Brief Communication: Some Critical Notes on the Czech System of Evaluation of Research

Roman Solc

Charles University, Faculty of Science, Departement of Anthropology and Human Genetics,
Vinicna 7, Prague 2, 128 43, Czech Republic,
<roman.solc@natur.cuni.cz>

Roman Solc (1986) is a member of the Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Prague. His specialization is human molecular cytogenetics. Further he deals with genetics of circadian rhythm, didactics of biology and science evaluation. He also devotes himself to popularizing science as an author, translator or lecturer.

Solc, Roman. 2019. "Brief Communication: Some Critical Notes on the Czech System of Evaluation of Research." *Knowledge Organiztion* 46)5): 387-397. 4 references. DOI:10.5771/0943-7444-2019-5-387.

Abstract: We try to critically analyze some aspects of the system of evaluation of scientific research in the Czech Republic. The government presents it as very objective, fair and motivating. Using the evaluating methodology of the state, data obtained from *Journal Citation Reports* and mathematical models we would like to demonstrate that the system is neither fair nor motivating.

Received: 7 June 201; Revised: 19 June 2019; Accepted: 21 June 2019

Keywords: RIV points, journals, categories, scientific outputs, Czech Republic

1.0 Introduction

The system of evaluation of scientific outputs in the Czech Republic has been redeemed by some changes in the last ten years or so and has become more complex and more differentiated. But the articles in the impacted journals are still considered the primary scientific outputs, associated mainly with the science financing, so there has to be a sophisticated method to their evaluation.

The whole system of evaluation of scientific outputs has been established by viewing all outputs that are registered in the database known as the *Information Register of R&D results* (further only "RIV"). This database is curated by The Office of the Government of the Czech Republic, and it is publicly accessible. Every output receives some points (further only RIV-points), which are allocated to individual authors and institutions. They are awarded according to a procedure that is defined in the material called *Methodology of Evaluation of Research Organizations and Evaluation of Finished Programmes* (valid for the current year or years) (further only "*Methodology*"), created by the governmental The Research, Development and Innovation Council (a body appointed by the government composed of politicians, scientists and industrialists). The obtained RIV-points are one of the essential indicators by which the money for science is distributed from the state budget to the institutions.

The impact factor of the journal is very important when evaluating these articles, but the position of the journal in the category (or categories) of ranking in the *Journal Citation Reports* (*JCR*)(when the order of journals is defined by their impact factors) is absolutely crucial. More points are given to articles published in the journals with the highest impact factor, more precisely to those occupying the first places in the category list. These points are awarded according to the mathematical algorithm which has slightly changed in the recent years. How one changeover in that algorithm, which was carried out between years 2009 and 2010, handicaps interdisciplinary science was described earlier (Solc 2014).

The algorithm is

$$\text{RIV-points} = 10 + 295*[(1-N)/(1+(N/0{,}057))]$$

where "N-value" is the normalized ranking of the periodical, and

$$N = (P-1)/(P_{max}-1)$$

where "P" is the periodical's ranking according to the *JCR* in a series sorted in the descending order by IF and where "P_{max}" is the total number of periodicals in the given field according to the *JCR* (VVI 2009; 2010; 2013).

The difference between *Methodology 2009* and *Methodology 2010* becomes evident when a single journal is rated in more than one category. The older version of *Methodology* (VVI 2009) states that "if the periodical is registered for several fields, that field will be used for evaluation, in which the periodical achieves better order in a relation to total number of periodicals in the field." Newer *Methodologies* (VVI 2010; 2013) say that "if the periodical is registered for several fields, the normalized ranking of the periodical 'N' will be calculated as the arithmetic average of the normalized rankings of the periodical in all fields where it is registered." One negative effect of that changeover associated with science interdisciplinarity was described earlier (Solc 2014). These conditions determined by the more recent *Methodologies* will be further analyzed in this study.

The newest *Methodology* (VVI 2013) brought one more change in that the *JCR* ranking should be adjusted by exclusion of a significant number of self-citations in journals. This changeover has had only small impact on the whole system, and it is not reflected. Moreover, in the Czech evaluation system, there are some small alterations of the final RIV-points counts, which are technically impossible to register, process and analyze, but their importance is quite marginal.

That system of evaluation has been presented as quite objective and motivating for scientists. In this article, we would like to show in some critical notes, that the system is neither fair nor motivating. In fact, it tends to worsen the quality of science outputs, and it is quite discriminatory among scientific disciplines.

2.0 Material and methods

To get some real and relevant data we have randomly chosen twenty-five categories from the *JCR* database (ten from *Social Science Citation Index* (*SSCI*), fifteen from *Science Citation Index* (*SCI*)), and we have acquired some basic characteristics of all the journals in these categories (valid for 2013). We have been able to find out how many journals are included in every category, in how many categories every journal is generally included, what is the impact factor (IF) of every journal, what is the position of every journal in the categories in which it is included, what is the highest IF in every category ("IF (max)") and what is the mean IF in every category ("IF (mean)").

Then we used the algorithm (described above) and calculated the number of RIV-points, which belongs to a hypothetical article published in that journal, for every journal. First, we calculated a hypothetical ("ideal") number for a situation when the journal was included only in the category just analyzed. Secondly, we calculated the "real" number of RIV-points for the differences caused by an average of N-values from different categories taken into account. From these data we were especially interested in the highest number of RIV-points in every category ("RIV-point (max)") and in the number of RIV-points belonging to the journal with the highest IF in the category which has the N-value equal to zero ("RIV-point (N=0)"). In the "ideal" calculations, all these RIV-point values are 305. If the "RIV-point (max)" and "RIV-point (N=0)" was not the same, we found out the N-value of the journal to which the "RIV-point (max)" belongs. Finally, we added up both "real" and "ideal" RIV-point values for these journals in each and every category (which corresponds with a hypothetical situation of publishing one article in every journal in every category) and calculated the ratio "real" / "ideal" RIV-point values. All these calculations are summarized in Table 1, some of them in Figure 1.

For further analyses we defined the concept of "RIV-points / work-unit" to find a way how to confront an amount of work embedded in creating an article with a number of acquired RIV-points. We took an article in a journal with N=1 as the base-line (the necessary amount of work taken as 1) and assumed that the embedded extra work is inversely proportional to the "N-value," so when N=0 the amount of work is 30.5, proportionally according to the minimal and maximal amount of RIV-points. Then we needed a ratio between the acquired RIV-points and theoretically constructed amount of work.

So, we proposed the algorithm

$$\text{RIV-points / work-unit} = \text{RIV-points} / ((1-N)*29{,}5+1)$$

which is applicable with N from the interval <0;1>.

Just as we can construct the ideal curve of the algorithm defining the relationship between acquired RIV-points and N-value of the journal in which an article has been published, we can construct the ideal curve of the algorithm describing the relationship between this N-value and the RIV-points acquired for the unit of work (see Figure 2).

category	journals	average number of categories in which a journal is included	IF (max)	IF (mean)	RIV-points (N=0)	RIV-points (max) / ratio to 305	N-value (RIV-points max)	real/ideal RIV-points (whole category)
SSCI (Social Science Citation Index)								
Anthropology	82	1.62	5.477	1.008	305	305 (100 %)	0	89.9 %
Criminology and penology	52	1.60	3.060	1.123	305	305 (100 %)	0	96.5 %
Cultural studies	38	1.50	1.969	0.342	176	176 (57.7 %)	0	73.0 %
Demography	25	1.68	2.631	1.107	305	305 (100 %)	0	94.5 %
Ethics	50	2.34	3.887	0.875	282	282 (92.5 %)	0	79.1 %
Ethnic studies	15	2.20	1.755	0.765	73	73 (23.9 %)	0	53.0 %
Geography	76	1.76	6.000	1.233	254	254 (83.3 %)	0	98.4 %
History	72	1.67	1.293	0.270	305	305 (100 %)	0	73.9 %
Linguistic	169	1.49	3.309	0.695	81	231 (75.7 %)	0.018	75.5 %
Sociology	138	1.70	4.266	0.961	305	305 (100 %)	0	97.1 %
SCI (Science Citation Index)								
Anatomy and morphology	20	1.80	9.800	2.119	157	157 (51.5 %)	0	70.3 %
Biology	85	1.85	11.771	2.118	222	251 (82.3 %)	0.012	92.4 %
Chemistry, medicinal	58	2.28	10.715	2.577	215	222 (72.8 %)	0.018	95.0 %
Chemistry, organic	58	1.72	16.333	2.550	305	305 (100 %)	0	100.5 %
Developmental biology	41	2.02	20.241	3.497	245	245 (80.3 %)	0	95.9 %
Evolutionary biology	46	2.30	15.353	3.662	275	275 (90.2 %)	0	113.0 %
Genetics and heredity	164	2.01	39.794	3.746	305	305 (100 %)	0	107.7 %
Geology	44	1.45	4.638	1.415	305	305 (100 %)	0	95.7 %
Logic	20	2.35	0.917	0.455	117	117 (38.4 %)	0	47.5 %
Multidisciplinary sciences	55	1.38	42.351	2.663	305	305 (100 %)	0	99.2 %
Mycology	23	1.48	9.296	2.395	305	305 (100 %)	0	96.8 %
Ophthalmology	58	1.24	9.897	2.031	305	305 (100 %)	0	98.6 %
Optics	83	2.18	29.958	2.087	286	286 (93.8 %)	0	93.4 %
Rheumatology	30	1.13	10.252	3.002	305	305 (100 %)	0	103.6 %
Zoology	153	1.59	4.726	1.255	132	247 (81 %)	0.013	69.3 %

Table 1. Data from *JCR* database for 2013.

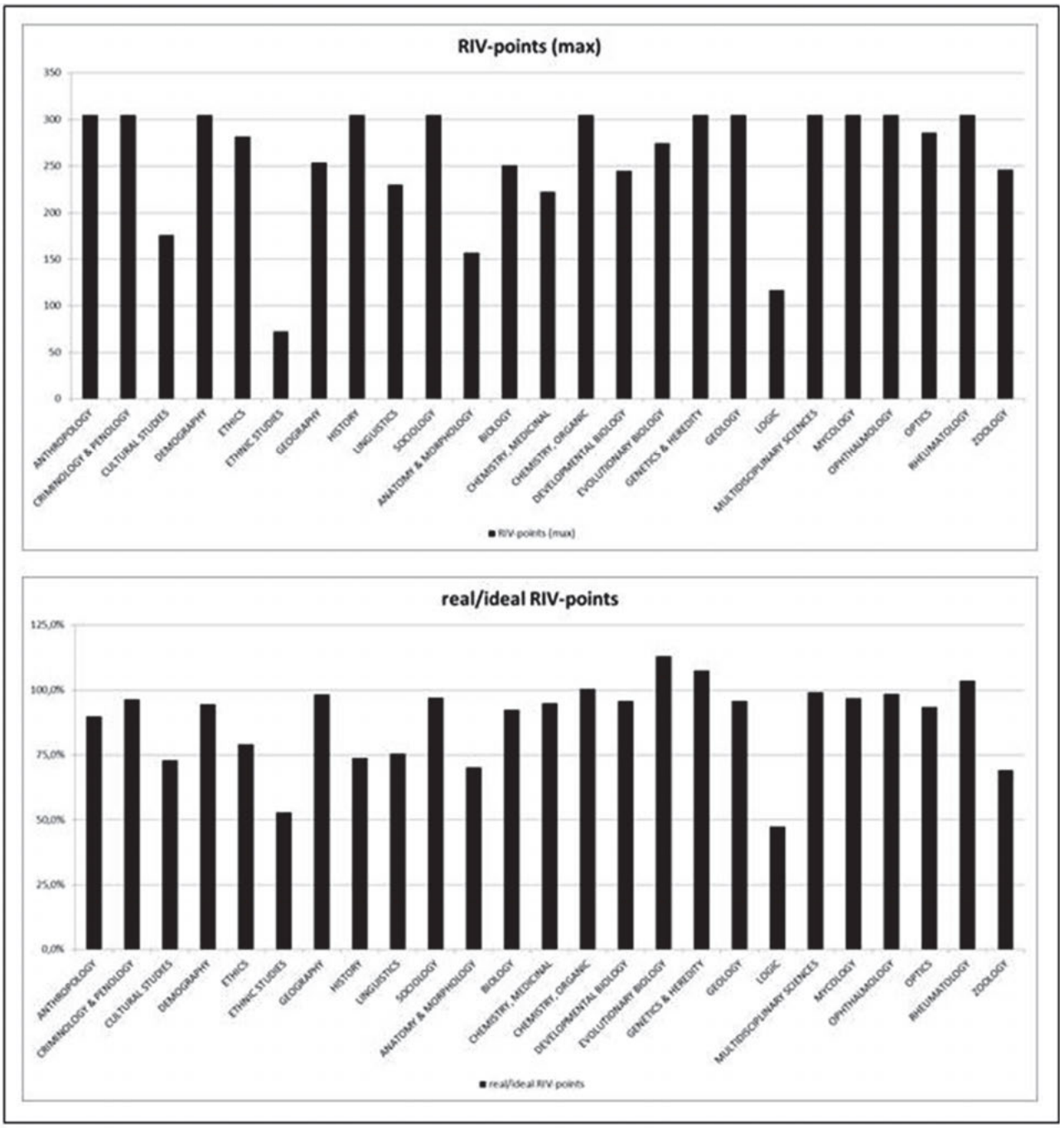

Figure 1. Upper graph: The highest number of RIV-points belonged to a journal in the category.
Lower graph: The ration between "real" and "ideal" added up RIV-point values for all journals in every category.

3.0 Results

First, we tried to analyze if there is any dissimilarity between journals either from *SSCI* or from *SCI*. We found out that from our characteristics significant differences are only in IF (max) ($p<0,01$) and IF (mean) ($p<0,0001$). In general, one can say that there are journals with a higher IF in *SCI* than in *SSCI*, which relates to a different citation culture in scientific disciplines, which are included either in *SCI* or in *SSCI*. But comparing other characteristics between categories regardless of classification in the database should be possible.

So, then we drew our attention to specific categories. It was obvious that only in twelve categories out of twenty-five (48 %) would it be possible to publish an article that could get a theoretically maximal amount of RIV-points (i.e., 305) (see Figure 1, the upper graph). When we compared "real" vs. "ideal" amounts of RIV-points obtained in

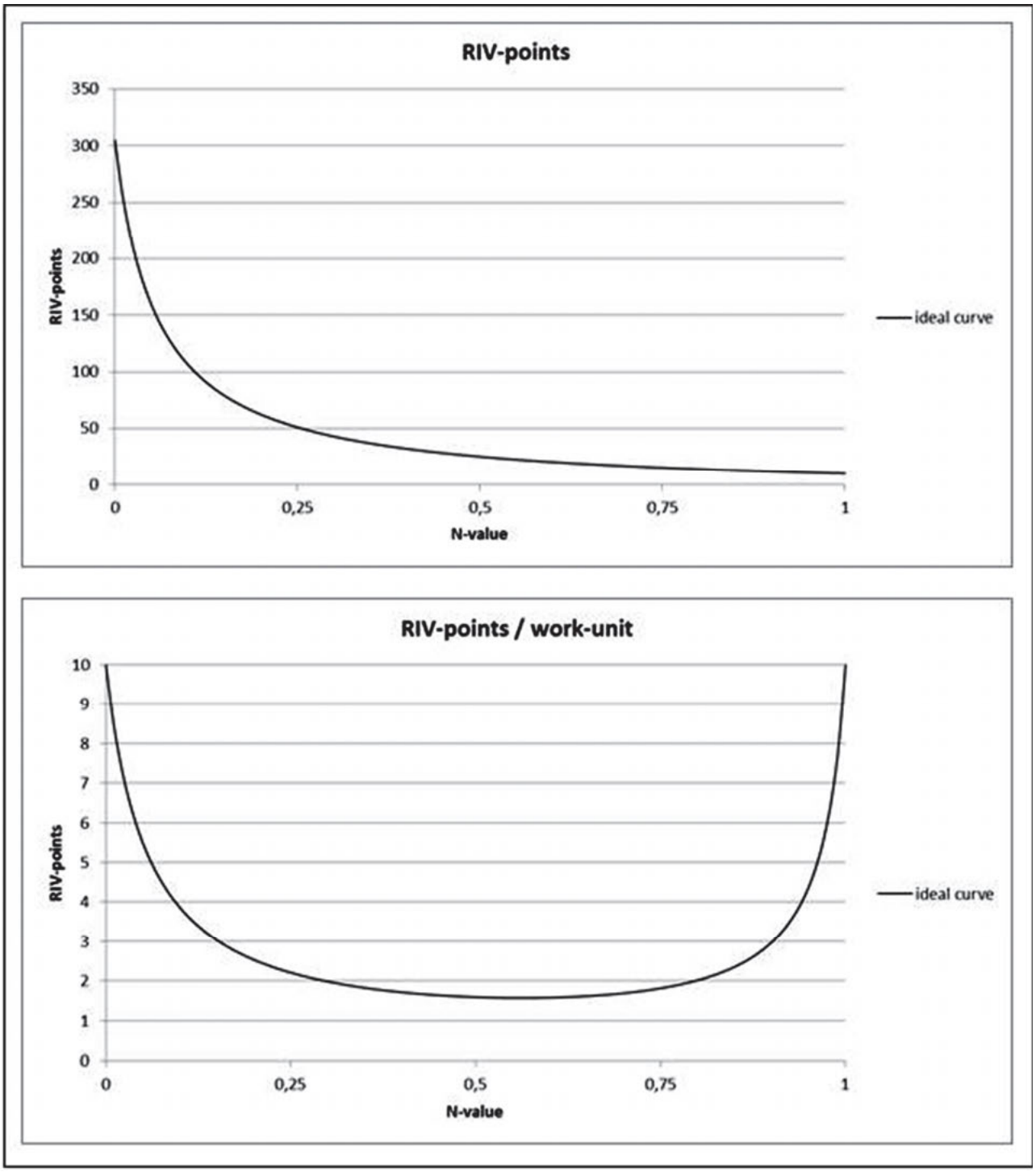

Figure 2. Upper graph: The ideal curve representing the relationship between acquired RIV-points and N-value of the journal in which an article has been published. Lower graph: The ideal curve representing the relationship between the RIV-points acquired for the unit of work and N-value of the journal in which an article has been published.

a hypothetical situation of publishing one article in every journal of the category, we found out that in twenty-one categories out of twenty-five (84 %) the "real" amount of RIV-points was lower than the "ideal" (from 47.5% to 99.2%). Only in four out of twenty-five (16 %) categories was the "real" amount of RIV-points higher than the "ideal" (from 100.5 % to 113.0 %) (see Figure 1, the lower graph).

To illustrate how the "real" distribution of RIV-points relates to the "ideal curve" in each category, there is a collection of graphs attached (see Figures 3-6).

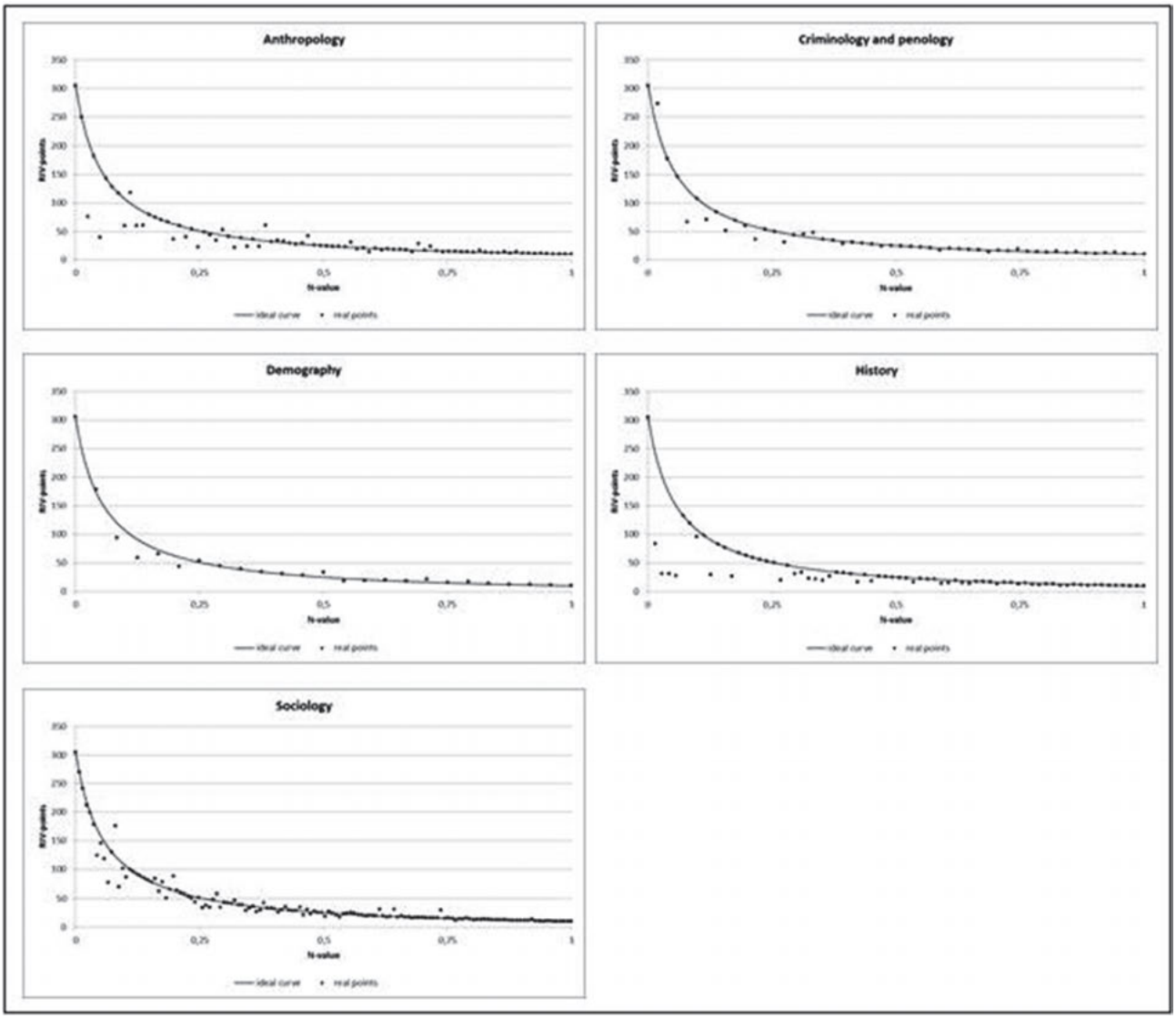

Figure 3. The comparison between the ideal allocation of RIV-points (curve) and the real number of RIV-points assigned to the journals in the category (points). These graphs represent a sample of categories from *SSCI* which have RIV-points (max) = 305.

Finally, we put together all journals from all the categories and focused on the amount of "real" RIV-points given to an article published in these journals and to our analytical value "RIV-point/work-unit" (see Figure 7). We can see that a lot of points ("real" RIV-points of the journal) lay under the ideal curve, so in fact the amount of RIV-points given to any article published in these journals is lower than it should be according to the ideal curve. Then we can see that many points ("RIV-point/work-unit" of the journal) lay under the ideal curve mainly on the left side (lower N-value). This means that the work on articles published in these journals is priced less than it should be in the ideal situation.

4.0 Discussion

In the discussion, we focused on two aspects—the fairness and the motivation. First, we could disclaim that the system of evaluation is fair on account of very big differences in the possibility of getting an equal amount of RIV-points in different disciplines. The average of N-value leads to its reduction in the case of articles published in many of the journals (journals indexed in more than one *JCR* category), mainly in those with interdisciplinary focus and mostly in journals standing on the border between science and humanities (indexed both in *SCI* and *SSCI*) (compare with Solc 2014). Because of that, some scientists cannot get the maximal amount of RIV-points even when publishing in the best journal of their respective discipline (imagine the "glass ceiling" in Figures 4 and 6). So, the amount of RIV-

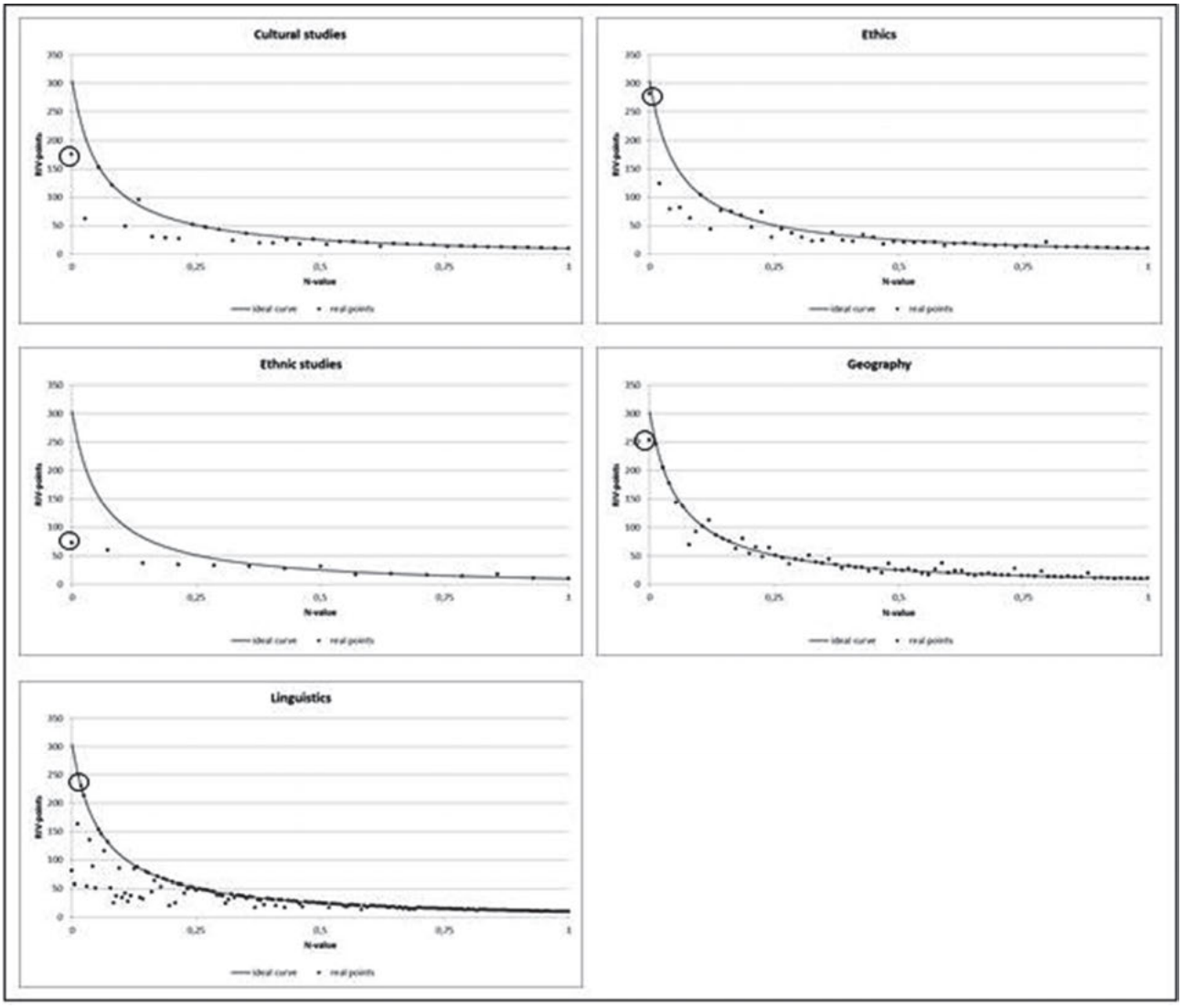

Figure 4. The comparison between the ideal allocation of RIV-points (curve) and the real number of RIV-points assigned to the journals in the category (points). These graphs represent a sample of categories from *SSCI* which have RIV-points (max) < 305 (the point in the ring).

points the scientist gets for his or her publication does not correspond with any objective criterion or parameter of quality, prestige or scientific level (neither IF) of neither a journal nor an article.

Second, we could observe, looking at our "RIV-points / work-unit" ideal curve, that the most effective way of applying the outputs of any scientific work is to publish them either in journals with a very high IF (very low N-value; e.g., the first quartile of *JCR* categories) or in journals with a very low IF (very high N-value; e.g., the fourth quartile of *JCR* categories). In fact, the most disadvantageous way is to publish in journals from the second and third quartile of *JCR* categories (the journals with an average IF). But when we look at the "real" data (see Figures 3, 4, 5 and mainly 6) we can see that lots of the journals from the first quartile are affected by the N-value average, and that these journals are relatively the most handicapped. So paradoxically, the most effective way to apply the scientific output is to publish them in journals with the lowest IF. There is no motivation by the system of evaluation to publish in the best journals or in journals with the highest IF (and lowest N-value) respectively. The system probably has just banked on the honour, self-respect and consciousness of the scientists aspiring to publish their work in the best quality and most prestigious journals of their discipline.

Hardly anyone knows what the future of the Czech system of evaluating of the scientific research will look like. The complex system described above was partly suspended in 2016, and a new one was announced. Since 2017 some governmental documents have been released, which describe the main general principles of the new system and

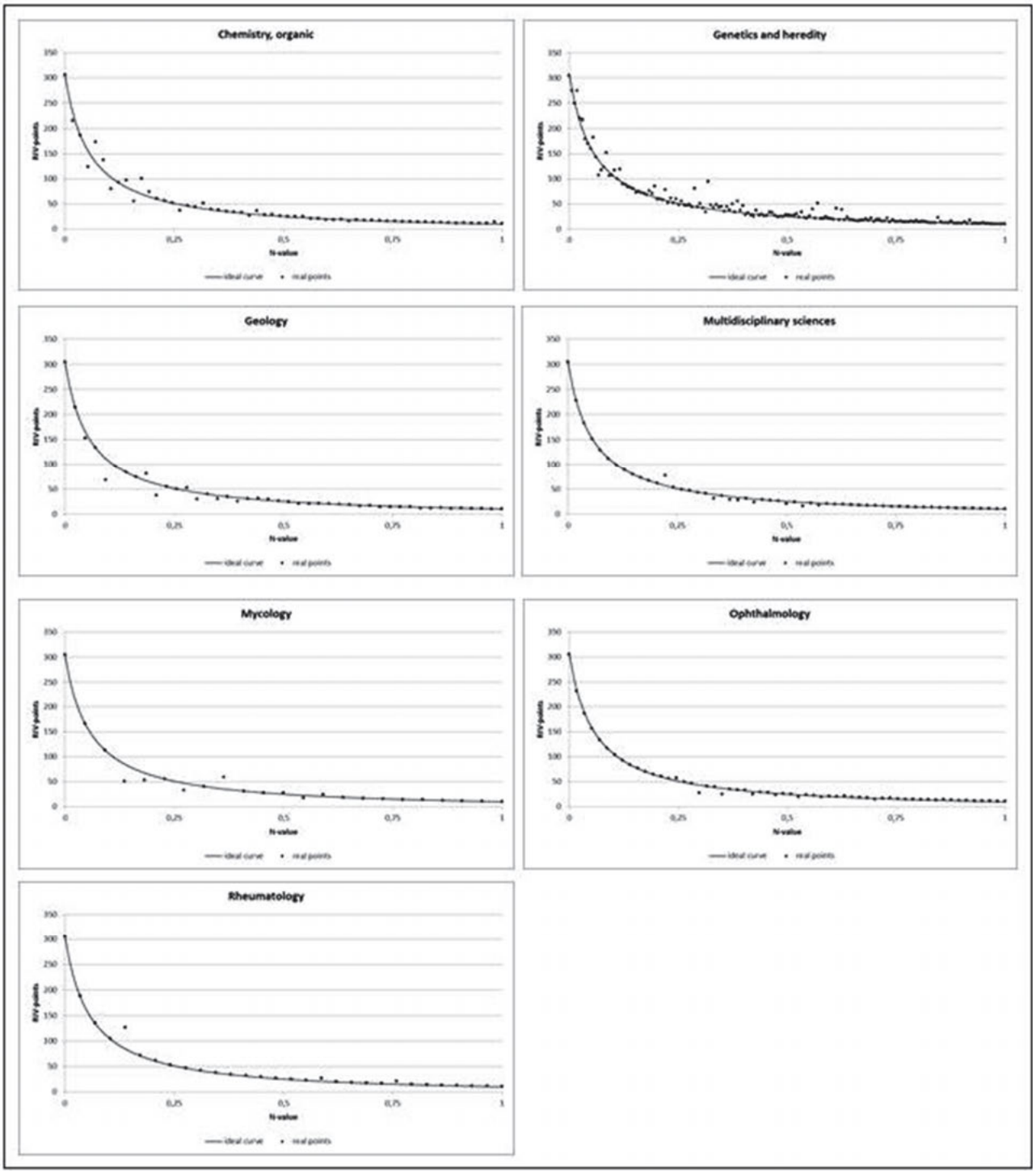

Figure 5. The comparison between the ideal allocation of RIV-points (curve) and the real number of RIV-points assigned to the journals in the category (points). These graphs represent a sample of categories from *SCI*, which have RIV-points (max) = 305.

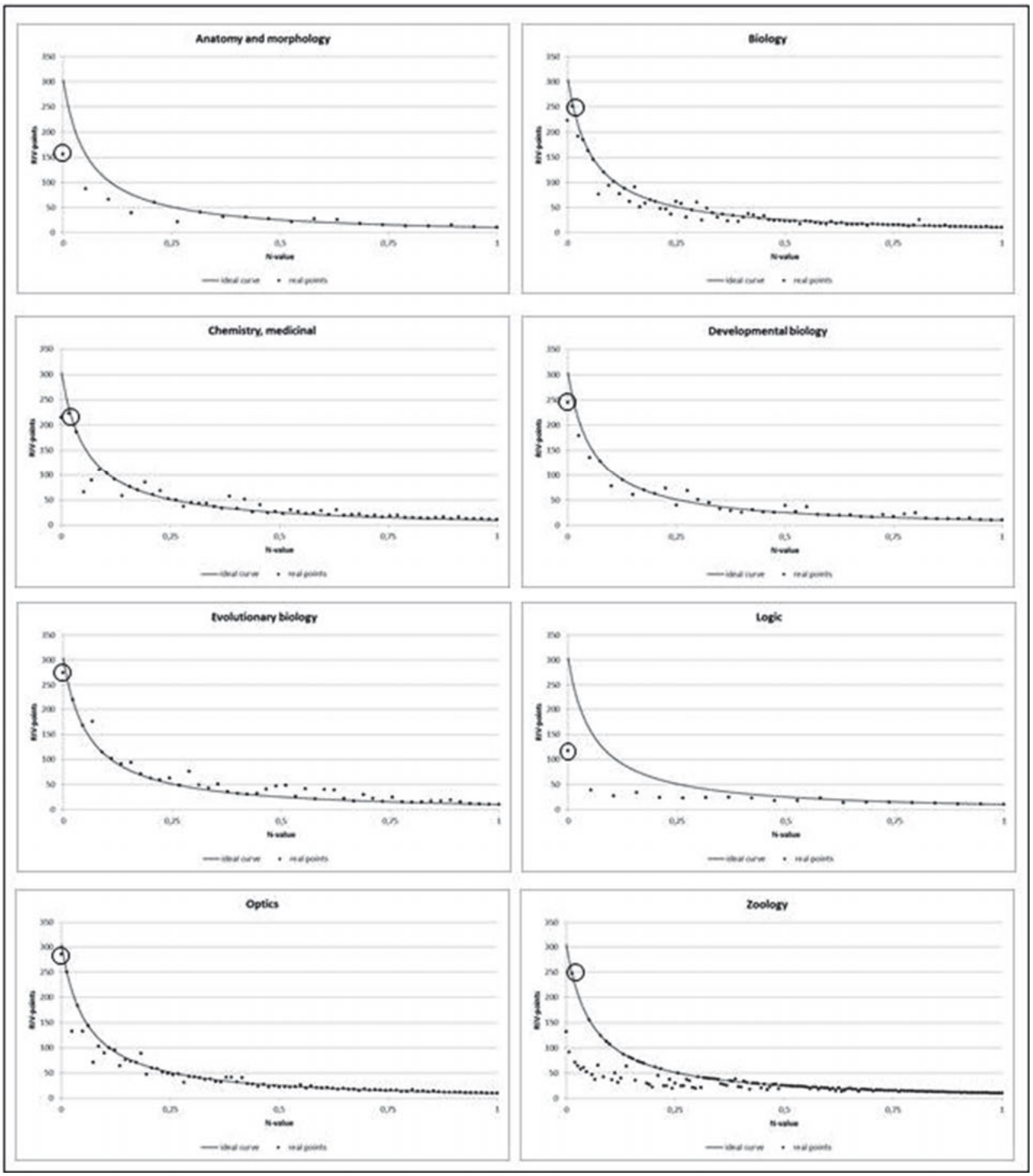

Figure 6. The comparison between the ideal allocation of RIV-points (curve) and the real number of RIV-points assigned to the journals in the category (points). These graphs represent a sample of categories from *SCI* which have RIV-points (max) < 305 (the point in the ring).

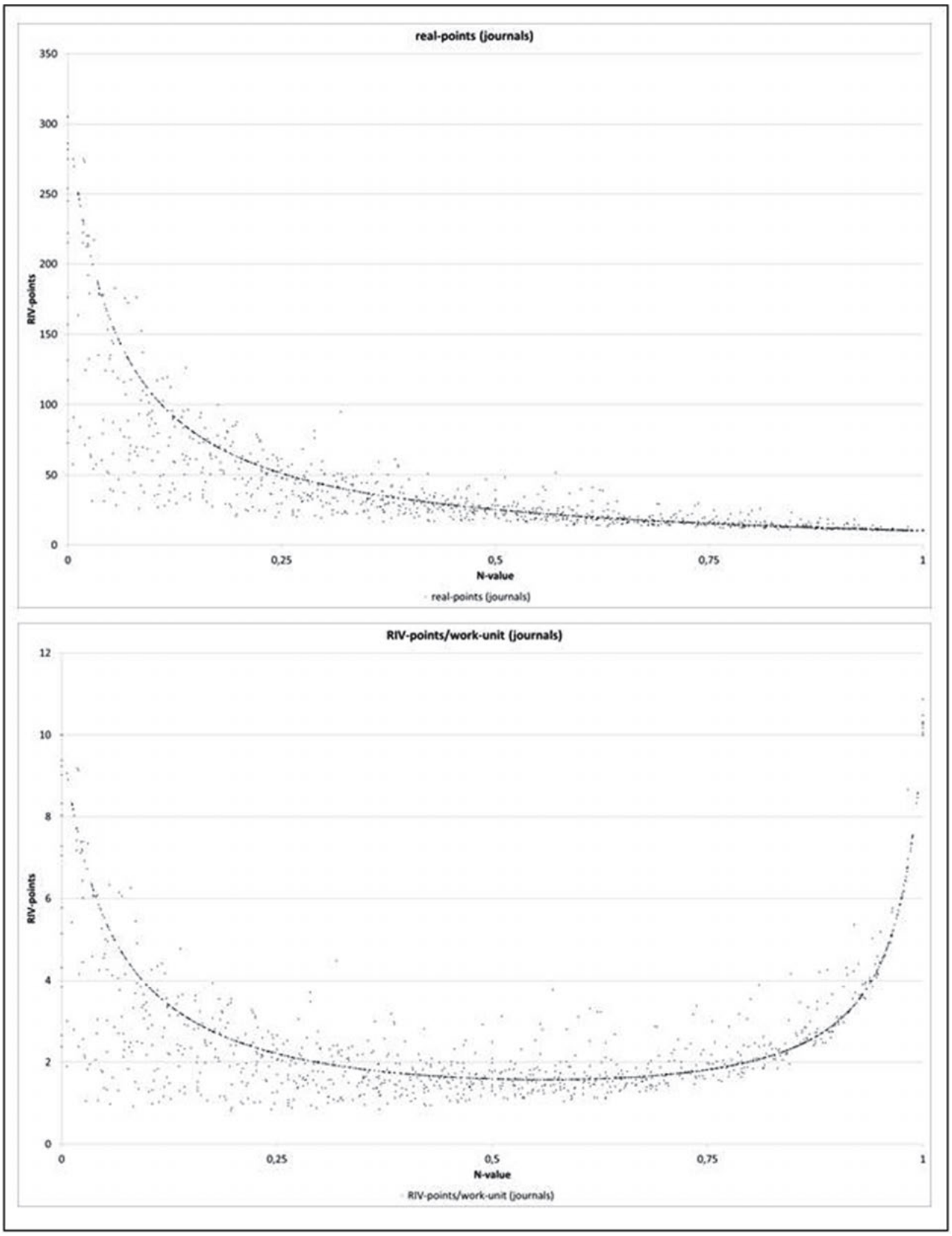

Figure 7. Upper graph: The "real" RIV-points that belong to all the journals in relation with their N-value. Lower graph: The "real" RIV-points/work-unit values of all the journals in relation with their N-values.

the process of its implementation that should be finished in 2020. But it has not been introduced to the scientific community in its entirety so far.

References

Solc, Roman. 2014. "Brief Communication: The Czech System of Evaluation of Science Research Handicaps Interdisciplinary Science." *Knowledge Organization* 41: 410-3.

VVI (Rada pro výzkum, vývoj a inovace) [Research, Development and Innovation Council]. 2009. "Hodnocení výsledků výzkumu a vývoje v roce 2009" [Methodology of Evaluation of Research and Development in year 2009]. https://www.vyzkum.cz/FrontClanek.aspx?idsekce=532412

VVI (Rada pro výzkum, vývoj a inovace). 2010. "Metodika hodnocení výsledků výzkumných organizací a hodnocení výsledků ukončených programů" [Methodology of Evaluation of Research Organizations and Evaluation of Finished Programmes (valid for years 2010 and 2011)]. https://www.vyzkum.cz/FrontClanek.aspx?idsekce=566918

VVI (Rada pro výzkum, vývoj a inovace). 2013. "Hodnocení VaVaI 2016" [Methodology of Evaluation of Research Organizations and Evaluation of Finished Programmes (valid for years 2013 to 2016)]." https://www.vyzkum.cz/FrontClanek.aspx?idsekce=799259