Journal of the American Medical Informatics Association, 27(11), 2020, 1641–1647
doi: 10.1093/jamia/ocaa206
Advance Access Publication Date: 14 October 2020
Research and Applications

Research and Applications

Underrepresented racial minorities in biomedical informatics doctoral programs: graduation trends and academic placement (2002–2017)

Kevin Wiley,[1] Brian E. Dixon ,[2] Shaun J. Grannis ,[3] and Nir Menachemi[2]

[1]Department of Health Policy and Management, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, Indiana, USA, [2]Regenstrief Institute, Inc., Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, Indiana, USA, and [3]Regenstrief Institute, Inc., School of Medicine, Indiana University, Indianapolis, Indiana, USA

Corresponding Author: Nir Menachemi PhD, MPH, Regenstrief Institute, Inc., Richard M. Fairbanks School of Public Health, Indiana University, Health Sciences Building (RG), 1050 Wishard Blvd. 5th Floor, Indianapolis, IN 46202-2872, USA (nirmena@iu.edu)

Received 2 June 2020; Revised 31 July 2020; Editorial Decision 3 August 2020; Accepted 10 August 2020

ABSTRACT

Objective: Biomedical informatics attracts few underrepresented racial minorities (URMs) into PhD programs. We examine graduation trends from 2002 to 2017 to determine how URM representation has changed over time. We also examine academic job placements by race and identify individual and institutional characteristics associated with URM graduates being successfully placed in academic jobs.

Materials and Methods: We analyze a near census of all research doctoral graduates from US-accredited institutions, surveyed at graduation by the National Science Foundation Survey of Earned Doctorates. Graduates of biomedical informatics-related programs were identified using self-reported primary and secondary disciplines. Data are analyzed using bivariate and multivariable logistic regressions.

Results: During the study period, 2426 individuals earned doctoral degrees in biomedical informatics-related disciplines. URM students comprised nearly 12% of graduates, and this proportion did not change over time (2002–2017). URMs included Hispanic (5.7%), Black (3.2%), and others, including multi-racial and indigenous American populations (2.8%). Overall, 82.3% of all graduates accepted academic positions at the time of graduation with significantly more Hispanic graduates electing to go into academia (89.2%; $P < .001$). URM graduates were more likely to be single (OR = 1.38; $P < .05$), have a dependent (1.95; $P < .01$), and not receive full tuition remission (OR = 1.37; $P = .05$) as a student. URM graduates accepting an academic position were less likely to be a graduate of a private institution (OR = 0.70; $P < .05$).

Discussion and Conclusion: The proportion of URM candidates among biomedical informatics doctoral graduates has not increased over time and remains low. In order to improve URM recruitment and retention within academia, leaders in biomedical informatics should replicate strategies used to improve URM graduation rates in other fields.

Key words: informatics, minority groups, education, doctoral training, academic placement

INTRODUCTION

The US is committed to developing trainees, especially those of underrepresented racial minorities (URMs), in science, technology, engineering, and mathematics (STEM) fields at all education levels.[1–3] Biomedical informatics, an interdisciplinary STEM field that develops and uses computational approaches to collect and analyze biological and health data, has struggled (similar to other STEM fields) to attract URMs, especially at the PhD level.[4–6] Whereas URMs make up approximately 30% of the US population, they comprised just 9% of science and engineering doctoral graduates in 2017.[7,8] URM representation in academia is especially important at a time when informatics tools, such as artificial intelligence and machine learning approaches, are being criticized for amplifying and codifying racial bias into mainstream applications.[9–11] Further, the shortage of URM individuals in academic positions makes it challenging to attract URM trainees at all levels given that trainees cite having relatable mentors as key to their educational decisions.[12,13]

URM trainees, especially in STEM programs, face challenges at every stage of the educational and professional "pipeline"; a metaphor that describes the pathway from PhD programs to entry into academic careers.[12,14–16] URM trainees are more likely to take longer to complete their training and experience higher program attrition rates due to educational disadvantages and socioeconomic status.[15–17] Thus, these individuals in STEM doctoral programs struggle to reach graduation which limits the potential rate of transition into academic positions,[6,15] even though numerous fellowships and other programs exist to support them within STEM PhD programs.[18] These resources, albeit limited, have a positive impact on URM doctoral students within biomedical sciences and biomedical informatics-related fields. Specifically, such trainees report increased scientific self-efficacy, positive self-identity, higher degree completion rates, and improved perceptions regarding their suitability for positions in academia.[19,20] Nevertheless, little is known about overall graduation trends of URM students within biomedical informatics research doctoral programs and what characteristics are associated with their academic job placement compared to their peers.

The purpose of the current study is to examine trends over time in URM graduates from biomedical informatics doctoral programs (eg, PhD, ScD, and other research degrees) in the US and determine the characteristics of URM doctoral graduates who transition into academic positions. Doing so will inform whether the number of URM students as a proportion of those completing research doctoral degrees in biomedical informatics has changed over time and might suggest the type of action needed to increase URMs in doctoral training. Further, our study yields insights that could inform how to improve the rates at which URM doctoral graduates can be successfully recruited into academic positions. As such, we believe our findings will be of interest to decision-makers interested in increasing the number of URMs in biomedical informatics and other STEM fields. Moreover, our findings may be of interest to those focused on assuring that the conditions where informatics tools are developed by researchers are more likely to best serve the population at large.

MATERIALS AND METHODS

We utilize repeated cross-sectional data from a near census of all research doctoral recipients from US accredited doctoral-granting institutions from 2002 to 2017. Specifically, we examine temporal trends among PhD graduates who trained in a biomedical informatics-related training program. In particular, we determine the number and proportion of all graduates who are URM and explore characteristics of URM graduates that accept academic positions.

Our data are derived from the restricted files of the National Science Foundation (NSF) Survey of Earned Doctorates (SED). The SED is an annual survey of approximately 54 000 research doctoral graduates in all fields of study and is administered on behalf of the NSF by universities. Individual universities decide what doctoral degrees at their institution are considered research-oriented which typically requires a dissertation or thesis where new contributions to knowledge in the field are expected. Thus, whereas doctorates of philosophy (PhD) and doctorates of science (eg, ScD or DSc) are almost always included; clinical (eg, MD, DDS, PsyD) and other doctorates (eg, JD) have historically been excluded from the SED. Doctoral recipients that are eligible for the SED receive the survey at approximately the time of their graduation. For example, at Indiana University, doctoral graduates complete the survey when they deposit their final, approved dissertation with the university library. The survey includes questions that ask about their field of study, experience in the doctoral program, previous predoctoral training, student debt, demographics, and career plans including whether they have secured a job, and if so, what type, at the time of the survey. The SED dataset also includes institutional characteristics provided by each university.

SED respondents are asked to select from among approximately 300 three-digit codes to indicate the primary and secondary (if applicable) fields of their study. Given the diversity in training represented among those in the field of biomedical informatics and the lack of a discrete code representing this broad field, we identified graduates using 3 different approaches from the SED, including: (1) individuals whose primary field was code "102, bioinformatics," regardless of their secondary field; (2) individuals whose primary field was any computer science field and whose secondary field was any health-related field (see Supplementary Table for SED codes); or (3) individuals whose primary selection was a health-related field and whose secondary selection was from any computer science field. We collectively refer to individuals identified using these 3 approaches as 'biomedical informatics-related' graduates. The Institutional Review Board (IRB) at our university classified this study as exempt from human subject oversight.

Our primary dependent variable was whether the biomedical informatics-related graduate was a member of a URM group as considered by the NSF based on the US Office of Budget and Management definition.[21] URMs include Black or African American, Hispanic, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and those with 2 or more races that include at least 1 underrepresented group. Because of small sample sizes, we created and combined 3 categories of URMs, including: (1) Black or African American, (2) Hispanic, and (3) all other URMs. We coded a second dependent variable from the SED that included graduates who were from a URM race or ethnicity and who accepted a position in an academic setting. Academic employment includes US 4-year academic institutions and their affiliated research organizations, US medical schools, 2-year colleges, and academic postdoctoral fellowships.[22] Independent variables in our study included characteristics of graduates and academic institutions from the SED. Graduate characteristics include gender, age group, marital status, number of dependents, citizenship status, whether the trainee received full tuition remission, and year of graduation. Institutional characteristics include whether the university was public or private, whether its tax status was nonprofit or for profit, and the Carnegie Classification of Institutions of Higher Education classification. Carnegie classification identifies the research intensiveness of a university by assigning institutions of higher education into 4 categories: highest or very high research activity (R1), higher research activity (R2), moderate research activity (R3), or other.[23]

Descriptive statistics were generated to examine temporal trends in the total number of URM graduates between the years 2002 and 2017. Independent variables also characterized URM graduates in all 3 biomedical informatics-related training categories. In our first bivariate analyses, we utilized Pearson chi-square tests to assess the relationship between URMs and 3 biomedical informatics-related training categories and to assess the relationship between graduate characteristics and URM status. In addition, Pearson chi-square tests were used to explore the associations between all race categories and those who reported accepting an academic position. Multivariable logistic regression models were used to examine the associations between institutional and graduate characteristics (as described above) and our 2 dependent variables. Because the SED began recording bioinformatics as a discrete primary field of study in 2007, as a sensitivity analysis, we re-estimated our multivariable regression models including only years 2007–2017 to examine whether any of our findings were sensitive to this change. Analyses were conducted using SAS version 9.4 statistical software.

RESULTS

Our final sample included 2426 graduates who completed doctoral degrees in 1 of 3 biomedical informatics-related training categories between 2002 and 2017. Demographic characteristics of the sample are displayed in Table 1. A majority of graduates were male (56.6%), US citizens (58%), were married or in a marriage-like relationship (60.8%), and had received their doctoral degree from an R1 institution (79.3%). For the entire sample, the most frequent race among respondents was White (48.1%) followed by Asian (40.2%). URM respondents made up 11.7% (284 of 2426) of all doctoral graduates. In the total sample, URMs included 139 Hispanic graduates (5.7%), 77 Black graduates (3.2%), and 68 graduates from other URM groups (2.8%).

Over time, the total number of biomedical informatics graduates increased from 54 in 2002 to 342 in 2017 (see Figure 1). The proportion of White doctoral graduates trended downward from 61.8% in 2002 to 45.3% in 2017 (see Figure 2). Asian doctoral graduates increased from 23.6% of all graduates in 2002 to 39.5% in 2017. In Figure 2, we depict the sum of all URMs (Black/African-American, Hispanic, and URM Other) as a proportion of all doctoral graduates by year. Sample sizes of URM graduates, by year, ranged between 10 and 15 percent of all biomedical informatics-related doctoral recipients in most years. The variability in the trend line from 2002 to 2006 in Figure 2 is possibly due to these small sample sizes in each race category during these years.

When examining respondents by the 3 categories used to identify biomedical informatics-related doctoral recipients (see Table 2), we observed differences by race. Given the proportion of each race in our cohort, Whites were most represented among those with a computer science-related primary field while Black URMs were most represented among those with a health-related primary field. Further, whereas Asians were most represented among those selecting bioinformatics as their primary field, Hispanics were least represented among this category (all *P* values < .01).

Table 1. Characteristics of biomedical informatics-related graduates, N = 2426 (2002–2017)

Variable	Frequency (%)
Race	
• White	1166 (48.1)
• Asian	976 (40.2)
• Total Underrepresented racial minority (URM)	284 (11.7)
• Hispanic	139 (5.7)
• Black	77 (3.2)
• Other URM	68 (2.8)
Sex	
• Male	1445 (56.6)
• Female	981 (40.4)
Age groups	
• Less than or equal to 30 years	137 (5.7)
• 31-40 years	1389 (57.3)
• 41-50 years	723 (29.8)
• Greater than or equal to 51 years	177 (7.3)
Marital Status	
• Single, never married	882 (36.4)
• Married	1475 (60.8)
• Separated/Divorced	69 (2.8)
Number of dependents aged <18 years	
• 0	2236 (92.2)
• 1	121 (5)
• 2	51 (2.1)
• 3 or more	18 (0.7)
US Citizen (yes)	1408 (58)
Received full tuition remission while a doctoral student (yes)	2007 (82.7)
Biomedical informatics-related training category	
• Bioinformatics primary field	1471 (60.6)
• Computer science field is primary with secondary health field	619 (25.5)
• Health field is primary with secondary computer science field	336 (13.9)
Carnegie classification	
• Highest research activity (R1)	1924 (79.3)
• Higher research activity (R2)	248 (10.2)
• Moderate research activity (R3)	29 (1.2)
• Other	225 (9.3)
Type of university	
• Public	1643 (67.7)
• Private	783 (32.3)
Total number of graduates	
• 2002 (first year of study period)	54 (2.2)
• 2017 (last year of study period)	342 (14.1)

Overall, 1998 of 2426 of respondents (82.3%) indicated that they had accepted an academic position at the time of the survey (see Table 3). Hispanic respondents (89.2%) indicated accepting an academic position at significantly higher rates than all other race categories ($P < .001$). Among respondents with a health-related primary field, Black doctoral recipients (71.4%, $P = .086$) were marginally less likely to report accepting an academic job at the time of the survey.

Results from our regression analysis (see Table 4 first column) indicate that URM graduates were more likely to be single (OR: 1.38; $P < .05$), have a dependent (OR: 1.95; $P < .01$); and/or be a US citizen (OR: 2.17; $P < .001$). URM doctoral recipients were more likely to have trained in either computer science (OR: 1.52; $P < .01$) or health as a primary field (OR: 1.79; $P < .01$) compared to those whose primary field was bioinformatics. The proportion of URM graduates over time did not significantly change when the time trend was measured categorically (data shown) nor when the time trend was measured linearly (data not shown).

Regression results that examined factors associated with URMs who accepted an academic position are displayed in Table 4 (second column). Specifically, URMs who accepted an academic position were more likely to have at least 1 dependent (one dependent OR: 1.82; $P < .05$; 2 dependents OR: 2.21; $P < .05$), and more likely to be a US citizen (OR: 1.82; $P < .001$). URM graduates whose primary field was from either the health (OR: 1.61 $P < .05$) or computer science (OR: 1.39; $P < .05$) categories were more likely to have accepted an academic position than those with primary training in bioinformatics. Further, URMs who accepted an academic position were less likely to graduate from a private (as opposed to public) university (OR: 0.70; $P < .05$). Lastly, the proportion of URM graduates who accepted an academic position did not change over time when the temporal trend was measured categorically (data shown) nor when measured continuously (data not shown).

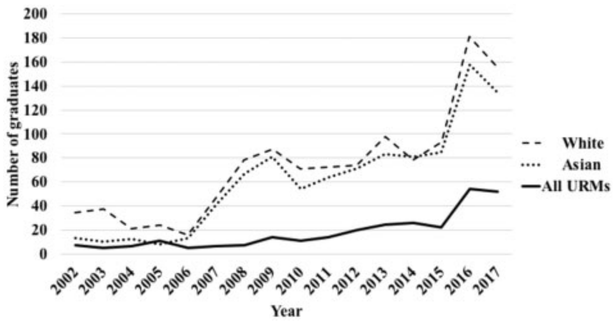

Figure 1. Total number of biomedical informatics-related graduates by race and year (2002–2017).

Note: Per the data use agreement, race categories (eg, Black/African-American, Hispanic, URM Other) are combined into a single category, All URMs, where frequencies are fewer than n = 5.

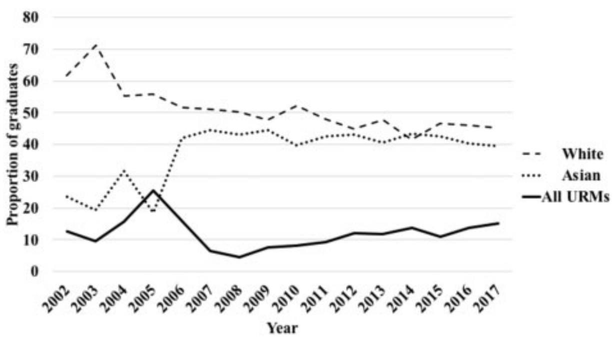

Figure 2. Proportion of biomedical informatics-related graduates by race and year (2002–2017).

Note: Per the data use agreement, race categories (eg, Black/African-American, Hispanic, URM Other) are combined into a single category, All URMs, where frequencies are fewer than n = 5.

Results from our sensitivity analysis that examined the addition of a bioinformatics primary category on the SED in 2007, indicate that the directions and magnitude of the effects among graduate and institutional characteristics were comparable to our primary analyses on the full sample (data not shown). Moreover, we observed no statistically significant interactions between doctoral graduation year and training category.

DISCUSSION

We found that URM graduates from biomedical informatics-related doctoral programs constitute an average of approximately 12% of all graduates annually and have proportionally not changed over time despite an observed yearly increase in the total number of graduates (including URMs) from 2002 to 2017. Hispanics, the largest URM group, comprised ∼6%, and Blacks comprised ∼3%, of all graduates between 2002 and 2017. The annual URM proportion of doctoral graduates in our study is similar to trends observed among doctoral trainees in both public health and other STEM fields.[24,25] Past research has attributed underrepresentation of minorities in doctoral training to deficiencies in preparation and a lack of relatable mentors in undergraduate and master-level training programs.[12,13,15] Despite philanthropic and federal investments in URM STEM education,[1–3,18] there remain gaps in student preparation which may have contributed to barriers in pursuing biomedical informatics-related doctoral degrees.

We found significant differences in URM graduates among the 3 biomedical informatics-related training categories. URM graduates, particularly Black graduates, were more likely to complete degrees with a health primary focus. Black Americans represent approximately 13.4% of the US population, however, they comprise 3.2% and 4.5% of biomedical informatics-related doctoral graduates and STEM doctoral graduates, respectively.[26–28] URM graduates were also less likely to complete training programs with a primary focus in bioinformatics (SED code 102). A consistent theme from the STEM and biomedical informatics education literature suggests that URM candidates lack research experience and requisite skills necessary to pursue technical doctoral degrees.[13,15,29] The literature also suggests that they have a lower likelihood of pursuing graduate training programs due to early education inadequacies and disinterest in STEM fields.[16,20]

Table 2. Biomedical informatics-related doctoral recipients by race and category

	Total (%)	Bioinformatics is primary field (n = 1471)	Computer science is primary field with health-related secondary field (n = 619)	Health field is primary with secondary computer science-related field (n = 336)	*P* value
White	1166 (48.1)	665 (45.2)	339 (54.8)	162 (48.2)	.0003
Asian	976 (40.2)	673 (45.8)	189 (30.5)	114 (33.9)	<.0001
Hispanic	139 (5.7)	60 (4.1)	50 (8.1)	29 (8.6)	<.0001
Black	77 (3.2)	35 (2.4)	21 (3.4)	21 (6.3)	.001
URM other	68 (2.8)	38 (2.6)	20 (3.2)	10 (2.9)	.70

Table 3. Biomedical informatics-related doctoral recipients who accepted an academic position by race and training category

				Under-represented minorities (URMs)			
	Accepted an Academic Position	White	Asian	Hispanic	Black	URM Other	*P*-value
Bioinformatics is primary field	1213 of 1471 (82.5%)	543 of 665 (81.7%)	552 of 673 (82.0%)	Not Shown (95.0%)	29 of 35 (82.9%)	32 of 38 (97.4%)	.865
Computer science primary field with secondary health-related field	520 of 619 (84%)	288 of 339 (84.9%)	157 of 189 (83.1%)	42 of 50 (84.0%)	Not Shown (90.5%)	Not Shown (70%)	.997
Health field is primary with secondary computer science-related field	265 of 336 (78.9%)	122 of 160 (75.3%)	95 of 114 (83.3%)	Not Shown (86.2%)	Not Shown (71.4%)	Not Shown (80%)	.086
Total	1998 of 2426 (82.3%)	953 of 1164 (81.9%)	804 of 976 (82.4%)	124 of 139 (89.2%)	63 of 77 (81.8%)	54 of 68 (79.4%)	<.0001

Note: Per the data use agreement, fractions are not shown when the difference between the numerator and denominator is fewer than n = 5.

Table 4. Factors associated with URM graduates and URMs who accepted academic positions (2002–2017)

	Odds of being in a URM group OR (95% CI)	URM graduates who accepted an academic position OR (95% CI)
Sex		
Male	Ref	Ref
Female	1.11 (0.85–1.44)	1.06 (0.79–1.39)
Age groups		
<30	0.78 (0.45–1.36)	0.71 (0.38–1.33)
31-40	Ref	Ref
41-50	0.96 (0.67–1.39)	0.92 (0.62–1.36)
>51	0.65 0.36–1.17)	0.52 (0.27–1.00)[a]
Marital Status		
Single, never married	1.38 (1.05–1.81)[a]	1.31 (0.98–1.76)
Married	Ref	Ref
Separated/Divorced	1.27 (0.64–2.54)	1.67 (0.84–3.33)
Number of dependents aged <18 years		
0	Ref	Ref
1	1.95 (1.14–3.33)[b]	1.82 (1.02–3.25)[a]
2	1.95 (0.91–4.18)	2.21 (0.99–4.89)[a]
>3	Not Shown	Not Shown
US Citizen		
Yes	Ref	Ref
No	0.46 (0.34–0.61)[c]	0.55 (0.41–0.75)[c]
Received full tuition remission while a doctoral student		
Yes	Ref	Ref
No	1.37 (1.00–1.88)[a]	1.34 (0.96–1.88)
Biomedical informatics-related training category		
Bioinformatics	Ref	Ref
Computer science primary field with secondary health field	1.52 (1.09–2.09)[b]	1.39 (0.98–1.97)
Health field is primary with secondary computer science field	1.79 (1.24–2.57)[b]	1.61 (1.12–2.39)[a]
Year of graduation		
2002–2006	Ref	Ref
2007–2011	0.71 (0.41–1.02)	0.59 (0.34–1.07)
2012–2017	1.22 (0.72–2.07)	0.98 (0.56–1.71)
Carnegie classification		
Highest research activity (R1)	Ref	Ref
Higher research activity (R2)	1.13 (0.75–1.68)	1.13 (0.73–1.75)
Moderate research activity (R3)	2.16 (0.87–5.37)	1.74 (0.63–4.81)
Other	1.11 (0.73–1.68)	0.98 (0.62–1.55)
Type of university granting degree		
Public	Ref	Ref
Private	0.81 (0.61–1.07)	0.70 (0.52–0.95)[a]

[a] $P < .05$, [b] $P < .01$, [c] $P < .001$.

Per the data use agreement, results were suppressed where observations and odds ratios could be combined to reveal a subgroup of fewer than n = 5.

Our data suggest that biomedical informatics-related doctoral programs have not graduated a proportional increase in students from racially diverse backgrounds. This is evidenced by the nonsignificant increase observed in the number and proportion of URMs in the field over time. Consequently, the field of biomedical informatics does not produce URM graduates and researchers at rates that are reflective of the populations they represent.[6,30] Our analysis provides insight into institutional barriers to completing biomedical informatics-related doctoral training programs. For example, we found that URM graduates received full tuition remission less often than their non-URM counterparts which has been documented in other STEM fields as a barrier to degree completion.[31] Moreover, our results are consistent with previous literature examining "leaky pipeline" issues for URM graduates in other fields.[15,24,32] Previous research demonstrates that funding challenges, lack of mentorship, and inadequate educational preparation drive attrition rates and worsen the pipeline of URM students in biomedical informatics-related programs.[15,20] For example, lack of financing may deter URM students from not only pursuing highly technical degrees but also from accepting academic positions upon graduating. Other institutional and individual factors among URM students may drive program delays and attrition rates.[33,34] Future research should explore characteristics that are associated with program attrition rates and time-to-graduation including familial responsibilities and institutional constraints (ie, tuition remission levels, publicly vs privately funded institutions).

Disparities in URM graduation rates from doctoral programs and among faculty members in academia may adversely affect the trajectory of racially and ethnically diverse students which in turn results in a dearth of research that is sensitive to the lived experiences of URMs. Some of these challenges can be resolved through increased funding for training programs, particularly for historically black colleges and universities that, for instance, produce approximately 30% of black graduates in science and engineering doctoral programs.[35]

Our study had several limitations. First, the SED is administered at or near the time of graduation when some graduates may not have yet received academic or other job offers. Thus, our findings are meant to be more comparative among the races/ethnicities examined rather than a definitive description of job placements for the field overall. Second, it is important to note that the SED only captures information about research doctoral graduates and does not capture information on applicants or matriculants that do not complete their degree (including those that receive terminal master degrees after enrollment in PhD programs). Similarly, the SED does not include information from graduates of professional doctoral programs (eg, MD, DMD, DNP) who might have advanced training in biomedical informatics. As such, our study should only be generalized to graduates from relevant doctoral programs and not current students or applicants. Third, we combined race and ethnicity categories where we found small sample sizes which may conflate results within URM groups. This affects a wide range of nationalities that are broadly grouped, such as Hispanics, where heterogeneity in lived experiences are expected. Likewise, this affects the broad range of URMs classified as 'other' in our study. Fourth, these data are cross-sectional, thus, our results can only be interpreted as associations that do not necessarily imply causal relationships. Nevertheless, our study is the first study to examine the proportions of biomedical informatics-related doctoral graduates that come from an URM background over a 16-year period. Further research should urgently focus on better understanding challenges in recruiting and/or retaining diverse doctoral applicants so that the proportion of URMs graduating and entering academic positions could be positively affected. There is an opportunity to use mixed methods to understand barriers to attracting and retaining URM students within PhD programs in our field. Moreover, future research should quantify the number of URM biomedical informatics related researchers with clinical doctorates who are not adequately captured in our data. Doing so could provide a more complete picture of URMs in the broader biomedical informatics community and academia. Lastly, there is an opportunity to consider whether a discrete SED code for biomedical informatics warrants further consideration.

CONCLUSION

During a multi-decade period, there was no significant change in URM graduates of bioinformatics-related doctoral programs and no change in the proportion of URM graduates offered academic jobs. Without special efforts, the field is likely to have difficulty attracting trainees at the undergraduate and master's level and insufficient diversity of lived experience influencing research at the population health level. Interventions used successfully in other fields to support URM students should be considered in biomedical informatics. For example, the PhD Project, whose mission is to increase diversity among business school faculty, has been successful in increasing URM graduation rates in the business disciplines.[36] Program and academic leaders, as well as professional societies within biomedical informatics, should identify ways to invest in increasing URMs in the field.

FUNDING

This work was supported by the National Library of Medicine of the National Institutes of Health (NIH) under award T15LM012502.

AUTHOR CONTRIBUTIONS

KW and NM conceived and drafted the study; NM acquired the data for the work; KW analyzed the data; all authors contributed to revisions and approved the final version. All authors assume public responsibility for the accuracy and integrity of the work.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Gonzalez HB, Kuenzi JJ. Science, technology, engineering, and mathematics (STEM) education: A primer. Congressional Research Service, Library of Congress Washington, DC. 2012. http://www.stemedcoalition.org/wp-content/uploads/2010/05/STEM-Education-Primer.pdf Accessed September 1, 2020.
2. GAO-18-290, Science, Technology, Engineering, and Mathematics Education: Actions Needed to Better Assess the Federal Investment. https://www.gao.gov/assets/700/690840.pdf Accessed September 1, 2020.

3. Revitalizing the Nation's Talent Pool in STEM. https://www.urban.org/sites/default/files/publication/50501/311299-Revitalizing-the-Nation-s-Talent-Pool-in-STEM.PDF Accessed September 1, 2020.
4. Mendez RG, Torres J, Ishwad P, *et al.* Assisting bioinformatics programs at minority institutions: needs assessment, and lessons learned – a look at an internship program. In: Proceedings of the XSEDE16 Conference on Diversity, Big Data, and Science at Scale; July 2016.
5. U.S. National Library of Medicine. Collection development manual. https://www.nlm.nih.gov/tsd/acquisitions/cdm/subjects12.html Accessed May 5, 2019.
6. Canner JE, McEligot AJ, Pérez M-E, *et al.* Enhancing diversity in biomedical data science. *Ethn Dis* 2017; 27 (2): 107–16.
7. Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019 | NSF - National Science Foundation. https://ncses.nsf.gov/pubs/nsf19304/digest/field-of-degree-minorities Accessed May 8, 2020.
8. United States Census Bureau QuickFacts. https://www.census.gov/quickfacts/fact/table/US/PST045218 Accessed May 8, 2020.
9. Char DS, Shah NH, Magnus D. Implementing machine learning in health care: addressing ethical challenges. *N Engl J Med* 2018; 378 (11): 981–3.
10. Rajkomar A, Hardt M, Howell MD, *et al.* Ensuring fairness in machine learning to advance health equity. *Ann Intern Med* 2018; 169 (12): 866–72.
11. Obermeyer Z, Powers B, Vogeli C, *et al.* Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; 366 (6464): 447–53.
12. Estrada M, Burnett M, Campbell AG, *et al.* Improving underrepresented minority student persistence in STEM. *CBE Life Sci Educ* 2016; 15 (3): es5. doi: 10.1187/cbe.16-01-0038.
13. Hurtado S, Cabrera NL, Lin MH, *et al.* Diversifying science: underrepresented student experiences in structured research programs. *Res High Educ* 2009; 50 (2): 189–214.
14. McGee R Jr, Saran S, Krulwich TA. Diversity in the biomedical research workforce: developing talent. *Mt Sinai J Med* 2012; 79 (3): 397–411.
15. Allen-Ramdial S-AA, Campbell AG. Reimagining the pipeline: advancing STEM diversity, persistence, and success. *Bioscience* 2014; 64 (7): 612–8.
16. Willcockson IU, Johnson CW, Hersh W, *et al.* Predictors of student success in graduate biomedical informatics training: introductory course and program success. *J Am Med Inform Assoc* 2009; 16 (6): 837–46.
17. Hurtado S, Newman C, Tran M, *et al.* Improving the Rate of Success for Underrepresented Racial Minorities in STEM Fields: Insights from a National Project. UCLA; 2016. https://www.researchgate.net/profile/Christopher_Newman6/publication/303480066_An_Inspirational_and_Onerous_Journey_from_the_Great_Migration_to_the_Academy/links/59dc2675458515e9ab452cd0/An-Inspirational-and-Onerous-Journey-from-the-Great-Migration-to-the-Academy.pdf Accessed May 3, 2019.
18. 26 STEM Graduate Fellowships for Minorities | ProFellow. ProFellow. 2019. https://www.profellow.com/fellowships/26-stem-graduate-fellowships-for-minorities-and-women/ Accessed October 10, 2019.
19. Woodcock A, Hernandez PR, Schultz PW. Diversifying science: intervention programs moderate the effect of stereotype threat on motivation and career choice. *Soc Psychol Personal Sci* 2016; 7 (2): 184–92.
20. Hurtado S, Newman CB, Tran MC, *et al.* Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions Inst Res* 2010; 2010 (148): 5–15.
21. Management and Budget Office. Standards for maintaining, collecting, and presenting federal data on race and ethnicity. *Fed Regist* 2016; 81: 67398–401.https://www.federalregister.gov/d/2016-23672 Accessed September 1, 2020.
22. nsf.gov—Survey of Earned Doctorates—NCSES - US National Science Foundation (NSF). https://www.nsf.gov/statistics/srvydoctorates/ Accessed April 24, 2019.
23. Carnegie Classifications | Basic Classification. http://carnegieclassifications.iu.edu/classification_descriptions/basic.php Accessed October 3, 2019.
24. Jackson JR, Holmes AM, Golembiewski E, *et al.* Graduation and academic placement of underrepresented racial/ethnic minority doctoral recipients in public health disciplines, United States, 2003-2015. *Public Health Rep* 2019; 134 (1): 63–71.
25. Underrepresented minorities—Differences between women and men—Field of degree: Women, men, and racial and ethnic groups—nsf.gov—Women, Minorities, and Persons with Disabilities in Science and Engineering—NCSES—US National Science Foundation (NSF). https://nsf.gov/statistics/2017/nsf17310/digest/fod-wmreg/minorities-by-degree-share.cfm Accessed May 21, 2019.
26. FACT SHEET: Spurring African-American STEM Degree Completion | U.S. Department of Education. https://www.ed.gov/news/press-releases/fact-sheet-spurring-african-american-stem-degree-completion Accessed January 16, 2020.
27. U.S. Census Bureau QuickFacts: United States. Census Bureau QuickFacts. https://www.census.gov/quickfacts/fact/table/US/PST045218 Accessed January 16, 2020.
28. Anderson BJ. Minorities and mathematics: the new frontier and challenge of the nineties. *J Negro Educ* 1990; 59 (3): 260–72.
29. Museus SD, Palmer RT, Davis RJ, *et al. Racial and ethnic minority students' success in STEM education*. Hoboken, NJ: Jossey-Bass; 2011.
30. Southerland WM, Swamidass SJ, Payne PRO, *et al.* The diversity and disparity in biomedical informatics (DDBI) workshop. *Pac Symp Biocomput* 2018; 23: 614–7.
31. Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline. Expanding Underrepresented Minority Participation. National Academy of Sciences, National Academy of Engineering (US), and Institute of Medicine (US) 2011. https://www.ncbi.nlm.nih.gov/books/NBK83377/pdf/Bookshelf_NBK83377.pdf Accessed September 1, 2020.
32. Huang G, Taddese N, Walter E. Entry and Persistence of Women and Minorities in College Science and Engineering Education. Research and Development Report. NCES 2000-601. *National Center for Education Statistics* Published Online First: 2000. https://nces.ed.gov/pubs2000/2000601.pdf Accessed September 1, 2020.
33. Wilson MA, DePass A, Bean AJ. Institutional interventions that remove barriers to recruit and retain diverse biomedical PhD students. *LSE* 2018; 17 (2): ar27.
34. A Broken Pipeline: Minority Students and the Pathway to the Ph.D. | News | The Harvard Crimson. https://www.thecrimson.com/article/2016/4/27/minority-students-PhD-pipeline/ Accessed February 17, 2020.
35. The Impact of HBCUs on Diversity in STEM Fields | UNCF. UNCF. 2019. https://www.uncf.org/the-latest/the-impact-of-hbcus-on-diversity-in-stem-fields Accessed Jan 16, 2020.
36. The PhD Project: About Us. The PhD Project. https://www.phdproject.org/our-success/about-us/ Accessed May 8, 2020.