

Anti-Racism and the Next Generation of Disease Diagnostic Tools

This manuscript ([permalink](#)) was automatically generated from [rreggiar/farr_manuscript@b1cd990](#) on October 6, 2022.

Authors

- **Roman E. Reggiardo**

 [0000-0001-6309-249X](#) ·  [rreggiar](#) ·  [rreggiar](#)

Department of Biomolecular Engineering, University of California at Santa Cruz

· Funded by UCSC Baskin Engineering FARR; NIDDK/KUH F99 DK131504

✉ — Correspondence possible via [GitHub Issues](#)

Abstract

Various forms of highly technical genetic analyses have a long history of being ignorant of race, ethnicity, class, and sex as well as being co-opted for explicitly eugenic, racist, homophobic, or otherwise hateful purposes. There is a critical ongoing discussion on how modern sequencing technologies can both contribute to reduction of suffering but also drive harmful narratives due to incomplete and/or unrepresentative results being interpreted as evidence for biological superiority or inferiority.

A new class of medical technologies, called 'liquid biopsies', are poised to dramatically impact how and when cancer diagnoses are made by harnessing genetic information. These diagnostic tests account for mutations in one's genome and/or changes in the abundance of certain expressed genes, measured from a simple blood draw.

While the promise of liquid biopsies is a future where diagnoses are accessible, accurate, and non-invasive, they face the same potential shortcomings as other genetic tests: without adequate representation during development, it is impossible to say whether diagnostic performance and relevance extends inclusively and equitably across our diverse society. This burgeoning technology is intertwined with an equally impactful, far more pervasive, and demonstrably biased approach to classification problems: machine learning.

In this report, I intend to both quantify and discuss the relevant barriers to equitable liquid biopsy diagnostics by surveying the existing published data and investigating the sources, details, and diversity of samples used to demonstrate liquid biopsy performance.

UC Santa Cruz Land Acknowledgement

> "The land on which we gather is the unceded territory of the Awaswas-speaking Uypi Tribe. > The Amah Mutsun Tribal Band, comprised of the descendants of indigenous people taken to > missions Santa Cruz and San Juan Bautista during Spanish colonization of the Central Coast, > is today working hard to restore traditional stewardship practices on these lands and heal from historical trauma."

Introduction

Diseases, particularly cancer, must be identified and characterized before effective therapies can be applied. To make a diagnosis as such, there must be clearly “defined characteristics”, also known as “biomarkers”, detected at abnormal abundance relative to a pre-defined standard. Historically, these biomarkers have enabled the detection

Genomics at the forefront of diagnostic technology

Molecular signatures secreted into extracellular space have long been utilized to reveal pathological events in diverse, potentially ill populations.

As the complexity of intended-to-diagnose diseases increases to multi-genic and multi-organ systems, the ability to acquire more complex and broadly informative extracellular signatures is paramount in the development of equivalent diagnostic approaches: Next Generation Sequencing (NGS) **liquid biopsies**.

Now, at the intersection of contemporary high-throughput NGS and robust machine-learning enabled classification models, liquid biopsies are being developed to sensitively and specifically diagnose diseases in broader at-risk populations. While the technologies are primed to deliver a new paradigm in diagnostics that reshape prognostic expectations, there are numerous challenges rooted in equitable representation and meaningful inclusion that require consideration. Among them are those questions that have dogged genomics-enabled genetics research in recent decades:

- *Who is data generated by? Who is data generated for?*
- *How is data stored and who owns it?*
- *Who benefits from the analyses and technologies built upon biomedical datasets?*
- *How are social labels used to contextualize analysis?*

This article seeks to contextualize the advent of minimally invasive, genetics- and genomics-based diagnostics within the current efforts to make genomics research and clinical development more equitable and inclusive endeavors. In particular, much has been made of the bias and danger inherent to proclamations derived from currently limited Human genomes and studies. Here, the focus will be placed on the noisier, more heterogeneous output of the genome: *Transcription*. As RNA expression becomes a critical tool in diagnosing disease, considering the intersection of this data-rich event with current challenges to genomic studies is critical to understanding its future role as a clinical diagnostic tool.

References
