# Publication Trajectories of Russian Cardiologists: Ways to Attain International Visibility

Dyachenko, E.; Mironenko, A.

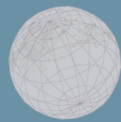**Note:** To cite this publication please use the final published version (if applicable).

**STI 2018 Conference Proceedings**

*Proceedings of the 23rd International Conference on Science and Technology Indicators*

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

**Chair of the Conference**

Paul Wouters

**Scientific Editors**

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

**Layout**

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

# Publication Trajectories of Russian Cardiologists: Ways to Attain International Visibility

Ekaterina Dyachenko[*], Asya Mironenko[**]

[*] *edyachenko@hse.ru*
Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, Moscow, Russia

[**] *amironenko@gmail.com*
Center for Institutional Analysis of Science & Education, European University at Saint Petersburg, Saint Petersburg, Russia

## Introduction

In this study we investigate the publication trajectories of medical researchers in Russia. According to Web of Science data, Russian researchers have contributed only to 0.7% of 2016 world research papers in medical and health sciences. Partly, but not entirely, this is because health sciences traditionally have not been among R&D priorities in Russia. However, even with the level of resources taken into account, the international visibility of Russian medical research seems to be low. Russian medical researchers continue to publish their results mostly in local journals, much as was the case in the Soviet era. Here we will focus on the early career publication behavior of researchers in one particular field, cardiology, to analyze their path toward internationally visible research.

Over the last few decades early career researchers (ECR) have been actively studied in different countries. There is a shared understanding that this is a difficult period of a research career. Young researchers struggle to gain jobs, resources, recognition, skills, and professional identity all at once. Issues discussed in various studies include the training of ECR; their professional socialization patterns; publication behavior; efficiency of support programs designed for young researchers, and so on [Jacob, 2007; Laudel, 2008; Besselaar, 2015; Lindahl, 2016]. Some of these studies are based on bibliometric methods, mainly on publication and citation counts. In this study, we use sequence analysis, an approach almost neglected in bibliometric research.

Sequence analysis was introduced to social sciences in the 1980s by sociologist Andrew Abbott. By the 1980s, the method was employed in natural science, mostly in bioinformatics. Abbott was a consistent proponent of using it in social sciences [Abbott, 1995]. These methods imply that development of a process is codified as a sequence of states. The sequence or set of sequences is then analyzed using various quantitative techniques. Usually the objective of sequence analysis is to classify the trajectories, to identify one or several "ideal type" scenarios, and to find the factors associated with these typical paths. Abbott's ideas that sequence analysis can provide new cognitive opportunities in addition to statistical

analysis seemed convincing to many. The approach was not just picked up by social researchers in its simplest forms, but was also elaborated [Aisenbrey, 2010].

In sociology, sequence analysis is most often used in career and life-course studies. There are also applications for the evolution of firms or whole states [Cornwell, 2015]. In this study, we apply it to publication records, which to our knowledge is new kind of object for this set of methods. One of the very few scientometric studies that used sequence analysis is the investigation of citation histories of publications [Bornmann, 2017]. We argue that considering a publication record a sequence can provide some new insights.

We are interested in identifying the patterns of early career publication behavior which could be associated with becoming an internationally recognized cardiology researcher in Russia. According to Web of Science, over the last decade the global research output in cardiology was about 20,000 papers annually. Russian researchers contribute to 1%-2% of papers in international cardiology journals per year. Most of these papers are published in low-ranked journals and do not gain many citations. This output seems far too low for a country with approximately 50 medical schools, several large cardiology centers, and several hundreds of PhDs in cardiology awarded annually.

The question we ask is, is there one or several patterns in publishing behavior at the early career stage that are typical of international-level researchers? Additionally, we are interested in the publishing behavior of researchers who start their career in different locations. Russia is characterized by a high level of R&D centralization. While Moscow and St Petersburg together employ about half of the researchers in Russia, in the rest of the country academic life is much more rarefied. There is a shared perception that Moscow and St Petersburg provide on average better opportunities for researchers, not only in terms of resources but also in terms of the professional environment. With this in mind, we are going to check if there is a difference in the patterns of internationalization between researchers beginning their career in the "center" and in the "periphery".

**Methodology**

For the analysis of publication trajectories, we selected all researchers who were awarded PhDs in medical sciences with a specialization in cardiology in Russia in 2005. In order to identify those researchers, we used the catalogue of the National Library of Russia, to which each PhD awardee is required to submit a thesis. The search for theses provided us with a list of 335 cardiology researchers. We gathered data on the papers they published in academic journals throughout their careers. We used three sources[1]:

    1) Scientific Electronic Library (https://elibrary.ru/) - online database which covers about 6000 Russian academic journals.
    2) Web of Science Core Collection (SCI-EXPANDED, SSCI, A&HCI).
    3) PhD theses indexed in the National Library of Russia.

Having checked these three sources, we discovered papers for 280 researchers. We assume that these databases allowed us to compile a rather complete journal publication record for each cardiologist. The use of online libraries is probably the best way to obtain these data because publishing CV's online still is not a wide-spread practice for Russian researchers. From PhD theses we also obtained data on the city where each thesis was completed.

---

[1] In first two databases only document types "article", "review" and ""proceedings paper" were taken into account.

In this study we will mostly (apart from one particular graph) use the so-called untimed-event recording, where it is not the time of the events that matters but only the order of the events. In order to create the untimed-event sequences the publication record of each researcher was coded with an alphabet of 4 elements:

IT – paper in an international top journal (journal from the top quartile of Journal Impact Factor ranking in any subject category of JCR 2016);

I – paper in an international non-top journal (any other non-Russian journal);

NT – paper in a national top journal (journal included to Russian Science Citation Index[2]);

N – paper in a national non-top journal (Russian journal not included to Russian Science Citation Index).

The data unit is the sequence of coded papers of a researcher, which can look as follows:
N  N  N  NT  N  N  I  NT  N  NT  I  IT  NT  IT

Further analysis was performed with TraMineR, which is a package for R (Trajectory Miner for R), developed in Institute of Demography and Socioeconomics, University of Geneva. The package offers opportunities for the visualization, description, classification and search for patterns in sequence data [Gabadinho, A., 2011].

**Results**

In total, 280 cardiologists published 3,868 papers in 493 different journals, of which 115 were international journals and 378 – national. It is important as the context of further observations that the vast majority of the papers are coauthored. The length of the publication records varies from 1 to 149, with the mean recorded at 13.8 papers, and the median – five papers. In the publication records under investigation the earliest paper was published in 1987, most recent – in 2018. About half of the researchers have journal papers published after 2006, which probably means that they were engaged in research after receiving their PhD. About 29% of researchers have papers in international journals, but only 9% have papers in international top journals.

Table 1. Characteristics of publication trajectories, by groups of researchers.

| | N of researchers | Mean number of papers per researcher | Mean number of papers per researcher, by type of paper | | | |
|---|---|---|---|---|---|---|
| | | | N | NT | I | IT |
| All researchers | 280 | 13.8 | 7.1 | 5.8 | 0.7 | 0.2 |
| Researchers from the center | 149 | 15.3 | 7.0 | 7.1 | 0.8 | 0.3 |
| Researchers from the periphery | 131 | 12.1 | 7.1 | 4.2 | 0.6 | 0.1 |
| National-level researchers | 199 | 9.8 | 5.7 | 4.2 | 0 | 0 |
| International-level researchers | 81 | 23.6 | 10.6 | 9.7 | 2.5 | 0.8 |

The group "Researchers from the center" consists of those who received their PhD in Moscow or St Petersburg, "Researchers from the periphery" are those who received PhD elsewhere, "National-level researchers" are those who have papers only in Russian journals, "International-level researchers" are those who have papers in international journals.

[2] The Russian Science Citation Index is a subset of Russian journals, consisting of "the most influential scholarly literature in Russia based on citation analysis" [https://clarivate.libguides.com/webofscienceplatform/rsci].

Table 1 contains descriptive characteristics of the publication records for the whole set of cardiologists and subgroups within the sample. One can see from the Table 1 that papers in international journals comprise less than 10% of all papers published by Russian cardiologists. Interestingly, even for international-level researchers, these papers account for less than 15% of total output. Researchers who obtained their PhD in central cities tend to publish more papers, among them those published in international journals, than researchers from other cities, but the difference is not so great.

One might expect that generally researchers start their career by publishing locally, then some of them "grow" into publication in international journals. Our data to some extent support this expectation. In the set of cardiologists under study the share of those publishing international papers slightly increases over time (Figure 1). Especially, the growth is noticeable for the recent years. Partly such dynamics could be due to recent R&D policy measures stimulating international publications. Still, the gentle slope allows one to suggest that the majority of international-level researchers started publishing in international journals at the very beginning of career.
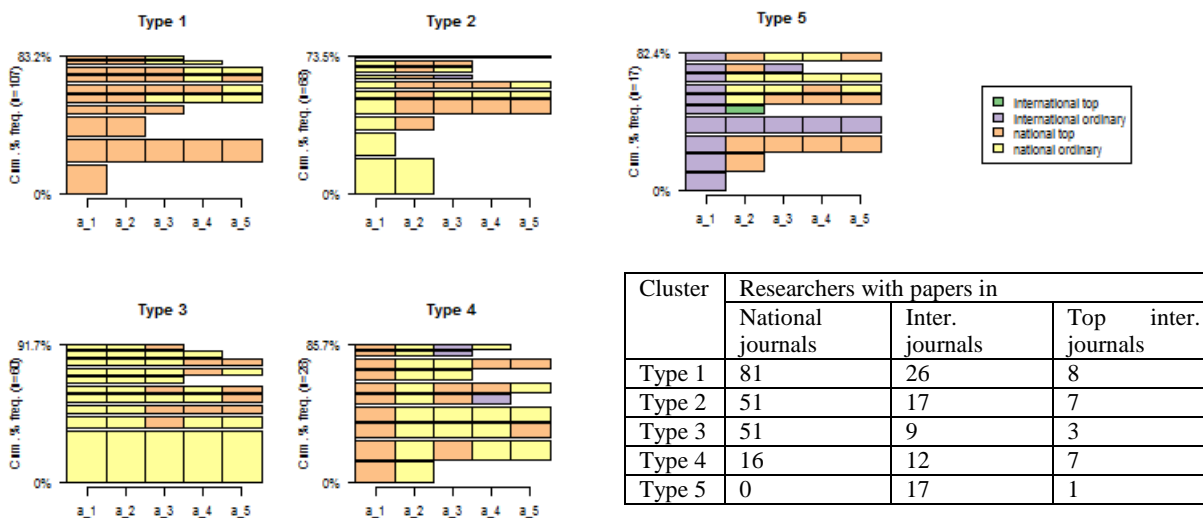
Figure 1. Frequency distribution of researchers according the category of their highest-ranked paper.
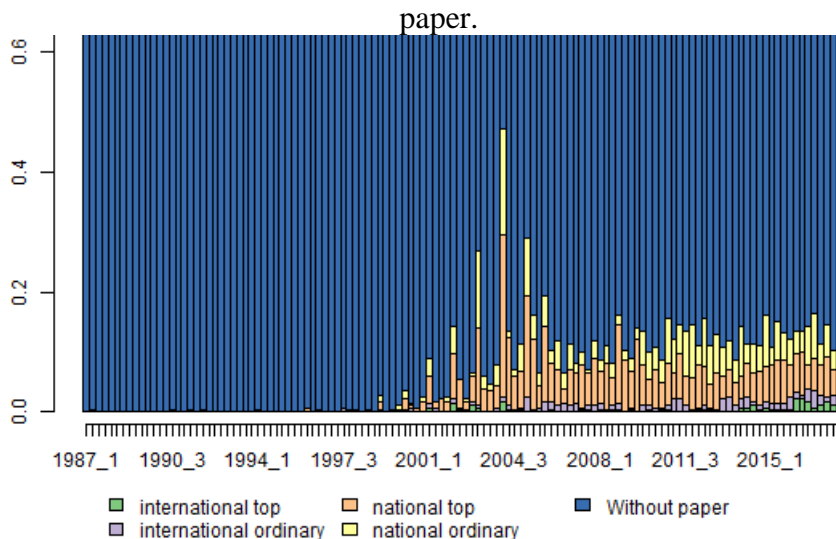


For this particular graph, which is called a chronogram, we used another mode of sequence data, state sequence over time. One time-stamp on the graph relates to one quarter of a year. The height of the green bar shows the share of researchers (% of 280) that have published papers in an international top journal in particular period, violet – those from the rest who have papers in international non-top journals, orange – those with papers published in national top journals, yellow – in national non-top journals.

In order to compare the publication trajectories at the early career stage for international-level and national-level researchers, we applied a standard sequence analysis approach which is to classify trajectories and identify their "types". We consider the first five papers published by a researcher as their early career publication record. The usual method of classifying sequences is to choose the measure of dissimilarity between sequences, to calculate the matrix of pairwise distances for the set of sequences, and then to apply a clustering algorithm to extract groups of "similar" sequences. Although the most frequently used method to assess sequence similarity is optimal matching, we used another method, called the Longest Common Prefix (LCP) [Gabadinho, 2011]. In this method, two sequences are considered more similar the longer the coincidence of their first elements is sustained. Having computed the distance matrix, we applied hierarchical clustering with the Ward method and chose five cluster solutions as the types of early career publication trajectories (Figure 2).

Figure 2.Early career publishing patterns. Ten most frequent publication sequences for five types of careers (a_i stands for the i-th paper).



| Cluster | Researchers with papers in | | |
|---------|---------------------------|---|---|
| | National journals | Inter. journals | Top inter. journals |
| Type 1 | 81 | 26 | 8 |
| Type 2 | 51 | 17 | 7 |
| Type 3 | 51 | 9 | 3 |
| Type 4 | 16 | 12 | 7 |
| Type 5 | 0 | 17 | 1 |

One can immediately see that among the five types of publication trajectories type 5 is associated with international publications. All 17 researchers in the related cluster published internationally from the very beginning of their career[3].Still, this cluster contains only a minority of international-level researchers, while the rest of them are distributed across other clusters, with their concentration particularly high in cluster 4 and notably low in cluster 3. Interestingly, when we consider early career papers for researchers who later published in international top-level journals, it is not even trajectories from cluster 5 that are typical for them, but those from cluster 4. We can conclude that, first, to reach the international level it is not necessary to start from it. About 35% of international-level researchers did not have papers in international journals among the first five. Second, we see that if a researcher starts a publication career from a top national journal and then switches to national non-top journals, it is not a sign of the "wrong path", rather quite the opposite.

Figure 3. Early career publishing patterns for international-level researchers, beginning their career in the center and periphery.

A) State distribution plot
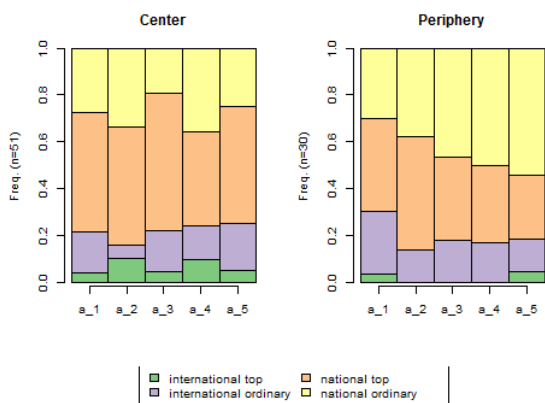
B) Publication sequences



---
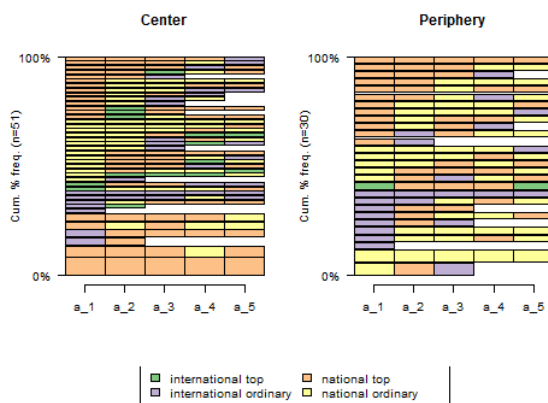
[3] It is then natural that chi-square test refutes the null hypothesis that national/international status of a researcher is independent of the type of early career trajectory.

Another question of the study was whether researchers starting their career in central cities reach international visibility by different paths compared to those starting at the periphery. We found that the chances of publishing in international journals depend on where you start. While in the whole set 53% of cardiologists are from the center, among those having international papers 63% are from the center[4], and among those with papers in top-level international journals – 85%[5] began in the center. To put figures differently, the chances of reaching the international level are higher for researchers from center (34%) than for those from periphery (23%).

We expected that researchers from Moscow and St Petersburg more often start their publication career from international papers, than regional researchers. Figures 3Aand 3B show the state distribution plot and publication sequences for the early career papers of the two subgroups of international researchers. We see that in both groups there is no predominant sequence. No matter where researcher starts his or her career, in the center or in the periphery, there are many paths that can lead to the international publications. While more than half international-level researchers began publishing internationally early, our hypothesis that this is more typical for researchers from center does not stand. Finally, we note that both groups publish most of their early career papers in national journals, but cardiologists from the center publish more often in top journals and those from the periphery publish more often in non-top journals.

**Conclusion**

The goal of this study was to describe the patterns of internationalization of Russian medical research through the lens of individual publication trajectories. We restricted our analysis to one particular field, cardiology, and traced the trajectories for a cohort of researchers awarded PhDs in 2005. Most of them were at the beginning of their academic career then. In 2005, there was less pressure on researchers in Russia to publish in academic journals, not to say international journals. Since then the professional environment has changed substantially for researchers. In the recent years, one of the priorities of R&D policy in Russia has been to increase the international visibility of research results. While for some fields one can observe progress in the integration of Russian research into international science, for medical research this progress remains limited. In case of cardiology researchers some positive dynamics of international publications can be observed. Still, it seems that the majority of them publish their research results only locally.

We were interested in finding early career patterns of becoming an international level researcher. There are fields of science where the distinction between international and local researchers is determined by the very content of the field of study. In the case of cardiology, the content is not highly localized, which is why the distinction between international and local researchers has to be social in nature. One of the important issues raised in this work is the difference in career opportunities and paths between researchers starting in the center and those beginning in the periphery of national science.

The findings of the study are the following.

---

[4] Chi-squared test p-value is about 0.05.
[5] Chi-squared test p-value is less than 0.01.

1) Among cardiology researchers that gained PhD in 2005 29% publish papers in international journals, and 9% – in top international journals. The share of such researchers publishing internationally is slightly increasing across time.

2) The majority of international-level researchers start publishing papers in international journals at the very beginning of career.

3) Researchers starting their careers not in Moscow or St Petersburg have lower chances of reaching the international level than researcher from central cities. Early career trajectories of international level researchers are diverse. Researchers from the periphery begin by publishing mostly in national non-top journals, while those from the center begin their career in national top journals.

## References

Abbott, A. (1995). Sequence analysis: new methods for old ideas. Annual review of sociology, 21(1), 93-113.

Aisenbrey, S., & Fasang, A. E. (2010). New life for old ideas: The" second wave" of sequence analysis bringing the" course" back into the life course. Sociological Methods & Research, 38(3), 420-462.

Van den Besselaar, P., & Sandstrom, U. (2015). Early career grants, performance, and careers: A study on predictive validity of grant decisions. Journal of Informetrics, 9(4), 826-838.

Bornmann, L., Adam, Y. Y., & Fred, Y. Y. (2017). Sequence analysis of annually normalized citation counts: an empirical analysis based on the characteristic scores and scales (CSS) method. Scientometrics, 113(3), 1665-1680.

Cornwell, B. (2015). Social sequence analysis: Methods and applications (Vol. 37). Cambridge University Press.

Gabadinho, A., Ritschard, G., Studer, M. & Mueller, N. S. (2011). Mining sequence data in R with the TraMineR package: A user's guide (for version 1.8). [http://traminer.unige.ch/documentation.shtml].

Jacob, J., Lehrl, S., & Henkel, A. (2007). Early recognition of high quality researchers of the German psychiatry by worldwide accessible bibliometric indicators. Scientometrics, 73(2), 117-130.

Laudel, G., & Gläser, J. (2008). From apprentice to colleague: The metamorphosis of early career researchers. Higher Education, 55(3), 387-406.

Lindahl, J., & Danell, R. (2016). The information value of early career productivity in mathematics: a ROC analysis of prediction errors in bibliometricly informed decision making. Scientometrics, 109(3), 2241-2262.