

The diversity and disparity in biomedical informatics (DDBI) workshop

William M. Southerland¹ and S. Joshua Swamidass²

¹*Department of Biochemistry & Molecular Biology*

Howard University College of Medicine

Washington, DC

Email: wsoutherland@howard.edu

²*Institute for Informatics*

Washington University in St. Louis

St. Louis, Missouri

Email: swamidass@gmail.com

Philip R. O. Payne

Institute for Informatics

Washington University in St. Louis

St. Louis, Missouri

Email: prpayne@wustl.edu

Laura Wiley

Division of Biomedical Informatics and Personalized Medicine

University of Colorado

Denver, Colorado

Email: laura.wiley@ucdenver.edu

ClarLynda Williams-DeVane

Biomedical Biotechnology Research Institute

North Carolina Central University

Durham, North Carolina

Email: clarlynda.williams@nccu.edu

The Diversity and Disparity in Biomedical Informatics (DDBI) workshop will be focused on complementary and critical issues concerned with enhancing diversity in the informatics workforce as well as diversity in patient cohorts. According to the National Institute of Minority Health and Health Disparities (NIMHD) at the NIH, diversity refers to the inclusion of the following traditionally underrepresented groups: African Americans/Blacks, Asians (>30 countries), American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Latino or Hispanic (20 countries). Gender, culture, and socioeconomic status are also important dimensions of diversity, which may define some underrepresented groups. The under-representation of specific groups in both the biomedical informatics workforce as well as in the patient-derived data that is being used for research purposes has contributed

to an ongoing disparity; these groups have not experienced equity in contributing to or benefiting from advancements in informatics research. This workshop will highlight innovative efforts to increase the pool of minority informaticians and discuss examples of informatics research that addresses the health concerns that impact minority populations. This workshop topics will provide insight into overcoming pipeline issues in the development of minority informaticians while emphasizing the importance of minority participation in health related research. The DDBI workshop will occur in two parts. Part I will discuss specific minority health & health disparities research topics and Part II will cover discussions related to overcoming pipeline issues in the training of minority informaticians.

Keywords: Diversity; Disparity; Biomedical Informatics.

1. Rationale for Diversity in Biomedical Informatics

The Diversity and Disparity in Biomedical Informatics (DDBI) workshop will be focused on complementary and critical issues concerned with enhancing diversity in the informatics workforce as well as diversity in patient cohorts. According to the National Institute of Minority Health and Health Disparities (NIMHD) at the NIH, diversity refers to the inclusion of the following traditionally underrepresented groups: African Americans/Blacks, Asians (>30 countries), American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Latino or Hispanic (20 countries). Gender, culture, and socioeconomic status are also important dimensions of diversity, which may define some underrepresented groups. The under-representation of specific groups in both the biomedical informatics workforce as well as in the patient-derived data that is being used for research purposes has contributed to an ongoing disparity; these groups have not experienced equity in contributing to or benefiting from advancements in informatics research. With the recent and growing importance of Precision Medicine initiatives in healthcare strategies and with the increasing contribution of biomedical informatics to Precision Medicine activities, it has become imperative that all segments of society become fully engaged and represented in biomedical informatics research efforts. Further, a diverse informatics workforce is an essential component to achieving diversity in patient cohorts.

2. The Need for Diversity in the Biomedical Informatics Workforce

We face a two-fold challenge: the lack of diversity in the biomedical and health informatics workforce and lack of diversity in study cohorts which leads to disparities in contributions from diverse groups which results in experimental results that are not fully reflective of demographics. These healthcare challenges present wide-ranging biomedical informatics opportunities that will be addressed in this workshop. These challenges are quite evident in the Precision Medicine (All of Us) initiative which emphasizes that the participation of individuals from diverse social, racial/ethnic, ancestral, geographic, and economic backgrounds is critical. It has become imperative that all segments of society become fully engaged and represented in informatics efforts since a diverse informatics workforce is essential to achieving diversity in patient cohorts and in scientific contributions. For example, this is especially true in the area of genomics (more specifically genome informatics) which is emerging as a guide in the treatment of disease. This workshop will highlight innovative efforts to increase the pool

of minority informaticians and discuss examples of informatics research that addresses the health concerns that impact minority populations. Collectively, these workshop topics will provide insight into overcoming pipeline issues in the development of minority informaticians while emphasizing the importance of minority participation in health related research.

3. Strategies for Addressing Diversity and Disparity in Informatics

The DDBI workshop will occur in two parts. Part I will discuss specific minority health & health disparities research topics and Part II will cover discussions related to overcoming pipeline issues in the training of minority informaticians.

3.1. *Minority Health and Health Disparities Research*

This phase of the workshop will focus on two scientific presentations, one presentation will focus on biomedical informatics and the other on clinical informatics. The biomedical presentation: “On Using Local Ancestry to inform eQTL mapping in African Americans” will present data on the impact of local ancestry on eQTL mapping of admixed populations such as African Americans. This presentation recognizes the importance of eQTL data in understanding the function of non-coding variants in European and Asian populations and explores the significance of this data in admixed populations such as African Americans that are disproportionately impacted by a variety of diseases. The clinical informatics presentation will focus on methodological approaches to Howard University Hospital (HUH) EHR data analysis. The HUH patient population is approximately 90% minority. Detailed analysis and understanding of this data promises to shed highly granular insights into unique parameters of minority health and health disparity in the Washington, DC metropolitan area. This presentation will focus on the development of an innovative interactive tool designed to probe HUH EHR data. HUeMR (Hoard University electronic Medical Records) is a secure web-based i2b2 plugin that supports complex Boolean search operations. It has a highly interactive user interface that allows rapid data analysis for cohort discovery. Data is displayed using editable interactive charts. Users can create multiple rows of charts that contain different types of data. Users can refine queries by clicking on the charts and then select one or more additional query parameters. Additionally, users can search for diagnosis by keyword or International Classification of Disease (ICD) codes. This presentation will feature use-case examples from the Howard University Hospital (HUH) de-identified EHR data.

3.2. *Overcoming Pipeline Issues in Developing Minority Informaticians*

The lack of diversity in the field of bioinformatics is a reflection of a larger problem of exposure of underrepresented students to opportunities and role models in fields that require computational ability. In order to increase the diversity of the informatics workforce there needs to be increased exposure of more diverse students to computational thinking and practices, role models, and opportunities. This exposure needs to be deliberate and targeted to ensure that students understand the career outlook and opportunities available in the field of bioinformatics. In fact, jobs for Computer and Information Research Scientist are projected to grow 11 percent faster than the average for all

occupations over the next seven years¹. In order to achieve this awareness researchers in the field of computer science, the natural sciences and industry partners are encouraged to form alliances that ensure that interventions are identified and implemented at critical junctures to increase the diversity of the bioinformatics workforce. Minority Serving Institutions (MSIs) should play a significant role in the advancement of this cause and need to take steps to develop their capacity to offer programs that prepare their students for this field².

This session will feature a panel presentation that will discuss challenges, opportunities, and success stories involving industry-academia partnerships focused on overcoming pipeline barriers in the training of minority informaticians. The Howard-Google collaboration, referred to as ‘Howard West’ will be a prominent discussion topic of the panel. The ‘Howard West’ initiative provides Howard University undergraduate students hands-on training and exposure in informatics at the Google campus in Mountain View, California with the aim of stimulating interest in computational methods as career options. The participating students are also accompanied by Howard faculty during the Google experience. The ‘Howard West’ initiative will be discussed from the perspectives of participating faculty, students, and Google partners. Positive outcomes, challenges and lessons learned will be discussed during this panel presentation. The panel will explore how partner stakeholders can work together to develop interdisciplinary programs to increase exposure of students at critical junctures. Additionally, the social, psychological and environmental issues related to ‘belonging’ that minorities face in the computing related workforce will be a topic of discussion. This panel will focus on collaborative and creative interdisciplinary pathways that serve to increase the diversity of the bioinformatics workforce. It is essential that promising interventions be shared, scaled and transferred across disciplines and institutions in educating and preparing diverse students. This panel comes at an opportune time to share challenges and solutions in the creation of approaches that address the lack of diversity in the bioinformatics workforce.

Acknowledgments

This project was supported (in part) by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under Award Number G12MD007597.

References

1. Computer and Information Research Scientists : Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics: <https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm>. Accessed: 2017-08-01.
2. Mendez, R.G., Torres, J., Ishwad, P., Nicholas, H.B., Jr. and Ropelewski, A. 2016. Assisting Bioinformatics Programs at Minority Institutions: Needs Assessment, and Lessons Learned – A Look at an Internship Program. *Proceedings of the XSEDE16 Conference on Diversity, Big Data, and Science at Scale* (New York, NY, USA, 2016), 52:1–52:8.