

# Demographic Bias in Human Cell Studies

Jessica Snyder<sup>1</sup>, Iva Bojic<sup>1,2</sup>, Aarom Gerow<sup>3</sup>, Carlo Ratti <sup>1,2</sup>

<sup>1</sup>*Massachusetts Institute of Technology, SENSEable City Lab*

<sup>2</sup>*Singapore-MIT Alliance for Research and Technology*

<sup>3</sup>*The University of Chicago, Knowledge Lab*

December 5, 2016

## Abstract

Health research has standardized breast cancer models, using a handful of cell lines, which has successfully generated survival rates near 90%, while failing to balance the demographics, and consequently tumor sub-types, of the patient population, possibly contributing to increased morbidity among demographic not represented in the most heavily used cell lines. Re-evaluating the use of cell lines are used to improve breast cancer treatment options, for a broader range of tumor sub-types, begins with a data-driven analysis of which and how cell lines have been used in published research and patents.

## Rationale

An estimated \$150 billion USD will be spent from 2010-2020 on nearly 20 million cancer patients [1]. The most common cancer in women, second overall to prostate cancer, is breast cancer, comprising 32% of the cost [1]. Breast cancer statistics published by the Center for Disease Control (CDC) since 1999 show that while white woman are more likely to develop breast cancer, black patients are less likely to survive than other races, exposing a well-acknowledged disparity in the efficacy of breast cancer treatment for black

women [2, 3]. The likely causes include societal, behavioral and biological factors, broadly categorized as extrinsic factors which are the responsibility of public policy and the individual, and intrinsic factors, such as genetics and cell structure within aggressive tumor sub-types, which are the responsibility of medical researchers to consider in designing fundamental research studies [4, 5].

Raising the survival rate of black breast cancer patients to the national average requires critical review of socioeconomic-dependence of high-quality health care access and inclusive diagnostic techniques, without question [6]. Failing to diagnosis at early stages, as well as faster growing, more aggressive tumor sub-types, disproportionately effect black patients [7].

Vetting the influence of intrinsic factors of breast cancer is critical to guide the action of fundamental medical research, especially the types of cell lines used in foundational research. Are there intrinsic factors of the tumor sub-types causing morbidity, which are disproportionately present in black breast cancer patients? Environmental exposure and genetic background determine the breast cancer's subtypes, aggressiveness, and ultimately, patient's prognosis. Highly aggressive tumor subtypes are correlated to risk factors such as obesity, which disproportionately effects black women. Self-reported race correlated to breast cancer properties, such as estrogen receptor status and stage of diagnosis, more than percent African ancestry, suggesting social determinants shape the disease physiology more than genetic predispositions [4]. Currently, alternative treatment options for patients with a genetic predisposition for hormone receptor-negative breast cancers, which are also likely black, is limited [8].

Research supports black breast cancer patients present elevated risk for aggressive, intrinsic factors of breast cancer which adversely effect prognosis. The next logical question is to what extent have black patients donated to cell lines used by medical researchers, and then, how well these lines are represented in published research and patents. The purpose of asking these questions is to determine any extent to under-representation, allowing for corrective action to simply include cell lines which represent the aggressive tumor sub-types causing the most morbidity. Data-driven retrospectives on disease burden and treatment effectiveness have been used in the past to guide medical research funding and health policy decisions towards the most productive population-level outcomes [9].

The long term consequences of medical research in demographics most heavily represented as donors to biopsies used in lab work have the best prognosis for treatment. Targeted breast cancer therapies and

preventative screening sensitive to various tumors are required to absolve demographic inequalities in breast cancer treatment [7]. Cell lines are a critical tool of scientific discovery for medical treatments. “Human cancer in a test tube” provides a standard for genotypic and phenotypic stability.

US CDC statistics publishes statistics on the number of women diagnosed with breast cancer divided by the population size of women multiplied by 100k, for all women and for each ethnicity. Similarly, the death rate is published as a function of the number of woman. White woman are most commonly diagnosed with breast cancer, however black woman are most likely die from the cancer, concluding medical treatment is more effective for white woman than black woman.

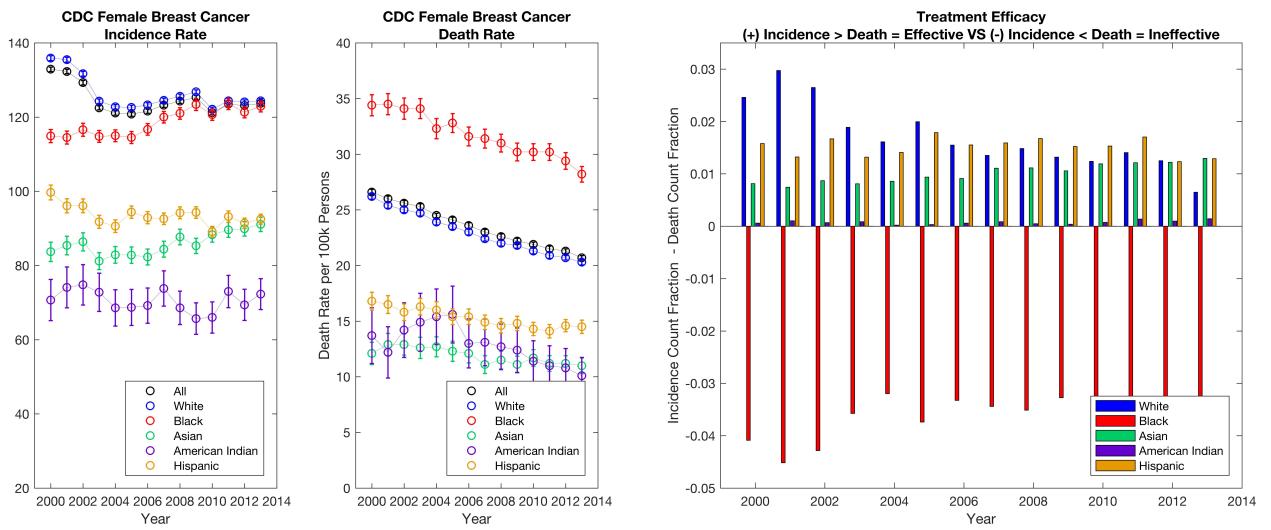


Figure 1: Rationale to investigate the representation of cell models derived from black patients in breast cancer research based on a persistent discrepancy in breast cancer treatment efficacy for black patients.

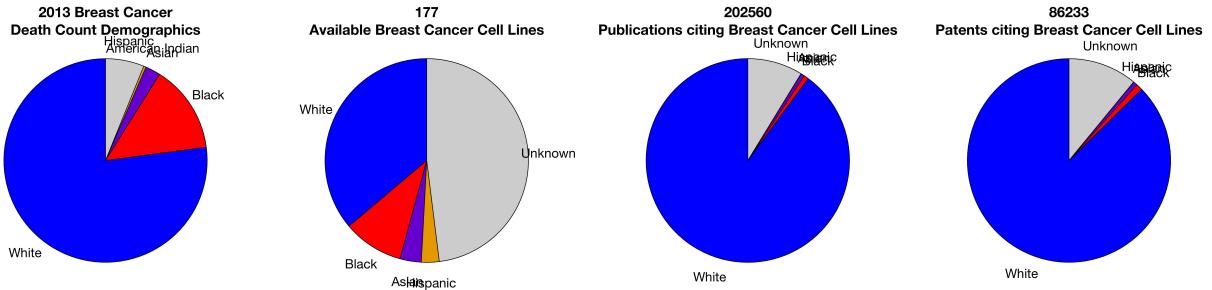


Figure 2: Comparison of the need for medical research, as defined by the most recent US CDC breast cancer death counts, to the demographic representation in the available cell models derived from breast tissue, and the use of the available cell lines in publications and US patents.

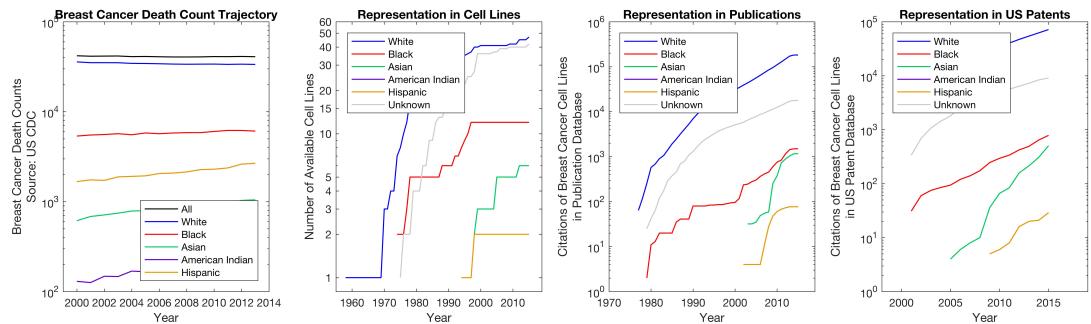


Figure 3: Demographic representation in available cell lines, publication and patents as compared to the death counts for breast cancer.

## The available human cell lines

### Demographic inclusiveness of the publication record

### Demographic inclusiveness of the US patent database

### Forging ahead with personalized medicine

Transformative approaches to health care, such as personalized medicine, authenticate the importance of genetic variation on the effectiveness of treatment, physicians use genomic analysis to determine the course of treatment, circumventing the limitations of generalized treatments with tailored treatments. Here, we survey

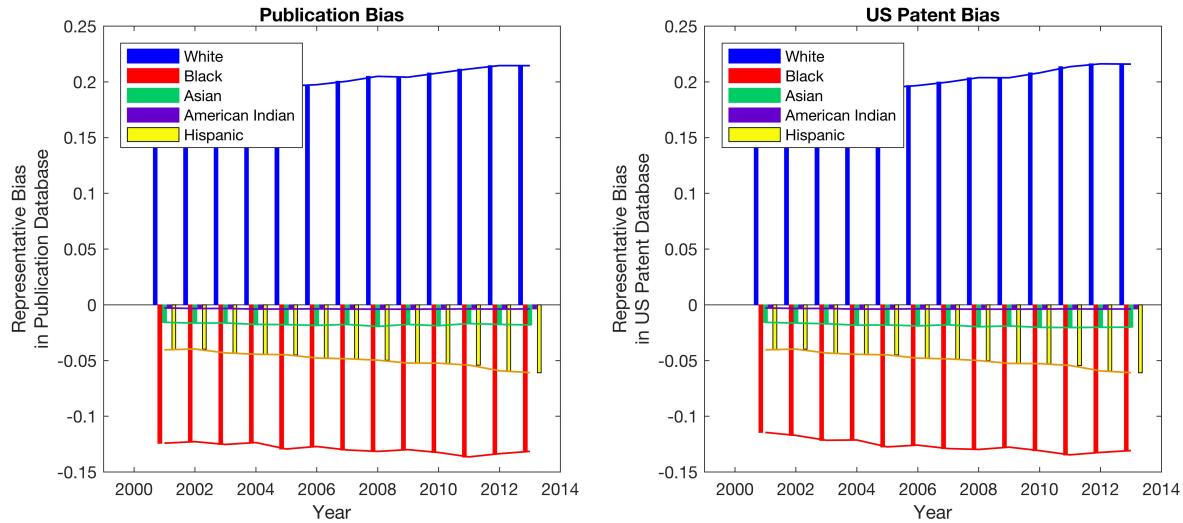


Figure 4: The representative bias in breast cancer medical research and therapeutic development, as determined by comparing the fraction of each demographic in the breast cancer death count to the number of publications and US patents citing cell lines generated from each ethnicity.

the genetic variation in breast cancer research with human cell models by characterizing the demographics of cell line donors represented in breast cancer research.

## References

- [1] Angela B Mariotto, K Robin Yabroff, Yongwu Shao, Eric J Feuer, and Martin L Brown. Projections of the cost of cancer care in the united states: 2010–2020. *Journal of the National Cancer Institute*, 2011.
- [2] Robert T Greenlee, Taylor Murray, Sherry Bolden, and Phyllis A Wingo. Cancer statistics, 2000. *CA: a cancer journal for clinicians*, 50(1):7–33, 2000.
- [3] Amy Berkman, Bernard F Cole, Philips A Ades, Samantha Dickey, Stephen T Higgins, Amy Trentham-Dietz, Brian L Sprague, and Susan G Lakoski. Racial differences in breast cancer, cardiovascular disease, and all-cause mortality among women with ductal carcinoma in situ of the breast. *Breast cancer research and treatment*, 148(2):407–413, 2014.
- [4] Kerryn W Reding, Christopher S Carlson, Orsalem Kahsai, Christina C Chen, Andrew McDavid, David R Doody, Chu Chen, Kimberly Lowe, Leslie Bernstein, Linda Weiss, et al. Examination of

- ancestral informative markers and self-reported race with tumor characteristics of breast cancer among black and white women. *Breast cancer research and treatment*, 134(2):801–809, 2012.
- [5] Kevin Brennan and James M Flanagan. Is there a link between genome-wide hypomethylation in blood and cancer risk? *Cancer prevention research*, 5(12):1345–1357, 2012.
- [6] Vickie L Shavers, Linda C Harlan, and Jennifer L Stevens. Racial/ethnic variation in clinical presentation, treatment, and survival among breast cancer patients under age 35. *Cancer*, 97(1):134–147, 2003.
- [7] Nataliya G Batina, Amy Trentham-Dietz, Ronald E Gangnon, Brian L Sprague, Marjorie A Rosenberg, Natasha K Stout, Dennis G Fryback, and Oguzhan Alagoz. Variation in tumor natural history contributes to racial disparities in breast cancer stage at diagnosis. *Breast cancer research and treatment*, 138(2):519–528, 2013.
- [8] Dezheng Huo, Francis Ikpatt, Andrey Khramtsov, Jean-Marie Dangou, Rita Nanda, James Dignam, Bifeng Zhang, Tatyana Grushko, Chunling Zhang, Olayiwola Oluwasola, et al. Population differences in breast cancer: survey in indigenous african women reveals over-representation of triple-negative breast cancer. *Journal of Clinical Oncology*, 27(27):4515–4521, 2009.
- [9] Jane J Kim, Anna NA Tosteson, Ann G Zauber, Brian L Sprague, Natasha K Stout, Oguzhan Alagoz, Amy Trentham-Dietz, Katrina Armstrong, Sandi L Pruitt, Carolyn M Rutter, et al. Cancer models and real-world data: Better together. *Journal of the National Cancer Institute*, 108(2):djh316, 2016.

## **Methods**

### **Definitions**

**Cell line nomenclature**

**Excluding misidentified cell lines**

**Building the database of publications**

**Building the database of patents**

**Comparison to incidence and morbidity counts**