Evaluation of the Processes and Outcomes of a Physician Leadership Program: The Continuous Feedback Loop Design

Gregg M. Gascon, PhD; Huey T. Chen, PhD; Liliana Morosanu, MPH; Victor H. Chen, PhD; Phil Cass, PhD; Robert Falcone, MD

Introduction: Physicians assume leadership roles in their health care organizations and practices often without support or training. The Physicians Leadership Academy provides integrated leadership seminars, mindfulness training, and executive coaching in a 10-month curriculum to physicians across area organizations.

Methods: Program evaluators responded to stakeholders' need for continuous program improvement by developing a continuous feedback loop evaluation design incorporating a program monitoring system and a theory-driven program evaluation. Given the size of the 2019 to 20 cohort ($n = 19$), a one-group pretest/posttest design was used to assess the mechanisms of the program (mindfulness and wellbeing) along with knowledge development, emotional intelligence, and personal and professional growth. The assessments used a combination of published and administrator-developed assessments to address the unique aspects of the program. Doing such ensured continuous improvement and sustainability for the program.

Results: The cohort of physicians demonstrated significant engagement and learning across the curriculum, improved mindfulness, and improved capacity of the providers to affect their health care system and communities.

Conclusions: The utility of the program was demonstrated through quantitative and qualitative analyses. Implications of the methodology for future evaluations of program developments are discussed.

Keywords: continuous program improvement, leadership, mindfulness, program evaluation, theory-driven evaluation

DOI: 10.1097/CEH.0000000000000436

Many physicians across the United States and other countries face challenges of high workload and increasing performance demands; as a result, some suffer from stress, depression, and burnout,[1,2] which gradually reduce the quality of their contributions to the health care community.[3] Regardless, physicians are often expected to assume leadership roles without having received adequate training. Leadership training is not uncommon in the health care industry. Systematic reviews point to the fact that leadership programs targeting physicians aim to strengthen their leadership competencies to improve organizational performance; however, very few are focused on personal growth and awareness

and even fewer use interactive learning and feedback or assess system-level outcomes.[4] One program—the Physicians Leadership Academy (PLA)—approaches physician leadership development differently, using mindfulness and well-being as mechanisms of transformational change that, while not targeted at addressing stress and burnout, hold the potential to address one of the key challenges to health care mental health today, that of mindfulness. Mindfulness programs seem effective in helping physicians cope with stress, anxiety, and burnout[5–7] while also improving attention, emotion, and self-regulation.[8,9] When mindfulness awareness is included, assessing the results of integrating mindfulness-awareness practice with the leadership skills of health care practioners[10] through a theory-driven evaluation would offer the opportunity to understand its mechanisms of change.

The purpose of this study is to describe the process of designing and evaluating a novel physician leadership program in which we mapped out the elements of the training program and tested each component in the chain of reasoning tying all elements together to investigate program effectiveness.

THE PROGRAM AND ITS THEORETICAL FOUNDATION

In 2012, the Columbus Medical Association conducted fifty two-hour interviews with member physicians to better understand the reality of their professional lives and identify what a regional medical society could offer of support and assistance to physicians. The interviews yielded two interrelated findings. First, many physicians were overwhelmed with the production demands being placed on them by the institutions they worked for, the electronic documentation required of them by state and

Dr. Gregg M. Gascon: *Adjunct Assistant Professor, Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, OH.* **Dr. Huey T. Chen:** *Professor of the Department of Public Health and Director of the Center for Applied Research and Evaluation in the College of Health Professions, Mercer University, Atlanta, GA.* **Ms. Liliana Morosanu:** *Instructor and Coordinator, Program and Student Academic Advancement in the College of Health Professions, Mercer University, Atlanta, GA.* **Dr. Victor H. Chen:** *Lecturer, Department of Management, University of North Texas, Denton, TX.* **Dr. Phil Cass:** *Director, Physicians Leadership Academy, Columbus, OH.* **Dr. Robert Falcone:** *CEO, Columbus Medical Association and its Affiliates, 1390 Dublin Road, Columbus, OH.*

Correspondence: *Gregg M. Gascon, PhD, Department of Biomedical Informatics, College of Medicine, The Ohio State University, 1800 Cannon Drive, 250 Lincoln Tower, Columbus, OH 43210; e-mail: gascon.2@osu.edu.*

federal regulations, the litigious nature of medicine, and myriad-related issues. These factors produced anxiety, depression, and overall life dissatisfaction which are now referred to as burnout. Second, many physicians felt powerless to change the system and often described themselves as victims of the system and not its leaders. In its current state, American medicine leaves few avenues for patients to advocate for themselves, so that responsibility often lands with the physician as well.

For the medical association this raised two questions: (1) Should the medical association create a well-being initiative? (2) Should the medical association also create a leadership program to facilitate capacity building? Realizing that these two questions required one approach, the PLA was founded in 2013 in Columbus, OH, and trained its first cohort of 17 physicians in 2014 to 15. Since then, 8 to 20 physicians have been trained annually (2015–2016, 15; 2016–2017, 16; 2017–2018, 8; 2018–2019, 18; 2019–2020, 19; 2020–2021, 20; and 2021–2022, 25). A process evaluation was conducted on the first two cohorts to identify and improve implementation fidelity. This evaluation focused on the 2019 to 20 cohort ($n = 19$). From that beginning through today, the goal of the program has been to reconnect practicing physicians with their original purpose for becoming physicians, by helping them grow personally and professionally, and ultimately apply their new strengths to the improvement of the health care system and the community they serve.

The program spans 10 months with a September orientation; 3-day October retreat; November–March seminars, practice, and coaching; 3-day April retreat; May–June seminars, practice, and coaching; June graduation ceremony; and award of 35 continuing medical education credits. Additional annual retreats are also available to alumni. Fellows are expected to attend each session, practice meditation, engage with their executive coach, and apply everything in their daily practice. The PLA offers continuous support for the practice of skills, application of knowledge, and development of attitudes that promote personal and professional growth, thus blending the three major components of the program discussed next. Given the resources involved, the program cost $11,500 per fellow in 2019 to 20, but 65% of the cost of the program is paid for by the Columbus Medical Association for the benefit of physicians and the communities they serve.

The Leadership Seminar

The leadership seminar provides the backbone for the program. Its content is grounded on Otto Scharmer Theory U. Theory U posits that the institutional failures seen in today's world are the result of our blindness to the deeper dimensions of leadership and transformational change, the place from where effective leadership and social action come into being.[11–13] At its root, Theory U focuses participants' attention on the connections between who they are as a person to how they respond to themselves and the world; this opens them up to the possibility of positive change and leadership. According to Scharmer, "…the way we pay attention to a situation, individually and collectively, determines the path the system takes and how it emerges" (Scharmer, 2007, p. 5). Each monthly seminar focuses on a different approach to leadership (eg, the leadership mindset, organizational culture and system change, managing in complex environment, introduction to systems and design thinking, understanding and making sense of data, advocacy to create change, etc.) that connects back to Theory U and lasts from 3:00

to 8:00 PM These sessions are expanded during the autumn and spring retreats.

Mindfulness Meditation Practice

The second foundation of the [program] program, mindfulness meditation, helps physicians develop emotional and social intelligence. According to Kabat-Zinn,[14] mindfulness is a state in which one is highly aware and focused on the reality of the present moment, accepting and acknowledging it, without getting caught up in one's thoughts or emotional reactions to that reality. Mindfulness training in health care has been shown to decrease stress, anxiety, depression, and burnout and increase empathy, relaxation, and well-being.[15,16] When practitioners used mindfulness techniques, they were more compassion and develop more effective communication with patients.[17] The PLA training follows the most common object-placement method, in which the object is the breath (or the experience of breathing). Emphasis on the breath allows participants to more easily be reminded of their mindfulness practice as they go about their daily lives—their breath is always with them. Physicians are encouraged to rest simply in their experience, in their basic nature, and experience what it is to be human.

Executive Coaching

The third foundation of the PLA program, executive coaching, is designed to guide physicians in taking steps toward achieving their personal and professional goals and further enhance leadership development and performance. The PLA uses three external coaches, trained and certified through the Hudson Institute of Coaching (Santa Barbara, CA) and credentialed through the International Coaching Federation (Lexington, KY). They engage physicians for 10 sessions by following the coaching process presented in Table 1. Fellows have access to coaching outside of the monthly sessions as well, given that the coaching process is by no means linear, as physicians could reformulate goals and strategies as they progress through training. Coaching has been found to have been used successful in improving well-being.[18]

In the present context, the executive coaching process encapsulates five steps. In the first, contracting, the participants establish the agenda, commitment, expectations objectives, and roles. As the participants enter the process of executive coaching, they develop a balanced view of the circumstances of the process through assessment and feedback (the second step). Developmental action plans are established in the third step, as the participants set the goals, strategies, and tactics for development although the coaching process in the PLA allows for goals, pacing, and direction may change over time. The fourth step is devoted to implementation. Here, the participants practice, refine, and reinforce new development tools, techniques, and/or behaviors. The fifth step, evaluation and follow-up, uses the measurement of progress and plans for ongoing development.

EVALUATION APPROACH AND DESIGN

Program Theory

Evaluation literature indicates it is essential for evaluators to help stakeholders clarify their program theory (or theory of change), to support communication about the program, and

TABLE 1.

Implementation and Outcome Measures as Viewed Through the Kirkpatrick Model

	Measure	Definition, Question Items, and Instrument	Scale
Reaction	Attendance	Primary measure of participant's engagement collected by sign-in sheets	Count
	Usefulness	The seminar topic is useful to me as a person	1 = strongly
	Relevancy	The seminar topic is relevant to me as a physician	disagree
	Satisfaction	Overall, I was satisfied with the content presented on the topic	5 = strongly agree
	Faculty evaluations	Five items assessing the faculty:	1 = strongly
		1. knowledgeable about the topic	disagree
		2. communicate clearly	5 = strongly
		3. encourage discussion on the topic	agree
		4. answered questions appropriately	
		5. overall satisfied with the manner topic was presented	
Learning	Knowledge/ understanding	The level of understanding the topic taught during seminars	1 = very low 5 = very high
	Practice	Over the past 30 d, how many times have you practiced mindfulness meditation?	Count
		Over the past 30 d, how many times have you met with your executive coach?	
	Homophily	Participants developed a community of practitioners by the end of the program—assessed the extent to which fellows:	1 = strongly
		1. trusted each other	disagree
		2. considered each other honest	5 = strongly
		3. considered each other supportive of colleagues maintained a respectful relationship with each other	agree
Behaviors	Mindfulness (mechanism)	Mindfulness Awareness Attention Scale (MAAS)	1 = almost always
		15 items	6 = rarely
	Well-being (mechanism)	World Health Organization Well-being Index (WHO-5)	0 = at no time
		Subjective measure of five items: feeling cheerful, relaxed, active, fresh, and rested	5 = all the time
	Emotional intelligence	Assessment of emotional self-awareness, emotional self-control, adaptability, positive outlook, and achievement orientation	0 = not at all 10 = significantly
		To what degree do you believe that the program has improved your ability to demonstrate emotional intelligence?	
	Social intelligence	Ability to empathize, show compassion, listen to understand, influence, and show inspirational leadership	0 = not at all
		To what degree do you believe that the program has improved your ability to demonstrate social intelligence?	10 = significantly
	Cognitive capacity	To what degree do you believe that the program has improved your cognitive capacities?	0 = not at all 10 = significantly
	Leadership capacity	To what degree do you believe that the program has improved your capacity for leadership?	0 = not at all 10 = significantly
	Personal growth	Six items assessing:	1 = strongly
		1. greater sense of calmness	disagree
		2. greater sense of self-confidence	5 = strongly
		3. more confidence in articulating decisions	agree
		4. better work/life balance	
		5. better physical health	
		6. better mental health	
	Professional growth	Six items assessing	1 = strongly
		1. patient centricity	disagree
		2. listening to patients	5 = strongly
		3. communicating effectively with patients	agree
		4. collaborating effectively with other clinicians in patient treatment	
		5. presence during the time with patients	
		6. seeking the advice of other professionals more readily	
Results	Health care system application	Physicians' ability and capacity to affect the health care system	1 = strongly
		Five items assessing ability to	disagree
		1. lead meetings that achieve results	5 = strongly
		2. be more creative in solving problems	agree
		3. create more sustainable and life-affirming solutions	
		4. help create more future-oriented solutions	
		5. lead more patient-centric systems of care	
	Community application	The extent to which physicians found themselves in a better position to build stronger communities from a health care leadership perspective.	1 = strongly disagree
		Four items assessing	5 = strongly
		1. enhanced understanding of the community	agree
		2. enhanced understanding of the health care system and its patients	
		3. preparedness to work for a more life-affirming community	
		4. preparedness to advocate for laws and policies that are more life affirming.	

guide the evaluation design.[19–21] Given that the PLA is a training program, the stakeholders' program theory fit the Kirkpatrick[22] training evaluation model with its four levels of outcomes: reaction, learning, behaviors, and results. In the PLA program, reaction gauges participants' satisfaction with classes; learning measures acquired knowledge and skills from the training, and expected behaviors are optimal mindfulness and well-being, social and emotional intelligence, and leadership capacity. Stakeholders believe that the program must affect mindfulness and well-being as a mechanism that enables the achievement of all other outcomes. Finally, results examine to what degree physicians apply the newly acquired skills within their organizations and communities.

The Kirkpatrick evaluation model also provided the base for selecting indicators for program monitoring activities (M&E system)—further discussed in the Methodology section—and allowed the establishment of two feedback loops to support continuous program improvement (Figure 1)—the second goal of evaluation stressed by stakeholders.

Evaluation Design

Communication with key sponsors indicates that they need an assessment of the program's effects for both accountability and ongoing program improvement purposes. The systems thinking literature illustrate the usefulness of feedback loops for improving the performance of a system.[23,24]

Two feedback loops were proposed to meet the stakeholders' needs: *feedback loop 1*: The monitoring and evaluation system (M&E) approach to program evaluation provided a platform for collecting implementation and participants' reaction and learning data, as activities unfolded.[25] These data were communicated to instructors to inform their preparation for the next class. *Feedback loop 2*: The theory-driven evaluation (TDE) approach to program evaluation provided the theoretical basis for assessing the relationships among the intervention, mechanisms, and outcomes.[26] Arguably, novel interventions should produce the results (or outcomes) that stakeholders desire if the theoretical foundation assumptions of the interventions are correct, and the proposed mechanisms are conducive to outcomes achievement. TDE was used to assess whether participants progressed toward desired outcomes at midcourse and graduation, by assessing the mechanisms and expected outcomes. The integration of these two feedback loops for this study is illustrated in Figure 1.
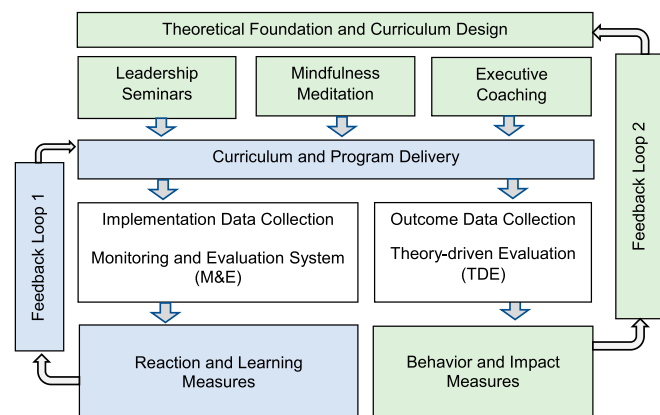

FIGURE 1. Feedback loop evaluation design.

Research Method, Measures, and Data Analysis

This section explains the rationale for the research methods adopted and provides an overview of the methods, measures, and analysis of the data.

Research Design and Data Analysis

The one-group pretest–post-test design was used to assess the mechanisms of the program (mindfulness and well-being) along with personal and professional growth at three points in time: baseline, to gather pretest data; midcourse at 5 months, to assess progress toward desired outcomes; and finally at 10 months, to assess final scores.[27] Data were analyzed through paired sample t tests. The one-group post-test design was used to assess all other outcomes at the end of the program through a one-group t test using an acceptable standard value for comparison as agreed on by stakeholders; research demonstrating the utility of the one-group t test for extremely small sample sizes has been widely cited.[28]

Power was calculated using the expected cohort size of the 2019 to 20 fellows. An a priori power analysis was conducted using G*Power version 3.1.9.4 to test the difference between matched pairs using a one-tailed test, a large effect size (d = 0.70), and 0.05 alpha. The results demonstrated that the 2019 to 20 cohort was sufficient to achieve a power of 0.80.[29,30]

Measures and Data Collection

All measures collected from participants are presented in Table 1. Implementation measures were collected monthly through self-administered questionnaires distributed electronically to inform the program by feedback loop 1 (M&E). Outcome measures were collected at the baseline, midpoint, and conclusion of the program through self-administered questionnaires distributed electronically to inform the program by feedback loop 2 (TDE). The other outcomes presented in Table 2 were measured at the end of the program and tested against a standard value as agreed on by stakeholders. Thus, Table 2 provides the outcomes associated with the measures articulated by a training model component in Table 1.

At baseline, the participants were asked to identify their age, sex, race, marital status, title, medical specialty, their graduation year, and the extent to which they were acquainted with their fellow cohort members (homophily). To test for the presence of homophily, fellows were provided with a list of their cohort members at the orientation and asked if they knew them sufficiently to give an informed opinion on their character.[31] These queries were repeated monthly to provide evidence of the development of a community of practitioners. Participants were also tested for initial levels of well-being (WHO-5) and mindfulness (MAAS) because these were the mechanisms expected to facilitate the achievement of the rest of the outcomes. At midterm (5 months), participants were tested for mindfulness, well-being, and personal and professional growth; additional final measures were collected again after the program (Table 2).

The evaluation approach and design were reviewed by the Ohio State University Institutional Review Board in August 2019 and found to constitute program evaluation and, hence, not human subjects research as defined under 45 CFR 46.102(d).

RESULTS

Among the 2019 to 20 cohort (*n* = 19), 52.6% were male, 63.2% were Caucasian, 15.8% were African American, 15.8%

TABLE 2.
Results of the Outcome Measures

	Outcome Measure	Across Sessions M(SD)	Baseline or Standard Value M(SD)	Midterm, M(SD)	Final, M(SD)	t-value	P	95% CI	Effect Size, Hedges g_{av}
Reaction	Attendance (n = 19)	18.5 (0.972)							
	Usefulness 1–5 scale	4.64 (0.118)							
	Relevancy 1–5 scale	4.64 (0.115)							
	Satisfaction 1–5 scale	4.55 (0.233)							
Learning	Knowledge and understanding	Before: 2.37 (0.602); after 3.99 (0.210)							
	Mindfulness meditation practice/month	10.2 (avg. Of averages)							
	Coaching sessions/month	0.78 (ibid)							
	Trust, respect, honesty, and support for cohort members (1–10 scale)	9.7 across all 4 attributes							
Behaviors	Mindfulness (mechanism)		53.11 (13.78)	54.56 (12.78)	61.42 (11.63)	−4.579	.000	−1.041, −0272	−0.66
	Well-being (mechanism)		16.0 (4.24)	13.44 (3.59)	16.84 (3.66)	-.831	.417	−0.771, −0.338]	-.22
	Emotional intelligence		8.0		8.84 (1.12)	3.281	.004	0.30, 1.38	0.72
	Social intelligence		8.0		8.95 (1.08)	3.288	.001	0.43, 1.47	0.84
	Cognitive capacity		8.0		8.16 (1.74)	0.395	.697	0.30, 1.38	
	Leadership capacity		8.0		9.05 (0.97)	4.729	.000	0.58, 1.52	1.04
	Personal growth			24.00 (4.07)	25.42 (3.32)	−2.112	.050	−0.823, 0.25	−0.40
	Professional growth			23.56 (4.33)	25.68 (3.00)	−2.610	.018	−1.06, −0.06	-.56
Results	Health care system application		4.0		21.68 (2.647)	2.774	.013	0.82, 0.59	-.61
	Community application		4.0		17.63 (2.50)	2.846	.011	0.11, 0.71	-.63

were Asian, and 5.3% reported multiple races. The mean (standard deviation) age was 41.2 years (7.3 years).

Implementation Measures
The participants' feedback on reaction and learning measures was high. Attendance was high throughout the program averaging 97% across 10 sessions, despite the normal workload of physicians and the appearance and escalation of a pandemic. Usefulness scores ranged from 4.37 to 4.75, relevancy from 4.42 to 4.81, and satisfaction from 4.21 to 4.81; there were no scores below 4 (agree) (Table 2).

Faculty evaluations were distributed, analyzed, and evaluated for accountability purposes and were not reported for the program evaluation but rather for program improvement. Participants' understanding of each seminar topic increased significantly in each session. Table 2 indicates the scores of topic understanding improved significantly after attending each of the sessions.

After each session, participants reported on the number of times they practiced mindful meditation and met with their executive coach. Participants reported practicing mindfulness meditation 6 to 12 times per month throughout the program and meeting monthly with their executive coach, thus meeting two of the objectives of the program. On a monthly cadence, this implementation information was shared with program stakeholders.

In contrast to the assumption of homophily, few knew each other before the first session: The median number of fellows known before the PLA was a median of 2 (n = 19) with an interquartile range of 3. By the end of the program, PLA cohort members had built meaningful relationships with one another, judging each other uniformly and strongly in positive terms. At the end of the program, [program] fellows characterized their relationships with one another on a 1—10 scale with 10 indicating the highest positive rating. The mean (standard deviation) scores were 9.61 (.608) for trusting, 9.67 (.594) for honest, 9.72 (.575) for supportive, and 9.78 (.548) for mutually respectful.

Outcome Measures
By the end of the program, PLA fellows improved mindfulness, personal and professional growth, leadership capabilities, and social and emotional intelligence by statistically significant margins, with small-to-large effect sizes but not well-being or cognitive capacities. By these measures and the development of a community of practitioners with strong relationships, the program was successful.

DISCUSSION
This study illustrated the application of the continuous feedback loop approach to evaluating the 2019 to 2020 cohort of the PLA. One of the core features of the program evaluation was the use of reliable, valid, and objective assessments drawn from the research literature and the use of targeted, program-developed assessments that targeted specific program outcomes. Continuing education programs rely on evidence-based assessments for reliable and valid measurement of progress

toward program goals. Finding those that meet that criteria and come without cost or copyright constraints are essential to keeping participant and program evaluation costs reasonable, a core tenet of the four standards (accuracy, feasibility, propriety, and utility) of program evaluation.[32]

The PLA represents a unique approach to physician leadership development offered and supported by a regional medical association for the direct benefit of physicians and the communities they serve. Its resource requirements are significant, and the time and personal time investments it requires of participants are substantial. To justify the costs, the program sponsors and directors have examined the program processes twice and the outcomes once. To continue measure outcomes continuously, it worked with evaluators to ensure that all evaluation and assessment procedures could be replicated into the future for continuous improvement.

There are limitations associated with this program evaluation that stem from the uniqueness of the program; those include the assessments used for the program that were developed by program administrators. Their face validity and internal consistency were noted above, but the specificity of their purpose may limit their generalizability. This is related to the fact that the evaluation was conducted on a single cohort and at a single location.

This study contributes to the literature of evaluation design. Health care and other decision makers have shown increasing attention to continuous quality improvement of programs.[33] Consequently, the evaluation literature has had to adapt evaluation approaches for addressing continuous improvement issues in health care training programs to keep up with the pace of changes in the field.[34,35] This study illustrated the use of the continuous feedback-loop design to inform program improvement for practitioners and theorists alike.

This study also contributes to the literature of theory-driven program evaluation by articulating how formal and stakeholder theories can be tested alongside one another. While the characteristics and merits of the two have been addressed in the literature,[19] this study offered examples of how these theories could be surfaced, examined, and tested in a way that offers unique insights into program stakeholders who simultaneously want to see their ideas tested in conjunction with the formal theories that they used as program mechanisms (ie, mindfulness and well-being). Clarifying the program theory in this way allows stakeholders to connect their program with wider audiences in a variety of settings.

Lessons for Practice

- The choice of approaches and methods for program evaluation should be situational, tailored to the nature of the program, the needs of stakeholders, and the purpose of the evaluation.

- Continuing education programs rely on assessments to measure progress toward program goals. Finding those that achieve those goals and come without cost or copyright constraints are essential to keeping participant and evaluation costs reasonable, a goal of program evaluation.

- To address the need for continuous improvement, evaluators need to adapt existing program evaluation approaches to provide program stakeholders with the information they need in real time.

REFERENCES

1. Dzau VJ, Kirch DG, Nasca TJ. To care is human—collectively confronting the clinician-burnout crisis. *N Engl J Med.* 2018;378:312–314.
2. West CP, Dyrbye LN,, Shanafelt TD. Physician burnout: contributors, consequences and solutions. *J Intern Med.* 2018;283:516–529.
3. Sklar DP. Leadership in academic medicine: purpose, people, and programs. *Acad Med.* 2018;93:145–148.
4. Frich JC, Brewster AL, Cherlin EJ, et al. Leadership development programs for physicians: a systematic review. *J Gen Intern Med.* 2015;30:656–674.
5. Goyal M, Singh S, Sibinga EMS, et al. Meditation programs for psychological stress and well-being: a Systematic Review and Meta-analysis. *JAMA Intern Med.* 2014;174:357–368.
6. Regehr C, Glancy D, Pitts A, et al. Interventions to reduce the consequences of stress in physicians: a Review and Meta-Analysis. *J Nerv Ment Dis.* 2014;202:353–359.
7. Fendel JC, Bürkle JJ,, Göritz AS. Mindfulness-based interventions to reduce burnout and stress in physicians: a systematic review and meta-analysis. *Acad Med.* 2020;96:751–764.
8. Ludwig DS, Kabat-Zinn J. Mindfulness in medicine. *JAMA.* 2008;300:1350–1352.
9. Reb J, Atkins PWB. *Mindfulness in Organizations: Foundations, Research, and Applications.* Cambridge, England: Cambridge University Press; 2015.
10. Wasylkiw L, Holton J, Azar R, et al. The impact of mindfulness on leadership effectiveness in a health care setting: a pilot study. *J Health Organ Manag.* 2015;29:893–911.
11. Scharmer CO. *Addressing the Blind Spot of Our Time. An Executive Summary of the New Book by Otto Scharmer, Theory U: Leading from the Future as it Emerges*; 2007. Available at: http://gudrunmiller.de/wp-content/uploads/2015/11/Adressing-the-blind-spot.pdf.
12. Scharmer O,, Yukelson A. Theory U: from ego-system to eco-system economies. *JCC.* 2015:35–39.
13. Scharmer O. *The Essentials of Theory U: Core Principles and Applications.* Oakland, CA: Berrett-Koehler Publishers; 2018.
14. Kabat-Zinn J. *Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life*: Hachette Books; 2009.
15. Krasner MS. Association of an educational program in mindful communication with burnout, empathy, and attitudes among primary care physicians. *JAMA.* 2009;302:1284–1293.
16. Pipe TB, Bortz JJ, Dueck A, et al. Nurse leader mindfulness meditation program for stress management: a randomized controlled trial. *JONA: J Nurs Adm.* 2009;39:130–137.
17. Grepmair L, Mitterlehner F, Loew T, et al. Promoting mindfulness in psychotherapists in training influences the treatment results of their patients: a randomized, double-blind, controlled study. *Psychother Psychosom.* 2007;76:332–338.
18. Theeboom T, Beersma B,, Vianen AEM. Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *J Posit Psychol.* 2014;9:1–18.
19. Chen H. *Practical Program Evaluation: Theory-Driven Evaluation and the Integrated Evaluation Perspective* (2nd ed). Thousand Oaks, CA: Sage Publications; 2015.
20. Chen HT. *Theory-driven Evaluations.* Thousand Oaks, CA: Sage Publications; 1990.
21. Donaldson SI. *Program Theory-Driven Evaluation Science: Strategies and Applications.* New York, NY: Lawrence Erlbaum; 2007.
22. Kirkpatrick D. *Evaluating Training Programs.* San Francisco: Berrett-Koehler Publishers, Inc; 1994.
23. Wolfe A. Institute of Medicine report: crossing the quality chasm: a new health care system for the 21st century. *Policy Polit Nurs Pract.* 2001;2:233–235.
24. Trochim WM, Cabrera DA, Milstein B, et al. Practical challenges of systems thinking and modeling in public health. *Am J Public Health.* 2006;96:538–546.
25. Mrazek PB, Biglan A,, Hawkins JD). *Community-monitoring Systems: Tracking and Improving the Well-Being of America's Children and Adolescents.* Washington, DC: Society for Prevention Research; 2007.
26. Chen HT, Morosanu L, Bellury LM, et al. Multiwave formative evaluation of a retention program for minority nursing students: intended effects, unintended consequences, and remedial actions. *Am J Eval.* 2020;41:71–88.

27. Shadish WR, Cook TD,, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Houghton Mifflin; 2002.

28. de Winter JCF. Using the Student's t-test with extremely small sample sizes. *Pract Assess Res Eval*. 2013;18.

29. Faul F, Erdfelder E, Buchner A, et al. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 2009;41:1149–1160.

30. Cumming G,, Calin-Jageman R. *Introduction to the New Statistics: Estimation, Open Science, and beyond*. New York, NY: Routledge; 2016.

31. McPherson M, Smith-Lovin L,, Cook JM. Birds of a feather: homophily in social networks. *Annu Rev Sociol*. 2001;27:415–444.

32. Yarbrough DB, Shulha LM, Hopson RK, et al. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users* (3rd ed). Thousand Oaks, CA: Sage Publications; 2010.

33. Zamboni K, Baker U, Tyagi M, et al. How and under what circumstances do quality improvement collaboratives lead to better outcomes? A systematic review. *Implement Sci*. 2020;15:1–20.

34. Schwartz AR, Siegel MD,, Lee AI. A novel approach to the program evaluation committee. *BMC Med Educ*. 2019;19:1–10.

35. Trombetta C, Capdeville M, Patel PA, et al. The program evaluation committee in the adult cardiothoracic anesthesiology fellowship – harnessing opportunities for program improvement. *J Cardiothorac Vasc Anesth*. 2020;34:797–804.