IDA | SCIENCE AND TECHNOLOGY POLICY INSTITUTE

March 24, 2021

MEMORANDUM

To: Ravi Basavappa and Rebecca Miller
Office of Strategic Coordination, National Institutes of Health

From: Sally S. Tinkle, Xueying "Shirley" Han, Lara L. Rubinyi, Erick D. Cohen
Science and Technology Policy Institute (STPI)

Through: Kristen M. Kulinowski
Director, STPI

Subject: Transformative Research Award Evaluation

The National Institutes of Health Office of Strategic Coordination (NIH/OSC) supports high-risk, high-reward research through targeted research programs, one of which is the NIH Director's Transformative Research Award (TRA). This award supports individuals or teams proposing transformative projects that are inherently risky and untested but have the potential to create or overturn fundamental scientific paradigms. NIH/OSC asked STPI to evaluate research outputs from the 2010–2012 TRA awardee cohorts and determine whether the research is more transformative, innovative, and impactful than research produced by comparison groups. STPI developed and implemented a multi-modal evaluation strategy that is the basis for recommendations to inform future NIH Director policy decisions. The final report contains details of methods, results, conclusions, and considerations.

The report of these findings is attached to this memo.

IDA SCIENCE AND TECHNOLOGY POLICY INSTITUTE

Evaluation of the Research Outputs from the 2010-2012 NIH Transformative Research Awards

Sally S. Tinkle

Xueying “Shirley” Han

Lara L. Rubinyi

Erick D. Cohen

March 24, 2021

Executive Summary

The NIH Director's Transformative Research Award (TRA) initiative is one of four components of the NIH Common Fund's High-Risk, High-Reward Research (HRHR) program. Launched in 2009, the TRA initiative is designed to support individuals or teams who propose biomedical research that is inherently risky, untested, and has the potential to create or overturn fundamental scientific paradigms.

The NIH Office of Strategic Coordination (OSC) asked the IDA Science and Technology Policy Institute (STPI) to evaluate the research outputs from the 2010–2012 TRA awardee cohorts and determine whether the research is *more innovative, impactful, and interdisciplinary* than research produced by comparison groups. Through consultation with OSC, STPI translated this overarching research question into two key study questions:

- Do the scientific outputs produced by TRA awardees represent a *paradigm shift* for biomedical research, that is, a significant change in a universally recognized scientific achievement?
- Are the outputs *more impactful* than research produced by comparison groups?

To address these questions, a multi-modal study design was developed for data collection, analysis, and integration. This included surveys of TRA and comparison group awardees, analyses of bibliometric and altmetric data, and a senior scientist review. Two comparison groups were used for this evaluation: recipients of the NIH Director's Pioneer Award (NDPA) who are presumed more likely to produce research outputs similar to TRA awardees, and R01 recipients who are presumed more likely to produce more traditional research outputs. As a result, two separate analyses were performed for this evaluation: the TRA-NDPA comparison and the TRA-R01 comparison. Key findings for each comparison from the awardee survey, bibliometric analysis, and senior scientist survey are presented separately below.

It is important to note that small sample sizes result in large standard errors, which can lead to imprecise estimates of the true effects between groups of interest. Therefore, the survey results described in this report should not be taken as firm conclusions representing the actual awardee populations.

Integration of Findings

Key study question: Do the scientific outputs produced by TRA awardees represent a *paradigm shift* for biomedical research?

TRA-NDPA comparison

- Results from the awardee and senior scientist reviewer surveys suggest that there are no significant differences in transformative potential between TRA and NDPA research outputs.

TRA-R01 comparison

- Results from the awardee and senior scientist reviewer surveys indicated several different areas in which TRA research was considered to have more transformative potential than R01 research. However, the small sample sizes for the number of awardee and senior scientist reviewer survey respondents precludes definitive conclusions.

Key study question: Are the outputs *more impactful* than research produced by comparison groups?

TRA-NDPA comparison

- TRA awards produced similar numbers of publications and received significantly fewer citations than NDPA awards once *award duration, total direct cost, whether an award had single or multi-PIs,* and the research *area of science* were factored into the analysis. In contrast, TRA publications received significantly more citations than NDPA publications once *year of publication, whether an award had single or multi-PIs,* and *number of authors* on a publication were taken into consideration. In addition, TRA publications had significantly higher Altmetric attention scores than NDPA publications. The lack of agreement among different publication and citation metrics precludes definitive conclusions on whether TRA research outputs were more impactful than NDPA research outputs.

TRA-R01 comparison

- TRA awards produced significantly more publications and received significantly more citations than R01 awards once *award duration, total direct cost, whether an award had single or multi-PIs*, and the research *area of science* were factored into the analysis. In addition, TRA publications received significantly higher citations than R01 awards once *year of publication, whether an award had single or multi-PIs,* and *number of authors* on a publication were taken into consideration. TRA publications also had significantly higher Altmetric

attention scores than R01 publications. Combined, these results suggest that TRA research outputs were more impactful than the R01 comparison group research outputs.

Summary and Recommendations

Transformative Research

The variability in terminology used by senior scientist reviewers to describe transformative research reflects the individuality of their interpretations of *transformative research*. While acknowledging the NIH interest in providing principal investigators with the flexibility to propose a broad array of transformative research ideas, STPI suggests NIH consider approaches to develop more specific Funding Opportunity Announcement language and review criteria to define transformative research, or the characteristics of transformative research. This more specific language could then be used in future program evaluations, perhaps increasing the likelihood of identifying differences between awardee groups.

Research Impact

The timeframe established for this evaluation begins 1 year after an award's project start date and ends 1 year after the project end date, which may be an insufficient amount of time to assess the transformative impact of the resulting research on the scientific paradigms of biomedical research. Furthermore, it is challenging to recognize when a specific research finding is transformative. Because scientific progress builds on previous research, very rarely—if ever—does an idea or theory emerge *de novo*. Moreover, current multi-disciplinary and interdisciplinary research make it less likely that a major scientific breakthrough will be made by a single researcher or small group of investigators.

Moving forward, NIH could evaluate TRA research over a timeline more reflective of the decade or more needed to determine that research is transformative; increase the TRA cohort size by adding awardees to the 2010–2012 group as they meet the criteria for *project end date plus 1 year;* or refocus the goal of the initiative and its concomitant evaluation from an emphasis on research outputs to spurring novel, paradigm-shifting thinking.

Programmatic Impact

Results of this evaluation suggest that transformative research is occurring in all three groups—TRA, NDPA, and R01—although to differing degrees. NIH might consider the relationship of the NDPA and R01 programs to the TRA initiative and determine whether greater distinction between the programs benefits the NIH mission or whether an emphasis on transformative potential regardless of mechanism would be more effective. In addition, an examination of the conceptual and operational similarities and differences between the

TRA and the high risk, high reward New Innovator Award program, both of which emphasize innovative research, could enhance the unique characteristics and biomedical benefit of each program.

Concluding Thoughts

Several components of this multi-modal analysis demonstrate that, despite the definitional challenges and limitations to the study design, the 2010–2012 TRA awardees have produced impactful biomedical research that aligns with the goals of the initiative. Determination of the degree of transformative impact will require the test of time; however, numerous TRA awardees acknowledged in the free response survey questions the importance of the TRA initiative in funding research they believed to be outside the parameters of the traditional R01 mechanism. These comments and the results of this evaluation confirm the role of the TRA initiative in spurring transformative research at NIH.

Contents

1. Introduction

A. Overview of the National Institutes of Health Director's Transformative Research Award Initiative

The National Institutes of Health (NIH) Director's Transformative Research Award (TRA) initiative is a component of the NIH Common Fund's High-Risk, High-Reward Research (HRHR) program. The HRHR program supports exceptionally creative scientists pursuing highly innovative research with the potential for broad impact in biomedical, biobehavioral, or social sciences within the NIH mission.[1] The TRA initiative, begun in 2009, supports individuals at all career stages or teams who propose research that is inherently risky, untested, and has the potential to create or overturn fundamental scientific paradigms.[2] Although the award uses the traditional NIH Research Project (R01) mechanism, the requirement for a detailed experimental plan has been replaced by a description of the scientific or technical challenge, the planned approach, and the transformative potential of the research results. No preliminary data are required, and budgets are flexible with no specified limit. The NIH Common Fund spent approximately $28 million on the initiative in FY2018, with a total NIH investment of $46 million.

B. Scope of the Evaluation

The NIH Office of Strategic Coordination (OSC) develops and coordinates Common Fund programs, including the TRA initiative. OSC asked the IDA Science and Technology Policy Institute (STPI) to evaluate research outputs from the 2010–2012 TRA awardee cohorts[3] and determine whether the research is *more innovative, impactful, and interdisciplinary* than research produced by comparison groups.[4] The results of this assessment, detailed here, are provided to the NIH Director to inform future TRA policy and investment decisions.

STPI initiated this effort by exploring more fully the characteristics of transformative research. STPI examined the descriptions of transformative research in the NIH 2010–2012

---

1 https://commonfund.nih.gov/highrisk

2 https://commonfund.nih.gov/tra

3 The criteria for the 2009 TRA FOA were modified prior to the release of the 2010 FOA; therefore, the 2009 awards were not included in the evaluation.

4 OSC SOO

TRA Funding Opportunity Announcements (FOAs), the National Science Foundation (NSF) and National Science Board (NSB) definitions of transformative research, and several publications in the peer-reviewed literature.[5,6,7,8,9] STPI's analysis identified *paradigm shift*, *impact*, and *innovation* as key descriptors of transformative research (Appendix A); however, STPI determined that survey options to assess *innovativeness* could not be clearly distinguished from those for *paradigm shift*. As a consequence, innovativeness was considered embedded in the survey options for paradigm shift.

To assess the transformative potential of the TRA awards and following consultation with OSC, STPI translated the overarching question in the statement of objectives into two key study questions:

- Do the scientific outputs produced by TRA awardees represent a *paradigm shift*[10] for biomedical research?
- Are the outputs *more impactful* than research produced by comparison groups?

STPI identified two comparison groups for this assessment: the NIH Director's Pioneer awardees who are presumed more likely to produce research outputs similar to TRA awardees and R01 awardees who are more likely to produce more traditional NIH research outputs. These groups are discussed more fully in the comparison group development section of the study design.

C. Study Design

To assess the multi-faceted questions outlined above, STPI conducted a retrospective cohort analysis using a multi-modal study design that included development of a logic model, comparison groups, surveys, and bibliometric analyses. Interviews would be conducted if clarification of survey responses was needed.

The logic model and comparison groups are foundational to the subsequent analyses and presented here in detail. Brief descriptions of the surveys and bibliometrics are also

---

5 2010 TRA FOA: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-09-022.html

6 2011 TRA FOA: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-10-010.html

7 2012 TRA FOA: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-11-006.html

8 https://www.nsf.gov/about/transformative_research/definition.jsp

9 National Science Board, *2020 Vision for the National Science Foundation*, 2005

10 A paradigm shift is a significant change in a universally recognized scientific achievement that for a time provide model problems and solutions to a community of practitioners (see Appendix A for additional information).

provided in this section to give a holistic overview of the study design, with detailed methodologies in the appropriate sections of the report.

1. Logic Model

a. Developing the Framework

A logic model is an evaluation tool that uses an implicit *if-then* construct to depict shared relationships among a program's resources, activities, outputs, and outcomes. The model is not intended to provide all details about a program, but instead focuses on those key aspects that are likely to influence observed outcomes.[11]

The primary components of a logic model that organizes information pertaining to NIH and to the awardee are presented in Figure 1. A logic model often includes the rationale and assumptions that were inherent in the program's development.[12] STPI identified publicly available information and employed subject matter expertise to generate elements of the rationale (a set of reasons for a course of action) and assumptions (concepts and ideas accepted as true without proof) that might underlie TRA initiative development. This exercise creates a knowledge framework within which each of the logic model components can be examined.

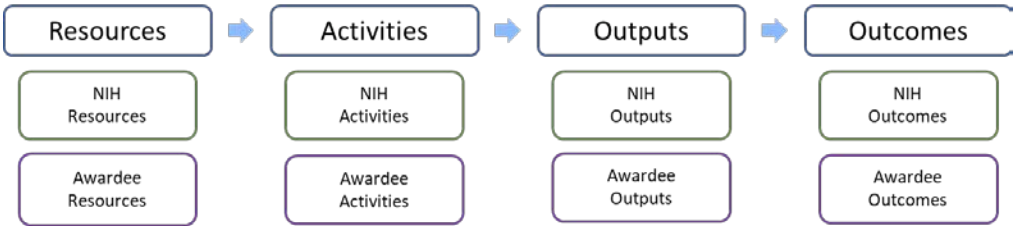


Figure 1. Primary Components of an NIH Logic Model

STPI identified high-level elements of a rationale for the development of the TRA initiative within the HRHR program. In the context of the STPI evaluation, funding fundamental biomedical research is part of the NIH mission to improve public health and well-being. Focused initiatives are a mechanism to support more of a specific type of research in a shorter time span, and a transformative research initiative will produce more novel, innovative, and paradigm-shifting discoveries, products, and tools faster than traditional research approaches. Transformative research outcomes will provide the

---

[11] Innovation Network Logic Model Workbook http://www.pointk.org/client_docs/File/logic_model_workbook.pdf

[12] Frechtling, JA (2015). "Logic Models". *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier. pp. 299–305.

American people better health outcomes, greater economic gain and more biomedical applications faster.

To understand the assumptions underlying the rationale in the context of public information on the TRA initiative and NIH processes and culture, STPI outlined a series of assumptions that align with this rationale. There is an overarching assumption that the TRA initiative would effectively solicit more transformative research proposals and generate more transformative research results. The program structure infers that a Common Fund initiative could use the traditional R01 mechanism for review of paradigm-shifting research although the process's established review criteria could present barriers to transformative research. R01 funding does include the potential for larger budgets. Finally, use of the R01 mechanism may imply that a successful TRA award would produce research results that could be used as preliminary data in a follow-on traditional R01 application process.

Within this framework of rationale and assumptions, STPI first examined NIH and awardee resources—those elements within NIH that support the TRA initiative. In the first column of Figure 2, the noteworthy resources are the Common Fund and the R01 mechanisms, which provide the reassurance of an established application process and funding. The HRHR program's goals are consistent with elements of the TRA initiative— an emphasis on high risk and innovative, high impact biomedical research. The primary awardee resources are transformative ideas and the professional training and experience to accomplish transformative research.

The anticipated *NIH activities* follow the NIH grant award process—publishing the TRA FOA and developing a review process that identifies transformative research potential. NIH also interacts with awardees formally through program officers and more informally through the annual HRHR research symposium, which brings together awardees from all four HRHR initiatives to build community and collaborations. Anticipated *awardee activities* focus primarily on the research process: conduct the experiments, publish the results, and apply for additional funding.

Near-term *outputs* from NIH focus on refining the FOA and awardee selection criteria, and optimizing the size of the program, all of which would lead to the outcome of a more effective TRA initiative and the potential for institutes within NIH to adopt or expand programs similar to the trans-NIH Common Fund TRA initiative. TRA researchers are likely to produce innovative, paradigm-shifting research outputs, which would enhance their careers and likelihood that they will continue to propose transformative research.

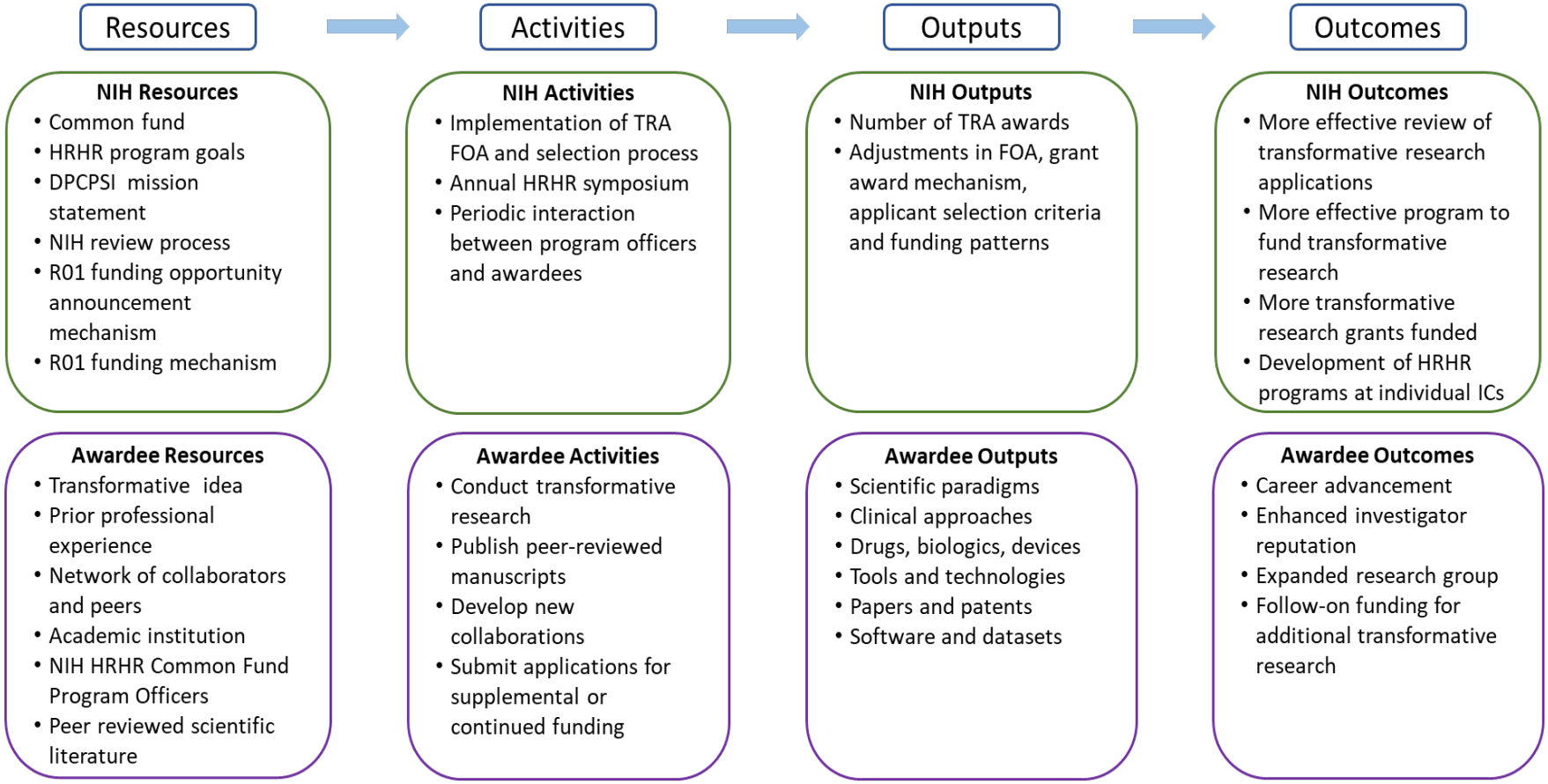


Figure 2. Transformative Research Award Logic Model

2. Comparison Groups

To assess the characteristics of transformative research—paradigm shifting, innovative, impactful—STPI identified two comparison groups: the NIH Director's Pioneer Award (NDPA) recipients and R01 recipients. Two separate analyses were performed for this evaluation: the TRA-NDPA comparison and the TRA-R01 comparison.

NDPA awardees are scientists with outstanding records of exceptional creativity pursuing new research directions to develop pioneering approaches to major challenges in biomedical, social science, and behavioral research.[13] Like the TRA initiative, this award is offered through the Common Fund HRHR program. The pioneering approach that it encourages suggests research outputs similar to those anticipated from the TRA award. The second comparison group is composed of investigators who received an R01 award during FY2010–2012. The R01 grant is the NIH award made to support a discrete, specified, circumscribed project that represents the investigator's specific interest and competencies as they relate to the mission of the NIH.[14] Research outputs from this award are anticipated to reflect the *discrete, specified, circumscribed* characteristics rather than the paradigm-shifting characteristics of transformative research.

[13] https://commonfund.nih.gov/pioneer

[14] https://grants.nih.gov/grants/funding/r01.htm

a. Considerations for Comparison Group Development

STPI recognized that the number of principal investigators (PIs) per award could be a potential confounding factor. Although TRA and R01 awards allow single and multiple PIs, NDPA awards are single PI only. Because the literature is inconclusive regarding the influence of single and multiple PIs on research output, STPI accounted for the difference between single and multi-PI awards by including it as an independent variable in a multi-variable analysis as part of the bibliometrics analyses.[15,16] The multiple PI condition also precluded the use of investigator-related demographic data (e.g., age, gender, years since PhD) to develop comparison groups, and STPI used the range of direct costs for each TRA award as an R01 selection criterion. STPI did not include no-cost extensions and budget supplements in data collection, and eliminated outliers whose award was greater than $10 million dollars in the final dataset. The resulting distribution of direct costs for the majority of the FY2010–2012 TRA awards ranged between $596,000 and $9,000,000 (Figure 3).

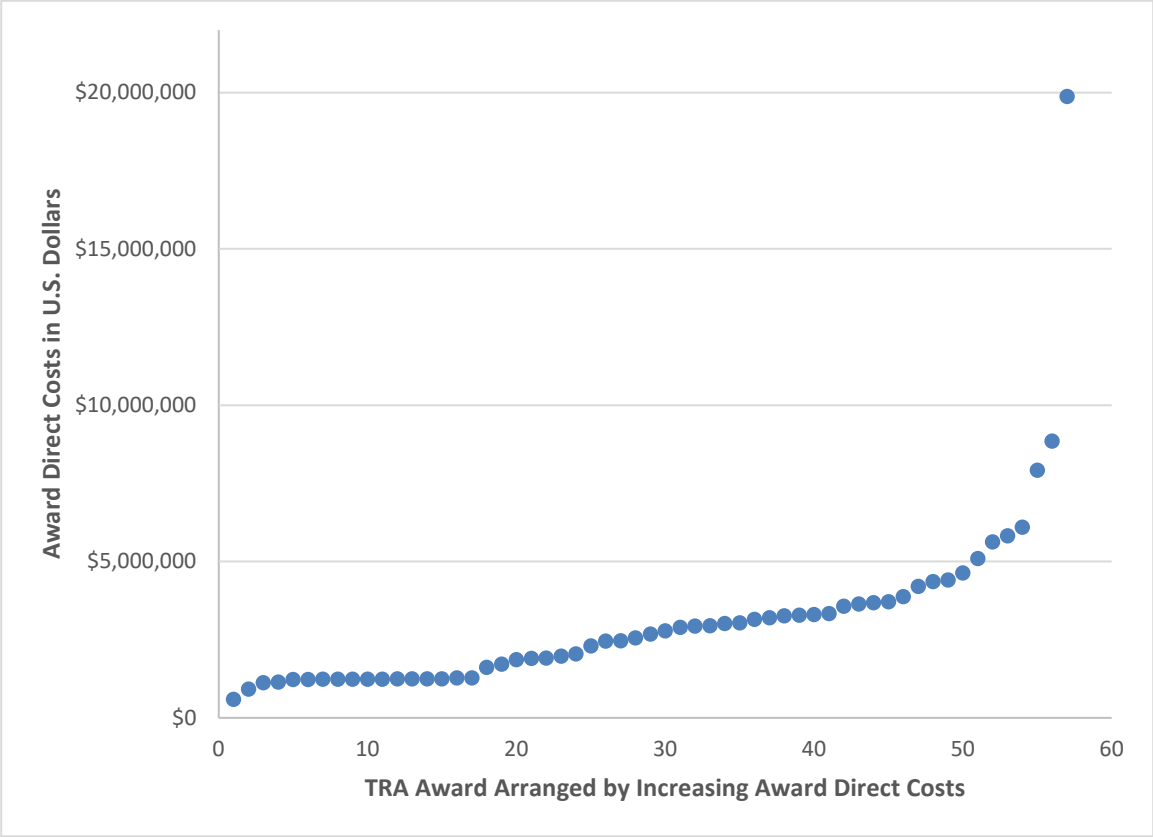


Figure 3. Direct Cost Analysis for TRA Awards

[15] https://www.sciencedirect.com/science/article/pii/S0362331917300708

[16] https://www.nap.edu/read/19007/chapter/14#209

b. Comparison Groups

1) TRA Group

NIH made 57 TRA awards between FY2010–2012: 20 awards in 2010, 17 in 2011, and 20 in 2012. For the TRA-NDPA comparison, STPI excluded five TRA investigators who received an NDPA during the same years for which they had a TRA award. Fifty-two awards are included in the TRA-NDPA comparison. All 57 TRA awards were included in the TRA-R01 comparison.

2) NDPA Group

NIH awarded 40 NDPAs in fiscal years 2010–2012: 17 awards in 2010, 13 in 2011, and 10 in 2012. To achieve closer parity with the TRA cohort, STPI and NIH agreed to include the 2009 NDPA cohort of 18 additional awards. Of the 58 awards, STPI excluded 6 NDPA investigators who received a TRA award during the same years that their NDPA was active. Fifty-two NDPA awards are included in the final comparison group.

3) R01 Comparison Groups

To identify a matched 2010–2012 R01 comparison group, STPI developed a sampling strategy that incorporated the NIH Query, View, and Report (QVR) search criteria: award year, award duration, and direct cost.

As outlined in Figure 4, STPI first selected all possible R01 recipients of awards in FY2010–FY2012 and obtained 11,348 matches. Of these awards, 10,831 had direct costs within the TRA range, and award duration reduced the R01 subset to 5,927 awards. STPI then employed Propensity Score Matching[17] to ensure proportionality between the TRA and R01 groups for award duration, distribution of single and multi-PIs, and total direct cost.

STPI identified 169 potential R01 awardees, approximately 3 times the number of TRA awardees, because comparison groups are often less incentivized to participate in surveys. This larger group size would allow STPI to compensate for a lower overall response rate and ensure that approximately the same number of R01 awardees responded to the survey as the more incentivized TRA awardee group.

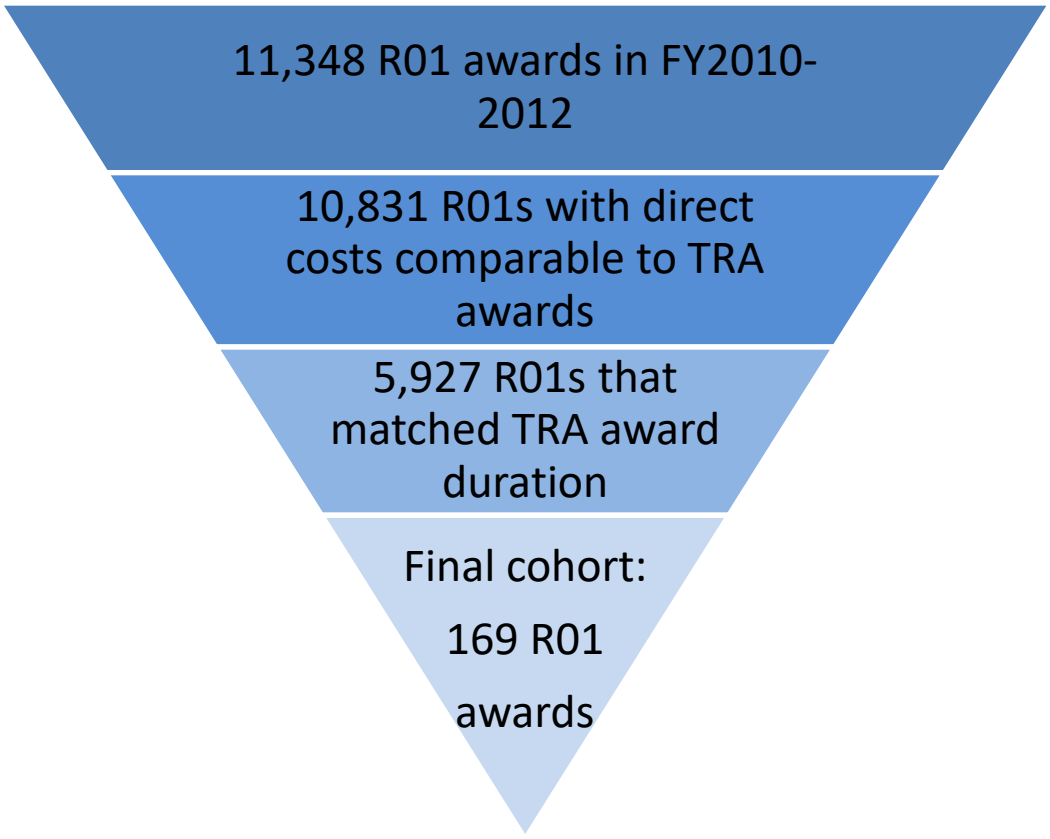


Figure 4. Development of 169 Matched R01 Awardees

[17] Austin, P.C. 2011. "An introduction to propensity score methods for reducing the effects of confounding in observational studies." *Multivariate Behavioral Research* 46(3): 399-424.

The peer-reviewed scientific literature has documented that publication and citation rates vary by field of science, a finding that could impact interpretation of the bibliometric analysis.[18] STPI examined the areas of science represented in the award abstracts to determine whether the distribution of the areas of science was proportional for the TRA-NDPA groups and the TRA-R01 groups.[19] A STPI subject matter expert reviewed each abstract manually and noted key terms. The key terms were then reviewed and four areas of science emerged: biomedical research, biobehavioral research, therapy, and tool development.

The number of awards in the four areas of science were approximately the same for each group, with the exceptions of TRA awardees receiving more tool development awards and R01 awardees proposing more biobehavioral research.

3. Surveys of Awardees

Surveys sample individuals from a population to make inferences about the population being studied. The awardee survey was administered to three groups: TRA awardees, NDPA awardees, and R01 awardees. Two separate analyses are performed on the data and should not be conflated. The TRA awardees are assessed against the NDPA awardees and provide one set of results, and in the second analysis, TRA awardees are assessed against R01 awardees to produce a second set of results.

a. Awardee Survey

STPI used the characteristics of transformative research (Appendix A) to develop the structure and content of a survey that assesses TRA research outputs. Briefly, STPI asked awardees whether their research was paradigm shifting, innovative and/or impactful; changed scientific thinking or research practice; developed new technologies; or had near-term or future impact.

STPI also developed a plan to maximize the survey response rate so that inferences about TRA research in FY 2010–2012 might be considered even if they are not statistically significant. The term *awardees* is used in this report when the entire comparison group is being analyzed or discussed, and *respondents* is used to delineate analyses and discussions that pertain solely to those who completed the survey.

---

[18] Piro, Fredrik Niclas, Dag W. Aksnes, and Kristoffer Rørstad. "A macro analysis of productivity differences across fields: Challenges in the measurement of scientific publishing." Journal of the American Society for Information Science and Technology 64, no. 2 (2013): 307-320.

[19] STPI first applied the INSPIRE cluster analysis program that was created by the Pacific Northwest National Laboratory (PNNL) to the content of the award abstracts and determined that the data set was too small for this approach to provide meaningful results.

b. Senior Scientist Review Survey

Expert review is the cornerstone of the NIH review process. STPI used an analogous process to perform expert review of the transformative potential of the research published by the TRA and comparison group survey respondents. To align the expert's area of science with the TRA outputs, STPI solicited the names of experts from the awardees through the survey and from NIH program officers managing the TRA awards.

Research outputs for survey respondents for all three groups were randomly assigned in the senior scientist review to provide each reviewer with a selection of transformative and non-transformative research outputs. The senior scientist review survey mirrored awardee survey content so that the expert responses could be aligned with the awardee responses.

4. Bibliometric Analyses

Bibliometric analyses provide a measure of quantity, quality, and impact of published scientific articles.[20] STPI assessed well-established publication and citation metrics, as well as more recent, alternative metrics (Table 1).[21] Although altmetrics is not accepted as a rigorous measure of quantity and quality on its own, it does provide an additional dimension to the understanding of scientific impact.

Table 1. Metrics of Productivity and Impact in the TRA Assessment

Publication Level Metrics	Citation Level Metrics	Alternative Metrics (altmetrics)
Total number of peer-reviewed publications	Total number of citations	Wikipedia
Number of publications per year	Number of citations per year	Public policy documents
Time to first publication	Time to first citation by publication	Research blogs
Time to first publication by project start date	Time to first publication by project start date	Mainstream media
Overall time to publication for each publication measured by publication start date	Overall time to each citation for each publication measured by publication start date	Social media
Cost per publication	Cost per publication	
	Relative Citation Ratio	

5. Organization of the Report

The remainder of this report is organized into the following sections: analysis of TRA-NDPA awardees (survey and bibliometric analyses); analysis of TRA-R01 awardees (survey and bibliometric analyses); senior scientist review; integration of findings; and summary and recommendations. Appendix A contains the characteristics of transformative research; Appendix B survey questions aligned with the concepts of paradigm shift and impact; Appendix C the awardee survey; Appendix D survey results for TRA, NDPA, and R01 awardees; Appendix E the senior scientist review survey questions; Appendix F the

[20] https://library.leeds.ac.uk/info/1406/researcher_support/17/measuring_research_impact

[21] Altmetric. 2020. "What are Altmetrics? An Introduction." Accessed 21 May 2020. Available at: https://www.altmetric.com/about-altmetrics/what-are-altmetrics/.

senior scientist review survey data tables; and Appendix G the complete list of senior scientist review descriptors for transformative research.

2. Results for the TRA-NDPA Analysis

A. Awardee Survey

To assess TRA's research impacts, STPI created an online survey to solicit awardees' perspectives on their TRA and NDPA awards. The survey questions and survey administration were nearly identical for each group, and differences between the two surveys are discussed below. All survey materials can be found in Appendix C.

1. Survey Development

The STPI team developed survey questions that addressed the multiple aspects of paradigm shifting, and innovative/impactful research. The survey items were iterated with the NIH HRHR team, and content and format were tested through a STPI focus group. A summary of the survey questions by *paradigm shifting* and *innovative/impactful* categories is provided in Appendix B.

2. Survey Administration

STPI developed and administered the surveys in Alchemer, a web-based survey platform. Each awardee received a personalized invitation to take the survey; responses were kept confidential, and only aggregate results are provided. In tandem with the survey, TRA recipients received an email from HRHR staff underlining the importance of the study and asking them to participate in the survey. The survey was sent to the 52 TRA and 52 NDPA awardees on December 3, 2019. The first reminder to this initial set was sent 1 week later, on December 10, 2019, and the second reminder was sent 1 week after that, on December 17, 2019. The last reminder was sent after the year-end holidays, on January 7, 2020. The survey remained open until February 19, 2020.

3. Survey Analysis

Analysis was performed only on completed surveys. Descriptive statistics such as the number of responses for each question and the percentage of survey respondents selecting each answer choice are provided for each question. Free response questions were coded into a series of qualitative categories and are reported as counts for each category. A comprehensive list of survey questions is in Appendix C.

Two-sample proportion tests[22] were used for questions with answer choices of *agree* and *disagree* to examine whether the percentage of survey respondents who selected *agree* differed between TRA and NDPA awardees. Because there are only two answer choices, STPI did not perform a two-sample proportion test for those who selected *disagree* as it is simply the complement of those who selected *agree* and would result in the same statistical significance. For survey questions that had 5-item Likert scale response choices (e.g., *strongly disagree* to *strongly agree*), positive answer choices were summed and a two-sample proportion test was performed to compare whether the percentage of survey respondents who responded positively differed between TRA and NDPA awardees.[23] A two-sample proportion test was also performed for *likely/unlikely* questions to compare the percentage of TRA and NDPA respondents who selected *likely* for that item, as well as for questions that were presented as *select all that apply* items.

All statistically significant findings in this analysis are significant at $p < 0.05$. Additional descriptive statistics such as counts and percentage of survey respondents responding to each answer choice are also included in the analysis. Descriptive statistics, including all statistical tests, can be found in Appendix D.

4. Results

a. Response Rate

Of the 52 TRA awardees eligible for the TRA-NDPA comparison, 23 completed the survey (44% response rate).[24,25] Of the 52 NDPA awardees eligible for the TRA-NDPA comparison, 32 completed the survey (62% response rate).[26] Full data tables of survey results are provided in Appendix D.

---

[22] Yates' continuity correction was applied automatically in R for all two-sample proportion tests to account for the small sample sizes of awardee respondents.

[23] No statistical analyses were performed on combining the *strongly disagree* and *disagree* answer choices because the number of respondents who selected these answer choices were too few for any statistical tests to be meaningful.

[24] For the TRA-NDPA comparison, STPI excluded five TRA investigators who received an NDPA during the same years for which they had a TRA award. Fifty-two awards are included in the TRA-NDPA comparison.

[25] Of the 57 total TRA awardees, 24 completed the survey. However, only 23 TRA awardees are eligible for the TRA-NDPA comparison because one of the TRA awardees was excluded from this comparison because the individual had received both TRA and NDPA awards. All 24 awardee surveys were used in the TRA-R01 comparison.

[26] There was one partial (i.e., incomplete) survey response from TRA awardees, and two partial survey responses from NDPA awardees. These three partial survey responses were not included in any of the survey analyses.

b. Descriptive Statistics

1) Paradigm Shifting

The first set of questions addressed the paradigm-shifting nature of the grantees' research. These questions included novel inventions, novel combinations of ideas and approaches, or use of technology that significantly changed research practice or thinking in their field (Figure 5). TRA and NDPA survey respondents reported at similar levels that they agreed with the statements *My research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice* (100% and 91% for TRA and NDPA survey respondents, respectively; $\chi_1^2 = 0.76, p = 0.38$); *One or more of my research ideas challenged existing science/technology paradigms.* (82% and 90%; $\chi_1^2 = 0.18, p = 0.67$); *My research led to a novel invention or a new technology which significantly improved current practices.* (95% and 87%; $\chi_1^2 = 0.3, p = 0.58$); *My research furthered existing practices or thinking in my field.* (81% and 63%; $\chi_1^2 = 1.1, p = 0.3$); *My research led to the development of a new methodology that significantly changed research in my field.* (70% and 74%; $\chi_1^2 = 0, p = 0.95$); *My research took the next steps in my established area of investigation.* (68% and 74%; $\chi_1^2 = 0.3, p = 0.87$); *My research required the use of equipment, technique or model that was novel and significantly changed research in my field.* (73% and 55%; $\chi_1^2 = 1.07, p = 0.3$); and *My research aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice* (71% and 56%; $\chi_1^2 = 0.68, p = 0.41$).

Figure 5. Paradigm-Shifting Research Outputs by Type

2) Innovative and Impactful

This question block explored the timing of impact (Figure 6). TRA and NDPA survey respondents reported at similar levels that they agreed with the statements *My research resulted in an immediate and significant shift in practices or thinking* (23% and 38% for TRA and NDPA survey respondents, respectively; $\chi_1^2 = 0.73$, $p = 0.39$); *My research resulted in a gradual but building shift in practices or thinking* (62% and 57%; $\chi_1^2 = 0$, $p = 0.97$); and *My research resulted in a delayed but significant shift in practices or thinking* (64% and 79%; $\chi_1^2 = 0.72$, $p = 0.4$).

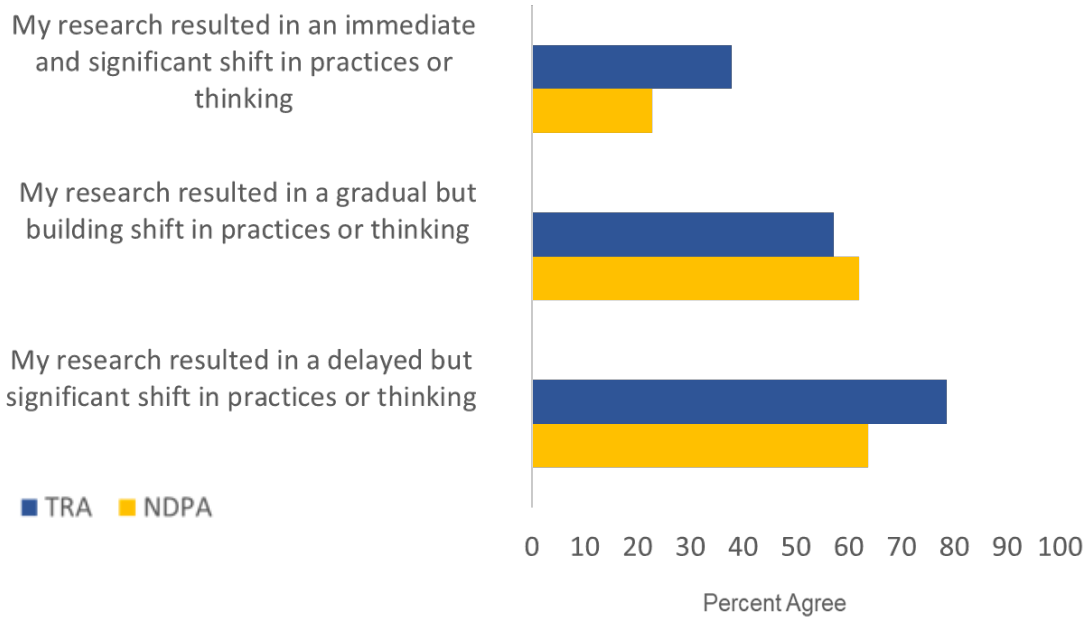


Figure 6. Timing of the Onset of Impact

For the next set of questions TRA and NDPA respondents were not significantly different in responding *likely* that their research would result in the following statements (Figure 7): *A new synthesis of disparate ideas* (48% and 58% for TRA and NDPA survey respondents responding *likely*, respectively; $\chi_1^2 = 0.22$, $p = 0.64$), *A change in how research is conducted* (87% and 63%; $\chi_1^2 = 2.9$, $p = 0.09$), *The discovery of new phenomenon or advancement of a theoretical concept* (70% and 68%; $\chi_1^2 = 0.1$, $p = 0.93$), *The development of new therapies, clinical tools, or strategies* (70% and 68%; $\chi_1^2 < 0.001$, $p = 1$), *The development of a new technology* (70% and 84%; $\chi_1^2 = 0.96$, $p = 0.33$), and *The development of a new methodology* (70% and 68%; $\chi_1^2 < 0.001$, $p = 1$).

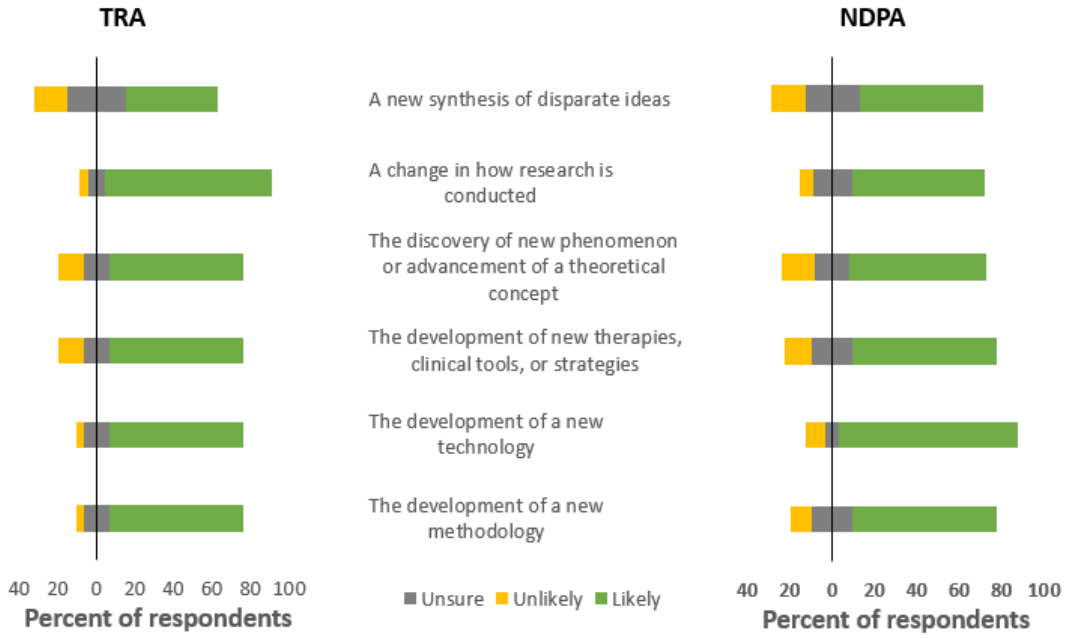


Figure 7. Likeliness of Research Resulting in Impact

This question block examined achievements and honors grantees received once their award-funded research was published (Figure 8). This was a "select all that apply" question, and TRA and NDPA survey respondents reported at similar levels that the following applied: *My research has been listed as a highlight in an academic journal* (77% and 81% for TRA and NDPA survey respondents, respectively; $\chi_1^2 < 0.001$, $p = 0.99$)*; I received an award/honor* (100% and 100%)[27]; *I have been invited to serve as a regular journal reviewer* (91% and 88%; $\chi_1^2 < 0.001$, $p = 1$); *I have been asked to give an invited presentation about my research* (64% and 75%; $\chi_1^2 = 0.35$, $p = 0.55$); *My research has been featured in the popular press/media* (73% and 78%; $\chi_1^2 = 0.017$, $p = 0.90$); and *I have been invited to be a keynote speaker to share my research* (46% and 69%; $\chi_1^2 = 2.05$, $p = 0.15$).

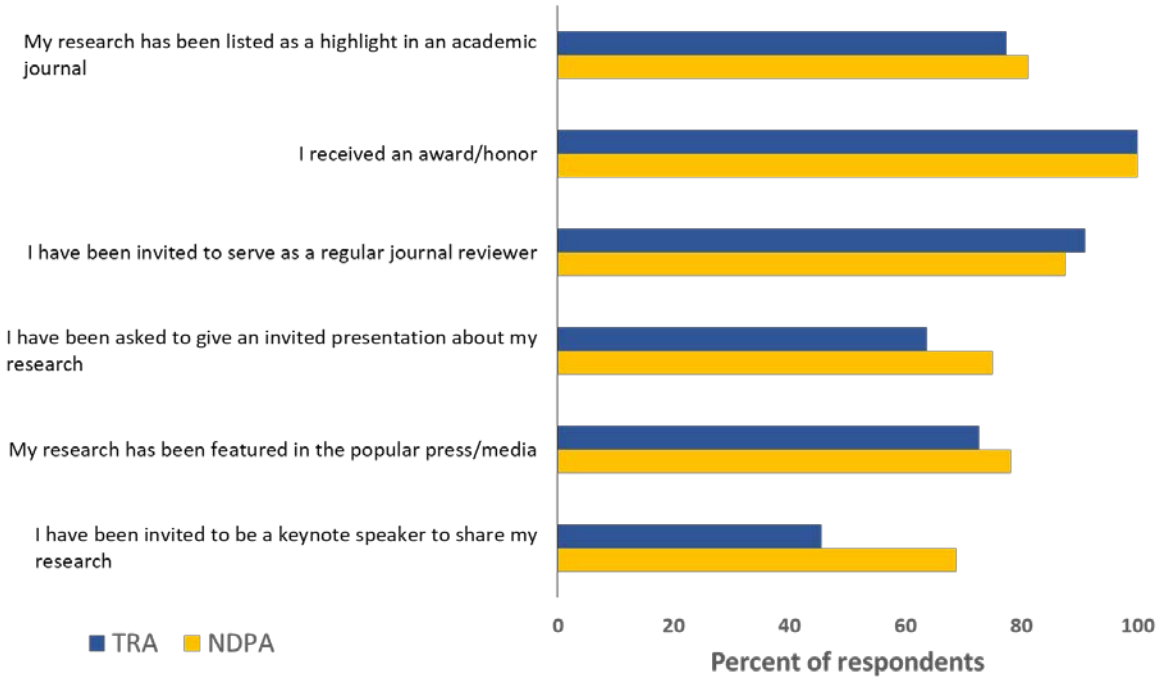


Figure 8. Honors or Awards as a Measure of Research Publication Impact

[27] A Chi-squared test statistic could not be calculated because there was no variability in the responses.

Next, respondents were asked whether these achievements were attributable to their award (Figure 9). This was a 'select all that apply' question that was populated by responses to the previous question. TRA and NDPA survey respondents reported at similar levels that the following applied: *My research has been listed as a highlight in an academic journal* (47% and 58% for TRA and NDPA survey respondents, respectively; $\chi^2_1 = 0.14,\ p = 0.71$); *I received an award/honor* (91% and 84%; $\chi^2_1 = 0.084,\ p = 0.77$); *I have been invited to serve as a regular journal reviewer* (90% and 93%; $\chi^2_1 < 0.001,\ p = 1$); *I have been asked to give an invited presentation about my research* (79% and 92%; $\chi^2_1 = 0.43,\ p = 0.51$); *My research has been featured in the popular press/media* (88% and 80%; $\chi^2_1 = 0.04,\ p = 0.84$); and *I have been invited to be a keynote speaker to share my research* (70% and 86%; $\chi^2_1 = 0.37,\ p = 0.54$).

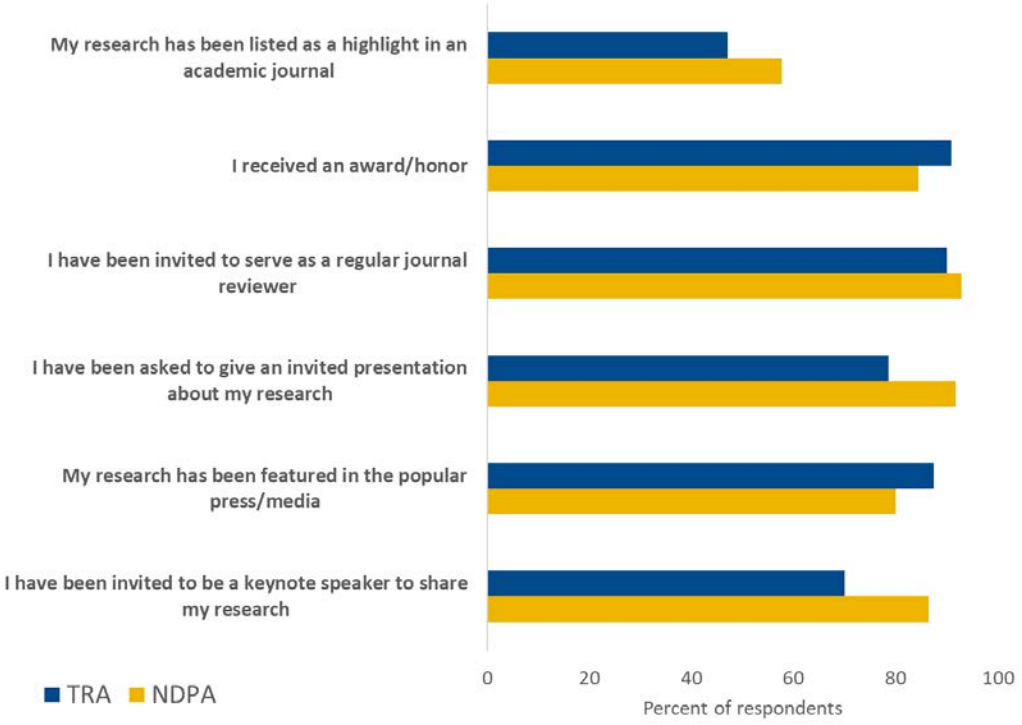


Figure 9. Honor or Awards Attributed to TRA/NDPA Research Outputs

This question block examined the flexibility of the grant mechanism (Figure 10). Respondents indicated whether they strongly disagree, disagree, neither agree nor disagree, agree or strongly agree. Significance tests were performed by combining agree and strongly agree items.[28] TRA and NDPA survey respondents selected agree or strongly agree at similar levels for the following statements: *The period of the grant was long enough to redirect research as ideas/methods evolved* (100% and 69% for TRA and NDPA survey respondents agreed, respectively ; $\chi_1^2 = 0.83$, $p = 0.36$), *The grant allowed me the freedom to pursue nontraditional research* (70% and 97%; $\chi_1^2 = 0.09, p = 0.77$); and *The grant allowed for flexibility in the use of funding* (96% and 97%; $\chi_1^2 = 1.8, p = 1$).

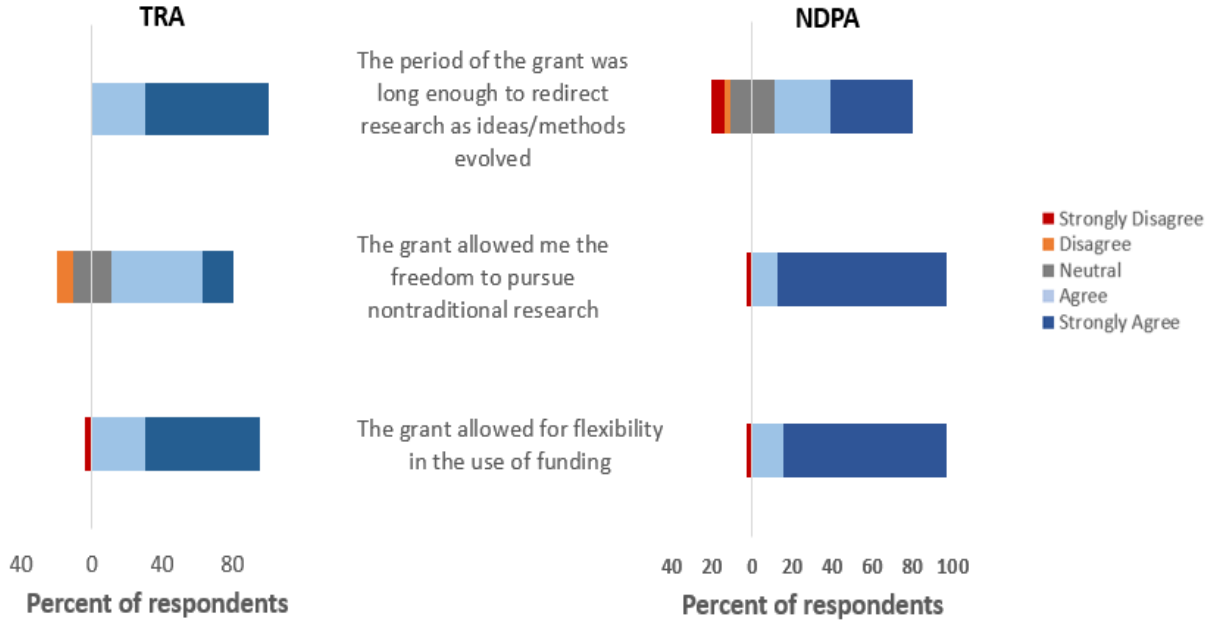


Figure 10. Flexibility of Grant Mechanisms

[28] No statistical analyses were performed on combining the *strongly disagree* and *disagree* answer choices because the number of respondents who selected these answer choices were too few for any statistical tests to be meaningful.

The next two questions ask about changes to their research plan, and whether or not the grantees are continuing the trajectory of their research (Figure 11). Both TRA and NDPA survey respondents agreed at similar levels that the following occurred: *My research plan changed significantly from what I originally proposed* (44% and 44% for TRA and NDPA survey respondents, respectively; $\chi_1^2 = 0$, $p = 1$), and *I am continuing or planning to continue the trajectory of my TRA- (or NDPA-) funded research post-award* (100% and 91%; $\chi_1^2 = 0.83$, $p = 0.36$).

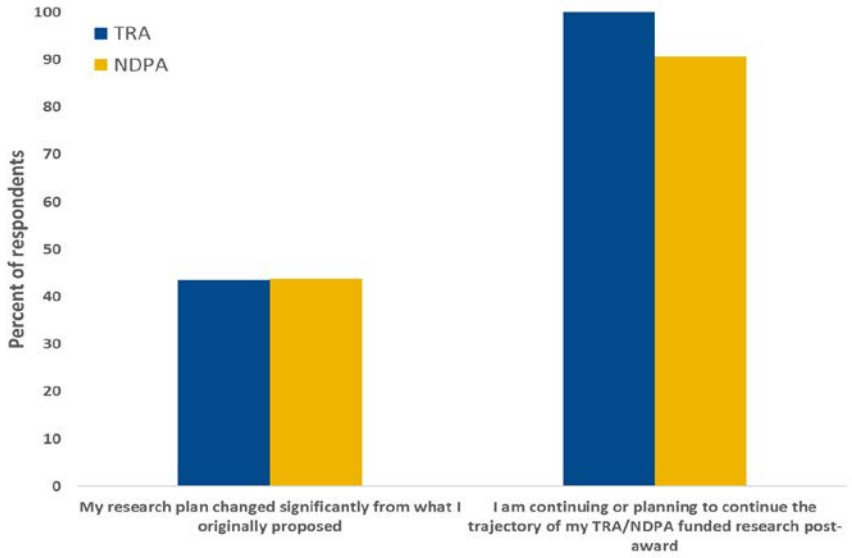


Figure 11. Research Direction

If a respondent indicated that they are planning to continue the trajectory of their research, the last question in this theme asks whether this research is funded (Figure 12). TRA and NDPA respondents (70% and 86%, respectively) reported similar levels of funding for their post-award research ($\chi_1^2 = 1.86, p = 0.17$): *You indicated you were continuing or planning to continue the trajectory of your TRA- (or NDPA-) funded research post-award. Please indicate whether this continuation is funded.*

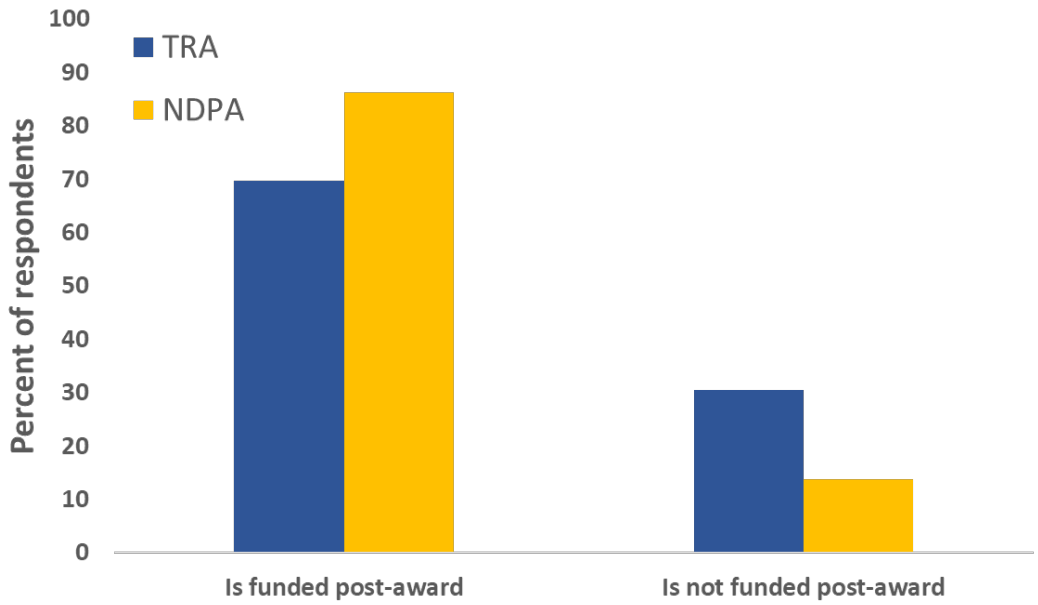


Figure 12. Funding Post-Award

c. Qualitative Analysis of Free Response Questions

In addition to the survey questions asked and examined above, respondents were asked to give their thoughts in free response at various stages of the survey. To analyze these data, STPI inductively coded responses. Below is an examination of responses.

1) Changes in Research Plan

If respondents indicated that they agree with the following statement, *My research plan changed significantly from what I originally proposed*, they were given the following open response question: *You indicated that your research plan changed significantly from what you originally proposed. Please indicate why those changes occurred.* Ten TRA respondents responded to this question, as did 14 NDPA respondents.

Several TRA (n = 4) respondents and NDPA (n = 6) respondents indicated that results or discoveries led to a new research direction. A TRA respondent said, "The work was substantially expanded from the original research plan. It was a combination of finding

exciting new results that we didn't anticipate, developing new experiments to test them further, and exploring additional multiple independent measures."

Two TRA and three NDPA respondents reported that their new direction was due to technologies that were developed during the course of their grant. Both a TRA and NDPA respondent indicated that as they advanced their research it improved the technology in tandem, with the NDPA respondent saying, "We needed new technology to achieve the goals of the project. We achieved those goals, but in addition, we created wholly new technology that I had not anticipated initially."

Other reasons given for the change in research plan included the following: the initial plan was no longer possible or the best direction, there needed to be a change in methods, and there were unforeseen challenges as well as unanticipated opportunities.

2) Research Trajectory Post-Award

If respondents indicated that they disagreed with the statement, *I am continuing or planning to continue the trajectory of my TRA/NDPA funded research post-award,* they were asked, *You indicated that you are not continuing or planning to continue the trajectory of your TRA/NDPA funded research post-award. Please describe what factors went into your decision.* One TRA recipient and three NDPA recipients responded to this request. The TRA recipient indicated that they "moved on to new ideas and questions." Two NDPA awardees indicated that they moved from academia into industry. One NDPA awardee described the difficulty they encountered when attempting to get continued funding, saying, "because the grant is a one-time award, I have encountered more difficulty than my peers who benefit from the renewal pipeline." The respondent goes on to suggest that, "NDPA should be guided towards specific award mechanisms than having to break into entirely new funding streams."

3) Other Information

Lastly, all respondents were given the opportunity to answer the question, *Is there any additional information you would like to share?* Fourteen TRA respondents and 12 NDPA respondents provided information under this question. Six TRA and seven NDPA respondents gave positive feedback about their program. A TRA respondent said that, "This is an amazing program that provides tremendous opportunities to move the needle on new and important research fields that run outside of the traditional bounds." Another TRA respondent indicated that the program "is invaluable to creative research and should be continued." A NDPA respondent indicated that, "in a way the pioneer 'ruined' me—the intellectual and experimental freedom was so delicious I found it hard to return to the conventional grant format."

One concern that was brought up by two TRA and five NDPA respondents was the difficulty encountered in renewing funding after receiving an HRHR award. One NDPA respondent summed it up, saying, "I regularly speak with other grant awardees, and one thing we all have noticed is that finding funding can be very difficult after the NDPA. Generally even after the award the projects are considered highly creative and cutting-edge, and federal funding in particular is very challenging. I received an NDPA early in my career, but I have *never* received an R01, notwithstanding many attempts and a great track record of achievement (several papers in nature, science, cell, etc). Unlike grant awards, pioneer awards are one-time, so for our most creative scientists, there remain few options for funding ground-breaking research." A TRA respondent suggested an option for an extension after describing difficulties in traditional mechanisms, "It would have been very nice, however, to have had the option to apply for an extension of the TRA research, as it was highly successful and it would have been nice to be able to continue within a program that funds high-risk/high-reward ideas that have proven successful."

d. Limitations to the Data

As a reminder, 23 TRA and 32 NDPA awardees responded to the survey. Based on these sample sizes, the estimated effect size for a two-tailed alternative hypothesis testing to detect mean differences in two independent groups is $d = 0.78$. With power set at 0.80, this means that STPI could only detect large differences between TRA and NDPA survey responses. Consequently, caution should be taken when interpreting survey results. Small sample sizes result in large standard errors, which lead to imprecise estimates of the true effects between groups of interest—in this case, the TRA and NDPA awardees. Therefore, the survey results described in this report should not be taken as firm conclusions representing the actual awardee populations. For instance, a lack of statistical significance does not mean there is no effect; it might be the case that there was insufficient power to detect the effect of interest. Similarly, small sample sizes can lead to false-positive results and an overestimation of the magnitude of the relationship between two variables. In other words, a result with statistical significance does not mean there is a true effect between the groups of interest. Overall, careful consideration should be taken to avoid making strong conclusions about the TRA and NDPA awardees, regardless of whether survey results yielded statistical significance.

B. Bibliometric Analysis

1. Methodology

To assess whether there are bibliometric differences between TRA and NDPA outputs, STPI examined a variety of publication and citation level metrics, detailed below.

In addition, STPI used altmetric data to complement traditional, citation-based analyses and provide a more comprehensive understanding and assessment of research influence.

As a reminder, only publications attributed to the TRA and NDPA awards of interest as designated in NIH's QVR system are included in the following analyses. In addition, publications were limited to those that were published 1 year after an award's project start date and within a year of an award's project end date to account for delays in publishing.

Similarly, NIH's iCite database only contains citation data for articles published between 1980 and 2019, and only has Relative Citation Ratio (RCR) data for articles published between 1980 and 2018. [29]

a. Publication Level Metrics

To assess whether publication output differed between TRA and NDPA awards, STPI considered the following metrics:

- *total number of publications produced per award,*
- *average number of publications produced per award per year*, and
- *average total direct cost spent per publication per award.*

To determine whether the *total number of publications produced per award* differed between TRA and NDPA awards, a generalized linear model (GLM) with a Poisson distribution was used with *total number of publications produced per award* as the dependent variable, and *group* (*i.e.,* TRA or NDPA) as the explanatory variable. In addition, because there are other variables that could influence the total number of publications produced for a given award, STPI performed a multi-variable GLM regression analysis to take into consideration the effects of the following variables in addition to *group*:

- *award duration* (continuous variable, units in years),
- *total direct cost* (*(per $100,000 spent;* continuous variable, units in dollars),
- *whether the award had multi-PIs* (binary categorical variable: 1 for multi-PI, and 0 for single PI awards), and
- *area of science* (categorical variable consisting of biobehavioral research, biomedical research, therapy intervention, and tool development).

A type-II sum of squares analysis of deviance was performed to assess which of the explanatory variables in the GLM significantly affected the *total number of publications*

[29] iCite can be accessed at https://icite.od.nih.gov/.

produced per award. To determine which areas of science are significantly different from one another, pairwise comparisons among the different areas of science were performed using a post-hoc Tukey test with the *glht* function from the *multcomp* package in R.[30] Awards with zero publications were included in both the single- and multi-variable regression analyses.

A non-parametric Kruskal-Wallis rank sum test was used to assess whether the average *number of publications produced per award per year* and the average *total direct cost spent per publication per award* differed by *group*. The number of publications produced per award per year was calculated by taking the total number of publications produced by an award divided by the award duration (in years). The total direct cost spent per publication per award was calculated by dividing the total direct cost for an award by the total number of publications produced by that award. Awards with zero publications were removed from both analyses.

Lastly, STPI also assessed whether *time to first publication* as well as *overall time to publication* (i.e., the rate at which articles were published) were significantly different by *group*. Publications that provided at least the month and year for the publication date were included in this analysis. *Time to first publication* was calculated as the number of days between the project start date and the publication date for the first article attributed to an award. Articles published after the first publication were removed from this analysis. *Overall time to publication* was calculated as the number of days between the project start date and the publication date for each article attributed to an award. For publications in which only the month and year were provided, the *day* unit was set to the first of the month. A Cox proportional hazard model was used to examine whether *group* was predictive of both *time to first publication* and *overall time to publication*. Awards with zero publications were removed from this analysis.

b. Citation Level Metrics

Differences between TRA and NDPA awards were assessed for the following citation level metrics:

- *total number of citations received per publication,*
- *total number of citations received per award,*
- *average number of citations received per publication per year,*

[30] Hothorn, T., Bretz, F., Westfall, P., Heiberger, R.M., Schuetzenmeister, A., Scheibe, S. 2020. "Package 'multcomp.' Simultaneous inference in general parametric models. Version 1.4-13." Available at http://multcomp.R-forge.R-project.org.

- *average total direct cost spent per citation received per award*, and
- *average RCR*
- *weighted RCR.*

The total number of citations received per publication, average number of citations received per publication per year, and RCR are data provided by iCite. An award level weighted RCR was calculated as the number of articles associated with the award multiplied by their average RCR. STPI calculated the average total direct cost spent per citation per award by dividing the total direct cost of an award by the total number of citations received across all publications that fell within the specified time frame described above for an award.

A GLM with a Poisson distribution was used to assess whether the *total number of citations received per publication* (dependent variable) differed by *group* (explanatory variable). Similarly, a GLM with a Poisson distribution was also used to assess whether the *total number of citations received per award* (dependent variable) differed by *group* (explanatory variable) at the grant level. STPI performed a multi-variable GLM regression analysis at the grant level to take into consideration the effects of the same variables listed in the publication analysis (*group, award duration, total direct cost (per $100,000 spent),* whether the award had *multi-PI,* and *area of science*). A separate multi-variable GLM regression analysis was performed at the publication level to account for *year of publication* (continuous variable) and *total number of authors* (continuous variable) as potential variables on top of *group* and whether the award had *multi-PI.* A type-II sum of squares analysis of deviance was performed for each GLM to assess which explanatory variable(s) is predictive of the total number of citations received. To determine which areas of science differed significantly froms one another, pairwise comparisons among the different areas of science were performed using a post-hoc Tukey test with the *glht* function from the *multcomp* package in R (Hothorn et al. 2020).

A non-parametric Kruskal-Wallis rank sum test was used to assess whether the *average number of citations received per publication per year, average total direct cost spent per citation received per award,* and *average RCR* differed by *group*.

STPI also assessed whether *time to first citation* differed significantly by *group*.[31] *Time to first citation* (in days) was calculated by (1) an award's project start date and (2) by the publication date for the article that received the citation. For an award's project start date, *time to first citation* was calculated as the number of days between the publication date of an award's first citation and an award's project start date. In other words, *time to first citation* was considered as the first citation of an award and not as the first citation of

[31] Self citations were included in the citation analyses.

the first publication of that award. In most cases, the first publication of an award received the first forward citation, but this was not true for all cases. For the article that received the first citation of an award, *time to first citation* was calculated as the number of days between an article's publication date and the publication date of an award's first citation.

Lastly, STPI considered the rate at which citations were accumulated (i.e., *time to citation*). *Time to citation* was calculated for each citation that an article received as the number of days between the publication date of the citation and (1) the award's project start date of the article that received the citation, and (2) the publication date of the article that received the citation.

All negative *time to first citation* and *time to citation* values (i.e., a study was cited before it was formally published) were removed from the analysis. For each case, a Cox proportional hazard model was used to assess whether the citation rate was significantly affected by *group*. Publications that provided at least the month and year for the publication date were included in this analysis.

Awards with zero publications were removed from all citation analyses. Publications with citation data listed as *NA* were removed on an individual analysis by analysis basis as some citation data would have a numeric value listed (e.g., total number of citations received) but others would not (e.g., RCR).

c. Altmetrics

Altmetric data complement traditional, citation-based metrics and include citations on Wikipedia as well as public policy documents; discussions on research blogs; coverage on mainstream media; and mentions on social media such as Twitter and Facebook (Altmetric 2020a). Whereas traditional citations provide information on the research impact of an article on the academic community, altmetric data also take into consideration how widely disseminated an article is beyond the publishing journal and immediate scientific community and how much attention an article receives from the public sphere (Altmetric 2020a). Similarly, because of the lag time between article submission and actual publication, it takes time for articles to accumulate citations and therefore, there is also an associated lag time in the ability to measure the immediate impact of an article using traditional, citation-based metrics. Altmetric data, by virtue of being sourced from the internet, allow for faster assessment of research impact.

There are, as with any metric, limitations to the use of altmetric data. Altmetric data are a complement to, and not a replacement for, traditional, journal-based citations. It is important to consider altmetric data in context such as understanding where the underlying data come from (e.g., which sources are discussing the article of interest, what the sources are saying about the article of interest). Lastly, altmetric data are still relatively new and more research is needed to better understand the use and interpretation of altmetrics. To

prevent individuals from artificially inflating the altmetric score for an article, companies that gather such data use algorithms to identify and correct for artificial inflation.

To gather altmetric data, STPI queried the Altmetric database using the rAltmetric package in R (Ram 2017). Articles were identified using their PubMedID (PMID). For the altmetric analysis, we focused on the Altmetric attention score as the response variable of interest. The Altmetric attention score is an automatically calculated, weighted count of all of the attention a research output has received and is based on three main factors: the volume of attention or mentions that a research output receives; the source that mentioned the research output; and how often authors of each mention talk about the scholarly articles. Each of these factors is weighted accordingly (Altmetric 2020b). For volume, a mention is only counted once from each person per source so that if the same person tweets about the same paper more than once, only the first mention will be counted towards the Altmetric score. Different sources contribute differently to the Altmetric attention score; reputable sources, such as a newspaper article, contribute more than a blog post, which contributes more than a tweet.[32,33] And lastly for authors, a mention about an article that is shared by a researcher to other researchers counts more than a journal account sharing the same article link automatically. More generally, the Altmetric attention score is a metric for the amount of online activity a research output receives and is not necessarily a metric for the quality of the research, or the researcher, as mentions may be both negative or positive. STPI used the Altmetric attention score as a metric for broad research impact and to identify articles that received exceptional online coverage that may be of interest to the NIH HRHR program.

A Kruskal-Wallis test was used to assess whether the Altmetric attention score differed significantly between TRA and NDPA awards and a quantile-quantile (QQ) plot was used to compare the distribution of Altmetric attention scores between the two groups. In addition, STPI tested whether the Altmetric attention scores were significantly different between TRA and NDPA publications for each area of science using a Kruskal-Wallis test.

2. Results

As a reminder, 57 TRA awards were granted between FY2010 and FY2012, and 58 NDPA awards were granted between FY2009 and FY2012. Five TRA and six NDPA awards were excluded from all analyses because they had PIs who had active, overlapping TRA and NDPA awards at the same time. All analyses are based on the 52 remaining TRA and NDPA awards.

---

[32] The standard weightings for each mention type used by Altmetric can be found here: https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-score-calculated-

[33] Ram, K. 2017. "rAltmetric v 0.7.0: retrieves altmetrics data for any published paper from Altmetric.com." Available at: https://github.com/ropensci/rAltmetric.

a. Publications

Overall, 51 of the 52 awards from both TRA and NDPA reported producing at least one publication. In total, the 51 TRA and NDPA awards produced 1,089 and 970 publications, respectively. After limiting publications to those published after a year of an award's project start date to within a year of an award's project end date, 50 of the 52 TRA and NDPA awards had at least one publication that fell within the specified time frame, resulting in a total of 882 TRA and 777 NDPA publications that were included in STPI's analyses (Table 2).

Table 2. Summary Statistics of Bibliometric Data by *Group*

Group	Total number of awards	Number (percent) of awards with at least one publication	Total number of publications ǂ	Mean (± SE) number of publications/award *
TRA	52	50 (96.2%)	882	17.0 (± 1.83)
NDPA	52	50 (96.2%)	777	14.9 (± 1.54)

Source: publications were downloaded from QVR (March 2020) and citations to those publications were downloaded from iCite (May 2020)

Note: ǂ Only publications that had publication dates after a year of an award's project start date and within a year of an award's project end date are included in STPI's analyses.

Note: * Includes awards with 0 publications.

Overall, TRA awards (including those with zero publications), on average (± SE), produced a significantly higher number of publications (17.0 ± 1.83) compared to NDPA (14.9 ± 1.54) awards ($\chi^2_1 = 6.65, p = 0.01$; Table 3). No significant difference was detected, however, between TRA and NDPA awards after STPI normalized the number of publications produced by award duration. The average (± SE) number of publications produced per award per year was 3.15 (± 0.30) for TRA awards and 2.77 (± 0.27) for NDPA awards ($\chi^2_1 = 0.76, p = 0.38$). Similarly, there was no significant difference in the average (± SE) total direct cost spent to produce a publication between TRA ($290,450 ± $64,512) and NDPA ($392,518 ± $87,233) awards ($\chi^2_1 = 0.24, p = 0.63$).

Table 3. Mean (± SE) Values on Bibliometric Publication Metrics by *Group*

Metric	TRA	NDPA	χ^2_1	P-value
Number of publications produced per award	17.0 ± 1.83	14.9 ± 1.54	6.65	0.01 *
Number of publications produced per award per year	3.15 ± 0.30	2.77 ± 0.27	0.76	0.38
Total direct cost spent per publication	$290,450 ± $64,512	$392,518 ± $87,233	0.24	0.63

Source*:* publication data were downloaded from QVR (March 2020)

* Significant at $p < 0.05$

** Significant at $p < 0.01$

*** Significant at $p < 0.001$

1) Multi-variable GLM Analysis of Publications at the Award Level

Results from the multi-variable GLM regression analysis and type-II sum of squares analysis of deviance showed that when all other variables were held constant, the total number of publications produced did not differ significantly between TRA and NDPA awards ($p = 0.14$) (Table 4). The number of publications produced, however, was significantly impacted by *award duration, total direct cost (per $100,000 spent)*, research *area of science* and whether an award had *multi-PIs* ($p < 0.001$ for each).

Table 4. Results from the Type-II Sum of Squares Analysis of Deviance for Publications at the Award Level

Variable of interest	Chi-squared test statistic (df)	P-value
Group	2.14 (1)	0.14
Award duration	31.7 (1)	< 0.001 ***
Total direct cost (per $100,000 spent)	55.4 (1)	< 0.001 ***
Multi-PI award	24.1 (1)	< 0.001 ***
Area of science	42.7 (3)	< 0.001 ***

* Significant at $p < 0.05$

** Significant at $p < 0.01$

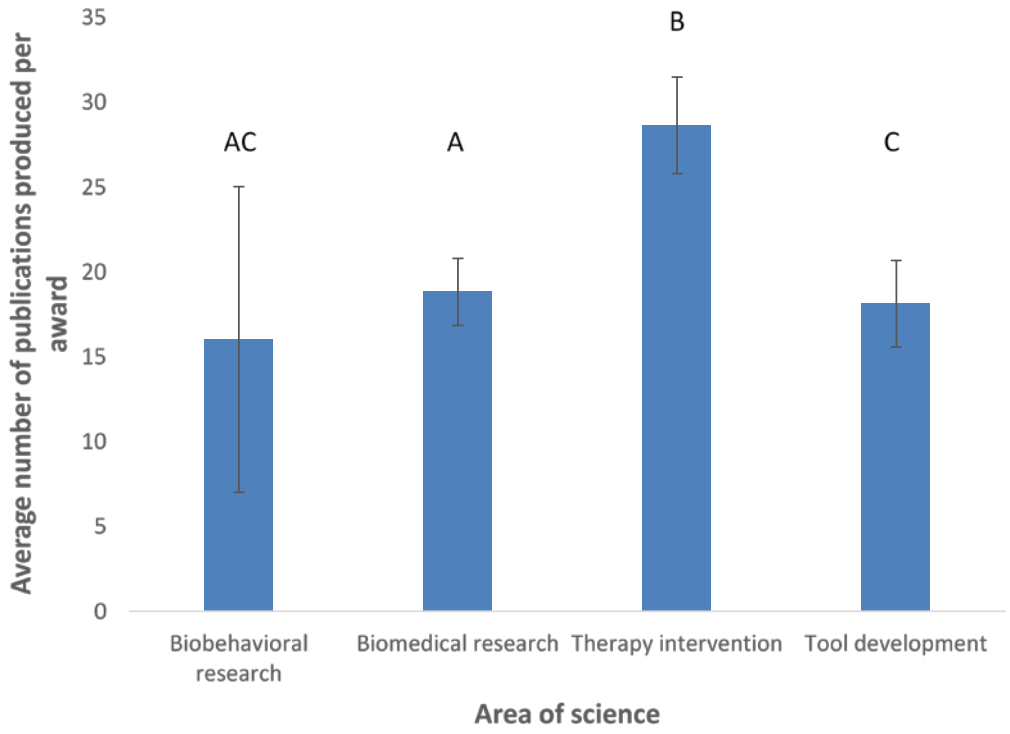
*** Significant at $p < 0.001$

Specifically, when all other variables are held constant, the expected number of publications produced by TRA awards is 8.6% lower relative to NDPA awards, though this decrease is not significant ($p = 0.14$); for every year increase in *award duration*, the expected number of publications produced per award increased by 22%; the expected number of publications produced by awards with multi-PIs increased by 40% compared to awards with single PIs; and the expected number of publications produced per $100,000 spent in total direct cost increased by 0.7% (Table 5). In the GLM analysis, behavioral research was arbitrarily set as the baseline for area of science against which all other areas of science were compared. To assess how each area of science compares to one another, a post-hoc Tukey test was performed.

Table 5. GLM Results for Publications at the Award Level

Variable of interest	Coefficient estimate	Expected increase or decrease in number of publications produced relative to baseline
Group		
NDPA (baseline)	NA	NA
TRA	-0.09	-8.6%
Award duration (continuous variable)	0.20	22%
Total direct cost (per $100,000 spent) (continuous variable)	0.007	0.7%
Multi/Single Award PI		
Single PI (baseline)	NA	NA
Multi-PI	0.34	40%
Area of science		
Behavioral research (baseline)	NA	NA
Biomedical research	0.07	7.8%
Therapy intervention	0.47	59%
Tool development	-0.06	-5.8%

Results from the post-hoc Tukey test on *area of science* showed that the average (± SE) number of publications produced did not differ significantly between awards focused on biobehavioral (16.0 ± 9.00) and biomedical research (18.8 ± 1.98; $p = 0.49$), nor between awards focused on biobehavioral research and tool development (18.1 ± 2.55; $p = 0.95$) (Figure 13). Awards focused on therapy intervention (28.6 ± 2.84) had a significantly higher number of publications than those focused on biobehavioral research ($p = 0.003$); awards focused on therapy intervention produced a significantly higher number of publications than those focused on biomedical research ($p < 0.001$); awards focused on tool development had a significantly higher number of publications than those focused on biomedical research ($p = 0.03$); and awards focused on therapy intervention produced a significantly higher number of publications than those focused on tool development ($p < 0.001$).



Different letters denote significant differences in the number of publications produced between areas of science.

Figure 13. Average (± 1 SE) Number of Publications Produced per Award by Area of Science

2) Time to Publication Analysis

For *time to first publication*, the Cox regression model indicated there was no significant difference in the rate at which the first article of an award was published between TRA and NDPA awards ($p = 0.6$). The hazard ratio[34] (95% confidence interval)[35] for TRA awards compared to NDPA awards was 1.12 (0.75 to 1.67) (Figure 14).

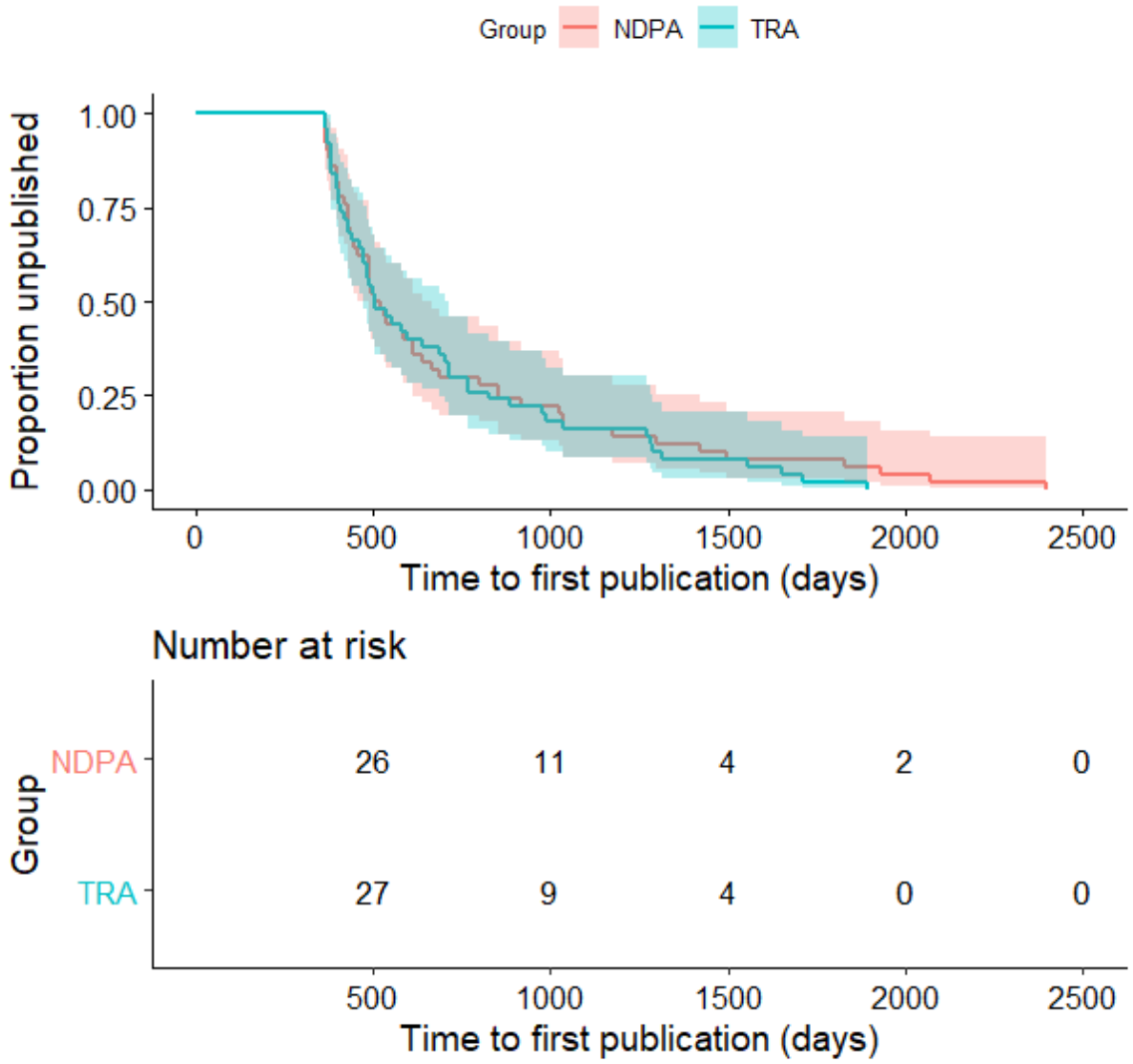


Figure 14. Proportion of Awards That Have Not Published at Least One Article over Time by *Group*

[34] The hazard ratio is the ratio of the probability of an event occurring in a treatment group compared to the probability of an event occurring in a control group.

[35] A confidence interval is a range of values with a specified probability that the true value of the parameter of interest lies within it.

Regarding the publication rate, the Cox regression model for *overall time to publication* indicated that there was no significant difference in the rate at which articles were published between TRA and NDPA awards ($p = 0.3$). The hazard r atio (95% confidence interval) for TRA awards compared to NDPA awards was 1.06 (0.96 to 1.17) (Figure 15).

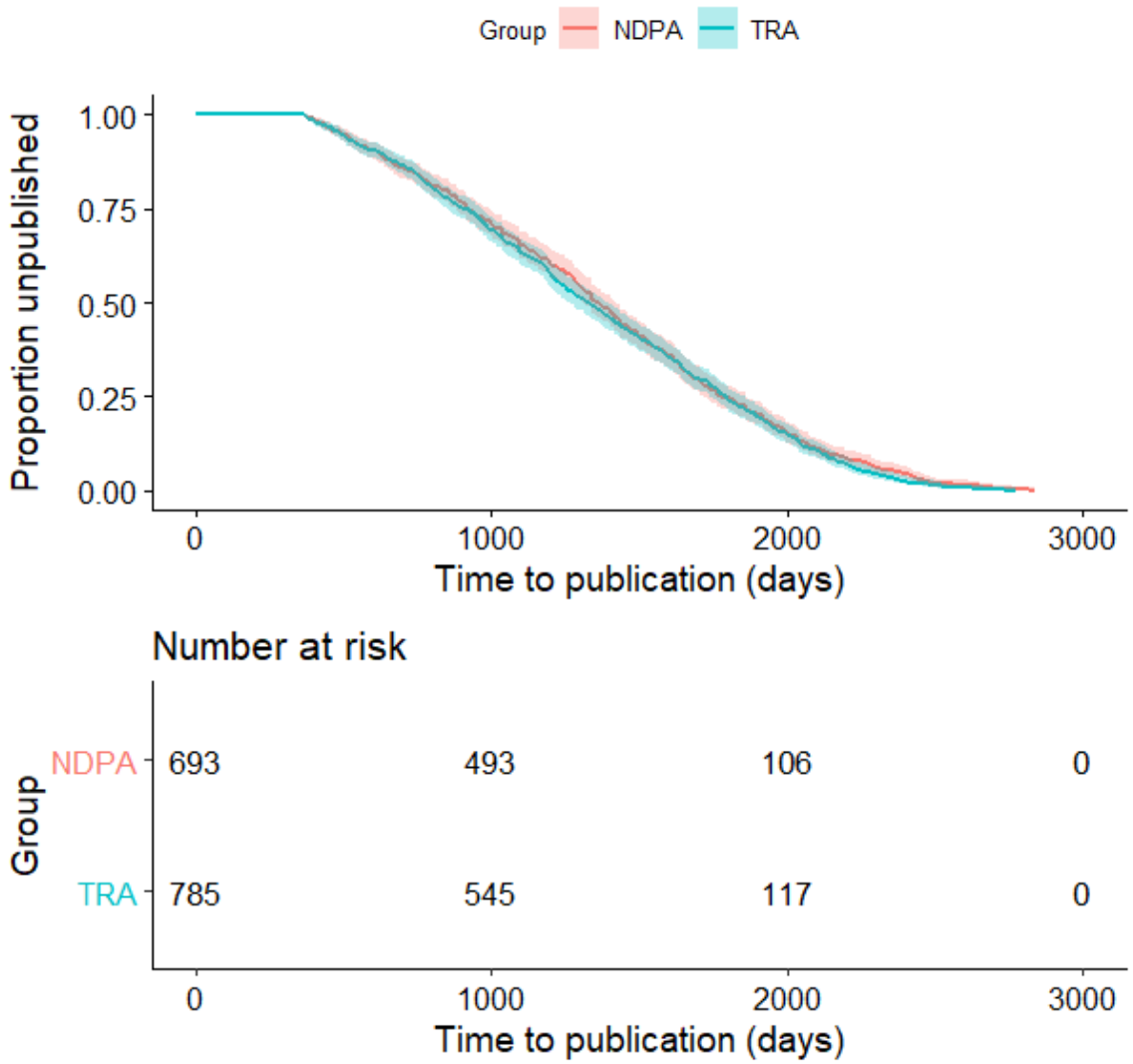


Figure 15. Proportion of Articles Not Published over Time by *Group*

b. Citations

At the time of STPI's analyses, the iCite database only contained citation data for articles published between 1980 and 2019. As a result, iCite was unable to provide citation data for two NDPA publications, which were removed from all citation analyses. Overall, 870 of the 882 TRA publications and 763 of the 777 NDPA publications have received at least one citation (Table 6). As of April 2020, TRA publications have accumulated 59,705 citations, and NDPA publications have accumulated 42,069 citations. On average (± SE), TRA award publications received significantly more citations (67.6 ± 6.04) than NDPA award publications (54.3 ± 3.71; $\chi^2_1 = 1{,}216, p < 0.001$; Table 6). This was true even after the data were normalized to number of citations received per publication per year. TRA award publications, on average (± SE), received 12.9 (± 0.99) citations per year, which was significantly higher than NDPA award publications (8.89 ± 0.53; $\chi^2_1 = 4.97, p = 0.03$). There was no significant difference, however, between TRA and NDPA awards for total direct cost spent for each citation received ($\chi^2_1 = 0.10, p = 0.75$). Each citation, on average (± SE), cost $9,834 (± $1,991) for TRA awards and $25,510 (± $16,679) for NDPA sa wards.

TRA award publications had significantly higher RCRs (4.09 ± 0.29) than NDPA award publications (2.90 ± 0.17; $\chi^2_1 = 9.38, p = 0.002$). However, TRA awards did not have significantly different weighted RCRs compared to NDPA awards (73.2 ± 19.4 and 43.5 ± 7.09 for TRA and NDPA awards, respectively; $\chi^2_1 = 0.78, p = 0.38$).

Table 6. Mean (± SE) Values on Bibliometric Citation Metrics by *Group*

Metric	TRA	NDPA	χ^2_1	P-value
Number of citations received per publication	67.6 ± 6.04	54.3 ± 3.71	1,216	< 0.001 ***
Number of citations received per publication per year	12.9 ± 0.99	8.89 ± 0.53	4.97	0.03 *
Total direct cost spent per citation received	$9,834 ± $1,991	$25,510 ± $16,679	0.10	0.75
Relative Citation Ratio	4.09 ± 0.29	2.90 ± 0.17	9.38	0.002 **
Weighted Relative Citation Ratio	73.2 ± 19.4	43.5 ± 7.09	0.78	0.38

Source: publication data were downloaded from iCite (May 2020)

* Significant at $p < .05$

** Significant at $p < .01$

*** Significant at $p < .001$

1) Multi-variable GLM Analysis of Citations at the Award Level

At the award level, results from the multi-variable GLM regression analysis and type-II sum of squares analysis of deviance showed that *group*, *award duration*, *total direct cost (per $100,000 spent)*, whether an award had *multi-PIs*, and *area of science* all significantly affected the total number of citations received ($p < 0.001$ for each; Table 7).

Table 7. Results from the Type-II Sum of Squares Analysis of Deviance for Citations at the Award Level

Variable of interest	Chi-squared test statistic (df)	P-value
Group	1,938 (1)	< 0.001 ***
Award duration	1,015 (1)	< 0.001 ***
Total direct cost (per $100,000 spent)	28,570 (1)	< 0.001 ***
Multi-PI award	7,323 (1)	< 0.001 ***
Area of science	22,802 (3)	< 0.001 ***

* Significant at $p < 0.05$

** Significant at $p < 0.01$

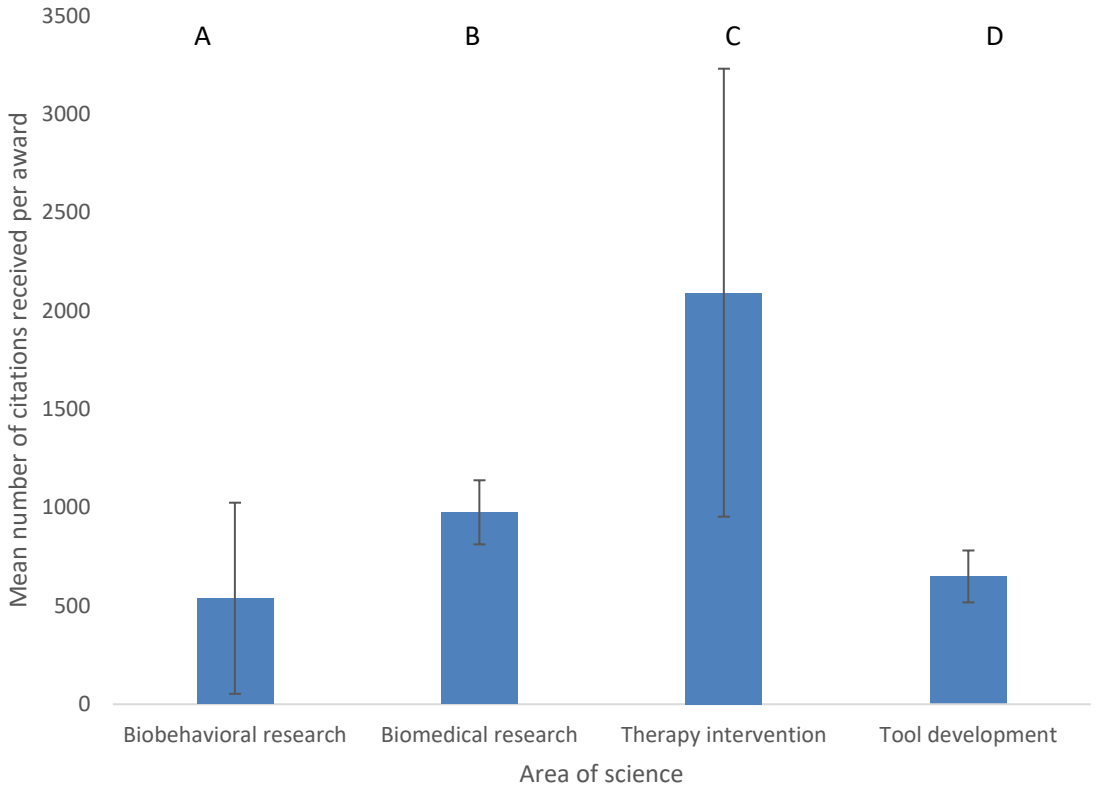
*** Significant at $p < 0.001$

Specifically, when all other variables are held constant, the expected number of citations decreased by 31% for TRA awards relative to NDPA awards; the expected number of citations increased by 17% for every year increase in award duration; and the expected number of citations increased by 106% for awards with multi-PIs relative to those with single PIs; and the expected number of citations received increased by 1.8% for every $100,000 spent in total direct cost (Table 8). In the GLM analysis, behavioral research was arbitrarily set as the baseline for area of science against which all other areas of science were compared. To assess how each area of science compares to one another, a post-hoc Tukey test was performed.

Table 8. GLM Results for Citations at the Award Level

Variable of interest	Coefficient estimate	Expected increase or decrease in number of citations received relative to baseline
Group		
NDPA (baseline)	NA	NA
TRA	-0.37	-31%
Award duration	0.15	17%
Total direct cost (per $100,000 spent)	0.02	1.8%
Multi/Single Award PI		
Single PI (baseline)	NA	NA
Multi-PI	0.72	106%
Area of science		
Behavioral research (baseline)	NA	NA
Biomedical research	0.76	114%
Therapy intervention	1.38	298%
Tool development	-0.19	-17%

Results from the post-hoc Tukey test on *area of science* showed that the average (± SE) number of citations received differed significantly across all areas of science ($p < 0.001$ for each pair-wise comparison; Figure 16). Specifically, awards focused on therapy intervention received, on average (± SE), the highest number of citations per award (2,092 ± 1139), followed by awards focused on biomedical research (976 ± 163), tool development (650 ± 132), and then biobehavioral research (539 ± 486).



Different letters denote significant differences in the number of citations received among different areas of science.

Figure 16. Mean Number of Citations Received per Award by Area of Science

2) Multi-variable GLM Analysis of Citations at the Publication Level

A multi-variable GLM regression analysis was also conducted to assess which factors influenced the number of citations received at the publication level (i.e., total number of citations received per publication). Results from this multi-variable GLM regression analysis and type-II sum of squares analysis of deviance showed that *group*, *year of publication*, whether an award had *multi-PIs*, and the total *number of authors* on a publication all significantly affected the number of citations received per publication ($p < 0.001$ for each; Table 9).

Table 9. Results from the Type-II Sum of Squares Analysis of Deviance for Citations at the Publication Level

Variable of interest	Chi-squared test statistic (df)	P-value
Group	1,384 (1)	< 0.001 ***
Year of publication	21,775 (1)	< 0.001 ***
Multi/Single PI award	1,558 (1)	< 0.001 ***
Number of authors	5,392 (1)	< 0.001 ***

* Significant at $p < 0.05$

** Significant at $p < 0.01$

*** Significant at $p < 0.001$

Specifically, the expected number of citations received per TRA award publication was 35% higher than those received per NDPA award publication; the expected number of citations received per publication was 38% higher for awards with multi-PIs compared to those with single PIs; for every year increase in year of publication, the expected number of citations received per publication decreased by 23%; and for every additional author listed on a publication, the expected number of citations received per publication increased by 2.1% (Table 10).

Table 10. GLM Results for Citations at the Publication Level

Variable of interest	Coefficient estimate	Expected increase or decrease in number of citations received relative to baseline
Group		
NDPA (baseline)	NA	NA
TRA	0.30	35%
Year of publication	-0.26	-23%
Multi/Single PI Award		
Single PI (baseline)	NA	NA
Multi-PI	0.32	38%
Number of authors	0.02	2.1%

3) Time to Citation Analysis

In total, TRA publications were cited by 45,283 unique publications and NDPA publications were cited by 33,619 unique publications. The Cox regression model for *time to first citation* by an award's project start date showed that there was no significant difference in rate at which first citations were received between TRA and NDPA awards ($p = 0.9$). The hazard ratio (95% confidence interval) for TRA awards compared to NDPA awards was 1.02 (0.68 to 1.52) (Figure 17).

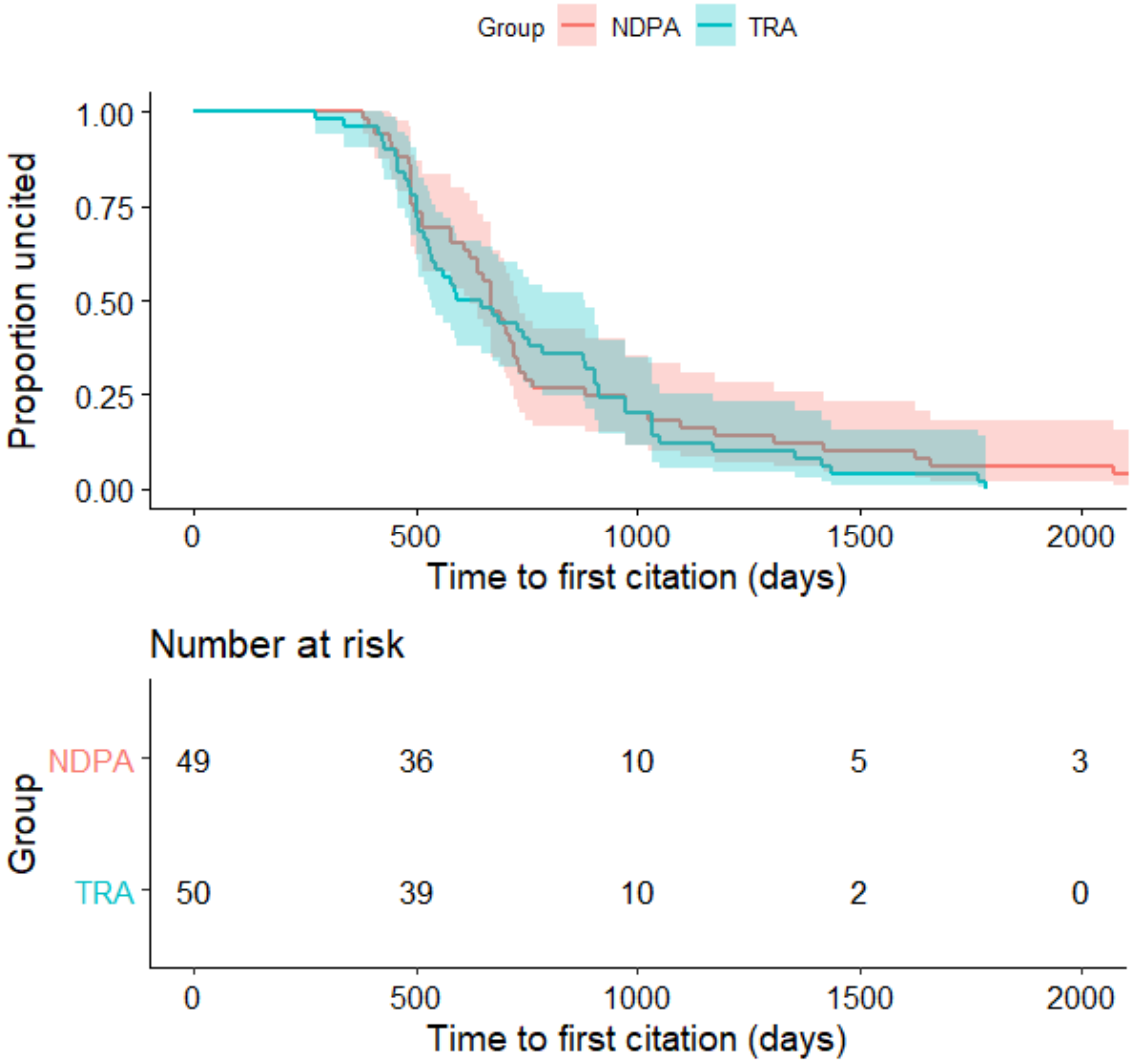


Figure 17. Proportion of Awards with Publications That Are Uncited over Time by *Group* Where Time to First Citation Is Calculated from an Award's Project Start Date

Similarly, the Cox regression model for *time to first citation* by the publication date of the article that received the first citation of an award showed that there was no significant difference in rate at which first citations were received between TRA and NDPA awards ($p = 0.6$) (Figure 18). The hazard ratio (95% confidence interval) for TRA awards compared to NDPA awards was 1.12 (0.75 to 1.66) meaning that TRA awards, on average, received their first citation 12% faster than NDPA awards. This difference, however, was not significant.

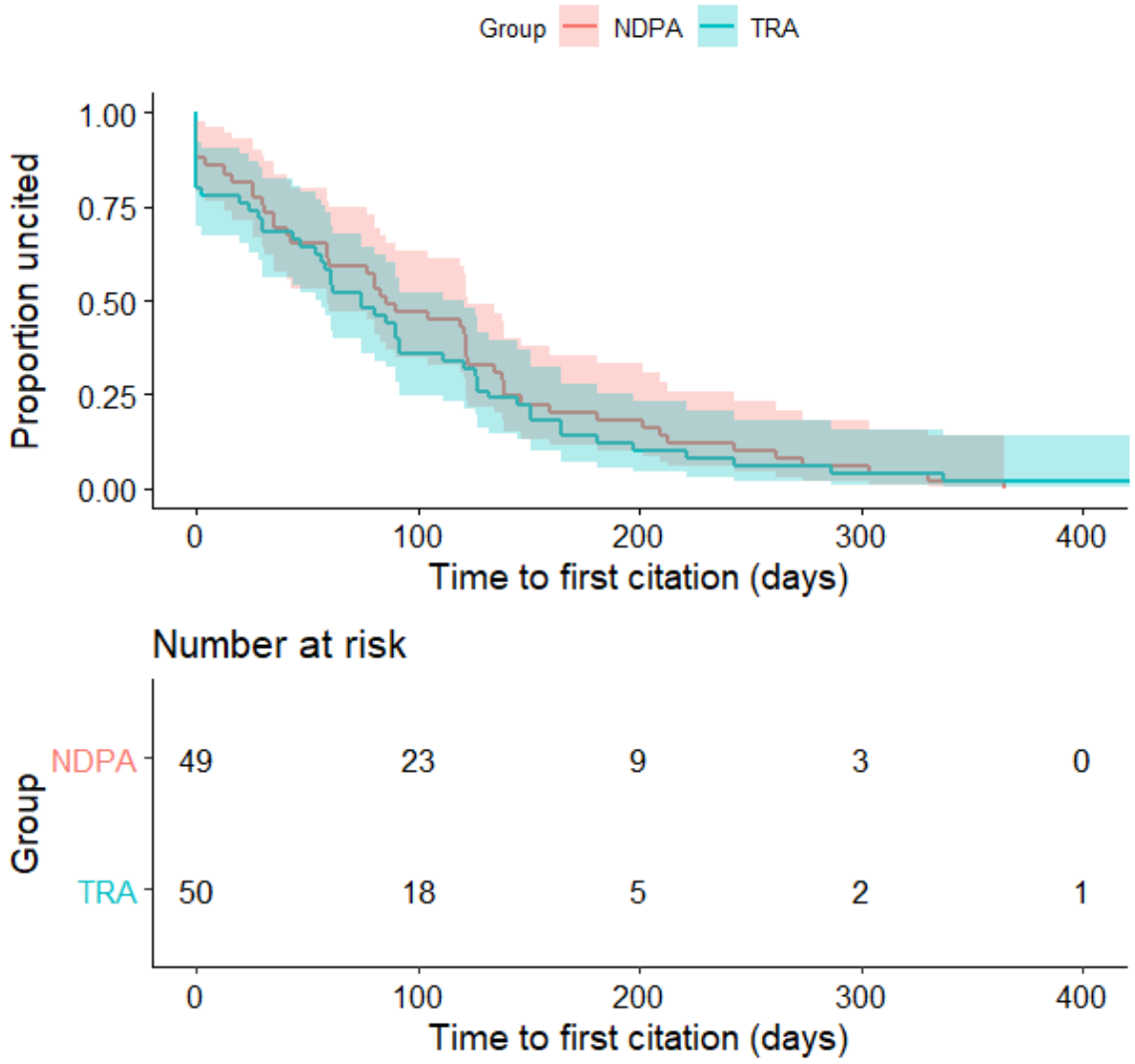


Figure 18. Proportion of Awards with Publications That Are Uncited over Time by *Group* Where Time to First Citation Is Calculated from the Publication Date of the Article That Received the Citation

For *time to citation* by project start date, the Cox regression model showed that the rate at which TRA publications accumulated citations was significantly faster than that of NDPA publications ($p < 0.001$) (Figure 19). The hazard ratio (95% confidence interval) for TRA publications compared to NDPA publications was 1.62 (1.60 to 1.64) meaning that TRA publications, on average, receive citations approximately 62% faster than NDPA publications.

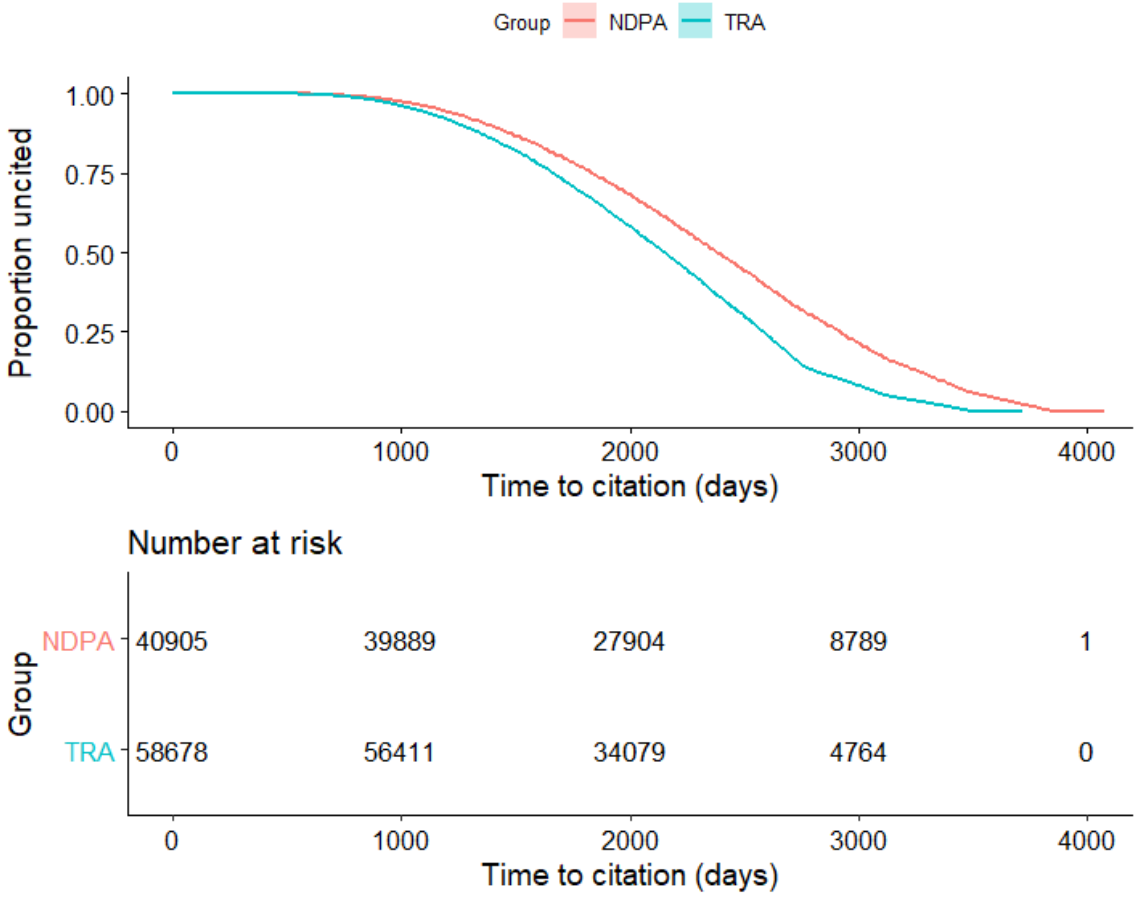


Figure 19. Proportion of Publications Uncited over Time by *Group* Where Time to Citation Is Calculated from an Award's Project Start Date

For time to citation by the publication date of the article that received the citation, the Cox regression model showed that the rate at which TRA publications accumulated citations was significantly faster than that of NDPA publications ($p < 0.001$) (Figure 20). The hazard ratio (95% confidence interval) for TRA publications compared to NDPA publications was 1.26 (1.24 to 1.27) meaning that TRA publications, on average, receive citations approximately 26% faster than NDPA publications.

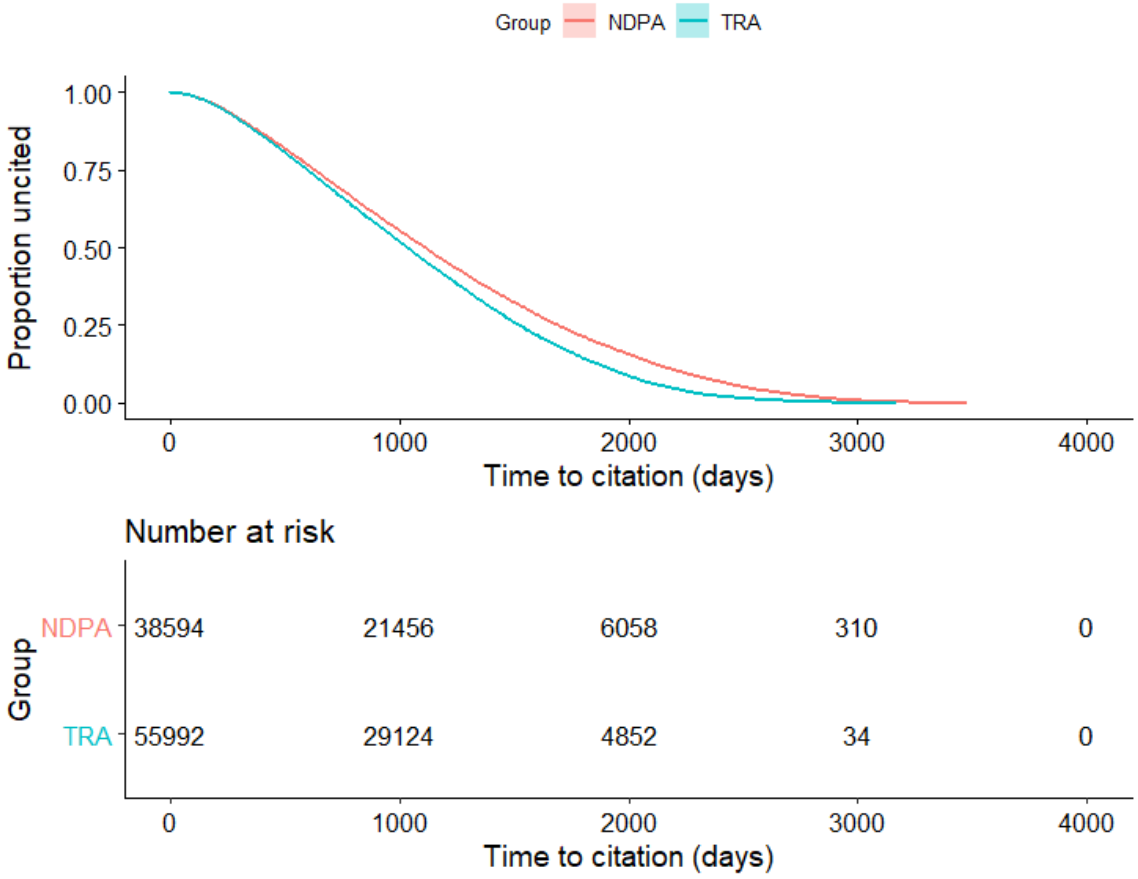


Figure 20. Proportion of Publications Uncited over Time by *Group* Where Time to Citation Is Calculated from the Publication Date of the Article That Received the Citation

c. Altmetrics

The mean (± SE) Altmetric attention score was 47.0 (± 4.10) among TRA publications and 38.6 (± 3.99) among NDPA publications. Results from the Kruskal-Wallis test indicate that TRA publications, on average, have significantly higher Altmetric attention scores than NDPA publications ($\chi_1^2 = 4.63, p = 0.03$). This is corroborated by the QQ plot of Altmetric attention scores for TRA and NDPA publications, which shows that while there are some NDPA publications that received higher than expected Altmetric attention scores, TRA publications, overall, tended to have higher scores than NDPA publications (Figure 21).

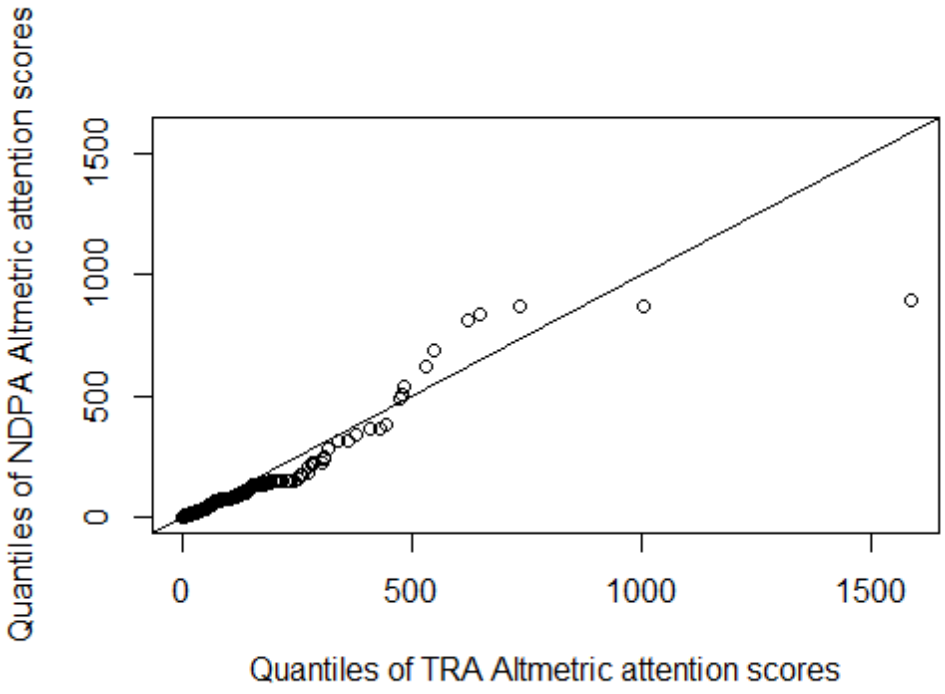


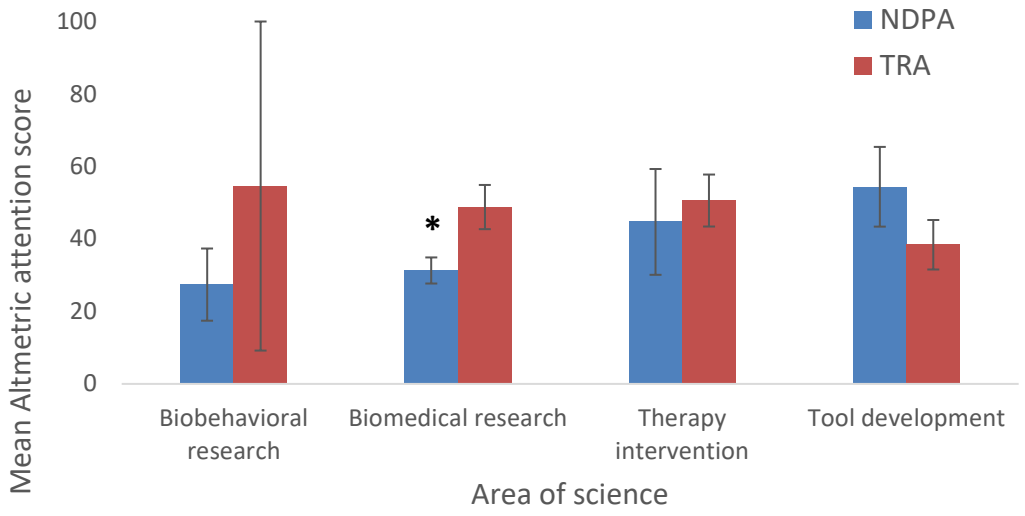
Figure 21. QQ Plot of Altmetric Attention Scores for TRA and NDPA Publications

The QQ plot shows one TRA publication that received much higher online and media attention than all other TRA and NDPA publications. Specifically, the publication of interest (PMID 26828196) is from grant number DK090989 (Disappearing Gastrointestinal Microbiota in Epidemic Obesity) and titled "Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer." This article had an Altmetric attention score of 1,584 and has received coverage by 142 news outlets as of May 2020—including Newsweek, CNN, PBS, the Washington Post, Wired, the Los Angeles Times, and CBS News. Comparatively, it has received 263 total citations since the article was published in 2016 and has a RCR of 25.7. Both the total number of citations and the RCR

value indicate that the article made substantial impact on the scientific community. However, when compared against all publications, there are 65 other publications that had a higher total number of citations, and 22 publications that had higher RCRs. In fact, one TRA publication (PMID 24157548: Genome engineering using the CRISPR-Cas9 system) had the highest total number of citations (3,262) as well as the highest RCR (129) among all TRA and NDPA publications. Conversely, this article has an Altmetric attention score of 180 and has, thus far, received coverage from 11 news outlets since its publication in 2013.

This finding highlights the importance of considering altmetric data when assessing research impact. From a traditional, citation-based viewpoint, the publication on clustered regularly interspaced short palindromic repeat (CRISPR) represents the highest level of scientific findings and research impact. The publication on vaginal microbial transfer, while also notable, would have been considered less significant in comparison when considering traditional metrics such as total citations received and RCRs. However, its broader impact on the general public could have only been brought to the forefront using altmetric data.

We found that Altmetric attention scores were not influenced by the area of science of an award ($\chi_3^2 = 5.23, p = 0.16$). The average Altmetric attention score is 32.3 (± 11.2) for behavioral research, 40.6 (± 3.7) for biomedical research, 48.4 (± 7.1) for therapy interventions, and 46.4 (± 6.5) for tool development publications. When considering area of science by group, we found no significant differences in Altmetric attention scores between TRA and NDPA for biobehavioral research ($\chi_1^2 < 0.01, p = 0.95$; average ± SE Altmetric attention scores were 54.6 ± 45.4 and 27.4 ± 10.0, respectively, for TRA and NDPA awards), therapy intervention ($\chi_1^2 = 3.16, p = 0.08$; 50.6 ± 7.2 and 44.7 ± 14.6, respectively), and tool development publications ($\chi_1^2 = 0.47, p = 0.49$; 38.4 ± 6.8 and 54.4 ± 11.0) (Figure 22). The only exception is in biomedical research, where TRA (48.8 ± 6.1) publications had significantly higher Altmetric attention scores than NDPA (31.3 ± 3.6) publications ($\chi_1^2 = 4.21, p = 0.04$).



* denotes significant difference (p < 0.05) in Altmetric attention score between TRA and NDPA publications

Figure 22. Mean (± SE) Altmetric Attention Score by Group and Area of Science

3. Results for the TRA-R01 Analysis

A. Awardee Survey

To assess TRA's research impacts, STPI created an online survey to solicit awardees' perspectives on their TRA and R01 awards. The survey questions and survey administration were nearly identical for each group, and differences between the two surveys are discussed below. All survey materials can be found in Appendix C.

1. Survey Development

The STPI team developed survey questions that addressed the multiple aspects of paradigm shifting, and innovative/impactful research. The survey items were iterated with the NIH HRHR team, and content and format were tested through a STPI focus group. A summary of the survey questions by *paradigm shifting* and *innovative/impactful* categories is provided in Appendix B.

2. Survey Administration

STPI developed and administered the surveys in Alchemer, a web-based survey platform. Each awardee received a personalized invitation to take the survey; responses were kept confidential, and only aggregate results were provided to NIH. In tandem with the survey, TRA recipients received an email from HRHR staff underlining the importance of the study and asking them to participate in the survey. The survey was sent to the 57 TRA and the first wave of 57 R01 awardees on December 3, 2019. The first reminder to this initial set was sent 1 week later, on December 10, 2019, and the second reminder was sent 1 week after that, on December 17, 2019. The last reminder was sent after the holidays, on January 7, 2020. Due to the lower response rate for R01 awardees, three rounds of invitations were sent. The invitation for the second wave of 55 R01 awardees was sent out on December 12, 2020, and the third wave of 56 R01 awardees was sent on January 7, 2020, with a similar spacing for their subsequent reminders as the first round of R01 invitations. The survey remained open until February 19, 2020.

3. Survey Analysis

Analysis was performed on completed surveys. Descriptive statistics such as the number of responses for each question and the percentage of survey respondents selecting each answer choice are provided for each question. Free response questions were coded

into a series of qualitative categories and are reported as counts for each category. A list of survey questions along with the respective response choices can be found inAppendix C.

Two-sample proportion tests[36] were used for questions with answer choices of *agree* and *disagree* to examine whether the percentage of survey respondents who selected *agree* differed between TRA and R01 awardees. Because there are only two answer choices, STPI did not perform a two-sample proportion test for those who selected *disagree* as it is simply the complement of those who selected *disagree* and would result in the same statistical significance. For survey questions that had Likert scale response choices (e.g., *strongly disagree* to *strongly agree*), positive answer choices were summed and a two-sample proportion test was performed to compare whether the percentage of survey respondents who responded positively differed between TRA and R01 awardees. A two-sample proportion test was also performed for *likely/unlikely* questions to compare the percentage of TRA and R01 respondents who selected *likely* for that item, as well as for questions that were select all that apply.

All statistically significant findings in this analysis are significant at $p < 0.05$. Additional descriptive statistics such as counts and percentage of survey respondents responding to each answer choice are also included in the analysis. Complete data, including all statistical tests, can be found in Appendix D.

4. Results

a. Response Rate

All 57 TRA awardees were eligible for the TRA-R01 comparison.[37] Twenty-four of the 57 TRA awardees completed the survey (42% response rate). One hundred and sixty-nine R01 awardees were eligible for the TRA-R01 comparison. Thirty-five of the 169 R01 awardees completed the survey (21% response rate).[38] More R01 awardees than TRA awardees received an invitation to participate in the survey to obtain approximately the same number of R01 respondents as TRA respondents. Full data tables of survey results are provided in Appendix D.

---

[36] Yates' continuity correction was applied automatically in R for all two-sample proportion tests to account for the small sample sizes of awardee respondents.

[37] As noted in the comparison group development section of the report, STPI excluded five TRA investigators who received an NDPA during the same years for which they had a TRA award. Fifty-two awards are included in this evaluation. For the TRA-R01 comparison, all 57 TRA awards were included.

[38] There was one partial (i.e., incomplete) survey response from TRA awardees, and nine partial survey responses from R01 awardees. These 10 partial survey responses were not included in any of the survey analyses.

b. Descriptive Statistics

1) Paradigm Shifting

The first set of questions addressed the paradigm-shifting nature of the grantees' research. These questions included novel inventions, novel combinations of ideas and approaches, or use of technology that significantly changed research practice or thinking in their field (Figure 23). TRA respondents were significantly more likely than R01 respondents to agree with the statements, *My research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice* (100% and 69% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 6.99$, $p = 0.01$); and *My research furthered existing practices or thinking in my field.* (82% and 47%; $\chi_1^2 = 5.39$, $p = 0.02$).

R01 respondents were significantly more likely than TRA respondents to agree with the statement, *My research aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice* (68% and 94% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 0.68$, $p = 0.02$).

TRA and R01 survey respondents reported at similar levels that they agreed with the statements, *One or more of my research ideas challenged existing science/technology paradigms.* (83% and 88% for TRA and R01 respondents, respectively; $\chi_1^2 = 0.02$, $p = 0.88$); *My research led to a novel invention or a new technology which significantly improved current practices*. (91% and 77%; $\chi_1^2 = 1.19$, $p = 0.28$); *My research led to the development of a new methodology that significantly changed research in my field.* (71% and 53%; $\chi_1^2 = 1.21$, $p = 0.27$); *My research took the next steps in my established area of investigation.* (70% and 57%; $\chi_1^2 = 0.46$, $p = 0.5$); and *My research required the use of equipment, technique or model that was novel and significantly changed research in my field.* (74% and 66%; $\chi_1^2 = 0.14$, $p = 0.71$).

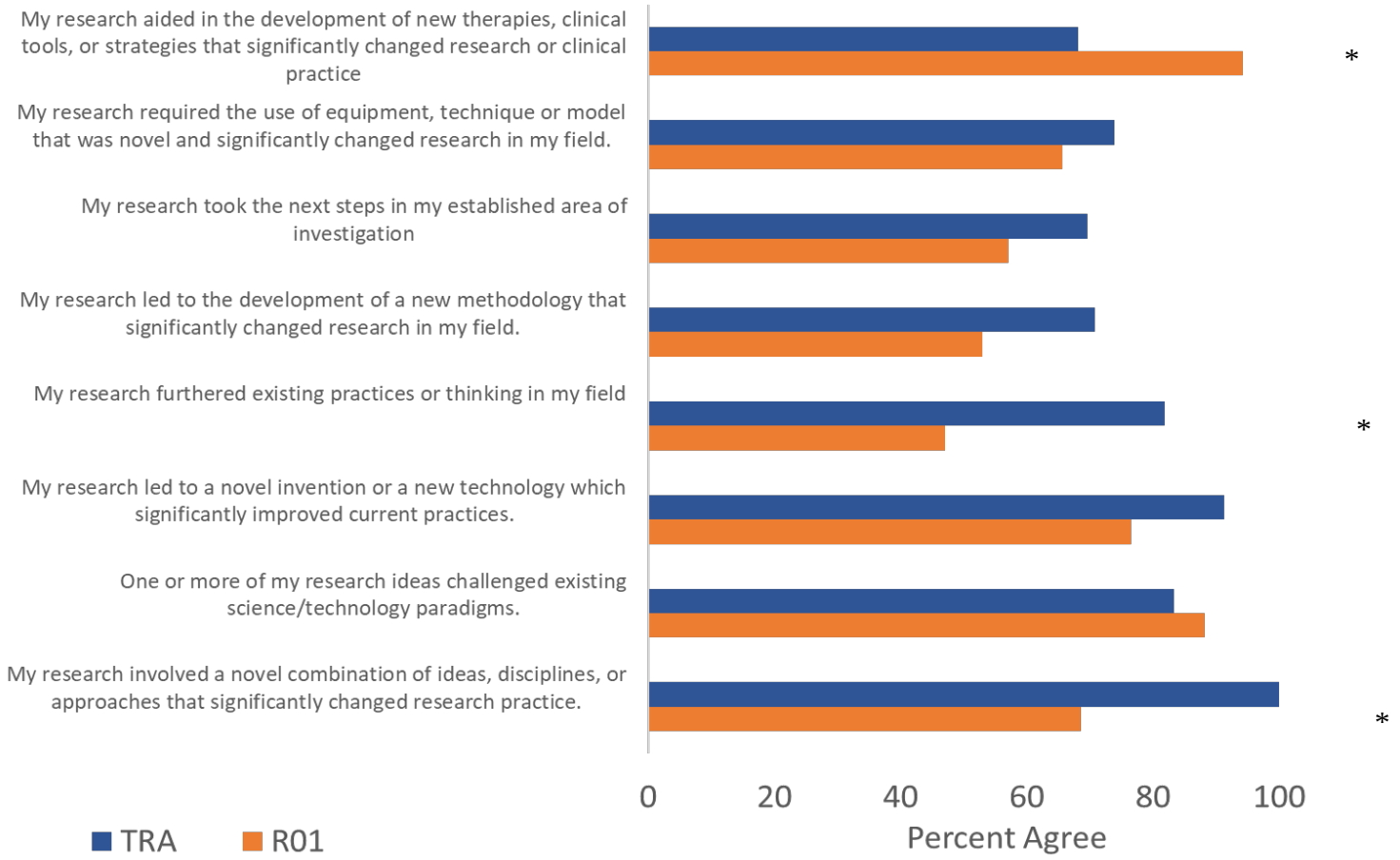


Figure 23. Paradigm-Shifting Research Outputs by Type

2) Innovative and Impactful

This question block explored the timing of impact (Figure 24). TRA respondents were significantly more likely than R01 respondents to agree with the statement, *My research resulted in a delayed but significant shift in practices or thinking* (64% and 30% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 4.68$, $p = .03$).

TRA and R01 survey respondents reported at similar levels that they agreed with the statements, *My research resulted in an immediate and significant shift in practices or thinking* (26% and 36% for TRA and R01 survey respondents, respectively; $\chi_1^2 = .27$, $p = .6$); *My research resulted in a gradual but building shift in practices or thinking* (62% and 80%; $\chi_1^2 = 1.22$, $p = .27$).

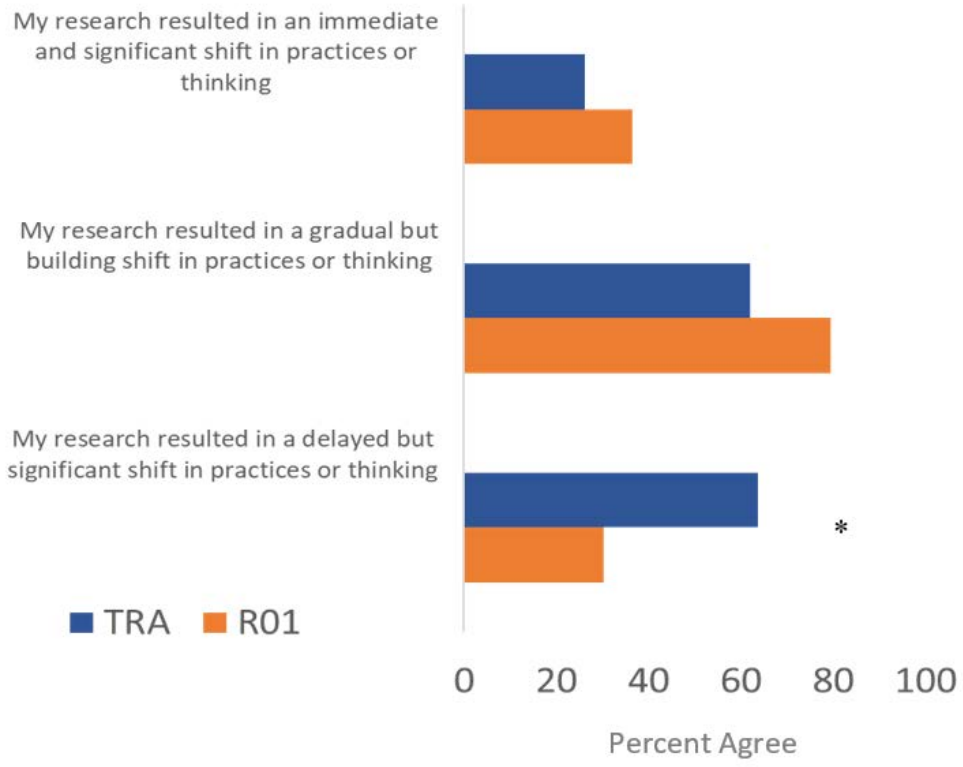


Figure 24. Timing of the Onset of Impact

The next set of questions asks respondents about the likelihood of their research resulting in several different outcomes (Figure 25). TRA respondents were more likely to indicate that their research would likely result in, *The discovery of new phenomenon or advancement of a theoretical concept* (71% and 37% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 5.19$, $p = .02$).

TRA and R01 respondents were not significantly different in responding "likely" that their research would result in the following statements, *A new synthesis of disparate ideas* (50% and 32% for TRA and R01 survey respondents responding "likely," respectively; $\chi_1^2 = 1.17$, $p = .28$), *A change in how research is conducted* (83% and 77%; $\chi_1^2 = 0.06$, $p = .8$), *The development of new therapies, clinical tools, or strategies* (71% and 54%; $\chi_1^2 = 1.02$, $p = .31$), *The development of a new technology* (71% and 66%; $\chi_1^2 = 0.13$, $p = .72$), and *The development of a new methodology* (70% and 68%; $\chi_1^2 = 0.02$, $p = .9$).

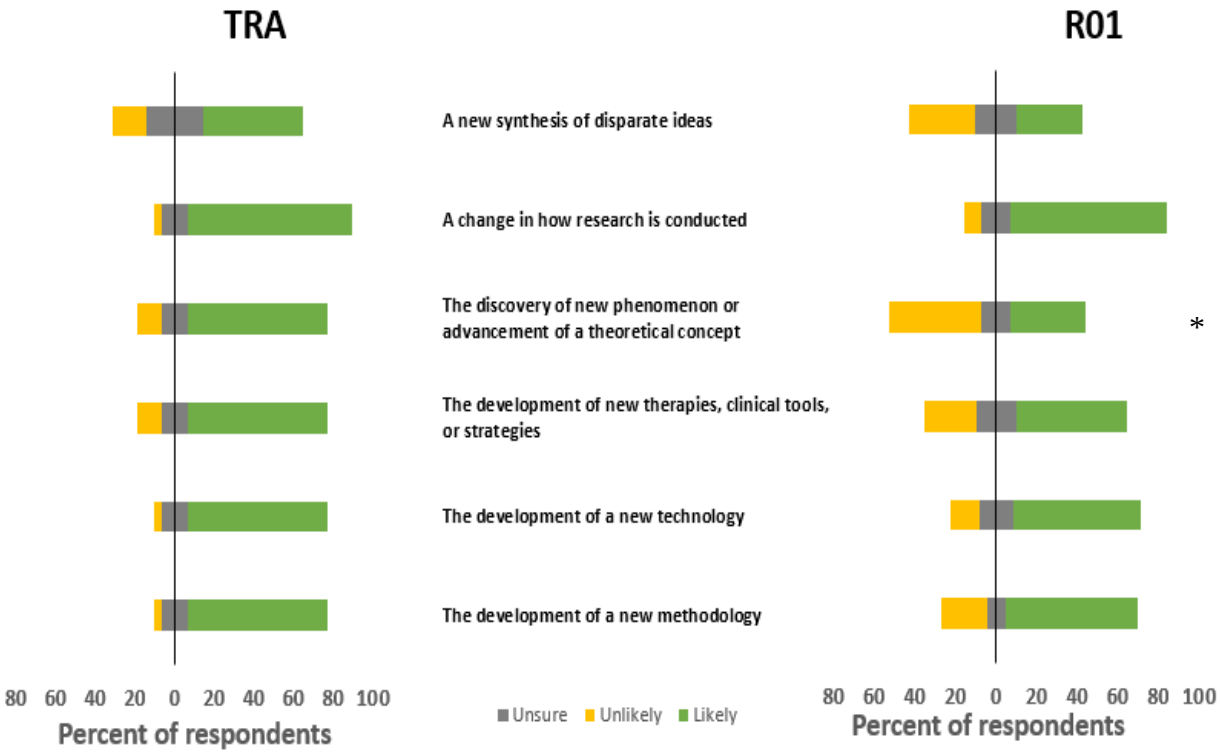


Figure 25. Likeliness of Research Resulting in Impact

This question block examined achievements and honors grantees received once their award funded research was published (Figure 26). This was a select all that apply question. TRA survey respondents were significantly more likely than R01 respondents to select *I have been invited to serve as a regular journal reviewer* (91% and 38% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 7.65, p = .005$).

TRA and R01 respondents survey respondents reported at similar levels that the following statements applied: *My research has been listed as a highlight in an academic journal* (78% and 74% for TRA and R01 survey respondents, respectively; $\chi_1^2 = .009, p = .93$)*; I received an award/honor* (100% and 85%; $\chi_1^2 = 2.10, p = .15$); *I have been asked to give an invited presentation about my research* (65% and 56%; $\chi_1^2 = .18, p = .67$); *My research has been featured in the popular press/media* (74% and 65%; $\chi_1^2 = .20, p = .66$); and *I have been invited to be a keynote speaker to share my research* (44% and 38%; $\chi_1^2 = .01, p = .90$).

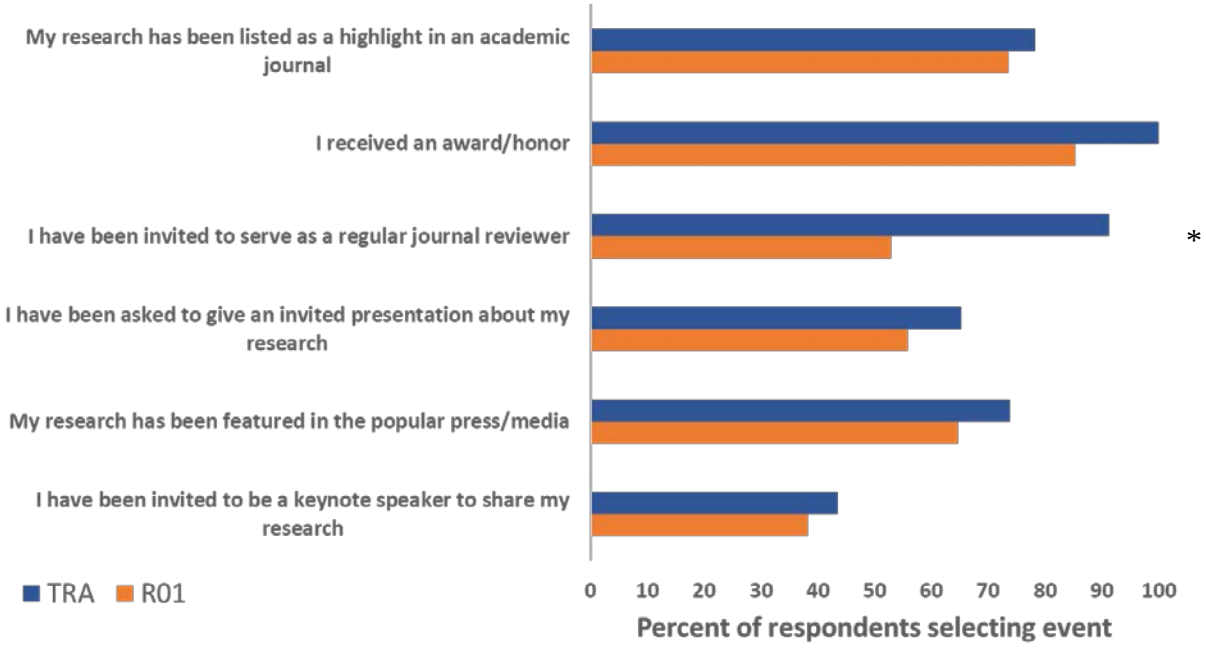


Figure 26. Honors or Awards as a Measure of Research Publication Impact

Next, respondents were asked whether these achievements were attributable to their award (Figure 27). This was a select all that apply question populated by the responses given in the previous question (Figure 26). TRA and R01 survey respondents reported at similar levels that the following applied: *My research has been listed as a highlight in an academic journal* (50% and 56% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 0.0006$, $p = 0.94$)*; I received an award/honor* (91% and 97%; $\chi_1^2 = 0.04$, $p = 0.84$); *I have been invited to serve as a regular journal reviewer* (91% and 94%; $\chi_1^2 < 0.001$, $p = 1$); *I have been asked to give an invited presentation about my research* (80% and 84%; $\chi_1^2 < .0001$, $p = 1$); *My research has been featured in the popular press/media* (88% and 91%; $\chi_1^2 < 0.001$, $p = 1$); and *I have been invited to be a keynote speaker to share my research* (70% and 85%; $\chi_1^2 = 0.11$, $p = 0.74$).

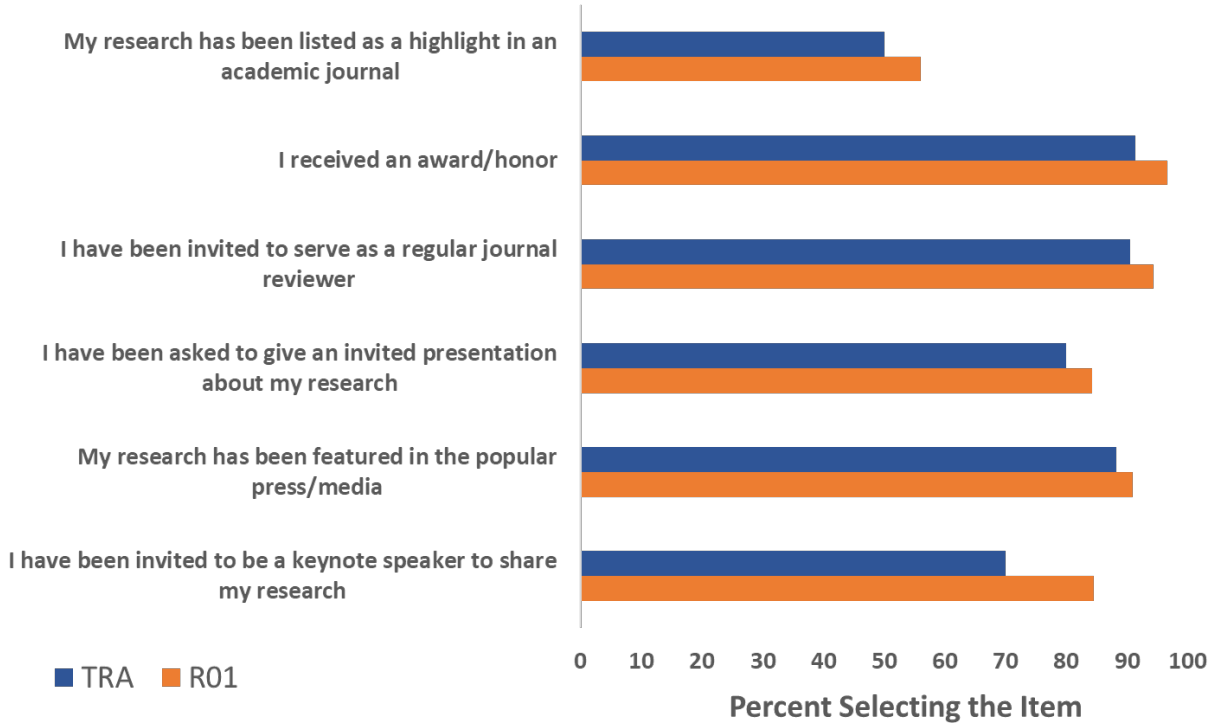


Figure 27. Honor or Awards Attributed to TRA/R01 Research Outputs

This question block examined the flexibility of the grant mechanism (Figure 28). Respondents indicated whether they strongly disagree, disagree, neither agreed nor disagree, agree or strongly agree. Significance tests were performed by combining agree and strongly agree items.[39] R01 respondents were significantly more likely than TRA respondents to select agree or strongly agree to the statement, *The period of the grant was long enough to redirect research as ideas/methods evolved* (96% and 54% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 6.35, p = 0.01$).

TRA and R01 respondents survey respondents selected agree or strongly agree at similar levels that the following statements: *The grant allowed me the freedom to pursue nontraditional research* (71% and 54% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 3.18, p = 0.07$)*;* and *The grant allowed for flexibility in the use of funding* (96% and 69%; $\chi_1^2 = 0.68, p = 0.41$).

[39] No statistical analyses were performed on combining the *strongly disagree* and *disagree* answer choices because the number of respondents who selected these answer choices were too few for any statistical tests to be meaningful

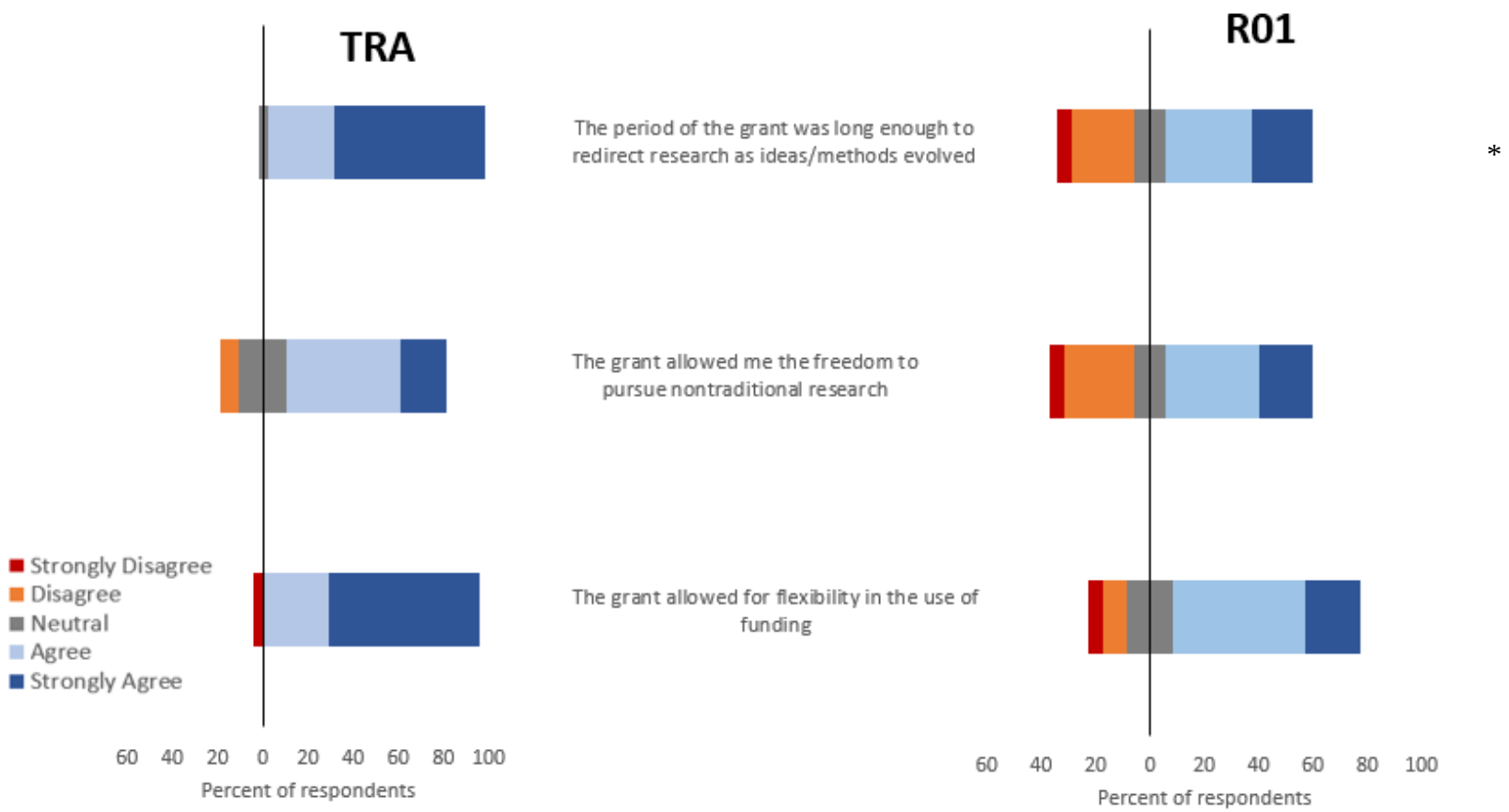


Figure 28. Flexibility of Grant Mechanisms

The next two questions ask about changes to their research plan, and whether the grantees are continuing the trajectory of their research (Figure 29). Both TRA and R01 survey respondents agreed at similar levels that the following occurred: *My research plan changed significantly from what I originally proposed* (42% and 40% for TRA and R01 survey respondents, respectively; $\chi_1^2 = 0$, $p = 1$), and *I am continuing or planning to continue the trajectory of my TRA/R01 funded research post-award* (96% and 83%; $\chi_1^2 = 1.22$, $p = 0.27$).

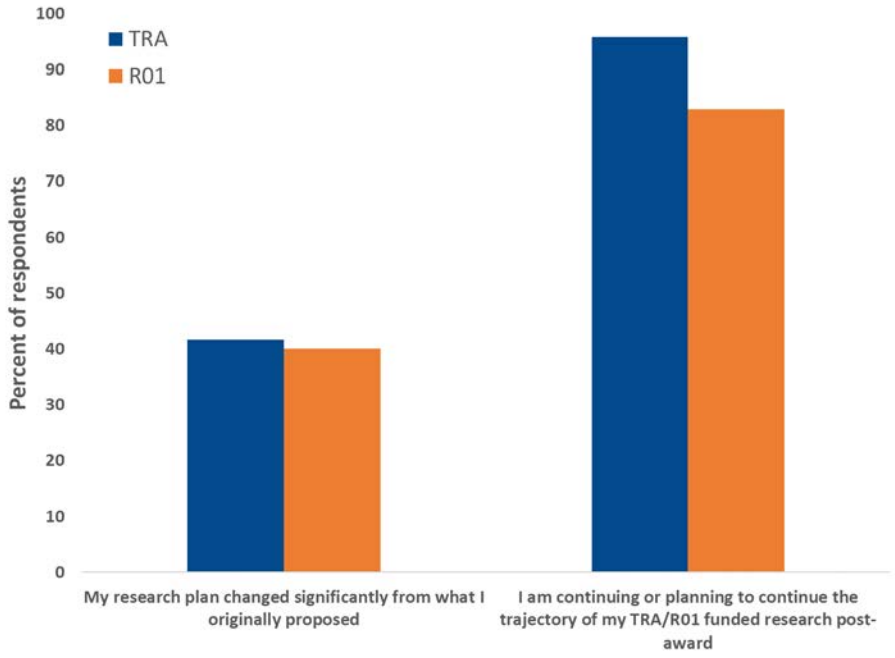


Figure 29. Research Direction

If a respondent indicated that they are planning to continue the trajectory of their research, the last question in this theme asks whether this research is funded (Figure 30). TRA (70%) and R01 respondents (69%) reported similar levels of funding ($\chi_1^2 < 0.001$, $p = 1$) for their post-award research, *You indicated you were continuing or planning to continue the trajectory of your TRA/R01 funded research post-award. Please indicate whether this continuation is funded.*

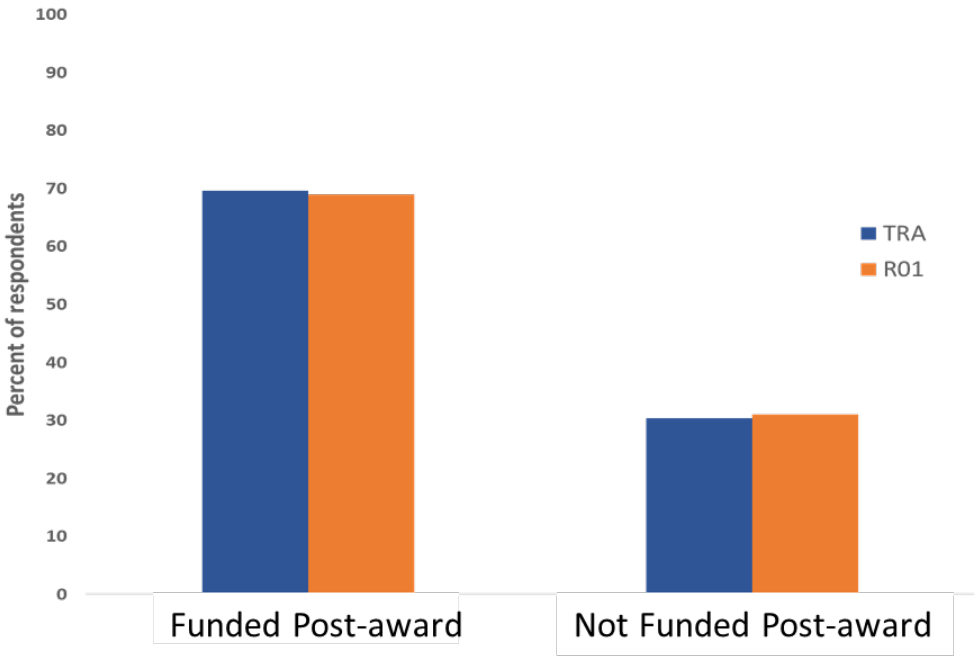


Figure 30. Funding Post-award

c. Qualitative Analysis of Free Response Questions

In addition to the survey questions asked and examined in the above section, respondents were asked to give their thoughts in free response at various stages of the survey. To analyze these data, STPI inductively coded responses. Below is an examination of responses.

1) Changes in Research Plan

If respondents indicated that they agree with the following statement, *My research plan changed significantly from what I originally proposed*, they were given the following open response question: *You indicated that your research plan changed significantly from what you originally proposed. Please indicate why those changes occurred.* Ten TRA awardees responded to this question, as did 13 R01 awardees.

Several TRA (n=4) respondents and R01 (n=7) respondents indicated that results or discoveries led to a new research direction. A TRA respondent reported, “The work was substantially expanded from the original research plan. It was a combination of finding exciting new results that we didn't anticipate, developing new experiments to test them further, and exploring additional multiple independent measures.” One R01 respondent indicated that, “during the course of my research, I developed novel techniques that were not anticipated at the time the grant was originally submitted.”

Both TRA and R01 respondents reported that their new direction was due to technologies that were developed during the course of their grant, (2 and 3, for TRA and R01, respectively). Two R01 respondents reported the development of new technology changed their research direction, with one saying, “Technology became cheaper allowing additional work to be done.” One TRA respondent indicated that the advancement of technology as they advanced the science allowed for new research possibilities, saying, “as we advanced the science the advances improved the technology and opened more doors allowing us to achieve proof of concepts faster.”

Other reasons given for the change in research plan included the following: the initial plan was no longer possible or the best direction, there needed to be a change in methods, and there were unforeseen challenges as well as unanticipated opportunities.

2) Research Trajectory Post-Award

If respondents indicated that they agree with the statement, *I am continuing or planning to continue the trajectory of my TRA/R01 funded research post-award,* they were asked, *You indicated that you are not continuing or planning to continue the trajectory of your TRA/R01 funded research post-award. Please describe what factors went into your decision.* One TRA recipient and six R01 recipients responded to this question. The TRA recipient indicated that they “moved on to new ideas and questions.” Four R01 awardees indicated that they moved from academia into industry. One R01 awardee described difficulty getting additional funding, and another moved to a new position that did not require winning R01s.

3) Other Information

Lastly, all respondents were given the opportunity to answer the question, *Is there any additional information you would like to share?* Fourteen TRA respondents and nine R01 respondents provided information under this question. Six TRA respondents gave positive feedback about their program. A TRA respondent said that, “This is an amazing program that provides tremendous opportunities to move the needle on new and important research fields that run outside of the traditional bounds.” Another TRA respondent indicated that the program “is invaluable to creative research and should be continued.”

One concern that was brought up by two TRA and three R01 respondents was the difficulty renewing funding. One R01 respondent described this frustration, citing the successful publications and resulting from their R01, saying they were, "now extremely frustrated—three failed efforts to renew r01." A TRA respondent suggested an option for an extension after describing difficulties in tradition mechanisms, "It would have been very nice, however, to have had the option to apply for an extension of the TRA research, as it was highly successful and it would have been nice to be able to continue within a program that funds high-risk/high-reward ideas that have proven successful."

Two R01 respondents and one TRA respondent reported issues with the review process for continued funding, with one R01 awardee saying, "(1) it is a lengthy process for continuing funding (2) the review process is sometimes unfair and tainted by the view of some reviewers (3) it appears that the review process is superficial (4) only 'safe' projects seem to receive funding." The TRA awardee commented on the difficulty finding proper review panels, "I believe this work put us far out in front of the field—far enough that it was challenging for review at typical study sections. It integrated multiple fields, which made it challenging to find an appropriate panel."

d. Limitations to the Data

As a reminder, 24 TRA and 35 R01 awardees responded to the survey. Based on these sample sizes, the estimated effect size for a two-tailed alternative hypothesis testing to detect mean differences in two independent groups is $d = 0.76$. With power set at 0.80, this means that STPI could only detect large differences between TRA and R01 survey responses. Consequently, caution should be taken when interpreting survey results. Small sample sizes result in large standard errors, which lead to imprecise estimates of the true effects between groups of interest—in this case, the TRA and R01 awardees. Therefore, the survey results described in this report should not be taken as firm conclusions representing the actual awardee populations. For instance, a lack of statistical significance does not mean there is no effect; it might be the case that there was insufficient power to detect the effect of interest. Similarly, small sample sizes can lead to false-positive results and an overestimation of the magnitude of the relationship between two variables. In other words, a result with statistical significance does not mean there is a true effect between the groups of interest. Overall, careful consideration should be taken to avoid making strong conclusions about the TRA and R01 awardees, regardless of whether survey results yielded statistical significance.

B. Bibliometric Analysis

1. Methodology

To assess whether there are bibliometric differences between TRA and R01 outputs, STPI examined a variety of publication and citation level metrics, detailed below. In addition, STPI used altmetric data to complement our traditional, citation-based analyses to provide a more comprehensive understanding and assessment of research influence.

All 57 TRA and 169 R01 awardees were included in the bibliometric analyses. Only publications attributed to the TRA and R01 awards of interest as designated in NIH's QVR system are included in the following analyses. In addition, publications were limited to those that were published 1 year after an award's project start date and within a year of an award's project end date to account for delays in publishing.

Similarly, NIH's iCite database only contains citation data for articles published between 1980 and 2019, and only has RCR data for articles published between 1980 and 2018. [40]

a. Publication Level Metrics

To assess whether publication output differed between TRA and R01 awards, STPI considered the following metrics:

- *total number of publications produced per award,*
- *average number of publications produced per award per year*, and
- *average total direct cost spent per publication per award.*

To determine whether the *total number of publications produced per award* differed between TRA and R01 awards, a GLM with a Poisson distribution was used with *total number of publications produced per award* as the dependent variable, and *group* (i.e., TRA or R01) as the explanatory variable. In addition, because there are other variables that could influence the total number of publications produced for a given award, STPI performed a multi-variable GLM regression analysis to take into consideration the effects of the following variables in addition to *group*:

- *award duration* (continuous variable, units in years),
- *total direct cost* (*per $100,000 spent*; continuous variable, units in dollars),
- *whether the award had multi-PIs* (binary categorical variable: 1 for multi-PI, and 0 for single PI awards), and

[40] iCite can be accessed at https://icite.od.nih.gov/.

- *area of science* (categorical variable consisting of biobehavioral research, biomedical research, therapy intervention, and tool development).

A type-II sum of squares analysis of deviance was performed to assess which of the explanatory variables in the GLM significantly affected the *total number of publications produced per award*. To determine which areas of science are significantly different from one another, pairwise comparisons among the different areas of science were performed using a post-hoc Tukey test with the *glht* function from the *multcomp* package in R (Hothorn et al. 2020). Awards with zero publications were included in both the single- and multi-variable regression analyses.

A non-parametric Kruskal-Wallis rank sum test was used to assess whether the average *number of publications produced per award per year* and the average *total direct cost spent per publication per award* differed by *group*. The number of publications produced per award per year was calculated by taking the total number of publications produced by an award divided by the award duration (in years). The total direct cost spent per publication per award was calculated by dividing the total direct cost for an award by the total number of publications produced by that award. Awards with zero publications were removed from both analyses.

Lastly, STPI also assessed whether *time to first publication* as well as *overall time to publication* (i.e., the rate at which articles were published) were significantly different by *group*. Publications that provided at least the month and year for the publication date were included in this analysis. *Time to first publication* was calculated as the number of days between the project start date and the publication date for the first article attributed to an award. Articles published after the first publication were removed from this analysis. *Overall time to publication* was calculated as the number of days between the project start date and the publication date for each article attributed to an award. For publications in which only the month and year were provided, the *day* unit was set to the first of the month. A Cox proportional hazard model was used to examine whether *group* was predictive of both *time to first publication* and *overall time to publication*. Awards with zero publications were removed from this analysis.

b. Citation Level Metrics

Differences between TRA and R01 awards were assessed for the following citation level metrics:

- *total number of citations received per publication,*
- *total number of citations received per award,*
- *average number of citations received per publication per year,*
- *average total direct cost spent per citation received per award,* and
- *average RCR*
- weighted RCR.

The total number of citations received per publication, average number of citations received per publication per year, and RCR are provided by iCite. An award level weighted RCR was calculated as the number of articles associated with the award multiplied by their average RCR. STPI calculated the average total direct cost spent per citation per award by dividing the total direct cost of an award by the total number of citations received across all publications that fell within the specified time frame described above for an award.

A GLM with a Poisson distribution was used to assess whether the *total number of citations received per publication* (dependent variable) differed by *group* (explanatory variable). Similarly, a GLM with a Poisson distribution was also used to assess whether the *total number of citations received per award* (dependent variable) differed by *group* (explanatory variable) at the grant level. STPI performed a multi-variable GLM regression analysis at the grant level to take into consideration the effects of the same variables listed in the publication analysis (*group, award duration, total direct cost (per $100,000 spent),*whether the award had *multi-PI,* and *area of science*). A separate multi-variable GLM regression analysis was performed at the publication level to account for *year of publication* (continuous variable) and *total number of authors* (continuous variable) as potential variables on top of *group* and whether the award had *multi-PI.* A type-II sum of squares analysis of deviance was performed for each GLM to assess which explanatory variable(s) is predictive of the total number of citations received. To determine which areas of science differed significantly from one another, pairwise comparisons among the different areas of science were performed using a post-hoc Tukey test with the *glht* function from the *multcomp* package in R (Hothorn et al. 2020).

A non-parametric Kruskal-Wallis rank sum test was used to assess if the *average number of citations received per publication per year, average total direct cost spent per citation received per award,* and average RCR differed by *group*.

STPI also assessed whether *time to first citation* differed significantly by *group*. *Time to first citation* (in days) was calculated by (1) an award's project start date and (2) by the publication date for the article that received the citation. For an award's project start date,

time to first citation was calculated as the number of days between the publication date of an award's first citation and an award's project start date. In other words, *time to first citation* was considered as the first citation of an award and not as the first citation of the first publication of that award. In most cases, the first publication of an award received the first forward citation but this was not true for all cases. For the article that received the first citation of an award, *time to first citation* was calculated as the number of days between an article's publication date and the publication date of an award's first citation.

Lastly, STPI considered the rate at which citations were accumulated (i.e., *time to citation*). *Time to citation* was calculated for each citation that an article received as the number of days between the publication date of the citation and (1) the award's project start date of the article that received the citation, and (2) the publication date of the article that received the citation.

All negative *time to first citation* and *time to citation* values (i.e., a study was cited before it was formally published) were removed from the analysis. For each case, a Cox proportional hazard model was used to assess whether the citation rate was significantly affected by *group*. Publications that provided at least the month and year for the publication date were included in this analysis.

Awards with zero publications were removed from all citation analyses. Publications with citation data listed as *NA* were removed on an individual analysis by analysis basis as some citation data would have a numeric value listed (e.g., total number of citations received) but others would not (e.g., RCR).

c. Altmetrics

Altmetrics provide complementary data to traditional, citation-based metrics, including citations on Wikipedia as well as public policy documents; discussions on research blogs; coverage on mainstream media; and mentions on social media such as Twitter and Facebook (Altmetric 2020a). Whereas traditional citations provide information on the research impact of an article on the academic community, altmetrics also consider how widely disseminated an article is beyond the publishing journal and immediate scientific community and how much attention an article receives from the public sphere (Altmetric 2020a). Similarly, because of the lag time between article submission and actual publication, it takes time for articles to accumulate citations and therefore, there is also an associated lag time in the ability to measure the immediate impact of an article using traditional, citation-based metrics. Altmetrics, by virtue of being sourced from the internet, allow for faster assessment of research impact.

There are limitations to the use of altmetric data. Altmetrics are a complement to, and not a replacement for, traditional, journal-based citations. It is important to consider altmetric data in context such as understanding where the underlying data come from (e.g.,

which sources are discussing the article of interest, what the sources are saying about the article of interest). Lastly, altmetric data are still relatively new and more research is needed to better understand the use and interpretation of altmetrics. To prevent individuals from artificially inflating the altmetric score for an article, companies that gather such data use algorithms to identify and correct for artificial inflation.

To gather altmetric data, STPI queried the Altmetric database using the rAltmetric package in R (Ram 2017). Articles were identified using their PubMedID (PMID). For the altmetric analysis, we focused on the Altmetric attention score as the response variable of interest. The Altmetric attention score is an automatically calculated, weighted count of all of the attention a research output has received and is based on three main factors: the volume of attention or mentions that a research output receives; the source that mentioned the research output; and how often authors of each mention talk about the scholarly articles. Each of these factors is weighted accordingly (Altmetric 2020b). For volume, a mention is only counted once from each person per source so that if the same person tweets about the same paper more than once, only the first mention will be counted towards the Altmetric score. Different sources contribute differently to the Altmetric score where reputable sources, such as a newspaper article, contributes more than a blog post, which contributes more than a tweet.[41] And lastly for authors, a mention about an article that is shared by a researcher to other researchers counts more than a journal account sharing the same article link automatically. More generally, the Altmetric attention score is a metric for the amount of online activity a research output receives and is not necessarily a metric for the quality of the research, or the researcher, as mentions may be both negative or positive. STPI used the Altmetric attention score as a metric for broad research impact and to identify articles that received exceptional online coverage that may be of interest to the NIH HRHR program.

A Kruskal-Wallis test was used to assess whether the Altmetric attention score differed significantly between TRA and R01 awards and a QQ plot was used to compare the distribution of Altmetric attention scores between the two groups. In addition, STPI tested whether the Altmetric attention scores were significantly different between TRA and R01 publications for each area of science using a Kruskal-Wallis test.

2. Results

As a reminder, 57 TRA awards and 169 R01 awards were used in this analysis unless indicated otherwise.

---

[41] The standard weightings for each mention type used by altmetric can be found here: https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-score-calculated-

a. Publications

Overall, 56 of the 57 TRA and 166 of 169 R01 awards reported producing at least one publication. In total, the 57 TRA and 169 R01 awards produced 1,255 and 3,123 publications, respectively. After limiting publications to those published after a year of an award's project start date to within a year of an award's project end date, 55 of the 57 TRA and 161 of 169 R01 awards reported producing at least one publication that fell within the specified time frame, resulting in a total of 1,007 TRA and 2,270 R01 publications that were included in STPI's analyses (Table 11).

Table 11. Summary Statistics of Bibliometric Data by *Group*

Group	Total number of awards	Number (percent) of awards with at least one publication	Total number of publications ǂ	Mean (± SE) number of publications/award *
TRA	57	55 (96.5%)	1,007	17.7 (± 1.93)
R01	169	161 (95.3%)	2,270	13.4 (± 0.94)

Source*:* publications were downloaded from QVR (March 2020) and citations to those publications were downloaded from iCite (May 2020)

Note*:* ǂ Only publications that had publication dates after a year of an award's project start date and within a year of an award's project end date are included in STPI's analyses.

Note*:* * Includes awards with 0 publications.

Overall, TRA awards (including those with zero publications), on average (± SE), produced a significantly higher number of publications (17.7 ± 1.9) compared to R01 (13.4 ± 0.9) awards ($\chi^2_1 = 50.46, p < 0.001$; Table 12). Significant difference was still detected, between TRA and R01 awards after STPI normalized the number of publications produced by award duration. The average (± SE) number of publications produced per award per year was 3.26 (± 0.32) for TRA awards and 2.49 (± 0.17) for R01 awards ($\chi^2_1 = 6.40, p = 0.01$). There was no significant difference in the average (± SE) total direct cost spent to produce a publication between TRA ($279,363 ± $58,919) and R01 ($471,255 ± $67,072) awards ($\chi^2_1 = 3.25, p = 0.07$).

Table 12. Mean (± SE) Values on Bibliometric Publication Metrics by *Group*

Metric	TRA	R01	χ^2_1	P-value
Number of publications produced per award	17.7 ± 1.9	13.4 ± 0.9	50.46	< 0.001 ***
Number of publications produced per award per year	3.26 ± 0.32	2.49 ± 0.17	6.40	0.01 *
Total direct cost spent per publication	$279,363 ± $58,919	$471,255 ± $67,072	3.25	0.07

Source: publication data were downloaded from QVR (March 2020)

* Significant at $p < 0.05$

** Significant at $p < 0.01$

*** Significant at $p < 0.001$

1) Multi-variable GLM Analysis of Publications at the Award Level

Results from the multi-variable GLM regression analysis and type-II sum of squares analysis of deviance showed that when all other variableswere held constant, the total number of publications produced differed significantly between TRA and R01 awards ($p < 0.001$; Table 13). The number of publications produced was also significantly impacted by *award duration*, *total direct cost (per $100,000 spent)*, whether an award had *multi-PIs*, and the research *area of science* ($p < 0.001$ for each).

Table 13. Results from the Type-II Sum of Squares Analysis of Deviance for Publications at the Award Level

Variable of interest	Chi-squared test statistic (df)	P-value
Group	41.9 (1)	< 0.001 ***
Award duration	70.1 (1)	< 0.001 ***
Total direct cost (per $100,000 spent)	54.3 (1)	< 0.001 ***
Multi-PI award	60.2 (1)	< 0.001 ***
Area of science	39.4 (3)	< 0.001 ***

* Significant at $p < 0.05$

** Significant at $p < 0.01$

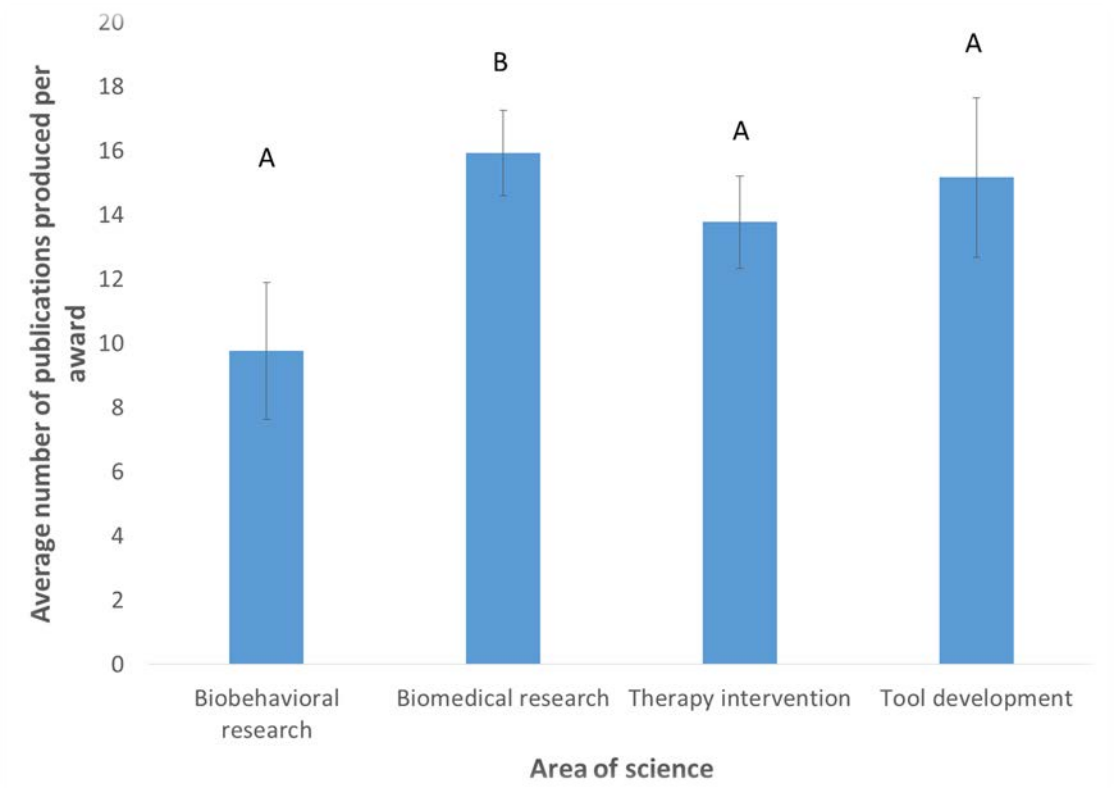
*** Significant at $p < 0.001$

Specifically, when all other factors are held constant, the expected number of publications produced by TRA awards is 30% more relative to R01 awards; for every year increase in *award duration*, the expected number of publications produced per award increased by 25%; the expected number of publications produced by awards with multi-PIs increased by 32% compared to awards with single PIs; and the expected number of publications produced increased by 0.6% for every $100,000 spent in total direct cost (Table 14). In the GLM analysis, behavioral research was arbitrarily set as the baseline for area of science against which all other areas of science were compared. To assess how each area of science compares to one another, a post-hoc Tukey test was performed.

Table 14. GLM Results for Publications at the Award Level

Variable of interest	Coefficient estimate	Expected increase or decrease in number of publications produced relative to baseline
Group		
R01 (baseline)	NA	NA
TRA	0.26	30%
Award duration	0.22	25%
Total direct cost (per $100,000 spent)	0.006	0.6%
Multi/Single Award PI		
Single PI (baseline)	NA	NA
Multi-PI	0.28	32%
Area of science		
Behavioral research (baseline)	NA	NA
Biomedical research	0.34	40%
Therapy intervention	0.18	19%
Tool development	0.15	16%

Results from the post-hoc Tukey test on *area of science* showed that average (± SE) number of publications produced from awards focused on biomedical research (15.9 ± 1.33) are significantly higher than publications than awards focused on any other area of science ($p < 0.001$ for each pair-wise comparison) (Figure 31). However, the average number of publications produced did not differ significantly between awards focused on therapy intervention (13.78 ± 1.43) and those focused on tool development (15.2 ± 2.49; $p = 0.97$). No significant difference was found between the number of publications from awards focused on biobehavioral research (9.8 ± 2.14) and tool development($p = 0.27$), nor between biobehavioral research and therapy intervention ($p = 0.05$).



Different letters denote significant differences in the number of publications produced between areas of science.

Figure 31. Average (± 1 SE) Number of Publications Produced per Award by Area of Science

2) Time to Publication Analysis

For *time to first publication*, the Cox regression model indicated that there was no significant difference in the rate at which the first article of an award was published between TRA and R01 awards ($p = 0.13$). The hazard ratio (95% confidence interval) for TRA awards compared to R01 awards was 1.27 (0.93 to 1.73) (Figure 32).

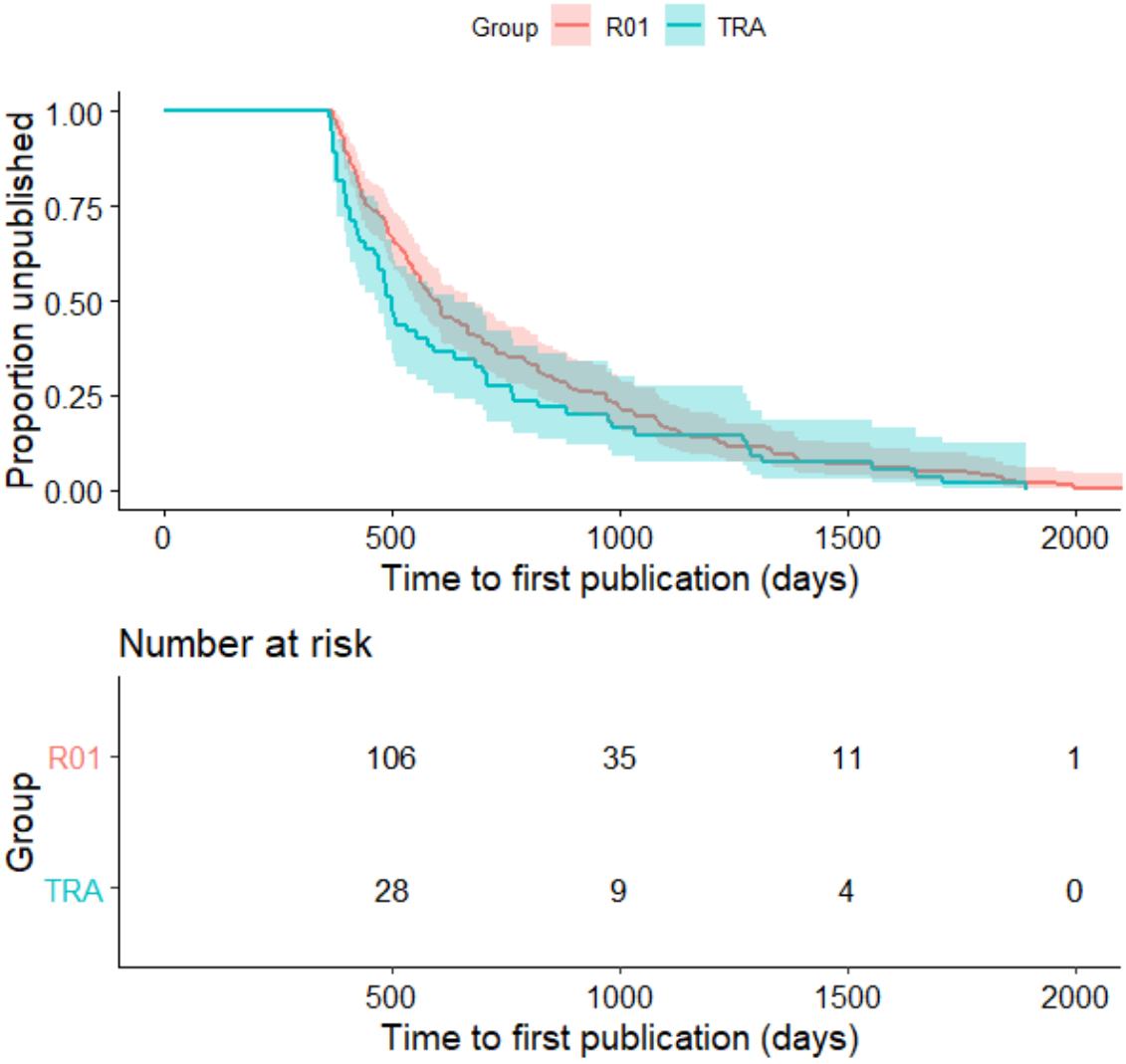


Figure 32. Proportion of Awards That Have Not Published at Least One Article over Time by *Group*

Regarding the publication rate, the Cox regression model for *overall time to publication* indicated that there was a significant difference in the rate at which articles were published between TRA and R01 awards ($p < 0.001$). The hazard ratio (95% confidence interval) for TRA awards compared to R01 awards was 1.28 (1.18 to 1.39) (Figure 33). TRA awards have a 28% faster overall time to publication compared to R01 awards.

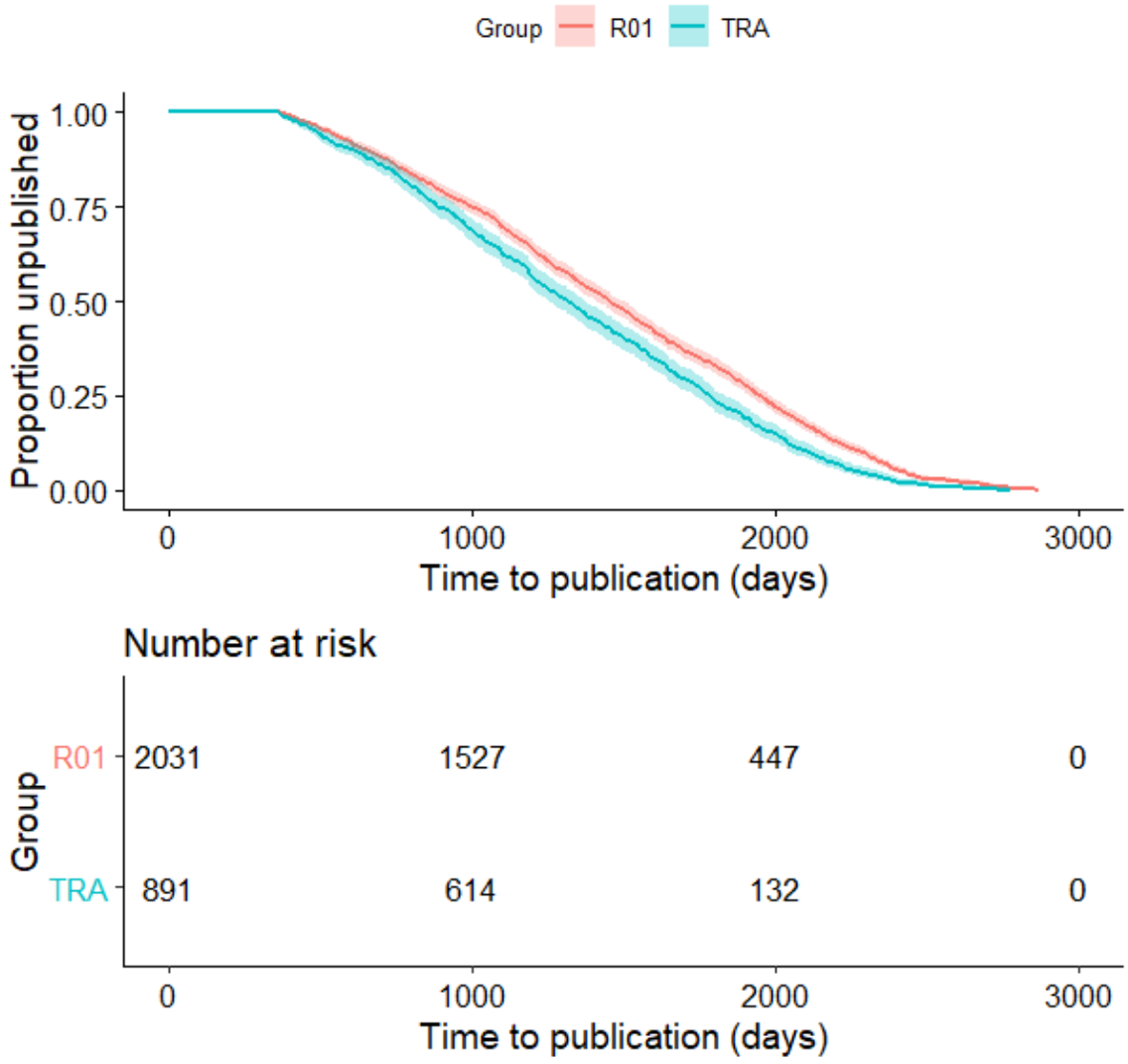


Figure 33. Proportion of Articles Not Published over Time by *Group*

b. Citations

At the time of STPI's analyses, the iCite database only contained citation data for articles published between 1980 and 2019. As a result, iCite was unable to provide citation data for six R01 publications, which were removed from all citation analyses. Overall, 986 of the 1,007 TRA publications and 2,197 of the 2,270 R01 publications have received at least one citation (Table 15). In total, TRA publications have accumulated a total of 77,167 citations as of April 2020, and R01 publications have accumulated a total of 65,715 citations during the same time. On average (± SE), TRA award publications received significantly more citations (76.6 ± 8.01) than R01 award publications (29.0 ± 2.02; $\chi_1^2 = 33{,}028,\ p < 0.001$; Table 15). This was true even after the data were normalized to number of citations received per publication per year. TRA award publications, on average (± SE), received 14.24 (± 1.23) citations per year, which was significantly higher than R01 award publications (5.54 ± 0.38; $\chi_1^2 = 185.73,\ p < 0.001$). There was also a significant difference between TRA and R01 awards for total direct cost spent for each citation received ($\chi_1^2 = 18.92, p < 0.001$). Each citation, on average (± SE), costed $9,304 (± $1,830) for TRA awards and $51,606 (± $11,118) for R01 awards.

TRA award publications had significantly higher RCRs (4.48 ± 0.37) than R01 award publications (2.26 ± 0.13; χ_1^2=87.80,p<0.001). TRA awards also had significantly higher weighted RCRs than R01 awards (82.1 ± 20.5 and 31.1 ± 3.34 for TRA and R01 awards respectively; χ_1^2=10.1,p<0.01).

Table 15. Mean (± SE) Values on Bibliometric Citation Metrics by *Group*

Metric	TRA	R01	χ_1^2	P-value
Number of citations received per publication	76.6 ± 8.01	29.0 ± 2.02	33,028	< 0.001 ***
Number of citations received per publication per year	14.24 ± 1.23	5.54 ± 0.38	185.73	< 0.001 ***
Total direct cost spent per citation received	$9,304 ± $1,830	$51,606 ± $11,118	18.92	< 0.001 ***
Relative Citation Ratio	4.48 ± 0.37	2.26 ± 0.13	87.80	< 0.001 ***
Weighted Relative Citation Ratio	82.1 ± 20.5	31.1 ± 3.34	10.1	< 0.01 **

Source*:* publication data were downloaded from iCite (May 2020)

* Significant at $p < 0.05$

** Significant at $p < 0.01$

*** Significant at $p < 0.001$

1) Multi-variable GLM Analysis of Citations at the Award Level

At the award level, results from the multi-variable GLM regression analysis and type-II sum of squares analysis of deviance showed that *group*, *award duration*, *total direct cost (per $100,000 spent)*, whether an award had *multi-PIs*, and *area of science* all significantly affected the total number of citations received ($p < 0.001$ for each; Table 16).

Table 16. Results from the Type-II Sum of Squares Analysis of Deviance for Citations at the Award Level

Variable of interest	Chi-squared test statistic (df)	P-value
Group	44,991 (1)	< 0.001 ***
Award duration	392 (1)	< 0.001 ***
Total direct cost (per $100,000 spent)	22,889 (1)	< 0.001 ***
Multi-PI award	10,423 (1)	< 0.001 ***
Area of science	19,873 (3)	< 0.001 ***

* Significant at $p < 0.05$

** Significant at $p < 0.01$

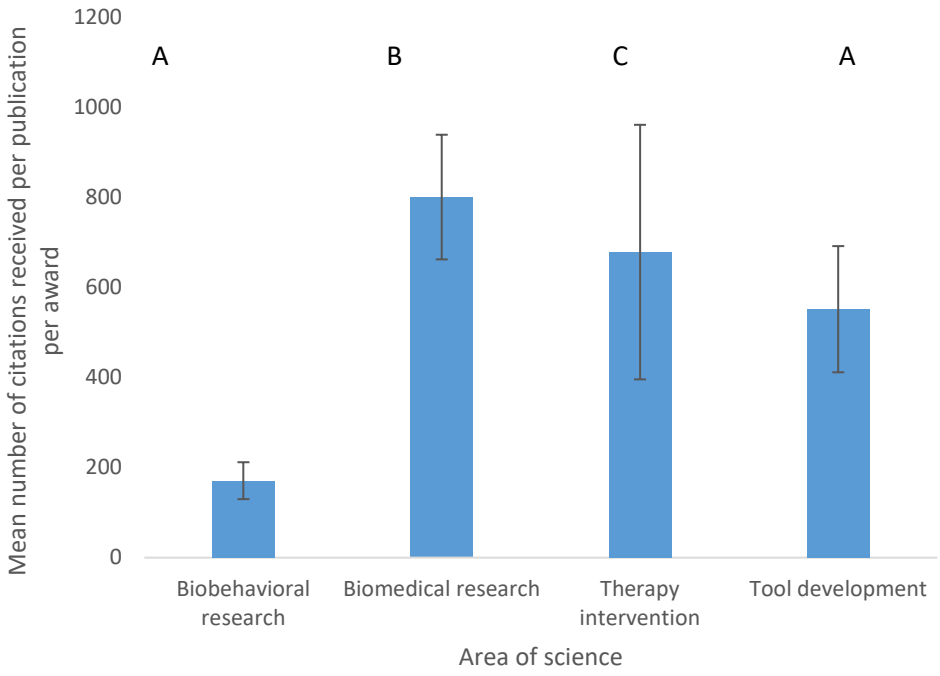
*** Significant at $p < 0.001$

Specifically, when all other variables are held constant, the expected number of citations increased by 231% for TRA awards relative to R01 awards; the expected number of citations decreased by 8% for every year increase in award duration; the expected number of citations increased by 74% for awards with multi-PIs relative to those with single PIs; and the expected number of citations increased by 1.5% for every $100,000 spent in total direct cost (Table 17). In the GLM analysis, behavioral research was arbitrarily set as the baseline for area of science against which all other areas of science were compared. To assess how each area of science compares to one another, a post-hoc Tukey test was performed.

Table 17. GLM Results for Citations at the Award Level

Variable of interest	Coefficient estimate	Expected increase or decrease in number of citations received relative to baseline
Group		
R01 (baseline)	NA	NA
TRA	1.20	231%
Award duration	-0.08	-8%
Total direct cost (per $100,000 spent)	0.01	1.5%
Multi/Single Award PI		
Single PI (baseline)	NA	NA
Multi-PI	0.55	74%
Area of science		
Behavioral research (baseline)	NA	NA
Biomedical research	1.11	205%
Therapy intervention	1.03	181%
Tool development	0.001	1%

Results from the post-hoc Tukey test on *area of science* showed that the average (± SE) number of citations received differed significantly across all areas of science ($p < 0.001$ for each pair-wise comparison; Figure 34) except for the pair-wise comparison between awards focused on tool development and awards focused on biobehavioral research ($p = 0.98$). Specifically, awards focused on biomedical research received, on average (± SE), the highest number of citations per award (801 ± 138), followed by awards focused on therapy intervention (679 ± 140), tool development (552 ± 140), and then biobehavioral research (172 ± 41).



Different letters denote significant differences in the number of citations received among different areas of science.

Figure 34. Mean Number of Citations Received per Award by Area of Science

2) Multi-variable GLM Analysis of Citations at the Publication Level

As a reminder, a multi-variable GLM regression analysis was also conducted to assess which factors influenced the number of citations received at the publication level (i.e., total number of citations received per publication). Results from this multi-variable GLM regression analysis and type-II sum of squares analysis of deviance showed that *group*, *year of publication*, whether an award had *multi-PIs*, and the total *number of authors* on a publication all significantly affected the number of citations received per publication ($p < 0.001$ for each; Table 18).

Table 18. Results from the Type-II Sum of Squares Analysis of Deviance for Citations at the Publication Level

Variable of interest	Chi-squared test statistic (df)	P-value
Group	34,575 (1)	< 0.001 ***
Year of publication	35,601 (1)	< 0.001 ***
Multi/Single PI award	1,604 (1)	< 0.001 ***
Number of authors	10,413 (1)	< 0.001 ***

* Significant at $p < 0.05$

** Significant at $p < 0.01$

s*** Significant at $p < 0.001$

Specifically, the expected number of citations received per TRA award publication was 172% higher than those received per R01 award publication; the expected number of citations received per publication was 24% higher for awards with multi-PIs compared to those with single PIs; for every year increase in year of publication, the expected number of citations received per publication decreased by 26%; and for every additional author listed on a publication, the expected number of citations received per publication increased by 1.9% (Table 19).

Table 19. GLM Results for Citations at the Publication Level

Variable of interest	Coefficient estimate	Expected increase or decrease in number of citations received relative to baseline
Group		
R01 (baseline)	NA	NA
TRA	1.0	172%
Year of publication	-0.30	-26%
Multi/Single PI Award		
Single PI (baseline)	NA	NA
Multi-PI	0.21	24%
Number of authors	0.02	1.9%

3) Time to Citation Analysis

In total, TRA publications were cited by 71,872 unique publications and R01 publications were cited by 62,798 unique publications. The Cox regression model for *time to first citation* by an award's project start date showed that there was significant difference in rate at which first citations were received between TRA and R01 awards ($p = 0.0135$). The hazard ratio (95% confidence interval) for TRA awards compared to R01 awards was 1.48 (1.08 to 2.01) (Figure 35). TRA awards receive their first citation 48% faster compared to R01 awards.

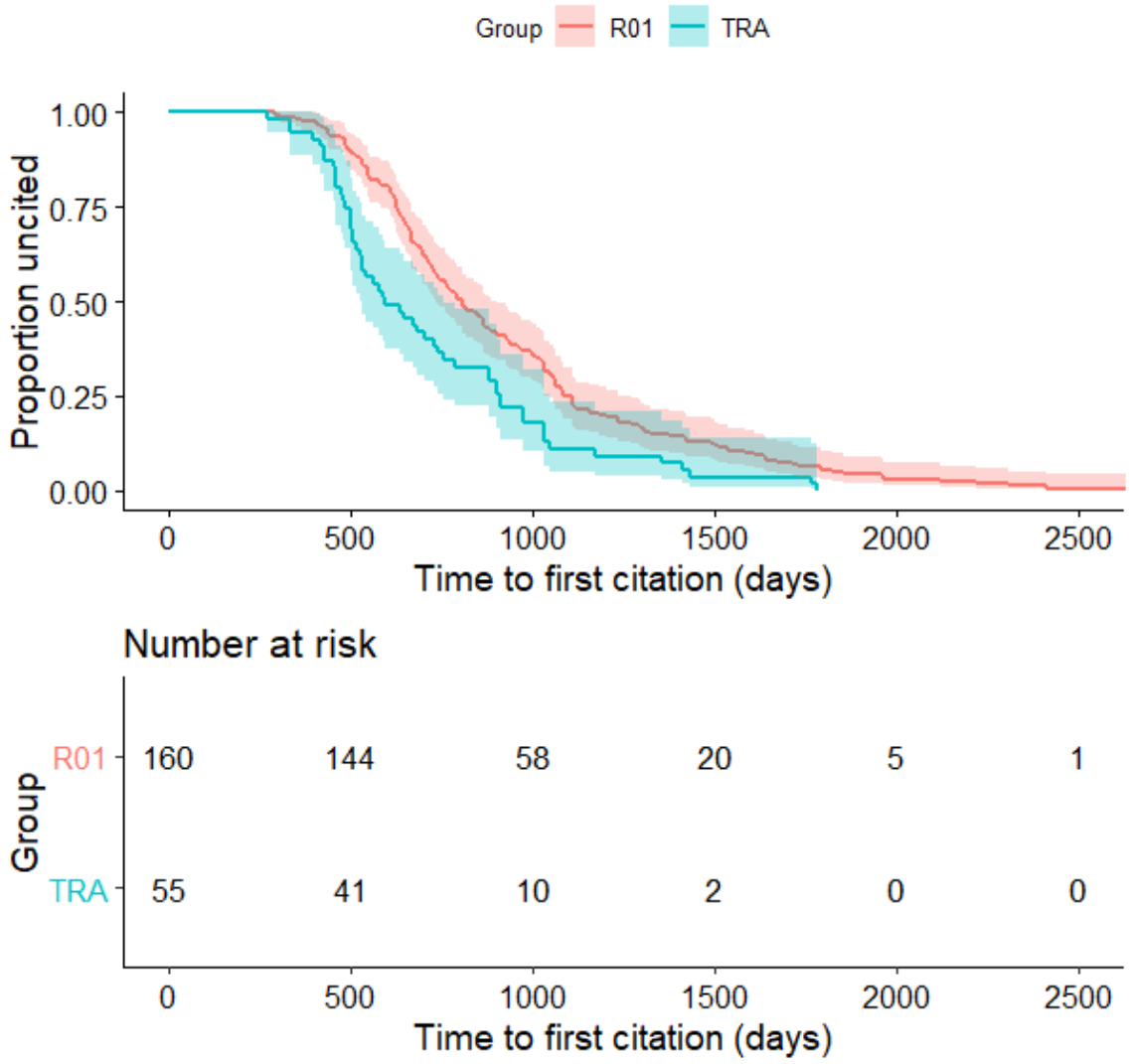


Figure 35. Proportion of Awards with Publications That Are Uncited over Time by *Group* Where Time to First Citation Is Calculated from an Award's Project Start Date

Similarly, the Cox regression model for *time to first citation* by the publication date of the article that received the first citation of an award showed that there was significant difference in rate at which first citations were received between TRA and R01 awards ($p < 0.001$) (Figure 36). The hazard ratio (95% confidence interval) for TRA awards compared to R01 awards was 1.70 (1.24 to 2.33) meaning that TRA awards, on average, received their first citation 70% faster than R01 awards.

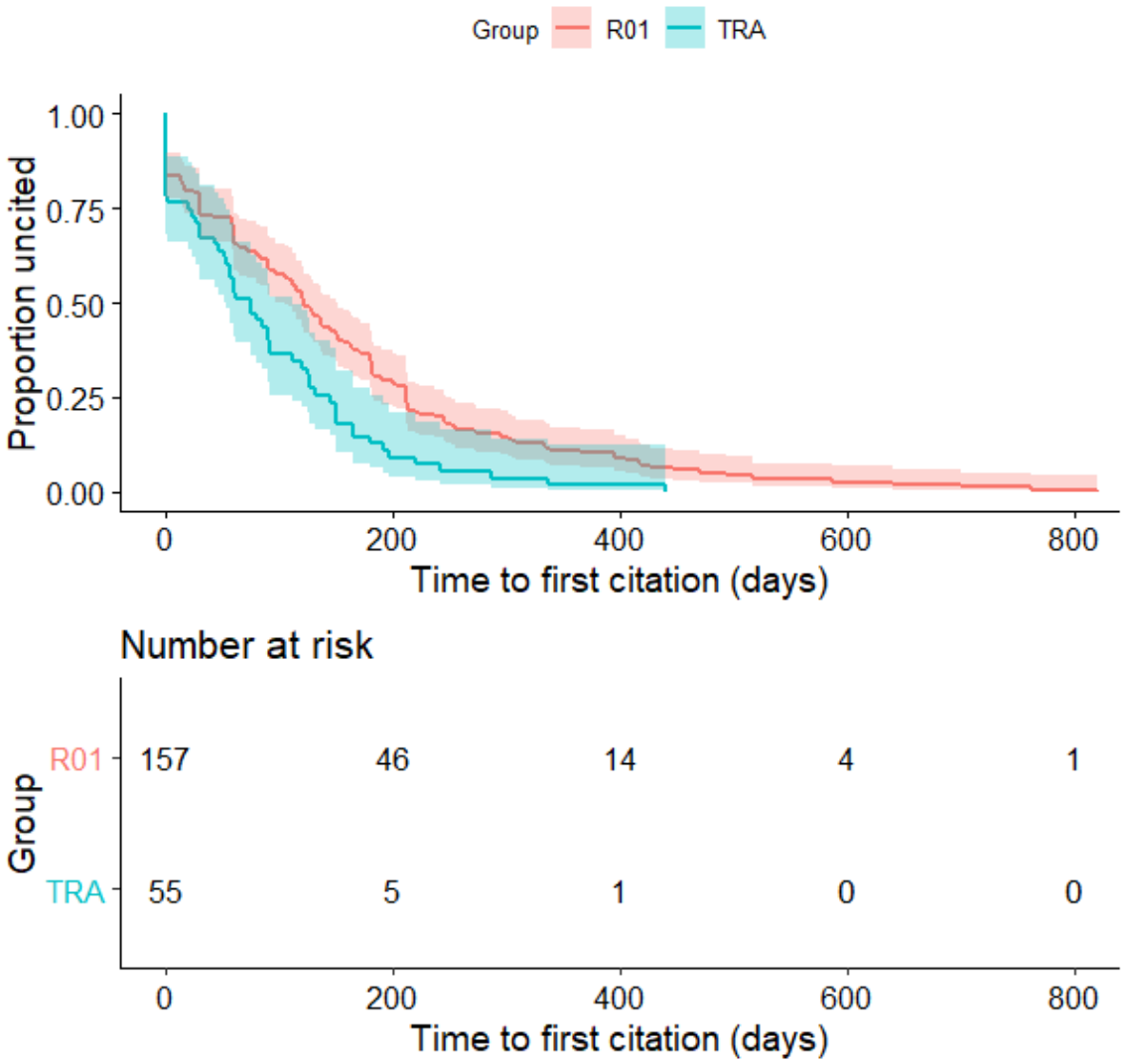


Figure 36. Proportion of Awards with Publications That Are Uncited over Time by *Group* Where Time to First Citation Is Calculated from the Publication Date of the Article That Received the Citation

For *overall time to citation* by project start date, the Cox regression model showed that the rate at which TRA publications accumulated citations was significantly faster than that of R01 publications ($p < 0.001$) (Figure 37). The hazard ratio (95% confidence interval) for TRA publications compared to R01 publications was 1.16 (1.14 to 1.17) meaning that TRA publications, on average, receive citations approximately 16% faster than R01 publications.

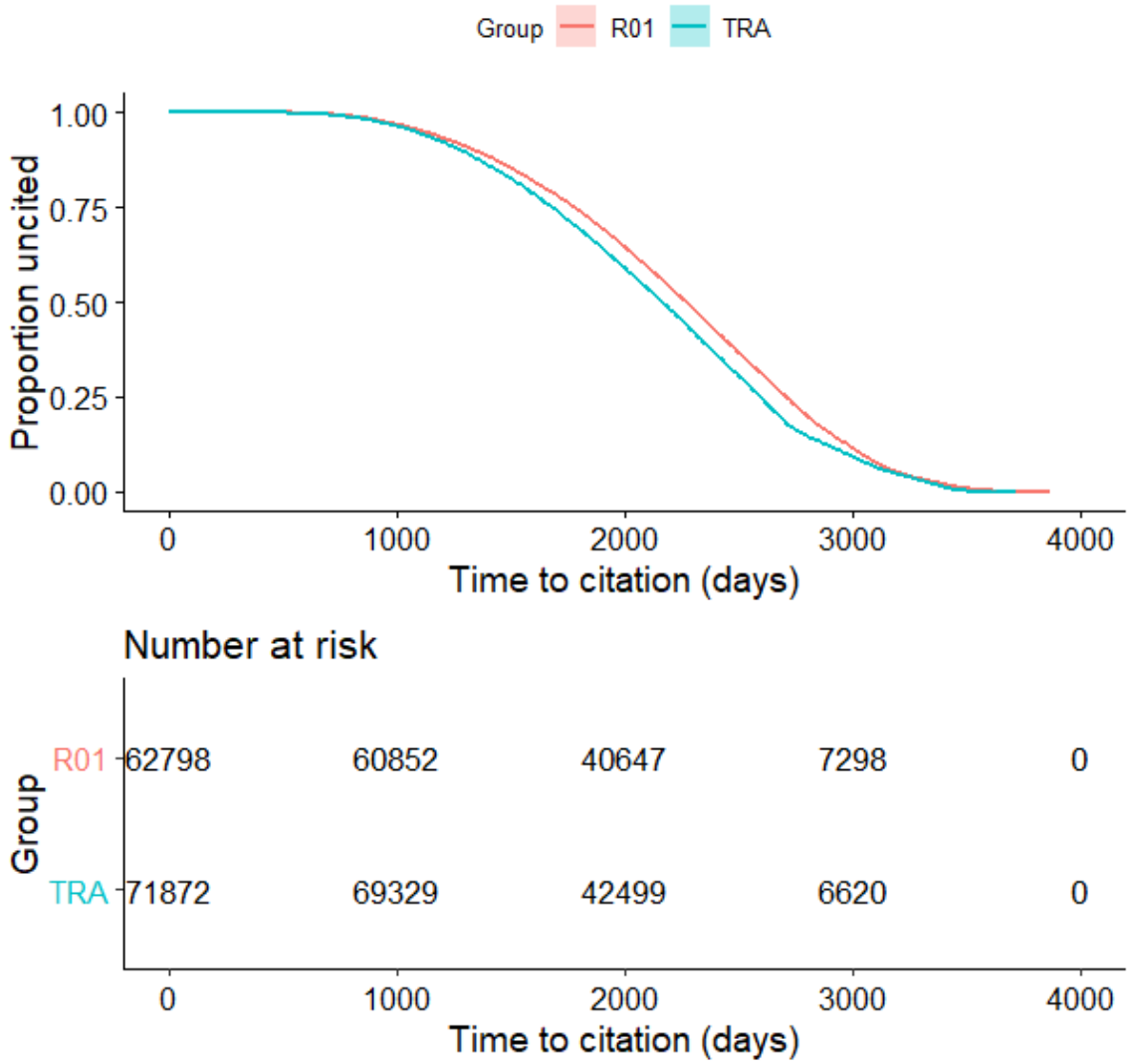


Figure 37. Proportion of Publications Uncited over Time by *Group* Where Time to Citation Is Calculated from an Award's Project Start Date

For overall time to citation by the publication date of the article that received the citation, the Cox regression model showed that the rate at which TRA publications accumulated citations was significantly slower than that of R01 publications ($p < 0.001$) (Figure 38). The hazard ratio (95% confidence interval) for TRA publications compared to R01 publications was 0.91 (0.90 to 0.92) meaning that TRA publications, on average, receive citations approximately 9% slower than R01 publications.

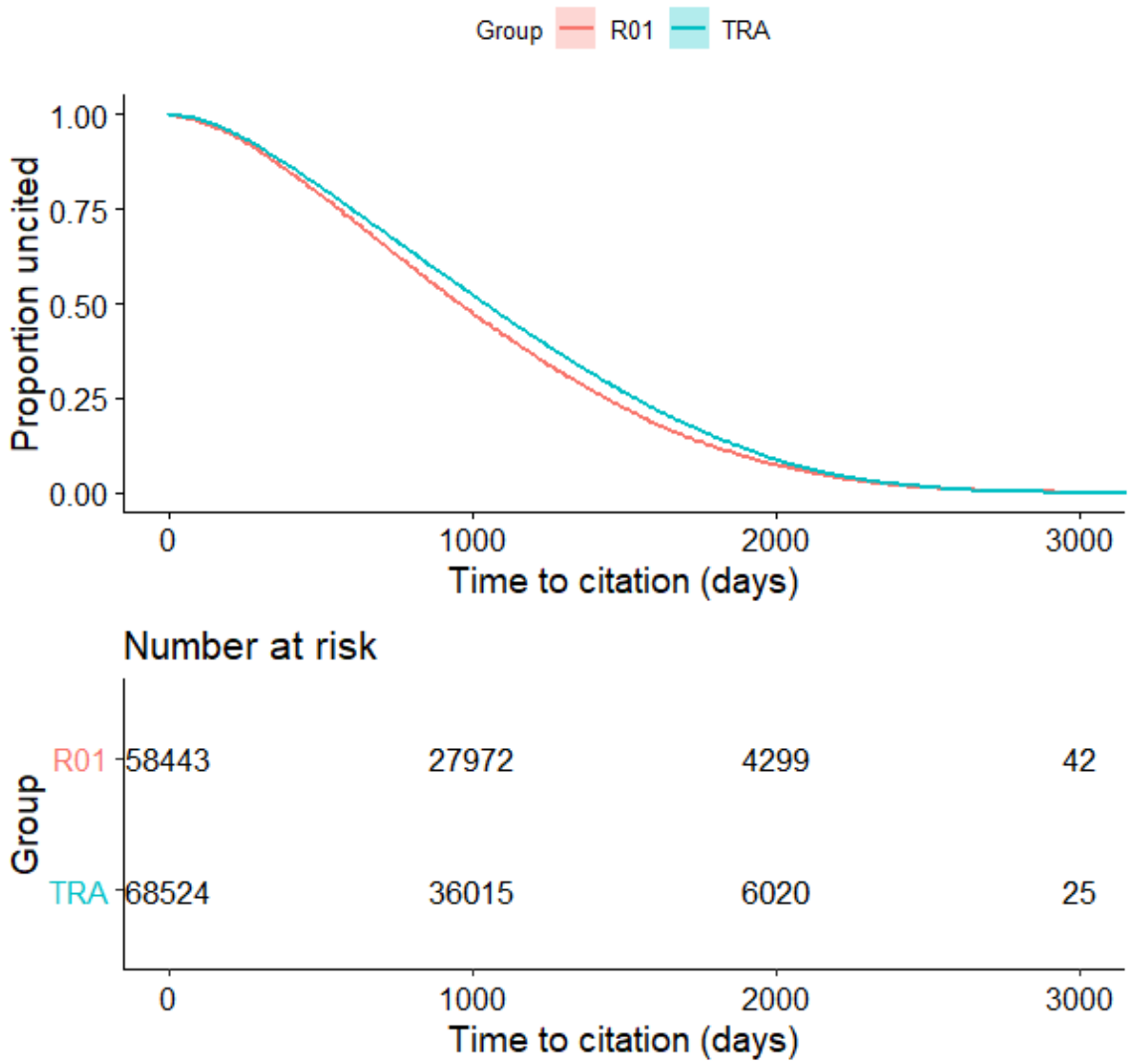


Figure 38. Proportion of publications Uncited over Time by *Group* Where Time to Citation Is Calculated from the Publication Date of the Article That Received the Citation

c. Altmetrics

The mean (± SE) Altmetric attention score was 52.1 (± 4.28) among 870 TRA publications and 32.3 (± 8.01) among 1171 R01 publications. Results from the Kruskal-Wallis test indicate that TRA publications, on average, have significantly higher Altmetric attention scores than R01 publications ($\chi_1^2 = 130.7\,, p < 0.001$). This is corroborated by the QQ plot of Altmetric attention scores for TRA and R01 publications, which shows that while there are some R01 publications that received higher than expected Altmetric attention scores, TRA publications, overall, tended to have higher scores than R01 publications (Figure 39).

A QQ plot is a visual tool to help assess whether two different data sets have the same distribution by plotting the quantiles (i.e., the fraction, or percent, of points below a given value) of each group against one another. In Figure 39, the estimated quantiles from the TRA and R01 Altmetric attention scores are represented by the x- and y-axes, respectively.[42] The 45-degree diagonal line represents the theoretical quantiles—if both the TRA and R01 Altmetric attention scores came from a population with the same distribution, the points should fall approximately along the diagonal line. The greater the departure from this diagonal line, the greater the evidence that the TRA and R01 Altmetric attention scores come from populations with different distributions.

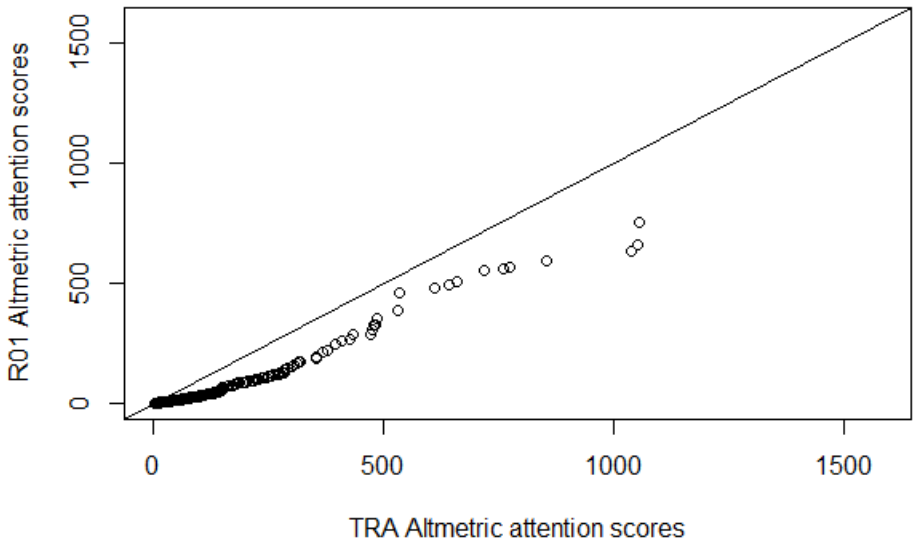


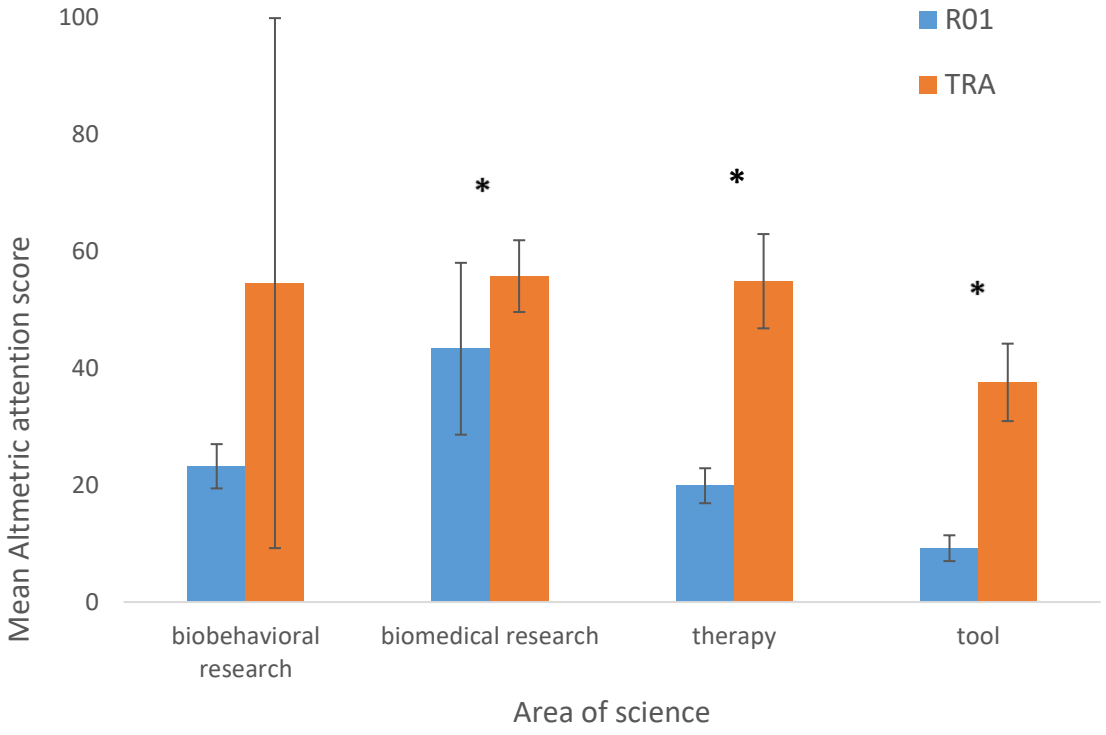
Figure 39. QQ Plot of Altmetric Attention Scores for TRA and R01 Publications

[42] For more information about QQ plots, please visit the Engineering Statistics Handbook at https://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.htm#:~:text=A%20q%2Dq%20plot%20is%20a,70%25%20fall%20above%20that%20value.

The QQ plot shows one TRA publication that received much higher online and media attention than all other TRA and R01 publications. Specifically, the publication of interest (PMID 26828196) is from grant number DK090989 (Disappearing Gastrointestinal Microbiota in Epidemic Obesity) and titled "Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer." This article had an Altmetric attention score of 1,584 and has received coverage by 142 news outlets as of May 2020—including Newsweek, CNN, PBS, the Washington Post, Wired, the Los Angeles Times, and CBS News. Comparatively, it has received 263 total citations since the article was published in 2016 and has a RCR of 25.7. Both the total number of citations and the RCR value indicate that the article made substantial impact on the scientific community. However, when compared against all publications, there are 65 other publications that had a higher total number of citations, and 22 publications that had higher RCRs. In fact, one TRA publication (PMID 24157548: Genome engineering using the CRISPR-Cas9 system) had the highest total number of citations (3,262) as well as the highest RCR (129) among all TRA and R01 publications. Conversely, this article has an Altmetric attention score of 180 and has, thus far, received coverage from 11 news outlets since its publication in 2013.

This finding highlights the importance of considering altmetric data when assessing research impact. From a traditional, citation-based viewpoint, the publication on CRISPR represents the cream of the crop of scientific findings and research impact. The publication on vaginal microbial transfer, while also notable, would have paled in comparison when considering traditional metrics such as total citations received and RCRs. However, its broader impact on the general public could have only been brought to the forefront using altmetric data.

TRA publications received significantly higher Altmetric attention scores (55.72 ± 6.13) compared to R01 publications for biomedical research (43.30 ± 14.69; $p < 0.001$; Figure 40). For therapy intervention research, TRA publications received significantly higher Altmetric attention scores (54.84 ± 8.06) compared to R01 publications (19.88 ± 2.99; $p < 0.001$). The same is true for tool development research, TRA publications received significantly higher altmetric attention scores (37.55 ± 6.62) compared to R01 publications (9.20 ± 2.21; $p < 0.001$). The only exception is in biobehavioral research, where there was no significant difference between TRA publications (54.52 ± 45.30) and R01 publications (23.21 ± 3.79; $p = 0.33$).



* Denotes significant difference (p < 0.05) in Altmetric attention score between TRA and R01 publications

Figure 40. Mean (± SE) Altmetric Attention Score by Group and Area of Science

4. Senior Scientist Review

To obtain an independent scientific assessment of TRA awardee survey results, STPI canvassed senior scientists about the transformative potential of the three most cited research publications for each TRA, NDPA, and R01survey respondent. Senior scientists were defined as mid-career scientists who had received an R01 within the last 10 years and performed at least 6 months of service to NIH, as reflected in their QVR record. The senior scientist survey was designed to mirror awardee survey content so that the senior scientist reviewers' responses could be aligned with awardee survey respondent responses.

A. Survey Development

STPI designed the senior scientist review structure so that each set of awardee research results was reviewed by three reviewers. A set of research results was defined as the three most cited publications attributed to the award. Each reviewer assessed three randomly assigned awardees that were some combination of TRA, NDPA, and R01 survey respondents.

1. Survey Respondents

As a reminder, a total of 91 individuals responded to the awardee survey (24 TRA awardees, 32 NDPA awardees, and 35 R01 awardees). Eleven of the 91 awards associated with the survey respondents did not have three publications attributed to their respective awards and were removed from the analysis (two TRA awards, three NDPA, and six R01 awards). As a result, 22 TRA, 29 NDPA, and 29 R01 awards were included in the analysis.

2. Senior Scientist Reviewers

STPI identified senior scientists using a two-pronged approach. First, all awardees were asked in the survey to identify three reviewers who they thought could contribute a knowledgeable and complete review of their respective research. This approach produced 237 potential reviewers. Second, STPI contacted the most recent program officer (PO) associated with each survey respondent's award and asked them to provide up to five reviewers who could serve as a subject matter expert for the award of interest. POs provided 179 potential reviewer names. STPI combined the two pools and, from the 416 recommendations, filtered out email addresses that were duplicative, missing, or did not have a proper format (e.g., @domain.edu). Fifty-seven names were removed from the list, and the final pool of potential reviewers contained 359 individuals.

A $500 honorarium was offered to reviewers who successfully completed their assignment. Receipt of the honorarium required review of research and completion of a survey for each of the three awardees and submission of all federally required paperwork.

3. Survey Content

The survey was composed of three question blocks that mirrored the grantee survey questions on paradigm shift and impact. Additional survey questions assessed the reviewer's confidence in their ability to assess the awardee's subject area and how the reviewer would describe the term "transformative." A full list of survey questions can be found in Appendix E.

B. Survey Administration

Three hundred and thirty-nine senior scientists received email invitations to participate in the review and instructions on the required paperwork on September 17, 2020. Eighty senior scientists agreed to participate by the deadline of October 7, 2020. Review materials were sent out, and those who had not completed their assignment by October 16, 2020 received a reminder email about the October 26, 2020 deadline. The survey was officially closed on October 29, 2020.

C. Survey Analysis

1. Statistical Methods

Descriptive statistics were used to compare TRA and NDPA award reviews, and TRA and R01 award reviews. Two-way chi-squared goodness of fit proportion tests were used to determine whether the percentage of reviewers who selected *agree* or *likely agree* for a survey question for TRA awards was significantly different from responses selected for NDPA or R01 awards.

2. Statistical Limitations

There are limitations to the use of Chi-squared analysis as used in this analysis, mainly the assumption of independent responses. Since each reviewer was randomly assigned three awardee packets, it is possible that the order in which the packets were reviewed influenced whether reviewers perceived the research output from an award as transformative. As a result, the assumption of independence across awardee packets may not hold.

D. Results

Seventy-three of the 80 reviewers completed their reviews (91% response rate). Because each reviewer was assigned a random R01, TRA, or NDPA awardee packet and each awardee packet was reviewed by multiple reviewers; 60 TRA, 79 NDPA, 83 R01 reviews were completed, and the number of reviews for each award varies. The results are presented here by survey question, and complete data tables are provided in Appendix F.

1. Descriptive Statistics for TRA and NDPA Awards

a. Paradigm Shift

The first set of questions addressed the paradigm-shifting nature of the awardees' research. These questions included novel inventions, novel combinations of ideas and approaches, or use of technology that significantly changed research practice or thinking in their field (Figure 41).

Reviewers agreed with all paradigm-shifting statements at similar levels for both TRA and NDPA awardees:[43] *One or more of the awardees research ideas challenged existing science technology paradigms* (61.0% and 64.5%; $\chi_1^2 = 0.05$, p = 0.82); *The awardees research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice* (69.5% and 63.6%; $\chi_1^2 = 0.28$, p = 0.59); *The awardees research furthered existing practices or thinking in their field* (93.3% and 93.6%; $\chi_1^2 < 0.001$, p = 1); *The awardees research led to a novel invention or a new technology which significantly improved current practices* (57.6% and 48.7%; χ_1^2 = 0.74, p = 0.39); *Led to the development of a new methodology that significantly changed research in their field* (57.6% and 49.3%; χ_1^2 = 0.61, p = 0.44); *Required the use of equipment technique or model that was novel and significantly changed research in their field* (64% and 54%; χ_1^2 = 1.1, p = 0.29); and *Aided in the development of new therapies clinical tools or strategies that significantly changed research or clinical practice* (35.6% and 33.8%; χ_1^2 < 0.001, p = 0.97); *Took the next steps in their established area of investigation* (90.0% and 88.5%; χ_1^2 < 0.001, p = 0.99).

[43] Data for descriptive statistics are presented as percentage of reviewers who agrees with the statement for TRA research outputs and percentage of reviewers who agrees with the statement for NDPA research outputs; χ_1^2, p value.

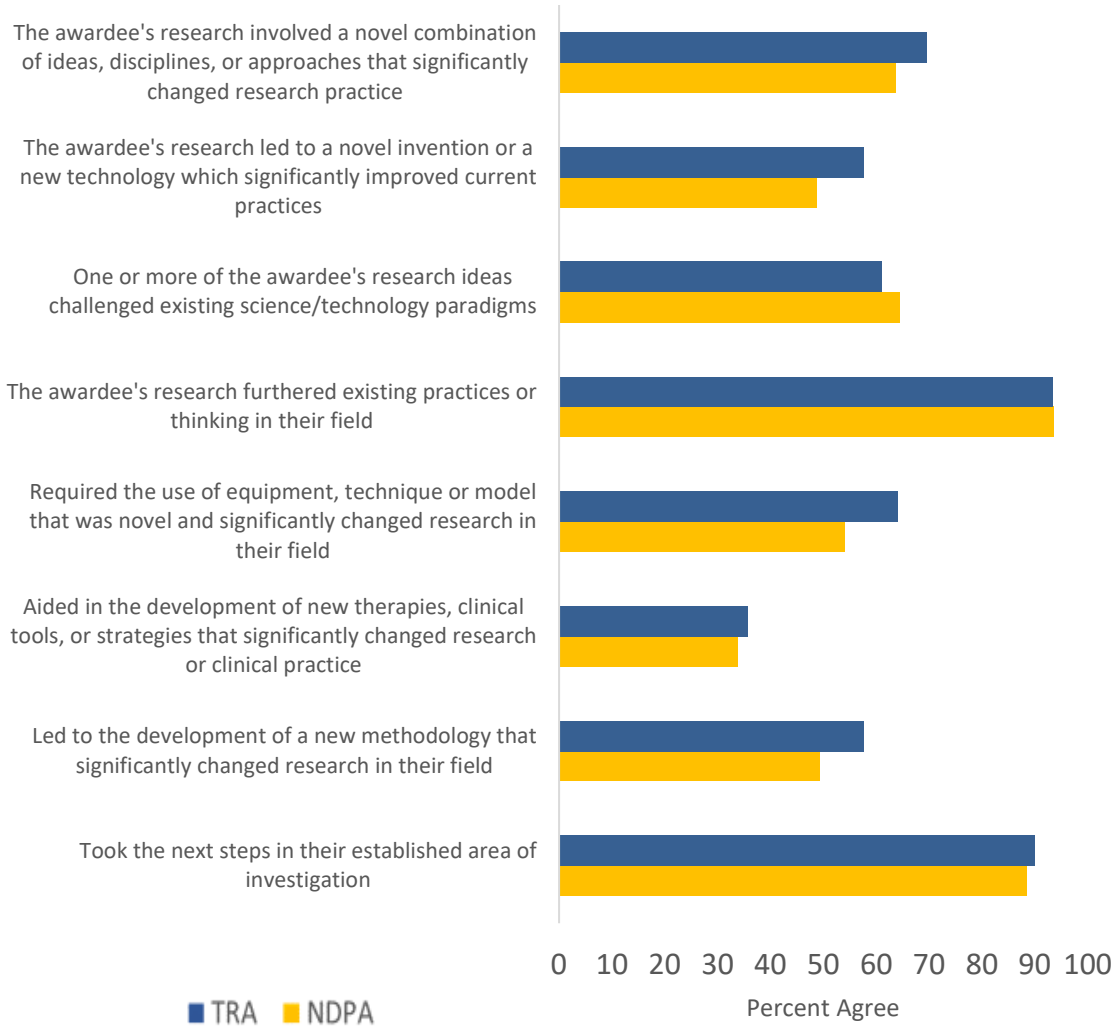


Figure 41. Senior Scientist Review: Paradigm-Shifting Research Outputs by Group

b. Impact

This question block explored the timing of impact (Figure 42). Reviewers agreed with all statements on impact at similar levels for both TRA and NDPA awardees: *A gradual but building shift in practices or thinking* (74.6% and 72.0%; χ_1^2 = 0.02, p = 0.89); *A delayed but significant shift in practices or thinking* (17.2% and 25.0%; χ_1^2 = 0.73, p = 0.39); and *An immediate and significant shift in practices or thinking* (32.2% and 23.0%; χ_1^2 = 0.99, p = 0.32) .

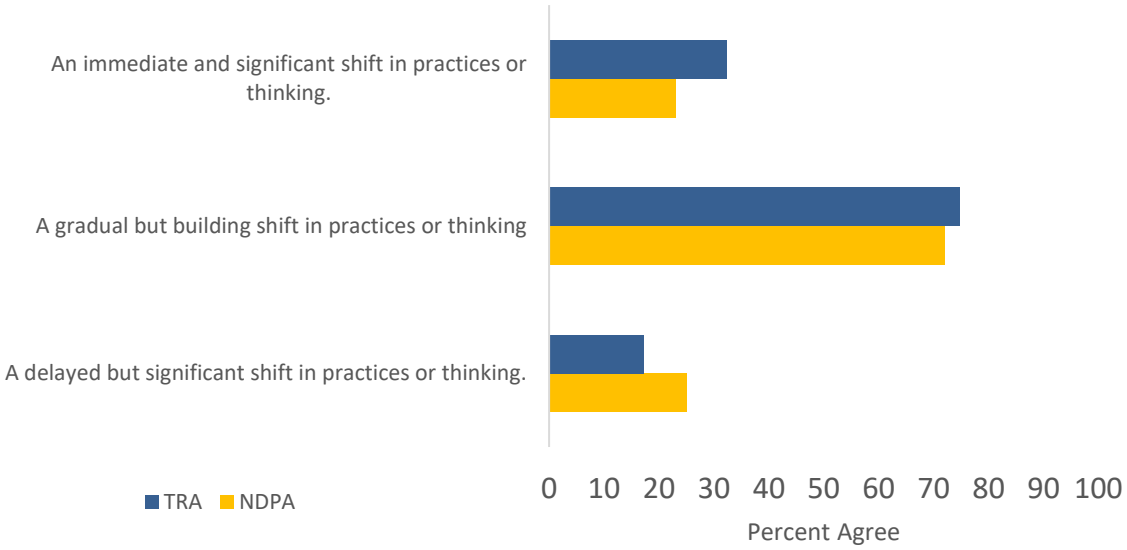


Figure 42. Senior Scientist Review: Timing of the Onset of Impact

For the next set of questions, reviewers were asked about the likelihood that each award demonstrated innovation. Reviewers reported "likely" significantly more for TRA recipients on the following question: *The development of new therapies clinical tools or strategies* (58% and 37%; χ_1^2 = 5.28, p = 0.02).

Reviewers reported that TRA and NDPA awardees were similarly likely to result in the following: *The discovery of new phenomenon or advancement of a theoretical concept* (65.0% and 69.2%; χ_1^2 = 0.12, p = 0.73); *A new synthesis of disparate ideas* (47% and 61%; χ_1^2 = 2.19, p = 0.14); *The development of a new methodology* (60.0% and 64.1%; χ_1^2 = 0.1, p = 0.75); *The development of a new technology* (53.3% and 48.7%; χ_1^2 = 0.13, p = 0.71); and *A change in how research is conducted* (50.0% and 41.0%; χ_1^2 = 0.77, p = 0.38).

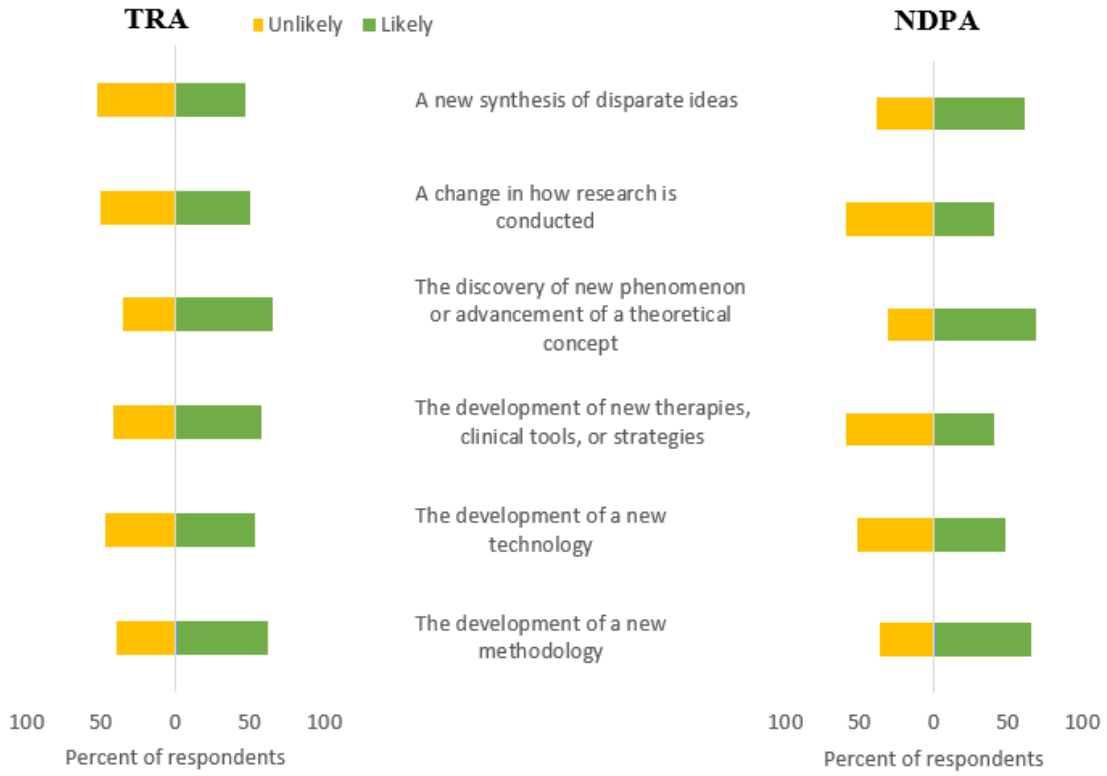


Figure 43. Senior Scientist Review: Likeliness of Research Resulting in Innovation

2. Assessment of Scientific Impact

Reviewers were asked to indicate whether they felt qualified to assess the scientific impact of each award they reviewed. Of the 60 and 78 individual TRA and NDPA reviews that were conducted, reviewers indicated that they did *not feel qualified to assess the scientific merit of the award* 13 (22%) and 24 (31%) times, respectively.

3. Descriptive Statistics for TRA and R01 Awards

a. Paradigm Shift

The first set of questions addressed the paradigm-shifting nature of the awardees' research. These questions included novel inventions, novel combinations of ideas and approaches, or use of technology that significantly changed research practice or thinking in their field (Figure 44).

Senior scientist reviewers were significantly more likely to agree that with the following statements for TRA awardees: *The awardees research involved a novel combination of ideas, disciplines or approaches that significantly changed research practice* (69% and 44% for TRA and R01 awardees, respectively; χ_1^2 = 8.06, $p < 0.001$); *The awardees research led to a novel invention or a new technology* (58% and 32% for TRA and R01 awardees, respectively; χ_1^2 = 8.4, $p < 0.01$); *Led to the development of a new methodology that significantly changed research in their field* (58% and 35%; χ_1^2 = 6.01, p = 0.01); and *Required the use of equipment technique or model that was novel and significantly changed research in their field* (64% and 40%; χ_1^2 = 7.08, p = 0.01).

Reviewers were likely to identify the TRA and R01 research outputs as similar: *One or more of the awardees research ideas challenged existing science technology paradigms* (61% and 48% for TRA and R01 awardees, respectively; χ_1^2 = 1.98, $p = 0.16$); *The awardees research furthered existing practices or thinking in their field* (93.3% and 91.6%; χ_1^2 = 0.01, p = 0.94); *Aided in the development of new therapies clinical tools or strategies that significantly changed research or clinical practice* (36% and 50%; χ_1^2 = 2.34, p = 0.13); *Took the next steps in their established area of investigation* (90.0% and 81.9%; χ_1^2 = 1.22, p = 0.27).

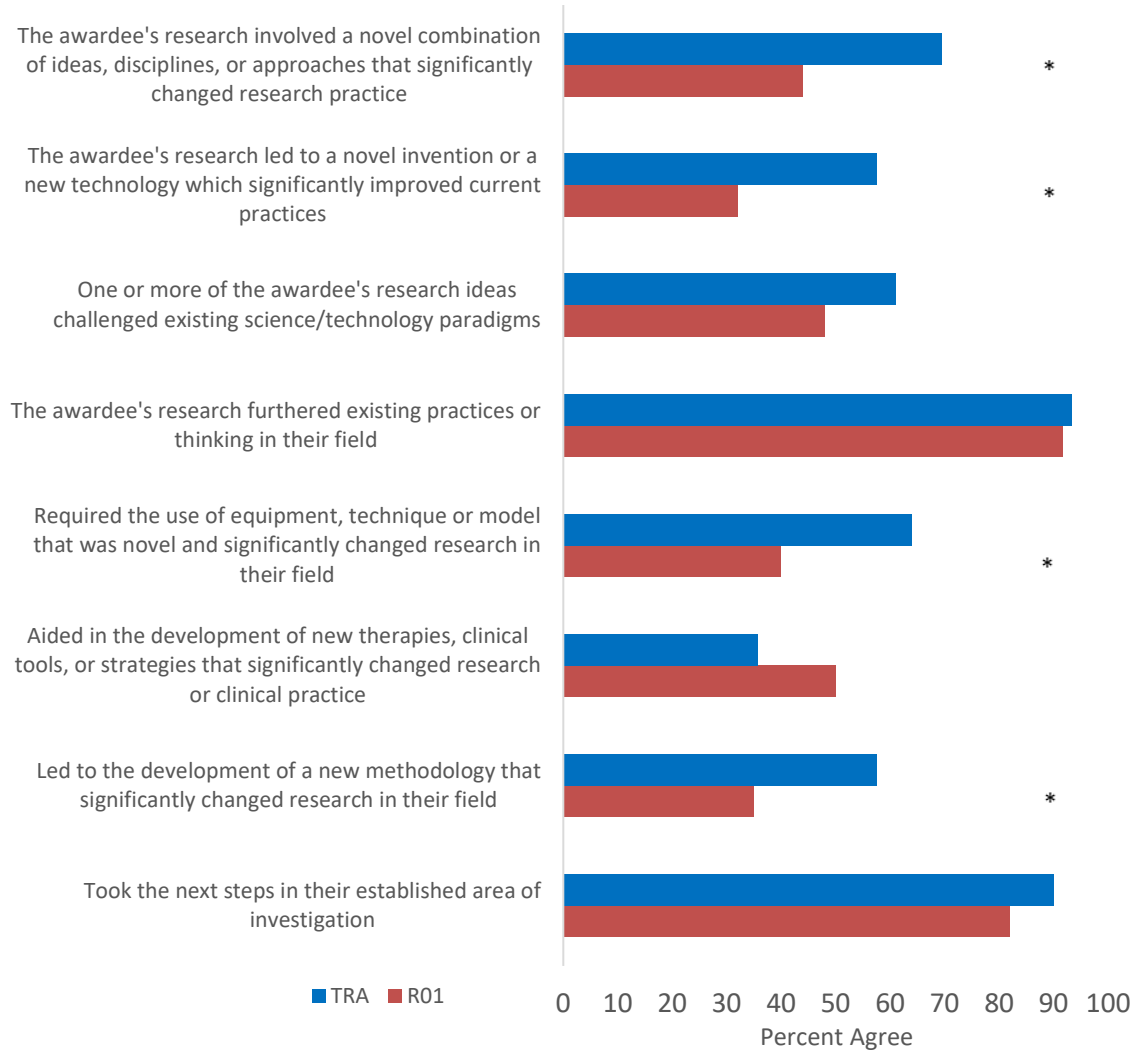


Figure 44. Senior Scientist Review: Paradigm-Shifting Research Outputs by Type

b. Impact

This question block explored the timing of impact (Figure 45). reviewers reported that TRA awardees were significantly more likely to result in: *An immediate and significant shift in practices or thinking* (32% and 15% for TRA and R01 awardees, respectively; χ_1^2 = 5.02, p = 0.03).

Reviewers also agreed with the statements at similar levels for both TRA and R01 awardees: *A gradual but building shift in practices or thinking* (74.6% and 74.7% for TRA and R01 awardees, respectively; $\chi_1^2 < 0.001$, p = 1); and *A delayed but significant shift in practices or thinking* (17.2% and 19.5%; χ_1^2 = 0.01, p = 0.9).

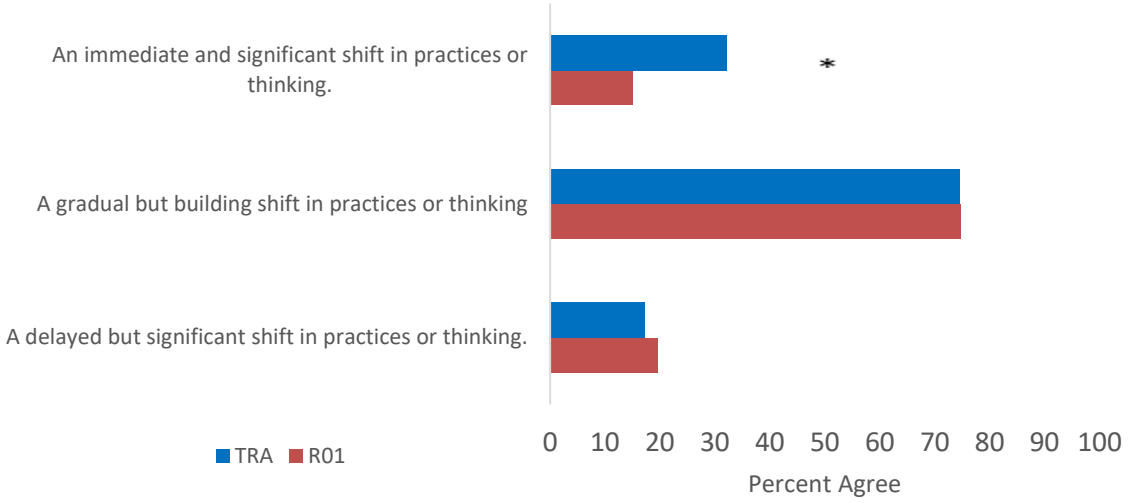


Figure 45. Senior Scientist Review: Timing of the Onset of Impact

For the next set of questions, reviewers selected whether they felt grantee's research was unlikely or likely to result in a series of statements characterizing transformative research (Figure 46). The following question was statistically greater for TRAs: *The discovery of new phenomenon or advancement of a theoretical concept* (65% and 31% for TRA and R01 awardees, respectively; $\chi_1^2 = 14.6$, $p < .0001$); *The development of a new methodology* (60% and 42% awardees; $\chi_1^2 = 3.74$ $p = 0.05$); The *development of a new technology* (53% and 30% awardees; $\chi_1^2 = 6.89$, $p = 0.01$).

Reviewers reported *likely* at similar levels for TRA and R01 awardees: *A new synthesis of disparate ideas looking into the future (47% and 34% for TRA and R01 awardees, respectively; $\chi_1^2 = 1.93$, $p = 0.16$); The development of new therapies clinical tools or strategies* (58.3% and 59% awardees; $\chi_1^2 < 0.001$, $p = 1$); and *A change in how research is conducted* (50% and 39% awardees; $\chi_1^2 = 1.42$, $p = 0.23$);

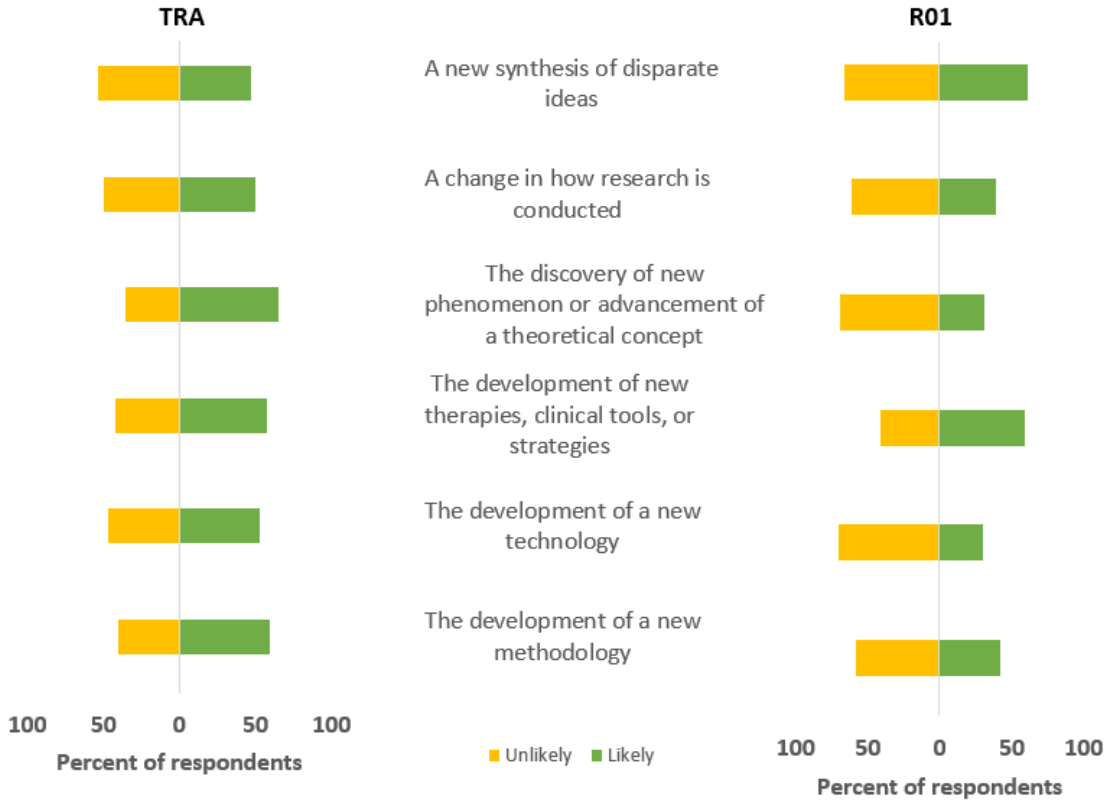


Figure 46. Senior Scientist Review: Likeliness of Research Resulting in Innovation

c. Senior Scientist Review of Scientific Impact

Reviewers were asked to indicate whether they felt qualified in assessing the scientific impact of each award they reviewed. Of the 60 and 83 individual TRA and R01 reviews that were conducted, reviewers indicated that they did *not feel qualified to assess the scientific merit of the award* 13 (22%) and 20 (24%) times, respectively.

4. Characteristics of Transformative Research

Reviewers were asked to provide up to five key words that they associate with the word *transformative*. One hundred and sixty-five unique words were provided to describe transformative research (please see Appendix G for full list). The top 10 words used to describe "transformative" were *paradigm* (n = 29), *impact* (n = 27), change (n = 23), shift (n = 23), innovative (n = 19), revolutionary (n = 19), technology (n = 7), creative (n = 6), groundbreaking (n = 6), and challenge (n = 5). A total of 58 words were provided more than once and 107 words were provided once by reviewers to describe "transformative" research.

5. Limitations to Data Interpretation

The solicitation of senior scientists has several limitations that require consideration when contextualizing the results. First, the reviewer evaluates, by award, the 3 research papers that received the most citations during the time frame for this study. It is possible that the 3 most highly cited papers do not reflect the most transformative research outputs for that award, especially if transformative impact builds slowly over time.

Second, experts note that, to add rigor to a subjective process, solicitations of subject matter experts should be probabilistic in nature. These elicitations, however, were not done in a probabilistic manner, and may be subject to uncontrollable biases including the anchoring heuristic (i.e., people tend to adjust their judgements based on initial given values) and the range-frequency heuristic (i.e., total probabilities tend to be assigned evenly between categories given).[44] Perhaps the strongest bias is the floating definition of "transformative." As Morgan notes:

> *There is clear evidence that without some quantification, the use of qualitative words such as "likely" and "unlikely" to describe uncertainty can mask important, often critical, differences between the views of different experts. The problem arises because the same words can mean very different things to different people, as well as different things to the same person in different contexts*

Because the survey did not explicitly instruct the reviewrs how to think of "transformative" research, the results may be biased and likely would not be reproducible.[45,46] These considerations are discussed in more detail in the interpretation of findings and Appendix G.

---

[44] Anthony O'Hagan, "Expert knowledge elicitation: subjective but scientific," *The American Statistician* 73, (Mar 20, 2019): 70-71.

[45] O'Hagan, "Expert knowledge," 70-71.

[46] Morgan M. Granger, "Use (and abuse) of expert elicitation in support of decision making for public policy," *PNAS* 110, no. 20, (May 20, 2014): 7177.

5. Integration of Findings

The multi-modal study design employed in this evaluation consists of a logic model, surveys of awardees and comparison groups, bibliometric analyses, and a senior scientist review. STPI integrated the results of these analyses and assessed the degree to which the TRA awardee research findings were more transformative, innovative and impactful—and more likely to introduce new, or evolve existing, scientific paradigms. The results are presented by study question followed by other considerations pertinent to the overall conclusions:

- Do the scientific outputs produced by TRA awardees represent a *paradigm shift* for biomedical research?
- Are the outputs *more impactful* than research produced by comparison groups?

The TRA-NDPA comparison and the TRA-R01 comparison are separate analyses, and the results should not be conflated. As noted in the study design section of this report, small sample size only allows for the detection of large differences between groups. Criteria for the NDPA and TRA awards are similar, and the R01 comparison group is funded through a mechanism described as the traditional NIH award made to support a discrete, specified, circumscribed project.[47]

For all tables in this section, the gray fill color indicates that the analytical results for the groups are not significantly different (NS); blue indicates TRA results are significantly different from the comparison group (TRA>comparison group); and the orange indicates that the comparison group results are significantly different from the TRA results (comparison group>TRA).

Study question 1: Do the scientific outputs produced by TRA awardees represent a *paradigm shift* for biomedical research?

In the absence of clear, quantifiable criteria for paradigm-shifting research, comparison groups were used to determine whether TRA awardee research was more paradigm shifting than traditional NIH biomedical research. As described in detail in the study design section, NDPA awardees are exceptionally creative scientists pursuing new research directions, suggesting research outputs more likely to be paradigm shifting than the traditional R01 award that supports discrete, specified, circumscribed research—which

[47] https://grants.nih.gov/grants/funding/r01.htm

is more likely to produce incremental advances in biomedical research. Additionally, STPI used the characteristics of transformative research to develop a series of survey options that survey respondents and senior scientist reviewers could select to describe the transformative potential of the research under consideration. The pairing of these two survey approaches provides the researcher's perspective on their research results, and the senior scientist review provides an external perspective on the three most cited research publications attributed to each survey respondent's award. The results are analyzed in aggregate.

In Table 20, the results for the TRA and NDPA awardee survey and the senior scientist review identify no significant differences between the TRA and NDPA awardees for any response option, suggesting that their research results have similar likelihood to shift a biomedical research paradigm. The results for the TRA-R01 comparison provide a mixed response pattern with agreement between the two surveys only for the option indicating that the TRA awardee research was more likely to involve a novel combination of ideas, disciplines, or approaches that changed research practices. Although TRA awardee research was assessed by senior reviewers as more likely to be paradigm shifting by four survey options and R01 awardees assessed their own research as more transformative by one criterion--more likely to produce new therapies, clinical tools, and strategies that could significantly change research or clinical practice—the mixed response pattern between the awardee and senior scientist surveys prevents definitive conclusions based solely on these data.

Table 20. Summary of Survey Results by Comparison Group and Survey Mechanism

	TRA and NDPA Comparison		TRA and R01 Comparison	
Survey Options for Paradigm shift:	**Awardee Survey**	**Senior Scientist Review**	**Awardee Survey**	**Senior Scientist Review**
Novel combination of ideas, disciplines or approaches that significantly changed research practices	NS	NS	TRA>R01	TRA>R01
Novel intervention or new technology which significantly improved current practices	NS	NS	NS	TRA>R01
Research ideas challenged existing science and technology paradigms	NS	NS	NS	NS
Furthered existing practices or thinking in my field	NS	NS	TRA>R01	NS
Required novel use of equipment, technique or model that significantly changed research in my field	NS	NS	NS	TRA> R01
Developed new therapies, clinical tools, strategies that significantly changed research or clinical practice	NS	NS	R01>TRA	NS
Developed new methodology that significantly changed research in my field	NS	NS	NS	TRA>R01
Took the next steps in established research field	NS	NS	NS	NS

Study question 2: Are the outputs *more impactful* than research produced by comparison groups?

To assess the impact of TRA research, STPI first evaluated the bibliometric factors most directly associated with publications and citations—and through the surveys, the timeframe across which impact might occur. Secondly, STPI evaluated the factors that influence publications. For the total number of publications or citations per award, TRA awardees produced significantly more publications and received significantly more citations than either comparison group (Table 21). This pattern held when publications and citations per award were normalized per year, with the exception of TRA and NDPA awardee publications, which were not significantly different. There was no significant difference in the TRA and NDPA awardee costs per publication or citation; however, the cost per R01 citation was significantly higher than for TRA awardees. The TRA awardees RCR, as a measure of productivity and impact, was significantly higher than NDPA awardees, although when the number of publications per award was factored in, there was no statistical difference in the weighted RCR values. TRA awardee RCR and weighted RCR values were significantly higher than R01values. As a final measure of impact, the TRA awardee altmetric attention scores were significantly higher than NDPA scores and R01 scores.

Table 21. Research Impact

	TRA and NDPA Comparison		TRA and R01 Comparison	
Bibliometric Measures of Impact	**Publications**	**Citations**	**Publications**	**Citations**
Total number per award	TRA>NDPA	TRA>NDPA	TRA>R01	TRA>R01
Total number per award per year	NS	TRA>NDPA	TRA>R01	TRA>R01
Cost per publication	NS	NS	NS	R01>TRA
RCR (productivity and impact)		TRA>NDPA		TRA>R01
Weighted RCR		NS		TRA>R01
Altmetrics score (article level metrics)		TRA>NDPA		TRA>R01

Closer examination of metrics for publications and citation at the award level demonstrated the influence of award duration, total direct costs, multiple or single PIs, and area of science on bibliometric measures. When these four additional factors are held constant, TRA and NDPA awardees produce similar numbers of publications per award; however, TRA awardees had significantly fewer citations per award than NDPA awardees. The lack of agreement among different publication and citation metrics precludes definitive

conclusions on whether TRA research outputs were more impactful than NDPA research outputs.

For the R01 comparison group, bibliometric analysis demonstrated that TRA awardees produced significantly more publications, received significantly more citations, and had higher Altmetric attention scores than R01 awards, suggesting that TRA research outputs were more impactful than R01 research outputs (Table 21). The bibliometric findings for publications and citations were corroborated by the multi-variable analysis.

When considering the amount of time that may be necessary for transformative research to impact a field of science or clinical practice, no significant difference between TRA and NDPA awardees is identified for either survey, even when the results for gradual and delayed impact are grouped together (Table 22). TRA awardees are more likely to report a delayed impact for their research, and reviewers considered the TRA research outputs to have significantly more immediate impact; however, no other measures demonstrated a significant difference. Additionally, nso significant difference was detected for TRA and R01 groups in either survey when gradual and delayed impact were combined.

Table 22. Timeframe for Impact as a Factor is Assessing Transformative Potential

	TRA and NDPA Comparison			TRA and R01 Comparison	
Survey Options on the Timing of Research Impact:	**Awardee Survey**	**Senior Scientist Review**		**Awardee Survey**	**Senior Scientist Review**
Immediate	NS	NS		NS	TRA>R01
Gradual but building	NS	NS		NS	NS
Delayed	NS	NS		TRA>R01	NS
Combined gradual and delayed	NS	NS		NS	NS

Interestingly, TRA awardees were more likely than R01 awardees to report a future research impact for discovery of a new phenomenon or advancement of a theoretical concept, a finding underscored by the senior scientist reviewers (Table 23). And while these reviewers were more likely to identify transformative impact for TRA research outputs of new clinical tools (TRA>NDPA) and new technologies (TRA>R01), these observations were not supported by the awardee survey analysis.

Table 23. Potential for Future Research Impact

	TRA and NDPA Comparison		TRA and R01 Comparison	
Survey Options for Future Research Impact:	**Awardee Survey**	**Senior Scientist Review**	**Awardee Survey**	**Senior Scientist Review**
New synthesis of disparate ideas	NS	NS	NS	NS
Change in how research is conducted	NS	NS	NS	NS
Discovery of a new phenomenon or advancement of a theoretical concept	NS	NS	TRA>R01	TRA>R01
Development of new therapies, clinical tools, strategies	NS	TRA>NDPA	NS	NS
Development of new technology	NS	NS	NS	TRA>R01
Development of new methodology	NS	NS	NS	NS

As an indirect measure of research impact, STPI asked awardees whether they received additional professional recognition during the timeframe of their award (Table 24). Awardees were queried about a general level of professional recognition and that which might be specifically attributed to their award. TRA and NDPA awardees did not differ significantly on any professional recognition options listed in the survey. TRA awardees were significantly more likely than R01 awardees to attribute invitations to serve as a journal reviewer to their award; however, there were no other significant differences.

Table 24. Professional Recognition as an Indirect Measure of Research Impact

	TRA and NDPA Comparison			TRA and R01 Comparison	
Survey Options for Professional Recognition:	**All Recognition**	**Recognition Attributed to Award**		**All Recognition**	**Recognition Attributed to Award**
Highlighted academic journal	NS	NS		NS	NS
Award/honor	NS	NS		NS	NS
Journal reviewer	NS	NS		NS	TRA>R01
Invited presentations	NS	NS		NS	NS
Popular press	NS	NS		NS	NS
Keynote speaker	NS	NS		NS	NS

Other Considerations

To examine additional factors that might influence a researcher's ability to produce transformative research results, STPI asked awardees whether they changed research direction during the course of their award (Table 25). Awardees were similarly likely to report a change in research direction during the award and an intention to continue their research post-award. They were similar in reporting that they received follow-on funding to do so.

Table 25. Continuation of Research Direction

Survey Options for Research Direction:	TRA and NDPA Comparison		TRA and R01 Comparison
Changed significantly from award application	NS		NS
Continuing this research post-award	NS		NS
Survey Options for Receiving Post-award Funding on Award Topic:			
I received follow on funding: yes/no	NS		NS

STPI also asked whether the grant mechanism—R01 for TRA and R01 awardees and DP1 for NDPA awardees[48]—provided the flexibility needed to perform transformative research (Table 26). TRA respondents were significantly more likely than R01 awardees to report that the R01 mechanism allowed them to redirect their research. Although the results did not achieve a significant difference, more TRA awardees agreed with the statements that the R01 mechanism supported non-traditional research and that funding was flexible.

Table 26. Supportive Grant Mechanism

Survey Options for Grant Mechanism:	TRA and NDPA Comparison		TRA and R01 Comparison
Support non-traditional research	NS		NS
Allow PI to redirect research	NS		TRA>R01
Allow flexible use of funding	NS		NS

[48] NIH Director's Pioneer Award (DP1) applications are meant to support individual scientists of exceptional creativity who propose pioneering—and possibly transformative—research that, if successful, will have a major impact on a broad area of relevance to the NIH. Pioneer Award applications do not require preliminary data, specific aims, or a detailed research plan. (https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-011.html)

6. Summary and Recommendations

TRA and NDPA survey respondents and senior scientist reviewers report no significant differences in transformative research outputs. Using the bibliometric and multi-variable analyses in this evaluation, STPI measured no significant difference between TRA and NDPA research impact; however, TRA awardees' RCR and altmetric scores were significantly higher than NDPA scores. In sum, these data suggest that TRA awardee research outputs and impact are comparable to NDPA research that is recognized by NIH as demonstrating exceptional creativity, new research directions, and pioneering approaches to major challenges in biomedical, social science, and behavioral research.[49]

The mixed response pattern for the TRA and R01 awardee comparison indicates several areas for which the TRA research was reported as more likely to be transformative; however, the small sample size precludes definitive conclusions. In contrast, bibliometric and multi-variable analysis, as well as RCR and Altmetric scores, demonstrate that overall, TRA research was considered more impactful than R01 research.

Because the purpose of this evaluation is to provide data and context to inform the NIH Director's future TRA policy and investment decisions, STPI identified three areas that have implications for policy: transformative research, research impact, and programmatic impact.

A. Transformative Research

Several factors could contribute to the differences between the awardee survey results and the senior scientist review results. While small sample size only allows for the detection of large differences between groups, individual perceptions of the characteristics of transformative research should be considered. Variability in perceptions of transformative research was underscored by the lengthy list of terms senior scientist reviewers used to describe transformative research. To understand the frequency and variability of the descriptors, STPI applied a word cloud algorithm to the list. STPI first removed stop-words,[50] words such as prepositions or conjunctions that do not add meaning

[49] https://commonfund.nih.gov/pioneer

[50] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, "Dropping Common Terms Stop Words." *Introduction to Information Retrieval*, Cambridge University Press. 2008. https://nlp.stanford.edu/IR-book/html/htmledition/dropping-common-terms-stop-words-1.html

to the text (e.g., “and,” “the,” and “is”), and stemmed words[51]—verbs with different tenses or words with different endings are counted as the same word (e.g., “train,” “training,” and “trained”)—before applying a word cloud algorithm using the wordcloud package[52] in R to the list. The result is a graphical representation of word frequency that gives greater prominence to words that appear more often (Figure 47).[53,54] Five terms were most commonly used—paradigm, impact, change, shift, innovative—and 16 additional words form a second ring of descriptors; however, an additional 144 words are cited 1–3 times by reviewers. The breadth of terms used by reviewers to describe transformative research further complicates an assessment of the differences in transformative potential between comparison groups. The full list of terms used by senior scientist reviewers to describe transformative research and the frequency at which they were used can be found in Appendix G.

Figure 47. Word Cloud Corresponding to the Words Used by Senior Scientist Reviewers

---

51 Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, “Stemming and Lemmatization.” *Introduction to Information Retrieval*, Cambridge University Press. 2008. https://nlp.stanford.edu/IR-book/html/htmledition/stemming-and-lemmatization-1.html

52 Ian Fellows (2018). wordcloud: Word Clouds. R package version 2.6. https://CRAN.R-project.org/package=wordcloud

53 F. Heimerl, S. Lohmann, S. Lange and T. Ertl, "Word Cloud Explorer: Text Analytics Based on Word Clouds," 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, 2014, pp. 1833-1842, doi: 10.1109/HICSS.2014.231.

54 https://www.betterevaluation.org/en/evaluation-options/wordcloud

The variability in terminology used by senior scientist reviewers in an open response survey question to describe transformative research reflects the individuality of their interpretations of *transformative research* and challenges the specificity through which TRA initiative research results can be evaluated. While acknowledging the NIH interest in providing PIs with the flexibility to propose a broad array of transformative research ideas, STPI suggests NIH consider approaches to develop more specific FOA language and review criteria to define transformative research, or the characteristics of transformative research. It is also likely that characteristics of transformative research might vary across different disciplines within biomedical research. An analysis of these differences in terminology would also contribute to a general and scientific discipline-specific understanding of transformative research. This more specific language could then be used in awardee and reviewer survey development, perhaps increasing the likelihood of identifying differences between awardee groups.

B. Research Impact

As described in the study design, research impact is generally assessed through an analysis of publication and citation metrics over a pre-determined period of time. The timeframe established for the evaluation of 5-year TRA awards made in 2010–2012 begins 1 year after an award's project start date and ends 1 year after the project end date. Recognizing that this may be insufficient time for transformative research to impact scientific direction, STPI performed an anecdotal review of several significant research achievements to examine the length of time between initial discovery, recognition of high potential impact, and the identification of what is often many discoveries contributing to the overall transformative research result. For example, 45% of Nobel prizes for physiology or medicine[55] were awarded more than 20 years after the original discovery because it often takes that long for a discovery to be recognized as having "changed the scientific paradigm and [to be] of great benefit for humankind." [56,57]

It is also challenging to recognize when a specific research finding is "transformative." Because scientific progress builds on previous research, very rarely—if ever—does an idea or theory emerge *de novo*. The discovery of CRISPR serves an example. The first recorded existence of CRISPR was in 1987 when an unusual repetitive

---

55 Fortunato, Santo. 2014. "Growing time lag threatens Nobels." *Nature* 508: 186.

56 Wilkins, Alasdair. 2016. "How long does it take to win a Nobel prize?" *Vocativ*, October 13. Accessed December 27, 2020. Available at: https://www.vocativ.com/366734/how-long-does-it-take-to-win-a-nobel-prize/index.html.

57 Nobel Foundation. 2020. "Nomination and selection of Medicine Laureates." Accessed December 27, 2020. Available at: https://www.nobelprize.org/nomination/medicine/.

DNA sequence was observed in *E. coli*.[58] An abridged timeline of discoveries surrounding CRISPR shows that it unfolded over more than 20 years with research performed across 9 countries with hundreds of researchers adding to the collective scientific knowledge and understanding of CRISPR.[59] Moreover, current multi-disciplinary and interdisciplinary research makes it less likely that a major scientific breakthrough will be made by a single researcher or small group of investigators. Biomedical research, like many other disciplines, has become more collaborative through time as researchers build collective scientific knowledge.

These data suggest several options for NIH consideration. NIH could evaluate TRA research over a longer timeline, perhaps assessing research impact for the 2010–2012 cohort 10 years after project end date and every 5 years after that. This approach would define a transformative research trajectory by tracking impact over a timeline that is consistent with paradigm-shifting discoveries. NIH could redefine the evaluation timeline to increase the TRA cohort size by adding awardees to the 2010–2012 group as they meet the criteria for *project end date plus 1 year*. For example, the evaluation could be done every 3 years so that the 2010–2012 cohort would evolve to the 2010–2015 cohort, followed by the 2010–2018 cohort. This approach would increase the statistical power to detect differences in research outputs between comparison groups and increase the timeframe for the analysis, thus increasing the likelihood that transformative impact will be measurable. Because TRA follow-on awards are funded through the traditional R01 review mechanism, this approach would also evaluate the likelihood of competitive renewal of transformative research grants. Finally, NIH might refocus the goal of the initiative and its concomitant evaluation from an emphasis on research outputs to spurring novel, paradigm-shifting thinking. This focus could be evaluated strategically and operationally through an assessment of the TRA application research strategy, and it would be independent of the time lag associated with publications and the often multi-decadal recognition of transformative research.

C. Programmatic Impact

Although challenged by the limitations of the study outlined in this report, results of the evaluation suggest that transformative research is occurring in all three comparison groups—TRA, NDPA, and R01—although to differing degrees. NIH might consider the relationship of the NDPA and R01 programs to the TRA initiative and determine whether greater distinction between the programs benefits the NIH mission or whether an emphasis on transformative potential regardless of mechanism would be more effective. This

---

[58] Ishino, Y., Krupovic, M., Forterre, P. 2018. "History of CRISPR-Cas from encounter with a mysterious repeated sequence to genome editing technology." *Journal of Bacteriology* 200(7): e00580-17.

[59] Lander, E.S. 2016. "The heroes of CRISPR. *Cell* 164:18-28.

consideration is further strengthened by the overall conclusion in this evaluation that transformative research was being performed though both the DP1 funding mechanism (NDPA awardees) and R01 funding mechanism (TRA and R01 awardees).

STPI also notes the use of the term *innovative* as a characteristic of transformative research (Appendix A) and that one of the four HRHR initiatives is the New Innovator Award (NIA). An examination of the conceptual and operational similarities and differences between these two programs could maximize their unique characteristics. In the 2016 STPI report, *An Outcome Evaluation of the National Institutes of Health Director's New Innovator Award Program for Fiscal Years 2007–2009*, innovative research is defined as duplicable knowledge considered new in the context in which it is introduced and demonstrated to be useful in practice. As noted in the scope of the TRA evaluation, STPI identified paradigm shift, impact, and innovation as key descriptors of transformative research. Exploration of the goals and characteristics of these research programs could enhance the biomedical benefit of each.

D. Concluding Thoughts

Several components of this multi-modal analysis demonstrate that, despite the definitional challenges and limitations to the study design that are detailed in the full report, the 2010–2012 TRA awardees have produced impactful biomedical research that aligns with the goals of the initiative. Determination of the degree of transformative impact will require the test of time; however, numerous TRA awardees acknowledged in the free response survey questions the importance of the TRA initiative in funding research they believed to be outside the parameters of the traditional R01 mechanism. These comments and the results of this evaluation confirm the role of the TRA initiative in spurring transformative research at NIH.

Appendix A. Characteristics of Transformative Research

To identify a core set of characteristics of transformative research that would inform the awardee survey, STPI examined the descriptions of transformative research in the 2010–2012 TRA FOA instructions to the applicant and review criteria and the definition and characteristics of transformative research offered by the National Science Foundation (NSF) and the National Science Board (NSB). The specific language used to inform STPI's work includes:

- NIH TRA FOA (2010-2012): ...*projects* [that] *must have the potential to create or overturn fundamental scientific paradigms through novel approaches, transform the way research is conducted through the development of novel tools or technologies, or lead to major improvements in health through the development of highly innovative therapies, diagnostic tools, or preventive strategies* and research that is *groundbreaking, exceptionally innovative, original, and/or unconventional*. [60]

- NIH TRA website (2020)*: transformative projects that are inherently risky and untested but have the potential to create or overturn fundamental paradigms...*[61]

- NSF (2019): ...*research* [that] *involves ideas, discoveries, or tools that radically change our understanding of an important existing scientific or engineering concept or educational practice or leads to the creation of a new paradigm or field of science, engineering, or education. Such research challenges current understanding or provides pathways to new frontiers*.[62] NSF also lists three characteristics of transformative research that further explicate the term. Transformative research challenges conventional wisdom, leads to unexpected insights that enable new techniques or methodologies, or redefines the boundaries of science, engineering, or education.[63]

- NSB (2005): ...*research that has the capacity to revolutionize existing fields, create new subfields, cause paradigm shifts, support discovery, and lead to*

[60] RFA-RM-10-010: NIH Common Fund Transformative Research Projects Program (R01)

[61] https://commonfund.nih.gov/tra

[62] https://www.nsf.gov/about/transformative_research/definition.jsp

[63] https://www.nsf.gov/about/transformative_research/characteristics.jsp

> *radically new technologies. The report further explains "In practice, distinguishing between innovative and transformative research is difficult at best and, some would argue, only possible in hindsight. Indeed, the two forms of scientific progress do exist side-by-side and, often, proceed hand-in-hand and overlap each other*.[64]

To understand the concept of paradigm shift, STPI considered Thomas Kuhn's foundational definition of paradigms as universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners.[65] Further literature review defined paradigm shift as radical generative (generating important new ideas) or radical destructive (making existing ideas obsolete or less salient), and added multi-disciplinarity and impact (high, broad, builds over time) to the list of characteristics.[66]

The TRA FOA also identifies several specific research outcomes: novel scientific paradigms, new and improved clinical approaches, highly innovative therapies, diagnostic tools, or preventive strategies, transformative technologies, approaches, tools, or technologies. FOA review criteria stipulate that the application should show clear transformative potential by describing research that has the potential for exceptional consequences for the field and size of community affected, and that proposes or challenges a fundamental paradigm in the field. The research proposal should demonstrate use of novel theoretical concepts, approaches or methodologies, instrumentation, or interventions.

---

64 National Science Board, *2020 Vision for the national science foundation*, 2005.

65 Thomas Kuhn, Structure of Scientific Revolutions

66 Pierre-Antoine Arrighi, Pascal Le Masson, Benoit Weil. Managing radical innovation as an innovative design process: generative constraints and cumulative sets of rules in Creativity and Innovation Management, Wiley, 2015, 24 (3), pp.373-390.

With this background information, STPI next developed a subset of transformative research characteristics and examined their relationship to several historical examples of transformative biomedical research (e.g., discovery of DNA, pig brain resuscitation experiments, Alzheimer sub-types and drug failures) to determine whether a characteristic was necessary and sufficient to define research outputs as transformative (Table 27). STPI determined that paradigm shift was a fundamental attribute of transformative research and both necessary and sufficient to determine that research is transformative. Impact and innovativeness are often corollaries of paradigm-shifting research but not sufficient for research to be considered transformative.

STPI also recognized that research impact, as measured through citation rates over time, can be immediate and high, broad with large cumulative impact, or build more slowly over time. Research with immediate and high impact could be recognized through a large number of citations in a single or limited number of scientific disciplines within a short time after publication of research findings. Research of broad cumulative impact could have citations in multiple disciplines, whereas research that builds over time might have limited citations following research publication that steadily increases over a longer period of time.

Table 27 - Transformative Research Attributes

Attribute	Necessary (Y/N)	Sufficient (Y/N)
Paradigm shift	Yes	Yes
Impactful	Yes	No
Innovative	Yes	No
Risky	No	No
Creative	No	No
Translational	No	No
Multidisciplinary	No	No

Appendix B. Survey Questions Aligned with the Concepts of Paradigm Shift and Impact

Survey Questions	Response Choices (If Applicable)
Paradigm Shift	
Considering your TRA/NDPA/R01 supported research, please select whether you agree/disagree with the following statements:	My research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice. One or more of my TRA/NDPA/R01 research ideas challenged existing science/technology paradigms. My research led to a novel invention or a new technology which significantly improved current practices. My research furthered existing practices or thinking in my field.
Considering your TRA/NDPA/R01 supported research, please select whether you agree/disagree with the following statements: My research…	Led to the development of a new methodology that significantly changed research in my field. Took the next steps in my established area of investigation. Required the use of equipment, technique or model that was novel and significantly changed research in my field. Aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice.
Impact/Innovation	
It is well established that some research has immediate impact, and other research builds gradually. Considering the potential or realized effects of your TRA/NDPA/R01 research, please select whether you agree/disagree with the following statements: My research resulted in…	A delayed but significant shift in practices or thinking. A gradual but building shift in practices or thinking. An immediate and significant shift in practices or thinking.
Looking into the future and considering the potential outcomes of your TRA/NDPA/R01 supported research, please select whether the outcome is likely/unlikely: My research could potentially result in…	The development of a new methodology. The development of a new technology. The development of new therapies, clinical tools, or strategies.

Survey Questions	Response Choices (If Applicable)
	The discovery of new phenomenon or advancement of a theoretical concept.
	A change in how research is conducted.
	A new synthesis of disparate ideas.
Please indicate whether or not the following took place for you once your TRA/NDPA/R01 research was published. Please select all that apply:	I have been invited to be a keynote speaker to share my research.
	My research has been featured in the popular press/media.
	I have been asked to give an invited presentation about my research.
	I have been invited to serve as a regular journal reviewer.
	I received an award/honor.
	My research has been listed as a highlight in an academic journal.
Describe any other recognition resulting from your TRA/NDPA/R01.	
Which of the changes that you selected do you think could be attributed to your TRA/NDPA/R01. Select all that apply.	I have been invited to be a keynote speaker to share my research.
	My research has been featured in the popular press/media.
	I have been asked to give an invited presentation about my research.
	I have been invited to serve as a regular journal reviewer.
	I received an award/honor.
	My research has been listed as a highlight in an academic journal.
Considering the various aspects of the TRA/NDPA/R01 program, please select the degree to which you agree/disagree with the following statements:	The TRA/NDPA/R01 allowed for flexibility in the use of funding.
	The TRA/NDPA/R01 allowed me the freedom to pursue nontraditional research.
	The period of the TRA/NDPA/R01 was long enough to redirect research as ideas/methods evolved.
My TRA/NDPA/R01 research plan changed significantly from what I originally proposed.	Agree
	Disagree
You indicated that your research plan changed significantly from what you originally proposed. Please indicate why those changes occurred.	
I am continuing or planning to continue the trajectory of my TRA funded research post-award.	Agree
	Disagree

Survey Questions	Response Choices (If Applicable)
You indicated that you are not continuing or planning to continue the trajectory of your TRA/NDPA/R01 funded research post-award. Please describe what factors went into your decision.	
You indicated you were continuing or planning to continue the trajectory of your TRA/NDPA/R01 funded research post-award. Please indicate whether this continuation is funded.	Continuation of my TRA/NDPA/R01 funded research IS funded post-award Continuation of my TRA/NDPA/R01 funded research IS NOT funded post-award
Is there any additional information you would like to share?	
An important component of the TRA/NDPA/R01 evaluation is the expert review panel. Please identify your top three TRA research outputs, including publications or any other outputs, that best represent your TRA/NDPA/R01. Feel free to copy and paste your three selections into the corresponding text boxes below from your CV.	Output 1 Output 2 Output 3
STPI will identify reviewers who can contribute a knowledgeable and complete review of your TRA/NDPA/R01-supported research. These individuals will represent a mix of perspectives on the science covered by your TRA/NDPA/R01.	Reviewer 1 Reviewer 2 Reviewer 3 Reviewer 4 Reviewer 5

Appendix C. Awardee Survey

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal government.

Purpose of the Survey

This survey solicits your perspectives on the Transformative Research Award (TRA)/NIH Director's Pioneer Award/R01 you received, specifically the TRA/NDPA/R01 related activities, outcomes, and elements of the award which influenced your ability to conduct impactful research.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to the NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

Instructions for the Survey

Please have a current, electronic version of your CV available for reference.

The survey is divided into two sections:

Section 1-Awardee Research Perspectives: Your perspective on the body of work that resulted from your award and the TRA/NDPA/R01 program.

Section 2-Research Program: Requested information on TRA/NDPA/R01 program outcomes.

The estimated survey completion time is 25 minutes. You will be able to move backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last submitted page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses.

While completing this survey, you will be asked several questions about your NIH application. You should only consider your TRA/NDPA/R01 award when answering these questions.

If you would like to review the relevant grant Funding Opportunity Announcement, please see the following links:

2010 TRA FOA: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-09-022.html

2011 TRA FOA: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-10-010.html

2012 TRA FOA: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-11-006.html

Follow-Up Interview

After submission, STPI staff may call you for a short (~30 minute) phone interview to discuss your responses.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at NIHgrantstudy@ida.org.

Your responses are invaluable to the study. Thank you for your participation.

1) Considering your TRA/NDPA/R01 supported research, please select whether you agree/disagree with the following statements:

	Agree	**Disagree**
One or more of my TRA/NDPA/R01 research ideas challenged existing science/technology paradigms.	()	()
My research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice.	()	()
My research furthered existing practices or thinking in my field.	()	()
My research led to a novel invention or a new technology which significantly improved current practices.	()	()

2) Considering your TRA/NDPA/R01 supported research, please select whether you agree/disagree with the following statements:

My research…

	Agree	**Disagree**
Led to the development of a new methodology that significantly changed research in my field.	()	()
Required the use of equipment, technique or model that was novel and significantly changed research in my field.	()	()
Aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice.	()	()
Took the next steps in my established area of investigation.	()	()

3) It is well established that some research has immediate impact, and other research builds gradually. Considering the potential or realized effects of your TRA/NDPA/R01 research, please select whether you agree/disagree with the following statements:

My research resulted in…

	Agree	**Disagree**
A gradual but building shift in practices or thinking.	()	()
A delayed but significant shift in practices or thinking.	()	()
An immediate and significant shift in practices or thinking.	()	()

4) Looking into the future and considering the potential outcomes of your TRA/NDPA/R01 supported research, please select whether the outcome is likely/unlikely:

My research could potentially result in…

	Unlikely	**Likely**	**Unsure**	**N/A**
The discovery of new phenomenon or advancement of a theoretical concept.	()	()	()	()
A new synthesis of disparate ideas.	()	()	()	()
The development of a new methodology.	()	()	()	()
The development of a new technology.	()	()	()	()
The development of new therapies, clinical tools, or strategies.	()	()	()	()
A change in how research is conducted.	()	()	()	()

5) Please indicate whether or not the following took place for you once your TRA/NDPA/R01 research was published. Please select all that apply:

[ ] I received an award/honor.

[ ] My research has been featured in the popular press/media.

[ ] My research has been listed as a highlight in an academic journal.

[ ] I have been invited to be a keynote speaker to share my research.

[ ] I have been asked to give an invited presentation about my research.

[ ] I have been invited to serve as a regular journal reviewer.

6) Describe any other recognition resulting from your TRA/NDPA/R01.

_____
_____
_____
_____

7) Which of the changes that you selected do you think could be attributed to your TRA/NDPA/R01. Select all that apply.

8) Considering the various aspects of the TRA/NDPA/R01 program, please select the degree to which you agree/disagree with the following statements:

	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**	**N/A**
The TRA/NDPA/R01 allowed me the freedom to pursue non-traditional research.	()	()	()	()	()	()
The period of the TRA/NDPA/R01 was long enough to redirect research as ideas/methods evolved.	()	()	()	()	()	()
The TRA/NDPA/R01 allowed for flexibility in the use of funding.	()	()	()	()	()	()

9) My TRA/NDPA/R01 research plan changed significantly from what I originally proposed.

() Agree

() Disagree

Logic: Shown if “Agree” to "My TRA/NDPA/R01 research plan changed significantly from what I originally proposed. "

10) You indicated that your research plan changed significantly from what you originally proposed. Please indicate why those changes occurred.

_____

_____

_____

_____

11) I am continuing or planning to continue the trajectory of my TRA/NDPA/R01 funded research post-award.

() Agree

() Disagree

Logic: Shown if “Disagree” to "I am continuing or planning to continue the trajectory of my TRA/NDPA/R01 funded research post-award. "

12) You indicated that you are not continuing or planning to continue the trajectory of your TRA/NDPA/R01 funded research post-award. Please describe what factors went into your decision.

_____

_____

_____

_____

Logic: Shown if “Agree” to "I am continuing or planning to continue the trajectory of my TRA/NDPA/R01 funded research post-award. "

13) You indicated you were continuing or planning to continue the trajectory of your TRA/NDPA/R01 funded research post-award. Please indicate whether this continuation is funded.

() Continuation of my TRA/NDPA/R01 funded research IS funded post-award

() Continuation of my TRA/NDPA/R01 funded research IS NOT funded post-award

14) Is there any additional information you would like to share?

_____

_____

_____

_____

15) An important component of the TRA/NDPA/R01 evaluation is the expert review panel. Please identify your top three TRA/NDPA/R01 research outputs, including publications or any other outputs, that best represent your grant.

Feel free to copy and paste your three selections into the corresponding text boxes below from your CV.

Output 1: _____

Output 2: _____

Output 3: _____

16) STPI will identify reviewers who can contribute a knowledgeable and complete review of your grant-supported research. These individuals will represent a mix of perspectives on the science covered by your grant.

	First Name	**Last Name**	**Contact Information**
Reviewer One	_____ _____	_____ _____ _____	_____ _____ _____

	_____	_____	_____
Reviewer Two	_____	_____	_____
Reviewer Three	_____	_____	_____
Reviewer Four	_____	_____	_____
Reviewer Five	_____	_____	_____

Thank You!

Appendix D. Survey Results for TRA, NDPA, and R01 Awardees

A. TRA-NDPA Comparison

Table 28. Considering your TRA/NDPA supported research, please select whether you agree/disagree with the following statements

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
My research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice.	TRA (n = 22)	22 (100.0%)	0 (0.0%)	0.76 (df = 1)	0.38
	NDPA (n = 32)	29 (90.6%)	3 (9.4%)		
One or more of my research ideas challenged existing science/technology paradigms.	TRA (n = 23)	19 (82.6%)	4 (17.4%)	0.18 (df = 1)	0.67
	NDPA (n = 31)	28 (90.3%)	3 (9.7%)		
My research led to a novel invention or a new technology which significantly improved current practices.	TRA (n = 22)	21 (95.5%)	1 (4.5%)	0.3 (df = 1)	0.58
	NDPA (n = 31)	27 (87.1%)	4 (12.9%)		
My research furthered existing practices or thinking in my field	TRA (n = 21)	17 (81.0%)	4 (19.0%)	1.1 (df = 1)	0.3
	NDPA (n = 30)	19 (63.3%)	11 (36.7%)		
My research led to the development of a new methodology that significantly changed research in my field.	TRA (n = 23)	16 (69.6%)	7 (30.4%)	0 (df = 1)	0.95

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
	NDPA (n = 31)	23 (74.2%)	8 (25.8%)		
My research took the next steps in my established area of investigation	TRA (n = 22)	15 (68.2%)	7 (31.8%)	0.03 (df = 1)	0.87
	NDPA (n = 31)	23 (74.2%)	8 (25.8%)		
My research required the use of equipment, technique or model that was novel and significantly changed research in my field.	TRA (n = 22)	16 (72.7%)	6 (27.3%)	1.07 (df = 1)	0.3
	NDPA (n = 31)	17 (54.8%)	14 (45.2%)		
My research aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice	TRA (n = 21)	15 (71.4%)	6 (28.6%)	0.68 (df = 1)	0.41
	NDPA (n = 32)	18 (56.2%)	14 (43.8%)		

Table 29. Research was considered to have gradual or delayed impact if survey respondents responded affirmatively to whether their research had "A gradual but building shift in practices or thinking" or "A delayed but significant shift in practices or thinking."

Group of interest (total number of respondents)	Delayed or gradual shift	Chisq (df)	p
TRA (N = 23)	19 (82.6%)	0 (df = 1)	1
NDPA (N = 32)	26 (81.2%)		

Table 30. It is well established that some research has immediate impact, and other research builds gradually. Considering the potential or realized effects of your TRA/NDPA research, please select whether you agree/disagree with the following statements.

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
My research resulted in a delayed but significant shift in practices or thinking	TRA (n = 22)	14 (63.6%)	8 (36.4%)	0.72 (df = 1)	0.4
	NDPA (n = 28)	22 (78.6%)	6 (21.4%)		
My research resulted in a gradual but building shift in practices or thinking	TRA (n = 21)	13 (61.9%)	8 (38.1%)	0 (df = 1)	0.97
	NDPA (n = 28)	16 (57.1%)	12 (42.9%)		
My research resulted in an immediate and significant shift in practices or thinking	TRA (n = 22)	5 (22.7%)	17 (77.3%)	0.73 (df = 1)	0.39
	NDPA (n = 29)	11 (37.9%)	18 (62.1%)		

Table 31. Looking into the future and considering the potential outcomes of your TRA/NDPA supported research, please select whether the outcome is likely/unlikely. My research could potentially result in:

Item of interest	Group	Unlikely	Likely	Unsure	N/A
The development of a new methodology	TRA	1 (4.3%)	16 (69.6%)	3 (13.0%)	3 (13.0%)
	NDPA	3 (9.7%)	21 (67.7%)	6 (19.4%)	1 (3.2%)
The development of a new technology	TRA	1 (4.3%)	16 (69.6%)	3 (13.0%)	3 (13.0%)
	NDPA	3 (9.4%)	27 (84.4%)	2 (6.2%)	NA
The development of new therapies, clinical tools, or strategies	TRA	3 (13.0%)	16 (69.6%)	3 (13.0%)	1 (4.3%)
	NDPA	4 (12.9%)	21 (67.7%)	6 (19.4%)	NA
The discovery of new phenomenon or advancement of a theoretical concept	TRA	3 (13.0%)	16 (69.6%)	3 (13.0%)	1 (4.3%)
	NDPA	5 (16.1%)	20 (64.5%)	5 (16.1%)	1 (3.2%)
A change in how research is conducted	TRA	1 (4.3%)	20 (87.0%)	2 (8.7%)	NA
	NDPA	2 (6.2%)	20 (62.5%)	6 (18.8%)	4 (12.5%)
A new synthesis of disparate ideas	TRA	4 (17.4%)	11 (47.8%)	7 (30.4%)	1 (4.3%)
	NDPA	5 (16.1%)	18 (58.1%)	8 (25.8%)	NA

Table 32. Please indicate whether or not the following took place for you once your TRA research was published. Please select all that apply

Item of interest	TRA (N = 22)	NDPA (N = 32)	Chi-square (df = 1)	p
I have been invited to be a keynote speaker to share my research	10 (45.5%)	22 (68.8%)	2.05	0.15
My research has been featured in the popular press/media	16 (72.7%)	25 (78.1%)	0.017	0.90
I have been asked to give an invited presentation about my research	14 (63.6%)	24 (75.0%)	0.35	0.55
I have been invited to serve as a regular journal reviewer	20 (90.9%)	28 (87.5%)	< 0.001	1
I received an award/honor	22 (100.0%)	32 (100.0%)	NA	NA
My research has been listed as a highlight in an academic journal	17 (77.3%)	26 (81.2%)	< 0.001	0.99

Table 33. Which of the changes that you selected do you think could be attributed to your TRA/NDPA. Select all that apply.

Item of interest	TRA	NDPA	Chi-square (df = 1)	p
I have been invited to be a keynote speaker to share my research	7 (70.0%)	19 (86.4%)	0.37	0.54
My research has been featured in the popular press/media	14 (87.5%)	20 (80.0%)	0.04	0.84
I have been asked to give an invited presentation about my research	11 (78.6%)	22 (91.7%)	0.43	0.51
I have been invited to serve as a regular journal reviewer	18 (90.0%)	26 (92.9%)	< 0.001	1
I received an award/honor	20 (90.9%)	27 (84.4%)	0.084	0.77
My research has been listed as a highlight in an academic journal	8 (47.1%)	15 (57.7%)	0.14	0.71

Table 34. Considering the various aspects of the TRA program, please select the degree to which you agree/disagree with the following statements

Item of interest	Group	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
The grant allowed for flexibility in the use of funding	TRA	1 (4.3%)	NA	NA	7 (30.4%)	15 (65.2%)
	NDPA	1 (3.1%)	NA	NA	5 (15.6%)	26 (81.2%)
The grant allowed me the freedom to pursue nontraditional research	TRA	NA	2 (8.7%)	5 (21.7%)	12 (52.2%)	4 (17.4%)
	NDPA	1 (3.1%)	NA	NA	4 (12.5%)	27 (84.4%)
The period of the grant was long enough to redirect research as ideas/methods evolved	TRA	NA	NA	NA	7 (30.4%)	16 (69.6%)
	NDPA	2 (6.2%)	1 (3.1%)	7 (21.9%)	9 (28.1%)	13 (40.6%)

Table 35. Please indicate whether you agree or disagree with the following statements

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
My research plan changed significantly from what I originally proposed	TRA (n = 23)	10 (43.5%)	13 (56.5%)	0 (df = 1)	1
	NDPA (n = 32)	14 (43.8%)	18 (56.2%)		
I am continuing or planning to continue the trajectory of my TRA funded research post-award	TRA (n = 23)	23 (100.0%)	0 (0.0%)	0.83 (df = 1)	0.36
	NDPA (n = 32)	29 (90.6%)	3 (9.4%)		

Table 36. You indicated you were continuing or planning to continue the trajectory of your TRA funded research post-award. Please indicate whether this continuation is funded

Item of interest	Group	Is funded post-award	Is not funded post-award
	TRA	16 (69.6%)	7 (30.4%)
Continuation of grant funded research	NDPA	25 (86.2%)	4 (13.8%)

Table 37. Looking into the future and considering the potential outcomes of your TRA/NDPA supported research, please select whether the outcome is likely/unlikely. My research could potentially result in:

Item of interest	Group	Unlikely	Likely	Chisq (df = 1)	p-value	Unsure	N/A
The development of a new methodology	TRA	1 (4.3%)	16 (69.6%)	< 0.001	1	3 (13.0%)	3 (13.0%)
	NDPA	3 (9.7%)	21 (67.7%)			6 (19.4%)	1 (3.2%)
The development of a new technology	TRA	1 (4.3%)	16 (69.6%)	0.96	0.33	3 (13.0%)	3 (13.0%)
	NDPA	3 (9.4%)	27 (84.4%)			2 (6.2%)	NA
The development of new therapies, clinical tools, or strategies	TRA	3 (13.0%)	16 (69.6%)	<0.001	1	3 (13.0%)	1 (4.3%)
	NDPA	4 (12.9%)	21 (67.7%)			6 (19.4%)	NA
The discovery of new phenomenon or advancement of a theoretical concept	TRA	3 (13.0%)	16 (69.6%)	0.01	0.93	3 (13.0%)	1 (4.3%)
	NDPA	5 (16.1%)	20 (64.5%)			5 (16.1%)	1 (3.2%)

Item of interest	Group	Unlikely	Likely	Chisq (df = 1)	p-value	Unsure	N/A
A change in how research is conducted	TRA	1 (4.3%)	20 (87.0%)	2.90	0.09	2 (8.7%)	NA
	NDPA	2 (6.2%)	20 (62.5%)			6 (18.8%)	4 (12.5%)
A new synthesis of disparate ideas	TRA	4 (17.4%)	11 (47.8%)	0.22	0.64	7 (30.4%)	1 (4.3%)
	NDPA	5 (16.1%)	18 (58.1%)			8 (25.8%)	NA

B. TRA-R01 Comparison

Table 38. Considering your TRA/R01 supported research, please select whether you agree/disagree with the following statements

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
My research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice.	TRA (n = 23)	23 (100.0%)	0 (0.0%)	6.99 (df = 1)	0.01
	R01 (n = 35)	24 (68.6%)	11 (31.4%)		
One or more of my research ideas challenged existing science/technology paradigms.	TRA (n = 24)	20 (83.3%)	4 (16.7%)	0.02 (df = 1)	0.88
	R01 (n = 34)	30 (88.2%)	4 (11.8%)		
My research led to a novel invention or a new technology which significantly improved current practices.	TRA (n = 23)	21 (91.3%)	2 (8.7%)	1.19 (df = 1)	0.28
	R01 (n = 34)	26 (76.5%)	8 (23.5%)		
My research furthered existing practices or thinking in my field	TRA (n = 22)	18 (81.8%)	4 (18.2%)	5.39 (df = 1)	0.02
	R01 (n = 34)	16 (47.1%)	18 (52.9%)		
My research led to the development of a new methodology that significantly changed research in my field.	TRA (n = 24)	17 (70.8%)	7 (29.2%)	1.21 (df = 1)	0.27
	R01 (n = 34)	18 (52.9%)	16 (47.1%)		
My research took the next steps in my established area of investigation	TRA (n = 23)	16 (69.6%)	7 (30.4%)	0.46 (df = 1)	0.5

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
	R01 (n = 35)	20 (57.1%)	15 (42.9%)		
My research required the use of equipment, technique or model that was novel and significantly changed research in my field.	TRA (n = 23)	17 (73.9%)	6 (26.1%)	0.14 (df = 1)	0.71
	R01 (n = 35)	23 (65.7%)	12 (34.3%)		
My research aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice	TRA (n = 22)	15 (68.2%)	7 (31.8%)	5.1 (df = 1)	0.02
	R01 (n = 35)	33 (94.3%)	2 (5.7%)		

Table 39. Research was considered to have gradual or delayed impact if survey respondents responded affirmatively to whether their research had "A gradual but building shift in practices or thinking" or "A delayed but significant shift in practices or thinking."

Group of interest (total number of respondents)	Delayed or gradual shift	Chisq (df)	p
TRA (N = 24)	19 (79.2%)	0.93 (df = 1)	0.33
R01 (N = 35)	32 (91.4%)		

Table 40. It is well established that some research has immediate impact, and other research builds gradually. Considering the potential or realized effects of your TRA/NDPA research, please select whether you agree/disagree with the following statements.

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
My research resulted in a delayed but significant shift in practices or thinking	TRA (n = 22)	14 (63.6%)	8 (36.4%)	4.68 (df = 1)	0.03
	R01 (n = 33)	10 (30.3%)	23 (69.7%)		
My research resulted in a gradual but building shift in practices or thinking	TRA (n = 21)	13 (61.9%)	8 (38.1%)	1.22 (df = 1)	0.27
	R01 (n = 34)	27 (79.4%)	7 (20.6%)		
My research resulted in an immediate and significant shift in practices or thinking	TRA (n = 23)	6 (26.1%)	17 (73.9%)	0.27 (df = 1)	0.6
	R01 (n = 33)	12 (36.4%)	21 (63.6%)		

Table 41. Looking into the future and considering the potential outcomes of your TRA/R01 supported research, please select whether the outcome is likely/unlikely. My research could potentially result in:

Item of interest	Group of interest	Unlikely	Likely	Unsure	N/A
	TRA	1 (4.2%)	17 (70.8%)	3 (12.5%)	3 (12.5%)
The development of a new methodology	R01	8 (22.9%)	23 (65.7%)	3 (8.6%)	1 (2.9%)
	TRA	1 (4.2%)	17 (70.8%)	3 (12.5%)	3 (12.5%)
The development of a new technology	R01	5 (14.3%)	22 (62.9%)	6 (17.1%)	2 (5.7%)
	TRA	3 (12.5%)	17 (70.8%)	3 (12.5%)	1 (4.2%)
The development of new therapies, clinical tools, or strategies	R01	9 (25.7%)	19 (54.3%)	7 (20.0%)	NA
	TRA	3 (12.5%)	17 (70.8%)	3 (12.5%)	1 (4.2%)
The discovery of new phenomenon or advancement of a theoretical concept	R01	16 (45.7%)	13 (37.1%)	5 (14.3%)	1 (2.9%)
	TRA	1 (4.2%)	20 (83.3%)	3 (12.5%)	NA
A change in how research is conducted	R01	3 (8.6%)	27 (77.1%)	5 (14.3%)	NA
	TRA	4 (16.7%)	12 (50.0%)	7 (29.2%)	1 (4.2%)

Item of interest	Group of interest	Unlikely	Likely	Unsure	N/A
A new synthesis of disparate ideas	R01	11 (32.4%)	11 (32.4%)	7 (20.6%)	5 (14.7%)

Table 42. Looking into the future and considering the potential outcomes of your TRA/R01 supported research, please select whether the outcome is likely/unlikely. My research could potentially result in:

Item of interest	TRA (N = 23)	R01 (N = 34)	Chi-square (df = 1)	P
I have been invited to be a keynote speaker to share my research	10 (43.5%)	13 (38.2%)	0.01	0.90
My research has been featured in the popular press/media	17 (73.9%)	22 (64.7%)	0.20	0.66
I have been asked to give an invited presentation about my research	15 (65.2%)	19 (55.9%)	0.18	0.67
I have been invited to serve as a regular journal reviewer	21 (91.3%)	18 (52.9%)	7.65	0.005
I received an award/honor	23 (100.0%)	29 (85.3%)	2.10	0.15
My research has been listed as a highlight in an academic journal	18 (78.3%)	25 (73.5%)	0.009	0.93

Table 43. Which of the changes that you selected do you think could be attributed to your TRA/R01. Select all that apply.

Item of interest	TRA	R01	Chi-square (df = 1)	p
I have been invited to be a keynote speaker to share my research	7 (70.0%)	11 (84.6%)	0.11	0.74
My research has been featured in the popular press/media	15 (88.2%)	20 (90.9%)	< 0.001	1
I have been asked to give an invited presentation about my research	12 (80.0%)	16 (84.2%)	< 0.001	1
I have been invited to serve as a regular journal reviewer	19 (90.5%)	17 (94.4%)	< 0.001	1
I received an award/honor	21 (91.3%)	28 (96.6%)	0.04	0.84
My research has been listed as a highlight in an academic journal	9 (50.0%)	14 (56.0%)	0.006	0.94

Table 44. Considering the various aspects of the TRA program, please select the degree to which you agree/disagree with the following statements

Item of interest	Group	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
The grant allowed for flexibility in the use of funding	TRA	1 (4.2%)	NA	NA	7 (29.2%)	16 (66.7%)	NA
	R01	2 (5.7%)	3 (8.6%)	6 (17.1%)	17 (48.6%)	7 (20.0%)	NA
The grant allowed me the freedom to pursue nontraditional research	TRA	NA	2 (8.3%)	5 (20.8%)	12 (50.0%)	5 (20.8%)	NA
	R01	2 (5.7%)	9 (25.7%)	4 (11.4%)	12 (34.3%)	7 (20.0%)	1 (2.9%)
The period of the grant was long enough to redirect research as ideas/methods evolved	TRA	NA	NA	1 (4.2%)	7 (29.2%)	16 (66.7%)	NA
	R01	2 (5.7%)	8 (22.9%)	4 (11.4%)	11 (31.4%)	8 (22.9%)	2 (5.7%)

Table 45. Please indicate whether you agree or disagree with the following statements

Item of interest	Group of interest (total number of respondents)	Agree	Disagree	Chisq (df)	p
My research plan changed significantly from what I originally proposed	TRA (n = 24)	10 (41.7%)	14 (58.3%)	0 (df = 1)	1
	R01 (n = 35)	14 (40.0%)	21 (60.0%)		
I am continuing or planning to continue the trajectory of my TRA funded research post-award	TRA (n = 24)	23 (95.8%)	1 (4.2%)	1.22 (df = 1)	0.27
	R01 (n = 35)	29 (82.9%)	6 (17.1%)		

Table 46. You indicated you were continuing or planning to continue the trajectory of your TRA funded research post-award. Please indicate whether this continuation is funded

Item of interest	Group	Is funded post-award	Is not funded post-award
Continuation of grant funded research	TRA	16 (69.6%)	7 (30.4%)
	R01	20 (69.0%)	9 (31.0%)

Table 47. Looking into the future and considering the potential outcomes of your TRA/R01 supported research, please select whether the outcome is likely/unlikely. My research could potentially result in:

Item of interest	Group of interest	Unlikely	Likely	Chisq (df = 1)	p-value	Unsure	N/A
The development of a new methodology	TRA	1 (4.2%)	17 (70.8%)			3 (12.5%)	3 (12.5%)
	R01	8 (22.9%)	23 (65.7%)	0.02	0.90	3 (8.6%)	1 (2.9%)
The development of a new technology	TRA	1 (4.2%)	17 (70.8%)			3 (12.5%)	3 (12.5%)
	R01	5 (14.3%)	22 (62.9%)	0.13	0.72	6 (17.1%)	2 (5.7%)
The development of new therapies, clinical tools, or strategies	TRA	3 (12.5%)	17 (70.8%)			3 (12.5%)	1 (4.2%)
	R01	9 (25.7%)	19 (54.3%)	1.02	0.31	7 (20.0%)	NA
The discovery of new phenomenon or advancement of a theoretical concept	TRA	3 (12.5%)	17 (70.8%)			3 (12.5%)	1 (4.2%)
	R01	16 (45.7%)	13 (37.1%)	5.19	0.02	5 (14.3%)	1 (2.9%)
A change in how research is conducted	TRA	1 (4.2%)	20 (83.3%)			3 (12.5%)	NA
	R01	3 (8.6%)	27 (77.1%)	0.06	0.80	5 (14.3%)	NA

Item of interest	Group of interest	Unlikely	Likely	Chisq (df = 1)	p-value	Unsure	N/A
A new synthesis of disparate ideas	TRA	4 (16.7%)	12 (50.0%)			7 (29.2%)	1 (4.2%)
	R01	11 (32.4%)	11 (32.4%)	1.17	0.28	7 (20.6%)	5 (14.7%)

Appendix E. Senior Scientist Review Survey

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) high-risk, high-reward research program that is part of the Common Fund. STPI is an independent, federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

The survey in which you are participating includes Transformative Research Award (TRA) awardees and two comparison groups. Specifically, the survey queries award-related activities and outputs to determine if these elements are scientifically or technically transformative, or not. This is done through a series of questions that reflect the characteristics: transformative, paradigm shifting, innovative and impactful.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to the NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

Instructions for the Survey

STPI is asking you to review three awardees who are matched to you based on the S&T topic represented in the grant. You are asked to review the top 3 cited publications for each of the awardees. Please have the packet provided for each awardee available for reference.

The estimated completion time is 90 minutes, and your time will be compensated with a $500 honorarium if all three reviews are completed.

You will be able to move backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last submitted page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses.

While completing this survey, you will be asked several questions about each awardee.

Please only consider that awardee as you respond to the corresponding questions. The questions will repeat for each awardee.

Follow-Up Interview

After submission of your reviews, STPI staff may call you for a short (~30 minute) phone interview to clarify your responses, if necessary for the analysis.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact STPI at NIHgrantstudy@ida.org for operational questions or Dr. Ravi Basavapa at ravikumar.basavappa@nih.gov for NIH concerns.

Compensation

To thank you for your time, and as noted above, at the completion of three reviews and a few legal forms, STPI will provide a $500 honorarium.

Your responses are invaluable to the study. Again, thank you for your participation.

Logic: The survey ends if responded, “No, I would not like to participate” to "Please indicate whether or not you would like to participate in this review. "

1) Please indicate whether or not you would like to participate in this review.

() Yes, I would like to participate

() No, I would not like to participate

NOTE: Senior scientist reviewers saw the following questions three times. Once for each award packet they received.

2) Please select which awardee you would like to review first. *

() [invite("custom 1")]

() [invite("custom 2")]

() [invite("custom 3")]

3) Considering [question('option title'), id='73']'s research, please select whether you agree/disagree with the following statements:

	Agree	**Disagree**
One or more of the awardee's research ideas challenged existing science/technology paradigms.	()	()
The awardee's research involved a novel combination of ideas, disciplines, or approaches that significantly changed research practice.	()	()
The awardee's research furthered existing practices or thinking in their field.	()	()
The awardee's research led to a novel invention or a new technology which significantly improved current practices.	()	()

4) Considering [question('option title'), id='73']'s research, please select whether you agree/disagree with the following statements:

The awardee's research…

	Agree	**Disagree**
Led to the development of a new methodology that significantly changed research in their field.	()	()
Required the use of equipment, technique or model that was novel and significantly changed research in their field.	()	()
Aided in the development of new therapies, clinical tools, or strategies that significantly changed research or clinical practice.	()	()
Took the next steps in their established area of investigation.	()	()

5) It is well established that some research has immediate impact, and other research builds gradually. Considering the potential or realized effects of [question('option title'), id='73']'s research, please select whether you agree/disagree with the following statements:

The awardee's research resulted in…

	Agree	**Disagree**
A gradual but building shift in practices or thinking.	()	()
A delayed but significant shift in practices or thinking.	()	()
An immediate and significant shift in practices or thinking.	()	()

6) Looking into the future and considering the potential outcomes of [question('option title'), id='73']'s research, please select whether the outcome is likely/unlikely:

The awardee's research could potentially result in…

	Unlikely	**Likely**	**Unsure**	**N/A**
The discovery of new phenomenon or advancement of a theoretical concept.	()	()	()	()
A new synthesis of disparate ideas.	()	()	()	()
The development of a new methodology.	()	()	()	()
The development of a new technology.	()	()	()	()
The development of new therapies, clinical tools, or strategies.	()	()	()	()
A change in how research is conducted.	()	()	()	()

7) Is there any additional information you would like to share about [question('option title'), id='73']?

_____

_____

_____

_____

8) Did you feel qualified to assess the scientific merit of [question("option title"), id="73"]?

() Yes, I felt qualified to assess the scientific merit of this awardee.

() No, I did not feel qualified to assess the scientific merit of this awardee.

Logic: the survey ends if responded, "Would you like to continue to the next awardee?" to "I would not like to participate beyond what I have contributed"

9) Would you like to continue to the next awardee?

() I would like to continue to the next awardee

() I would not like to participate beyond what I have contributed

NOTE: The previous questions are repeated for each of the awardees the senior scintist reviewer reviewed.

In the box below, please write up to five key words that you associate with the term "transformative."

_____

_____

_____

_____

Thank You!

Appendix F. Senior Scientist Review Survey Data Tables

A. TRA-NDPA Comparison

Table 48. Please select whether you agree/disagree with the following statements.

Item of interest	Group of interest (total number of reviews)	Agree	Disagree	Chisq (df)	p
one or more of the awardees research ideas challenged existing science technology paradigms considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 59)	36 (61.0%)	23 (39.0%)	0.05 (df = 1)	0.82
	NDPA(n = 76)	49 (64.5%)	27 (35.5%)		
the awardees research involved a novel combination of ideas disciplines or approaches that significantly changed research practice considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 59)	41 (69.5%)	18 (30.5%)	0.28 (df = 1)	0.59
	NDPA(n = 77)	49 (63.6%)	28 (36.4%)		
the awardees research furthered existing practices or thinking in their field considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 60)	56 (93.33%)	4 (6.67%)	0 (df = 1)	1.00
	NDPA(n = 78)	73 (93.59%)	5 (6.41%)		
the awardees research led to a novel invention or a new technology which significantly improved current practices considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 59)	34 (57.6%)	25 (42.4%)	0.74 (df = 1)	0.39
	NDPA(n = 76)	37 (48.7%)	39 (51.3%)		

Item of interest	Group of interest (total number of reviews)	Agree	Disagree	Chisq (df)	p
led to the development of a new methodology that significantly changed research in their field considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 59)	34 (57.6%)	25 (42.4%)	0.61 (df = 1)	0.44
	NDPA(n = 75)	37 (49.3%)	38 (50.7%)		
required the use of equipment technique or model that was novel and significantly changed research in their field considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 59)	38 (64%)	21 (36%)	1.1 (df = 1)	0.29
	NDPA(n = 76)	41 (54%)	35 (46%)		
aided in the development of new therapies clinical tools or strategies that significantly changed research or clinical practice considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 59)	21 (35.6%)	38 (64.4%)	0 (df = 1)	0.97
	NDPA(n = 77)	26 (33.8%)	51 (66.2%)		
took the next steps in their established area of investigation considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 60)	54 (90.0%)	6 (10.0%)		0.99
	NDPA(n = 78)	69 (88.5%)	9 (11.5%)		

Table 49. Considering the potential or realized effects of the research you reviewed, please select whether you agree/disagree that the awardee's research resulted in the following statements.

Item of interest	Group of interest (total number of reviews)	Agree	Disagree	Chisq (df)	p
a gradual but building shift in practices or thinking it is well established that some research has immediate impact and other research builds gradually considering the potential or realized effects of the awardee's research please select whether you agree disagree with the following statements the awardees research resulted in	TRA(n = 59)	44 (74.6%)	15 (25.4%)	0.02 (df = 1)	0.89
	NDPA(n = 75)	54 (72.0%)	21 (28.0%)		
a delayed but significant shift in practices or thinking it is well established that some research has immediate impact and other research builds gradually considering the potential or realized effects of the awardee's research please select whether you agree disagree with the following statements the awardees research resulted in	TRA(n = 58)	10 (17.2%)	48 (82.8%)	0.73 (df = 1)	0.39
	NDPA(n = 72)	18 (25.0%)	54 (75.0%)		
an immediate and significant shift in practices or thinking it is well established that some research has immediate impact and other research builds gradually considering the potential or realized effects of the awardee's research please select whether you agree disagree with the following statements the awardees research resulted in	TRA(n = 59)	19 (32.2%)	40 (67.8%)	0.99 (df = 1)	0.32
	NDPA(n = 74)	17 (23.0%)	57 (77.0%)		

Table 50. Please select whether the awardee's research could potentially result in the following outcomes is likely or unlikely.

Item of interest	Group of interest (total number of reviews)	Likely	Unlikely	Chisq (df)	p
the discovery of new phenomenon or advancement of a theoretical concept looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	39 (65.0%)	21 (35.0%)	0.12 (df = 1)	0.73
	NDPA(n = 78)	54 (69.2%)	24 (30.8%)		
a new synthesis of disparate ideas looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	28 (47%)	32 (53%)	2.19 (df = 1)	0.14
	NDPA(n = 79)	48 (61%)	31 (39%)		
the development of a new methodology looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	36 (60.0%)	24 (40.0%)	0.1 (df = 1)	0.75
	NDPA(n = 78)	50 (64.1%)	28 (35.9%)		
the development of a new technology looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	32 (53.3%)	28 (46.7%)	0.13 (df = 1)	0.71
	NDPA(n = 78)	38 (48.7%)	40 (51.3%)		
the development of new therapies clinical tools or strategies looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	35 (58%)	25 (42%)	5.28 (df = 1)	0.02
	NDPA(n = 78)	29 (37%)	49 (63%)		
a change in how research is conducted looking into the future and considering the potential outcomes of the awardee's research please	TRA(n = 60)	30 (50.0%)	30 (50.0%)	0.77 (df = 1)	0.38

Item of interest	Group of interest (total number of reviews)	Likely	Unlikely	Chisq (df)	p
select whether the outcome is likely unlikely the awardees research could potentially result in	NDPA(n = 78)	32 (41.0%)	46 (59.0%)		

B. TRA-R01 Comparison

Table 51. Please select whether you agree/disagree with the following statements.

Item of interest	Group of interest (total number of reviews)	Agree	Disagree	Chisq (df)	p
one or more of the awardees research ideas challenged existing science technology paradigms considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 59)	36 (61%)	23 (39%)	1.98 (df = 1)	0.16
	R01(n = 82)	39 (48%)	43 (52%)		
the awardees research involved a novel combination of ideas disciplines or approaches that significantly changed research practice considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 59)	41 (69%)	18 (31%)	8.06 (df = 1)	0.00
	R01(n = 82)	36 (44%)	46 (56%)		
the awardees research furthered existing practices or thinking in their field considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 60)	56 (93.3%)	4 (6.7%)	0.01 (df = 1)	0.94
	R01(n = 83)	76 (91.6%)	7 (8.4%)		

Item of interest	Group of interest (total number of reviews)	Agree	Disagree	Chisq (df)	p
the awardees research led to a novel invention or a new technology which significantly improved current practices considering the awardee's research please select whether you agree disagree with the following statements	TRA(n = 59)	34 (58%)	25 (42%)	8.4 (df = 1)	0.00
	R01(n = 82)	26 (32%)	56 (68%)		
led to the development of a new methodology that significantly changed research in their field considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 59)	34 (58%)	25 (42%)	6.01 (df = 1)	0.01
	R01(n = 82)	29 (35%)	53 (65%)		
required the use of equipment technique or model that was novel and significantly changed research in their field considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 59)	38 (64%)	21 (36%)	7.08 (df = 1)	
	R01(n = 82)	33 (40%)	49 (60%)		
aided in the development of new therapies clinical tools or strategies that significantly changed research or clinical practice considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 59)	21 (36%)	38 (64%)	2.34 (df = 1)	0.13
	R01(n = 82)	41 (50%)	41 (50%)		
took the next steps in their established area of investigation considering the awardee's research please select whether you agree disagree with the following statements the awardees research	TRA(n = 60)	54 (90.0%)	6 (10.0%)	1.22 (df = 1)	0.27
	R01(n = 83)	68 (81.9%)	15 (18.1%)		

Table 52. Considering the potential or realized effects of the research you reviewed, please select whether you agree/disagree that the awardee's research resulted in the following statements.

Item of interest	Group of interest (total number of reviews)	Agree	Disagree	Chisq (df)	p
a gradual but building shift in practices or thinking it is well established that some research has immediate impact and other research builds gradually considering the potential or realized effects of the awardee's research please select whether you agree disagree with the following statements the awardees research resulted in	TRA(n = 59)	44 (74.58%)	15 (25.42%)	0 (df = 1)	1.00
	R01(n = 83)	62 (74.70%)	21 (25.30%)		
a delayed but significant shift in practices or thinking it is well established that some research has immediate impact and other research builds gradually considering the potential or realized effects of the awardee's research please select whether you agree disagree with the following statements the awardees research resulted in	TRA(n = 58)	10 (17.2%)	48 (82.8%)	0.01 (df = 1)	0.90
	R01(n = 82)	16 (19.5%)	66 (80.5%)		
an immediate and significant shift in practices or thinking it is well established that some research has immediate impact and other research builds gradually considering the potential or realized effects of the awardee's research please select whether you agree disagree with the following statements the awardees research resulted in	TRA(n = 59)	19 (32%)	40 (68%)	5.02 (df = 1)	0.03
	R01(n = 81)	12 (15%)	69 (85%)		

Table 53. Please select whether the awardee's research could potentially result in the following outcomes is likely or unlikely.

Item of interest	Group of interest (total number of reviews)	Likely	Unlikely	Chisq (df)	p
the discovery of new phenomenon or advancement of a theoretical concept looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	39 (65%)	21 (35%)	14.6 (df = 1)	0.00
	R01(n = 83)	26 (31%)	57 (69%)		
a new synthesis of disparate ideas looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	28 (47%)	32 (53%)	1.93 (df = 1)	0.16
	R01(n = 83)	28 (34%)	55 (66%)		
the development of a new methodology looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	36 (60%)	24 (40%)	3.74 (df = 1)	0.05
	R01(n = 83)	35 (42%)	48 (58%)		
the development of a new technology looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	32 (53%)	28 (47%)	6.89 (df = 1)	0.01
	R01(n = 83)	25 (30%)	58 (70%)		
the development of new therapies clinical tools or strategies looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	35 (58.33%)	25 (41.67%)	0 (df = 1)	1.00
	R01(n = 83)	49 (59.04%)	34 (40.96%)		
a change in how research is conducted looking into the future and considering the potential outcomes of the awardee's research please select whether the outcome is likely unlikely the awardees research could potentially result in	TRA(n = 60)	30 (50%)	30 (50%)	1.42 (df = 1)	0.23
	R01(n = 83)	32 (39%)	51 (61%)		

Appendix G. Senior Scientist Reviewers' Descriptors for Transformative Research

As a reminder, reviewers were asked to provide up to five key words that they associate with the term "transformative." The full list of words provided by the reviewers are provided below in Table 54.

Table 54. Words Associated with "Transformative" Research by Senior Scientist Reviewers

Word Describing Transformative Research	Number of Occurrences	Word Describing Transformative Research	Number of Occurrences	Word Describing Transformative Research	Number of Occurrences
paradigm	29	field	5	grind	3
impact	27	influential	5	integrative	3
change	23	unexpected	5	break	3
shift	23	breakthrough	4	direction	3
innovative	19	concept	4	discovery	3
revolutionary	10	enable	4	excite	3
technology	7	innovation	4	game	3
creative	6	insight	4	practice	3
groundbreaking	6	significant	4	radical	3
challenge	5	unique	4	treatment	3
disruptive	5	approach	3	idea	2

Word Describing Transformative Research	Number of Occurrences	Word Describing Transformative Research	Number of Occurrences	Word Describing Transformative Research	Number of Occurrences
bold	2	insightful	2	_change	1
changer	2	inspirational	2	_novel	1
highly	2	leap	2	accuracy	1
_new	2	pioneer	2	adopt	1
advance	2	prize	2	alter	1
application	2	reframe	2	efficient	1
biological	2	fundamental	2	blow	1
surprise	2	set	2	crucial	1
technique	2	breadth	1	definitive	1
therapeutic	2	broad	1	derivative	1
translational	2	build	1	diagnostic	1
clinical	2	care	1	disciplinary	1
conceptually	2	central	1	disrupt	1
cut	2	cite	1	dissemination	1
dogma	2	conceptual	1	distinct	1
edge	2	conventional	1	earthshaking	1
exist	2	foundational	1	influencer	1
emergent	1	fresh	1	inspire	1
empower	1	human	1	interdisciplinary	1
endure	1	imaginative	1	invisible	1
enhance	1	implement	1	last	1
evalutating	1	implementation	1	lead	1
exceptionally	1	implication	1	life	1

Word Describing Transformative Research	Number of Occurrences	Word Describing Transformative Research	Number of Occurrences	Word Describing Transformative Research	Number of Occurrences
eye	1	improve	1	lucid	1
forward	1	inaccessible	1	major	1
measurement	1	outcome	1	quo	1
mechanistic	1	patient	1	reimagining	1
mind	1	perspective	1	research	1
model	1	persuasive	1	result	1
move	1	practical	1	revolution	1
multi	1	precedent	1	risk	1
nobel	1	previously	1	robust	1
noteworthy	1	provoke	1	science	1
original	1	question	1	scientific	1
seminal	1	target	1	ultimate	1
sound	1	technical	1	unconventional	1
spread	1	technological	1	understand	1
standard	1	term	1	uniquely	1
status	1	theory	1	unorthodox	1
strategic	1	think	1	unprecedented	1
success	1	tool	1	validate	1
survival	1	translation	1	visible	1
synergy	1	trend	1	wide	1
widely	1	wolf	1		