

Rapid COVID-19 Prognostic Blood Test for Disease Severity Using Epigenetic Immune System Biomarkers

Adam G. Marsh;¹ G. Mark Anderson;² Erich J. Izdepski³

1. University of Delaware, Center for Bioinformatics and Computational Biology; Genome Profiling, LLC

2. Genome Profiling, LLC

3. BTS Software Solutions

Abstract

Objective. To develop a novel whole-blood epigenetic biomarker of immune system status, or EpiMarker, that would indicate whether a person with a recent COVID-19 diagnosis is at risk for severe symptoms including Acute Respiratory Distress Syndrome. **Methods.** Using a novel methyl-sensitive restriction endonuclease approach to measure site-specific DNA methylation profiles, immune system phenotype EpiMarkers are identified using a machine-learning computational bioinformatics platform. The result is a diagnostic network of 20 to 40 immune DNA methylation sites having the greatest predictive power for identifying patients whose COVID-19 disease will likely progress to ARDS requiring ICU/intubation care. **Results.**

Immune system status in peripheral whole blood provides a sensitive and responsive sentinel signal reflecting how different functional pathways are currently being regulated in a subject. Deciphering this signal status of how immune cells are set to respond provides deep functional information regarding patient health and potential disease phenotypes resulting from a cytokine storm characteristic of a hyper immune inflammatory response to COVID-19 infection.

Conclusions. The ability to identify future potential changes in patient health using this novel EpiMarker technology opens new avenues for defending populations from severe disease risks of Acute Respiratory Distress Syndrome. **Policy Implications.** A successful EpiMarker Assay for COVID-19 disease severity risk would allow for two important applications: (1) patients could be triaged early in the course of infection to allow for critical decisions for allocating resources, both in terms of hospital infrastructure (ICU beds, ventilators) and therapeutic drug treatments; and (2) pre-infection, individuals could be screened to identify personnel at low-risk for mission critical assignments (first responders, doctors, nurses, military personnel, etc.) during future pandemics and ongoing battles with viral pathogens like influenza.

Introduction

The COVID-19 pandemic is causing enormous patient suffering and death, and unprecedented economic and operational disruptions. A key challenge to managing the pandemic is not knowing which patients will experience severe disease and which will experience no symptoms or only mild disease. Consequently, limited medical resources cannot be focused on the approximately 20% of COVID-19 infected patients whose lives will be most threatened by the infection. This is a critical issue now, and with no anti-viral therapeutic or vaccine availability projected until 2021 or later, it will continue to be a critical issue going forward.

Although it is early in the medical field's understanding of COVID-19 disease etiology, one pattern has emerged: the leading cause of mortality in patients with COVID-19 is hypoxic

respiratory failure from acute respiratory distress syndrome (ARDS).^{1,2} We know that the lung epithelium is a primary site of infection with a complex histopathology. Complications generally present as an immune system overreaction and dysfunction with severe inflammation and leukocyte infiltration.³ In these cases there is likely a malfunction in the patient's ability to appropriately regulate their immune response to the viral challenge. Severe tissue inflammation arises from a "cytokine storm" of signaling pathways that disintermediate normal controls on immune system function.⁴ Cytokine signaling pathways are an important molecular regulator of the human adaptive immune system but can produce deleterious side effects when unbalanced (e.g., auto-immune diseases).

To fight COVID-19, we need a better understanding of immune system response to this infection. It is likely that an immune system pre-disposition contributes to the ability of some people to easily fight the infection while others experience a dysfunctional viral infection response that leads to ARDS. Recent work has shown the importance of epigenetic control mechanisms in response to viral antigen vaccines.⁵ The adaptive antigen production response involves many pathways in circulating white blood cells and DNA methylation is an important epigenetic mechanism that is involved in determining vaccine efficacy in subjects. Thus, specific pre-infection pathway structures or regulatory controls (embodied by epigenetic DNA methylation patterns) may exist in patients that ultimately suffer ARDS that can serve as early biomarkers of mortality and morbidity risks in response to COVID-19 infection.

Technology

DNA methylation is an important epigenetic mechanism used by cells to control gene expression. Exogenous stressors, such as disease intrusion or immune system changes, trigger unique and specific DNA methylation responses, or signatures, that are encoded across the 28 million genomic methylation sites in every cell's DNA.

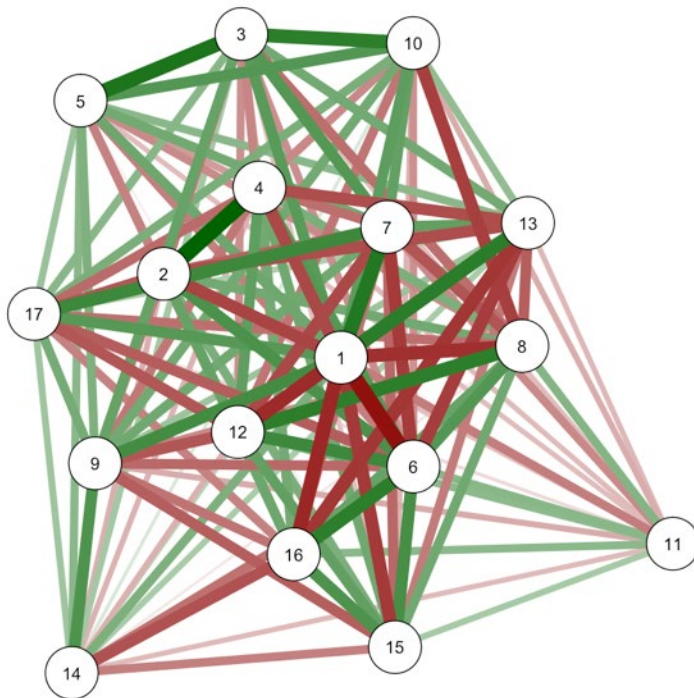
The GenPro EpiMarker Platform measures genome-wide DNA methylation from which it discovers complex patterns of methylation in genomic DNA that can distinguish between two functional phenotypes such as the presence or absence of disease or likelihood of responding to a particular drug. In this proposed use, pattern analyses of CpG methylation would be focused on identifying putative prognostic biomarkers of immune-system status and provide functional information for genes and pathways associated with COVID-19 disease progression in different patients.

Genome Profiling's technology for identifying epigenetic biomarkers (EpiMarkers) in peripheral blood immune cells has been effective in diagnostic applications for cerebral palsy,⁶ early breast cancer risk assessment, Parkinson's disease,⁷ leukemia and immuno-oncology (IO) applications to stratify clinical trial subjects. The nature of IO drugs provides an ideal confirmation of the predictive power of Genome Profiling's EpiMarker platform. Recently completed retrospective studies with a top ten biopharma immuno-oncology translational research team has resulted in the discovery and blind validation of a novel EpiMarker that successfully stratifies responder vs non-responder solid tumor cancer patients in clinical trials of IO checkpoint inhibitor therapies. Here, EpiMarker assays for immune system pathway status can identify patients that are more likely to respond to an IO therapy, increasing the response rate in a stratified cohort by 50% to 60% above the unstratified response rate. These results are robust even with patient cohorts in late stage trials with extensive prior treatments and different tumor histology types.

The EpiMarker is a collection of CpG sites (dinucleotide sequence of cytosine and guanosine) and the algorithmic models that connect or define their associations and synergistic predictive power (see Figure 1). The representation in Figure 1 shows the EpiMarker as a network with the CpG sites as the nodes and the model equations relating them together as the edges.

Conceptually, the collection of both CpG sites and equations can be thought of as a network, yet the physical embodiment of the EpiMarker is more like a high-dimensional database where over 500,000 model equations and results derived from the set of CpG sites is stored for retrieval when classification calls are needed for new, unknown samples.

Figure 1. EpiMarker Network-like Topology. From a methylome profile of 2 million CpG sites, a predictive set of 20 to 40 CpG sites is identified by the EpiMarker Discovery platform. In this representation, the CpG sites comprise the nodes (circles) and the algorithmic, numerical relationships among CpG node combinations are the red and green edges (lines). The structure is high-dimensional with over 500,000 equations described by the connecting lines.



Genome Profiling's novel application of immuno-methylome profiles holds promise for discovering why some patients are much more susceptible to adverse outcomes arising from COVID-19 infection. The greater accuracy with Genome Profiling's methyl-sensitive restriction endonuclease metrics provides unparalleled analytical power in discriminating subtle methylome EpiMarker patterns that are diagnostic of disease and health among different patients.⁶ Further, the machine-learning discovery platform delivers methylome profiles that are incredibly rich in information for assessing the status and activities of a patient's immune system. Combining this level of immuno-phenotypic data with the platform's statistical pattern analysis and machine-learning capabilities to derive novel diagnostic/prediction models, opens new frontiers for research and discovery in clinical epigenetic applications.

Approach

For COVID-19 disease progression and the onset of medical complications, it is likely that patients who resist infection and experience only mild symptoms do so because of an immune system reaction that is predisposed or preconditioned to combating coronaviruses. Such predisposition could also involve greater regulatory controls over innate and/or adaptive immune responses to prevent tissue inflammation in lung epithelia. By comparing patients with different disease symptoms (mild vs severe), the functional state differences in their pre-symptomatic immune systems will likely identify an EpiMarker signature in peripheral blood. In addition, the EpiMarker discovery process will illuminate biological information (e.g. genes, pathways) that are sensitive to the infection microenvironment and early pathology in the lung, as well as possible markers for the potential for immune cells to drive harmful overproduction of cytokines.

The development of a diagnostic assay would be executed in two stages. First, a retrospective clinical trial for Discovery & Blind Validation of the EpiMarker. Here, the EpiMarker Platform will be used to profile the immune system status of subjects who are known to have experienced either: mild disease symptoms requiring minimal hospital care, or severe disease symptoms requiring advanced, specialized hospital care and ventilator intubation.

The initial work is based on whole genome sequencing using Genome Profiling's novel approach to measure DNA methylation via NGS. Once quantified, a Training Set of 50 vs 50 (mild:severe) patients will be used for the machine learning platform to identify the EpiMarker set of 20 to 40 CpG sites that has the most predictive power to identify the symptom group-type of a blind Validation Set of samples (100 total, 50:50 mild:severe). The first working components of the EpiMarker bioinformatic predictive algorithms are produced at this stage and assembled into a structure for easy retrieval and execution.

The second stage would focus on translating the NGS derived EpiMarker into a targeted, high-throughput plate assay that can be executed at a fraction of the cost of whole-genome sequencing per patient. Once the EpiMarker set is identified, then a targeted sequencing panel-capture assay is designed and developed for each of the CpG sites in the EpiMarker. Off-the-shelf technology using Thermo Fisher's AmpliSeq platform can allow for the rapid development of the specific molecular reagents required to measure CpG methylation at each of the target sites. The AmpliSeq technology allows for the multiplexing of hundreds of target sites if needed and is well suited for a rapid, high-throughput assay format. In addition, this assay platform is well-suited for CLIA approvals and rapid clinical implementation as Laboratory Developed Test (LDT). During development, 200 initial samples would serve as the material that would guide and test each of the CpG assays. Assay development would be executed in a CLIA certified laboratory. Once completed, assay reagents can be ordered directly from Thermo Fisher including the AmpliSeq plate format sample holders that would be required to execute the assay.

The assay development plan would meet all regulatory and quality requirements described in the FDA Emergency Use Authorization guidance document. As soon as a clinical partner and funding source are identified, we would schedule a discussion with the FDA regarding Test and Criteria for Emergency Use Authorization Issuance. A prognostic assay for COVID-19 infection and associated ARDS clearly meets the criterion of a serious or life-threatening disease. Evidence of effectiveness will be obtained by applying the final assay format to approximately 200 blood samples from COVID-19 patients with known clinical outcome in a fully blinded validation round to demonstrate and confirm the accuracy and performance of the test.

In terms of risk and benefit to the patient this analysis poses minimal direct risks because only a small, standard blood sample draw is required to perform the test. More importantly, at present there are no alternatives to the proposed prognostic assay to triage patients for COVID-19 risk for ARDS or other severe outcomes. Patients falling into broad categories based on age or co-morbidities such as obesity or diabetes seem to be at higher risk, but the majority of patients in these classes still do not experience severe COVID-19 symptoms.

Public Health Implications

Establishing a link between an immune system status EpiMarker and the functional ability of patients to successfully fight a COVID-19 infection would have large ramifications for pursuing new therapeutic agents and diagnostic/prognostic strategies. This is critical not just for better management of the COVID-19 threat but also in preparing global health systems for future pandemics, as well as aiding in our current battles against non-pandemic but ever-present viral pathogen diseases, like seasonal influenza. If an EpiMarker assay can demonstrate that there is a defined immuno-methylome imprint that predisposes patients to fight viremia infections, then this assay could be a game changer in terms of opening new options for how we as a society respond to infectious diseases.

Diagnostic immuno-methylome EpiMarker assays hold promise for identifying subjects at risk. The most immediate would be in health care triage. Aggressive treatment options could be pursued earlier in patients at high risk. This likely has the most promise for decreasing the COVID-19 mortality rate at the source. By not treating patients at low risk of severe disease in hospitals, precious resources of beds, equipment, medications would be saved, as well as reducing the overall exposure risk to doctors and nurses by only hospitalizing those patients that are likely to need such acute care.

In addition to health care resource triage and planning, this EpiMarker assay has a large potential impact in terms of the ability of local and city governments to pre-screen police and first-responders for disease severity risks. Screened low-risk personnel can be dispatched to public-facing, mission critical jobs knowing that the mortality and morbidity risks from performing these jobs and concomitant exposure to COVID-19 may be ameliorated for them. In addition, hospital intensive care staffing decisions for doctors, nurses, orderlies could also be informed by each individual's personal risk to severe disease symptoms. And by extension, if successful, every large business could have employees screened, allowing formulation of a pandemic critical-mission personnel plan for continuity of economic activities without having to shut-down operations to the extent experienced in the spring of 2020.

Importantly, the EpiMarker Platform is not disease or application specific. It can be trained to provide severity prognostics on other infectious diseases. Adding support for analyzing case severity of a new disease follows the same process; namely, collecting data from new patients with varying demographics and disease response severity, training a new machine learning model, and deploying the model. The resulting panel assay is based on a standard blood draw and analyzed using standard laboratory equipment. Novel EpiMarker-enabled blood tests will reduce death rates and are much faster to develop and deploy than either a new vaccine or immunity test. Moreover, if there are vaccine or anti-viral drugs, these tests will efficiently prioritize which patients should be treated first.

Within a military context, availability of advanced medical care requires special planning in deployed or other extremis military situations. Prognostic testing such as this can help avoid mission compromise by identifying at risk subjects a priori. Further expansion and development of this diagnostic platform strategy could become a frontline defense for future pandemics, current battles with viral pathogens like influenza, and be extended to other biological threats where early assessment of patient risk will lead to more effective allocation of scarce medical care and resources.

A simple EpiMarker blood-based assay has significant potential for the Army, other military branches and the civilian contractors that support their missions. Military operations and unit effectiveness can be compromised by disease, natural or weaponized. Knowing which individuals would be most susceptible to serious, incapacitating illness and which individuals would be moderately inconvenienced by an illness provides a powerful planning and operational tool. Imagine if you knew in advance which personnel would be more susceptible to a serious, life threatening disease than others? They could be assigned accordingly, perhaps in closer proximity to medical facilities, rather than deployed to areas of high disease risk or into remote areas without medical facilities.

Competing Interests

AGM is the inventor of the EpiMarker technology and a Co-Founder of Genome Profiling LLC. He holds an equity position in this company and is a member of the Board of Directors. GMA is VP for EpiMarker Solutions to Genome Profiling. EJI is Senior VP for Research and Development at BTS Software Solutions.

References

1. Teuwen, L. A., Geldhof, V., Pasut, A., & Carmeliet, P. (2020, July). COVID-19: The vasculature unleashed. *Nature Reviews. Immunology*, 20(7), 389–391. [PubMed https://doi.org/10.1038/s41577-020-0343-0](https://doi.org/10.1038/s41577-020-0343-0)
2. Wu, Z., & McGoogan, J. M. (2020, April 7). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA*, 323(13), 1239–1242. [PubMed https://doi.org/10.1001/jama.2020.2648](https://doi.org/10.1001/jama.2020.2648)
3. Fung, S. Y., Yuen, K. S., Ye, Z. W., Chan, C. P., & Jin, D. Y. (2020, March 14). A tug-of-war between severe acute respiratory syndrome coronavirus 2 and host antiviral defence: Lessons from other pathogenic viruses. *Emerging Microbes & Infections*, 9(1), 558–570. [PubMed https://doi.org/10.1080/22221751.2020.1736644](https://doi.org/10.1080/22221751.2020.1736644)
4. Maticic, C. (2020, Jun 2), Blood vessel attack could trigger coronavirus' fatal 'second phase'. *Science*. Retrieved from: <https://doi.org/doi:10.1126/science.abd1296>
5. Arts, R. J. W., Moorlag, S. J. C. F. M., Novakovic, B., Li, Y., Wang, S. Y., Oosting, M., . . . Netea, M. G. (2018, January 10). BCG vaccination protects against experimental viral infection in humans through the induction of cytokines associated with trained immunity. *Cell Host & Microbe*, 23(1), 89–100.e5. [PubMed https://doi.org/10.1016/j.chom.2017.12.010](https://doi.org/10.1016/j.chom.2017.12.010)

6. Crowgey, E. L., Marsh, A. G., Robinson, K. G., Yeager, S. K., & Akins, R. E. (2018, June 21). Epigenetic machine learning: Utilizing DNA methylation patterns to predict spastic cerebral palsy. *BMC Bioinformatics*, 19(1), 225. [PubMed https://doi.org/10.1186/s12859-018-2224-0](https://doi.org/10.1186/s12859-018-2224-0)
7. Marsh, A. G., Cottrell, M. T., & Goldman, M. F. (2016, November 2). Epigenetic DNA methylation profiling with MSRE: A quantitative NGS approach using a Parkinson's Disease test case. *Frontiers in Genetics*, 7, 191. [PubMed https://doi.org/10.3389/fgene.2016.00191](https://doi.org/10.3389/fgene.2016.00191)

Copyright (c) 2020 Delaware Academy of Medicine / Delaware Public Health Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.