

# A Statistical Analysis of Competitive Research Contracting in the Field of Life Sciences

E. V. Poverennaya<sup>1\*</sup>, A. V. Lisitsa<sup>1</sup>, A. N. Petrov<sup>2</sup>, A. A. Makarov<sup>3</sup>, N. G. Luzgina<sup>4</sup>

<sup>1</sup>Orekhovich Institute of Biomedical Chemistry, Russian Academy of Medical Sciences

<sup>2</sup>Directorate of the CNTP at the Russian Ministry of Education and Science

<sup>3</sup>Engelhardt Institute of Molecular Biology, Russian Academy of Sciences

<sup>4</sup>Research Center of Clinical and Experimental Medicine, Siberian Branch, Russian Academy of Medical Sciences

\*E-mail: k.poverennaya@gmail.com

Copyright © 2011 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Assessment of applications in open state calls for a competitive contracting is regulated by the Federal Law 94-FZ. Ratios of different criteria in assessing applications (price, terms, and quality of proposal) are set in the legislation and make it possible, as a result of calculations using a universal formula, to assign a score for each application. Using these rating scores, winners are selected. In this paper, we have attempted to characterize the correlation between the price and quality of the research proposed by the applicants and estimate how efficient it is to use the legally set weighted values of different criteria in assessing applications.

Russian Federal R&D Program the Federal Law 94-FZ [1] prescribes that for the state as a customer, the top priority is given to the lower costs of project implementation. In the framework of this principle was used for assessment and comparison of research applications. Normally, low price and quick performance are more important than the quality; however, such approach is poorly effective for research contracting.

From December 2010 to January 2011 in the framework of the Federal R&D Program two types of research lots were opened for applications. Lots for investigative research were opened with the maximum price tag of 6 mln rubles per 2-years, whereas applied research

calls were at the price 12 mln rubles for the same period of time. Each lot enabled to conclude research contracts with several top-scoring applicants. As scorings were dependent on the price diminishing, afterwards it has been concluded that price dumping in scientific competitions has reached 60% [3].

In terms of the development of the scientific-and-technological potential of the country, the calls were first opened under condition in which, to comply with the requirements of Resolution of the Russian Government No. 722 dated September 10, 2009 [4], the specific weight of the price in the gross assessment of the application comprised 55%. It should be noted that all the research contract calls were announced in a relatively short pe-

riod of time, i.e. the applicants had limited opportunity to alter their tactics in the competition.

We have analyzed the following data taken from the protocols of assessment and comparison of research proposals [5]:

- Maximum lot price;
- Name of the applicant;
- Expert assessment of application by criterion “scientific and technical level of research” (maximum 35 scores);
- Expert assessment by “professional reputation of the applicant” (maximum 40 scores);
- Expert assessment by “prior R&D experience” (maximum 25 scores);
- and
- Project implementation cost, demanded by the applicant.

A total of 41 lots were analyzed. In the contest for these lots, 183 scientific organizations took part, and each organization filed at least one application. A total of 421 applications were filed and included in this analysis. The results of the calls were taken from official protocols published by the Ministry of Education and Science of the Russian Federation. The price demanded by the participant was normalized to the maximum price of the lot, and the obtained value of the *price-drop* was expressed in percentage

points. Given the fact that expert assessments based on the criteria “scientific and technical level of research,” “professional reputation of the,” and “prior R&D experience” correlated with each other (the correlation coefficient exceeded than 0.73), the sum of the above-said criteria was considered the integral measure of assessment, which was further designated as *application quality*.

To calculate the statistical correction, we analyzed the dependence between the assessment of the application quality  $V$  and the price-drop  $C$ . To reveal this dependence, applications were selected, which came out of the 49 (from 183 organizations-participants) most active participants, i.e. those that filed at least three applications for different lots, and at least one application from those filed was a winner.

In the case if three or more applications were filed by one participant, we averaged the assessment of the application quality and price-drop, and, in accordance with the obtained means  $V$  and  $C$ , a point was plotted on the graph. If several data points (from different participants) corresponded to one value of price-drop, we excluded, before the averaging, outliers using the t-test ( $p \geq 0.95$ ). The obtained scatter plot data was approximated using the linear equation  $V = aC + b$ , where the values of constants  $a$  and  $b$  were obtained with the help of the least square method (built-in tool of Microsoft Excel).

The value of constant  $a$ , defining the slope angle of the approximation linear equation curve, was used as a statistical correction, given the expression (1):

$$a = \frac{dV}{dC} = \frac{V - V'}{C' - C}, \quad (1)$$

where values  $V$  and  $C$  were obtained from averaged data presented in official protocols, and  $V'$  and  $C'$  were values corrected under

the assumption that the research quality must not be affected by the price. Assuming that the minimum value of the reduction of the price, not leading to a decrease in quality, comprises  $C' = 0$ , in order to calculate the *adjusted assessment of the application quality*  $V'$ , expression (1) was transformed into:

$$V' = V - |a|C. \quad (2)$$

(The module is used as the coefficient  $a$  is negative due to the reciprocal dependency between the application quality and price-drop.)

Using formula (2), we carried out a retrospective calculation of the adjusted values of assessments of the scientific quality for 421 applications. For each contest lot, the final score of applications was calculated using the formula listed in the official call documentation:

$$R = 0.45V' + 0.55C. \quad (3)$$

The lots were sorted (ranked) by descending degree of the final score  $R$ .

In order to compare the ratings of applications, the applications were sorted according to the final score, which accounts for the estimations of application quality. A series of model ratings were generated by changing the weight of the application quality from 0.45 to 1.00 with a step of 0.05. Modeled ratings were compared with the expert rating, obtained by formula (3). In agreement with different weights of the quality criterion, we changed the specific weight of the cost so that the sum of both criteria equaled one.

In order to compare the artificially generated model and expert ratings, we used the Spearman coefficient of rank correlation (CRC):

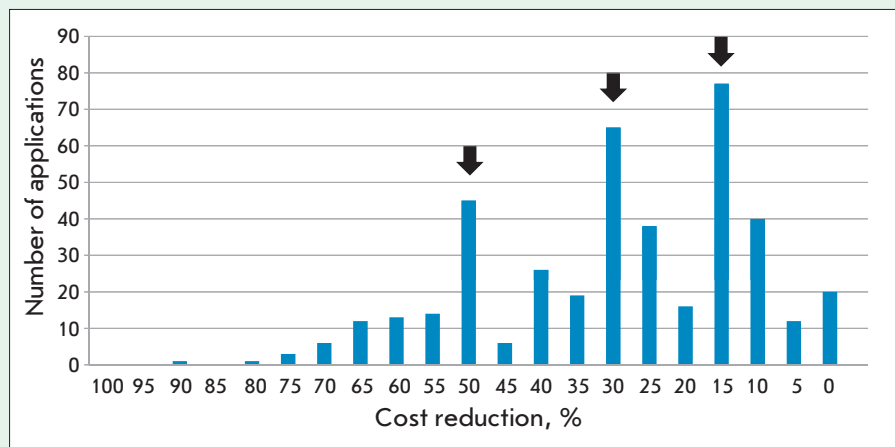
$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}, \quad (4)$$

where  $d$  is the difference between

the ranks for each application in the lot and  $n$  is the number of applications in the lot. CRC was computed for each lot separately; thereat, five lots were excluded from the analysis as they contained less than five applications. For the remaining 36 lots, we assessed the reliability of CRC using the critical value  $r_s$  at a given number of applications per lot. We selected the values of rank correlation which were reliable at a significance level  $p \geq 0.95$ , and the selected CRC values were averaged for all lots. The mean CRC values were put on the graph in accordance with the specific weights of the criterion “application quality,” for which ratings were computed. On the graph, using the least square method, we found two approximating equations, corresponding to linear sections of the CRC dependency. Equations were solved to define the optimal weight of the “application quality” criterion, exceeding which CRC did not significantly increase at the increase of the specific weight.

The data, presented in the protocols of assessment and comparison of applications, are summarized in *Fig. 1* in the form of a histogram indicating the number of filed applications depending on the price-drop intervals. Less than 5% of the applications (21 out of 421 applications) entered the contest with zero price-drop, i.e. the cost of the project equaled the initial (maximum) cost of the call. The mean price-drop was 30% for all applications surveyed.

As is seen from the histogram (*Fig. 1*), the distribution of applications has three clear peaks corresponding to price-drops of 15, 30, and 50%. The presence of these peaks allows for a discussion of the general considerations underlying a participant’s decision on the cost of their work. It can be presumed that a fall in the cost of up to 15%, which was the case in 155 applications, is



**Fig. 1.** Distribution of research applications depending on the degree of price-drop. Peaks are marked by arrows.

not critical for the research quality and can be achieved by a reduction in overhead expenses (coverage of overhead expenses from the applicant's internal resources). A 30% price reduction may point at sufficient applicant groundwork, e.g., a stock of consumables or reagents, readily available for the project implementation. A total of 119 applications (i.e. 28% of the total number) fell in the range of costs reductions of 15 to 30%. The next category includes 96 applications in which the cost reduction ranged from 30–50%. In this case, it could be suggested that the fall in costs will adversely affect the quality of the project implementation. Finally, in the histogram, a group is seen, with costs falling by more than 50%. This group includes 51 applications, which number is twice less than for the previously described categories of pricing. Lower number indicates that “raider” tactics drops out of the general concept pursued by the applicants when estimating the cost of their research work.

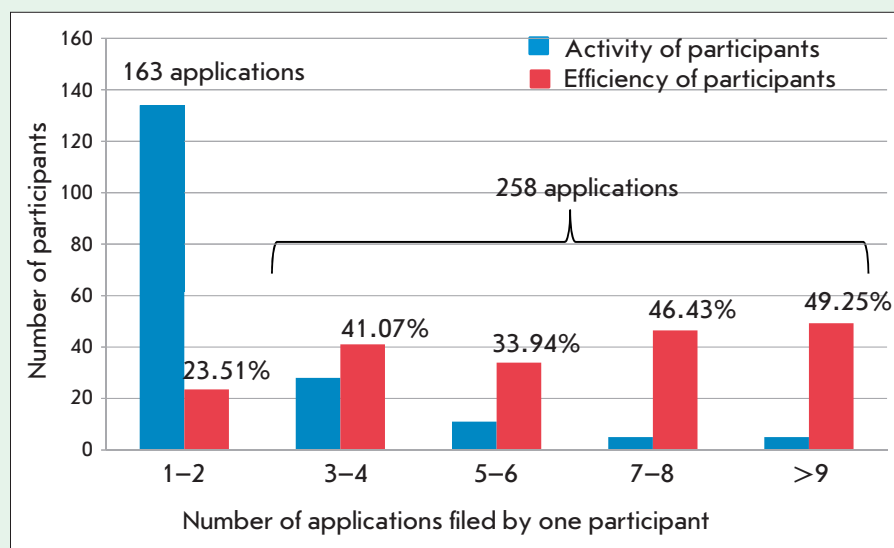
An analysis of the data published in the protocols of the research calls allows one to estimate the activity and efficiency of the participants. Most applications were filed by low-activity participants who put in most cases one, and sometimes,

two applications. As is seen from the histogram in *Fig. 2*, there were 132 such entities, i.e. substantially over half of the applications. The data obtained for one or two applications are not sufficient to draw statistical generalizations. Besides, as can be seen from *Fig. 2*, low-activity participants rarely win. Probably, sole applications are filed by institutions whose profile of activity is not fully compliant with general trends in

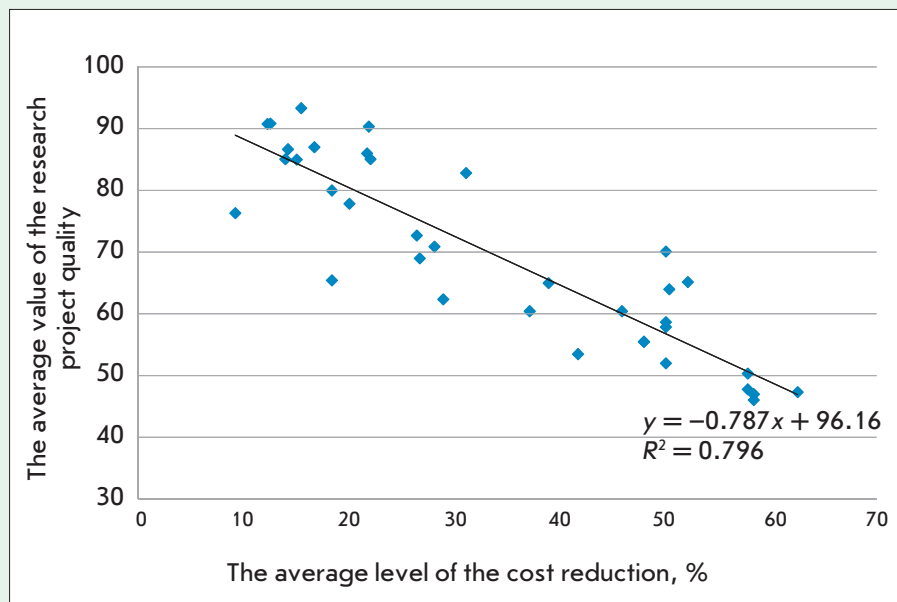
life science. Adaptation to a limited number of specific tasks generally lowers the competitiveness, which is reflected in the low efficiency of the participants who filed only one application – 23,5%.

High efficiency is typical for the participants who put in three or more applications (*Fig. 2*). For the participants who filed three to four applications, the efficiency exceeded 40%, i.e. nearly half of the applications awarded a contract. Hence, one can conclude that the work of the institution in several directions increases competitiveness and, most likely, such participants exert the greatest influence on the development of life sciences in Russia. Given the above considerations, we further analyzed the data on the quality of the applications of the participants who filed three and more applications. The number of such participants was 49.

Before we proceed to the analysis of the applications quality, we should note the tight interconnection between the categories of expert quality scorings. For instance,



**Fig. 2.** Distribution of values of activity and efficiency of participants. Activity is calculated as the number of applications filed by the participant; efficiency, the ratio of the number of “winning” applications to the total number of applications filed by the participant.



**Fig. 3.** The linear approximation of the dependence of the experts' quality assessment (V) and the price-drop in the application (C). Each point corresponds to one participant, and respective values of the price-drop and application quality are averaged by the winning applications.

the level of the research correlates well with the professional reputation of the applicant ( $R^2 = 0.74$ ;  $n = 421$ ), while the reputation, in turn, depends on experience ( $R^2 = 0.73$ ;  $n = 421$ ). Such a correlation is quite regular, which further permitted to use the sum of the scores in all three expert categories as the score of the quality of the research project submitted for the call.

Figure 3 shows data demonstrating the correlation between the application quality of the projects and reduction in the demanded price. The dependence has been obtained for 105 winning applications, which came up from the selected 49 active participants. If quality assessments coincided in several applications, the relevant values of cost reduction for these applications were averaged. Therefore, for example, the lowest reduction in cost according to the data in Fig. 3 amounts to 9% (not 0%, as follows from the histogram in Fig. 1).

Approximation of the data by the

linear dependence, in agreement with formula (2), allowed us to reveal the value of the absolute term  $V = 96$  and adjustment multiplier  $a = -0.787$ , and include these terms for the calculation of the adjusted assessment of the application quality  $V'$ :

$$V' = 96 - 0.787 \times C. \quad (5)$$

The absolute term in the formula defines the limit for the score assessment of the research quality. The value obtained on the basis of the approximation results almost coincides with the maximum of the scores 100, set in the call official documentation. Substituting the values shown in Fig. 3 in formula (5), one can find the point  $V = 47$ ;  $C = 60\%$ , at which the adjusted assessment of the research quality equals zero. This means that for the applications whose price is reduced by 60% and more, the risk of poor quality of results is so high that, in the expert opinion, it is impos-

sible in principle to implement the project.

The adjusted assessment of the research quality, obtained as a result of approximation, was substituted in formula (3) and a new total score was calculated for each lot, which was the weighted sum total of the indicators "application quality" and "price-drop." Applications filed in each lot were rearranged in accordance with the values of the total score. The obtained ranking of the applications was considered an expert rating, i.e. the rating in which the order of applications encapsulates the collective opinion of the experts about the link between the application's cost and its quality.

The expert rating was then compared with the model ratings obtained at different values of the weights of the cost and quality criteria, e.g., the values 0.4 and 0.6 were taken for the weights, respectively. The weights were used to compute the new value of the final score assessment for all applications, after which the applications were positioned within each lot in descending order. Such a model rating was compared with the expert rating previously obtained with the use of the adjusted assessment. The comparison was carried out with the use of the Spearman rank correlation coefficient (CRC), which is sensitive to the order of elements in sorted lists. In the described example case, when the weight of the quality criterion was set at 0.6, the value of CRC was 0.865. This value indicates that the order of applications in the expert rating matches well the model rating with the preset quality-to-cost ration 0.6/0.4/

On opposite case, when, for instance quality-to-cost was 0.4/0.6, the CRC value equaled only 0.75.

In Fig. 4, values of CRC are shown for 11 model ratings obtained at different quality-to-cost ratios. Along the ordinate axis the

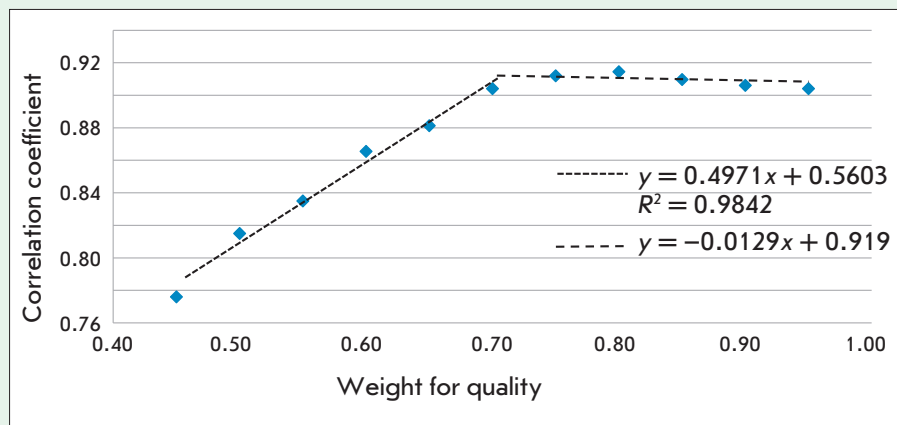


Fig. 4. Dependence of rank correlation of the expert rating and model ratings, obtained at different weights of the criterion “application quality.”

weights of the application quality used to generate the model ratings are plotted, and along the abscissa axis, values of CRC obtained as a result of comparing model ratings with the expert rating. It is seen from Fig. 4 that there is a clear dependence between the CRC and the weight of the quality criteria. The increase in the value of the correlation coefficient continues up to the level of values of the quality criterion of 0.7, at which CRC goes beyond 0.9. Further increase in the weight

of the quality criterion did not affect in any significant manner the CRC value.

For purposes of illustration, Fig. 4 shows approximation lines crossing at a point with the abscissa coordinate 0.703 (3). The slope of the straight line, approximating the plateau on the graph, may vary depending on the number of points used to draw the equation. The dynamics of the CRC increase changes up to the level of 0.7; therefore, it is purposeful to take this value for

the minimum weight at which optimal coincidence of the model and expert ratings is achieved.

## CONCLUSIONS

There exists a reciprocal dependency between the quality of the research project and reduction in its cost. If the price of the project drops by more than 60% against the initial estimation of the call budget the proper implementation of research becomes impossible. The legally established weights for the price and quality are currently inappropriate for the competitive research contracting, as these weights almost neglect the expert opinion while comparing the applications merely by the price. Using heuristic modeling, we have shown that optimal ranking of research applications can be achieved by using the quality to-cost ratio 0.7/0.3. ●

*The authors are grateful to the members of the Live Systems Workgroup and K. Timirbaev (the MetaSyntez Corporation) for participation in the discussion of the paper.*

## REFERENCES

1. On the placement of orders for the supplies of goods, execution of works, and rendering of services for state and municipal needs: Federal Law dated July 21, 2005 No. 94 – FZ. The Code of Laws of the Russian Federation. 2005. No. 30. Clause 3105.
2. The Federal Target Program “Research and development in the priority directions of the development of the Russian scientific-and-technological complex for 2007–2012”: Resolution of the Government of the Russian Federation dated July 6, 2006 No. 977-r. The Code of Laws of the Russian Federation. 2006. No. 29. Clause 3258.
3. “The system of state purchases in the sphere of science and education requires radical changes”. Polit.ru. URL: <http://www.polit.ru/news/2011/08/05/sg/>
4. On the approval of the Rules of assessing applications for participation in the competition for the right to sign the state or municipal contract for the supplies of goods, execution of works, and rendering of services for state or municipal needs: Resolution of the Government of the Russian Federation dated September 10, 2009 No. 722. The Code of Laws of the Russian Federation. 2009. No. 38. Clause 4477.
5. Arrangement 1.2. “Carrying out problem-oriented investigative studies and creating the scientific-and-technical groundwork in technologies in the field of live systems”. Federal Target Program “Research and development in the priority directions of the development of the Russian scientific-and-technological complex for 2007–2013.” URL: <http://www.fcpir.ru/catalog.aspx?CatalogId=1044>