

Oznaka poročila: ARRS-RPROG-ZP-2015/22



## ZAKLJUČNO POROČILO O REZULTATIH RAZISKOVALNEGA PROGRAMA

(za obdobje 1. 1. 2009 - 31. 12. 2014)

### A. PODATKI O RAZISKOVALNEM PROGRAMU

#### 1. Osnovni podatki o raziskovalnem programu

<b>Šifra programa</b>	P3-0154
<b>Naslov programa</b>	Metodologija za analizo podatkov v medicini Methodology for data analysis in medical sciences
<b>Vodja programa</b>	8992 Janez Stare
<b>Obseg raziskovalnih ur (vključno s povečanjem financiranja v letu 2014)</b>	21763
<b>Cenovni razred</b>	
<b>Trajanje programa</b>	01.2009 - 12.2014
<b>Izvajalci raziskovalnega programa (javne raziskovalne organizacije - JRO in/ali RO s koncesijo)</b>	381 Univerza v Ljubljani, Medicinska fakulteta
<b>Raziskovalno področje po šifrantu ARRS</b>	3 MEDICINA 3.08 Javno zdravstvo (varstvo pri delu)
<b>Družbeno-ekonomski cilj</b>	07. Zdravje
<b>Raziskovalno področje po šifrantu FOS</b>	3 Medicinske vede 3.03 Zdravstvene vede

### B. REZULTATI IN DOSEŽKI RAZISKOVALNEGA PROGRAMA

#### 2. Povzetek raziskovalnega programa<sup>1</sup>

SLO

Predmet raziskovalnega programa so metode za odkrivanje znanja ali generiranje novih hipotez, ki do znanja lahko pripeljejo. Od tega smo daleč največ ponovno naredili na področju biostatistike, kjer sma zlasti na področju analize preživetja in

analize večrazsežnih podatkov objavili več pomembnih člankov v najbolj uglednih svetovnih revijah.

Tako smo v reviji Biometrics objavili novo mero pojasnjene variabilnosti, ki ima vse potrebne lastnosti ter članek v reviji Applied Statistics, ki v osnovi spreminja pogled na obnašanje takšnih mer ob krnjenih podatkih.

Na področju relativnega preživetja smo objavili novo, predvsem pa prvo sploh pravilno, metodo za ocenjevanje čistega preživetja. Članek je postal najbolj citiran članek revije Biometrics v letih 2012 in 2013. Poleg tega smo objavili novo metodo ocenjevanja aditivnega modela, ki je bila objavljena v eni najuglednejših revij Biostatistics.

Na področju analize visoko razsežnih podatkov smo se ukvarjali z uvrščanjem, s posebnim poudarkom na uvrščanju neuravnoteženih podatkov. Naši rezultati so pokazali, da ima prisotnost velikega števila spremenljivk dodatno negativen vpliv na napovedno natančnost za manjši razred in da so rezultati uvrščanja, ki jih dobimo z naivno uporabo klasifikatorjev za neuravnotežene podatke, pristranski. Identificirali smo dejavnike, ki povzročajo to pristranost in preučili učinkovitost različnih pristopov za popravilo uvrščevalnih metod za neuravnotežene podatke. Predlagali smo tudi novo metodo za uvrščanje, ki je manj občutljiva na neravnotežje. Rezultati so bili objavljeni v člankih v reviji BMC Bioinformatics.

Na področju odkrivanja zakonitosti z analizo biomedicinske strokovne literature (literature-based discovery ali na kratko LBD) je del raziskav potekal na razvoju nove metodologije in orodij za integracijo rezultatov poskusov z mikromrežami (microarray) s semantičnimi relacijami iz strokovne literature. Razvili smo metodologijo in računalniška orodja, ki omogočajo pomoč pri interpretaciji mikromrež ter pri generiranju novih hipotez. Izpopolnili smo metodologijo za razdvoumljanje simbolov genov ter začeli razviti novo metodo za povzemanje informacij in ekstrakcijo znanja iz strokovne literature, ki smo jo poimenovali "semantična asociacijska pravila".

Dalje smo se lotili iskanja novih načinov zdravljenja neovaskularizacije pri diabetični retinopatiji. Pri tem smo uporabili tako orodje SemBT, ki smo ga razvili na našem inštitutu, kot tudi podatke raznih relevantnih omičnih raziskav. Z uporabo tega pristopa smo uspeli ne samo neodvisno odkriti nekatera zdravila, ki se trenutno uporabljajo pri kliničnih študijah, temveč tudi postaviti nekaj hipotez za nove načine zdravljenja.

ANG

The subject of this research programme is the methodology for discovery of existing or potentially new knowledge, trends, and relations in medical data. We are most productive in the area of biostatistics in which we published several important papers in top biostatistical journals.

In a paper in the journal Biometrics we introduced a new measure of explained variation for survival models which has all required statistical properties. This was later followed with a paper in Applied Statistics where we introduced a correct method of estimation of a measure of information gain.

In the area of relative survival, where we published a few important papers (one in Biostatistics), but by far the most important was published in Biometrics, introducing a new, and the only correct, method for estimating net survival.

The other field of research in survival analysis is the usage of pseudo values for the analysis of censored data. Our research was published in the journal Statistical Methods in Medical Research.

In the area of high-dimensional data analysis we worked on classification, with a special focus on class-imbalanced data. Our results showed that the inclusion of a large number of variables additionally weakens the predictive power, and that the classification

results, based on naive usage of classifiers, are biased. We identified the factors causing this bias and studied different procedures to correct it. We also proposed a new classification method that is less sensible to class-imbalance. Our results were presented in papers published in BMC Bioinformatics.

In the field of literature-based discovery (LBD) we addressed the problem of the integration of results from microarray experiments with the semantic relationships from literature. We developed a methodology and a tool that make easier the interpretation of the results obtained with microarray experiments and that can be used to generate new hypotheses. We improved the methodology for the disambiguation (razdvoumljanje) of gene symbols and started developing a new method for the synthesis of information and the extraction of knowledge from literature, which we named "semantic association rules".

We also investigated new methods for the treatment for neovascularization in retinopathy for diabetic patients. We used the SemBT tool, which was developed in our Institute, and data from relevant -omics studies. Using this approach we were able to discover independently some drugs, which are currently under investigation in clinical studies, and to propose new hypotheses for new treatment approaches.

### **3. Poročilo o realizaciji predloženega programa dela na raziskovalnem programu, (vključno s predloženim dopolnjenim programom dela v primeru povečanja financiranja raziskovalnega programa v letu 2014)<sup>2</sup>**

SLO

Člani naše programske skupine so v letih 2009-2014 objavili 203 izvirne znanstvene članke, od tega 153 v revijah, ki so indeksirane v SCI ali SSCI.

Podrobnejša vsebinska poročila smo pisali vsako leto posebej, tukaj jih na kratko povzemam:

Biostatistika:

Analiza preživetja

Daleč najpomembnejša dosežka na tem področju sta članka v reviji Biometrics (2011 in 2012), ki uvajata novo mero za pojasnjeno variabilnost v analizi preživetja in novo metodo za ocenjevanje čistega preživetja. Revija Biometrics je članek o pojasnjeni variabilnosti izbrala za vzorčno predavanje na njeni konferenci maja 2011 na Kreti, drugi članek pa je postal najbolj citiran članek te revije med članki, objavljenimi v letih 2012 in 2013. Obema so sledila vabila na več vabljenih predavanj na mednarodnih konferencah in delavnicah ter na več tujih univerzah.

Poleg tega smo na področju relativnega preživetja objavili novo metodo ocenjevanja koeficientov v aditivnem modelu. Metoda se izogne vsakršnim predpostavkam o bazičnem tveganju in tako omogoča zanesljivejše ocenjevanje v mnogih praktičnih situacijah, ko predpostavke obstoječih metod niso izpolnjene oz. jih je nemogoče preverjati. Pristop temelji na EM algoritmu in uporabniku omogoča, da pridobi dodatno informacijo o podatkih ter s tem obogati interpretacijo. Metoda je bila objavljena v ugledni statistični reviji Biostatistics.

Analiza večrazsežnih podatkov

Tu smo se osredotočili na uvrščanje visoko-razsežnih podatkov, s posebnim poudarkom na uvrščanju neuravnoteženih podatkov, kjer je število enot med razredi zelo različno. Problem je že dokaj dobro poznan za uvrščanje neuravnoteženih podatkov, kjer je število spremenljivk majhno, za visoko-razsežne podatke pa je novost. Zato smo preučili obnašanje najbolj pogosto uporabljenih klasifikatorjev na visoko-dimenzionalnih neuravnoteženih podatkih s področja biomedicine in predlagali rešitve, s katerimi bi lahko zmanjšali problem neravnotežja. To je bila tudi tema doktorske naloge Roka Blagusa, ki je bil mladi raziskovalec na IBMI in je doktoriral leta 2011. Glavni metodološki rezultati so bili predstavljeni v treh člankih, ki so bili objavljeni v reviji BMC Bioinformatics. Rezultati, ki so nastali iz sodelovanja na visokorazsežnih podatkih z Onkološkim inštitutom iz Milana, pa so bili objavljeni v reviji PLOS One.

#### Kakovost v zdravstvu

Dolgoletno strokovno delo na področju statističnega nadzora kakovosti v zdravstvu smo nadgradili z izvirnim pristopom k odkrivanju odstopajočih podatkov o deležih (dvojno-kvadratnokorenski grafikon z mejami nadzora na podlagi regresije skozi izhodišče). Članek, ki je pokazal prednosti našega pristopa v primerjavi z uveljavljenimi splošnimi statističnimi metodami prepoznavanja odstopajočih podatkov, je bil objavljen v reviji Applied Stochastic Models in Business and Industry.

#### Bibliometrija

Tudi na tem področju smo objavili 4 pomembne metodološke članke, dva v reviji Journal of Informetrics, ki je na svojem področju 2. od 84 revij, kar po merilih ARRS šteje za izjemen dosežek. Druga dva članka sta bila objavljena v reviji Scientometrics, prav tako ugledni reviji s tega področja.

Poleg tega smo s partnerji iz Slovenije, Belgije ter Velike Britanije analizirali porabo denarja za medicinske raziskave v Evropski Uniji (EU). Razvili smo računalniške programe, ki avtomatično, za tisoče projektov finansiranih s strani EU, preveri raziskovalne rezultate teh projektov v zbirkah Google Scholar (GS) ter v PubMed (PM). Rezultati analize so pokazali, da v poprečju vsaka raziskovalna publikacija stane davkoplačevalce s področja EU okoli €140,000 (GS) do €225,000 (PM).

#### Odkrivanje znanja z analizo strokovne literature

Na tem področju smo nadaljevali z razvojem lastne programske opreme (na primer, izboljšali smo metodologijo z vpeljavo pristopa "pregledovanja odkritij", ki temelji na teoriji grafov z uporabo semantičnih relacij iz biomedicinske literature), vse več pa programska orodja uporabljamo za praktične raziskave. Tako smo kot del mednarodne zasedbe iskali razlago za zmanjšano kvaliteto pri starajočih se moških ter postavili novo hipotezo da je omenjena zmanjšana kvaliteta spanja posledica zmanjšanega nivoja testosterona, ki inhibira kortisol. Rezultate raziskave smo objavili v reviji Sleep. Začeli smo se ukvarjati tudi z iskanjem razlage za stranske učinke zdravil, in sicer tako, da avtomatično iščemo gene ali proteine, ki vzročno povezujejo zdravila s stranskimi učinki. Drugače povedano, iščemo farmakogenomsko razlago za stranske učinke zdravil.

#### **4. Ocena stopnje realizacije programa dela na raziskovalnem programu in zastavljenih raziskovalnih ciljev<sup>3</sup>**

*SLO*

Običajno je že, da na področju statistike (preživetje in analiza večrazsežnih podatkov) cilje dosegamo in presehamo. Ocenjujem, da smo za tako majhno ekipo izjemno uspešni, zlasti glede odkrivanja novih pristopov, ki jih tudi objavljamo v vrhunskih revijah. Poleg tega dosegajo naši članki v statistični literaturi (sedaj tudi že v drugih medicinskih revijah) izjemno odmevnost.

Zelo veseli smo rezultatov na področju bibliometrije oz. scientometrije, kjer smo s štirimi objavami v vrhunskih revijah presegli pričakovanja.

Pa tudi z raziskovanji na področju podatkovnega rudarjenja v bibliografskih zbirkah smo zelo zadovoljni, saj imamo med drugim članek kategorije 1A1 in pa kratek znanstveni prispevek v eni najuglednejših revij Lancet, tudi kategoriziran kot 1A1.

Ukvarjanje s kakovostjo v zdravstvu se je prikradlo v našo programsko skupino skozi zanimanje za statistične metode v zdravstvu nasploh, tudi tukaj pa imamo že originalen metodološki prispevek, objavljen v ugledni reviji.

Ob tem je izšlo veliko člankov, kjer so naši raziskovalci soavtorji člankov o raziskavah v medicini. S tem naša skupina izpolnjuje svoje temeljno poslanstvo, da so namreč njeni napori

usmerjeni v podporo raziskavam v slovenski medicini.

Če povzamem zgornji povzetek, lahko rečem, da smo po našem mnenju program v celoti realizirali, na nekaterih področju vsekakor nad pričakovanji.

#### 5. Utemeljitev morebitnih sprememb programa raziskovalnega programa oziroma sprememb, povečanja ali zmanjšanja sestave programske skupine v letu 2014<sup>4</sup>

SLO

Ni bilo sprememb.

#### 6. Najpomembnejši znanstveni rezultati programske skupine<sup>5</sup>

		Znanstveni dosežek	
1.	COBISS ID	28569561	Vir: COBISS.SI
	Naslov	SLO	O ocenjevanju relativnega preživetja
		ANG	On estimation in relative survival
	Opis	SLO	Osnovni cilj analize preživetja bolnikov z rakom je izračun krivulje čistega preživetja, torej krivulje, ki jo določa le smrtnost zaradi raka, ne pa tudi drugih možnih vzrokov smrti. Izkazalo se je, da metode, ki so v ta namen rutinsko uporabljane, dajejo pristranske rezultate, ki so odvisni od splošne umrljivosti in zato neprimerljivi med državami, interpretacija dobljenih krivulj pa je nejasna in zavajajoča. Članek teoretično primerja obstoječe metode, razloži kaj posamezna metoda ocenjuje in kako jo je potrebno interpretirati. Članek tudi predlaga nov pristop, ki omogoča nepristransko oceno čistega preživetja in s tem zeleno primerjavo med državami.
		ANG	Estimation of relative survival has become the first and the most basic step when reporting cancer survival statistics. Standard estimators are in routine use by all cancer registries. However, it has been recently noted that these estimators do not provide information on cancer mortality that is independent of the national general population mortality. Thus they are not suitable for comparison between countries. Furthermore, the commonly used interpretation of the relative survival curve is vague and misleading. The present article attempts to remedy these basic problems. The population quantities of the traditional estimators are carefully described and their interpretation discussed. We then propose a new estimator of net survival probability that enables the desired comparability between countries. The new estimator requires no modeling and is accompanied with a straightforward variance estimate. The methods are described on real as well as simulated data.
	Objavljeno v	Biometric Society [etc.]; Biometrics; 2012; Vol. 68, št. 1; str. 113-120; Impact Factor: 1.412; Srednja vrednost revije / Medium Category Impact Factor: 1.089; A': 1; WoS: CU, MC, XY; Avtorji / Authors: Pohar Perme Maja, Stare Janez, Esteve Jacques	
	Tipologija	1.01 Izvirni znanstveni članek	
COBISS ID	27655129	Vir: COBISS.SI	
Naslov	SLO	Mera pojasnjene variabilnosti v analizi zgodovine dogodkov	
	ANG	A measure of explained variation for event history data	
Opis	SLO	V članku je definirana nova mera pojasnjene variabilnosti v analizi zgodovine dogodkov (event history analysis). Mera ima ustrezne statistične lastnosti in se jo lahko uporabi za vse vrste modelov in podatkov (npr. časovna odvisnost, ponavljajoči dogodki ...)	
		There is no shortage of proposed measures of prognostic value of survival models in the statistical literature. They come under different names,	

		ANG	including explained variation, correlation, explained randomness, and information gain, but their goal is common: to define something analogous to the coefficient of determination R <sup>2</sup> in linear regression. None however have been uniformly accepted, none have been extended to general event history data, including recurrent events, and many cannot incorporate time-varying effects or covariates. We present here a measure specifically tailored for use with general dynamic event history regression models. The measure is applicable and interpretable in discrete or continuous time; with tied data or otherwise; with time-varying, time-fixed, or dynamic covariates; with time-varying or time-constant effects; with single or multiple event times; with parametric or semiparametric models; and under general independent censoring/observation. For single-event survival data with neither censoring nor time dependency it reduces to the concordance index. We give expressions for its population value and the variance of the estimator and explore its use in simulations and applications. A web link to R software is provided.
	Objavljeno v		Biometric Society [etc.]; Biometrics; 2011; Vol. 67, issue 3; str. 750-759; Impact Factor: 1.827; Srednja vrednost revije / Medium Category Impact Factor: 1.046; A': 1; WoS: CU, MC, XY; Avtorji / Authors: Stare Janez, Pohar Perme Maja, Henderson Robin
	Tipologija		1.01 Izvirni znanstveni članek
3.	COBISS ID	27503321	Vir: COBISS.SI
	Naslov	SLO	Napovedovanje razreda za visoko razsežne, po razredih neuravnotežene, podatke
		ANG	Class prediction for high-dimensional class-imbalanced data
	Opis	SLO	Standardne metode za razvrščanje podatkov, ki so neuravnoteženi, pogosto manj natančno napovedujejo pripadnost za manjši razred; napoved je pristrana k večjemu razredu. V tem članku ocenjujemo ali visoka razsežnost prinaša dodatne izzive, ko gre za razvrščanje neuravnoteženih podatkov. Identificirali smo dodatni vir pristranosti, ko analiziramo visoko-raksežne podatke; raziskovali smo tudi učinkovitost pristopov, ki so na voljo za zmanjšanje učinka neuravnoteženosti.
		ANG	The goal of class prediction studies is to develop rules to accurately predict the class membership of new samples. The rules are derived using the values of the variables available for each subject: the main characteristic of high-dimensional data is that the number of variables greatly exceeds the number of samples. Frequently the classifiers are developed using class-imbalanced data, i.e., data sets where the number of samples in each class is not equal. Standard classification methods used on class-imbalanced data often produce classifiers that do not accurately predict the minority class; the prediction is biased towards the majority class. In this paper we investigate if the high-dimensionality poses additional challenges when dealing with class-imbalanced prediction. We evaluate the performance of six types of classifiers on class-imbalanced data, using simulated data and a publicly available data set from a breast cancer gene-expression microarray study. We also investigate the effectiveness of some strategies that are available to overcome the effect of class imbalance. Results Our results show that the evaluated classifiers are highly sensitive to class imbalance and that variable selection introduces an additional bias towards classification into the majority class. Most new samples are assigned to the majority class from the training set, unless the difference between the classes is very large. As a consequence, the class-specific predictive accuracies differ considerably. When the class imbalance is not too severe, down-sizing and asymmetric bagging embedding variable selection work well, while over-sampling does not. Variable normalization can further worsen the performance of the classifiers. (Abstract truncated

			at 2000 characters)
	Objavljeno v		BioMed Central; BMC bioinformatics; 2010; Letn. 11; str. 523 (1-27); Impact Factor: 3.028; Srednja vrednost revije / Medium Category Impact Factor: 1.868; A': 1; WoS: CO, DB, MC; Avtorji / Authors: Blagus Rok, Lusa Lara
	Tipologija		1.01 Izvirni znanstveni članek
4.	COBISS ID	31282137	Vir: COBISS.SI
	Naslov	SLO	O standardizaciji indeksa aktivnosti
		ANG	On standardization of the activity index
	Opis	SLO	Relativni indeks specializiranosti (RSI) je definiran kot preprosta standardizacija indeksa aktivnosti (AI), ki naj bi olajšala interpretacijo AIja. RSI naj bi imel vrednosti med -1 in +1, ki pomenita ničelno aktivnost na določenem področju oz. aktivnost samo na danem področju. Čeprav RSI očitno ne more biti 1, pa je žal tudi njegova zgornja meja odvisna od področja in lahko bistveno manjša od 1. To sledi iz dejstva, da ima AI zelo različne zgornje meje za različna področja. Kar pa pomeni, da je primerjava teh indeksov med področji nesmiselna oz. zavajajoča. Zato menimo, da se RSI ne bi smelo uporabljati. Poleg tega predlagamo ustrezen način standardizacije AIja.
		ANG	Relative Specialization Index (RSI) was introduced as a simple transformation of the Activity Index (AI), the aim of this transformation being standardization of AI, and therefore more straightforward interpretation. RSI is believed to have values between -1 and 1, with -1 meaning no activity of the country (institution) in a certain scientific field, and 1 meaning that the country is only active in the given field. While it is obvious from the definition of RSI that it can never be 1, it is less obvious, and essentially unknown, that its upper limit can be quite far from 1, depending on the scientific field. This is a consequence of the fact that AI has different upper limits for different scientific fields. This means that comparisons of RSIs, or AIs, across fields can be misleading. We therefore believe that RSI should not be used at all. We also show how an appropriate standardization of AI can be achieved.
	Objavljeno v		Elsevier; Journal of informetrics; 2014; Vol. 8, iss. 3; str. 503-507; Impact Factor: 3.580; Srednja vrednost revije / Medium Category Impact Factor: 1.031; A'': 1; A': 1; WoS: NU; Avtorji / Authors: Stare Janez, Kejžar Nataša
	Tipologija		1.01 Izvirni znanstveni članek
5.	COBISS ID	30069465	Vir: COBISS.SI
	Naslov	SLO	Primerjava porazdelitev citatov in h indeksa med skupinami različnih velikosti
		ANG	Comparison of the citation distribution and h-index between groups of different sizes
	Opis	SLO	Na primerjave porazdelitev citiranosti močno vpliva število objavljenih člankov, kar je vezano na velikost primerjanih skupin. V članku uvedemo metodo za primerjanje skupin različnih velikosti. Porazdelitev citiranosti večje skupine je zmanjšana za ustrezen faktor in primerjana s porazdelitvijo manjše skupine. Podane so pričakovane vrednosti in tolerančni intervali za reducirano množico citatov. Obe porazdelitvi lahko tudi grafično primerjamo. Izračunan je tudi reduciran Hirschev h indeks. Metodo lahko uporabljamo skupine ali enote, ki se ločijo glede na velikost financiranja, v številu raziskovalcev, lahko primerjamo države ob upoštevanju velikosti populacije ali BDPja in podobno. Pokažemo, da je za računanje reduciranega h indeksa dovolj poznavanje zgornjega dela porazdelitve pri večji skupini. Metode ilustriramo z različnimi primeri.

	ANG	Evaluating the performance of institutions with different resources is not easy, any citation distribution comparisons are strongly affected by the differences in the number of articles published. The paper introduces a method for comparing citation distributions of research groups that differ in size. The citation distribution of a larger group is reduced by a certain factor and compared with the original distribution of a smaller group. Expected values and tolerance intervals of the reduced set of citations are calculated. A comparison of both distributions can be conveniently viewed in a graph. The size-independent reduced Hirsch index - a function of reducing factor that allows the comparison of groups within a scientific field - is calculated in the same way. The method can be used for comparing groups or units differing in full-time equivalent, funding or the number of researchers, for comparing countries by population, gross domestic product, etc. It is shown that for the calculation of the reduced Hirsch index, the upper part of the original citation distribution is sufficient. The method is illustrated through several case comparisons.
Objavljeno v		Elsevier; Journal of informetrics; 2012; Vol. 6, iss. 4; str. 712-720; Impact Factor: 4.153; Srednja vrednost revije / Medium Category Impact Factor: 0.989; A'': 1; A': 1; WoS: NU; Avtorji / Authors: Pohar Perme Maja, Stare Janez, Žaucer Rok, Žaucer Matjaž
Tipologija		1.01 Izvirni znanstveni članek

### 7. Najpomembnejši družbeno-ekonomski rezultati programske skupine<sup>6</sup>

	Družbeno-ekonomski dosežek	
1.	COBISS ID	268736768 Vir: COBISS.SI
	Naslov	SLO Izvlečki in program ANG Abstracts and program
	Opis	SLO Janez Stare je od leta 2004 naprej (vsako leto) predsednik programskega odbora mednarodne konference Applied Statistics (glej <a href="http://conferences.nib.si/AS2012/default.htm">http://conferences.nib.si/AS2012/default.htm</a> ). Skupaj z Lara Luso urejata zbornik konference. ANG Janez Stare has been chair of the Programme committee of the Applied Statistics (see <a href="http://conferences.nib.si/AS2012/default.htm">http://conferences.nib.si/AS2012/default.htm</a> ) since 2004. He and Lara Lusa they are editors of the conference proceedings.
	Šifra	B.02 Predsedovanje programskemu odboru konference
	Objavljeno v	Statistical Society of Slovenia; 2013; 83 str.; Avtorji / Authors: Lusa Lara, Stare Janez
	Tipologija	2.25 Druge monografije in druga zaključena dela
2.	COBISS ID	215795712 Vir: vpis v poročilo
	Naslov	SLO Metodološki zvezki ANG Advances in methodology and statistics
	Opis	SLO Lara Lusa je od leta 2013 sourednica revije Metodološki zvezki ANG Since 2013 Lara lusa is the co-editor of the journal Advances in methodology and statistics
	Šifra	C.04 Uredništvo mednarodne revije
	Objavljeno v	Metodološki zvezki = Advances in methodology and statistics
	Tipologija	4.00 Sekundarno avtorstvo
3.	COBISS ID	20359385 Vir: vpis v poročilo



	Naslov	SLO	Acta informatica medica
		ANG	Acta informatica medica
	Opis	SLO	Gaj Vidmar je urednik revije
		ANG	Gaj Vidmar is the editor of the journal.
	Šifra	C.05 Uredništvo nacionalne revije	
	Objavljeno v	Acta informatica medica	
	Tipologija	4.00 Sekundarno avtorstvo	
4.	COBISS ID	30371289	Vir: COBISS.SI
	Naslov	SLO	Akademski dosežki kot posledica 9-letnega investiranja EU v raziskave v zdravstvu
		ANG	Academic output of 9 years of EU investment into health research
	Opis	SLO	Evropska unija je skozi svoje okvirne programe glavni podpornik raziskovalnega dela v Evropi. Proračun se je od 19•1 milijard evrov v letih 2002–06 povečal 53 milijard v letih 2007–13. Pri takšnem obsegu financiranja je seveda pomembno vedeti, kako se denar porablja in kaj se producira. Mi smo uporabili informacijski pristop k evalviranju znanstvenih dosežkov, ki so bili posledica financiranja raziskovanja v zdravstvu v okvirnih programih 5 in 6. V ta namen smo razvili računalniški program, ki je preiskal bazo Google Scholar glede na kode posameznih financiranih projektov. Naši rezultati kažejo, da se na področju zdravja več kot polovica denarja porabi v projektih, ki ne dajo nobenega bibliografsko merljivega znanstvenega rezultata.
		ANG	The European Union's (EU's) Framework Programmes (FP) for Research and Technological Development are major funders of research in Europe. The European Commission clearly values its science programme: the budget rose from €19•1 billion under FP6 (2002–06) to €53 billion2 in FP7 (2007–13). With billions of taxpayers' Euros invested, it is desirable to know where the money is being spent and what is produced. We attempted to use an informatics approach to identify academic publications resulting from health-related research funded under FP5 and FP6. We therefore wrote a program to data-mine Google Scholar with EU grant codes, counting papers returned. Our findings show that more than half of the money of health-related research resulted in no detectable academic output.
	Šifra	F.23 Razvoj novih sistemskih, normativnih, programskih in metodoloških rešitev	
	Objavljeno v	Lancet Publishing Group; The Lancet; 2012; Letn. 380, št. 9846; str. 971-972; Impact Factor: 39.060; Srednja vrednost revije / Medium Category Impact Factor: 2.548; A": 1; A': 1; WoS: PY; Avtorji / Authors: Galsworthy Michael J., Hristovski Dimitar, Lusa Lara, Ernst Kelly	
	Tipologija	1.03 Kratki znanstveni prispevek	
5.	COBISS ID	32893696	Vir: vpis v poročilo
	Naslov	SLO	urednik
		ANG	editor
	Opis	SLO	Maja Pohar Perme je sourednica revije Zdravniški vestnik
		ANG	Maja Pohar Perme is coeditor of the journal Zdravniški vestnik
	Šifra	C.05 Uredništvo nacionalne revije	
	Objavljeno v	Zdravniški vestnik. Pohar, Maja (urednik 2010-). [Tiskana izd.]. Ljubljana: [Slovensko zdravniško društvo], 1992-. ISSN 1318-0347. [COBISS.SI-ID 32893696]	

Tipologija	4.00	Sekundarno avtorstvo
------------	------	----------------------

## 8. Drugi pomembni rezultati programske skupine<sup>Z</sup>

1. število čistih citatov znanstvenih del v zadnjih 10 letih (2005-2014) je 4512
2. najodmevnejše delo v zadnjih 10 letih (2005-2014) ime 807 citatov
3. imeli smo vrsto vabljenih predavanj na mednarodnih konferencah, mednarodnih delavnicah in univerzah, zapisi so v COBISSu.

## 9. Pomen raziskovalnih rezultatov programske skupine<sup>8</sup>

### 9.1. Pomen za razvoj znanosti<sup>2</sup>

SLO

Poleg splošnih ugotovitev, navedenih v programu (delo na področju pojasnjene variabilnosti, relativnega preživetja, analize večrazsežnih podatkov in programa za kreiranje hipotez) naj pomen naših raziskav za razvoj (biostatistične) znanosti, kot se je kazal v obdobju od 1.1.2010 do danes, ilustriram z vabljenimi predavanji nekaterih članov skupine:

Janez Stare - vabljena predavanja na

- strokovnem srečanju ožjega kroga strokovnjakov s področja napovedovanja (vsa predavanja so vabljena) 2010, Oberwolfach, Nemčija

[http://www.mfo.de/admin/workshops/1662/www\\_view](http://www.mfo.de/admin/workshops/1662/www_view)

- konferenci International Biometric Conference, Florianopolis, Brazilija

na univerzah

- Lyon 2010 (2 predavanji)

- London 2010 (2 predavanji)

- Milano 2010 (2 predavanji)

Lara Lusa

na univerzah

- Torino 2009

- Neapelj 2010

naj dodam, da smo v letu 2011 že imeli tri vabljena predavanja, imeli bomo vsaj še dve (na pomembnih mednarodnih konferencah), dobivamo že vabila za leto 2012 (Stockholm).

ANG

ANG

Besides general facts, mentioned in the program (our work in explained variation, relative survival, multidimensional data and programme for creation of hypothesis), the importance of our research for the development of biostatistical science is best illustrated by numerous invited lectures that we had on different, always international, occasions. In the last year, so from Jan 1 2010 these were:

Janez Stare - invited talks at

- workshop on individual prediction 2010 (all lectures were invited), Oberwolfach, Germany

[http://www.mfo.de/admin/workshops/1662/www\\_view](http://www.mfo.de/admin/workshops/1662/www_view)

- International Biometric Conference, Florianopolis, Brazilija

at universities

- Lyon 2010, two invited talks

- London 2010, two invited talks

- Milano 2010, two invited talks

Lara Lusa

at universities

- Torino 2010
- Naples 2010

This year (2011) we already had 3 invited lectures, at least two more at important international conferences (EMR-IBS on Crete nad ISCB in Ottawa), and we are already receiving invitations for year 2012.

## 9.2. Pomen za razvoj Slovenije<sup>10</sup>

SLO

Vse naše raziskave izhajajo iz osnovne usmeritve našega inštituta, to je nuditi kar najkvalitetnejšo podporo raziskovalnemu delu v slovenski medicini. Brez teh raziskav bi bila naša biostatistična in znanstveno informacijska podpora bistveno revnejša.

Naše delo nudi podporo raziskovalnemu in strokovnemu delu v medicini in preko tega tudi tehnološkemu razvoju v slovenski medicini.

Vsaka mednarodno odmevna dejavnost utrjuje nacionalno identiteto. Naše raziskave vsekakor odmevajo v krogu ljudi, ki ga pred nami iz Slovenije ni nihče naslavljal.

ANG

All our research is based on our goal to provide adequate support to research in medicine in Slovenia. And without our own research our support would not have been what it is. Through our support to research in medicine, we participate in its technological development. Our collaboration with researchers in medicine has in hundreds of papers over the years in journals with impact factors.

Any activity with international impact strengthens national identity. And our research certainly becomes known to people that no one in Slovenia has addressed before.

## 10. Zaključena mentorstva članov programske skupine pri vzgoji kadrov v obdobju 1.1.2009-31.12.2014<sup>11</sup>

### 10.1. Diplome<sup>12</sup>

vrsta usposabljanja	število diplom
bolonjski program - I. stopnja	
bolonjski program - II. stopnja	
univerzitetni (stari) program	

### 10.2. Magisterij znanosti in doktorat znanosti<sup>13</sup>

Šifra raziskovalca	Ime in priimek	Mag.	Dr.	MR	
		<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	

Legenda:

- Mag.** - Znanstveni magisterij
- Dr.** - Doktorat znanosti
- MR** - mladi raziskovalec

### 11. Pretok mladih raziskovalcev – zaposlitev po zaključenem usposabljanju<sup>14</sup>

Šifra raziskovalca	Ime in priimek	Mag.	Dr.	Zaposlitev	
		<input type="radio"/>	<input type="radio"/>	<input type="text"/>	

Legenda zaposlitev:

- A** - visokošolski in javni raziskovalni zavodi
- B** - gospodarstvo
- C** - javna uprava
- D** - družbene dejavnosti
- E** - tujina
- F** - drugo

### 12. Vključenost raziskovalcev iz podjetij in gostovanje raziskovalcev, podoktorandov ter študentov iz tujine, daljše od enega meseca, v obdobju 1.1.2009-31.12.2014

Šifra raziskovalca	Ime in priimek	Sodelovanje v programski skupini	Število mesecev	
		▼		

Legenda sodelovanja v programski skupini:

- A** - raziskovalec/strokovnjak iz podjetja
- B** - uveljavljeni raziskovalec iz tujine
- C** - študent - doktorand iz tujine
- D** - podoktorand iz tujine

### 13. Vključevanje v raziskovalne programe Evropske unije in v druge mednarodne raziskovalne in razvojne programe ter drugo mednarodno sodelovanje v obdobju 1.1.2009-31.12.2014<sup>15</sup>

SLO

EURETHNET. European Information Network Ethics in Medicine and Biotechnology. 5. Okvirni program EU. Trajanje: 2002 - 2005. IBMI vodi delovno skupino (Workpackage, WP) za gradnjo informacijskega sistema, namenjenega zbiranju bibliografskih zapisov v državah, novih pritopticah v EU.

ETHICSWEB. Inter-connected European information and documentation system for ethics and science: European ethics documentation centre. 7. Okvirni program EU. Coordination and Support Actions. Si7-CT-2008-217817. Trajanje: 2008 - 2011. IBMI vodil delovno skupino (WP), ki je razvila spletišče projekta, sistem za upravljanje s podatki, podatkovne zbirke, iskalnike projekta in nudila tudi siceršnjo informacijsko podporo.

ERANet ERASysBioPlus. Skupni razpis MIZŠ RS in EU. Projekt BioMod UE\_PTL - Biophysical Modelling of the Uterine Electromyogram for understanding and preventing Pre-term Labor. Trajanje: 2010 - 2013. IBMI vodil delovno skupino (WP), ki je vodila razvoj informacijskega sistema projekta in upravljanje s podatki, nastalimi v projektu.

ITECH (Roadmap for Research and Innovation in Health Technology). 7. Okvirni program EU FP7-HEALTH-2013-INNOVATION-1 (Proposal No: 602667-2). Trajanje: 2013 - 2016. IBMI sodeluje pri zbiranju podatkov na nacionalnem nivoju.

ELIXIR - A sustainable European infrastructure for biological information, supporting life science research and its translation to medicine, agriculture, bioindustries and society. Trajna prioriteta infrastruktura ESFRI. Infrastrukturo podpirata MIZŠ RS in EU. Nacionalno slovensko vozlišče vodi Medicinska fakulteta UL (IBK/CFGBC in IBMI).

ANG

EURETHNET. European Information Network Ethics in Medicine and Biotechnology. 5th EU Framework. Duration: 2002 - 2005. IBMI was a workpackage leader responsible for the development of an information system for collecting bibliographic records in New EU

## Member States.

ETHICSWEB. Inter-connected European information and documentation system for ethics and science: European ethics documentation centre. 7th EU Framework. Coordination and Support Actions. SiS7-CT-2008-217817. Duration: 2008 - 2011. IBMI was a workpackage leader responsible for the information support to project activities, development of the project's website and portal, project's content management system, databases, and search engines.

ERANet ERASysBioPlus. Joint call of Slovenian Ministry of Education, Science and Sport and 7th EU Framework. Project title: BioMod UE\_PTL - Biophysical Modelling of the Uterine Electromyogram for understanding and preventing Pre-term Labor. Duration: 2010 - 2013. IBMI was a workpackage leader responsible for the information support, development of the project's information system and data management.

ITECH (Roadmap for Research and Innovation in Health Technology). 7th EU Framework FP7-HEALTH-2013-INNOVATION-1 (Proposal No: 602667-2). Duration: 2013 - 2016. IBMI co-operates in data collection on the national level.

ELIXIR - A sustainable European infrastructure for biological information, supporting life science research and its translation to medicine, agriculture, bioindustries and society. A permanent ESFRI priority. The infrastructure is jointly supported by Slovenian Ministry of Education, Science and Sport and EU Horizon2020. The Slovenian National Elixir Node is led by Faculty of Medicine (IBK/CFGBC and IBMI).

**14. Vključenost v projekte za uporabnike, ki so v obdobju trajanja raziskovalnega programa (1.1.2009–31.12.2014) potekali izven financiranja ARRS<sup>16</sup>**

SLO

--

**15. Ocena tehnološke zrelosti rezultatov raziskovalnega programa in možnosti za njihovo implementacijo v praksi (točka ni namenjena raziskovalnim programom s področij humanističnih ved)<sup>17</sup>**

SLO

Kot že omenjeno zgoraj, naše raziskave koristijo raziskavam v medicini, nimajo pa tehnoloških ambicij.
--

**16. Ocenite, ali bi doseženi rezultati v okviru programa lahko vodili do ustanovitve spin-off podjetja, kolikšen finančni vložek bi zahteval ta korak ter kakšno infrastrukturo in opremo bi potrebovali**

možnost ustanovitve spin-off podjetja	<input type="radio"/> DA <input checked="" type="radio"/> NE
potrebni finančni vložek	EUR
ocena potrebne infrastrukture in opreme <sup>18</sup>	

**17. Izjemni dosežek v letu 2014<sup>19</sup>**

**17.1. Izjemni znanstveni dosežek**

Po merilih ARRS sta izjemna dosežka oba članka, objavljena v reviji Journal of Informetrics. Po
---

naših merilih sta precej večja dosežka članka, objavljena v reviji Biometrics, zlasti članek, naveden zgoraj pod zaporedno številko 1, ki je najbolj citiran članek te revije v zadnjih letih.

## 17.2. Izjemni družbeno-ekonomski dosežek

## C. IZJAVE

Podpisani izjavljam/o, da:

- so vsi podatki, ki jih navajamo v poročilu, resnični in točni;
- se strinjamo z obdelavo podatkov v skladu z zakonodajo o varstvu osebnih podatkov za potrebe ocenjevanja in obdelavo teh podatkov za evidence ARRS;
- so vsi podatki v obrazcu v elektronski obliki identični podatkom v obrazcu v papirnati obliki;
- so z vsebino poročila seznanjeni in se strinjajo vsi izvajalci raziskovalnega programa.

### Podpisi:

*zastopnik oz. pooblaščen oseba  
matične RO (JRO in/ali RO s  
koncesijo):*

in

*vodja raziskovalnega programa:*

Univerza v Ljubljani, Medicinska  
fakulteta

Janez Stare

### ŽIG

Kraj in datum:

Ljubljana

26.2.2015

### Oznaka poročila: ARRS-RPROG-ZP-2015/22

<sup>1</sup> Napišite povzetek raziskovalnega programa v slovenskem jeziku (največ 3.000 znakov vključno s presledki – približno pol strani, velikost pisave 11) in angleškem jeziku (največ 3.000 znakov vključno s presledki – približno pol strani, velikost pisave 11). [Nazaj](#)

<sup>2</sup> Napišite kratko vsebinsko poročilo, v katerem predstavite raziskovalno hipotezo in opis raziskovanja. Navedite ključne ugotovitve, znanstvena spoznanja, rezultate in učinke raziskovalnega programa in njihovo uporabo ter sodelovanje s tujimi partnerji. V primeru odobrenega povečanja obsega financiranja raziskovalnega programa v letu 2014 mora poročilo o realizaciji programa dela zajemati predložen program dela ob prijavi in predložen dopolnjen program dela v letu 2014. Največ 12.000 znakov vključno s presledki (približno dve strani, velikosti pisave 11). [Nazaj](#)

<sup>3</sup> Realizacija raziskovalne hipoteze. Največ 3.000 znakov vključno s presledki (približno pol strani, velikosti pisave 11). [Nazaj](#)

<sup>4</sup> V primeru bistvenih odstopanj in sprememb od predvidenega programa dela raziskovalnega programa, kot je bil zapisan v predlogu raziskovalnega programa oziroma v primeru sprememb, povečanja ali zmanjšanja sestave programske skupine v zadnjem letu izvajanja raziskovalnega programa, napišite obrazložitev. V primeru, da sprememb ni bilo, navedite: "Ni bilo sprememb.". Največ 6.000 znakov vključno s presledki (približno ena stran, velikosti pisave 11). [Nazaj](#)

<sup>5</sup> Navedite znanstvene dosežke (največ pet), ki so nastali v okviru izvajanja raziskovalnega programa. Raziskovalni dosežek iz obdobja izvajanja programa vpišete tako, da izpolnite COBISS kodo dosežka – sistem nato sam izpolni naslov objave, naziv, IF in srednjo vrednost revije, naziv FOS področja ter podatek, ali je dosežek uvrščen v A'' ali A'. [Nazaj](#)

<sup>6</sup> Navedite družbeno-ekonomske dosežke (največ pet), ki so nastali v okviru izvajanja raziskovalnega programa. Družbeno-ekonomski dosežek iz obdobja izvajanja programa vpišete tako, da izpolnite COBISS kodo dosežka – sistem nato sam izpolni naslov objave, naziv, IF in srednjo vrednost revije, naziv FOS področja ter podatek, ali je dosežek uvrščen v A'' ali A'.

Družbeno-ekonomski dosežek je po svoji strukturi drugačen kot znanstveni dosežek. Povzetek znanstvenega dosežka je praviloma povzetek bibliografske enote (članka, knjige), v kateri je dosežek objavljen.

Povzetek družbeno-ekonomskega dosežka praviloma ni povzetek bibliografske enote, ki ta dosežek dokumentira, ker je

dosežek sklop več rezultatov raziskovanja, ki je lahko dokumentiran v različnih bibliografskih enotah. COBISS ID zato ni enoznačen, izjemoma pa ga lahko tudi ni (npr. prehod mlajših sodelavcev v gospodarstvo na pomembnih raziskovalnih nalogah, ali ustanovitev podjetja kot rezultat programa ... - v obeh primerih ni COBISS ID). [Nazaj](#)

<sup>7</sup> Navedite rezultate raziskovalnega programa iz obdobja izvajanja programa v primeru, da katerega od rezultatov ni mogoče navesti v točkah 6 in 7 (npr. ker se ga v sistemu COBISS ne vodi). Največ 2.000 znakov vključno s presledki (približno 1/3 strani, velikost pisave 11). [Nazaj](#)

<sup>8</sup> Pomen raziskovalnih rezultatov za razvoj znanosti in za razvoj Slovenije bo objavljen na spletni strani: <http://www.sicris.si/> za posamezen program, ki je predmet poročanja. [Nazaj](#)

<sup>9</sup> Največ 4.000 znakov vključno s presledki (približno 2/3 strani, velikost pisave 11). [Nazaj](#)

<sup>10</sup> Največ 4.000 znakov vključno s presledki (približno 2/3 strani, velikost pisave 11). [Nazaj](#)

<sup>11</sup> Upoštevajo se le tiste diplome, magisteriji znanosti in doktorati znanosti (zaključene/i v obdobju 1.1.2009–31.12.2014), pri katerih so kot mentorji sodelovali člani programske skupine. [Nazaj](#)

<sup>12</sup> Vpišite število opravljenih diplom v času izvajanja raziskovalnega programa glede na vrsto usposabljanja. [Nazaj](#)

<sup>13</sup> Vpišite šifro raziskovalca in/ali ime in priimek osebe, ki je v času izvajanja raziskovalnega programa pridobila naziv magister znanosti in/ali doktor znanosti ter označite doseženo izobrazbo. V primeru, da se je oseba usposabljala po programu Mladi raziskovalci, označite "MR". [Nazaj](#)

<sup>14</sup> Za mlade raziskovalce, ki ste jih navedli v tabeli 11.2. točke (usposabljanje so uspešno zaključili v obdobju od 1.1.2009 do 31.12.2014), izberite oz. označite, kje so se zaposlili po zaključenem usposabljanju. [Nazaj](#)

<sup>15</sup> Navedite naslove projektov in ime člana programske skupine, ki je bil vodja/koordinator navedenega projekta. Največ 6.000 znakov vključno s presledki (približno ena stran, velikosti pisave 11). [Nazaj](#)

<sup>16</sup> Navedite naslove projektov, ki ne sodijo v okvir financiranja ARRS (npr: industrijski projekti, projekti za druge naročnike, državno upravo, občine idr.) in ime člana programske skupine, ki je bil vodja/koordinator navedenega projekta. Največ 6.000 znakov vključno s presledki (približno ena stran, velikosti pisave 11). [Nazaj](#)

<sup>17</sup> Opišite možnosti za uporabo rezultatov v praksi. Opišite izdelke oziroma tehnologijo in potencialne trge oziroma tržne niše, v katere sodijo. Ocenite dodano vrednost izdelkov, katerih osnova je znanje, razvito v okviru programa oziroma dodano vrednost na zaposlenega, če jo je mogoče oceniti (npr. v primerih, ko je rezultat izboljšava obstoječih tehnologij oziroma izdelkov). Največ 3.000 znakov vključno s presledki (približno pol strani, velikosti pisave 11). [Nazaj](#)

<sup>18</sup> Največ 1.000 znakov vključno s presledki (približno 1/6 strani, velikost pisave 11) [Nazaj](#)

<sup>19</sup> Navedite en izjemni znanstveni dosežek in/ali en izjemni družbeno-ekonomski dosežek raziskovalnega programa v letu 2014 (največ 1000 znakov, vključno s presledki, velikost pisave 11). Za dosežek pripravite diapozitiv, ki vsebuje sliko ali drugo slikovno gradivo v zvezi z izjemnim dosežkom (velikost pisave najmanj 16, približno pol strani) in opis izjemnega dosežka (velikost pisave 12, približno pol strani). Diapozitiv/-a priložite kot prilonko/-i k temu poročilu. Vzorec diapozitiva je objavljen na spletni strani ARRS <http://www.arrs.gov.si/sl/gradivo/>, predstavitve dosežkov za pretekla leta pa so objavljena na spletni strani <http://www.arrs.gov.si/sl/analize/dosez/>. [Nazaj](#)

Obrazec: ARRS-RPROG-ZP/2015 v1.00a

F1-3B-ED-11-5C-5E-93-10-30-5C-A6-57-86-6D-DC-4C-51-4C-2A-46