A New Way of Evaluating Effectiveness of URM Summer Pipeline Programs

Krystal Hill\textsuperscript{1}  
Catherine Raney\textsuperscript{2}  
Kelli Jackson\textsuperscript{3}  
H Moses Murdock\textsuperscript{4}  
Erika Dawson\textsuperscript{5}  
Roy Hamilton\textsuperscript{2}  
Horace DeLisser\textsuperscript{2}  
Mira Mamtani\textsuperscript{6}  
Jaya Aysola\textsuperscript{2}  

\textsuperscript{1}Department of Internal Medicine, Baylor College of Medicine, Houston, TX, USA; \textsuperscript{2}Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; \textsuperscript{3}Department of Internal Medicine, University of Pennsylvania, Philadelphia, PA, USA; \textsuperscript{4}Brigham and Women’s Hospital, Boston, MA, USA; \textsuperscript{5}Perelman School of Medicine Office for Diversity and Inclusion, University of Pennsylvania, Philadelphia, PA, USA; \textsuperscript{6}Department of Emergency Medicine, Perelman School of Medicine, Philadelphia, PA, USA

**Purpose:** Many academic medical centers fund educational opportunities (pipeline programs) for students who are underrepresented in medicine (URM). However, there is a sparsity of published literature on pipeline programs and an even smaller body of published literature that investigates program effectiveness.

**Methods:** In a retrospective cohort study (n=12) of the Provost’s Summer Mentorship Program-Medicine (SMPM), we evaluated students’ rating of program effectiveness, students’ rating of the program’s impact on their mindsets, and SAT scores. Several program mindsets, including sense of belonging (inclusiveness) in the health professions and connection to mentors in the medical field, reflect common barriers that prevent URM students from pursuing careers in medicine as outlined in pipeline literature. We describe program effectiveness using mean and median ratings of SMPM effectiveness, ratings of mindsets, and SAT scores. We used Wilcoxon Rank Sum to assess pre and post program differences in ratings of mindsets and SAT scores.

**Results:** SMPM was effective for learners. The overall mean rating for SMPM effectiveness was 4.27. Mindsets for confidence, interest, sense of belonging, college mentorship, and physician mentorship were statistically different from the start to the end of SMPM (p<0.05), with mean improvement of about 34%, 41%, 44%, 180%, and 140% respectively. The mean pre and post SMPM SAT scores as well as 4-month follow-up SAT mean scores were 713 (SD:155), 813 (SD:83), and 1058 (SD:147), respectively. There was a statistically significant difference between all three SAT scores (p<0.05).

**Conclusion:** In addition to providing educational support, our pipeline program effectively increased students’ sense of belonging in the medical field and their connections to physician mentors, which are two common barriers for URM students who are interested in medicine.

**Keywords:** education, underrepresented in medicine, minority, socioeconomic status

**Introduction**

While Black, Indigenous, and people of color (BIPOC) are projected to comprise a majority of our population by 2050, several racial/ethnic minorities, including Blacks and Latinos, continue to be underrepresented in the physician workforce.\textsuperscript{3,4} Blacks and Latinos comprise only 5.0% and 5.8% of the physician workforce compared to their representations of 12.8% and 18.4% in the general US population, respectively.\textsuperscript{5,6} Asians as a broader category are well represented, but certain subpopulations, such as Southeast Asians (Laotian, Indonesian, Cambodian) continue to be underrepresented, comprising less than 1% of Asian medical students. Additionally, individuals from low socioeconomic backgrounds are also underrepresented. Research from the AAMC shows that about a quarter of medical
school enrollees reported parental income in the top 5% of US households and students from families with incomes in the lowest quintile make up 6% of medical students. This trend has not changed in the last three decades. The term Underrepresented in Medicine (URM) defines the aforementioned groups. Underrepresentation of minority physicians is also seen locally in Philadelphia. While Philadelphia’s estimated population of 1,580,863 consists of 42.6% Black, 14.1% Hispanic, and 0.4% American Indian and Alaskan Native, the total percent of URM physicians in Pennsylvania was about 4.5% in 2014.

There are numerous reasons why demographic disparities in medical school enrollees and the physician workforce exist, one being inequality in educational opportunities and resources that begin in early childhood and continue across the lifespan. Of note school system inequalities are downstream effects of historical discriminatory policies and social practices. The upstream legacies of income inequality, de-jure and later de-facto segregation, redlining of neighborhoods, and a public school system funded by neighborhood tax revenue all factor into downstream inequality in school systems. For URM students, other commonly reported barriers to matriculation and success in college and medical school include lack of mentorship and training in non-inclusive environments.

Pipeline programs, which provide educational opportunities via resources and mentorship to low-income and/or BIPOC students, represent one part of the solution for improving physician workforce disparities. At the University of Pennsylvania, efforts to increase the percentage of minority physicians have been underway for many years. One pipeline program—the Provost’s Summer Mentorship Program—is a month-long program comprised of five graduate school-based pipelines including medicine, nursing, law, engineering, and dentistry. All students are URM and attend Philadelphia public schools. Since its inception in 2006, about 255 out of 276 SMP alumni—from all five schools—are either currently enrolled at or have graduated from a postsecondary institution; 21 total SMP alumni are either currently enrolled at or have graduated from the University of Pennsylvania.

In our analysis, we present an overview of the Summer Mentorship Program Medicine (SMPM) curriculum. We examine the 2018 cohort of the SMPM (n=12). We do not include additional analysis of prior cohort years because metrics used to measure program effectiveness were not collected. We found a paucity of literature with information on pipeline curriculum and effectiveness. Our paper adds to this small body of literature. In this paper, we provide (1) an overview of the curriculum, (2) students’ rating of pipeline effectiveness, (3) impact on their mindsets and (4) SAT score improvements.

Methods
Study Design
We conducted a retrospective cohort study to determine the preliminary effectiveness of our program intervention using short-term outcomes of students’ rating of program domains, students’ rating of the pipeline’s impact on their mindsets, and SAT score improvement. We evaluated the effectiveness of the pipeline program using the Donald Kirkpatrick Model.

Participants
The University of Pennsylvania institutional review board gave the research exemption status after it was determined to be an educational quality improvement initiative. Participants verbally consented to inclusion in the study. Twelve students comprised the 2018 SMPM cohort and participated in the study.

The Pipeline Program Intervention
We developed SMPM’s curriculum using the conceptual framework of Kern’s six-step model of curricular development consisting of identification of the problem with general needs assessment, needs assessment for targeted learners, goals and objectives, educational strategies, implementation, and evaluation and feedback. The identified problem is that a small percentage of URM students apply and graduate from college, medical school, and residencies due to a variety of structural societal barriers, which includes educational disparity, lack of inclusive environments, and the absence of mentors in medicine. Consequently, there is a need for educational interventions known as pipeline programs. We determined the needs of our targeted learners—high school students from Philadelphia public schools—in the context of extreme educational disparities in Philadelphia’s school system.

SMPM’s curriculum is comprised of six goals/objectives: academic enhancement, career discussion with medical students and physicians, health science lessons, direct experiences with patients including medical procedures, mentorship, and introduction to public health science with community-based problem solving. We chose these domains based on literature review of the most efficacious...
Educational strategies included didactic lectures and opportunities for experiential learning. Examples included lessons on the periodic table, cell biology, and weekly SAT preparation. Students committed at least one day a week to SAT preparation and took a pre- and post SMPM SAT diagnostic exam. For career mentoring, the students learned first-hand about the career paths of prominent URM doctors and medical students at Penn. Health science lessons involved lessons on diabetes and hypertension, field trips, laboratory exercises in microbiology, urinalysis, gross anatomy dissections, physical exam and medical history-taking workshops, and CPR training. Patient care and procedures included practice with laparoscopic surgery equipment, phlebotomy, and lumbar punctures. Mentorship included shadowing physicians from various sub-specialities. For public health and community-based problem-solving, students developed solutions to community-based health problems, which they presented to their peers and parents at the program’s closing ceremony. The program was implemented in 2018 with support and funding from the University of Pennsylvania’s Vice Provost Office and Perelman School of Medicine.

Data Collection
We analyzed data from the following sources: post-program survey in which students used a Likert scale (1–5) to rate activities and pre- and post-program mindsets, pre and post SMPM SAT diagnostic, and self-reported 4-month post-SMPM SAT scores. Pre and post program mindsets involved ratings of four different mindsets—confidence in doing well in challenging science courses, interest in becoming a doctor, sense of belonging in the health professions, and connection to mentors in the medical field. The last two mindsets have been identified in the pipeline literature as reasons URM students who are interested in careers in medicine ultimately decide not to become medical professionals.1,2 The SMPM pre and post SAT diagnostics were administered during the program and were given to us by the central summer mentorship program office. Four-month follow up SAT scores were voluntarily self-reported to us by the students.

Student Selection and Characteristics
Participants were selected from 50 applications with preference given to underrepresented groups. Selection was based on a rating system partly aligned with cumulative disadvantage factors well documented in the literature to score and choose applicants.24,25 We collected disadvantage factors such as race/ethnicity, gender, socioeconomic status, and high school/college readiness index. The college readiness index (CRI) is a score developed by US News and World Report that assesses the effectiveness of a public high schools’ college preparatory curriculum.26 All students in the cohort attended public schools with very low CRI scores except for one.

All students were sophomores who attended Philadelphia public schools and were URM by socioeconomic and BIPOC status (All below 130% of the poverty line; 7 Black, 2 Latino, 3 Asian: 1 Cambodian, 2 Pakistani). 33% were male (4 Male, 8 Female). 58% (7) submitted a letter of recommendation from a science teacher. All showed interest in the medicine pipeline and ranked medicine amongst their top three pipelines with 75% (9) ranking medicine as their number one choice. Table 1 displays baseline student characteristics including GPA and pre-program mindsets rated on a Likert scale, 1–5. Pre-program mindsets involved ratings of five different mindsets—confidence in doing well in challenging science courses, interest in becoming a doctor, sense of belonging in the health professions, connection to college/medical student mentors, and connection to physician mentors.

Success Outcome Measures and Justification
We evaluated the effectiveness of the curriculum using a validated and reliable model, named Donald Kirkpatrick’s model for evaluating an educational intervention.27 Kirkpatrick’s model includes Level One:

Table 1 Baseline Characteristics of Program Participants

| Variable                                              | Mean (SD) | Median (IQR) |
|-------------------------------------------------------|-----------|--------------|
| GPA                                                   | 3.50 (0.68)| 3.84 (0.43) |
| Pre-program Mindsets (Likert scale 1–5: 1=poor, 5=excellent) | 3.17 (1.03) | 3 (1)       |
| Confidence in doing well in challenging science courses | 3.08 (1.38) | 3.5 (1.5)   |
| Interest in becoming a doctor                         | 2.83 (1.47)| 3 (2.25)     |
| Sense of belonging in the medical field               | 1.67 (0.98)| 1 (1)        |
| Connection to college/med school mentors (Other mentorship) | 1.83 (1.34) | 1 (1)        |

Note: *N = 12 for all variables.*
Reaction, Level Two: Learning, Level Three: Behavior, and Level Four: Results. We defined our outcomes based on Donald Kirkpatrick’s (DK) model for evaluating an educational intervention and learners achieved a Level One as measured by outcome 1 and 2 and achieved a Level Two as measured by outcome 3.

- Outcome 1: Student rating of SMPM effectiveness (aligned with DK level One reaction).
- Outcome 2: Post-SMPM mindset score (aligned with DK level One reaction).
- Outcome 3: Immediate post-program and 4-month post-SMPM SAT scores (aligned with DK level Two learning).

Statistical Analysis
Our analysis had four main objectives. First, we tabulated the frequencies of demographic and socioeconomic characteristics of our participants as well as the distributions of their baseline characteristics (Table 1). Second, we evaluated the distributions of the students’ perceptions of program effectiveness by domain (Table 2). Third, we tabulated and compared mean and median pre- and post-program mindset scores (Figure 1) and SAT scores (Table 3), using Wilcoxon Sum Ranking to assess for significant differences. Our threshold for statistical significance was 0.05. We used JMP 16.

Results
Students’ Rating of Program Effectiveness
We present mean (SD) scores as well as median (IQR) for each program domain to assess students’ rating of program effectiveness (Table 2). The overall mean rating for SMPM effectiveness was 4.27 (Likert scale 1–5). Program domains with the highest mean ratings were patient care and procedures (4.71), health science lessons (4.39), mentorship (4.22), and career discussions with URM physicians (4.22). The highest rated activities within these domains were cadaver dissections, simulations (placing a nasogastric tube, performing a lumbar puncture), learning cardiopulmonary resuscitation), and mentorship. The rest of domains were academic enhancement (3.8), and public health/community-based problem solving (3.92).

Students’ Rating of Program’s Impact on Their Mindsets
Mindsets for confidence in doing well in challenging science courses, interest in becoming a doctor, sense of belonging in the medical profession, and connection to mentors were statistically different form the start to end of SMPM (p<0.05). The mean scores for pre-program mindsets of confidence in doing well in challenging science courses, interest in becoming a doctor, sense of belonging in the medical profession, connection to college/med school mentors, and connection to physician mentors were 3.17, 3.08, 2.83,1.67, and 1.83 respectively. The mean scores for post-program mindsets– confidence in doing well in challenging science courses, interest in becoming a doctor, sense of belonging in the medical profession, connection to college/med school mentors, and connection to physician mentors– were 4.25, 4.33, 4.08, 4.67, and 4.42 respectively, with mean improvement of about 34%, 41%, 44%, 180%, and 140% respectively. Figure 1 illustrates pre and post program mindsets.

SAT Scores
We used three SAT test scores in our analysis: pre- and post-program and four-month follow-up (Table 3). The pre and post SMPM SAT scores as well as 4-month follow-up SAT mean scores were 713 (SD:155), 813 (SD:83), and 1058 (SD:147), respectively. For four-month follow-up scores, two data points were never reported (n=10). There was a statistically significant difference between all three SAT scores (p<0.05).

For reference, the 4-month post-SMPM SAT interquartile ranges were similar to interquartile ranges of admission SATs at five of six local Philadelphia colleges– La Salle, St. Joseph, University of the Sciences, Temple, and Drexel. Of note, the University of Pennsylvania’s 25th
percentile SAT admission score is 1420, about 110 points higher than the maximum 4-month post-SMPM SAT score.

**Discussion**

African Americans/Blacks, Latinos, Native Americans, people from low-income backgrounds, and certain groups of Asian descent make up a small percentage of the physician workforce in comparison to their percentages in the US population. To address one component of this multifaceted issue, medical schools have created intensive educational initiatives under the pipeline program label. Although many high school pipeline programs exist, demonstrating their effectiveness using long-term outcomes, including enrollment in medical school, is costly.

Table 3 Mean and Median Pre and Post Program SAT Diagnostic Scores as Well as 4-Month Post Program SAT Scores (Outcome 3)

|                      | Pre-Program SAT Diagnostic Scores | Post-Program SAT Diagnostic Scores | 4-Month Post-Program SAT Scores | Statistical Significance |
|----------------------|----------------------------------|-----------------------------------|--------------------------------|--------------------------|
| N                    | 12                               | 12                                | 10                             | p<0.05                   |
| Mean (SD)            | 713 (155)                        | 813 (83)                          | 1058 (147)                     | p<0.05                   |
| Median (IQR)         | 663 (175)                        | 813 (94)                          | 1040 (150)                     | p<0.05                   |

Notes: aPre-program SAT Diagnostic was administered during first week of program. bPost-program SAT diagnostic was administered during last week of program, and c4-month post-program SAT was self-reported 4 months after program end; the four-month follow-up SAT score is missing two data points (n=10). dStatistically significant difference between the three tests.
and takes many years to track. Additionally, pipeline programs’ enrollment numbers are relatively low due to resource limitations. Thus, it is critical to analyze short-term meaningful outcomes that programs can more easily collect.

It is important to note that societal disparities and biases influence student mindsets, often leading to internalized biases by students. For URM students, internalized biases take the form of feeling a sense of exclusion (not belonging), isolation, or lacking confidence, and often contributes to URM students’ decision to not pursue medicine. Thus, we strongly encourage pipelines to work to increase students’ sense of belonging and confidence via connecting students with role models in medicine, and positive reinforcement. Additionally, students’ sense of belonging should be tracked to assess the effectiveness of a pipeline program.

Of note, the pre-program mindset of connection to mentors in the medical field (MCMM) was significantly lower for all students compared to other pre-program mindsets. This is unsurprising given that disadvantaged students have disparate access to mentors in the medical field compared to non-disadvantaged students. Importantly, like sense of belonging, URM students’ cite low connection to mentors as a contributor to their decision not to pursue medical careers. Further work must be performed to understand the relationship between disadvantaged students, mentorship, and decision to pursue medical careers.

Limitations of our study include a small sample size (n=12), lack of distal outcomes, absence of controls, and potential selection bias. The significant improvements seen in mindset, sense of belonging, and SAT scores is promising given the small sample size and suggests the utility of measuring such outcomes in future cohorts and other programs as proximal endpoints. Longitudinal data, such as college and medical school enrollment rates, are necessary for understanding whether the intervention addresses the larger goal of increasing the number of URM physicians in the medical workforce. While future work on long-term impact of pipeline programs with comparative controls is ideal, such evaluations are challenging to design and suffer from attribution biases. As is the case in selecting students for admission to selective programs, selection bias needs to be considered. Students may have been just as successful without the intervention. Therefore, short term improvements in noted barriers to success may prove to be good indicators of the program effects on career trajectories.

Conclusion
This description of our pipeline program and evaluation demonstrates that preliminary effectiveness can be characterized by using meaningful short-term outcomes. Pipeline programs should be the focus of future investigation and would benefit from a unified way of measuring short-term and long-term success. This study offers critical insights into how to evaluate a program for intermediate outcomes with demonstrated associations with long term measures of success. We hope that the implementation of highly effective pipeline programs will aid in the goal of generating a physician workforce best able to deliver high-quality, culturally sensitive care to all patients, regardless of demographic factors.

Acknowledgments
We would like to acknowledge the University of Pennsylvania’s Vice Provost Office, SMP Program Coordinator Don Harrison, Perelman School of Medicine’s Program for Diversity and Inclusion, Office of Inclusion and Diversity, Eve J. Higginbotham MD (Dean of OID), Gregg Lipschik MD, and Anna Delaney for their efforts in either the pipeline program or research project. This research paper is dedicated to the memory of Gail Oberton (former SMP Coordinator).

Disclosure
All authors report no conflicts of interest in this work.

References
1. Fritz CDL, Press VG, Nabers D, Levinson D, Humphrey H, Vela MB. SEALS: an innovative pipeline program targeting obstacles to diversity in the physician workforce. J Racial Ethn Health Disparities. 2016;3(2):225–232. doi:10.1007/s40615-015-0131-x
2. Patel SI, Rodriguez P, Gonzales RJ. The implementation of an innovative high school mentoring program designed to enhance diversity and provide a pathway for future careers in healthcare related fields. J Racial Ethn Health Disparities. 2015;2(3):395–402. doi:10.1007/s40615-015-0086-y
3. Ortman JM, Guarneri CE. United States Population Projections: 2000 to 2050. United States Census Bureau; 2009.
4. AAMC. Diversity in medical education: facts & figures 2014. Washington, D.C.: American Association of Medical Colleges; 2014. Available from: https://www.aamc.org/data-workforce/reports/439214/ workforcediversity.html. Accessed December 4, 2018.
5. Figure 18. Percentage of all active physicians by race/ethnicity. AAMC; 2018. Available from: https://www.aamc.org/data-reports /workforce/interactive-data/figure-18-percentage-all-active-physicians-race-ethnicity-2018. Accessed February 8, 2021.
6. U.S. Census Bureau. United States Geography Profile. U.S. Census Bureau; 2019. Available from https://data.census.gov/cedsci/pro file?q=United%20States%26g%3D0100000US&g=0100000US. Accessed July 30, 2021.
11. Smedley BD, Stith AY, Colburn L, Evans CH. Rationed to Students of Color in America: Inequality in Teaching and Schooling: How Opportunity is Rationed to Students of Color in America. National Academies Press Book: The Right thing to do, the Smart thing to do; 2001.

12. Why segregation matters: poverty and educational inequality — the civil rights project at UCLA. Available from: https://www.civilrightproject.ucla.edu/research/k-12-education/integration-and-diversity/why-segregation-matters-poverty-and-educational-inequality. Accessed May 15, 2021.

13. Rury JL. Education and Social Change: Contours in the History of American Schooling. Routledge; 2012.

14. Estrada M, Hernandez PR, Schultz PW, Herrera J. A longitudinal study of how quality mentorship and research experience integrate underrepresented minorities into STEM careers. CBE Life Sci Educ. 2018;17(1):ar9. doi:10.1187/cbe.17-04-0066

15. Orom H, Semalulu T, Underwood W. The social and learning environments experienced by underrepresented minority medical students: a narrative review. Acad Med. 2013;88(11):1765–1777. doi:10.1097/ACM.0b013e3182a7a3af

16. Summer mentorship program. University of Pennsylvania. Available from: https://www.vpul.upenn.edu/eap/smp/. Accessed December 17, 2018.

17. Kirkpatrick DL. Evaluation, training programs: the four levels, San Francisco: Berrett-Koehler publishers. Learning transfer. Int J Train Dev. 1994;6(1):36–48.

18. Thomas PA, Kern DE, Hughes MT, Chen BY. Curriculum Development for Medical Education: A Six-Step Approach. JHU Press; 2016.

19. Center EL. Money Matters in Education Justice: Addressing Racial and Class Inequities in Pennsylvania’s School Funding System. 2017.

20. Wiley A, Wyatt J, Camara WJ. The development of a multidimensional college readiness index. Research report 2010-3. College Board. 2011. Available from: https://files.eric.ed.gov/fulltext/ED563050.pdf. Accessed July 30, 2021.

21. Smith SG, Nsiah-Kumi PA, Jones PR, Pamies RJ. Pipeline programs in the health professions, part 1: preserving diversity and reducing health disparities. J Natl Med Assoc. 2009;101(9):836–840, 845. doi:10.1016/S0027-9684(15)31030-0

22. Winkleby MA. The stanford medical youth science program: 18 years of a biomedical program for low-income high school students. Acad Med. 2007;82(2):139–145. doi:10.1097/AM.0b013e31802d8de6

23. Winkleby MA, Ned J, Ahn D, Koehler A, Kennedy JD. Increasing diversity in science and health professions: a 21-Year Longitudinal Study documenting college and career success. J Sci Educ Technol. 2009;18(6):535–545. doi:10.1007/s10956-009-9168-0

24. Bauman LJ, Silver EJ, Stein REK. Cumulative social disadvantage and child health. Pediatrics. 2006;117(4):1321–1328. doi:10.1542/peds.2005-1647

25. Nurius PS, Prince DM, Rocha A. Cumulative disadvantage and youth well-being: a multi-domain examination with life course implications. Child Adol Soc Work J. 2015;32(6):567–576. doi:10.1007/s10560-015-0396-2

26. Dalton B; U.S. News & World Report, Research Triangle Institute (RTI) International. Best U.S High Schools - U.S News & World Report. U.S News & World Report; 2018.

27. Smidt A, Balandin S, Sigafoos J, Reed VA. The Kirkpatrick Report (RTI) International. doi:10.1007/s10560-015-0396-2

28. Huber LP, Johnson RN, Kohli R. Naming racism: a conceptual look at internalized racism in U.S. schools. Chicano Latino L Rev. 2006;26(1):183. doi:10.5070/C7261021172

29. Gonzalez D. Internalized Racism: Biases Children and Adults Hold. The Cupola: Scholarship at Gettysburg College; 2018.

30. Putnam RD. Our Kids: The American Dream in Crisis. Simon and Schuster; 2015. Illustrated.