2014 EVALUATION OF THE PROVOCATIVE QUESTIONS INITIATIVE (2011 AND 2012 PQ RFAS)

EXECUTIVE SUMMARY

http://provocativequestions.nci.nih.gov/

May 2014

Center for Strategic Scientific Initiatives
National Cancer Institute
National Institutes of Health

Prepared by:
Thomson Reuters IP & Science
Rockville MD, 20850
www.thomsonreuters.com

EXECUTIVE SUMMARY

PROGRAM OVERVIEW & EVALUATION GOALS

Initiated in fiscal year (FY) 2011, the National Cancer Institute's (NCI's) Provocative Questions (PQ) Initiative provides support for cancer research addressing important questions that are broadly considered challenging or understudied. The PQ Initiative complements the NCI's broader funding portfolio with a more flexible Funding Opportunity Announcement (FOA) design, a focus on asking difficult questions, and an effort to solicit new approaches from diverse scientific disciplines.

In its first two issuances, the PQ Initiative funded 9.7% of applications (149/1,531), for a total of $60.8 million in new awards. R01s comprise a majority of grant applications (58%) and awards (64%) to date (Table 1).

Table 1. Applications, awards and funding percentages for the PQ Initiative Issuances 1 and 2

	R01			R21			All: R01 & R21			
Issuance	Applications	Awards	Total Cost	Applications	Awards	Total Cost	Applications	Awards	% Awarded	Total Cost
1	422	38	$17.9M	332	18	$3.6M	754	56	7.4%	$21.5M
2	460	58	$31.8M	317	35	$7.5M	777	93	12.0%	$39.2M
Total	882	96	$49.7M	649	53	$11.1M	1,531	149	9.7%	$60.8M

The NCI's Center for Strategic Scientific Initiatives (CSSI), has carried out an evaluation of the 2011 and 2012 PQ RFAs, with a focus on quantitative indicators of early progress. In addition to descriptive analyses of the research portfolio and volume of early publications from the portfolio, the following programmatic goals were evaluated:

1. Has there been an increase in the volume of research publications and grants within the targeted PQ research areas that corresponds with the launch of the initiative?
   (Findings #1-3)
2. Is the PQ Initiative attracting new ideas within the specific question areas?
   (Finding #4)
3. Has the PQ Initiative been effective at attracting and retaining new investigators to the NCI and NIH? (Findings #5-6)

The following executive summary provides the highlights of that evaluation. A full report is available upon request.

KEY FINDINGS

1. Direct Impact on Volume of Publications. Based on the early publication record of the first issuance (2011 RFAs) two question areas have shown an early sign of productivity in publications. (#1 Obesity & Cancer and #21 Therapy Resistance)

   The timing of this early evaluation prohibits a full bibliometric analysis because most grantees have not yet written and submitted publications. Still, there are early signs of trends within the publication record from the first RFA issuance. PQs #1 and #21 lead in the total number of publications, as well as the normalized count of publications per grant (Figure 1).

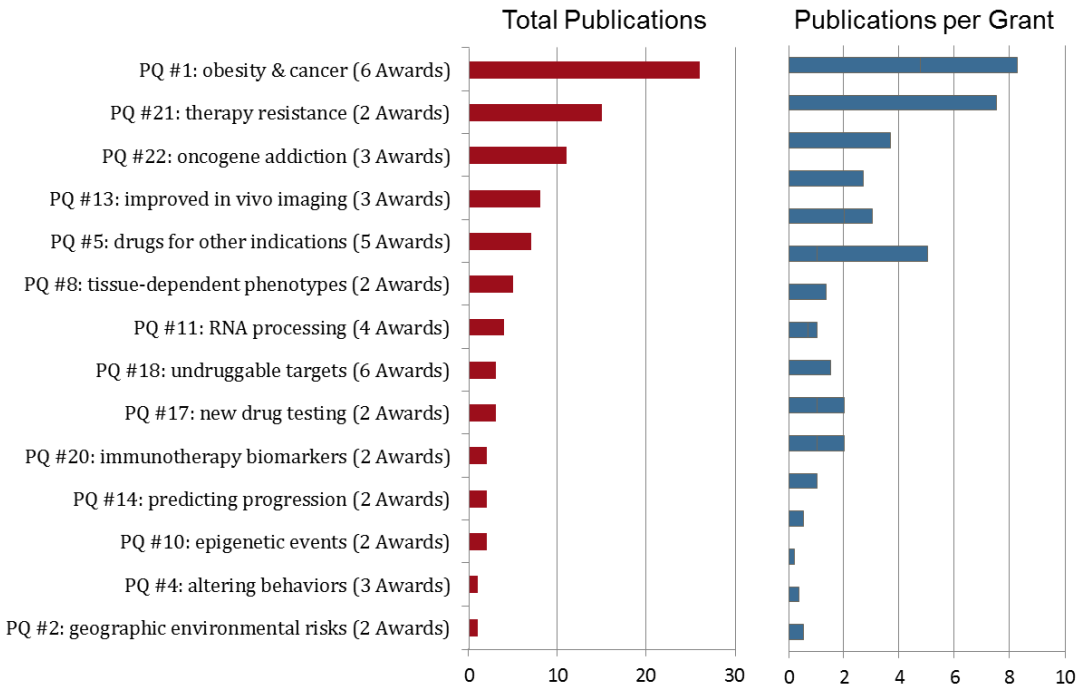

Figure 1. Total publication counts by PQ (left) and average publication counts per funded grant (right), excluding PQs that have not yet resulted in publications

2. Direct Impact on Volume of Grants (Direct Relevance): 33% of PQ grant applications failed to meet relevance criteria, potentially highlighting differences in the specificity of the questions posed.

Anomalies of low relevance measures for five question areas may be due to the open-ended nature of the questions. Further review is needed to assess the low relevance scores for the following questions:

- PQ# 3 and A4- measuring risk exposure
- PQ #4 and A3- altering behaviors / cognitive processes for behavior change
- PQ #20- immunotherapy biomarkers

Methods. This analysis was intended to develop a quick and early indication of the relevance of the PQ grant applications received to date. The results inform program staff, allowing them to adjust the process to increase the rate of relevant applications if manual review finds that necessary. We examined relevance using the text similarity approach described in subsequent sections (see "Compelling New Research Ideas: Methods"). Here, we compared the text of PQ applications found in the Specific Aims section to the text of the RFA (background, feasibility, implications for success). The tool used identifies general text similarity (similar to that used in NIH reporting websites). In some cases, the wording of the RFA might lead to grants that are identified by the tool as less relevant than would be determined by a manual review. To reduce this bias, we augmented the reference text (RFA text) with titles and abstracts of six relevant publications per PQ that were reviewed by subject matter experts. The final relevance score corresponds to the maximum score found.

Caveats. Relevance to the selected PQ and overall grant quality are measured through the peer review and funding plan development processes. This automated approach is not intended to replace manual review. Results from this analysis are likely indicators of both the degree of specificity of the question in addition to the relevance of the grant application (The caveats listed under "Compelling New Research Ideas" also apply here.)

3. Change in Volume of Research Publications and Grants (Indirect Relevance):

(a) Overall, there have been small increases in the volume of research related to PQ question areas when comparing the pre- and post-PQ years.[1]

a. When grouping together all PQ question areas (including Issuance 1 and 2), there was an increase (5.2%) in the proportion of cancer-related publications and a larger corresponding increase in the proportion of authors (11.2%) associated with those publications. Both differences were statistically significant ($p<0.0001$).

b. While there was an increase (6.5%) in the proportion of relevant grant applications (excluding applications to the PQ Initiative), the change varied significantly by question area.

Methods. We measured the volume of relevant research through a keyword-based literature and grant search that mimics a manual approach taken to address this question. Articles were considered relevant if they met the following criteria: (1) they used keywords indicating cancer relevance; (2) at least one term from two separate lists was present to ensure that different attributes of the PQ question area were generally addressed; and (3) a composite score based on keyword match statistics was found to be above a threshold set by a manual review of over 200 sample results.

Figure 2. Illustration of keyword approach to identify relevant publications.

Caveats. The success of this method in accurately identifying and ranking publications and grants relevant to a given PQ is variable by question; questions with a very specific nuance in an area that is generally well studied are challenging for an automated method to extract the small subset of relevant publications that address this specific nuance. After subject matter expert review of the initial results, further refinement to the keyword lists might improve the recall of this method. Also, because the scientific topics covered by the PQs are generally understudied areas, for some questions there is only a small body of literature that is truly relevant to the question, making automated recall of these publications more difficult.

(b) Four PQs stand out due to significant increases in publication volume in the post-PQs years and two stand out as showing significant decreases:

Increases: #1-obesity and cancer, #4- altering behaviors, #16- metastases clinical significance, #24- metastasis study techniques

Decreases: #11-RNA processing, #15-second primary cancers

(c) It is too early to assess the quality of the research findings, but some published research in high visibility journals may signal early indications of progress.

4. Attracting New Research Ideas:

(a) Roughly one-half of the PQ applications submitted to both the 2011 and 2012 RFAs were judged to be novel[2], with one-third of the PQ applications showing strong similarity to a given investigator's prior NIH grants.

- 47% of the PQ applications were judged to be novel relative to prior NIH grant applications. The distribution of novelty scores was on par with the trends seen in the comparison groups.[3]
- 32% of the PQ applications resembled the Principal Investigator's (PI's) own prior work.
- 21% of the PQ applications were similar to prior grant applications by other investigators.

[1] Pre- and post-PQ years were defined as the three calendar years immediately preceding the initiative (2008-2010) and the 3½ years following the launch of the initiative (2011- March 2014).
[2] See 3(a) Methods section for local definition of "novel" used within this evaluation.
[3] Definitions for comparison initiatives are provided in **Appendix A**.

Methods: We examined whether grant applications were unique by comparing them to a corpus of prior grant applications submitted to the NIH, including prior applications by any PI in the PQ cohort and comparison initiatives used for this analysis. It was not feasible to compare every PQ application to all other NIH grant applications. We therefore used an automated approach that measured the similarity of the Specific Aims section based on several factors. These included the number of terms in common between each pair of applications relative to the frequency of those terms within NIH grants generally. We selected a subset of grant applications (97,000) from the early 2000s through 2013 with a bias toward applications to NCI (the "companion cohort"). We then deemed "novel" those applications for which we found no prior similar applications within the companion cohort (maximum similarity below a threshold selected by blind subject matter expert review).

Caveats: This automated approach is generally found to be effective in identifying cases of very similar and very distinct text, but is not expected to provide insight into subtle differences in the language that might differentiate scientific approaches. The method is intended as a guide for further manual review of extreme cases.

Table 2. Statistically significant novelty effects.

Higher (+)		Lower (-)	
PQ	précis	PQ	précis
A3	cognitive processes for behavior change	18	undruggable targets
A4	measuring risk exposure	B3	immune response
B5	mutation/epigenetic change order	9	driver mutations
C5	improved in vivo imaging	10	epigenetic events
D6	cachexia	17	new drug testing
2	geographic environmental risks	23	tumor indolence
4	altering behaviors	24	metastasis study techniques
11	RNA processing		

(b) Differences in novelty scores were observed based on the question to which the PQ application responded, suggesting that some question areas are more likely to compel new research ideas relative to others.

- Applications submitted to PQ Question Groups A (cancer prevention and risk) and D (cancer therapy and outcomes) were found to be more novel in comparison with those submitted to B (mechanisms of tumor development or recurrence) and C (tumor detection, diagnosis and prognosis).

5. Attracting New Researchers.

(a) While the PQs Initiative ranks in the ~43rd percentile of NCI FOAs in attracting New Investigators using the formal definition established by NIH, it ranks in the ~54th percentile in attracting researchers with no prior NCI applications (New to NCI) and the ~58th percentile in attracting researchers with no prior NIH applications (New to NIH).

- The formal definition of new investigators allows for researchers to have been funded by select smaller grant mechanisms. The median percentage of New Investigators for PQs RFAs in the first two issuances is 50% (R21s) and 31% (R01s). Compared to the 2010 R01/R21 NCI FOAs, these percentages correspond to the 28th and 47th percentiles, respectively.
- While lower in percentage, the PQs rate of attracting new researchers to NCI of 19% (R21) and 15% (R01) places the PQs in the 56th and 54th percentile, respectively.
- Still lower percentages of new researchers attracted to NIH, 8% (R21) and 6% (R01) places the PQs in the 52nd and 63rd percentiles, respectively.

(b) The R21 mechanism has been more effective than the R01 in attracting new researchers by all measures used (New Investigator, New to NCI and New to NIH), despite a reduced requirement for preliminary data for PQ R01 grant applications.

6. Early Retention of New Researchers. **New Investigator awardees of the PQ Initiative have been active in applying for subsequent NIH RPGs, but differences relative to comparison initiatives were not found to be statistically significant. This is true for New Investigator applicants as well.**

- New investigators who applied for PQ grants subsequently applied for RPG grants at a higher rate than comparison initiatives (79% compared to 61%), but have been awarded at a similar rate (12% compared to 15%).
- New Investigators who received PQ awards have subsequently been more active in submitting NIH grant applications (74% compared to 70%) and more successful in receiving NIH awards (26% compared to 20%).

CONCLUSIONS

General

- Across a number of initiative goal areas, PQ applications and investigators appear to be performing at the level of comparison groups with similar innovative goals. Differences seen to-date are small, but promising.
- This early evaluation effort indicates that R21s may have favorable outputs, but more time is required to measure long term impact of R01s (with larger funding amounts). In terms of relevance of applicant text, and uniqueness of application text, no statistically significant difference has been measured when comparing R01s and R21s.

Direct Relevance

- More time must elapse prior to carrying out rigorous quantitative bibliometric analyses (at least two years of data following funding is suggested before formal comparisons can be made).
- 30% of grant applications found not to meet "relevance" criteria, potentially highlighting differences in specificity of questions posed.

Indirect Relevance

- 5.2% and 6.5% increase in the proportion of cancer-related research that focused on PQ question areas post-PQ launch, as measured by publication and grant activity, respectively.
- ~11% increase found in proportion of researchers authoring publications in PQ question areas post-PQ launch.

Compelling New Research Ideas

- This early evaluation suggests that approximately 50% of the applications that responded to the PQs Initiative were not substantially different from work that was previously proposed to the NIH. This trend was the same for the comparison initiatives evaluated here.

Attracting New Researchers and Early Retention

- These early results suggest that the initiative has been an effective tool in attracting high-quality researchers to the NIH who continue to seek and obtain NIH research funding.
    - PQ Initiative has been effective thus far in attracting new researchers to the NCI and to NIH broadly.
    - PQ Initiative has granted awards to New Investigators who subsequently apply for and receive RPGs at a relatively high rate (26%).
- In spite of the reduced requirement for preliminary data in R01s, PQs still saw greater proportions of new researchers applying to R21s than to R01s.

RECOMMENDATIONS

- To obtain a broader sense of the impact of the program, we recommend interviewing applicants, and particularly focusing on some of the new investigators who have subsequently sought NIH funding.

- Scientific responsiveness determinations made by program staff prior to peer review to withdraw applications that are not relevant to the selected PQ.

APPENDIX A: STUDY DESIGN

Illustrated in Figure 3 is a schematic of the semi-automated approaches used to give early indicators of the degree to which PQ applications are "novel" relative to prior NIH applications and the change in the volume of research during the pre- and post-PQ time periods.

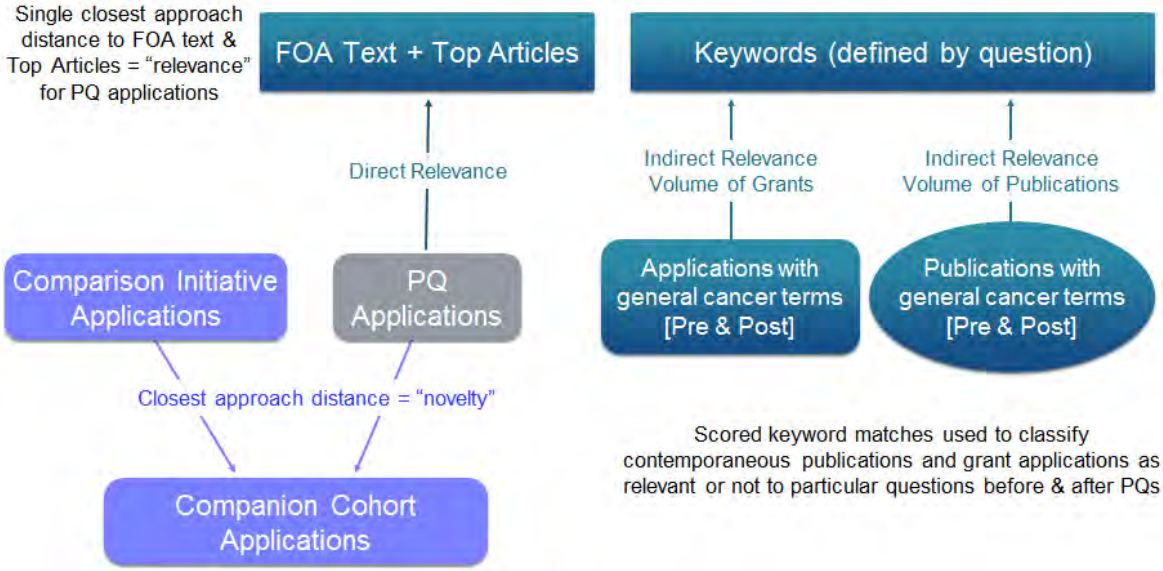

Figure 3. Schematic of comparisons used to estimate "novelty" (purple) and "relevance" (green).

Comparison Groups. Comparison groups were used to provide context when applicable. The comparator initiatives were selected based on matched qualitative criteria, including: (1) a focus on understudied problems; (2) targeting diverse research backgrounds; (3) issuing RFAs with diverse research topics; and (4) reducing the requirement for preliminary data. Based on these criteria, the following comparison initiatives were selected:

1. New Innovator (NIH Directors New Innovator Award Program)
2. Transformative (NIH Director's Transformative Research Awards)
3. Eureka (Exceptional, Unconventional Research Enabling Knowledge Acceleration)
4. Pioneer (NIH Director's Pioneer Award Program)
5. High-Impact (High-Impact Studies on Cancer Biology)
6. IMAT (Innovative Technologies for the Molecular Analysis of Cancer Phased Innovation Award)
7. 2010 NCI RFAs (R01/R21): A subset of NCI RFAs representing an average performance of initiatives within the same funding mechanisms used for the PQs. The dataset was defined as NCI averages for RFAs with R01 and R21 mechanisms calculated for competing awards only. (PA07-070 and RFAs with fewer than five applicants were excluded.)

A subset of each comparison group was selected to match on characteristics of the PQs, including fiscal years and, where possible, activity codes.

APPENDIX A: PQ APPLICATIONS AND AWARDS

Table 3. Success rate and funding amounts for the first two issuances of the PQs RFAs

Mechanism	Issuance 1		Issuance 2
R01	Applications	422	460
	RFAs	RFA-CA-11-011	RFA-CA-12-015(A),RFA-CA-12-017(B) RFA-CA-12-019(C),RFA-CA-12-021(D)
	Awards	38	58
	% Awarded	9%	13%

	Total cost	$17,866,028	$31,841,188
	Average cost per award	$470,159	$548,986
R21	Applications	332	317
	RFAs	RFA-CA-11-012	RFA-CA-12-016(A),RFA-CA-12-018(B) RFA-CA-12-020(C),RFA-CA-12-022(D)
	Awards	18	35
	% Awarded	5%	11%
	Total cost	$3,651,121	$7,450,788
	Average cost per award	$202,840	$212,880
All: R01 & R21	Applications	754	777
	RFAs	2	8
	Awards	56	93
	% Awarded	7%	12%
	Total cost	$21,517,149	$39,291,976
	Average cost per award	$384,235	$422,494

APPENDIX B: PROVOCATIVE QUESTIONS

Provocative Questions for Issuance 1					
PQ	Précis	Succeeded by PQ	Applications	Awards	% awarded
1	obesity & cancer	A2	84	6	7%
2	geographic environmental risks		15	2	13%
3	measuring risk exposure	A4	12	1	8%
4	altering behaviors	A3	15	3	20%
5	drugs for other indications	A1	67	5	7%
6	disease correlation		31	1	3%
7	age dependence	B4	19	1	5%
8	tissue-dependent phenotypes		19	2	11%
9	driver mutations		31	1	3%
10	epigenetic events	B2	27	2	7%
11	RNA processing		50	4	8%
12	novel infectious agents		28	5	18%
13	improved in vivo imaging	C5	22	3	14%
14	predicting progression	C3	50	2	4%
15	second primary cancers	B1	8	0	0%
16	metastases clinical significance	C4	9	0	0%
17	new drug testing	D5	32	2	6%
18	undruggable targets		69	6	9%
19	chemotherapy sensitivity	D2	9	0	0%
20	immunotherapy biomarkers		31	2	6%
21	therapy resistance	D1	42	2	5%
22	oncogene addiction		24	3	13%
23	tumor indolence	C1	23	0	0%
24	metastasis study techniques	B6	37	3	8%
Provocative Questions for Issuance 2					
PQ	Précis	Preceded by PQ	Applications	Awards	% awarded
A1	drugs for other indications	5	84	10	12%
A2	obesity & cancer	1	67	11	16%
A3	cognitive processes for behavior change	4	33	6	18%
A4	measuring risk exposure	3	11	3	27%
A5	physical activity & cancer		13	2	15%
A6	susceptibility during development		10	0	0%

B1	second primary cancers	15	16	0	0%
B2	epigenetic events	10	23	1	4%
B3	immune response		50	4	8%
B4	aging & cancer	7	37	3	8%
B5	mutation/epigenetic change order		28	4	14%
B6	metastasis study techniques	24	45	4	9%
C1	tumor indolence	23	18	0	0%
C2	physical properties		48	4	8%
C3	predicting progression	14	51	4	8%
C4	metastases clinical significance	16	12	1	8%
C5	improved in vivo imaging	13	19	5	26%
C6	dormancy and recurrence		32	2	6%
D1	evolution of drug resistance	21	36	6	17%
D2	chemotherapy sensitivity	19	26	3	12%
D3	long survivors		12	3	25%
D4	cancer field effect		18	3	17%
D5	new drug testing	17	62	9	15%
D6	cachexia		26	5	19%