



# Evaluation of the Implementation Research Institute: Trainees' Publications and Grant Productivity

Ana A. Baumann<sup>1</sup> · Bobbi J. Carothers<sup>2</sup> · John Landsverk<sup>3</sup> · Emily Kryzer<sup>1</sup> · Gregory A. Aarons<sup>4</sup> · Ross C. Brownson<sup>5,6</sup> · Charles Glisson<sup>7</sup> · Brian Mittman<sup>8</sup> · Enola K. Proctor<sup>1</sup>

Published online: 30 October 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

With growth in the field of dissemination and implementation (D&I) research, there has been growth in capacity building, with many training opportunities. As such, it is important to continue to evaluate D&I research training programs. This paper reports the results of an evaluation of the Implementation Research Institute (IRI), a R25 funded by the National Institute of Mental Health with additional funding by the Department of Veterans Affairs (VA). The fourth cohort also had a supplement from the National Institute on Drug Abuse. Using bibliometrics data, we report on a quasi-experimental retrospective cohort study assessing whether the rates of scholarly productivity in D&I science of IRI fellows (those who applied and were accepted to the training) were greater than those who applied but were not accepted to IRI. Our findings show that Selected Applicants' odds of publishing in implementation science were higher for earlier alumni, starting at 12% 1 year out and increasing to 94% for those who were 4 years out from starting training. Chances for Non-Selected Applicants remained relatively stable, starting at 47% at 1 year and going to 33% at 4 years since their application, a pattern that was stable even after controlling for demographic characteristics. These results support the hypothesis that IRI is increasing the D&I research productivity of those selected to the program, and that our fellows are advancing the field of D&I compared to those investigators not selected to our institute. Our finding also indicates the importance of a 2-year training.

**Keywords** Training in D&I · Evaluation · Bibliometrics

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10488-019-00977-4>) contains supplementary material, which is available to authorized users.

---

✉ Enola K. Proctor  
ekp@wustl.edu

<sup>1</sup> Brown School of Social Work, Washington University in St. Louis, 1 Brookings Drive, Campus Box 1196, St. Louis, MO 63130, USA

<sup>2</sup> Center for Public Health Systems Science, Washington University in St. Louis, 1 Brookings Drive, Campus Box 1196, St. Louis, MO 63130, USA

<sup>3</sup> Oregon Social Learning Center, 10 Shelton Mcmurphy Blvd, Eugene, OR 97401, USA

<sup>4</sup> Department of Psychiatry, Child and Adolescent Services Research Center (CASRC), UC San Diego, 9500 Gilmar Drive 0812, La Jolla, CA 92093-0812, USA

## Introduction

With robust growth in the field of dissemination and implementation research, there also has been growth in capacity building, with increasing training opportunities for dissemination and implementation (D&I) research that

<sup>5</sup> Prevention Research Center in St. Louis, Brown School at Washington University in St. Louis, 1 Brookings Drive, Campus Box 1196, St. Louis, MO 63130, USA

<sup>6</sup> Department of Surgery (Division of Public Health Sciences) and Alvin J. Siteman Cancer Center, Washington University School of Medicine, Washington University in St. Louis, St. Louis, USA

<sup>7</sup> Center for Behavioral Health Research, College of Social Work, The University of Tennessee, Knoxville, 1618 W. Cumberland Ave, 201 Henson Hall, Knoxville, TN 37996-332, USA

<sup>8</sup> Kaiser Permanente Southern California Department of Research and Evaluation, 100 S. Los Robles Ave, Pasadena, CA 91101-2453, USA

vary in length, depth, home institution, and target audience (Chambers et al. 2016; Proctor and Chambers 2016). Among the many trainings are the NIH Training Institute for Dissemination and Implementation Research in Health (TIDIRH) (Meissner et al. 2013), the National Cancer Institute Mentored Training for Dissemination and Implementation Research in Cancer (MT-DIRC) (Padek et al. 2015a, b, 2018), the National Cancer Institute's Training Institute for Dissemination and Implementation Research in Cancer (TIDIRC), several recent NHLBI K12 programs (National Heart, Lung, and Blood Institute 2017) and the National Institute of Mental Health/National Institute on Drug Abuse/VA funded Implementation Research Institute (IRI) (Proctor et al. 2013; Luke et al. 2016). In addition to these U.S. based trainings are international programs (e.g., Carlford et al. 2017; Straus et al. 2011; Ulrich et al. 2017). While the number of opportunities for D&I training have increased, leaders of these programs have recognized the challenge to ensure that training in D&I is both foundational and responsive to the fast evolving D&I field (Chambers et al. 2016). A similar concern has been voiced by trainees in the D&I field (Stamatakis et al. 2013). This is a characteristic problem for providing training in a fast developing field of research.

Given the rapid and dynamic evolution of the field, it is important to continue to evaluate D&I research training programs. The 2008 NIMH Investing in the Future Report recommends that training programs implement rigorous monitoring and evaluation systems in a timely manner to gauge the impact of each new program implemented on its target population. This report suggested that addressing training quality through objective methods is important but all too rare, and often missing in the field (National Institute of Mental Health 2008). Some training programs have published data showing general high rates of acceptance of the training model in terms of organization, quality of lectures, format, and quality of mentoring (Gonzales et al. 2012; Meissner et al. 2013; Proctor et al. 2013). The TIDIRH and the IRI programs have also shown that their trainees have high scholarly productivity (Luke et al. 2016; Meissner et al. 2013). A recent study reports on a number of specific competencies associated with the MT-DIRC training (Padek et al. 2018). For example, competency domains include basic understanding of D&I concepts, understanding of theories, D&I research designs and analytic approaches, and practice-based issues. While these data are useful and represent participant perspectives, they do not allow us to specify the extent to which the aforementioned institutes are objectively facilitating scholarly productivity in the D&I field compared to other non-formal D&I training. This paper reports the results of a comparative evaluation design for the National Institute Mental Health (NIMH) IRI (selected vs. non-selected applicants) with assessment of D&I research

productivity by means of the increasingly employed bibliometric methods.

## Description of the Implementation Research Institute (IRI) Training

The IRI is based on a learning collaborative approach to training implementation researchers. The National Institute of Mental Health (NIMH) has consistently funded eight fellows through the competitive R25 mechanism, and the Department of Veterans Affairs (VA) has funded two fellows during all the years of IRI. Two fellows from the last cohort included in this analysis was also funded by the National Institute on Drug Abuse (NIDA). The IRI training program's overall purpose is to assist early career scholars in launching a career in D&I science via experiential learning, didactic training, faculty mentoring and supporting for developing grant writing skills. Institute goals are to (1) strengthen human capital in the field of D&I science to address the challenges of mental health care through the training of a new generation of D&I researchers, and (2) advance intellectual capital for the still-developing field through stimulating the production of scholarly products such as papers, books, and curriculum models (Proctor et al. 2013).

## Multiple Training Components

Since its inception, the IRI's basic components are an initial 1-week, on-site training followed by a second consecutive annual on-site training at Washington University in St. Louis, paired with on-going mentoring with a core faculty member over a minimum 2 year period. In addition, participating IRI fellows are required to complete a 1–2 day learning site visit to a federally funded D&I science research project, hosted by an experienced principal investigator who has granted access and agrees to host the site visit. Travel funding is provided to each fellow for this component (see Proctor et al. 2013 for more detailed discussion of the IRI multiple components).

## IRI Selection Criteria and Process

The IRI's first cohort of fellows was selected in 2010. The IRI seeks to select a pool of early career investigators—diverse in discipline and area of expertise—who seek to advance their training for D&I science in mental health services. Applicants are asked to describe their career goals in the form of a research agenda and to propose an initial concept paper for an implementation research project. Applications are rated by the core faculty and a program representative of the funding agencies (NIMH and VA for the initial

three cohorts and NIDA for subsequent cohorts) according to the following criteria:

- (1) expressed interest in implementation research (IR), as reflected in a letter of application and described in a concept paper as part of the application;
- (2) prior or concurrent experience relevant to IR such as intervention development and/or testing, mental health services research, or study of organizational factors in mental health service delivery;
- (3) prior experience writing an NIH F31, R03, R34/R21, K award, or R01 proposal, or funding applications to other federal agencies or a foundation. Receipt of grant funding is not required. Recruiting fellows with some grant experience reduced our need to train in the fundamentals of grant writing;
- (4) based in a home institution with an on-site mentor supportive of the fellow's progress in grant writing and scholarly publication; local mentors did not need to be an expert in IR, but need to have a successful record of NIH or VA funding to support trainees in grant-writing fundamentals; and
- (5) access to a clinic/service settings willing to serve as pilot site for the fellow's implementation research.

Core faculty members independently rate each application on a scale from 1 to 9, based on NIH guidelines, where a score of 1 indicates an exceptionally strong application, and a score of 9 indicates an application with serious and substantive weaknesses. Raters are encouraged to note their rationale for ratings, and where appropriate to declare any conflicts of interest. Raters are excluded from rating applications where this is a real or perceived conflict of interest. Applicant overall scores are compiled in rank order and shared with the full team of raters. Rating discussion and acceptance decisions are made in a group phone or in-person meeting.

For this paper, each of the first three cohorts were comprised of a total of ten selected applicants, eight applicants funded by NIMH and 2 applicants funded by VA. The 4th cohort had a total of 12 selected applicants: 8 funded by NIMH, 2 funded by VA and 2 funded by NIDA.

## Evaluation Methods

### Design

Given the goal of the IRI to help early career scholars launch careers in mental health D&I science, we focused our evaluation on measures of scholarly productivity specific to D&I science, defined as a rate or proportion relative to their overall scientific productivity. Scholarly products are defined as

publications reporting on D&I studies and successful applications for funding of further D&I studies.

Training programs typically use pre/post designs based on outcomes measured for the applicants who were selected to participate in the training program. That design is weakened by its failure to include a comparison to outcomes for subjects who did not receive the training intervention. Furthermore, research training programs do not use randomization to select applicants, given their required reliance in specific selection criteria.

Our evaluation is atypical in its use of a comparison condition. Namely, we report on an evaluation assessing whether the rates of scholarly productivity in D&I science of IRI fellows (those who applied and were accepted and completed the fellowship) were greater than those who applied but were not accepted to IRI, controlling for differences in applicants that were standardized in the application materials between the two groups. We hypothesized that IRI fellows will be more likely than non-accepted applicants to receive grants and publish in the field of D&I science after their participation in the IRI institute.

This design is strengthened by its two comparisons: pre-post scholarly productivity, and between the selected and non-selected applicants. We note that using only applicants to the IRI training program provides an additional design strength in that the applicant pool only includes individuals with sufficient interest in D&I science such that they submitted a complete application. While control groups for science training institutes are not feasible because of the need to select on competitive criteria, one can conduct comparison analysis with those who applied but were not selected to the training. Technically, this is a quasi-experimental design retrospective cohort study with comparison of the conditions of those selected and not selected for training on the primary outcome of scholarly productivity related to D&I science. As indicated in the methods sections, this design was feasible because of the rapidly expanding availability of bibliometric methods that permitted examination of pre/post markers of scholarly productivity through public sources not constrained by problems in follow-up of subjects in the two conditions.

### Participants

Study participants were all applicants for the 2010, 2011, 2012, and 2013 cohorts. We had a total of 124 applications, 43 were applicants invited to IRI ("Selected Applicants") and 81 were applicants not invited to the Institute ("Non-Selected Applicants"). Through bibliometric methods we identified publications from applicants from all four cohorts. Because this study was conducted in May of 2014, and the fourth cohort had not finished their second Institute, and the window for funding exceed this timeframe, we included

grant data from all but the fourth cohort in the analysis. Our rationale is that grant awards take time and a 1 year span was not long enough to give time for the fourth cohort to be awarded D&I science grants. The total number of grants included in our analysis therefore is 82, where 31 are from Selected Applicants and 51 from Non-Selected Applicants.

All applicants accepted to IRI attended the first summer institute and completed their Learning Site Visit (see above); however a few varied in the timing and completion of the second summer institute. Specifically, one Selected Applicant did not attend the second summer institute due to family demands; therefore we consider that this person did not complete our training and was dropped from the comparisons of the two conditions. Three Selected Applicants from different cohorts were granted maternity leave, and therefore skipped their second year of training, returning during the following year to complete their second summer institute. All three were included in the analyses.

### Outcome Measurement

Publications and grant funding are key metrics in evaluating research training because they likely reflect the impact of the formal training on trainees' subsequent work in the field (Carpenter et al. 2014). Grants for which fellows were the PIs comprises the primary outcome used by NIMH to evaluate the quality and effectiveness of research training programs (National Institute of Mental Health 2008). We did not consider grants where fellows were co-Is or part of a team. Moreover, number of publications and grants awarded have been linked to increased applications for funding and to academic advancement including faculty promotion (Akl et al. 2012; Cramer et al. 2000). However, the measurement challenge for this specific training program was to not only count number of publications and number of grants funded but to isolate counts of those scholarly products that were objectively related to the field of D&I science.

### Bibliometrics

To overcome potential biases of using applicant produced curriculum vitae and self-selection of D&I content in scholarly products, we captured applicant publications through bibliometrics, a methodology increasingly used to measure research productivity in various scientific fields (O'Leary and O'Sullivan 2012). Bibliometrics was first coined by Alan Pritchard in 1962, and its use by several governments and other training programs indicates the emerging importance of this methodology as an evaluation tool (Durieux and Gevenois 2010; Eloy et al. 2012; Froghi et al. 2012; Giles 2004; Huang et al. 2015; O'Leary and O'Sullivan 2012). The advantage of using bibliometrics is that it allows a systematic counting of publications using search engines in

public databases thus providing an evaluation effort with an objective sample of peer-reviewed journal articles from selected and non-selected applicants to training institutes. To our knowledge, this is the first evaluation of D&I training using bibliometric data analysis. Additionally, the objectivity of the research productivity count is not dependent on the subject self-reporting selectively on their grants and publications. Analysis of publication activity was obtained via bibliometric methods performed by trained Washington University library scientists who were not formally affiliated with the IRI.

### Publicly Available Grant Data Bases

Grants abstracts were obtained via a search of all applicants' names on the NIH RePORTer and the VA Query websites. Publications of each Selected or Non-Selected Applicant were determined by searching authors' names using Elsevier Scopus (<http://www.scopus.com>), a comprehensive database with over 69 million records that provides access to peer-reviewed literature from 21,950 journals, including 100% coverage of MEDLINE titles. Since grants submitted but not funded are not available and privacy rules do not allow them to come into the public domain, we only considered grant applications funded which are in the public domain.

### Coding of Research Products for D&I Content

Names on all publications and grants were blanked out prior to coding. In addition, coders were blind to condition—selected or not selected for the IRI training. The second author who had no experience or involvement in the original application process for the four cohorts coded all products while the first author coded 20% of blinded publications and grants randomly selected for the purpose of reliability checking.

First, products qualified as having D&I science content if they had the word “implementation” in the title/abstract and if they reflected an implementation-centered hypothesis, design, or framework, focused on assessing the implementation climate of an organization, or described the implementation processes for a particular intervention. Initial independent coding reached 89% of agreement. Discrepancies were discussed between the coders and 100% were resolved through these discussions. Key words that led to products being coded as implementation research were, in order of frequency: *implementation* (n = 270), then within all grants and publications with *implementation*, *provider* (n = 59), *organization* (47), *strategies* (n = 34), *collaboration* (n = 18), *sustainability* (n = 13), and *implementation outcome* (n = 8). We then totaled the numbers of implementation science grants and publications for each fellow and applicant for analysis.

## Analysis

We compared grants and publications for Selected Applicants versus Non-Selected Applicants who applied but were not selected to the IRI training. Due to the highly non-normal distributions of the proportions of D&I grants and publications, it was necessary to dichotomize whether or not Selected Applicants and Non-Selected Applicants had received a grant or published in the implementation science field and then perform logistic regression for significance testing. For both grants and publications, three models were examined: (1) main effects (whether the person was a Selected Applicant or a Non-Selected Applicant and the number of years since IRI attendance), (2) controlling for having a D&I grant or publication at the time of application, (3) controlling for other demographic characteristics: gender, background in child mental health (compared to “not child mental health”), academic discipline, rank, number of recommendations letters in the applications, and time since receiving their undergraduate degree (as a proxy for general experience).

One additional analysis was considered, namely to conduct comparison analyses by cohort to test whether there was an increase in the training effect associated with greater experience in the later cohorts. However, these analyses were dropped from consideration due to the small size of the cells at the cohort level.

## Results

Table 1 shows Selected Applicants and Non-Selected Applicants demographics at the time of application. The majority of Total Applicants were female, did not focus on child mental health, and had a Clinical discipline at the time of application. Assistant Professors were the most common rank. Selected Applicants were more likely to have a focus on child mental health than Non-Selected Applicants, but the groups did not differ in other demographics. Overall, 34.7% (43/124) were selected from the total applicant pool. Selection rates varied across the four cohorts, from higher rates of 40% (10/25, cohort 2011) and 55.6% (10/18, cohort 2012) to 28.2% (11/39, cohort 2010) and 28.6% (12/42, cohort 2013). This pattern reflects higher selection rates in the years of low application numbers given the fixed number of slots allotted in the grant and selection committee ability to fill all slots with highly qualified individuals.

Table 2 shows that Selected Applicants were significantly more likely to have pre and post implementation publications as well as implementation related grants as compared with Non-Selected Applicants, likely due to the selection decisions for pre research productivity related to implementation research. As shown in Table 3, Selected Applicants had significantly higher odds of being awarded a D&I grant after IRI than Non-Selected Applicants, even when controlling for other variables available at application (*OR* 6.16,

**Table 1** Fellows and applicants' demographic characteristics at the moment of application

	Selected		Non-selected		Total		$\chi^2$	
	N	%	N	%	N	%	Value	p
Gender								ns
Male	10	23.3	24	29.6	34	27.4		
Female	33	76.7	57	70.4	90	72.6		
Child mental health							4.48	< .05
No	22	51.2	57	70.4	79	63.7		
Yes	21	48.8	24	29.6	45	36.3		
Year applied to IRI								ns
2010	11	25.6	28	34.6	39	31.5		
2011	10	23.3	15	18.5	25	20.2		
2012	10	23.3	8	9.9	18	14.5		
2013	12	27.9	30	37.0	42	33.9		
Discipline								ns
Allied health	12	27.9	17	21.0	29	23.4		
Clinical	23	53.5	47	58.0	70	56.5		
Social science	8	18.6	17	21.0	25	20.2		
Rank								ns
Instructor/lecturer/adjunct/post-doc/fellow/research scientist	13	30.2	32	39.5	45	36.3		
Assistant professor	20	46.5	35	43.2	55	44.4		
Associate/full professor	10	23.3	14	17.3	24	19.4		

**Table 2** Implementation publications and grants at application and post IRI (in 2014) for fellows and applicants

	Selected		Non-selected		Total		$\chi^2$	
	N	%	N	%	N	%	value	p
Had implementation publications at application							29.37	< .001
No	14	32.6	66	81.5	80	64.5		
Yes	29	67.4	15	18.5	44	35.5		
Had implementation publications from application to 2014							9.07	< .01
No	16	37.2	53	65.4	69	55.6		
Yes	27	62.8	28	34.6	55	44.4		
Had implementation grant at application							6.60	< .05
No	20	64.5	45	88.2	65	79.3		
Yes	11	35.5	6	11.8	17	20.7		
Had implementation grant from application to 2014							10.50	< .01
No	15	48.4	42	82.4	57	69.5		
Yes	16	51.6	9	17.6	25	30.5		

Grants from the 2013 cohort are not included in the analysis (see “Evaluation Methods” sections for details)

$SE = 0.62$ ; Model 3). Selected Applicants were therefore approximately six times more likely to be awarded a D&I grant than Non-Selected Applicants.

As shown in Table 4, IRI fellowship was associated with higher odds of having an implementation publication after IRI ( $OR\ 3.32$ ,  $SE = 0.40$ ; Model 1). This main effect was qualified by an IRI  $\times$  Time interaction ( $OR\ 3.99$ ,  $SE = 0.45$ ; Model 2), and this interaction remained significant ( $OR\ 5.98$ ,  $SE = 0.48$ ) after controlling for having an implementation publication at the time of application (Model 3). Figure 1 shows the findings from Model 3 and shows that Selected Applicants’ odds of publishing in implementation science were higher for earlier alumni, starting at 12% 1 year out and increasing to 94% for those who were 4 years out from starting training. Chances for Non-Selected Applicants remained relatively stable, starting at 47% at 1 year and going to 33% at 4 years since their application. This pattern was stable when adding demographic characteristics, and none of the demographic characteristics were related to implementation science publication (Model 4).

## Discussion

Our analyses indicate that training by IRI significantly increased the odds of D&I science research productivity. Selected Applicants were more likely than their Non-Selected counterparts to receive D&I science grants and contribute scholarly works to the D&I field following training. Specific to scholarly publication, the impact of training was greater the more time Selected Applicants had from the time of application, while this increase over time was not observed for Non-Selected Applicants. This supports the hypothesis laid out in the design section that the probability

of scholarly productivity in D&I science of IRI fellows (those who applied and were accepted and completed the fellowship) would be significantly greater than those who applied but were not accepted to IRI. This support for the hypothesis is strengthened measurably when the differences between the two conditions continued to be observed even after the differences on research productivity related to D&I science between the two conditions at the time of application were controlled in the analytic models.

Also observed is clear evidence that the Selected Applicants had a higher probability of research productivity related to D&I science at the time of application. This supports the tenet that competitive science training programs contain a built-in selection factor related to outcomes targeted for the training, which in this case was research productivity specific to D&I science. Of course, there is the possibility that applicants to the training program who are selected are simply “cherry picked” at the time of selection, explaining the downstream outcome differences between the two conditions, rather than exposure to the fellowship training making the difference. However, this alternative explanation is not supported in this evaluation because even when controlling for pre-application research productivity related to D&I science in the two conditions, the pre-post outcome differences were significantly greater in the training condition and the differences between the two conditions also appear to become greater over the 4 year observation period after application.

Several features of methodology strengthen this evaluation. First, using only applicants to the IRI training program for the evaluation pool strengthens the design, as trainee outcomes are compared to individuals with strong enough interest in and trajectory toward implementation science that they completed an IRI application. The application process

**Table 3** Implementation grant awarded between application and 2014

	Model 1		Model 2		Model 3					
	95% CI		95% CI		95% CI					
	B (S.E.)	OR	Upper	Lower	B (S.E.)	OR	Upper	Lower	OR	Upper
Constant	-0.47 (1.12)	0.63	0.63	-0.42 (1.13)	0.66	0.66	2.42 (2.17)	11.29		
Time	-0.32 (0.32)	0.39	0.73	1.37	1.37	0.72	1.367	0.67	1.38	1.38
IRI										
Non-selected (ref)										
Selected	<i>1.51 (0.52)</i>	<i>1.62</i>	<i>4.52</i>	<i>12.61</i>	<i>4.87</i>	<i>14.246</i>	<i>1.82 (0.62)</i>	<i>6.16</i>	<i>20.87</i>	<i>20.87</i>
Implementation grant at time of application										
No (ref)										
Yes					0.74	2.602	-0.68 (0.73)	0.12	0.51	2.10
Gender										
Male (ref)										
Female							0.35 (0.7)	0.36	1.42	5.61
Child mental health										
No (ref)										
Yes							-0.15 (0.61)	0.26	0.86	2.84
Discipline										
Allied health (ref)										
Clinical							0.61 (0.76)	0.41	1.83	8.15
Socialsci/stat							-0.65 (1.05)	0.07	0.52	4.06
Rank							-0.66 (0.49)	0.20	0.52	1.34
Recommendation count							-0.31 (0.37)	0.35	0.73	1.53
Time since undergraduate degree							-0.07 (0.06)	0.84	0.93	1.04
Fit										
Block change			$\chi^2(1) = 0.23$ , ns						$\chi^2(7) = 8.90$ , ns	
Model		$\chi^2(2) = 11.36$ , $p < .01$	$\chi^2(3) = 11.59$ , $p < .01$						$\chi^2(10) = 20.49$ , $p < .05$	
Nagelkerke R <sup>2</sup>		.18	.19						.31	

Terms significant at  $p < .05$  are italicized

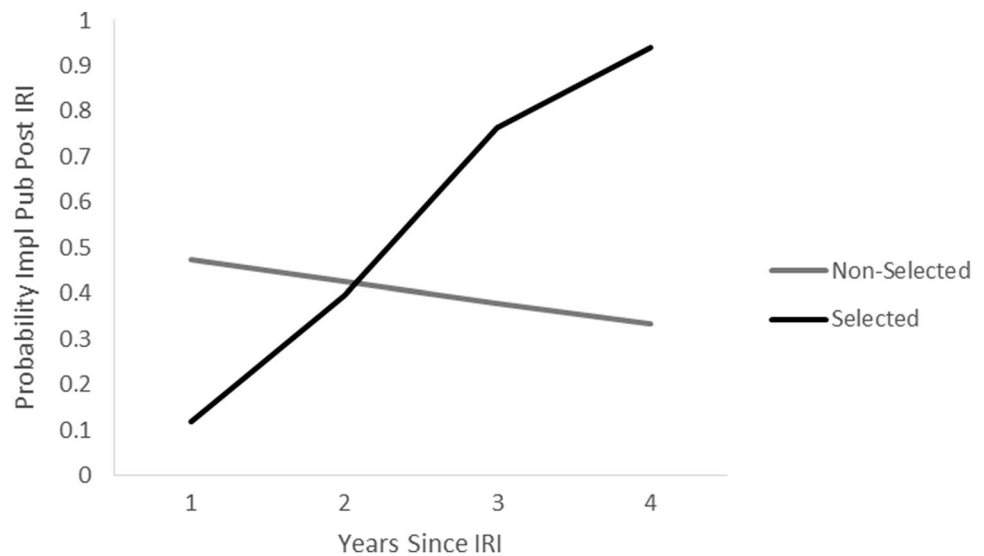
**Table 4** Implementation publication between application and 2014

	Model 1		Model 2		Model 3		Model 4	
	95% CI		95% CI		95% CI		95% CI	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
Constant	-1.31 (0.47)	0.27	-0.46 (0.50)	0.63	-0.49 (0.52)	0.61	1.07 (1.26)	1.26
Time	0.26 (0.15)	0.96	-0.07 (0.18)	0.65	-0.20 (0.20)	0.56	-0.20 (0.21)	0.54
IRI								
Non-selected (ref)								
Selected	<i>1.20 (0.40)</i>	<i>1.52</i>	<i>-1.97 (1.05)</i>	<i>0.02</i>	<i>-3.71 (1.25)</i>	<i>0.00</i>	<i>-4.41 (1.39)</i>	<i>0.01</i>
IRI × time			<i>1.39 (0.45)</i>	<i>1.67</i>	<i>1.79 (0.48)</i>	<i>2.34</i>	<i>2.06 (0.55)</i>	<i>2.68</i>
Implementation publication at time of application								
No (ref)								
Yes					1.65 (0.54)	1.81	1.80 (0.59)	1.90
Gender								
Male (ref)								
Female							0.08 (0.50)	0.41
Child mental health								
No (ref)								
Yes							0.76 (0.47)	0.85
Discipline								
Allied health (ref)								
Clinical							-0.83 (0.57)	0.14
Socialsci/stat							-0.03 (0.71)	0.24
Rank							-0.36 (0.37)	0.34
Recommendation count							-0.48 (0.29)	0.35
Time since undergraduate degree							0.02 (0.04)	0.95
Fit								
Block change					$\chi^2(1) = 12.63, p < .001$		$\chi^2(7) = 9.90, ns$	
Model					$\chi^2(3) = 24.64, p < .001$		$\chi^2(4) = 34.86, p < .001$	
Nagelkerke R <sup>2</sup>		.12		.24				.33
								.41

Terms significant at  $p < .05$  are italicized



**Fig. 1** Selected Applicants' chances of publishing an implementation paper by 2014 increased with time after IRI, where chances did not increase for Non-Selected Applicants. Having an implementation publication prior to training/application was set to .355 (% of all who had a prior implementation publication)



is not trivial in that it involves developing and honing a competitive research proposal concept, garnering institutional support, gathering supporting materials, identifying local mentors, and obtaining letters of support and recommendation. While randomizing for science training institutes is not feasible, one can conduct comparison analysis with those who applied but were not selected to the training. This analysis used a stronger comparison pool than, for example, a demographically adjusted pool of mental health researchers. The evaluation therefore contained a retrospective cohort study design with comparison of the conditions of selected and not selected for training on the primary outcome of scholarly productivity related to implementation science.

A second strength is the approach to measuring outcomes. Rather than relying on self-report, we measured the primary outcome indicator for research productivity via bibliometric methods performed by trained Washington University library scientists who were blinded to condition. Grants were measured from publicly available databases, also blinded to condition for coding. These measurement strategies overcome potential biases inherent in applicant-produced curriculum vitae and applicant self-identified products. Moreover, the reliance on publically available databases enabled us to identify scientific products from those who applied but were not selected for the IRI. However, we acknowledge that the search strategies employed did not capture all research activities, for example those in non-peer-reviewed journals, conference proceedings, lectures given, or gray literature; nor the number of grants submitted or under review. Success, nevertheless, in peer-reviewed publications and funded grants are much more accepted metrics for academic success.

Overall, the results support a strong answer of “yes” to the question of whether truly new D&I investigators are being produced during the IRI training. The design and

measurement features employed in the evaluation provide greater credibility for the conclusion that participation in the IRI training program was the mechanism that produced the increase in research productivity related to D&I science rather than attributes and accomplishments observed in the applicants to the program. The use of multi-variate modeling allowed for control of attribute and accomplishment differences between the two conditions at the time of application.

## Limitations to the Study

We acknowledge several limitations to this study. First, our analyses cannot illuminate what aspects of the training explain the results; that is, our evaluation did not intend to, nor can it, identify the mechanisms of change in selected applicants' research productivity. Trainee feedback and core faculty observation lead us to suggest that the intense exposure to D&I science research content, both substantively and methodologically, the 2 week-long immersion in the summer institutes, and the regular mentoring provided over 2 years explain the successful outcomes. The relative effectiveness of various training components may be testable in future trials. Another IRI evaluation, using social network analysis of the IRI network, revealed that mentoring, and specifically mentoring dosage, was found to be significantly related to future scientific collaborations 2 years later in the form of published papers and new grant submissions (Luke et al. 2016). Over time, it is also important to capture the practice and policy impacts of IRI trainees, likely via case studies.

Second, it could be that Selected Applicants were more productive than the Non-Selected Applicants because they became part of strong IRI network. The collaboration between Selected Applicants with mentors and their peers is, in fact, one of the tenets of IRI (Luke et al. 2016,

Proctor et al. 2013): the summer institute fostered formal networking through peer review of scholarly products as well as informal gatherings during the week. Additionally, informal gatherings were scheduled during D&I conferences during the year (Proctor et al. 2013). Our analysis of IRI network show, in fact, that Selected Applicants reported increased scientific collaboration over time (Luke et al. 2016). The networking among trainees and their mentors in fostering learning and scholarly production (Luke et al. 2016; Rienties and Kinchin 2014) indicate the importance of widening the D&I learning experiences through formal and informal networking opportunities (Proctor et al. 2013).

A third limitation is the absence of any assessment of applicant research context. One of our selection criteria was the availability of a NIH-funded mentor at their home institution, or of a mentor with proven relationship with the applicant if not at the same institution. The mentor did not need to be an expert in D&I but needed to have considerable NIH funding to be able to support the fellow. As with other NIH training programs, the bios and CVs of mentors were requested. While we could have developed a score for strength in research context at the moment of selection, we did not, and doing retrospectively would be biased. Hopefully future training programs can develop such scores to examine the effect of research context in the selection of trainees in their programs.

## Conclusion

It is clear that our Selected Applicants are advancing the field and leading the next generation of D&I researchers compared to those investigators not selected to our institute, as shown by their scholarly production in D&I science. Time is an important variable when planning D&I trainings. The results of the present study as well as our previous data (Luke et al. 2016) show that scholarly productivity increases with every 2 years of IRI mentoring. We believe that our mentoring model will have significant impact in the planning of future D&I trainings. It is our hope that the methodology used to measure the outcomes of IRI, namely, bibliometrics, will also foster further objective and longitudinal measurement of other D&I trainings.

**Funding** Funding was provided by National Institute of Mental Health (Grant No. NIMH R25 MH080916), U.S. Department of Veterans Affairs (Grant No. VA24016D0017), National Institute on Drug Abuse (Grant No. R25 MH080916-07S1), National Heart, Lung, and Blood Institute (Grant No. 3U01HL133994-03S1), National Heart and Lung Institute (Grant No. 5U24HL136790-02) and National Center for Advancing Translational Sciences (Grant no. UL1 TR002345).

## References

- Akl, E. A., Meerpohl, J. J., Raad, D., Piaggio, G., Mattioni, M., Paggi, M. G., et al. (2012). Effects of assessing the productivity of faculty in academic medical centres: A systematic review. *CMAJ: Canadian Medical Association Journal*, *184*, E602–E612. <https://doi.org/10.1503/cmaj.111123>.
- Carlford, S., Roback, K., & Nilsen, P. (2017). Five years' experience of an annual course on implementation science: An evaluation among course participants. *Implementation Science*, *12*(1), 101. <https://doi.org/10.1186/s13012-017-0618-4>.
- Carpenter, C. R., Cone, D. C., & Sarli, C. C. (2014). Using publication metrics to highlight academic productivity and research impact. *Academic Emergency Medicine*, *21*(10), 1160–1172. <https://doi.org/10.1111/acem.12482>.
- Chambers, D. A., Proctor, E. K., Brownson, R. C., & Straus, S. E. (2016). Mapping training needs for dissemination and implementation research: Lessons from a synthesis of existing D&I research training programs. *Translational Behavioral Medicine*, *7*, 593–601. <https://doi.org/10.1007/s13142-016-0399-3>.
- Cramer, J. S., Ramalingam, S., Rosenthal, T. C., & Fox, C. H. (2000). Implementing a comprehensive relative-value-based incentive plan in an academic family medicine department. *Academic Medicine*, *75*(12), 1159–1166. <https://doi.org/10.1097/00001888-200012000-00006>.
- Durieux, V., & Gevenois, P. A. (2010). Bibliometric indicators: Quality measurements of scientific publication. *Radiology*, *255*(2), 342–351. <https://doi.org/10.1148/radiol.09090626>.
- Eloy, J. A., Svider, P. F., Mauro, K. M., Setzen, M., & Baredes, S. (2012). Impact of fellowship training on research productivity in academic otolaryngology. *The Laryngoscope*, *122*(12), 2690–2694. <https://doi.org/10.1002/lary.23749>.
- Froghi, S., Ahmed, K., Finch, A., Fitzpatrick, J. M., Khan, M. S., & Dasgupta, P. (2012). Indicators for research performance evaluation: An overview. *BJU International*, *109*(3), 321–324. <https://doi.org/10.1111/j.1464-410X.2011.10856.x>.
- Giles, J. (2004). Thumbs up for fresh formula to gauge university funding. *Nature*, *427*(6976), 667. <https://doi.org/10.1038/427667b>.
- Gonzales, R., Handley, M. A., Ackerman, S., & O'Sullivan, P. S. (2012). A framework for training health professionals in implementation and dissemination science. *Academic Medicine*, *87*(3), 271–278. <https://doi.org/10.1097/ACM.0b013e3182449d33>.
- Huang, G., Fang, C. H., Lopez, S. A., Bhagat, N., Langer, P. D., & Eloy, J. A. (2015). Impact of fellowship training on research productivity in academic ophthalmology. *Journal of Surgical Education*, *72*(3), 410–417. <https://doi.org/10.1016/j.jsurg.2014.10.010>.
- Luke, D. A., Baumann, A. A., Carothers, B. J., Landsverk, J., & Proctor, E. K. (2016). Forging a link between mentoring and collaboration: A new training model for implementation science. *Implementation Science*, *11*(1), 137. <https://doi.org/10.1186/s13012-016-0499-y>.
- Meissner, H. I., Glasgow, R. E., Vinson, C. A., Chambers, D., Brownson, R. C., Green, L. W., et al. (2013). The US training institute for dissemination and implementation research in health. *Implementation Science*, *8*(1), 12. <https://doi.org/10.1186/1748-5908-8-12>.
- National Heart, Lung, and Blood Institute. (2017). Building the workforce to translate discoveries into health. Retrieved January 2018, from <https://www.nhlbi.nih.gov/news/2017/building-workforce-translate-discoveries-health>.
- National Institute of Mental Health. (2008). *Investing in the Future: National Advisory Mental Health Council Workgroup on Research Training*. Retrieved November 2017, from [http://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/investing-in-the-future\\_42525.pdf](http://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/investing-in-the-future_42525.pdf).

- O’Leary, J. D., & O’Sullivan, O. (2012). Research productivity among trainee anaesthetists in Ireland: A cross-sectional study. *Scientometrics*, *93*(2), 431–438. <https://doi.org/10.1007/s11192-012-0684-y>.
- Padek, M., Brownson, R., Proctor, E., Colditz, G., Kreuter, M., Dobbins, M., et al. (2015a). Developing dissemination and implementation competencies for training programs. *Implementation Science*, *10*(S1), A39. <https://doi.org/10.1186/1748-5908-10-s1-a39>.
- Padek, M., Colditz, G., Dobbins, M., Koscielniak, N., Proctor, E. K., Sales, A. E., et al. (2015b). Developing educational competencies for dissemination and implementation research training programs: An exploratory analysis using card sorts. *Implementation Science*, *10*(1), 114. <https://doi.org/10.1186/s13012-015-0304-3>.
- Padek, M., Mir, N., Jacob, R. R., Chambers, D. A., Dobbins, M., Emmons, K. M., et al. (2018). Training scholars in dissemination and implementation research for cancer prevention and control: A mentored approach. *Implementation Science*, *13*(1), 18. <https://doi.org/10.1186/s13012-018-0711-3>.
- Proctor, E. K., & Chambers, D. A. (2016). Training in dissemination and implementation research: A field-wide perspective. *Translational Behavioral Medicine*, *7*(3), 624–635. <https://doi.org/10.1007/s13142-016-0406-8>.
- Proctor, E. K., Landsverk, J., Baumann, A. A., Mittman, B. S., Aarons, G. A., Brownson, R. C., et al. (2013). The implementation research institute: Training mental health implementation researchers in the United States. *Implementation Science*, *8*(1), 105. <https://doi.org/10.1186/1748-5908-8-105>.
- Rienties, B., & Kinchin, I. (2014). Understanding (in) formal learning in an academic development programme: A social network perspective. *Teaching and Teacher Education*, *39*, 123–135.
- Stamatakis, K. A., Norton, W. E., Stirman, S. W., Melvin, C., & Brownson, R. C. (2013). Developing the next generation of dissemination and implementation researchers: Insights from initial trainees. *Implementation Science*, *8*(1), 1. <https://doi.org/10.1186/1748-5908-8-29>.
- Straus, S. E., Brouwers, M., Johnson, D., Lavis, J. N., Légaré, F., Majumdar, S. R., et al. (2011). Core competencies in the science and practice of knowledge translation: Description of a Canadian strategic training initiative. *Implementation Science*, *6*(1), 127. <https://doi.org/10.1186/1748-5908-6-127>.
- Training Institute for Dissemination and Implementation Research in Cancer (TIDIRC). Retrieved from <https://cancercontrol.cancer.gov/IS/training-education/tidirc/index.html>.
- Ullrich, C., Mahler, C., Forstner, J., Szecsenyi, J., & Wensing, M. (2017). Teaching implementation science in a new Master of Science Program in Germany: A survey of stakeholder expectations. *Implementation Science*, *12*(1), 55. <https://doi.org/10.1186/s13012-017-0583-y>.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.