

BIBLIOMETRICKÁ ANALÝZA PUBLIKAČNÍ AKTIVITY VÝZKUMNÝCH ORGANIZACÍ

JIŘÍ MÁLEK

Univerzita Pardubice, Fakulta chemicko-technologická, Studentská 573, 532 10 Pardubice

Celostátní hodnocení výsledků výzkumu v rámci platného Zákona o podpoře výzkumu (130/2002 Sb.) doposud realizované v ČR je založeno na Metodice hodnocení výsledků výzkumných organizací, kterou lze chápat jako prováděcí předpis schvalovaný vládou. Posledním předpisem, podle kterého proběhla hodnocení H2014, H2015 a H2016, je tzv. Metodika 2013 (cit.¹). Tento i předchozí způsob hodnocení výzkumných organizací byl založen na výpočtu sumární bodové hodnoty výsledků vložených do registru informací o výsledcích (RIV) v rámci informačního systému výzkumu, experimentálního vývoje a inovací (IS VaV). Do hodnocení jsou vždy zahrnuty všechny záznamy o výsledcích zařazené v RIV, bez ohledu na typ aktivity nebo zdroj financování, a to vždy za pětileté období, které končí na konci kalendářního roku, který hodnocení předcházelo. Rozhodným ukazatelem pro zařazení výsledku do hodnocení je rok uplatnění výsledku uvedený v RIV.

Metodika podrobně popisuje pravidla při hodnocení výsledků, včetně řešení sporů a nesrovnalostí v předaných datech. Výsledkem je nejen přiřazení bodového ohodnocení výsledků, ale také podílů na výsledcích jednotlivých předkladatelů. Na tuto druhou, ale podstatnou část hodnocení výsledků se často zapomíná. Bodování výsledků však bylo a stále je mnohými kritizováno. Nejčastější kritika spočívá v několika hlavních bodech: (1) bodování výstupů neodráží rozdílné publikační zvyklosti jednotlivých oborů, (2) bodová hodnota výsledků údajně nezobrazuje dostatečně věrně kvalitu vědeckých výstupů, (3) tento systém může vést k nesprávným motivacím. Třetí bod lze asi použít pro charakterizaci jakéhokoli systému hodnocení, který běží dostatečně dlouhou dobu, aby se jej mnozí naučili obcházet či zneužívat a také aby získal své hlasité odpůrce. Příznivci obvykle mlčí. Druhý bod je přinejmenším nepřesný. Bodová hodnota poměrně věrně zrcadlí kvalitu výstupů. To platí zejména pro publikace v impaktovaných časopisech. Jinou věcí ovšem je, že celkový součet bodových hodnot nemusí říkat nic o tom, jak jej bylo dosaženo. Stejně číslo může být výsledkem agregace několika velmi kvalitních výsledků s velkými bodovými hodnotami nebo naopak mnoha výsledků s nepatrnou hodnotou. První bod je však zásadní námitkou, kterou je třeba podrobněji prozkoumat.

Nedávno bylo popsáno, jakým způsobem lze veřejně přístupné výsledky Hodnocení 2014 (H2014) analyzovat, tak aby byla respektována oborová specifika. Konkrétně se

jednalo o data v oborové skupině 09-Chemie (cit.²). Tentokrát se blíže podíváme na celkem čtyři oborové skupiny:

08-Fyzikální vědy zahrnují RIV obory: BE-Teoretická fyzika, BF-Elementární částice a fyzika vysokých energií, BG-Jaderná, atomová a molekulová fyzika, urychlovače, BH-Optika, masery a lasery, BI-Akustika a kmity, BJ-Termodynamika, BK-Mechanika tekutin, BL-Fyzika plazmatu a výboje v plynech, BM-Fyzika pevných látek a magnetismus, BN-Astronomie a nebeská mechanika, BO-Biofyzika.

09-Chemické vědy zahrnují RIV obory: CA-Anorganická chemie, CB-Analytická chemie, separace, CC-Organická chemie, CD-Makromolekulární chemie, CE-Biochemie, CF-Fyzikální chemie a teoretická chemie, CG-Elektrochemie, CH-Jaderná a kvantová chemie, fotochemie a CI-Průmyslová chemie a chemické inženýrství.

10-Biologické vědy zahrnují RIV obory: DN-Vliv životního prostředí na zdraví, EA-Morfologické obory a cytologie, EB-Genetika a molekulární biologie, EC-Imunologie, ED-Fyziologie, EE-Mikrobiologie, virologie, EF-Botanika, EG-Zoologie, EH-Ekologie–společnost, EI-Biotechnologie a bionika.

11-Lékařské vědy zahrnují RIV obory: FA-Kardiovaskulární nemoci včetně kardiologie, FB-Endokrinologie, diabetologie, metabolismus, výživa, FC-Pneumologie, FD-Onkologie a hematologie, FE-Ostatní obory vnitřního lékařství, FF-ORL, oftalmologie, stomatologie, FG-Pediatry, FH-Neurologie, neurochirurgie, neurovědy, FI-Traumatologie, ortopedie, FJ-Chirurgie včetně transplantologie, FK-Gynekologie a porodnictví, FL-Psychiatrie, sexuologie, FM-Hygiena, FN-Epidemiologie, infekční nemoci a klinická imunologie, FO-Dermatovenerologie, FP-Ostatní lékařské obory, FQ-Veřejné zdravotnictví, sociální lékařství, FR-Farmakologie a lékárnická chemie.

Analýza dat H2014 je založena na podrobné bilanci impaktovaných publikací evidovaných v RIV (období 2009–2013). Pro každou výzkumnou instituci (VŠ po fakultách, AV po ústavech) byl podle Metodiky 2013 (cit.¹) určen publikační podíl, každé jednotlivé impaktované publikace. Zaokrouhlené celkové součty těchto publikačních podílů v jednotlivých oborových skupinách jsou uvedeny v tab. I.

K vyjádření kvality publikačního výstupu z praktických důvodů je využito tzv. normované oborově specifické pořadí (NOP) uvedené v databázi Web of Science, které se určí podle vztahu:

$$\text{NOP} = (P-1) / (P_{\max} - 1)$$

kde P je pořadí časopisu v daném oboru podle Journal Citation Reports (JCR) v řadě seřazené sestupně podle impaktního faktoru (IF) a P_{\max} odpovídá celkovému počtu

Tabulka I

Celkové počty a podíl impaktovaných publikací v kvalitnější polovině časopisů

Oborová skupina	N_{os}	Y_{os} [%]
08-Fyzikální vědy	5681	74,92
09-Chemické vědy	6449	71,83
10-Biologické vědy	6161	60,89
11-Lékařské vědy	6091	47,83

časopisů v daném oboru. Pokud je časopis zařazen do více oborů, je NOP určeno jako aritmetický průměr NOP ve všech oborech, kam je časopis v JCR zařazen. Stejný postup byl použit v Metodice 2013 (cit.¹) a v předchozích hodnoceních výsledků výzkumných organizací v ČR (včetně H2014).

Jednotlivé časopisy lze podle hodnoty NOP rozdělit do decilů. Následně můžeme vyjádřit procentuální publikační podíl instituce v kvalitnější polovině časopisů Y_i , který se vypočte podle vztahu:

$$Y_i = 100 \cdot \left(\frac{N_{1D}}{N_i} + \frac{N_{2D}}{N_i} + \frac{N_{3D}}{N_i} + \frac{N_{4D}}{N_i} + \frac{N_{5D}}{N_i} \right)$$

kde N_{kD} je počet přepočtených impaktovaných publikací instituce v k-tém decilu pro posuzovanou oborovou skupinu, N_i odpovídá součtu N_{kD} ve všech decilech pro posuzovanou instituci a oborovou skupinu:

$$N_i = \sum_{k=1}^{10} N_{kD}$$

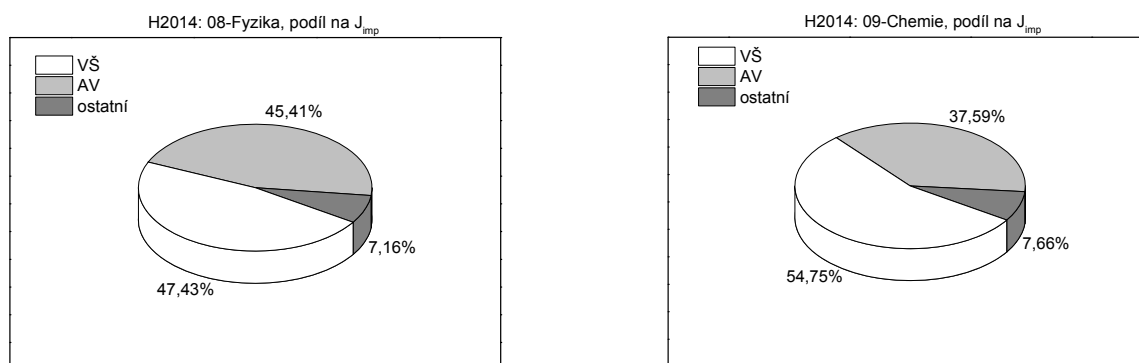
Jako kritérium publikačního výkonu instituce lze například definovat součin publikačního podílu instituce v první polovině kvalitnějších časopisů Y_i a relativního podílu počtu impaktovaných publikací této instituce v dané oborové skupině (N_i) vztaženého k celkovému počtu publikací v oborové skupině (N_{os}):

$$P(Y_i) = Y_i \left(\frac{N_i}{N_{os}} \right)$$

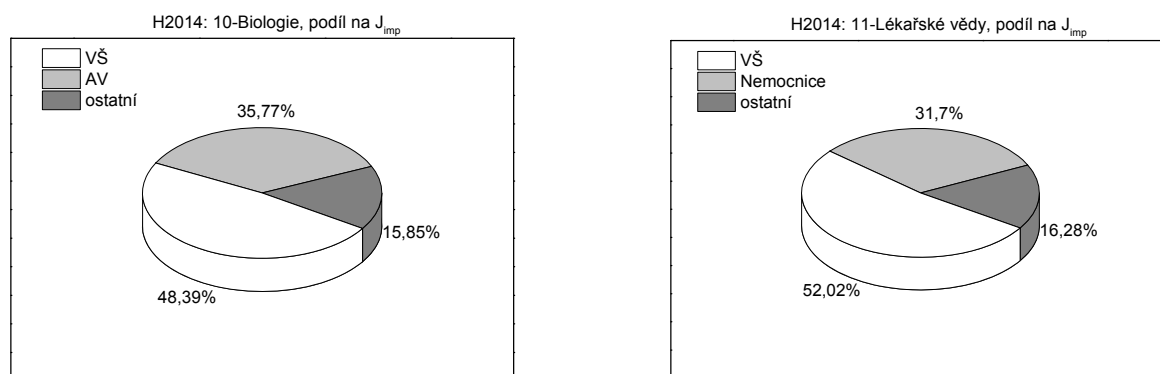
Takto definované kritérium již poměrně výrazným způsobem zohledňuje kvalitu časopisu, v němž je publikace zveřejněna. Prahová hodnota publikačního podílu instituce pro stanovení tohoto kritéria pro konkrétní výzkumnou instituci je z praktických důvodů nastavena jako $(N_i/N_{os}) > 3 \cdot 10^{-3}$.

Procentuální podíl publikací v kvalitnější polovině časopisů Y_{os} se pro každou oborovou skupinu poměrně výrazně liší (tab. I). Podobné rozdíly lze očekávat i mezi RIV obory, ale ty mohou v některých případech být příliš malé a pro statistické zpracování nevhodné. Oborové skupiny z hlediska zpracování dat představují určité optimum a celkový počet přepočtených publikací v jednotlivých oborových skupinách (N_{os}) je přibližně stejný. Podíly publikací v sektorech výzkumu jsou graficky znázorněny na obr. 1 a obr. 2. Je zřejmé, že ve všech oborových skupinách (v rámci H2014) byla většina impaktovaných publikací vytvořena na vysokých školách, za kterými s určitým odstupem následují ústavy AV. To je však čistě kvantitativní pohled bez zohlednění kvality.

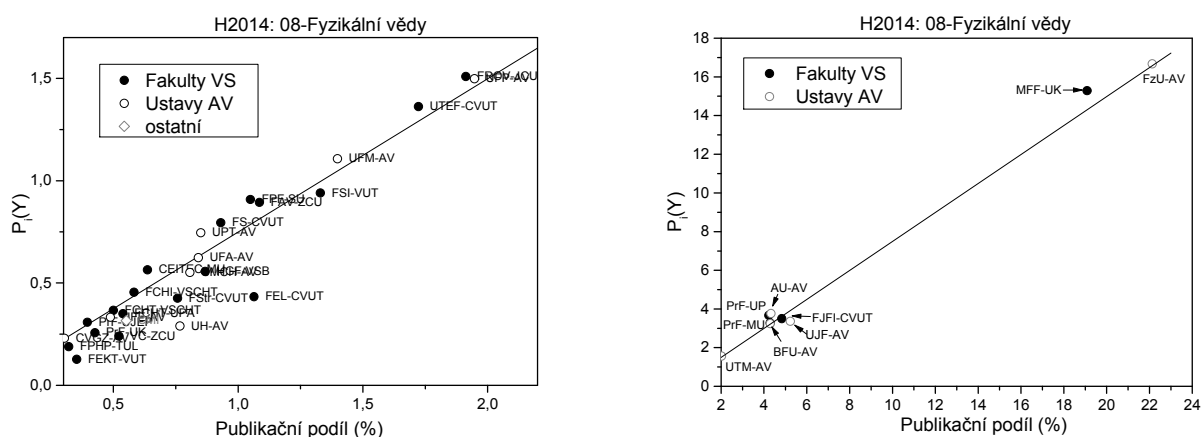
Pojďme se nyní podívat na podrobnější ilustraci publikačního výkonu a publikační kvality na jednotlivých institucích. Na obr. 3 je graficky znázorněn publikační výkon fakult, ústavů AV a dalších institucí pro oborovou skupinu 08-Fyzikální vědy vypočtený na základě dat H2014. Levá část zachycuje výzkumné instituce s menším podílem impaktovaných publikací v této oborové skupině ($N_i/N_{os} < 0,02$). Pravá část pak odpovídá institucím s největším publikačním podílem v oborové skupině. Body odpovídají pozici jednotlivých výzkumných organizací podle publikačního podílu a velikosti kritéria $P(Y_i)$. Plná čára reprezentuje medián oborové skupiny. Z těchto ilustrací je zřejmé, jakou měrou se publikační aktivita posuzované instituce shoduje s celkovým trendem oboru v ČR (v rámci dat H2014). Je též dobře patrné, že dvě největší



Obr. 1. Podíl jednotlivých sektorů výzkumu na impaktovaných publikacích v oborových skupinách 08-Fyzikální vědy a 09-Chemické vědy v rámci datového souboru H2014



Obr. 2. Podíl jednotlivých sektorů výzkumu na impaktovaných publikacích v oborových skupinách 10-Biologické vědy a 11-Lékařské vědy v rámci datového souboru H2014



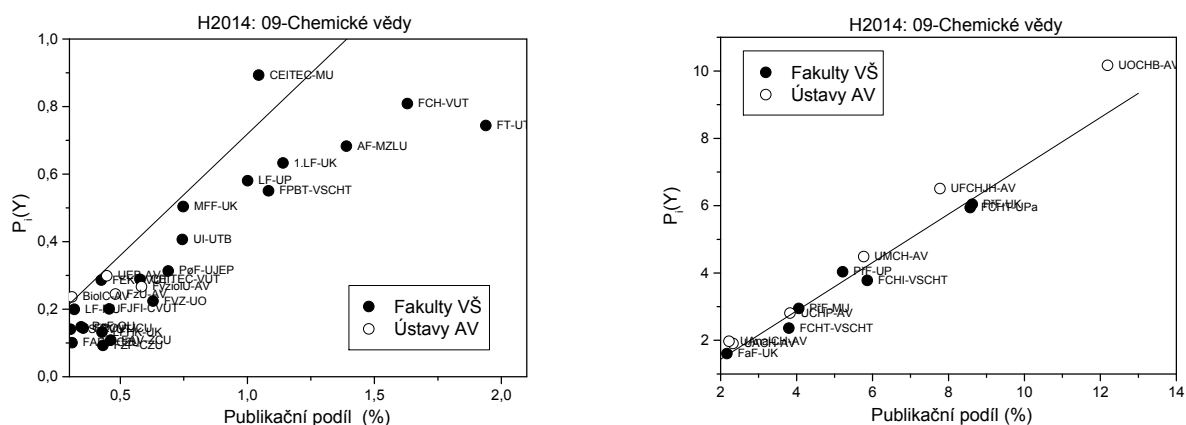
Obr. 3. Publikační výkon fakult a ústavů v oborové skupině 08-Fyzikální vědy (data H2014)

institute, které publikují v oborové skupině 08-Fyzikální vědy (tj. FzÚ-AV a MFF-UK) v publikačním výkonu násobně převyšují institute ostatní. Je zde však určitý rozdíl. Zatímco FzÚ-AV prakticky kopíruje medián oboru, tak MFF-UK je umístěna výrazněji nad mediánem, což ukazuje na četnější publikace v kvalitnější polovině impaktovaných časopisů. Podobný trend vykazují také například AU-AV a PřF-UP.

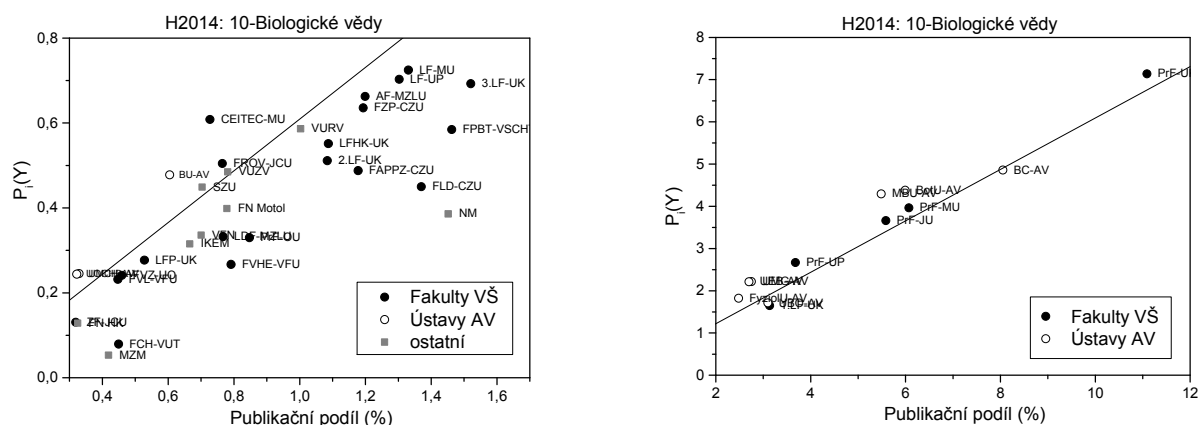
Na obr. 4 je graficky znázorněn publikační výkon fakult, ústavů AV a dalších institucí pro oborovou skupinu 08-Chemické vědy vypočtený na základě dat H2014. Podobně jako v předchozím případě, levá část zachycuje výzkumné institute s menším procentuálním podílem impaktovaných publikací v této oborové skupině ($< 2\%$) a pravá část pak odpovídá institucím s dominantním publikačním podílem v oborové skupině. Pozice těchto publikačně významnějších institucí jsou poměrně symetricky rozloženy podél mediánu oborové skupiny s výjimkou

některých ústavů AVČR, které zřetelně výrazněji publikují v kvalitnější polovině časopisů (UOCHB-AV a UFCHJH-AV, UAnalCH-AV). Zcela jiná situace je patrná pro institute s menším publikačním podílem. Zde je naopak patrná (s výjimkou CEITEC-MU) nezanedbatelná tendence publikovat v horší polovině impaktovaných časopisů.

Publikační výkon fakult, ústavů AV a dalších institucí pro oborovou skupinu 10-Biologické vědy vypočtený na základě dat H2014 je zobrazen na obr. 5. Podobně jako v předchozím případě, levá část zachycuje výzkumné institute s menším podílem impaktovaných publikací v této oborové skupině ($< 2\%$) a pravá část pak odpovídá institucím s dominantním publikačním podílem v oborové skupině. Pozice těchto publikačně významnějších institucí jsou opět poměrně symetricky rozloženy podél mediánu oborové skupiny s výjimkou některých ústavů a jedné fakulty, které výrazněji publikují v kvalitnější polovině impaktovaných časopisů (BotU-AV, MBU-AV, PřF-UP, UMG-AV,



Obr. 4. Publikační výkon fakult a ústavů v oborové skupině 09-Chemické vědy (data H2014)



Obr. 5. Publikační výkon fakult a ústavů v oborové skupině 10-Biologické vědy (data H2014)

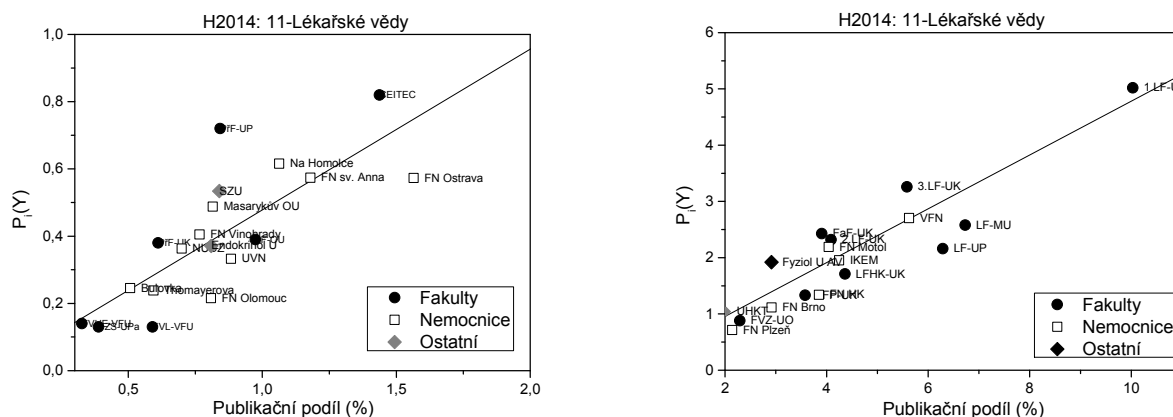
UEB-AV, FyziolU-Av a UEM-AV). Jinak je tomu pro instituce s menším publikačním podílem. Zde je naopak patrná nezanedbatelná tendence publikovat v horší polovině impaktovaných časopisů. Výjimkou je např. CEITEC-MU, BU-AV, FROV-JCU a VUZV. Je však třeba vzít v úvahu, že zejména pro velmi malé instituce, kde publikační výkon je na úrovni přibližně 1% podílu v oborové skupině, již posuzujeme pouze 60 impaktovaných publikací za 5 let (H2014), což může někdy odpovídat publikačnímu výkonu jedné výzkumné skupiny.

Publikační výkon fakult, ústavů AV a dalších institucí pro oborovou skupinu 11-Lékařské vědy vypočtený na základě dat H2014 je zobrazen na obr. 6. Body odpovídají pozici jednotlivých výzkumných organizací podle publikačního podílu a kritéria $P(Y_i)$. Plná čára reprezentuje medián oborové skupiny. Pozitivní odchylka od mediánu upozorňuje na pracoviště, která výrazněji publikují v kvalitnější polovině impaktovaných časopisů (3.LF-UK, FaF-UK, 2.LF-UK, FyziolU-AV, UEM-AV, PŘF-UP, PŘF-

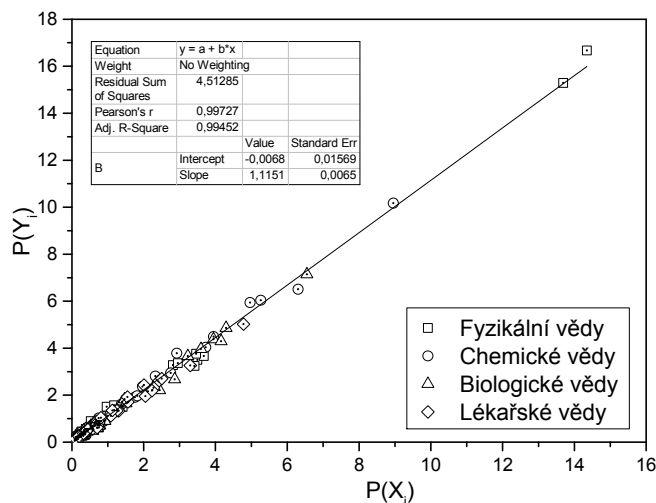
UK a SZU). Podobným způsobem lze nalézt pracoviště, která naopak dominantně publikují v horší polovině časopisů a nacházejí se tedy pod mediánem oborové skupiny.

Výhodou popisovaného kritéria $P(Y_i)$ je možnost přímého porovnání kvality publikačního výkonu institucí s mediánem oborové skupiny. Zřejmou nevýhodou je naopak to, že nezahrnuje publikace v druhé polovině časopisů řazených sestupně podle NOP. Pokud bychom je chtěli zahrnout, máme například možnost využít kritéria $P(X_i)$, které bylo popsáno dříve². Určitou výhodou kritéria $P(X_i)$ je navíc jeho aditivita. Výslednou hodnotu za celou instituci lze například složit součtem příslušných fakult. Číselně se obě tato kritéria výrazně neliší. Z obr. 7 je navíc patrné, že pro všechny posuzované instituce v rámci oborových skupin 08 – 11 (H2014), lze nalézt poměrně dobrou korelaci obou kritérií $P(X_i)$, $P(Y_i)$.

Popsaný způsob porovnání publikačního výkonu fakult, ústavů a dalších výzkumných institucí nelze považovat za hodnocení vědecké produktivity. K tomu by bylo



Obr. 6. Publikační výkon fakult, ústavů a nemocnic v oborové skupině 11-Lékařské vědy (data H2014)

Obr. 7. Korelace kritérií $P(X_i)$ a $P(Y_i)$ pro oborové skupiny 08 – 11

třeba vztáhnout číselné hodnoty kritéria $P(Y_i)$ například k počtu pracovníků, kteří se podíleli na získání posuzovaných bibliometrických výsledků. S využitím funkční databáze RIV v rámci IS VaVal je to technicky proveditelné.

LITERATURA

1. Metodika hodnocení výsledků výzkumných organizací a hodnocení výsledků ukončených programů (platná pro léta 2013 – 2015), Úřad vlády České republiky, Č.j.:1417/2013-RVV.
2. Málek J.: Chem. Listy 111, 681 (2017).