju v preventivni dejavnosti na področju SB2, ki se izvaja v okviru referenčnih ambulant. Aplikacija diplomirani medicinski sestri oz. diplomiranemu zdravstveniku omogoča učinkovit prikaz tveganja in oceno posrednega vpliva spremembe življenjskega sloga na tveganje za prisotnost preddiabetesa ali SB2. Študija prikazuje praktično uporabo dveh enostavnih napovednih modelov za zgodnje odkrivanje SB2 v

referenčni ambulantni. Preverjanje veljavnosti obeh modelov kaže na veliko povečanje napovedne učinkovitosti v primerjavi z obstoječim vprašalnikom za presejanje na področju SB2.

Uroš Rajkovič: Možnosti uporabe verig podatkovnih blokov v zdravstvu

Tehnologija veriženja podatkovnih blokov je nastala skupaj z razvojem kriptovalute Bitcoin, ki je neločljivo povezana z omenjeno tehnologijo. V tem oziru je potrebno proučevati in razumeti razloge za razvoj tovrstne tehnologije skupaj z njenimi prednostmi in omejitvami. Čeprav se tehnologija širi tudi na druga področja, je realizirana najpogosteje v finančnih okoljih. Zaradi njenih prednosti se pojavlja vprašanje, ali jo lahko uporabimo za zagotavljanje varnosti zapisov v zdravstvu. Na primerih izbranih zdravstvenih sistemov prispevek prikazuje omejitve, ki ne omogočajo enostavne preslikave tehnologije v novo okolje. Tehnologijo lahko uporabimo v zdravstvu, vendar to ni edina tehnologija, ki zagotavlja integriteto podatkov. Poleg tega je njena uporaba vezana na visoko procesorsko moč, ki ni poceni. Novo tehnologijo zato velja razumeti kot spodbudo, da ponovno razmišljamo o varnosti podatkov v zdravstvu.

Anja Brčan, Katja Krt, Maša Pintar: Model za pomoč patronažnim medicinskim sestram pri kategorizaciji rane

Rana je širok pojem, ki opredeljuje prekinjeno kontinuiteto kože in globljih struktur. Rane so neizogibni del v našem življenju. Nastanejo zaradi različnih poškodb, nesreč, obolenj in po operativnih posegih. Oskrba akutnih in kroničnih ran zahteva veliko strokovnega znanja in izkušenj. Šele pravilna ocena in oskrba rane omogočata, da rano privedemo do zgodnejše zacelitve brez zapletov in posledic. Zavedati se moramo, da lahko z nepravilno oskrbo rane njeno stanje celo poslabšamo.

Oblikovali smo odločitveni model, ki razvrsti rano glede na njene splošne in specifične značilnosti. Z odločitvijo želimo medicinskim sestram pomagati, da bodo pri svojem delu hitreje in bolj natančno opredelile, za kakšno vrsto rane gre, posledično pa jo bodo znale tudi primerno oskrbeti. Trenutno zdravljenje ran temelji predvsem na znanju in izkušnjah medicinske sestre. Prizadevamo si, da bi v oskrbi ran veljale poenotene smernice, ki bi se jih držale vse medicinske sestre in s tem omogočili poenoteno oskrbo.

Z uporabo računalniškega programa Dexi smo dosegli kategorizacijo rane od najzahtevnejše za

oskrbo proti manj zahtevni. Določili smo kriterije, funkcije koristnosti in izbrali štiri variante, ki jih je možno umestiti v kategorije. S pomočjo odločitvenega modela za razvrstitev rane bo izbrana primerna odločitev za oskrbo posamezne rane.

Alenka Žibert, Bernarda Lončar, Saša Šajn Lekše: Model za oceno vedenja nosečnice v podporo zdravju

Tvegana vedenja, ki lahko vplivajo na zdravje nosečnice in prihajajočega otroka, moramo pravočasno prepoznati, da lahko načrtujemo intervencije, ki bodo najbolj učinkovite za preprečevanje negativnih zdravstvenih izidov. Predmet odločanja je stopnja tveganega vedenja nosečnice. Problem je kompleksen, saj zadeva tako bodočo mater, kot tudi še nerojenega otroka in je velik izziv za primarno raven zdravstvenega varstva. S pomočjo odločitvenega modela želimo ločevati med nosečnicami z nizko, srednjo in visoko stopnjo tveganja z namenom prilagajanja vsebin individualne zdravstvene vzgoje in drugih intervencij zdravstvene in babiške nege v predporodni oskrbi. Za oceno tveganja smo uporabili večparametrski model za odločanje, ki je izdelan s pomočjo programa DEXi.

S pomočjo večparametrskega odločitvenega modela DEXi smo parametre, ki so pomembni za spremljanje obnašanja nosečnice v podporo zdravju, razvrstili v ustrezno drevesno strukturo. Za oceno stopnje tveganega vedenja nosečnice smo uporabili več kriterijev, ki smo jih strukturirali po pomembnosti (od spodaj navzgor). Glavni kriteriji v odločitvenem modelu se nanašajo na vitalne znake in način življenja nosečnice.

Izbrali smo si štiri variante, označene kot S, C, A in B, ki v našem primeru predstavljajo različne nosečnice. Vrednosti variant so predvidoma izmerjene ob kontrolnem obisku nosečnice v ginekološki ambulanti.

Varianta A nima nobene slabosti, a je odkritih nekaj pomanjkljivosti v znanju, ki pa nimajo bistvenega pomena na oceno pri skrbi za zdravje in potek nosečnosti. Sledi ji varianta S. Varianti C in B predstavljata velik problem. Varianta A je v vseh opisanih kriterijih nedvomno boljša od vseh ostalih variant in je skladna s pričakovanji. Odstopa le v primerjavi z varianto B, ki ima boljše ocene pri intenzivnosti in pogostosti telesne aktivnosti. Varianta B je po našem mnenju najmanj ustrežna, čeprav tako B kot C predstavljata nosečnico z velikim primanjkljajem v znanju in veščinah.

Obravnavna pacientov naj bi slonela na partnerskem odnosu z medicinsko sestro in na holističnem pristopu do pacienta. Pozornost velja nameniti vsem nosečnicam, saj jih moramo obravnavati individualno. Po pregledanemu odločitvenem modelu ugotovljamo, da je posebno pozornost potrebno posvetiti vsem naštetim kriterijem, saj vsak na svoj način vplivajo na končni rezultat. Odločitveni model lahko pomaga pri odkrivanju nosečnic, ki potrebujejo dodatna navodila in pomoč zdravstvenih delavcev.

Vladislav Rajkovič: Pomanjkanje kompetenc za digitalizacijo zdravstva

V tem prispevku razumemo kompetence kot zmožnost uspešne uporabe znanja in veščin za obvladovanje izzivov digitalizacije zdravstva. Zastavlja se vprašanje ustreznih kompetenc zdravstvenih delavcev v procesih obravnave pacientov. Se te nanašajo na računalništvo in informatiko? Morda celo na umetno inteligenco in podatkovno analitiko? Ta znanja so lahko pomemben sestavni del kompetenc, tako v procesu zdravljenja kot tudi procesu zdravstvene nege. Niso pa ta znanja edina, pa tudi ne najpomembnejša. Digitalizacija prinaša pomembne spremembe v procese stroke same. Digitalizirati le obstoječe procese je podobno, kot bi transportni sistem posodabljali le tako, da bi konja zamenjali z motorjem z notranjim izgorevanjem. Zato bomo razgrnili splet strokovnih in tehničnih kompetenc za spremembo procesov v zdravstvu zdravnikov, medicinskih sester in drugih zdravstvenih delavcev. Pomembno je razmišljanje zunaj ustaljenih okvirjev.

Maruša Kebe, Lucija Roblek, Brigita Jeretina: Prepoznavanje fizične in spolne zlorabe predšolskega otroka

Prepoznavna slabega ravnanja s predšolskim otrokom je lahko očitna pri vidnih sledih telesnega kaznovanja in spolnega nadlegovanja, lahko pa je zelo negotova. Razlaga, ki jo poda starš ali skrbnik, je pogosto netočna, poleg tega je otrok lahko premlad ali zaradi posledic slabega ravnanja z njim preveč prizadet, da bi lahko opisal dogodek. Za predšolske otroke ni nujno, da vsi obiskujejo vrtec ali druge oblike varstva, zato so zdravstveni delavci tisti, ki tekom obravnave otroka v zdravstvenem sistemu lahko prepoznajo sumljive poškodbe. Poleg tega zdravstvene delavce zavezuje zakon in etika, da si ob nasilju nad otroci ne zatiskajo oči, temveč da aktivno sodelujejo v prid žrtvi. Njihova osnovna naloga je, da zaščitijo žrtev. Raziskave kažejo, da se zdravstveni delavci odločajo za prijavo zlorabe otroka v večji meri, če imajo na voljo določena merila ali kriterije, po katerih naj bi se sum na zlorabo utemeljil. V ta namen smo s pomočjo metode DEX

in pripadajočega računalniškega programa Dexi razvili večkriterijski model, ki je zdravstvenim strokovnjakom v pomoč pri prepoznavi ogroženega predšolskega otroka za kakršnokoli obliko fizičnega ali spolnega nasilja. Obenem je pravočasno prepoznavanje nasilja tudi cilj odločitvenega procesa in zdravstvene obravnave otrok. Uporaba modela je namenjena predvsem ambulantam predšolske medicine. Medicinski sestri, ki se sreča s potencialno žrtvijo zlorabe, omogoča, da se opre na določene točke, ki ji omogočajo prepoznavo morebitne fizične ali spolne zlorabe. Pri tem je pomembno izpostaviti, da ni vsak fizični znak poškodbe tudi znak zlorabe. Pozorni pa moramo biti tudi na starše oz. skrbnike, ki spremljajo otroka. S tem modelom medicinska sestra lažje pride do zaključkov, ali je otrok zlorabljen ali ne, in temu primerno ukrepa.

Dragana Pejnović, Tanja Ritonja: Model za pomoč pri odločanju ob nakupu reanimacijskega vozička

Reanimacijski voziček je namenski voziček, ki ga uporabljamo pri oskrbi življenjsko ogroženih pacientov. Opremljen je z zdravili in pripomočki za oživljanje. Vsak bolnišnični oddelek mora imeti reanimacijski voziček z določeno opremo, ki si jo lahko prilagodijo glede na potrebe oddelka. Vsebinsko reanimacijskega vozička morajo poznati tako medicinske sestre kot zdravniki. Stalno pripravljeno zdravil in pripomočkov za oživljanje zagotovimo s strokovno in enotno urejenostjo reanimacijskega vozička. Medicinska sestra z vsakodnevno kontrolo in podpisom v evidenčni list prevzame odgovornost za njegovo brezhibnost. Oblikovali smo odločitveni model, ki razvrsti reanimacijski voziček glede na njegove komercialne pogoje, tehnične značilnosti in opremo. Z odločitvenim modelom želimo pomagati pri izbiri ob nakupu reanimacijskega vozička. Z dobro ocenjenim reanimacijskim vozičkom omogočimo kakovostno zdravstveno oskrbo življenjsko ogroženih pacientov. Za pomoč pri odločanju smo uporabili računalniško metodo DEX in pripadajoči program DEXi, s katerim smo ocenili reanimacijske vozičke od zelo ustreznega do neustreznega. Določili smo kriterije, funkcije koristnosti in ocenili štiri variante. Odločitveni model pripomore k izboru oziroma nakupu najboljše variante reanimacijskega vozička.

**Simona Kavčič, Anamarija Šalehar,
Gordana Lokajner: Informacijska podpora
zdravstvene nege in oskrbe na Onkološkem
inštitutu Ljubljana – na poti do kakovostne
in varne celostne obravnave onkoloških
pacientov**

Čas, v katerem živimo, in zdravstveni sistem, v katerem delamo, predstavljata zdravstvenim delavcem veliko izzivov. Sodobna tehnologija ni prisotna samo v našem osebnem življenju, temveč zelo hitro vstopa ter nas intenzivno spremlja tudi na profesionalnem področju.

V raziskavi smo uporabili tematsko vsebinsko analizo besedila, ki je temeljila na opisni raziskovalni metodi, metodi pregleda in primerjanju domače in tuje strokovne in znanstvene literature, iskane v mednarodnih bazah.

Na Onkološkem Inštitutu Ljubljana smo se odločili, da s podporo nove informacijske tehnologije dvignemo kakovost in varnost zdravstvene nege in oskrbe. Prispevek predstavlja teoretično podlago in primere dobre prakse iz kliničnih okolij pri uvajanju novega informacijskega sistema.

Predstavljamo prednosti negovalnega modula ter načrte za izboljšanje informacijskih procesov, ki bodo

pripomogli h kakovostnejši in celostni obravnavi onkoloških pacientov.

Zaključek

V zaključku srečanja je potekal sestanek članov SIZN, kjer je predsednica sekcije dr. Ema Dornik podala poročilo o delu SIZN za leto 2017, prav tako so bile opredeljene načrti in usmeritve za nadaljnje delo Sekcije za informatiko v zdravstveni negi. Zaključki letošnjega srečanja so:

- izkazalo, da je SIZN prava dodana vrednost zdravstveni negi in SDMI;
- potrebno je evidentiranje delovnih mest za potrebe e-ZN v digitalnem zdravstvu;
- v SIZN se vključuje vse več strokovnjakov, ki bodo tudi v prihodnje pripomogli k uresničitvi programa SIZN;
- s pomočjo družabnih omrežij bomo poskrbeli za boljšo prepoznavnost SIZN SDMI.

Zahvala

Zahvaljujemo se SDMI, ki je omogočilo naše srečanje. Hvala tudi članom SIZN, ki sodelujejo v naših aktivnostih, ter avtorjem, ki so pripravili povzetke predstavitev.

- **Infor Med Slov** 2017; 22(1-2): 33-37