Journal of Systemics, Cybernetics and Informatics (2023) 21(2), 13-20
https://doi.org/10.54808/JSCI.21.02.13

Teaching Health Informatics in Middle School: Experience from an NIH AIM-AHEAD pilot

Gregory TARDIEU
Alexandria City Public Schools
Alexandria, VA 22314, USA

Senait TEKLE
Biomedical Informatics Center, The George Washington University
Washington, D.C. 20037, USA

Linda ZANIN
Alexandria City Public Schools
Alexandria, VA 22314, USA

Teri L. CAPSHAW
Alexandria City Public Schools
Alexandria, VA 22314, USA

Alexander LIBIN
Georgetown-Howard Universities Center for Clinical and Translational Science
Washington, D.C. 20057, USA

Qing ZENG-TREITLER*
Biomedical Informatics Center, The George Washington University
Washington, D.C. 20037, USA

**Correspondence: Qing Zeng-Treitler, PhD, Professor & Director, The George Washington University, Biomedical Informatics Center, 2600 Virginia Ave NW, Suite 300, Washington, DC 20037, zengq@gwu.edu, 202-994-3143*

ABSTRACT [1]

Health informatics as a topic is rarely introduced to middle school students due to their age and insufficient background knowledge in computing and healthcare. At the same time, it has been observed that many students have lost interest in science and technology when they reach high school. Funded by the NIH AIM-AHEAD initiative, we embarked on a project to create a health informatics after-school initiative focused on AI. We recognize that students who identify as racial or ethnic minorities are less likely to be introduced to and less prepared for a range of STEM-H careers. Limited diversity in the life sciences and health professions has significant consequences for access to healthcare services. Preparing diverse students for the future "digitally proficient" healthcare workforce is fundamental to addressing health disparities, increasing cross-cultural communication, and positively impacting health equity. We acknowledge that students are more likely to thrive academically in areas of STEM- H when they have access to instructors from diverse races, ethnicities, and backgrounds who understand their experiences and perspectives.

Keywords: Health informatics education, middle schoolers, AI, healthcare, disparity, digitally proficient.

1. INTRODUCTION

The overall approach of our project is focused on transforming student lives and increasing diversity in the digitally proficient healthcare workforce. According to one study, women make up 15% of AI researchers at Facebook and 10% at Google [1]. It has also been reported that less than 3% of workers at technology companies are Black [2]. The healthcare workforce also faces challenges: A cross-sectional study using data from the 2019 American Community Survey (ACS) to compare the

[1] The authors wish to thank and acknowledge Dr. Stuart J. Nelson for his peer-editing of this paper.

diversity of 10 healthcare occupations (advanced practice registered nurses, dentists, occupational therapists, pharmacists, physical therapists, physician assistants, physicians, registered nurses, respiratory therapists, and speech-language pathologists) reported that Blacks and Hispanics are underrepresented in these occupations that have higher education requirements [3]. On the surface, Asians are overrepresented in many technology and healthcare fields. Studies, however, revealed strong discrimination against Asian workers. For example, among all races and ethnicities, Asians are most likely to be hired by a technology company and are the least likely to be promoted to senior leadership positions and receive lower performance ratings [4].

Diversity also goes beyond race and gender [5]. Young people from immigrant backgrounds, speaking English as a second language, growing up in low-income neighborhoods, etc., are disadvantaged when it comes to early exposure to advanced technology fields such as AI. Our project aims to provide a solution to this problem by bringing AI in a medical training program to a diverse group of middle school students in the form of AIM-AHEAD after-school clubs using a state-of-the-art mobile classroom.

Theorem/Proof/Lemma: The pilot objective is promoting diversity, equity, and inclusion in healthcare and informatics education. Students in US schools are becoming increasingly racially and ethnically diverse as a group; teachers and school leaders, for the most part, do not reflect this diversity [6]. To address the unique needs of students, our project is guided by the Culturally Responsive Teaching (CRT) model, an approach to teaching based on the idea that students learn best when their cultural backgrounds and experiences are taken into account. It is a student-centered approach focused on building relationships, creating a safe and supportive learning environment, and engaging students in meaningful learning experiences. CRT also emphasizes the importance of intrinsic motivation, the idea that students are more likely to be successful when they are motivated by their own interests and curiosity [7].

Our overarching hypothesis is that personalized and representative informatics will increase children's knowledge, interest, and future participation in a STEM-H workforce (Figure 1).

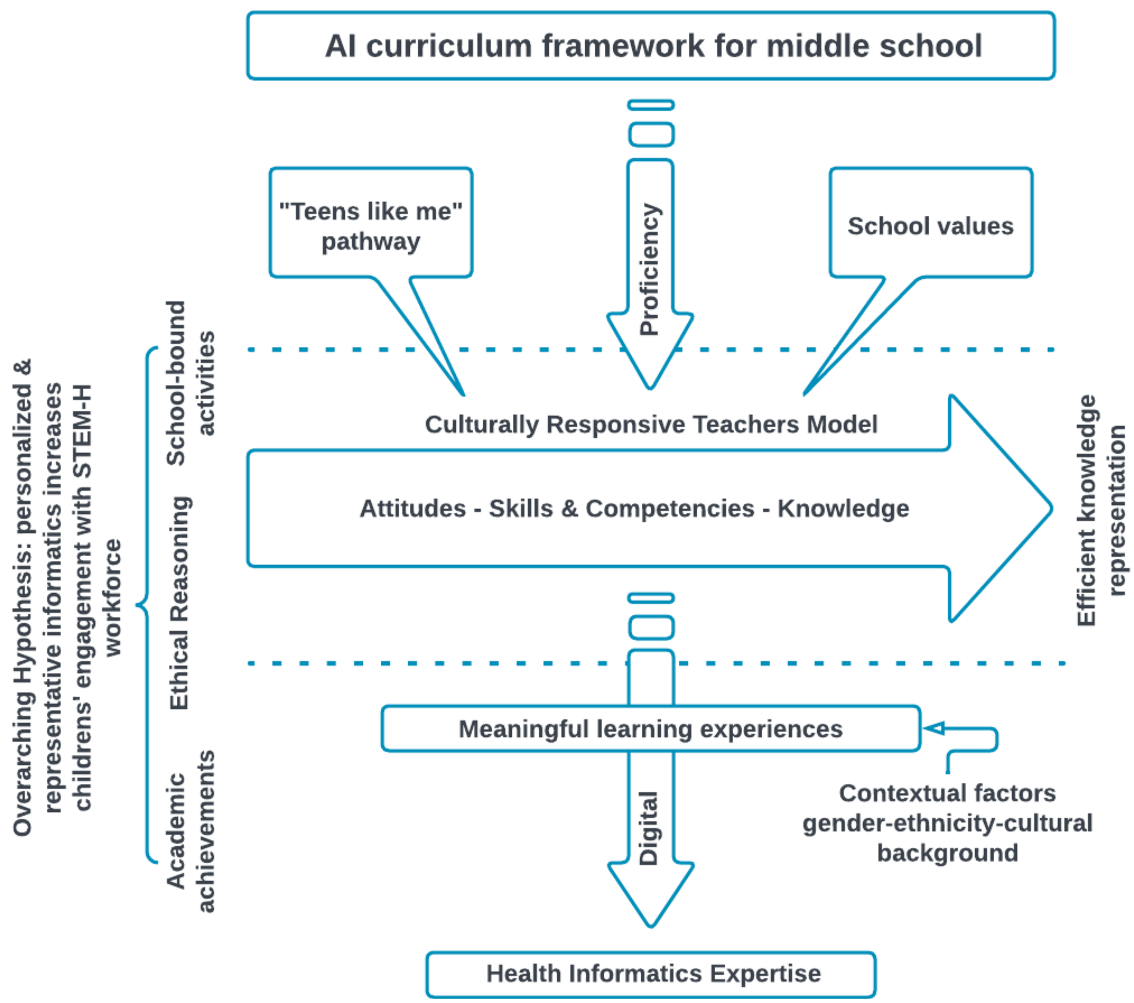


Figure 1: Conceptual framework for diversified AI-focused medical training program for middle school students.

The pilot target population is Alexandria City Public Schools (ACPS) in Virginia, serving more than 15,700 students who hail from more than 119 countries and speak 124 languages. ACPS has 18 schools, including two middle schools (6-8 grade), two K- 8 schools, one pre-K school, and the internationally recognized Alexandria City High School. Middle school challenges include providing students access to advanced technology resources, such as AI, and providing early exposure to technology to enhance learning experiences. Introducing this technology allows students the ability to apply, evaluate, and improve their skills in a controlled learning environment, without the risk of real-world consequences. Created learning scenarios are repeatable to allow students to improve performance through carefully analyzing mistakes and being free to make them without real-world risks. The immersive nature of our approach adds an experiential and spatial component that helps learners to retain information.

The after-school STEM-H programming is an effort to provide students with hands-on collaborative interactive activities to learn the basics of health informatics, computing, and AI.

2. BIG DATA APPROACH TO HEALTH INFORMATICS EDUCATION IN CHILDREN AND ADOLESCENTS

The AIM-AHEAD after-school club takes place in the Immersive Learning Center (Figures 2 and 3) and has 3 key components:

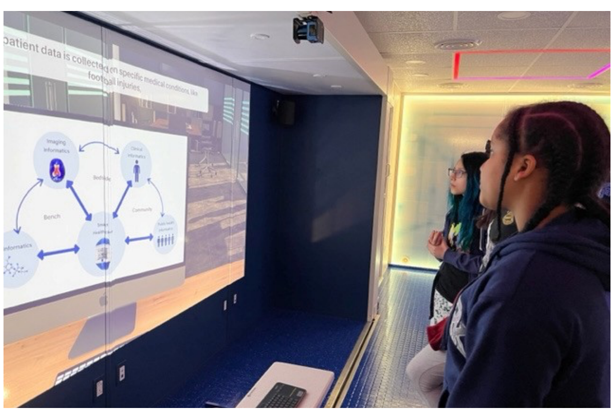


Figure 2: Sixth- and seventh-grade students are learning about health informatics in the Immersive Learning Center.

Figure 3: The Immersive Learning Center.

I. Mini "Tech Talks" on AI in medicine

We recruited a wide range of speakers from pioneers in AI in medicine and health informatics experts who are developing cutting-edge solutions in academia and industry to middle- high- school/college students who are learning about AI. We asked each speaker to record a 15–20-minute talk to inspire students with a focus on AI applications rather than the technical details of methodology. By providing concrete and compelling examples of AI applications that have an impact on health and healthcare, we hope to introduce the field of AI and health informatics as both exciting and important.

One example of a "Tech Talk" is a lesson by Marius George Linguraru, DPhil, MA, MSc, Children's National Hospital/The George Washington University, on the uses of advanced AI in facial recognition for early detection of Down syndrome in diverse populations and in developing countries. Image recognition is a powerful technology with significant implications in everyday life. AI image recognition picks up hidden patterns, even those not apparent to the human eye (Figure 4). While the technology is sophisticated, even middle school students can appreciate the disparity problem and the benefit of image recognition solutions.

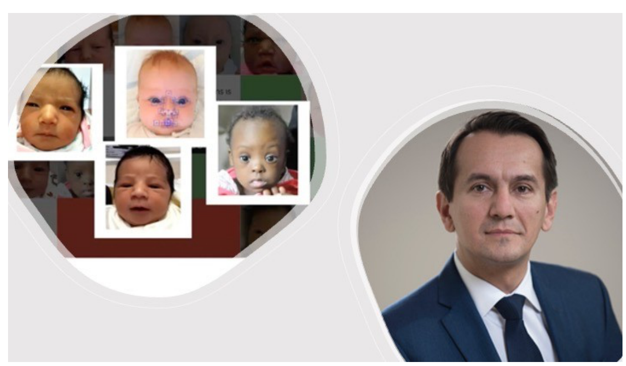


Figure 4: Dr. Marius George Linguraru of Children's National Hospital/The George Washington University gave a "Tech Talk" on "AI for Children with Rare Diseases."

II. Hands-on programming using Lego Robots

Since most of our ACPS middle school population is students from diverse backgrounds with no prior exposure to programming, we decided to teach programming using Lego Inventor Robot in an after-school club for 6-8th grade students. Considering the limited classroom time, we recruited a team of middle and high school students to pre-assemble the robots, allowing the after-school participants to jump into the programming components directly. There were four after-school 6-8th grade student clubs, with 12-26 students participating in each club. (Figures 5 and 6).

The first part of the programming lessons is focused on the basic commands, control structures, use of variables, and sensors. The second part dives into machine learning. The lessons give students an opportunity to experiment with learning from audio and visual data. The students are also able to use the trained model to make the robots more "intelligent."

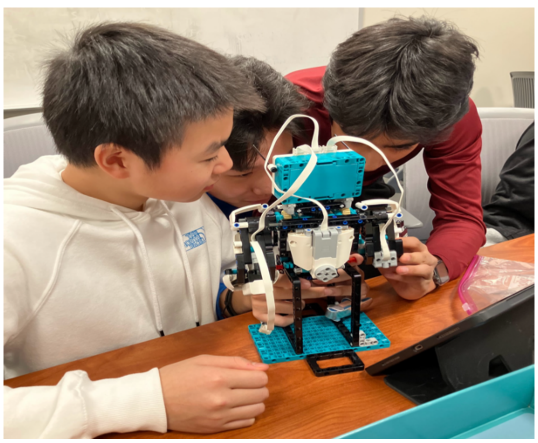


Figure 5: Students building and programming Lego robots.

Figure 6: Students in middle school after-school club using code to control the Lego robots.

III. Experiment with big, real word healthcare data in a "Teens Like Me" app

To introduce the middle school students to big data in healthcare, we curated two adolescent-focused datasets using the Cerner Real World Data (RWD) and the National Health and Nutrition Examination Survey (NHANES) (Table 1). We have experience working with both datasets. RWD is a huge, national electronic health record (EHR) dataset from more than 100 healthcare facilities across the US [8]. As its name suggests, it contains valuable real world clinical data that can be used to demonstrate the power of AI. NHANES, on the other hand, is a yearly survey conducted by the US Centers for Disease Control (CDC). NHANES collects cross-sectional data from a nationally representative sample, while over-samples certain minority groups [9]. We identified subjects in their teens from both datasets, to allow the "like me" analyses that are more relatable for the students. One analysis we have conducted is to characterize the nutrient patterns of food items in adolescent diets using data from a novel citizen science project we have previously carried out with the ACPS and the NHANES. In this analysis, we used two different unsupervised learning methods to discover the common nutrient profiles of food items consumed by teenagers and visualized the deficiencies and excesses according to the FDA-recommended daily values (DV) [10] (Figure 7).

Table 1: The demographic characteristics of the CDW and NHANES datasets.

	Cerner CDW (N=1,450,782)	NHANES (N=7,110)
	Mean/N	**Mean/N**
Age		
	14.74	15.4
Gender		
Female	673,417	3,460
Male	776,482	3,650
Unknown	883	0
Race		
White American	904,844	2,034
African American	184,865	1,752
Asian	28,643	N/A
Others	0	965
Unknown	196,032	0
Ethnicity		
Non-Hispanics	967,974	4,751
Hispanics	308,496	2,359

% of DV

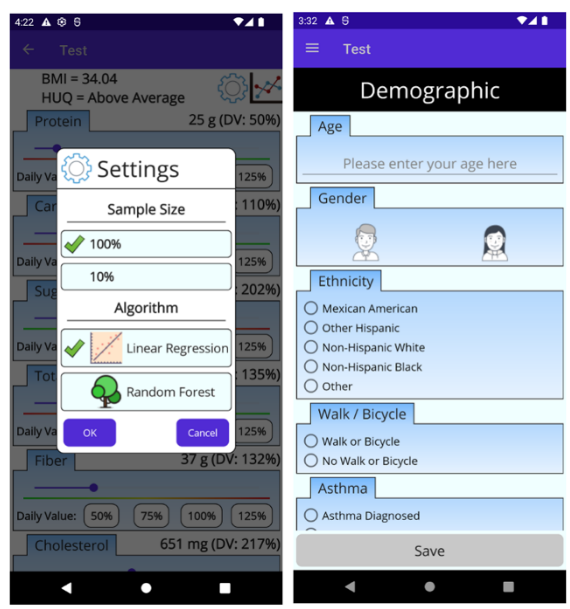
K-Means Cluster Size	Energy	Protein	Total Lipid	Carbohydrates	Dietary Fiber	Total Sugars	Calcium	Iron	Potassium	Sodium	Cholesterol
28%	3%	2%	0%	4%	0%	5%	0%	0%	0%	1%	0%
16%	3%	5%	1%	4%	4%	8%	2%	2%	4%	2%	0%
14%	12%	21%	14%	8%	5%	4%	5%	7%	0%	20%	9%
14%	18%	13%	2%	23%	7%	0%	0%	13%	0%	16%	0%
11%	6%	40%	5%	0%	0%	0%	0%	4%	0%	8%	20%
7%	22%	32%	33%	2%	0%	0%	0%	7%	0%	34%	13%
5%	20%	23%	12%	26%	32%	30%	5%	17%	5%	11%	0%
4%	18%	8%	10%	25%	5%	79%	0%	4%	0%	9%	0%
<1%	14%	25%	3%	7%	0%	0%	0%	1%	18%	257%	0%

GMM Cluster Size	Energy	Protein	Total Lipid	Carbohydrates	Dietary Fiber	Total Sugars	Calcium	Iron	Potassium	Sodium	Cholesterol
32%	2%	2%	0%	4%	0%	9%	1%	0%	2%	1%	0%
21%	11%	21%	9%	7%	4%	3%	2%	6%	0%	19%	12%
11%	9%	6%	11%	10%	9%	3%	2%	5%	0%	8%	0%
10%	18%	13%	0%	28%	0%	0%	0%	18%	2%	0%	0%
8%	10%	29%	19%	1%	0%	0%	0%	3%	0%	16%	0%
8%	10%	23%	14%	1%	0%	2%	2%	2%	3%	14%	28%
6%	14%	7%	6%	23%	6%	48%	0%	3%	1%	12%	0%
3%	20%	15%	18%	21%	19%	49%	5%	14%	4%	16%	0%
1%	10%	16%	12%	9%	5%	24%	0%	0%	0%	31%	10%

Figure 7: This figure shows the nutrient profiles of 9 clusters of food items generated using two different methods (K-Means and GMM). It suggests that the largest clusters are low in many nutrients, with sugar and protein being the exceptions when compared to the DV.

Given that most students do not have the level of programming skills to train realistic machine learning models, we've created a "Teens Like Me" app to allow the students to experiment with risk prediction models. The user-friendly app serves as an alternative to more complex programming languages such as R or Python. It allows students to dive into unsupervised learning analyses of the nutrients' datasets and observe the effect of using different clustering methods and varying the number of clusters.

Developed by the George Washington University Biomedical Informatics Center, this app offers students a practical opportunity to gain hands-on experience with predictive modeling techniques (Figure 8). In its current version, the app enables students to experiment with fundamental machine learning models, such as linear regression, logistic regression, and random forest models. These models are applied to predict the user's Body Mass Index (BMI) and their overall satisfaction with their health. Furthermore, students have the flexibility to customize these models for different sample sizes. This hands-on experience reinforces the essential concept model performance generally improves with larger datasets, highlighting the significance of data quantity in modeling.



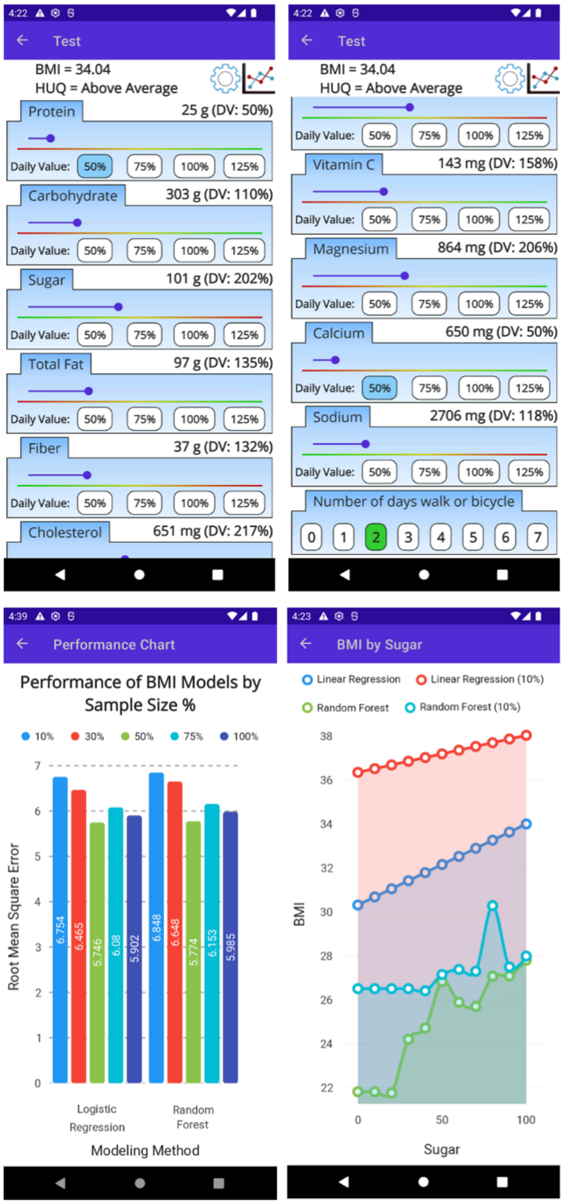


Figure 8: Teens Like Me app images

Evaluation results: Both students and teachers were asked to participate in a brief survey in order to gain valuable insights into their overall experiences of the after-school club. A sample of 29 students and 9 teachers provided responses, contributing to the analysis and findings. The survey results for the students indicate a high level of satisfaction with the activities with 94% expressing satisfaction and 100% expressing a likelihood to revisit the club. In terms of overall satisfaction, 38% rated it as excellent, and 42% as very good. Among the nine teacher respondents, 100% expressed satisfaction with the activities, with 75% very likely to revisit. 56% rated overall satisfaction as excellent, and 44% as very good. All teachers (100%) felt strongly comfortable with technology, and all agreed that technology improved their interest in activities and educational content. Valuable feedback from teachers highlighted the need for more time and increased availability of technology resources. Overall, the feedback affirms the success of the after-school club in engaging and enhancing the educational experience for students.

3. DISCUSSION POINT: PM METHODOLOGY OF HEALTH INFORMATICS EDUCATION FOR MIDDLE SCHOOL LEARNERS

The development of AI-based technologies calls for a new social contract between the computer sciences (CS) community, encompassing information technologies (ITs), information and communication technologists (ICTs), and programmers – often referred to as IT specialists. This agreement must also involve digital society stakeholders to be successful with adolescents, most of whom have been immersed in computer-mediated interactions since early childhood. Traditional education is transformed by new learning practices focusing not only on knowledge acquisition, but efficient knowledge implementation expressed through learning and professional competencies. In our pilot, 6-8th grade students from diverse backgrounds and knowledge levels joined the after-school club to learn more about health informatics (Figure 9).

An important part of the after-school club was to give students the chance to have hands-on learning experiences with health informatics. A unique combination of this technology promotes interactivity, and engagement, allowing students to enhance critical thinking and communication skills, team building and collaboration, and technical knowledge. The cumulative learning outcomes can be defined in our conceptual framework as generalized intellectual capital in education (InC^{ED}), expressed through an algorithm comprising four different factors: ethical reasoning, academic achievements, coping with complex school-bound activities and related everyday life situations, and, most importantly, digital proficiency. It is hypothesized that InC^{ED} has a direct impact on the learning competencies of students, warranting future successful management of professional and life challenges [11].

Figure 9: 8th grade students participating in the after-school program in the Immersive Learning Center.

Our conceptual framework recognizes certain limitations of the approach undertaken. Teaching and learning health informatics is challenging in the middle school 6-8th grades. Both teachers and students require a great deal of support.

We recognize that youth who identify as racial or ethnic minorities are less likely to be exposed to and less prepared for a range of STEM-H careers. Limited diversity in the life sciences and health professions has significant consequences for access to healthcare services. Preparing diverse students for the future digitally proficient healthcare workforce is fundamental to addressing health disparities, increasing cross-cultural communication, and positively impacting health equity. We acknowledge that students are more likely to thrive academically in areas of STEM-H when they have access to instructors from diverse races, ethnicities, and backgrounds who understand their experiences and perspectives.

The proposed AI curriculum framework for middle schoolers addresses the existing gap in teaching digital core competencies as an integrated discipline for middle school students. Education through children-centered health informatics focuses on providing a methodology and metrics for aligning AI-based technologies with school values and enhancing primary education with research and development components across traditional and innovative teaching disciplines.

The process of piloting and expanding initiatives like this is challenging. However, educating diverse students for careers in health informatics is fundamental to addressing disparities in healthcare, increasing cross-cultural communication, and positively impacting equity in health career pathways. Ultimately, to achieve change and to prepare a diverse healthcare workforce, ongoing commitments from secondary and higher education, as well as community collaborations are required. We will continue to work closely with teachers, leadership, and community partners to ensure that our middle school after-school programs have the support needed to take root and thrive. We are also planning to create an online curriculum that can be shared with other school districts.

4. ACKNOWLEDGMENT

We would like to thank Alexandria City Public Schools, The George Washington University School of Medicine and Health Sciences Immersive Learning Center, Dr. Anil Shanker, Senior Vice President for Research and Innovation Professor of Biochemistry, Cancer Biology, Neuroscience, and Pharmacology, Meharry Medical College, and the citizen scientists for their support and participation in this project.

AIM AHEAD Acknowledgement. Findings reported in this publication were supported by the Artificial Intelligence and Machine Learning Consortium funded by the National Institutes of Health under Award Number 1OT2OD032581. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. 2GHUCCTS Acknowledgement. Findings reported in this publication were partially supported by the National Center For Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR001409. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

5. REFERENCES

[1] S. M. West, M. Whittaker, and K. Crawford, “Discriminating systems: Gender, Race and Power in AI,” AI Now Institute, Apr. 2019.

[2] A. Brown, “Fixing AI’s diversity crisis through Edtech: An in-depth interview with Tony Effik, co-founder of The Black and Brilliant Advocacy Network,” Forbes. Accessed: Jul. 24, 2022. [Online]. Available: https://www.forbes.com/sites/anniebrown/2021/07/23/fixing-ais-diversity-crisis-through-edtech-an-in-depth-interview-with-tony-effik-co-founder-of-the-black-and-brilliant-advocacy-network/

[3] E. Salsberg *et al.*, “Estimation and comparison of current and future racial/ethnic representation in the US health care workforce,” *JAMA Network Open*, vol. 4, no. 3, pp. e213789–e213789, Mar. 2021, doi: 10.1001/jamanetworkopen.2021.3789.

[4] Centers for Disease Control and Prevention, “Framework for program evaluation.” Accessed: Jul. 24, 2022. [Online]. Available: https://www.cdc.gov/evaluation/framework/index.htm

[5] T. Faist, “Diversity – a new mode of incorporation?,” *Ethnic and Racial Studies*, vol. 32, no. 1, pp. 171–190, Jan. 2009, doi: 10.1080/01419870802483650.

[6] K. Schaeffer, “America’s public school teachers are far less racially and ethnically diverse than their students,” Pew Research Center. Accessed: Feb. 28, 2023. [Online]. Available: https://www.pewresearch.org/fact-tank/2021/12/10/americas-public-school-teachers-are-far-less-racially-and-ethnically-diverse-than-their-students/

[7] G. Gay and J. A. Banks, *Culturally responsive teaching: Theory, research, and practice*, 2nd edition. New York: Teachers College Press, 2010.

[8] L. Ehwerhemuepha *et al.*, “Cerner real-world data (CRWD) - A de-identified multicenter electronic health records database,” *Data Brief*, vol. 42, p. 108120, Jun. 2022, doi: 10.1016/j.dib.2022.108120.

[9] L. J. Akinbami *et al.*, “National Health and Nutrition Examination Survey, 2017-March 2020 Prepandemic File: Sample Design, Estimation, and Analytic Guidelines,” *Vital Health Stat 1*, no. 190, pp. 1–36, May 2022.
[10] J. T. Treitler *et al.*, “Characterizing nutrient patterns of food items in adolescent diet using data from a novel citizen science project and the US National Health and Nutrition Examination Survey (NHANES),” *Frontiers in Nutrition*, 2023.
[11] E. Libin, “Coping intelligence: Efficient life stress management,” *Frontiers in Psychology*, vol. 8, p. 302, 2017.