

A Randomized Controlled Trial to Improve the Success of Women Assistant Professors

Jeanne Ann Grisso, MD, MSc,¹ Mary Dupuis Sammel, ScD,² Arthur H. Rubenstein, MBCh,³ Rebecca M. Speck, PhD,⁴ Emily F. Conant, MD,⁵ Patricia Scott, BA,² Lucy Wolf Tuton, PhD,⁶ Alyssa Friede Westring, PhD,⁷ Stewart Friedman, PhD,⁸ and Stephanie B. Abbuhl, MD⁹

Abstract

Background: Given the persistent disparity in the advancement of women compared with men faculty in academic medicine, it is critical to develop effective interventions to enhance women's careers. We carried out a cluster-randomized, multifaceted intervention to improve the success of women assistant professors at a research-intensive medical school.

Materials and Methods: Twenty-seven departments/divisions were randomly assigned to intervention or control groups. The three-tiered intervention included components that were aimed at (1) the professional development of women assistant professors, (2) changes at the department/division level through faculty-led task forces, and (3) engagement of institutional leaders. Generalized linear models were used to test associations between assignment and outcomes, adjusting for correlations induced by the clustered design.

Results: Academic productivity and work self-efficacy improved significantly over the 3-year trial in both intervention and control groups, but the improvements did not differ between the groups. Average hours worked per week declined significantly more for faculty in the intervention group as compared with the control group (−3.82 vs. −1.39 hours, respectively, $p=0.006$). The PhD faculty in the intervention group published significantly more than PhD controls; however, no differences were observed between MDs in the intervention group and MDs in the control group.

Conclusions: Significant improvements in academic productivity and work self-efficacy occurred in both intervention and control groups, potentially due to school-wide intervention effects. A greater decline in work hours in the intervention group despite similar increases in academic productivity may reflect learning to “work smarter” or reveal efficiencies brought about as a result of the multifaceted intervention. The intervention appeared to benefit the academic productivity of faculty with PhDs, but not MDs, suggesting that interventions should be more intense or tailored to specific faculty groups.

Keywords: academic medicine, women in medicine, faculty development, women's careers in medicine

Introduction

SINCE THE MID-1990s, women have comprised nearly half of the medical school graduates in the United States and yet women are still underrepresented in senior ranks and leadership positions in academic medicine. Numerous stud-

ies have documented the slow advancement of women compared with men faculty in medical and biomedical careers as well as the stubborn persistence of this disparity over the past two decades.^{1–8}

The National Academies' extensive analysis of gender disparity concluded that “Transforming institutional structures

¹Perelman School of Medicine and School of Nursing, University of Pennsylvania, Philadelphia, Pennsylvania.

²Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.

Departments of ³Medicine, ⁴Anesthesiology and Critical Care, ⁵Radiology, and ⁶Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.

⁷Department of Management, DePaul University, Chicago, Illinois.

⁸Department of Management, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania.

⁹Department of Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.

and procedures to eliminate gender bias is a major national task that will require strong leadership and continuous attention, evaluation, and accountability.”⁹ Although many institutions have developed and implemented programs to promote the careers of women faculty, there is little empirical evidence regarding the effectiveness of such programs. Most interventions in academic medicine have been limited to pre-post evaluations,^{10–20} or focused on a single causal problem, such as implicit bias.^{21–23} Only a few have had a randomized design.²³ A recent review of faculty development interventions found that only 2 out of 41 studies included a comparison group and most relied only on post-intervention questionnaires.²⁴

Research in innovative interventions is essential to drive evidence-based change efforts to remediate gender disparities. In response to this critical gap in research, the National Institutes of Health (NIH) developed a funding opportunity to investigate causal factors and interventions that affect the careers of women in science. Funded through the NIH initiative, this study, the NIH-TAC (Transforming Academic Culture) Trial, evaluated the impact of a 3-year cluster-randomized intervention designed to improve key indicators of academic success among women assistant professors as well as to drive broader changes in culture at the department and institutional level of a research-intensive medical school.

Materials and Methods

Sample

Of all clinical and basic science departments at the institution, eligible departments were required to have three or more women assistant professors. For the departments of medicine and pediatrics, divisions were selected as the unit of randomization, because each of these departments had more than 300 faculty (compared with an average of 46 faculty in other departments) and the main governance, rewards, incentives, and expectations of the faculty were determined at the division level. In early 2010, all of the 27 eligible chairs and chiefs (16 departments, 11 divisions) agreed to participate. The protocol was approved by the University of Pennsylvania Institutional Review Board.

Randomization

Randomization was instituted after departments/divisions, and faculty were recruited. To ensure balance between the two arms, randomization was stratified by type of department (*i.e.*, basic science vs. clinical) and type of division (medicine vs. pediatrics). The primary unit of randomization was the department or division; thus, this study represents a cluster-randomized trial. Thirteen departments/divisions were randomized to the intervention group, and 14 were randomized to the control group. Within the intervention group, there were 11 clinical departments/divisions and 2 basic science departments/divisions, and in the control group there were 12 clinical and 2 basic science departments/divisions. Women assistant professors were assigned according to the intervention status of their respective department/division.

Intervention

The intervention was multifaceted and aimed at three levels of the organization for each of the 13 intervention departments/divisions and their women assistant professors. The first level

focused on professional development for women assistant professor faculty. The second level consisted of department/division faculty task forces and their department-specific change initiatives. The third level engaged department/division chairs and chiefs and other institutional leaders, including the School of Medicine Dean and Vice-Deans. We describe each of the three levels of the intervention in greater detail below.

(1) Professional development of women assistant professor faculty. Prior research on the causal factors in gender disparities has identified academic productivity and work-life integration as two of the key barriers to women’s career success.^{8,25} As a result, we designed the professional development level of the intervention to target these two factors affecting career progression. To enhance academic productivity, intervention faculty participated as a cohort in a semester-long, intensive Manuscript Writing Program of eight workshop sessions. Two senior women medical faculty taught evidence-based skills for overcoming common barriers to scholarly productivity combined with peer and senior mentoring.²⁶ To develop skills related to managing work and life demands, the cohort of intervention faculty also participated in the nationally recognized Total Leadership Program.²⁷ The Total Leadership Program consisted of four seminars with intensive activities between sessions, including experiments to improve performance in all spheres of life—work, home, community, and self—by finding mutual value among them through engagement with key stakeholders, reflection, and coaching. In the second and third years of the trial, supplemental professional development sessions were held to maintain peer support, offer additional mentoring, and promote ongoing skill development. Attendance at all faculty development sessions was tracked.

(2) Department/division faculty task force initiatives²⁸. Within each of the 13 intervention departments/divisions, the respective chair or chief each appointed a faculty member to chair their task force and also selected a group of associate and full professor men and women faculty members to serve on the task force. Each task force was charged with creating and implementing a specific department/division initiative to improve their local environment for women’s career success. Through 6 months of intensive strategic planning meetings with a recruited outside facilitator with expertise in leadership and innovation in healthcare, each task force designed an initiative and developed metrics of success. In the subsequent 2 years, the initiatives were implemented and evaluated. Progress and challenges were discussed in annual summits to which members of the 13 intervention task forces, respective chairs and chiefs, the School of Medicine Dean, and senior leadership were invited to attend. In the final summit of February 2013, all faculty in the intervention arm were encouraged to attend as each task force presented their accomplishments and discussed future goals. Examples of task force initiatives include the implementation of home reading work stations in the Department of Radiology, an overhaul of on-call scheduling for junior faculty in a pediatric division, and a grant development initiative in a basic science department that included a mock study section for grant review.

(3) Department/division and institutional leader engagement. The joint principal investigators of the trial worked

closely with the dean, vice-deans, chairs, and chiefs to generate top-down support and accountability for change. An NIH-TAC Trial research team investigator met with each chair or chief of the intervention departments/divisions every 6 months for the duration of the trial to elicit their feedback and continued full participation. The two primary responsibilities of these intervention senior leaders within their departments/divisions were (1) to encourage the full participation of intervention women assistant professors in the trial faculty professional development initiatives and (2) to oversee and support the task force process. In addition, an intervention summit was held each year in which intervention chairs/chiefs were brought together with task force members along with senior leaders in the medical school, to report on progress and learn from each other.

Control group

All chairs/chiefs of eligible departments/divisions agreed to participate before randomization. Both control and intervention leaders understood the randomization scheme and trial methods. After obtaining their consent, the research team did not contact the control chairs/chiefs during the ensuing 3-year trial. There were no interventions in the 14 departments/divisions in the control group or for the women assistant professors in those units. The research team's only contact with the control faculty after randomization was at the end of the study when post-intervention questionnaires and updated curricula vitae (CV) with grant information were requested from the women assistant professors in both control and intervention groups.

Outcomes

Publications and grants. Intervention and control group women assistant professors submitted updated CVs at the beginning and end of the trial. Research assistants, blinded to participant status, abstracted information that included duration in rank, academic track, and information on publications and grants. Two study authors reviewed random samples of CVs and assessed abstraction and coding accuracy. The few differences (<10%) were resolved by consensus. Information on publications included year, first-author status, and whether the publication was classified as peer reviewed. Information on grants included funding organization, dates, and faculty role (principal investigator vs. other). A participant's grant status was classified as improved if any of the following criteria were met comparing the baseline (pre-intervention) and post-intervention CVs: (1) increase in total number of grants; (2) increase in federally funded grants; or (3) increase in number of grants in which the participant was the principal investigator.

Questionnaire measures

Baseline and post-intervention questionnaires were completed by women assistant professor participants through a confidential online survey. The survey instrument included validated measures of socio-demographic factors, education, academic work history, marital status, children, work hours, work self-efficacy, work-family conflict (WFC), and department/division culture.²⁹ A subset of these measures

(described below) were identified as additional key outcomes for the success of women faculty.

Work hours. Work hours were indicated by an open-ended response to the question, "How many hours, per week, do you spend on work-related activities?" If participants provided a range of hours, the mid-point of that range was utilized.

Work-family conflict. WFC is defined as a form of inter-role conflict in which the demands of work and family are "at odds."³⁰ Meta-analytic evidence indicates that WFC is linked to important career-related outcomes such as turnover, burnout, well-being, and job and family satisfaction.^{31,32} Two dimensions of WFC were assessed. *Time-based work-interference-with-family* (TWIF) is the perception that the time demands of work interfere with participation in the family role. *Strain-based work-interference-with-family* (SWIF) is the perception that the stress from work has a negative impact on family life.³³ Each dimension was assessed with three items and scored on a five-point scale from 1 (Strongly Disagree) to 5 (Strongly Agree). Prior research has indicated that these scales are distinct, reliable, and valid measures of facets of WFC.³³ In the current study, the Cronbach's alpha reliability of the sub-scales were $\alpha=0.87$ for strain-based WFC and $\alpha=0.69$ for time-based WFC.

Work self-efficacy. Work self-efficacy was measured by using two items adapted from Riggs et al.³⁴ and scored on a 1–5 Likert-type scale. An average of the two items was used to create the scale ($\alpha=0.80$).

Department/division culture. To assess the specific dimensions of department/division culture that are the most relevant to the experiences of women faculty, we developed a measure for the purposes of this research assessing the *culture that is conducive to women's academic success* (CCWAS). CCWAS is defined as the shared perceptions regarding the extent to which the department/division culture is supportive of women's careers (See Westring et al.²⁹ for an in-depth description of measure development, reliability, and validity). Briefly, CCWAS is a higher-order culture measure that is indicated by four dimensions of culture for women's careers: (1) support for work-life balance; (2) equal access to opportunities; (3) freedom from gender bias; and (4) chair/chief support. The measure consists of 45 items rated on a scale from 1 (Strongly Disagree) to 5 (Strongly Agree), with higher scores indicating a more supportive department/division culture. The overall scale had a reliability of $\alpha=0.95$.

The CCWAS measure of each department/division was determined by calculating the average ratings reported by the women assistant professors whose primary appointments were in that department/division. Our analysis of the CCWAS measure provided empirical support for the equivalence of the measure for both department and division units, and they are, therefore, treated interchangeably.

Non-trial faculty development participation. All participants were asked to report whether they had participated in faculty development sessions during the previous 3 years. Intervention faculty were asked about activities other than those developed for the intervention. Control faculty were

asked whether they had participated in any faculty development program and, if so, how many. No questions were included about the duration or type of educational activity. Co-intervention in the control group was analyzed according to whether they had participated in any faculty development programs and if so, how many, during the period of the trial.

Statistical analysis and sample size

Generalized linear models were used to estimate group averages and to test associations between each outcome and intervention. Key outcomes included individual faculty change in hours worked, work self-efficacy, WFC, grants, and publications. In addition, changes in department/division culture were assessed. Statistical tests were adjusted to account for correlation induced by the clustered design using generalized estimating equations.³⁵ Both within-person factors (age, years in rank, race) and unit-level factors (intervention assignment) were modeled simultaneously. Overall, an intention-to-treat analysis was used. To further clarify our results, we investigated whether level of participation in the faculty development programs (Manuscript Writing Program and Total Leadership Program) within the intervention condition was associated with degree of change for the individual outcomes. The level of participation was assessed by testing for associations with outcomes in high (attending the majority of sessions) versus low (attending less than the majority of sessions) participation levels. A pre-specified subgroup analysis was also conducted based on educational degrees. We tested for associations of intervention condition with outcomes separately by type of degree. We developed three categories of education degrees: (1) MD only—faculty did not have other advanced degrees; (2) MD+ included MD faculty who had other advanced degrees, including PhDs; and (3) PhD faculty with or without other advanced degrees except for MDs. MD/PhDs were included in the MD+ category.

Sample size was estimated based on a hypothesized increase in peer-reviewed publications. We assumed a type 1 error of 5% and a power of 80%. Estimates for expected averages as well as standard deviations (SDs) were gathered from a sample set of CVs previously provided by nonparticipant women faculty in which the average increase was 1.5 publications over 2 years ($SD=3$). We hypothesized that the intervention group would have a rate of increase that was 1/2 SD higher than women in the control group.

Results

In this cluster-randomized trial, all of the 27 departments/divisions participated; 13 departments/divisions were randomized to the intervention group, and 14 were randomized to the control group (Fig. 1). Seventy-five percent ($N=134$) of the 178 eligible women assistant professors participated. Women who participated were similar in academic track distribution to those who did not. During the 3-year period of the trial, one participant withdrew from the study due to health reasons. Of the 133 faculty who participated throughout the trial, baseline questionnaires and CVs were obtained from 132 out of 133 (99.2%); follow-up questionnaires and updated CVs were provided by 126 out of 133 participants (94.7%). Overall, 21 out of 132 (16%) of the women faculty left the university during the 3-year course of the trial. There were no differences in the rates of departure between the experimental

and the control group faculty ($p=0.308$). Follow-up (post-intervention) surveys were sent to the faculty who left the university. Of note, 100% of those who had left the university completed the follow-up questionnaires and their data are included in the analyses provided here.

Table 1 shows the characteristics of the 132 women assistant professors who completed questionnaires and submitted CVs at baseline. There were no statistically significant differences between the intervention and control groups. Overall, 85% were married or in a domestic partnership, and 75% had one or more children living at home. Of the participants, 72% were physicians and most (71%) were in the clinician-educator track, a track with rigorous scholarship expectations within a defined probationary period (“up or out” policy).

Trial outcomes

There were substantial improvements in major outcomes within both the intervention and control groups. Specifically, we tested pre/post differences within the intervention and control groups separately and found statistically significant improvements in both groups in grants, total peer-reviewed publications, work self-efficacy scores, and hours worked per week (Table 2). However, there were few statistically significant differences in the intervention group compared with the control group (see Table 2). No differences were detected between intervention and control faculty in any measure of academic productivity (*i.e.*, publications and grant status). When we examined publications by year, though not statistically significant, there was a trend toward greater numbers of publications in the intervention group in 2012 and 2013. We did not find statistically significant differences in any of the other key outcome measures except for hours worked per week. Participants in the intervention group reported a significantly larger decrease in work hours compared with the control group (-3.82 hours vs. -1.39 hours per week, respectively; $p=0.006$).

The overall evaluation was based on an intention-to-treat analysis. However, in supplemental analyses, we assessed whether there were differences within the intervention group assistant professor women based on their level of participation in the Manuscript Writing Program and Total Leadership Program. We found that those with high participation rates (attending more than half of sessions) had more than a two-fold increase in peer-reviewed publications (rate ratio = 2.15, 95% confidence interval [CI]: 1.25–3.39, $p=0.006$) compared with those who participated less often (Fig. 2).

We explored the possibility that the impact of the intervention varied by type of academic degree. In faculty with PhD degrees (but without MD degrees), the intervention group had a 92% increase in first-author publications and a 131% increase in first-author peer-reviewed publications compared with control faculty with similar degrees (first-author publication rate ratio = 1.92, 95% CI: 1.13–3.29, $p=0.02$; first-author peer-reviewed publication rate ratio = 2.31, 95% CI: 1.34–3.99, $p=0.003$). Among faculty with MD degrees, there were no differences between the intervention and control groups in any measure. We did not find statistically significant differences by educational degree for changes in work hours or for levels in participation in the intervention sessions.

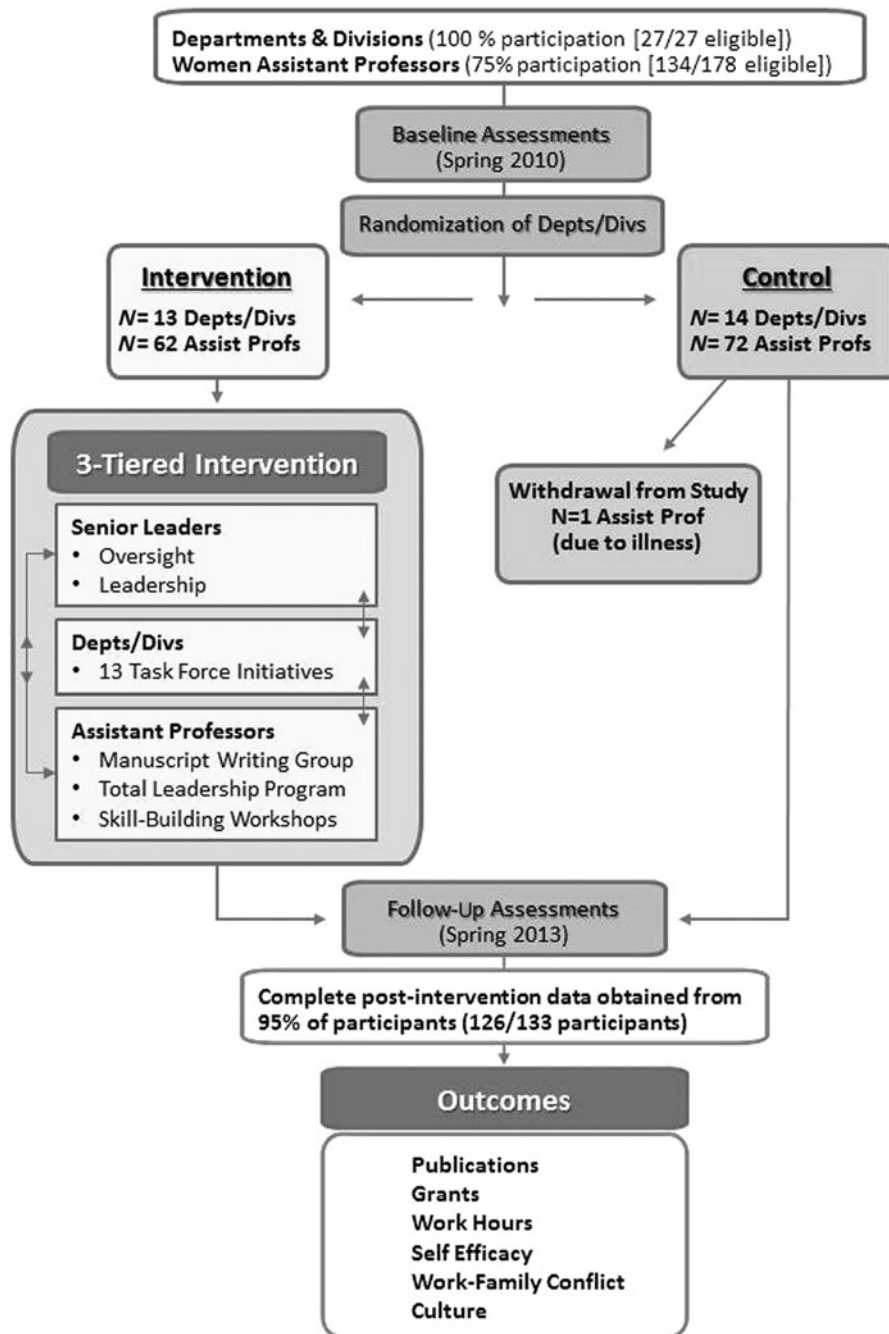


FIG. 1. Overview of the NIH-TAC trial. NIH-TAC, National Institutes of Health-Transforming Academic Culture.

We asked all faculty about their participation in non-trial faculty development programs. We asked intervention faculty about activities in addition to those provided through the intervention. Although both intervention and control faculty reported high rates of attending one or more non-trial faculty development programs, there were no statistically significant differences in participation rates between the groups. Of note, more than 65% of women faculty in the control group reported having participated in one or more faculty development programs during the trial period, with nearly half (49%) participating in two or more programs. However, we cannot compare the characteristics of the faculty development programs (e.g., duration, type, intensity, etc.) between the groups, as we did not collect that information.

Discussion

We evaluated the impact of a 3-year multifaceted intervention in a cluster-randomized trial at a research-intensive academic health center on the career success of women assistant professors. By implementing a multilevel approach to institutional change, we demonstrated the feasibility of a school-wide effort, including both clinical and basic science departments. At the same time, our trial demonstrates the challenges involved in conducting “real world” research of professional development programs and institutional change efforts when blinding of the intervention is not possible and when outcomes are best measured over many years, possibly even decades.

TABLE 1. CHARACTERISTICS OF WOMEN ASSISTANT PROFESSORS AT STUDY ONSET

	<i>Total cohort</i>		<i>Intervention group</i>		<i>Control group</i>		<i>p-value</i>
Number	132		62		70		
Avg. years as asst. professor (SD)	4.36 (2.59)		4.24 (2.72)		4.46 (2.48)		0.63
Academic track, <i>N</i> (%)							
Tenure	18	13.6%	8	12.9%	10	14.3%	0.42
Clinician-educator	93	70.5%	42	67.7%	51	72.8%	
Research	21	15.9%	12	19.4%	9	12.9%	
Degree, <i>N</i> (%)							0.31
MD only	45	34.9%	21	33.9%	24	35.8%	
MD+ (includes other advanced degrees, including PhDs)	48	37.2%	20	32.2%	28	41.8%	
PhD only and PhD + (includes other advanced degrees but excludes MDs)	36	27.9%	21	33.9%	15	22.4%	
Median age (years) (Q1–Q3)	40 (37–44)		41 (39–44)		40 (37–42)		0.07
Race, <i>N</i> (%)							0.37
African American	10	7.6%	4	6.5%	6	8.7%	
White	79	60.3%	42	67.7%	37	53.6%	
Asian	36	27.5%	13	21.0%	23	33.3%	
Hispanic/other	6	4.6%	3	4.8%	3	4.4%	
Marital status, <i>N</i> (%)							0.80
Single/other	20	15.3%	10	16.1%	10	14.5%	
Married or in domestic partnership	111	84.7%	52	83.9%	59	85.5%	
% with children living at home	98	74.8%	45	72.6%	53	76.8%	0.58

SD, standard deviation.

We found important but comparable levels of improvement in academic productivity and work self-efficacy in both intervention and control groups. If the trial had not included a concurrent control group, the improvements in our outcomes would have likely been attributed to the intervention.

One interpretation of our results is that our intervention was not effective in improving these outcomes. However, despite a rigorous randomized controlled study design, there are alternative explanations that may suggest that, in fact, the multifaceted intervention was potentially effective. An alternative interpretation is that the major improvements in both groups resulted from the trial itself, that is, the Haw-

thorne effect. The trial was highly visible at our institution, and it is likely that extensive knowledge of the trial and strong support of the senior leadership at multiple levels affected both control and intervention departments/divisions and the women assistant professors in those departments/divisions. A notable strength of the trial was the full engagement of all eligible departments and divisions (100% participated). In addition, the Dean of the School of Medicine chaired the trial's National Advisory Committee and the President of the University served as one of the members. The legitimacy and prestige of NIH funding contributed, at least in part, to the commitment to the trial that was demonstrated at all levels of

TABLE 2. OUTCOMES: COMPARISONS OF INTERVENTION VERSUS CONTROL GROUP

<i>Outcome</i>	<i>Intervention</i>		<i>Control</i>		<i>Intervention vs. control</i>	<i>p-value</i>
Work self-efficacy change	0.18*		0.24**		-0.06 (difference)	0.642
Work-family conflict change						
TWIF	-0.13		-0.05		-0.08 (difference)	0.541
SWIF	-0.20		-0.23*		0.03 (difference)	0.879
CCWAS culture score change (<i>n</i> = 109)	0.03		0.13*		-0.09 (difference)	0.274
Hours worked per week (change)	-3.82***		-1.39*		-2.43 (difference)	0.006
% with improved grants	41.6%***		55.7%***		RR = 0.75 (0.54, 1.03)	0.08
<i>Number of publications, Avg. No.</i>	<i>2009</i>	<i>2012</i>	<i>2009</i>	<i>2012</i>	<i>Rate ratio^a (95% CI)</i>	
Total	3.98	4.58	4.72	6.79***	0.80 (0.63, 1.02)	0.07
First author	1.41	1.28	1.57	1.42	1.00 (0.67, 1.50)	0.99
Peer review total	2.44	3.48***	3.25	4.86**	0.95 (0.68, 1.33)	0.78
Peer review first author	0.76	0.95	0.66	0.69	1.06 (0.58, 1.95)	0.85

^aRate ratio (RR)=ratio of increase in publications from 2009 to 2012 in intervention over controls.

p-values for within group change over time, **p* < 0.05, ***p* < 0.01, and ****p* < 0.001.

TWIF, Time-based work-interference-with-family; SWIF, Strain-based work-interference-with-family; CCWAS, culture conducive to women's academic success; CI, confidence interval.

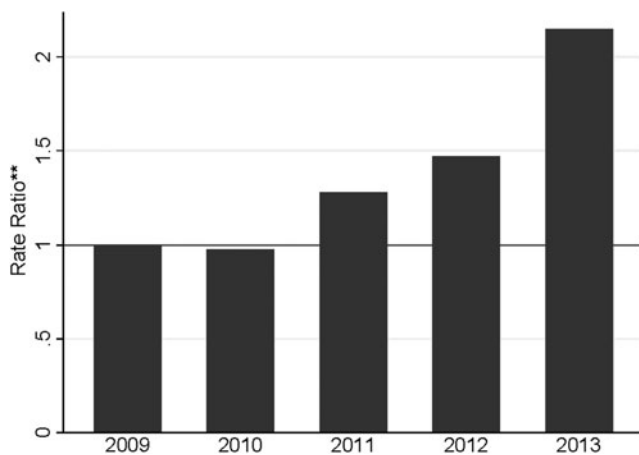


FIG. 2. NIH-TAC efficacy analysis: impact of level of participation (High participation: attended $\geq 50\%$ versus low participation: attended $< 50\%$ of intervention sessions) on number of peer-reviewed publications. *Vertical columns* indicate the ratio of the increase in the number of publications in high versus low levels of participation for each year. **Rate ratio values > 1 indicate that the number of publications was greater in the high participation group compared with the low participation group. Models adjusted for promotion track, years as asst. professor, education level, work self-efficacy, and clustering effect, using general linear models.

the medical school. We believe that this large-scale effort over 3 years may have brought about some measure of widespread institutional change, thus resulting in the reported improvements in both the intervention and control groups.

Other results support the interpretation that the intervention was at least partially effective. We found that PhD faculty in the intervention group appeared to benefit in increasing the number of publications whereas intervention MD faculty did not. It is possible that the MD faculty, who have rigorous clinical requirements, less flexible schedules, and “divided attention,” were, therefore, less able to focus on academic productivity. These results suggest that the MD faculty may need a more intensive or different intervention. Interventions tailored to specific faculty groups may be the most effective.

Another positive finding was that intervention faculty who attended half or more of the sessions increased their publication rates significantly more than those who participated less. This may indicate that a sustained level of participation in the interventions was required to demonstrate an impact on academic productivity. It is also possible that faculty with sufficient time and/or flexibility to fully participate had the motivation, resources, and support to increase their productivity.

The academic productivity measures evaluated in this study represent key outcomes assessed for promotion, tenure, and attainment of leadership positions. However, given the relatively long time frame required to prepare and revise manuscripts and grants, the fact that CVs were collected shortly after the 3 year intervention concluded may not have allotted adequate time to detect intervention effects. The trend toward greater productivity in the intervention group toward the final year of the trial may indicate a need for a

longer assessment period. A longer time frame would arguably be better for measuring other outcomes too, such as the culture change in the departments/divisions that were being impacted by all levels of the intervention. The Task Force initiatives, in particular, may have needed more time to fully evolve to have a visible impact on the measurement of culture.

Yet another promising finding was that the hours worked per week declined significantly more in the intervention group compared with the control group. Given equivalent levels of increases in the number of grants and publications, it appears that intervention faculty attained comparable levels of academic productivity in fewer work hours. This is not surprising given the purpose and methods of the intervention professional development programs for the assistant professors (*i.e.*, the Manuscript Writing Program and the Total Leadership Program). In Total Leadership, participants designed and implemented experiments that were intended to better align their actions and their values. The goal of the program is that participants gain a more effective integration of work and the other parts of their lives and, in doing so, may also develop a more personally effective use of time. In addition, the Manuscript Writing Program may have enhanced efficiency in writing, thus also reducing the hours required for equivalent levels of productivity. It is also possible that the Task Force initiatives were addressing department/division-specific issues that enabled more efficient use of faculty time. Given the high work demands placed on academic faculty, programs that provide tools for working “smarter” and for reducing, even by a relatively small amount, the average work week hours are important in addressing the work-life challenges of academic careers.

There were several limitations to our study. First, the study was limited to women assistant professors. This focus was predicated on the well-established fact that women have long been underrepresented in senior ranks and leadership positions in academic medicine.⁹ It was beyond the scope or required intent of the NIH funding opportunity to carry out interventions and measure outcomes in both men and women faculty. Second, it is important to note that the results from this trial may not be generalizable to other academic medical centers with different metrics of success for faculty advancement or without an “up or out” system.

Conclusion

The NIH-TAC Trial was a multifaceted cluster-randomized intervention trial involving an entire medical school over 3 years with the overarching goals of improving the academic success of women assistant professors and improving the culture that was conducive to women’s success in individual departments and divisions. We suspect that our intervention, which was highly visible and included the strong support of senior leadership, may have impacted the entire institution with significant improvements in both intervention and control groups. The greater decrease in reported levels of work hours may suggest efficiencies gained in intervention faculty without declines in academic productivity.

The results of this trial provide important insights into the challenges of studying large-scale interventions that are aimed at both individual and institutional outcomes with the goal of creating environments where women faculty can

succeed fully in their careers. We recommend that further randomized trials be designed to study interventions targeting specific subgroups of faculty with sufficiently long follow-up periods to fully assess important outcomes. Indeed, the generation, implementation, and evaluation of evidence-based change initiatives for academic medicine are essential for institutional transformation.

Acknowledgments

The authors acknowledge the contributions and leadership of Susmita Pati, MD, and Josef Reum, PhD, for their early work in the Task Force Initiative. They also thank Karin McGowan, PhD, and the late Seema Sonnad, PhD, for their expertise and outstanding commitment to this trial in their roles as leaders of the Manuscript Writing Program. Dr. Sonnad's sudden death in May 2015 was a great loss, and the authors dedicate this article to her memory.

Clinical Trial Registration

This study was not registered as a clinical trial, as the interventions did not target clinical or health outcomes. The study meets the exclusion criteria listed in: <https://clinicaltrials.gov/ct2/managerecs/fdaaa#WhichTrialsMustBeRegistered>.

Funding Source

This research was funded by the National Institutes of Health (grant no. R01-NS069793-03). The NIH had no role in the design or implementation of this investigation.

IRB Approval

The University of Pennsylvania Institutional Review Board approved this study protocol before data collection.

Author Disclosure Statement

No competing financial interest exist.

References

1. Tesch BJ, Wood HM, Helwig AL, Nattinger AB. Promotion of women physicians in academic medicine. Glass ceiling or sticky floor? *JAMA* 1995;273:1022–1025.
2. Nonnemaker L. Women physicians in academic medicine: New insights from cohort studies. *N Engl J Med* 2000;342:399–405.
3. Sexton KW, Hocking KM, Wise E, et al. Women in academic surgery: The pipeline is busted. *J Surg Educ* 2012;69:84–90.
4. Wong CA, Stock MC. The status of women in academic anesthesiology: A progress report. *Anesth Analg* 2008;107:178–184.
5. Stapleton FB, Jones D, Fiser DH. Leadership trends in academic pediatric departments. *Pediatrics* 2005;116:342–344.
6. Rayburn WF, Schrader RM, Fullilove AM, Rutledge TL, Phelan ST, Gener Y. Promotion rates for assistant and associate professors in obstetrics and gynecology. *Obstet Gynecol* 2012;119:1023–1029.
7. Lautenberger DM, Dandar MA, Raezer CL, Sloane RA. The state of women in academic medicine: The pipeline and pathways to leadership, 2013–2014. Washington, DC: Association of American Medical Colleges, 2014.
8. Jena AP, Khullar D, Ho O, Olenski AR, Blumenthal DM. Sex differences in academic rank in US medical schools in 2014. *JAMA* 2015;314:1149–1158.
9. Committee on Maximizing the Potential of Women in Academic Science and Engineering NAOs, National Academy of Engineering, and Institute of Medicine. Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering. Washington, DC: The National Academy Press, 2007.
10. Fried LP, Francomano CA, MacDonald SM, et al. Career development for women in academic medicine: Multiple interventions in a department of medicine. *JAMA* 1996;276:898–905.
11. Morahan PS, Voytko ML, Abbuhl S, Wara DW, Thorson J, Cotsonas CE. Ensuring the success of women faculty at AMCs: Lessons learned from the National Centers of Excellence in Women's Health. *Acad Med* 2001;76:19–31.
12. Welch JL, Jimenez HL, Walthall J, Allen SE. The women in emergency medicine mentoring program: An innovation approach to mentoring. *J Natl Med Assoc* 2009;4:362–366.
13. Von Feldt JM, Bristol M, Sonnad S, Abbuhl S, Scott P, McGowan KL. The brief CV review session: One component of a mosaic of mentorship for women in academic medicine. *J Natl Med Assoc* 2009;101:873–880.
14. Emans SJ, Goldberg CT, Milstein ME, Dobriner J. Creating a faculty development office in an academic pediatric hospital: Challenges and successes. *Pediatrics* 2008;121:390–401.
15. McDade SA, Richman RC, Jackson GB, Morahan PS. Effects of participation in the Executive Leadership in Academic Medicine (ELAM) program on women faculty's perceived leadership capabilities. *Acad Med* 2004;79:302–309.
16. Richman RC, Morahan PS, Cohen DW, McDade SA. Advancing women and closing the leadership gap the Executive Leadership in Academic Medicine (ELAM) program experience. *J Womens Health Gen Based Med* 2001;10:271–277.
17. Carnes M, VandenBosche G, Agatista PK, et al. Using women's health research to develop women leaders in academic health sciences: The national centers of excellence in women's health. *J Womens Health Gen Based Med* 2001;10:39–47.
18. Jaggi R, Butterton JR, Starr R, Tarbell NJ. A targeted intervention for the career development of women in academic medicine. *Arch Intern Med* 2007;167:343–345.
19. Bickel J, Atkinson BF, Cohen LS, et al. Increasing women's leadership in academic medicine: Report of the AAMC project implementation committee. *Acad Med* 2002;77:1043–1061.
20. Valantine HA, Grewal D, Ku MC, Shih MC, Stevenson D, Pizzo PA. The gender gap in academic medicine: Comparing results from a multifaceted intervention for Stanford faculty to peer and national cohorts. *Acad Med* 2014;89:904–911.
21. Girod S, Fassiotto M, Grewal D, et al. Reducing implicit gender leadership bias in academic medicine with an educational intervention. *Acad Med* 2016;91:1143–1150.
22. Carnes M, Devine PG, Isaac C, et al. Promoting institutional change through bias literacy. *J Divers High Educ* 2012;5:63–77.
23. Carnes M, Devine PG, Baier Manwell L, et al. The effect of an intervention to break the gender bias habit for faculty at

- one institution: A cluster randomized, controlled trial. *Acad Med* 2015;90:221–230.
24. Steinert Y, Maismith L, Mann K. Faculty development initiatives designed to promote leadership in medical education. A BEME systematic review. *Med Teach* 2012;34:483–503.
 25. Strong EA, De Castro R, Sambuco D, et al. Work–life balance in academic medicine: Narratives of physician-researchers and their mentors. *J Gen Intern Med* 2013;28:1596–1603.
 26. Sonnad SS, Goldsack J, McGowan KL. A writing group for female assistant professors. *J Natl Med Assoc* 2011;103:811–815.
 27. Friedman SD. Total leadership: Be a better leader, have a richer life. Boston, MA: Harvard Business School Publishing, 2008.
 28. Pati S, Reum J, Conant E, et al. Tradition meets innovation: Transforming academic medical culture at the University of Pennsylvania's Perelman School of Medicine. *Acad Med* 2013;88:461–464.
 29. Westring AF, Speck RM, Sammel MD, et al. A culture conducive to women's academic success: Development of a measure. *Acad Med* 2012;87:1622–1631.
 30. Greenhaus JH, Beutell NJ. Sources of conflict between work and family roles. *Acad Manage Rev* 1985; 10:76–88.
 31. Amstad FT, Meier LL, Fasel U, Elfering A, et al. A meta-analysis of work-family conflict and various outcomes with a special emphasis on cross-domain versus matching-domain relations. *J Occup Health Psychol* 2011;16:151–169.
 32. Ford MT, Heinen BA, Langkamer KL. Work and family satisfaction and conflict: A meta-analysis of cross-domain relations. *J Appl Psychol* 2007;92:57–80.
 33. Carlson DS, Kacmar KM, Williams LJ. Construction and initial validation of a multidimensional measure of work–family conflict. *J Vocat Behav* 2000;56:249–276.
 34. Riggs M, Warka J, Babasa B, Betancourt R, Hooker S. Development and validation of self-efficacy and outcome expectancy scales for job-related applications. *Educ Psychol Meas* 1994;54:793–802.
 35. Liang KY, Zeger SL. Regression analysis for correlated data. *Annu Rev Publ Health* 1993;14:43–68.

Address correspondence to:
Stephanie B. Abbuhl, MD
Department of Emergency Medicine
Perelman School of Medicine
University of Pennsylvania
Emergency Medicine Administrative Suite
Ground Ravdin Building
3400 Spruce Street
Philadelphia, PA 19104

E-mail: abbuhls@uphs.upenn.edu