Targeting the Peripheral Immune System as a Source of Precision Stroke Biomarkers: Implications for Clinical Diagnostics and Immunotherapeutics

Grant C. O'Connell

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
O'Connell, Grant C., "Targeting the Peripheral Immune System as a Source of Precision Stroke Biomarkers: Implications for Clinical Diagnostics and Immunotherapeutics" (2016). Graduate Theses, Dissertations, and Problem Reports. 6338.
https://researchrepository.wvu.edu/etd/6338
Targeting the Peripheral Immune System as a Source of Precision Stroke Biomarkers: Implications for Clinical Diagnostics and Immunotherapeutics

Grant C. O’Connell

Dissertation submitted
to the School of Pharmacy
at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in
Pharmaceutical and Pharmacological Sciences

Paul D. Chantler, Ph.D., Chair
Taura L. Barr, Ph.D.
Lori A. Hazlehurst, Ph.D.
Jason D. Huber, Ph.D.
Gordon P. Meares Ph.D.

Department of Pharmaceutical Sciences

Morgantown, West Virginia
2016

Keywords: Cardiovascular Disease, Machine Learning, Pattern Recognition, Cerebrovascular Disease, Systems Biology

Copyright 2016 Grant C. O’Connell
ABSTRACT

Targeting the Peripheral Immune System as a Source of Precision Stroke Biomarkers: Implications for Clinical Diagnostics and Immunotherapeutics

Grant C. O'Connell

Early and accurate diagnosis of stroke improves the probability of positive outcome, however the tools available to clinicians for the identification of stroke are limited. It is well established that the peripheral immune system responds robustly to ischemic brain injury, and thus may be a viable source of stroke biomarkers. The objective of this work was to utilize high throughput transcriptomics in combination with machine learning techniques to identify patterns of gene expression in peripheral whole blood with the potential to diagnose ischemic stroke, and to further determine the role of such gene expression patterns within the context of stroke immunopathology. A discovery cohort was recruited comprised of 39 ischemic stroke patients and 24 controls, and peripheral blood was sampled at emergency department admission. Genome-wide expression profiling was performed via microarray and a machine learning technique known genetic algorithm k-nearest neighbors (GA/kNN) was then used to identify ten genes (ANTXR2, STK3, PDK4, CD163, MAL, GRAP, ID3, CTSZ, KIF1B, and PLXDC2) whose coordinate pattern of expression was able to discriminate between groups with 98.4% accuracy. The diagnostic robustness of this pattern of expression was subsequently tested in an independent validation cohort which included an additional 39 ischemic stroke patients and 50 controls, and was able to discriminate between groups with 95.3% accuracy. Furthermore, leukocyte phenotyping of validation cohort subjects suggested that this pattern of differential expression is a marker of stroke-induced shifts in the cellular composition of the peripheral blood, as many of these ten genes exhibited a strong pattern of lineage-specific expression on isolated leukocyte subpopulations, and whole blood expression levels which were highly correlated with post-stroke white blood cell differential. However, results of in-vitro experiments using primary human leukocyte cultures suggested that one of these ten genes, CD163, which encodes a lymphoinhibitory peptide known as soluble cluster of differentiation 163 (sCD163), may play a role in modulation of stroke-induced peripheral immune phenotype, as monocyte-derived sCD163 present in the serum of ischemic stroke patients demonstrated an ability to inhibit the proliferation of lymphocytes isolated from healthy donors. Due to the robust nature of these findings, the ten transcriptional biomarkers identified by GA/kNN in our analysis warrant further evaluation to determine their true diagnostic efficacy in the clinical setting. Furthermore, future exploration into the role of sCD163 as a post-stroke immunomodulatory agent could provide novel insight regarding the underlying mechanisms which drive stroke immunopathology.
Targeting the peripheral immune system as a source of precision stroke biomarkers: Implications for clinical diagnostics and immunotherapeutics.

Grant C. O’Connell¹,²

¹Department of Pharmaceutical Sciences, School of Pharmacy, West Virginia University
²Center for Translational Stroke Research, Robert C. Byrd Health Sciences Center, West Virginia University
# Table of Contents:

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction and specific aims</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Background and significance</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Stroke</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>A need for stroke biomarkers</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Early stroke biomarker studies</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>The peripheral immune response as a source of biomarkers</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Preliminary stroke biomarker studies targeting the peripheral blood transcriptome</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Machine learning for biomarker discovery</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Machine learning for identification of candidate ischemic stroke biomarkers</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Role of stroke-identifying biomarkers in immunopathology</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>CD163 as a potential modulator of the stroke-induced peripheral immune response</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>Discussion</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Machine learning for identification of candidate ischemic stroke biomarkers</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Role of stroke-identifying biomarkers in immunopathology</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>CD163 as a potential modulator of the stroke-induced peripheral immune response</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>Experimental procedures</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Discovery cohort patients</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Validation cohort patients</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Neuroradiological imaging</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Blood collection</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Isolation of serum and plasma</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Isolation of leukocyte subpopulations for RNA analysis</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>RNA extraction</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>RNA amplification and microarray</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>GA/kNN analysis</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Quantitative reverse transcription PCR</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>White blood cell differential and NMLR</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Isolation of primary leukocytes for cell culture</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Neutrophil culture</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Monocyte culture</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>sCD163 depletion from human serum</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Lymphocyte proliferation assay</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>sCD163 ELISA</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>ADAM17 activity assay</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Phase Contrast Microscopy</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Statistical analysis</td>
<td>38</td>
</tr>
<tr>
<td>6</td>
<td>References</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>Supplemental data</td>
<td>45</td>
</tr>
<tr>
<td>Appendix</td>
<td>Associated manuscripts</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Manuscript 1</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Manuscript 2</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Manuscript 3</td>
<td>103</td>
</tr>
</tbody>
</table>
Chapter 1.

**Introduction and specific aims:**

There are currently 800,000 strokes reported annually in the United States, and stroke is responsible for 1 in every 19 mortalities. For those who survive, stroke constitutes a life altering event, with over 50% of patients requiring significant inpatient rehabilitation or skilled nursing assistance upon hospital discharge. It is well established that early and accurate diagnosis improves acute ischemic stroke (AIS) outcome by increasing the probability of successful intervention. Unfortunately, the current tools available to clinicians for diagnosis of AIS are limited; diagnosis via assessment of overt patient symptoms lacks in accuracy, while diagnosis by neuroradiological imaging is inaccessible early in the acute phase of care. As a result, there has been a recent push for the identification of reliable peripheral blood stroke biomarkers which could be rapidly measured in either the field or emergency department to expedite diagnosis and improve access to interventional treatment.

The earliest stroke biomarker studies focused on the detection of brain-specific proteins which are released into circulation as a result of brain damage. Unfortunately, a majority of biomarkers examined in these studies failed to demonstrate clinically useful levels of sensitivity and specificity at early enough time points in AIS progression to provide actionable diagnostic information. The reason this strategy may have failed to identify biomarkers which are diagnostically robust with regards to acute triage may lie in that the presence of brain-specific peptides in peripheral circulation following ischemic insult is likely largely dependent on apoptotic and necrotic neuronal cell death, processes which do not fully progress until as late as 24 hours after symptom onset. These early studies highlighted the necessity to target biological processes which show peripheral evidence of stroke-induced alterations relatively early in AIS progression.

Stroke triggers an immediate systemic inflammatory response which results in a dramatic shift in the phenotype of the peripheral immune system. The innate arm of the peripheral immune system undergoes rapid expansion, as monocyte and neutrophil counts increase significantly in the hours following stroke onset. Simultaneously, the adaptive arm of the immune system shifts into a state of suppression characterized by lymphopenia and limited responsiveness. The relatively rapid nature of this response suggests that the peripheral immune system may be an ideal source of biomarkers for early identification of AIS. Expectedly, proof of concept genome-wide transcriptomic studies have demonstrated that gene expression in peripheral blood is altered in the stroke pathology as a result of the peripheral immune response. However, these early stroke transcriptomics studies relied on methods commonly employed for basic differential expression analysis which may not be well suited for biomarker selection; while candidate markers identified in these studies showed promise, they failed to demonstrate levels of diagnostic robustness which would suggest they could be clinically useful. The biomarker panels identified in these early studies were comprised of candidate biomarkers discretely selected from genome-wide expression data based on arbitrary fold change cut-offs and statistical significance in basic parametric testing; such strategies may be inadequate for the identification of biomarkers, as such selection criteria do not directly infer discriminatory or prognostic
performance. Thus, strategies which select groups of candidate biomarkers based directly on discriminatory or prognostic ability may prove more effective in identifying transcriptional biomarkers for the diagnosis of AIS. Machine learning strategies may represent an optimal approach for biomarker identification based on such selection criteria, as many foundational machine learning algorithms were developed for classification and pattern recognition tasks.

Analysis of high-dimensional gene expression data using a pattern recognition approach known as genetic algorithm k-nearest neighbors (GA/kNN) has been successfully used in a small number of cancer studies to identify diagnostically relevant biomarker panels with strong discriminatory ability. While this strategy has proven robust in its applications in the field of cancer, it has yet to be utilized for biomarker discovery in the realm of cardiovascular disease. We hypothesized that a similar approach combining high-throughput transcriptomic profiling of the peripheral immune system with machine learning analysis would allow for the identification of candidate AIS biomarkers with strong enough diagnostic performance to have potential for future clinical implementation. Furthermore, we felt that exploration into the pathological role of candidate markers identified with such an approach could not only increase their clinical utility, but also provide valuable insight into the mechanisms which drive stroke immunopathology. Therefore, this work had two aims:

**Aim 1:** Use high-throughput transcriptomics in tandem with GA/kNN analysis to identify a pattern of gene expression in peripheral blood capable of accurately detecting ischemic stroke during the acute phase of care.

**Aim 2:** Evaluate the role of stroke-specific peripheral blood gene expression patterns within the context of stroke immunopathology.
Chapter 2.

Background and significance:

Stroke:

Stroke is the overarching classification for a cluster of cardiovascular pathologies in which cerebral blood supply becomes disrupted; if this disruption in blood flow becomes prolonged, neural tissues experience oxygen and glucose deprivation, which can ultimately result in neuronal cell death and permanent loss of function. About 10% of stroke cases are hemorrhagic in nature, in which blood flow disruption is a result of a ruptured intracerebral blood vessel and subsequent intracerebral bleeding. The remaining number of stroke cases are ischemic in nature, in which blood flow is disrupted as a result of a thrombotic or embolic cerebrovascular occlusion. Collectively, stroke is the fifth leading cause of death in the United States, and the leading cause of disability. Currently, the direct health care costs associated with stroke are approaching 80 billion dollars annually, a large proportion of which are expenditures associated with post-hospital rehabilitation.\(^1\)

A need for stroke biomarkers:

It is well established that the primary mechanism by which ischemic stroke outcome can be improved is through timely diagnosis and reperfusion of ischemic tissue via interventional therapies. Currently, there are only two FDA approved interventional strategies for AIS which have shown significant patient benefit: thrombolysis via intravenous administration of recombinant tissue plasminogen activator (rtPA), and physical recannulation via endovascular mechanical thrombectomy.\(^{24}\) Despite the fact that these interventions are known to limit primary tissue damage and reduce rates of morbidity and mortality, less than five percent of AIS patients receive treatment.\(^{25}\) One of the most promising avenues by which treatment rates could be increased is through improvements in the diagnostic process.

Due to the risk of life threatening adverse events, a definitive diagnosis of ischemic stroke is required prior to either thrombolytic or endovascular intervention.\(^{26}\) Currently, the only way to obtain a definitive diagnosis is through neuroradiological imaging via computerized tomography (CT) or magnetic resonance imaging (MRI). Unfortunately, these imaging techniques are inaccessible in the field, and often not immediately available in hospitals without dedicated stroke centers, such as smaller facilities and those which serve rural areas.\(^{27}\) This is problematic in that the benefits of interventional therapies are time limited, and they are only accessible at a select number of hospitals. Because the efficacy of rtPA decreases and the risk of adverse events increases as ischemic stroke pathology progresses, it is currently recommended that rtPA be administered within three and a half hours of symptom onset.\(^{28}\) While the therapeutic time window is not as limited for endovascular intervention is for thrombolytic therapy, endovascular thrombectomy still must be performed within seven hours of symptom onset to show significant benefit.\(^{29}\) Furthermore, because both thrombolytic and endovascular interventions require specialized personal, they are only performed at a limited number of hospitals which carry primary and
Thus, for patients to receive access to treatment, they must arrive at a stroke-ready hospital and receive a definitive diagnosis within a relatively short period of time.

In the prehospital setting, this can be facilitated through quick and confident recognition of stroke by first responders and emergency medical technicians, allowing for patients to be transported directly to hospital facilities which are designated for stroke care, increasing the chances of appropriate treatment and improved outcome. With regards to self-admitted patients and patients who are already in hospital care, the ability of hospital clinicians to confidently recognize stroke allows for immediate referral to a stroke team in the case of hospitals which have stroke center designations, or transfer to such a facility in the case of hospitals which are not equipped to deal with stroke. Unfortunately, inaccessibility to neuroradiological imaging means that these triage decisions are based on the assessment of overt patient symptoms using stroke recognition and severity scales such as the Cincinnati prehospital stroke scale (CPSS) and National Institutes of Health stroke scale (NIHSS). In the hospital setting, the ability to identify stroke with such assessments is highly inconsistent, with an estimated sensitivity ranging from 44-85%, and specificity ranging between 64-98%, depending on the scale used and experience of the clinician. The sensitivity and specificity of these assessments is even lower in the pre-hospital setting, where in some scenarios the ability to recognize stroke has been shown to be only marginally better than random.

The development of an accurate imaging-independent tool for diagnosis of AIS would allow clinicians to make more informed pre-hospital and in-hospital triage decisions, ultimately improving patient outcomes by expediting access to interventional treatment. One potential way such a tool could be developed is through the discovery of stroke-specific peripheral blood biomarkers which could be incorporated into rapid point of care tests. For this reason, there has been a substantial push for the identification of peripheral blood biomarkers with the potential for stroke diagnosis.

Early stroke biomarker studies:

Initial stroke biomarker studies focused largely on plasma detection of brain-specific proteins which are released into circulation as a result of brain damage, along with protein markers of thrombosis. Some of the most widely characterized of these candidate biomarkers include calcium-binding protein B (S100B), a glial-specific calcium binding protein, glial fibrillary acidic protein (GFAP), a filamentous protein produced predominantly by astrocytes, Von-Willebrand factor (vWF), a circulating glycoprotein involved in coagulation, and Fibrin D-dimer, a fibrin degradation product associated with thrombosis. Unfortunately, these and similar markers have failed to demonstrate clinically useful levels diagnostic accuracy at early enough time points in AIS progression to provide actionable diagnostic information, with even the most promising panels failing to achieve sensitivities and specificities approaching 90% when tested in a clinically relevant manner.

The reason this strategy may have been unsuccessful in identifying biomarkers which are diagnostically robust with regards to acute triage likely lies in that the presence of brain-specific peptides in peripheral
circulation following ischemic insult is likely largely dependent on apoptotic and necrotic neuronal cell death, processes which do not fully progress until as late as 24 hours after symptom onset. Thus, these early studies highlighted the necessity to target biological processes which show peripheral evidence of stroke-induced alterations relatively early in AIS progression.

The peripheral immune response as a source of stroke biomarkers:

It is becoming increasingly evident that ischemic stroke produces a robust systemic inflammatory response which results in a dramatic shift in the complexion of the peripheral immune system. Upon the onset of ischemia, signals initiated within the ischemic brain region result in near-immediate activation of innate peripheral immune system and simultaneous suppression of the adaptive peripheral immune system. Due to the rapid nature of this response, and the fact that it has a peripheral component, make it an appealing biological process to target in terms a potential source of stroke biomarkers.

At the onset of AIS, the local inflammatory cascade is initiated within the occluded cerebrovasculature. Vascular occlusion immediately results in intravascular ischemia, alterations in shear stress, and production of reactive oxygen species. Collectively, these physical and molecular changes trigger platelet activation, upregulation of adhesion molecules on endothelial cells, and the rapid release of proinflammatory cytokines by the cells of the neurovascular unit. Innate peripheral leukocyte populations in the immediate area of the insult respond to these collective stimuli with the release of matrix metalloproteinases (MMPs), while endothelial cells coordinately downregulate the expression of junctional proteins, ultimately resulting in increases in pericellular permeability of the endothelium. This increase in pericellular permeability allows for the extravasation of peripheral immune cells and peripheral proteins into the perivascular space. Within the perivascular space, hypoxia and increased accumulation of proinflammatory cytokines trigger the activation of local macrophages and mast cells, which then perpetuate the inflammatory cascade, promoting further infiltration of peripheral immune populations. While the immune response observed within the cerebrovasculature and perivascular space occurs almost immediately upon blood flow disruption, the immune response which occurs within the affected parenchymal brain tissue is slightly delayed, as it is highly dependent on the progression of ischemia-induced cell damage and activation of local microglial populations.

As ischemia initially spreads within the parenchyma, nucleotides such as adenosine triphosphate (ATP), as well as neurotransmitters, are released from neurons due to stress-induced depolarization. As a result, extracellular ATP and neurotransmitter levels rise dramatically within minutes of ischemic onset, priming local microglial populations. As ischemia continues to progress, prolonged glucose and oxygen deprivation trigger the onset of neuronal cell death, during which the structural integrity of the plasma membrane is compromised, and intracellular molecules which are associated with cellular damage are released into the extracellular space. These damage-associated molecules provide activating signals to microglia via various pattern recognition and scavenger receptors; in addition to ATP, other intracellular molecules containing purines such as uracil triphosphate (UTP) bind and activate purinoreceptors. Furthermore, intracellular proteins such as high mobility
group box 1 (HMGB1), amyloid beta (Aβ), and heat shock proteins (HSPs) bind and activate toll-like receptors (TLRs). The cumulative effect of these activating signals is increased activity of pro-inflammatory intracellular signaling pathways such as the nuclear factor kappa-B (NF-kB) pathway, which drive M1 microglial polarization and production of proinflammatory cytokines and chemokines. Pro-inflammatory cytokines and chemokines released by microglia activate local astrocytes and promote the migration of peripheral immune cells from the perivascular space into the damaged parenchyma.

Infiltrating innate immune cells, including neutrophils and monocyte-derived M1 activated macrophages, provide a population of phagocytes which will ultimately play a role in the clearance of apoptotic and necrotic debris, which is necessary for later repair and remodeling of the lesion. The inflammatory milieu within and around the damaged brain tissue initiates sympathetic nervous system signaling to peripheral immune organs, allowing for a coordinated peripheral response which can further support the innate immune activities occurring within the ischemic region. Sympathetic innervation of bone marrow triggers the release of newly-formed myeloid immune cells into peripheral circulation while sympathetic signaling to the spleen promote the release of sequestered neutrophils and monocytes. These newly-circulating innate immune cells are then free migrate to the ischemic region and contribute to the population of infiltrating peripheral innate immune cells, driven by chemotactic gradients originating from the ischemic tissue.

While infiltrating innate peripheral immune populations play a necessary role in the acute response to injury, excessive infiltration of peripheral adaptive immune populations can be pose a threat which has the potential to dramatically exacerbate the injury. In the absence of neurodegenerative disease or neural trauma, the peripheral adaptive immune system does not come in contact with central nervous system (CNS) antigens. Disruption of the blood brain barrier and the glia limitans during the acute inflammatory phase of stroke, as well as chemotactic gradients generated via the local immune response, not only allow for the recruitment of peripheral innate immune populations into the infarct region, but also allow for the infiltration of peripheral lymphocytes. Major histocompatibility (MHC) molecules and costimulatory proteins required for antigen presentation are upregulated on activated macrophages and microglia within the damaged tissue; phagocytosis of neuronal debris and subsequent presentation of neuronal antigens by these innate immune populations has the potential to activate infiltrating peripheral lymphocytes and trigger an adaptive autoimmune response. The presence of lymphocyte infiltrates in the brain and elevated titers of antibodies against brain-specific antigens have been observed months after injury in some cases. In these scenarios, this long-term adaptive response helps maintain an inflammatory environment at the site of injury, which promotes secondary tissue damage via prolonged innate immune system activity.

As a neuroprotective mechanism to limit such an adaptive immune response, the peripheral immune system shifts into a state of suppression. Sympathetic signaling initiated in the ischemic brain to peripheral lymphoid organs promotes the sequestration of adaptive immune populations, and circulating lymphocyte counts drop in the hours following ischemic onset. In addition to sequestration, peripheral lymphocytes
undergo functional changes as a result of ischemic insult which make them resistant to antigen-stimulated activation; lymphocytes obtained from stroke patients have been shown to exhibit impaired proliferation and altered cytokine production when stimulated with common pathogen-derived antigens.\textsuperscript{14,49–51} While the exact molecular mechanisms which drive these functional changes remain unclear, they are believed to be at least in part driven by stroke-induced alterations in circulating factors, as serum from stroke patients has been shown to inhibit the proliferation of lymphocytes derived from healthy donors.\textsuperscript{52}

This stroke induced-shift in the state of the peripheral immune system towards an innate-activated, adaptive-suppressed phenotype occurs relatively quickly following ischemic onset, as alterations in peripheral white blood cell differential are often already observable at emergency department admission.\textsuperscript{12–14} The rapidness of this response, in combination with the fact that peripheral immune cells are readily assessable via liquid biopsy, makes the peripheral immune system a potential source of robust ischemic stroke biomarkers.

**Preliminary stroke biomarker studies targeting the peripheral blood transcriptome:**

Initial studies investigating the peripheral immune response to stroke as a potential source of biomarkers chose to do so via microarray-based genome-wide expression profiling. This represents strong strategy, in that it allows for the parallel analysis of thousands of potential RNA biomarkers, increasing the likelihood of identifying transcripts with strong diagnostic potential.

In the first of these studies, Moore \textit{et al.}\textsuperscript{17} identified hundreds of transcripts which were differentially regulated in isolated peripheral blood mononuclear cells (PBMCs) obtained from a small cohort of AIS patients relative to those obtained from neurologically asymptomatic healthy control subjects. Furthermore, they were able to demonstrate that the coordinate expression levels of 22 of these differentially regulated genes could discriminate between AIS patients and control subjects with 78% sensitivity and 80% specificity. While the results of this study demonstrated strong proof of principle, the fact that gene expression profiling was performed in PBMCs limited potential translation of the identified panel for diagnostic use in the acute care setting, as the preparation of PBMCs is time consuming and requires laboratory equipment.

A pilot study performed by Tang \textit{et al.}\textsuperscript{18} and a subsequent validation study performed by Stamnova \textit{et al.}\textsuperscript{19} addressed this limitation by performing gene expression profiling in whole blood samples obtained from AIS patients and neurologically asymptomatic healthy control subjects at 3, 5, and 24 hours post symptom onset. Unfortunately, gene expression data from a single time point was not adequate to differentiate between AIS patients and control subjects with noteworthy levels of accuracy, however, a panel of 18 candidate genes was described which demonstrated the ability to discriminate between groups with 93.5% sensitivity and 89.5% specificity using combined expression data generated from all three blood draws. Several of these 18 genes were further confirmed as being differentially regulated in AIS in an independent investigation by Barr \textit{et al.},\textsuperscript{16} however their diagnostic utility was not assessed. While these studies once again provided promising proof of
principle, the necessity to obtain multiple blood samples to achieve high levels of sensitivity and specificity limited this candidate biomarker panel with regards to implementation for acute stroke triage.

Collectively, these preliminary studies demonstrated that the transcriptome of the peripheral immune system is responsive to stroke, and that stroke-induced changes in peripheral blood gene expression can be implemented for diagnostic purposes. Technical limitations aside, in terms of diagnostic accuracy, the candidate biomarker panels identified these early transcriptomic studies dramatically outperformed a wide majority of the candidate protein biomarkers which had been investigated previously. However, these panels were still far from being diagnostically perfect, despite only being tested against healthy control subjects absent of cardiovascular disease (CVD) risk factors.

**Machine learning for biomarker discovery:**

One major reason these early panels of transcriptional biomarkers may have failed to demonstrate robust diagnostic ability may lie in how they were selected. The panels identified in these initial stroke transcriptomics studies were comprised of candidate biomarkers discretely selected from genome-wide expression data based on arbitrary fold change cut-offs and statistical significance in parametric testing. Such strategies are commonly employed in differential expression studies, however may be inadequate for the identification of biomarkers, as such selection criteria do not directly infer discriminatory or prognostic performance. Furthermore, such selection techniques do little to inform about the collective diagnostic potential the candidate markers, as each candidate marker of the resultant panel is being selected as part of a discrete evaluation. Thus, strategies which select groups of candidate biomarkers based directly on collective discriminatory or prognostic ability may prove more effective in identifying transcriptional biomarkers for the diagnosis of AIS. Machine learning strategies may represent an optimal approach for biomarker identification based on such selection criteria, as many foundational machine learning algorithms were developed for classification and pattern recognition tasks.²⁰

Analysis of high-dimensional gene expression data using a pattern recognition approach known as genetic algorithm k-nearest neighbors (GA/kNN) has been successfully used in a small number of cancer studies to identify diagnostically relevant biomarker panels with strong discriminatory ability.²¹–²³ The GA/kNN approach combines a powerful search heuristic, GA, with a non-parametric classification method, kNN. In GA/kNN analysis, a small combination of genes (referred to as a chromosome) is generated by random selection from the total pool of gene expression data (Figure 1A). The ability of this randomly generated chromosome to discriminate between sample classes is then evaluated using kNN. In this evaluation, each sample is plotted as a vector in a multi-dimensional feature space where the coordinates of the vector are comprised of the expression levels of the genes of the chromosome. The class of each sample is then predicted based on the majority class of the nearest neighbors, or other samples which lie closest in Euclidian distance within the feature space (Figure 1B). The ability of the chromosome to discriminate between classes is quantified as a fitness score, or the proportion of samples which the chromosome is correctly able to classify. A termination cutoff (minimum proportion of correct classifications) determines the level of fitness required to pass evaluation. A chromosome
which passes kNN evaluation is labeled as a near-optimal solution and recorded, while a chromosome which fails undergoes repeated cycles of mutation and re-evaluation until a near-optimal solution is found (Figure 1A). This entire search paradigm is performed multiple times (typically hundreds of thousands) to generate a heterogeneous pool of near-optimal solutions (Figure 1C). The discriminatory ability of each gene is then ranked according to the number of times it appears in the near-optimal solution pool (Figure 1D), and the collective discriminatory ability of the top ranked genes can then be tested via kNN in a leave one out cross validation (Figure 1E). This approach has been utilized to generate biomarker panels capable of optimally discriminating...
between cancerous and non-cancerous colon biopsies, \textsuperscript{23} primary and metastatic melanoma tumors, \textsuperscript{21} as well as between B-cell lymphoma sub-types, \textsuperscript{22} all with accuracies ranging between 95 and 100%.

While this approach combining high-throughput transcriptomics with GA/kNN analysis has proven robust in several applications in the field of cancer, it has yet to be utilized for biomarker discovery in the realm of cardiovascular disease (CVD). We hypothesized that a similar approach combining high-throughput transcriptomic profiling of the peripheral immune system with machine learning analysis would allow for the identification of candidate AIS biomarkers with strong enough diagnostic performance to have potential for future clinical implementation. Furthermore, we felt that exploration into the pathological role of candidate markers identified with such an approach would not only increase their clinical utility, but also provide valuable insight into the mechanisms which drive stroke immunopathology.
Chapter 3.

Results:

Machine learning for identification of candidate ischemic stroke biomarkers:

In order to identify transcriptional signatures within the peripheral immune system with the potential for identification of AIS, we first recruited a modestly-sized discovery cohort comprised of 39 imaging-confirmed AIS patients and 24 neurologically asymptomatic control subjects (Table 1). Peripheral whole blood was sampled from patients at emergency department admission, and genome-wide expression profiling was performed via microarray. Gene expression data were subjected to GA/kNN analysis, and genes were ranked based on the ability of their expression levels to discriminate between AIS patients and controls, according to the number of times they were selected as part of a near-optimal solution (Figure 2A). The expression levels of top 50 genes identified by GA/kNN displayed a strong ability to discriminate between groups using kNN in leave one out cross validation; the combined expression levels of as few as four of the top 50 ranked genes were able to identify subjects in the discovery cohort the with levels of accuracy in excess of 98% (Figure 2B).

We then wanted to confirm that the top genes identified in our GA/kNN analysis that were truly optimal in terms of discriminatory ability. To do this, we compared the ability of the expression levels of top 50 genes selected by GA/kNN to differentiate between stroke patients and controls to that of genes selected at random. Specifically, we compared the accuracy of GA/kNN-selected genes to the accuracy of 50 sets of 50 genes randomly generated from the total pool of gene expression data, as well to the accuracy of 50 sets of 50 genes randomly selected from a sub-pool of genes which displayed greater than 1.7 fold differential regulation between groups. The top genes selected by GA/kNN performed significantly better than genes selected at random genome-wide, as well as significantly better than genes selected at random from those which were differentially regulated greater than 1.7 fold (Figure 2C). Collectively, the results of this analysis, in combination with the levels

Table 1. Discovery cohort clinical and demographic characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic (n=24)</th>
<th>Ischemic stroke (n=39)</th>
<th>Stat (df)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>59.9 ± 9.7</td>
<td>73.1 ± 14.0</td>
<td>t = -4.40 (61)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>14 (58.3)</td>
<td>22 (56.4)</td>
<td>χ² = 0.12 (1)</td>
<td>0.731</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>0 ± 0.0</td>
<td>5.3 ± 6.4</td>
<td>t = 5.17 (38)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>4 (16.7)</td>
<td>15 (38.5)</td>
<td>χ² = 7.02 (1)</td>
<td>0.008*</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>7 (29.2)</td>
<td>25 (64.1)</td>
<td>χ² = 11.2 (1)</td>
<td>0.001*</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>0 (0.00)</td>
<td>18 (46.2)</td>
<td>χ² = 15.5 (1)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>2 (8.30)</td>
<td>11 (28.2)</td>
<td>χ² = 3.58 (1)</td>
<td>0.058</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>2 (8.30)</td>
<td>6 (15.4)</td>
<td>χ² = 0.67 (1)</td>
<td>0.414</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>0 (0.00)</td>
<td>6 (15.4)</td>
<td>χ² = 4.08 (1)</td>
<td>0.043*</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>0 (0.00)</td>
<td>6 (15.4)</td>
<td>χ² = 4.08 (1)</td>
<td>0.043*</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>8 (33.3)</td>
<td>29 (74.4)</td>
<td>χ² = 10.3 (1)</td>
<td>0.001*</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>1 (4.20)</td>
<td>7 (17.9)</td>
<td>χ² = 2.55 (1)</td>
<td>0.111</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>5 (20.8)</td>
<td>17 (43.6)</td>
<td>χ² = 3.39 (1)</td>
<td>0.066</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>1 (4.20)</td>
<td>20 (51.3)</td>
<td>χ² = 14.9 (1)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>0 (0.00)</td>
<td>9 (23.1)</td>
<td>χ² = 6.46 (1)</td>
<td>0.011*</td>
</tr>
<tr>
<td>rtPA n (%)</td>
<td>2 (8.30)</td>
<td>2 (5.13)</td>
<td>χ² = 0.26 (1)</td>
<td>0.612</td>
</tr>
</tbody>
</table>

Continuous variables compared with student’s t-test (t), categorical variables compared with chi-squared test (χ²).
of accuracy observed, suggest that our biomarker discovery strategy was effective at selecting genes with optimal diagnostic potential in terms of the subjects of the discovery cohort. Because the use of genes beyond the top ten did not appear to improve overall accuracy (Figure 2B), and displayed diminishing diagnostic robustness relative to genes selected at random (Figure 2C), we chose to focus on only the top ten genes (ANTXR2, STK3, PDK4, CD163, MAL, GRAP, ID3, CTSZ, KIF1B, and PLXDC2) for the remainder of our analysis.

Figure 2. Top 50 genes selected by GA/kNN for identification of AIS. (A) The top 50 peripheral blood transcripts ranked by GA/kNN based on their ability to discriminate between AIS patients and neurologically asymptomatic controls in the discovery cohort. (B) Combined ability of the expression levels of top 50 genes selected by GA/kNN to discriminate between AIS patients and neurologically asymptomatic controls in the discovery cohort using kNN. (C) Ability of the expression levels of the top 50 genes selected by GA/kNN to discriminate between neurologically asymptomatic controls and AIS patients via kNN compared to the expression levels of genes selected at random. The accuracy of the top 10 genes selected by GA/kNN was specifically tested against the accuracy of randomly selected genes using single sample two-way t-test.
When comparing the peripheral blood expression levels of the top ten genes between AIS patients and controls, the magnitude of differential expression was modest in terms of fold change in the case of most genes, however differences in expression levels between groups were highly consistent across all subjects, which was reflected by high levels of statistical significance in parametric statistical testing (Figure 3A). The combined discriminatory power of the top ten genes was evident when their coordinate expression levels were plotted on a continuum for each individual subject; the overall pattern of expression was strikingly different between AIS patients and controls, and it was clear that the overall pattern of expression was more diagnostically powerful than the expression levels of any given gene on its own (Figure 3B).

Next, we wanted to determine whether potentially confounding intergroup differences in clinical and demographic characteristics (Table 1) influenced the expression levels of the top ten genes. In order to more intuitively explore the relationship between the pattern of gene expression observed across the top ten genes and these potentially confounding factors, we first used principal components analysis to describe the expression levels of the top ten genes as single composite RNA expression variable. The expression levels of the top ten genes were highly correlated, and a single principal component was able to capture 70% of the collective variation.
The result component scores (composite RNA expression) were strongly correlated with the expression levels of each of the individual candidate genes (Figure 4B), and visually appeared to summarize the gene expression pattern well (Figure 3C). Stroke, age, anticoagulant status, hypertension, antihypertensive status, dyslipidemia, history of myocardial infarction, and history of atrial fibrillation were then regressed against the composite RNA expression levels of the top ten genes using multiple regression. We subsequently performed variance decomposition via the Lindeman-Merenda-Gold (LMG) method to estimate the relative contributions of each regressor to the total variance in composite RNA expression explained by the resultant regression model. 

Figure 4. Use of principal components analysis for generation of a composite RNA expression variable summarizing the coordinate expression levels of the top ten genes. (A) Correlation matrix depicting the correlations between the expression levels of the top ten genes identified by GA/kNN in the discovery cohort. Strength of correlations were tested via Spearman’s rho. (B) Principal component loadings and coefficients used for generation of the composite RNA expression variable.

Figure 5. Influence of potentially confounding clinical and demographic characteristics on the expression levels of the top ten genes. (A) Multiple regression model generated by regressing potentially confounding clinical and demographic characteristics against the composite RNA expression levels of the top ten genes selected by GA/kNN in the discovery cohort. (B) Graphical representation of the relative contribution of each regressor towards the total variance in composite RNA expression explained by the model.
Stroke remained significantly associated with the composite RNA expression levels of the top ten genes after accounting for all potentially confounding factors included in the model (Figure 5A), and was responsible for a majority of the explained variance (77.9%, Figure 5B). In terms of potentially confounding factors, both anti-hypertensive status and anticoagulant status were significantly associated with the composite RNA expression levels of the top ten genes after accounting for all other regressors (Figure 5A), however these associations only accounted for a small amount of the variance in composite RNA expression explained by the model (6.5% and 4.5% respectively, Figure 5B). Results of this multiple regression analysis were supported by the results of a more traditional logistic regression analysis in which the composite RNA expression levels of the top ten genes were identified as the only significant predictor of stroke when considering the same potentially confounding covariates (Supplemental Table 1). Taken as a whole, these findings suggest that the pattern of differential expression observed across the top ten genes between groups is highly associated with stroke independently of the assessed potential confounding factors. While these findings do suggest that antihypertensive status and anticoagulant status may influence the expression levels of the top ten genes, the effect of this influence on expression levels are likely minimal relative to the effect of stroke, and intergroup differences in these factors were likely not significant drivers of the selection of these genes by GA/kNN.

We additionally used this composite RNA expression variable to examine the potential influence of stroke severity and time to blood draw on the pattern of gene expression observed across the top ten genes in the discovery cohort. The composite RNA expression levels of the top ten genes displayed a significant positive association with stroke severity as assessed by the NIHSS (Figure 6A), suggesting that the expression levels of the top ten genes are likely directly responsive to stroke pathology. We observed a weak nonsignificant negative relationship between the composite RNA expression levels of the top ten genes and the time from symptom

![Figure 6. Influence of stroke severity and time to blood draw on the coordinate expression levels of the top ranked genes in discovery cohort AIS patients.](image)

(A) Relationship between stroke severity, as assessed by NIHSS, and composite RNA expression levels of the top ten genes in discovery cohort AIS patients. (B) Relationship between time from symptom onset to blood draw and composite RNA expression levels of the top 10 genes in discovery cohort AIS patients, with indication of stroke severity. Strength of correlations were tested via Spearman’s rho.
onset to blood draw (Figure 6B). However, this negative relationship was likely driven by the influence of stroke severity, given that the composite expression levels of these genes were positively associated with stroke severity, and patients undergoing more severe strokes generally presented to the emergency department earlier than patients undergoing less severe strokes (Figure 6B). Collectively, these observations suggest that the stroke-induced differential expression of the top ten genes may have additional utility for the stratification of stroke severity, and is relatively temporally stable during the acute phase of care.

In order to evaluate the generalizability of the pattern of differential expression identified by GA/kNN in the discovery cohort, we next tested the diagnostic ability of the top ten ranked genes in an independent validation cohort enrolled via a second geographically and socioeconomically distinct clinical site (see experimental procedures section). This validation cohort was comprised of an additional 39 AIS patients along with two different control groups, one consisting of 30 neurologically asymptomatic controls, and the other consisting of 20 acute stroke mimics (Table 2). Peripheral blood samples were once again obtained from patients at emergency department admission, and the expression levels of the top ten genes identified by GA/KNN in the discovery cohort were measured via qRT-PCR.

The overall pattern of differential expression between AIS patients and asymptomatic controls observed across the top 10 genes in the discovery cohort was also seen when comparing AIS patients and asymptomatic controls in the validation cohort (Figure 7A). The strong ability of the top ten genes to differentiate between stroke patients and asymptomatic controls in the discovery cohort using kNN was also recapitulated in the validation cohort; the expression levels of the top ten genes used in combination were able to classify 95.6% of subjects correctly with a sensitivity of 92.3% and a specificity of 100% (Figure 7B).

When comparing AIS patients to stroke mimics, the overall pattern of differential expression observed across the top ten genes was identical to that observed when comparing AIS patients to asymptomatic controls,
however the magnitude of these expression differences was smaller in the case of several genes (Figure 7C). Despite this reduction in the magnitude of differential expression, the expression levels of the top ten genes used in combination were still able to accurately discriminate between AIS patients and stroke mimics, classifying 94.9% of subjects correctly with a sensitivity of 97.4% and a specificity of 90.0% (Figure 7D). However, it is important to note that it was evident that all ten genes were required to achieve high levels of diagnostic accuracy when comparing AIS patients to stroke mimics (Figure 7D), whereas similar levels of accuracy could be achieved with as few as the top four markers when comparing AIS patients to neurologically asymptomatic controls in both the discovery cohort (Figure 2B) and the validation cohort (Figure 7B). Despite this, the collective validation cohort results supported those of the discovery cohort, and provide further evidence that the top ten markers selected by GA/kNN in our analysis have high potential performance for identification of AIS.

Note: Full manuscript associated with these results can be found in Appendix, Manuscript 1.

Role of stroke-identifying biomarkers in immunopathology:

In order to gain a better understanding regarding the potential role of the top ten genes in the stroke-induced peripheral immune response, we next explored the relationship between their peripheral blood
expression levels and white blood cell differential within the validation cohort AIS group. For this analysis, we choose to summarize white blood cell differential in terms of the neutrophil-monocyte to lymphocyte ratio (NMLR), as it provides a single variable which can intuitively be used to infer peripheral immune status within the context of the known stroke-induced shift towards innate immune activation and adaptive immune suppression.

The whole blood expression levels of each of the seven genes which we observed as up-regulated in AIS were to some degree positively associated with NMLR (Figure 8A, 8B, 8C, 8D, 8H, 8I, 8J). Conversely, the whole blood expression levels of all three genes which we observed as being down-regulated in AIS were negatively associated with NMLR (Figure 8E, 8F, 8G). We hypothesized that this phenomenon was a result of two not-necessarily distinct possibilities: that changes in the expression of these genes observed with AIS at the whole blood level are simply a marker of stroke-induced shifts in white blood cell differential, or that changes in the transcription of these genes at the level of individual cells help drive the stroke-induced alterations in the phenotype of the peripheral immune system.

To begin to understand which of these of possibilities underlie the relationship between the expression levels of the top ten genes and white blood cell differential following AIS onset, we looked to identify where these genes are expressed in the peripheral immune system. To do this, we measured the basal expression levels of the ten candidate genes on six subpopulations of immune cells isolated from the peripheral blood of three healthy
donors. Interestingly, almost every gene exhibited some degree of lineage-specific expression. Each of the seven genes which we observed as being up-regulated in the peripheral blood in response to stroke were all predominantly expressed in cell populations of myeloid origin (Figure 9A, 9B, 9C, 9D, 9H, 9I, 9J). Conversely, all of genes which we observed as being downregulated were predominantly expressed in cells of lymphoid origin (Figure 9E, 9F, 9G). In several cases, the difference in expression between myeloid and lymphoid populations was quite substantial, in excess of hundred-fold.

Taken as a whole, the relationships we observed with NMLR, along with the strong pattern of lineage-specific expression, suggest that the expression levels of these genes in whole blood likely serve as markers of the collective cellularity of the peripheral immune system. This insinuates that the changes in the expression of these genes which we observed in AIS in whole blood are reflective of stoke-induced shifts in white blood cell count, and likely not a result of alterations in transcription at the level of individual cells. This ultimately suggests

Figure 9. Expression levels of the ten candidate transcriptional biomarkers on isolated leukocyte sub-populations. Expression levels of (A) ANTXR2, (B) STK3, (C) PDK4, (D) CD163, (E) MAL, (F) GRAP, (G) ID3, (H) CTSZ, (I) KIF1B, and (J) PLXDC2 in peripheral whole blood and isolated leukocyte populations obtained from one male and two female healthy donors. Expression levels are presented as fold change relative to the population of lowest expression, and were statistically compared via one-way ANOVA.
that many of these genes may not play a substantial mechanistic role as drivers of stroke immunopathology, as genes which mechanistically drive the changes observed in the peripheral immune system following stroke are likely ones which undergo true stroke-induced differential transcription. A scenario in which this may not hold true however, is in the case of genes which encode for secreted proteins; circulating concentrations of secreted lineage-specific proteins may become altered by virtue of changes in cell count alone, independently of any change in transcription or translation at the level of individual cells.

Note: Full manuscript associated with these results can be found in Appendix, Manuscript 2.

**CD163 as a potential modulator of the stroke-induced peripheral immune response:**

While a majority of the top ten genes identified in our analysis encode for intracellular proteins (Summarized in Table 3), one gene in particular, **CD163**, encodes for a protein known as cluster of differentiation 163 (CD163) which has the potential for secretion.\(^{54}\) CD163 is a membrane-bound scavenger receptor for extracellular hemoglobin which is believed to be expressed exclusively within the innate immune system,\(^ {55}\) where it is predominantly found on monocytes and macrophages,\(^ {56}\) and to a lesser extent on neutrophils.\(^ {57}\) Various stimuli can trigger CD163 ectodomain shedding via cleavage by the metalloprotease ADAM metallopeptidase domain 17 (ADAM17),\(^ {58,59}\) resulting in generation of a soluble truncated peptide (sCD163) which has been shown in multiple studies to directly interact with lymphocytes and inhibit antigen-induced proliferation.\(^ {60–62}\) Interestingly, other groups have reported elevations in ADAM17 activity in animal models of ischemic brain injury,\(^ {63–65}\) as well as stroke-induced increases in peripherally circulating levels of various ADAM17 substrates such as tumor necrosis factor alpha (TNFα) in human subjects.\(^ {66,67}\) Therefore, it is possible that stroke induces a rise in peripherally circulating levels of sCD163 via coordinate increases in the number of CD163 expressing innate immune cells and elevations in ADAM17 activity; such a rise in sCD163 levels could subsequently contribute to suppression of the adaptive immune system via the inhibitory effects of sCD163 on lymphocyte activity.

In order to address this hypothesis, we first assessed peripheral blood ADAM17 activity and circulating sCD163 levels in the peripheral blood of validation cohort subjects via enzyme activity assay and ELISA, respectively. ADAM17 activity was significantly elevated in the peripheral whole blood of AIS patients relative to that of both neurologically asymptomatic controls and acute stroke mimics (Figure 10C). Likewise, plasma levels
of sCD163 were significantly higher in AIS patients relative to both control groups (Figure 10D). Furthermore, both ADAM17 activity and plasma sCD163 levels displayed positive associations with stroke severity in terms of NIHSS (Figure 10E, 10F) and infarct volume (Supplemental Figure 1A, 1B); while our relatively small sample size and the variability inherent to such measures of injury status kept these associations from reaching statistical significance, the relationships observed provide tentative evidence that both factors are potentially directly responsive to stroke pathology.
To explore whether the elevations in sCD163 levels which we observed in AIS potentially play a role in modulation of the stroke-induced peripheral immune response, we next assessed the relationship between circulating soluble sCD163 levels and white blood cell counts within the validation cohort AIS group. Plasma sCD163 levels were significantly negatively associated with absolute lymphocyte counts (Figure 11A), as well as significantly positively associated with both absolute monocyte counts (Figure 11B) and absolute neutrophil counts (Figure 11C). Collectively, these results provided in vivo associative evidence which supported our hypothesis that stroke-induced increases in myeloid-derived circulating sCD163 may contribute to post-stroke suppression of the adaptive immune system.

Next, we wanted to determine the direct effects of the post-stroke peripheral inflammatory milieu on ADAM17-mediated sCD163 shedding by peripheral blood innate immune cell populations. To do this, primary neutrophil and monocyte cultures generated from the peripheral blood of twelve healthy donors were treated with pooled serum samples derived from a subset of ten AIS patients, ten neurologically asymptomatic controls, and ten acute stroke mimics which were relatively well matched in terms of clinical and demographic characteristics (Table 4). Serum incubation was performed in both the presence and absence of the ADAM17 inhibitors TAPI-0 and Marimastat, and cellular ADAM17 activity and sCD14 production were assessed following treatment.

Neutrophil cultures across all treatment conditions appeared largely phenotypically identical when visually inspected following treatment (Figure 12A). ADAM17 activity did appear to be elevated in neutrophil cultures treated with serum obtained from AIS patients relative to those treated with serum obtained from both neurologically asymptomatic controls and acute stroke mimics, however this increase was not statistically significant (Figure 12B). Overall, neutrophil cultures appeared to generate relatively little sCD163, although a limited number of cultures treated with AIS serum did appear to exhibit increased sCD163 production in response to treatment. This effect was not widespread however, and no statistically significant differences were observed

Table 4. Clinical and demographic characteristics of subject sub-populations used for in vitro experiments.

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic (n=10)</th>
<th>Stroke mimic (n=10)</th>
<th>Ischemic stroke (n=10)</th>
<th>Stat (df)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>57.5 ± 11.1</td>
<td>56.4 ± 17.1</td>
<td>63.5 ± 14.3</td>
<td>F= 0.63 (2, 27)</td>
<td>0.536</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>8 (80.0)</td>
<td>5 (50.0)</td>
<td>5 (50.0)</td>
<td>χ² = 2.50 (2)</td>
<td>0.266</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>0.0 ± 0.0</td>
<td>1.6 ± 2.4</td>
<td>6.6 ± 6.9</td>
<td>F= 5.96 (2, 27)</td>
<td>0.007*</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>5 (50.0)</td>
<td>2 (20.0)</td>
<td>4 (40.0)</td>
<td>χ² = 2.01 (2)</td>
<td>0.366</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>9 (90.0)</td>
<td>8 (80.0)</td>
<td>8 (80.0)</td>
<td>χ² = 0.48 (2)</td>
<td>0.766</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>5 (50.0)</td>
<td>7 (70.0)</td>
<td>4 (40.0)</td>
<td>χ² = 1.87 (2)</td>
<td>0.391</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>χ² = 0.00 (2)</td>
<td>1.000</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>χ² = 0.57 (2)</td>
<td>0.749</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>0 (00.0)</td>
<td>2 (20.0)</td>
<td>3 (30.0)</td>
<td>χ² = 3.36 (2)</td>
<td>0.186</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>0 (00.0)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>χ² = 2.22 (2)</td>
<td>0.329</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>9 (90.0)</td>
<td>8 (80.0)</td>
<td>7 (70.0)</td>
<td>χ² = 1.25 (2)</td>
<td>0.535</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>χ² = 0.00 (2)</td>
<td>1.000</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>2 (20.0)</td>
<td>7 (70.0)</td>
<td>3 (30.0)</td>
<td>χ² = 5.83 (2)</td>
<td>0.054</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>1 (10.0)</td>
<td>5 (50.0)</td>
<td>6 (60.0)</td>
<td>χ² = 5.83 (2)</td>
<td>0.054</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>χ² = 0.57 (2)</td>
<td>0.749</td>
</tr>
<tr>
<td>tPA n (%)</td>
<td>0 (00.0)</td>
<td>0 (00.0)</td>
<td>3 (30.0)</td>
<td>χ² = 6.67 (2)</td>
<td>0.035*</td>
</tr>
</tbody>
</table>

Continuous variables compared with one-way ANOVA (f), categorical variables compared with chi-squared test (χ²)
in sCD163 production between cultures treated with AIS serum and those treated with serum from control groups (Figure 12C). A second independent experiment using pooled serum samples derived from a separate sub-set

Figure 12. Effects of ischemic stroke serum on neutrophil ADAM17-dependant sCD163 production. (A) Morphology of healthy donor-derived neutrophils following three hours incubation with 10% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. (B) ADAM17 activity in neutrophil lysates collected following treatment. (C) Neutrophil-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferonni post-hoc.

Figure 13. Effects of ischemic stroke serum on monocyte ADAM17-dependant sCD163 production. (A) Morphology of healthy donor-derived monocytes following three hours incubation with 20% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. White arrowheads indicate cytoplasmic vacuoles. (B) ADAM17 activity in monocyte lysates collected following treatment. (C) Monocyte-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferonni post-hoc.
of subjects (Supplemental Table 2) yielded similar results (Supplemental Figure 2).

Counter to what was observed in terms of neutrophil cultures, there were striking visual differences between monocyte cultures treated with serum derived from AIS patients and those treated with serum derived from control groups. Cultures treated with control serum generally displayed a mix of adherent and semi-adherent cells, however several cultures incubated with AIS serum contained large patches of cells which appeared highly adherent with prominent cytoplasmic vacuoles characteristic of a shift towards a macrophage phenotype (Figure 13A). ADAM17 activity was significantly elevated in monocytes treated with AIS serum relative to those treated with serum from both control groups (Figure 13B). Furthermore, supernatants recovered from cultures treated with AIS serum contained significantly higher concentrations of monocyte-derived sCD163 than those recovered from cultures treated with serum from either control group, and this effect was largely ablated by inhibition of ADAM17 (Figure 13C). Once again, similar results were observed in a second independent experiment (Supplemental Figure 3) using pooled serum samples derived from a separate sub-set of subjects (Supplemental Table 2). Collectively, these observations demonstrate that soluble factors present in peripheral circulation following ischemic stroke have the capacity to trigger ADAM17-dependant sCD163 shedding from peripheral blood monocytes, however likely not from neutrophils.

Lastly, we wanted to determine the direct effect of post-stroke-elevations in circulating sCD163 levels on the capacity of the peripheral blood to support lymphocyte proliferation. To do this, PHA-stimulated proliferation of healthy-donor derived lymphocytes was assessed via BrdU incorporation in the presence of pooled serum obtained from control groups, pooled serum from AIS patients, and pooled serum obtained from AIS patients which had been depleted of sCD163 by immunoprecipitation (Supplemental Figure 4). Consistent with prior

Figure 14. Influence of post-stroke peripheral blood sCD163 levels on the capacity to support lymphocyte proliferation. (A) Morphology of healthy donor-derived lymphocytes following 72 hours of PHA-stimulated proliferation in the presence of 20% serum obtained from AIS patients and controls, which was either unmanipulated in terms of sCD163 levels, or depleted of sCD163 via immunoprecipitation using anti-CD163 monoclonal antibody. White arrowheads indicate clusters of blasting lymphocytes. (B) BrdU incorporation over the final 24 hours of treatment. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferroni post-hoc.
visual observations suggested decreased proliferative activity in lymphocyte cultures treated with unmanipulated AIS serum relative to those treated with serum from control groups, as they appeared to contain smaller and more diffuse clusters of blasting lymphocytes (Figure 14A). However, visual indications suggested that the inhibitory effect of AIS serum was at least moderately counteracted by the depletion of cCD163 (Figure 14A). In agreement with these visual observations, we observed significantly lower levels of BrDU incorperation in lymphocytes stimulated in the presence of AIS serum relative to those which were stimulated in the presence of serum derived from control groups, an effect which was partially ablated as a result of sCD163 depletion (Figure 14B). A second independent experiment using pooled serum samples derived from a separate sub-set of subjects (Supplemental Table 2) yielded similar results (Supplemental Figure 5).

Taken together, these results novelly suggest that post-stroke elevation in circulating levels of monocyte-derived sCD163 can have direct inhibitory effects on the proliferative capacity of lymphocytes, inferring a unique synergistic mechanism directly linking stroke induced activation of the peripheral innate immune system and suppression of the peripheral adaptive immune system.

Note: Full manuscript associated with these results can be found in Appendix, Manuscript 3.
Chapter 4.

Discussion:

The primary objective of this work was to implement a biomarker discovery strategy employing high throughput transcriptomics in tandem with machine learning analysis to identify patterns of gene expression in peripheral whole blood with the potential to detect AIS, and to further determine the role of such patterns within the context of stroke immunopathology. The ten transcriptional markers identified by GA/kNN in our analysis proved robust in their combined ability to differentiate between AIS patients and controls in both the discovery cohort and the independent validation cohort; not only did these markers display levels of diagnostic accuracy which exceed those reported in a majority of previous stroke biomarker studies, they also demonstrated characteristics which suggest they have the potential to be clinically useful. Furthermore, we were able to discern that many if not all these ten markers are likely differentially regulated in the peripheral whole blood of AIS patients as markers of stroke-induced shifts in white blood cell differential. However, our findings did suggest that one of these ten markers, CD163, may play a role in modulation of stroke-induced peripheral immune phenotype via a unique mechanism linking the coordinate responses of the innate and adaptive immune systems.

Machine learning for identification of candidate ischemic stroke biomarkers:

With regards to the countless number peripheral blood biomarker explorations which have been performed to date,5–9 to our knowledge, only one prior investigation has reported similar levels of diagnostic accuracy to those which we observed in this study in terms of discriminating between stroke patients and clinically relevant control populations. Dambinova et al.68 recently reported that plasma levels of brain-derived NR2 peptide, a degradation product of N-methyl-D-aspartate (NMDA) receptor cleavage, could be used to differentiate between stroke patients and a combination of acute stroke mimics and neurologically asymptomatic controls with 92% sensitivity and 96% specificity. However, a majority of blood samples in this prior study were obtained between 24 and 72 hours post symptom onset, and it is currently unknown if NR2 peptide would exhibit an equivalent level of diagnostic performance early in the acute phase of care. The ten marker panel identified in our analysis was tested earlier in the progression of pathology, and thus exhibits an obvious advantage in that it has the potential to provide actionable diagnostic information at an early enough time point to influence critical triage decisions which impact outcome.

The ten marker panel identified in our analysis displayed several favorable characteristics which could make it well suited for identification of ischemic stroke in the acute care setting. Most notably, the pattern of differential expression we observed between AIS patients and controls appeared to be relatively temporally stable. This is of clinical relevance from the standpoint that it is well established that acute stroke patients tend to arrive to the emergency department in two waves, the first within four hours from symptom onset (typically patients with more severe overt symptoms), and the second more than eight hours from symptom onset (typically patients with milder symptoms).69 For this reason, a potential diagnostic for identification of acute stroke needs
to be diagnostically robust across a wide time window with regards to the progression of stroke pathology. Another diagnostically beneficial characteristic we observed was that the stroke-associated pattern of expression across these ten markers was positively correlated with the NIHSS. Thus, these markers may have utility in stratifying injury severity, information which is commonly considered when making decisions regarding the prescription of interventional treatment.\(^{26}\) These characteristics, along with the fact that we observed levels of sensitivity and specificity which well exceed those achievable via the tools currently available to clinicians for the identification of stroke during acute triage, suggest that the ten marker panel identified in our analysis has legitimate potential for future clinical implementation.

While the findings reported here are exciting, it is important to note that this work was not without limitations. Of these, the one which had the greatest potential to confound the interpretation of our results was the fact that AIS patients and neurologically asymptomatic controls in our discovery cohort were not well matched with regards to several clinical and demographic characteristics; thus, intergroup differences in these factors had the potential to confound the selection of stroke specific genes in our GA/kNN analysis. To account for this possible limitation, we utilized a relatively high termination cut-off for optimal solution selection; under these conditions, a confounding factor would have to be almost ubiquitously present in one group, and nearly ubiquitously absent in the other, for it to influence the selection of candidate genes. The results of our multiple regression analysis suggest that this strategy was largely successful, however, they did infer that medication status may influence the expression of the candidate genes. Despite this, the ten candidate genes were still able to demonstrate high levels of diagnostic accuracy when discriminating between groups which were better matched in terms of these factors in the validation cohort.

Furthermore, there were additional limitations regarding the design of this study which would need to be addressed in future studies in order to move the biomarker panel identified in this investigation closer to clinical use. The first of these limitations regards the timing of blood sampling. One area in which the panel of biomarkers identified in our analysis would potentially have the greatest utility is in the pre-hospital setting, however, blood in our study was sampled at emergency department admission. While few stroke biomarker studies have sampled blood at as early of a time in AIS progression, blood would be sampled even earlier in the case of pre-hospital utilization. While our findings did suggest that the differential expression of the ten candidate markers is relatively temporally stable, blood was sampled from very few patients in under two hours from time of symptom onset. Thus, it would be encouraging to see in future work that the ten candidate genes exhibit a similar pattern of differential expression at even earlier time points in AIS progression.

Another limitation of this study regarding translation to pre-hospital use was our method of blood sampling. Blood was sampled via venipuncture, however venipuncture is rarely performed in the prehospital setting. Instead, blood sampling is typically performed via finger stick, which yields samples comprised predominantly of capillary blood. This could be problematic from that standpoint that recent studies have suggested that gene expression can differ significantly between whole venous and whole capillary blood.\(^{70}\)
possibly due to differences in cellular composition. Thus, for future pre-hospital implementation, it would need to be determined in future studies that the pattern of differential expression which we observed in peripheral venous blood across the ten candidate genes is equally observable in capillary blood.

An additional aspect of this study which could be improved upon in future work was the fact that our analysis only included ischemic stroke patients, and not patients diagnosed as other stroke subtypes such as intracerebral hemorrhage (ICH) and transient ischemic attack (TIA). Thus, it is currently unclear whether the pattern of differential expression we observed in this analysis is specific to AIS, or is a general response to brain injury. While the pattern of expression which was identified here would have clinical utility in either scenario, it is a distinction which would need to be made prior to clinical use. This, and the aforementioned limitations would perhaps be best addressed through implementation of a pilot clinical trial in which the expression levels of the candidate genes are measured in capillary blood samples obtained from consecutive suspected stroke patients in the pre-hospital setting. Such a study would provide valuable information regarding the true clinical potential of the gene expression pattern identified in this discovery study; collectively, it would potentially address current questions regarding the candidate pattern of differential expression regarding its temporal robustness, its detectability in capillary blood, and its specificity in terms of stroke-subtypes.

Another potential hurdle regarding the clinical implementation of the panel identified in this study is the fact that it is comprised of transcriptional biomarkers. There is currently no point of care diagnostic platform which is FDA approved for detection of endogenous human RNA, as almost all point of care technologies which are in clinical use have been developed for detection of proteins. However, research into rapid detection of nucleic acids is ongoing, and promising new advances such as those regarding direct RNA nanodetection and thermoneutral amplification suggest that suitable technologies will be available in the near future. Furthermore, it is likely that proper clinical validation of the markers identified this discovery study would take several years, and it is not improbable to think that suitable point of care detection platforms will be close to clinical use by that time.

Limitations aside, the robustness of our results demonstrate the power of the biomarker discovery paradigm employed in this investigation. Our findings support the notion that the stroke-induced peripheral immune response is viable source of biomarkers, and highlight that utility of high-throughput omics and machine learning techniques for candidate marker selection. While this study was aimed specifically at discovery of biomarkers for acute detection of stroke, it is likely that a similar approach could be used to identify biomarkers which address other clinical diagnostic needs within the realm of cerebrovascular disease. With this in mind, we are currently employing similar strategies to identify candidate biomarkers which could potentially be used to predict the efficacy of thrombolytic therapy and flag patients who at high risk for post-stroke BBB complications.
Role of stroke-identifying biomarkers in immunopathology:

Collectively, the relationships we observed with NMLR within the AIS group, along with the strong trend of lineage-specific expression which we observed on isolated leukocyte populations from healthy donors, suggest that the pattern of differential expression which we observed across the top 10 genes in whole blood in AIS likely serves simply as a marker of the shift in white blood cell differential which is known to occur in response to stroke. While this does not diminish the diagnostic potential of these genes as biomarkers, it suggests that many of them do not play a substantial mechanistic role in driving stroke immune pathology, as genes which drive the peripheral immune response are more likely to be transcriptionally differentially regulated at the level of individual cells. These findings not only provide a mechanistic explanation for the results of our microarray analysis, but they also provide insights into prior studies by other groups which have analyzed the transcriptome of peripheral whole blood in stroke.

Two previous studies have used microarray for expression profiling of whole blood in an attempt identify mechanisms which drive stroke-induced immunopathology. By extension, our findings suggest that it is plausible that many of the genes identified as being differentially regulated in these studies exhibit altered expression in stroke as markers of changes in white blood cell differential as well. This would infer that the biological mechanisms implied by these studies may in some ways be flawed. To assess this possibility, we measured the leukocyte sub-population expression levels of three genes identified by these studies as potentially playing a mechanistic role in driving post stroke immune suppression. Expression levels of ARG1 and NPL were demonstrated to be upregulated in whole blood in response to stroke in both studies; expectedly, we observed hundreds to thousands fold higher expression levels of these genes in cells of myeloid origin than in cells of lymphoid origin (Supplemental Figure 6A, 6B). These studies identified expression levels of CCR7 as being downregulated in whole blood in response to stroke; expectedly, we observed hundreds fold higher expression levels on cells of lymphoid origin than on cells of myeloid origin (Supplemental Figure 6C). While not definitively, these observations provide reason to believe that at least some of the genes suggested by these whole blood expression profiling studies as being mechanistic drivers of stroke immunopathology exhibit altered expression in whole blood simply as markers of changes in white blood cell differential, and may not play as mechanistic of a role as previously suggested. Our findings broadly suggest that the changes in white blood cell differential which occur in response to stroke makes reliably identifying pathologically relevant genes in whole blood unrealistic. Future studies which aim to use expression profiling of the peripheral immune system should take care to avoid this confound by measuring stroke induced changes in isolated leukocyte sub-populations or on single cells.

CD163 as a potential modulator of the stroke-induced peripheral immune response:

Of the top ten genes identified in our GA/kNN analysis, we felt that CD163 had the greatest potential to play a dual role not only as a marker of post-stroke white blood cell differential, but also as a mechanistic modulator of the stroke-induced peripheral immune response. We hypothesized that coordinate increases in the
number of CD163 expressing innate immune cells and elevations in ADAM17 activity induced by stroke could drive increases in peripherally circulating levels of sCD163, which could ultimately contribute to suppression of the adaptive immune system via the known inhibitory effects of sCD163 on lymphocyte activity.

Observations of elevated levels of sCD163 and ADAM17 activity in the peripheral blood of AIS patients, along with the fact that these elevations in sCD163 were negatively associated with lymphocyte counts, provided in vivo associative evidence supporting our hypothesis. Furthermore, our in vitro observations suggested that this stroke-induced increase in circulating sCD16 most likely originates from peripheral blood monocytes, as serum from AIS patients stimulated ADAM17-dependant production shedding in monocyte cultures obtained from healthy individuals. Our in vitro observations further suggested that stroke-induced elevations in circulating sCD163 elicit direct suppressive effects on the adaptive immune system, as depletion of sCD163 from AIS patient-derived serum was able to partially rescue its capacity to support lymphocyte proliferation. Collectively, these findings novelly suggest that sCD163 plays a mechanistic role in modulation of peripheral immunity following AIS, as a factor directly linking innate immune system activation and adaptive immune suppression.

To our knowledge, our results are the first to demonstrate a direct role for the innate immune system in mediating stroke-induced suppression of the peripheral adaptive immune system. While our results highlighted sCD163 as a means of such cross-communication, it is likely other innate-derived factors function to mediate the adaptive immune response to stroke via similar mechanisms. This notion is accentuated by the fact that complete ablation of sCD163 from stroke-derived serum was not adequate to fully rescue its capacity to support lymphocyte proliferation, inferring that there are a multitude of other soluble factors present in post-stroke peripheral circulation which have similar inhibitory effects. While the origins of these factors are inevitably diverse, it is likely that some arise via a similar mechanism as sCD163, as biological mechanisms are often redundant.75 Interesting from this regard is cluster of differentiation 14 (CD14), a membrane-bound pattern recognition receptor which is largely exclusive to the innate immune system.76,77 Much like CD163, CD14 can be shed via an MMP-dependent mechanism to generate a soluble protein (sCD14) which has been shown to directly interact with lymphocytes and exert suppressive effects.78,79 Preliminary results from our laboratory suggest that circulating levels of sCD14 are also elevated in response to stroke, and further exploration may reveal that sCD14 plays similar role as sCD163 in modulation of the peripheral adaptive immune response.

While this study identified a clear role for sCD163 in stroke immunopathology in terms of its end-action, our experiments did not look to identify the soluble factors present in peripheral circulation following stroke which are responsible for triggering the shedding of CD163 from monocytes. It is plausible to believe that this phenomenon is likely driven by a synergistic combination of stroke-induced DAMs, cytokines, and hormones, as members of the aforementioned families of factors are known to stimulate CD163 shedding. Notable from this standpoint is extracellular hemoglobin, a DAM which has been previously demonstrated as being acutely elevated in peripheral circulation following stroke as a result of intravascular and extravascular hemolysis.80 Free hemoglobin has been characterized as a potent inducer of CD163 shedding,81 and thus, may help drive

30
production of sCD163 in response to stroke. Along similar lines, levels of interleukin-6, cortisol, and reactive oxygen species, are all known to be elevated acutely in stroke pathology, and all have been shown to promote sCD163 production. Further exploration into these and other potential molecular signals which may trigger sCD13 generation within the context of stroke could not only provide a more complete picture of the mechanisms identified in this study, but a better understanding of post-stroke adaptive immune suppression as a whole, as the factors which drive post-stroke sCD163 production likely exert suppressive effects on the peripheral immune system via multiple parallel mechanisms.

While our findings are exciting, it is important to note that potentially confounding the interpretation of our results is the fact that elevations in sCD163 have been previously reported in atherosclerosis as a result of excessive intravascular macrophage activity within atherosclerotic plaques. As atherosclerosis is closely linked with stroke as one of its primary pathogenic drivers, it leaves open the possibility that our results were atherosclerosis driven, and not by the acute event of stroke itself. However, we find this scenario unlikely, as subjects in our control groups were well matched with AIS patients in terms of the prevalence of cardiovascular disease risk factors associated with atherosclerosis. Supporting our notion that the elevations in sCD163 which we observed in AIS were a direct result of the stroke-induced neurological insult is the fact that it has recently been reported that levels of macrophage-derived sCD163 are elevated in the cerebrospinal fluid (CSF) of pediatric traumatic brain injury (TBI) patients, providing concurrent evidence that cerebral trauma has the capacity to directly trigger increased sCD163 production. Furthermore, our findings, taken with theirs, suggest the possibility that stroke-induced elevations in sCD163 may not only modulation adaptive immune activity in the periphery, but could also serve a similar function within the lesion in terms of adaptive cellular infiltrates.

From a broader perspective, our findings are intriguing as they demonstrate a departure from the compartmentalized view of stroke immunopathology. While there are some exceptions, research regarding stroke immunopathology has long and too often treated the adaptive and innate immune responses as independent entities, however our results and those of select others suggest that these responses are likely synergistically linked via direct mechanisms. This realization is of utmost importance with regards to the current push towards the development of novel stroke therapeutics which target the peripheral immune system. Currently, experimental immune therapies are being developed which largely aim to improve stroke outcome via one of two often discrete avenues: either by inhibiting the acute innate immune response to stroke as means of limiting excessive tissue damage and mitigating the risk of negative secondary cerebrovascular events such as edema and hemorrhagic transformation, or by post-acutely stimulating the peripheral adaptive immune system as a means of limiting the risk of complications induced by post-stroke infection. However, immunotherapeutics proposed along this line of thinking often fail to account the possible interplay between the adaptive and innate immune systems, and the potential repercussions that could stem from modulating one response without consideration for the other.
A perfect example in this regard is the fact that ADAM17 inhibitors are often cited as a potential stroke therapeutic under the premise that they could ultimately limit excessive innate-driven inflammation and its associated complications via a reduction in TNFα production. However, our results clearly demonstrate that inhibition of ADAM17 within the innate immune system could directly interfere with mechanisms which act to suppress the adaptive immune system, an unintended effect which could put patients at elevated risk for the development of long-term autoimmune complications associated with loss of CNS self-tolerance. Such an example highlights the necessity for a global view regarding the peripheral immune system in terms of the development of novel stroke immunotherapeutics. Unfortunately, therapeutic design from such a perspective is currently unrealistic, as the mechanisms which drive the stroke-induced peripheral immune response, particularly that of the adaptive immune system, are still largely unknown. This current gap in knowledge may underlie the reality that stroke immunotherapeutics have thus-far been largely unsuccessful. Further future work which aims to characterize the peripheral immune response to stroke and its underlying mechanisms from a panoptic systems biology approach, preferably within human-centric models, would likely generate the foundational basis for the development of stroke immunotherapeutics with higher chances for clinical success.

Conclusions:

Taken as a whole, the results of this work demonstrate that a highly accurate RNA-based companion diagnostic for AIS is plausible using a relatively small number of markers, and also highlight the potential power of machine learning approaches for biomarker discovery in the realm of CVD. The ten transcriptional biomarkers identified in our GA/kNN analysis displayed levels of diagnostic performance which well exceed those reported in a majority of previous stroke biomarker investigations, as well as several characteristics which suggest they may have true clinical utility for identification of ischemic stroke during the acute phase of care; due to the robust nature of these findings, these markers warrant further evaluation to determine their true diagnostic efficacy in the clinical setting. While our results suggested that a majority of these ten markers likely do not play a role as mechanistic drivers of the stroke immunopathology, they did imply that one of them, CD163, may play a role in modulation of stroke-induced peripheral immune phenotype via a unique mechanism linking the coordinate responses of the innate and adaptive immune systems. From this regard, our results highlight a need for future research which examines the interplay between the innate and adaptive peripheral immune responses within the context of stroke, as such research could provide a better understanding of stroke pathology as a whole, and lead to the development of more effective immunotherapeutics.
Experimental procedures:

Discovery cohort patients:

Discovery cohort AIS patients and neurologically asymptomatic controls were recruited at Suburban Hospital, Bethesda, MD, which serves an upper-class metro area bordering Washington DC. AIS cases were of mixed etiology, and diagnosis was confirmed by MRI according to the established criteria for diagnosis of acute ischemic cerebrovascular syndrome (AICS). The median time from symptom onset to blood draw was 5.3 hours, as determined by the time the patient was last known to be free of AIS symptoms. In the case of patients who received thrombolytic therapy, blood samples were collected prior to the administration of recombinant tissue plasminogen activator (rtPA). Injury severity was determined according to NIHSS at the time of blood draw. Control subjects were deemed neurologically normal by a trained neurologist at the time of enrolment. Demographic information was collected from either the subject or significant other by a trained clinician. All procedures were approved by the institutional review boards of the National Institute of Neurological Disorders/National Institute on Aging at National Institutes of Health and Suburban Hospital. Written informed consent was obtained from all subjects or their authorized representatives prior to any study procedures.

Validation cohort patients:

Validation cohort AIS patients, acute stroke mimics, and neurologically asymptomatic controls were recruited at Ruby Memorial Hospital, Morgantown, WV, which serves an impoverished rural region of West Virginia which displays some of the highest CVD rates in the nation. As with the discovery cohort, AIS cases were of mixed etiology, and diagnosis was confirmed via neuroradiological imaging. Patients admitted to the emergency department as suspected strokes based on the overt presentation of stroke-like symptoms, but receiving a negative diagnosis for stroke upon imaging according to the established AICS diagnostic criteria were identified as acute stroke mimics. Discharge diagnoses of stroke mimics included cases of seizures, complex migraines, and other conditions which induce neurological symptoms such as hypertensive encephalopathy. The median time from symptom onset to blood draw was 4.6 hours and all blood was sampled before the administration of rtPA. Assessment of injury severity, screening of neurologically asymptomatic controls, and collection of demographic information was performed in an identical manner as in the discovery cohort. All procedures were approved by the institutional review boards of West Virginia University and Ruby Memorial Hospital. Written informed consent was obtained from all subjects or their authorized representatives prior to study procedures.

Neuroradiological imaging:

Neuroradiological imaging was performed using either MRI or CT within 24 hours of symptom onset. CT imaging was performed on either a Toshiba Aquilion 64 or Toshiba Aquilion 320 scanner (5 mm slices acquired at 120 KVp, 250 mA). MRI imaging was performed using either a Siemens Aera or Siemens Verio scanner (5
mm slices, 5 mm inter-slice gap, acquired via diffusion-weighted echoplanar imaging with a B-factor of 1000 s/mm²). The BrainLAB iPlan Neuroradiology software package (BrainLAB, Westchester, Ill) was used to calculate infarct volume via manual tracing as previously described, and all infarct volume calculations were verified by an experienced neuroradiologist.

**Blood collection:**

Parallel peripheral intravenous whole blood samples were collected from subjects via PAXgene RNA tubes (Qiagen, Valencia, CA), K₂EDTA Vacutainers (Becton Dickenson, Franklin Lakes, NJ), and serum separator tubes (Becton Dickenson). PAXgene RNA tubes were frozen immediately and stored at -80ºC until RNA extraction. K₂EDTA tubes were stored at room temperature until white blood cell differential (less than 30 minutes), plasma isolation (less than 30 minutes), or leukocyte isolation (performed immediately). Serum separator tubes were stored at room temperature until serum isolation (less than 30 minutes).

**Isolation of serum and plasma:**

For plasma isolation, EDTA-treated blood was spun at 2,000*g for 10 minutes to sediment hemocytes. Plasma was collected for ELISA, and the total hemocyte fraction was retained for subsequent protein extraction. Plasma was additionally spun at 10,000*g for 10 minutes to remove any residual blood cells or debris. For serum isolation, serum separator tubes were incubated at room temperature for a minimum of 15 minutes to allow for clot formation, and subsequently spun at 2,000*g for 10 minutes to sediment the resulting thrombus. Resultant samples were stored at -80ºC until analysis.

**Isolation of leukocyte subpopulations for RNA analysis:**

Leukocytes were isolated from healthy donor-derived EDTA-treated blood via immuno-magnetic negative selection (EasySep Direct, StemCell Technologies). Neutrophils, monocytes, CD4+ T-lymphocytes, CD8+ T-lymphocytes, B-lymphocytes, and NK-cells were isolated from 4 mLs of blood per population according to manufacture recommended protocols. Isolated cells were rinsed once in PBS and lysed in Qiagen buffer RLT containing beta-mercaptoethanol (BME), flash frozen in liquid nitrogen, and stored at -80°C until RNA extraction.

**RNA extraction:**

Whole blood RNA was extracted from PAXgene tubes via the PreAnalytiX PAXgene blood RNA kit (Qiagen) and automated using the QIAcube system (Qiagen). RNA was isolated from leukocyte lysates via the RNeasy Micro kit (Qiagen). Quantity and purity of isolated RNA was determined via spectrophotometry (NanoDrop, Thermo Scientific, Waltham, MA). Quality of RNA was confirmed by chip capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA).
RNA amplification and microarray:

RNA was amplified and biotinylated using the TotalPrep RNA amplification kit (Applied Biosystems, Grand Island, NY). Samples were hybridized to HumanRef-8 expression bead chips (Illumina, San Diego, CA) containing 25,000 unique probes and scanned using the Illumina BeadStation. Raw probe intensities were background subtracted, quantile normalized, and then summarized at the gene level using Illumina GenomeStudio. Sample labeling, hybridization, and scanning were performed per standard Illumina protocols. Raw data are assessable through the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) via accession number GSE16561.

GA/kNN analysis:

Normalized microarray data were filtered based on absolute fold difference between stroke and control; genes exhibiting a greater than 1.7 absolute fold difference in expression between AIS and control were retained for analysis. Filtered gene expression data were z-transformed and GA/kNN analysis was performed using C source code developed by Li et al., compiled in Linux Mint. Two-thousand near-optimal solutions were collected per sample using five nearest neighbors, majority rule, a chromosome length of five, and a termination cutoff of 0.97. Leave one out cross validation was performed using the top 50 ranked genes. The top 50 genes were tested against random gene combinations which were selected using the R sample() function (R 2.14, R Project for Statistical Computing).

Quantitative reverse transcription PCR:

cDNA was generated from purified RNA using the Applied Biosystems high capacity reverse transcription kit. For qPCR, target sequences were amplified from 10 ng of cDNA input using sequence specific primers (Supplemental Table 3) and detected via SYBR green (PowerSYBR, Thermo-Fisher) on the RotorGeneQ (Qiagen). Raw amplification plots were background corrected and CT values were generated via the RotorGeneQ software package. All reactions were performed in triplicate. Transcripts of B2M, PPIB, and ACTB were amplified as references and normalization was performed using the NORMAgene data-driven normalization algorithm (validation of this normalization strategy can be found in Appendix, Manuscript 7). 102

White blood cell differential and NMLR:

Complete blood count was obtained from EDTA-treated blood via combined optical flow cytometry and cellular impedance using an automated clinical hematology analyzer (Cell-Dyn, Abbott Diagnostics, Santa Clara, CA). NMLR was calculated as the geometric mean of absolute neutrophil and absolute monocyte count divided by absolute lymphocyte count (NMLR = sqrt[ neutrophil count * monocyte count ] / lymphocyte count).

Isolation of primary leukocytes for cell culture:

Starting with 30 mLs of healthy donor-derived EDTA-treated blood, peripheral blood mononuclear cells (PBMCs) were separated from polymorphonuclear cells (PMNs) and RBCs via centrifugation over a
polysaccharide-sodium diatrizoate density gradient (Lymphoprep, 1.077 g/mL, StemCell Technologies) for 30 minutes at 400*g. Resultant interphase PBMC fractions were collected, resuspended in RPMI 1640 (Life Technologies, Grand Island, NY) containing 10% autologous serum, separated into two aliquots, and set aside at room temperature. Pelleted PMN/RBCs were immediately placed on ice for PMN enrichment via RBC lysis induced by two 5 minute incubations with a 1:10 volume of ice cold ammonium chloride (ACK) buffer. The resultant PMNs were rinsed in ice cold PBS, counted via automated cytometer (Cellmeter X1, Nexcelom Bioscience, Lawrence, MA), and immediately plated for experiments.

Following PMN isolation, monocytes were further enriched from one aliquot of PBMCs via a second round of density gradient centrifugation as described by Menck et al.103 Briefly, PBMC suspensions were layered over an iso-osmotic 46% Percoll (1.131 g/mL, Sigma Aldrich) solution prepared in RPMI 1640 containing 10% autologous serum and centrifuged at 500*g for 30 minutes. Monocytes were collected from the resultant interphase, rinsed in PBS, counted, and immediately plated for experiments.

Following PMN and monocyte isolation, the remaining aliquot of PBMCs was rinsed in PBS and seeded in T75 flasks for lymphocyte expansion. Lymphocytes were expanded for 72 hours under standard mammalian cell culture conditions in a growth media comprised of RPMI 1640 containing 2% PHA-M (Gibco, Grand Island, NY), 2 mM L-Alanyl-Glutamine (GlutaMAX, Gibco), 25 mM HEPES (Gibco), 50 uM BME (Gibco), and 1% antibiotic antimycotic (Gibco), supplemented with 10% autologous serum. Following expansion, actively proliferating lymphocytes were rinsed in PBS, counted, and plated for experiments.

Neutrophil culture:

Neutrophils from healthy donors were seeded in 12-well plates at a density of 7.5×10^5 cells per well in 250 uL of RPMI 1640 containing 10% patient serum along with either Marimastat (10 uM, Sigma Aldrich, St-Louis, MO), TAPI-0 (10 uM, Sigma Aldrich), or vehicle (DMSO, Fisher Scientific). Following three hours of incubation under standard mammalian culture conditions, cells were lysed in NP40 lysis buffer (Life Technologies) and cellular ADAM17 activity was assessed via enzyme activity assay. Cell culture supernatants were collected and concentrations of sCD163 were quantified via ELISA; levels of neutrophil-derived sCD163 were determined by subtracting the concentration of sCD163 in media incubated without cells from the concentrations of sCD163 observed in cell culture supernatants.

Monocyte culture:

Monocytes from healthy donors were seeded in 12-well plates at a density of 1×10^6 cells per well in 250 uL of RPMI 1640 containing 20% patient serum along with either Marimastat (10 uM), TAPI-0 (10 uM), or vehicle (DMSO). Following three hours of incubation under standard mammalian culture conditions, cells were lysed in NP40 lysis buffer and cellular ADAM17 activity was assessed via enzyme activity assay. Cell culture supernatants were collected and concentrations of sCD163 were quantified via ELISA; levels of monocytes-
derived sCD163 were determined by subtracting the concentration of sCD163 in media incubated without cells from the concentrations of sCD163 observed in cell culture supernatants.

**sCD163 depletion from human serum:**

sCD163 was depleted from pooled serum samples via immunoprecipitation using biotinylated goat polyclonal antibody raised against the extracellular region of CD163 (BAF1607, R&D Systems, Minneapolis, MN). Biotinylated normal goat IgG (BAF108, R&D Systems) was used as an immunoprecipitation control. Immunoglobulins were conjugated to polymer-coated superparamagnetic beads (Dynabeads Streptavidin T1, Thermo Fisher, Waltham MA) at a ratio of 40 ug of immunoglobulins per 1 mg of beads. For immunoprecipitation, 5 mL of pooled serum was precleared with 2.5 mg of unconjugated beads, and then incubated with 2.5 mg of conjugated beads for 12 hours at 4ºC. Beads were subsequently separated from serum samples via multiple rounds of magnetic separation. Following immunoprecipitation, pooled serum was filtered at 14 microns to remove potential microbial contaminants prior to cell culture. Depletion of sCD163 was subsequently confirmed via ELISA (Supplemental Figure 4).

**Lymphocyte proliferation assay:**

Expanded lymphocytes from healthy donors were seeded in 48 well plates at 5*10⁴ cell per well in 200 uL growth media supplemented 20% pooled patient serum. Following 48 hours of culture, BrdU (Roche Life Sciences, Indianapolis, IN) was added to cultures at a final concentration of 20 uM. After allowing for 24 hours for BrdU incorporation, lymphocytes were adhered via centrifugation and fixed for BrdU quantification. Incorporated BrdU was quantified via colorimetric ELISA (Roche Life Sciences) per manufacture instructions.

**sCD163 ELISA:**

sCD163 was measured in plasma samples and cell culture supernatants using a commercially available ELISA assay (RAB0082, Sigma-Aldrich). Plasma samples were diluted 1:50, neutrophil cell culture supernatants were diluted 1:5, and monocyte cultures were diluted 1:10 prior to analysis. Due to the potential confound of hemolysis on concentrations of leukocyte derived analytes, plasma samples were screened for hemolysis prior to ELISA. Plasma absorbance was measured at 385 and 414 nm via spectrophotometry (NanoDrop, Thermo Scientific, Waltham, MA) and used to calculate a hemolysis score as described by Appierto et. al. Hemoglobin-depleted plasma (HemogloBind, Biotech Support Group, Monmouth Junction, NJ) spiked with serial dilutions of sonicated red blood cells were used as a positive control. Plasma samples with detectable hemolysis were excluded from analysis.

**ADAM17 activity assay:**

Total hemocyte fractions were thawed, mixed with NP-40 lysis buffer (Thermo Fisher) at a 1 to 1 ratio, and depleted of hemoglobin via the HemogloBind hemoglobin removal kit (Biotech Support Group, Monmouth Junction, NJ) as described by Park et al. Protein concentrations of hemoglobin-depleted total hemocyte
lysates and cell culture lysates were determined via DC-protein assay (Bio-Rad, Hercules, CA). ADAM17 activity was measured in 25 μg of total hemocyte lysate, or 10 μg of cell culture lysate, via the InnoZyme TACE activity kit (EMD Millipore, Temecula, CA) per manufacture instructions. Due to the reversible nature of ADAM17 inhibition by both marimastat and TAPI-0, experimental concentrations of drug were maintained in assay buffers when testing samples from inhibitor-treated conditions.

**Phase Contrast Microscopy:**

Phase contrast images of primary leukocyte cultures were obtained from an AxioObserver inverted microscope (Ziess, Thornwood, NY) equipped with a AxioCam mR5 digital camera (Ziess) using the AxioVision imaging software suite (Ziess). Scale bars were generated via imaging of a 0.07-1.50 mm scale calibration slide (Motic, Richmond, BC, CA).

**Statistical analysis:**

Parametric statistical analysis was performed using SPSS (IBM, Chicago, Ill) in combination with R 2.14 via the SPSS R integration plug-in. Chi-squared tests were used for comparison of dichotomous variables. Student t-tests or one-way ANOVA were used for comparison of continuous variables. Spearman’s rho was used to assess the strength of correlational relationships. Multiple regression analysis and subsequent variance decomposition was performed using the relaimpo R package. Penalized logistic regression was performed using the logistf R package. The level of significance was established at 0.05 for all parametric statistical testing. In the cases of multiple comparisons, p-values were adjusted using Holm’s Bonferroni method.
Chapter 6.

References:


92. Meisel C, Meisel A. Clinical implications of basic research Suppressing Immunosuppression after Stroke. 2015.


Supplemental data:

Supplemental Table 1. Influence of potentially confounding clinical and demographic characteristics on the predictive power of the top ten genes identified by GA/kNN. Firth’s penalized logistic regression was performed with stroke as the dependent variable and composite RNA expression, age, anticoagulant status, hypertension, antihypertension medication status, dyslipidemia, history of myocardial infarction, and history of atrial fibrillation as regressors. In the resultant model, composite RNA expression remained a highly significant predictor of stroke after considering all other regressors. Furthermore, none of the potentially confounding factors were significant predictors of stroke in the model.

Supplemental Figure 1. Relationships between post-stroke peripheral blood cellular ADAM17 activity and plasma sCD163 levels and infarct volume. (A) Relationship between peripheral blood cellular ADAM17 activity and infarct volume within the validation cohort AIS group. (B) Relationship between plasma sCD163 levels and infarct volume within the validation cohort AIS group. Strength of correlational relationships were tested via Spearman’s rho.

Supplemental Table 2. Clinical and demographic characteristics of subject sub-populations used for independent replicate in vitro experiments.
Supplemental Figure 2. Effects of ischemic stroke serum on neutrophil ADAM17-dependant sCD163 production in an independent replicate experiment. (A) ADAM17 activity in healthy-donor neutrophil lysates collected following three hours incubation with 10% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. (B) Neutrophil-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferroni post-hoc.

Supplemental Figure 3. Effects of ischemic stroke serum on monocyte ADAM17-dependant sCD163 production in an independent replicate experiment. (A) ADAM17 activity in healthy-donor monocyte lysates collected following three hours incubation with 10% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. (B) Monocyte-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferroni post-hoc.
Supplemental Figure 4. Confirmation of serum sCD163 depletion using immunoprecipitation. (A) Concentrations of sCD163 in pooled serum samples obtained from validation cohort AIS patients and controls measured with ELISA following immunoprecipitation via either anti-CD163 polyclonal antibody or isotype control immunoglobulin.

Supplemental Figure 5. Influence of post-stroke peripheral blood sCD163 levels on the capacity to support lymphocyte proliferation in an independent replicate experiment. (A) BrdU incorporation in healthy donor-derived lymphocytes over the final 24 of 72 hours PHA-stimulated proliferation in the presence of 20% serum obtained from AIS patients and controls, which was either unmanipulated in terms of sCD163 levels, or depleted of sCD163 via immunoprecipitation using anti-CD163 monoclonal antibody.
Supplemental Figure 6. Cell population-specific expression levels of genes highly cited as being potential drivers of stroke immunopathology based on the finding of prior stroke microarray investigations performed in whole blood. Expression levels of (A) NPL, (B) ARG1, and (C) CCR7 in peripheral whole blood and on isolated leukocyte populations obtained from one male and two female healthy donors. Expression levels are presented as fold change relative to the population of lowest expression, and were statistically compared via one-way ANOVA.
Supplemental Table 3. Primers and thermocycling conditions used for qRT-PCR.

<table>
<thead>
<tr>
<th>GENE</th>
<th>TRANSCRIPTS</th>
<th>PRIMERS (5' to 3')</th>
<th>PRODUCT (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTXR2</td>
<td>NM_058172.5</td>
<td>FOR: GATCTCTACCTCGTGCTGTGACA</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>NM_001145794.1</td>
<td>REV: AAATCCTCCGCAAGTTTAGTG</td>
<td></td>
</tr>
<tr>
<td>STK3</td>
<td>NM_006281.3</td>
<td>FOR: CGATGTTGGAATCCGAATTG</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>XM_011517258.1</td>
<td>REV: GTCTTTGTACTTGTGAGGTGTT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517255.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517254.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517253.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517252.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517251.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517250.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517249.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_011517247.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NM_001256312.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NM_001256313.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDK4</td>
<td>NM_002612.3</td>
<td>FOR: GACCCAGTCAACCAATCAAAATCT</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>REV: GGTCTACGAGATCGAGTGTA</td>
<td></td>
</tr>
<tr>
<td>CD163</td>
<td>NM_004244.5</td>
<td>FOR: GCAGGGAAATGGAATGGAAG</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>XM_005253529.3</td>
<td>REV: GTCAAAATTCAGAGGGAGCT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_005253528.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NM_203416.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAL</td>
<td>NM_002371.3</td>
<td>FOR: GCCCTCTTTTACCTCAGCG</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>NM_022439.2</td>
<td>REV: GCAATGTTTTTCATGGTAGCTGCT</td>
<td></td>
</tr>
<tr>
<td>GRAP</td>
<td>NM_006613.3</td>
<td>FOR: AGCCCTTGTCAGCTGACC</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td></td>
<td>REV: CGTAACCTCGTGGAAGAGAC</td>
<td></td>
</tr>
<tr>
<td>ID3</td>
<td>NM_002167.4</td>
<td>FOR: GAGAGGCAGTCAGTTAAGGC</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td></td>
<td>REV: TCTTTTGTCTGAGATGAC</td>
<td></td>
</tr>
<tr>
<td>GTSZ</td>
<td>NM_001336.3</td>
<td>FOR: GACGGGATCTGCCAAGAG</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td></td>
<td>REV: CGATGACGTTCTTGACGGGA</td>
<td></td>
</tr>
<tr>
<td>PLXDC2</td>
<td>NM_0032812.8</td>
<td>FOR: ACTCAAGTCGAGGAGGATACAGA</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>XM_011519750.1</td>
<td>REV: CCGGGTGCAAGAATGAGTAG</td>
<td></td>
</tr>
<tr>
<td>KIF1B</td>
<td>NM_015074.3</td>
<td>FOR: AAACAAAGGTATTGCGTGTCG</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>NM_183416.3</td>
<td>REV: GTAACTGCAACTTGGACAGAT</td>
<td></td>
</tr>
<tr>
<td>PPPIB</td>
<td>NM_000942.4</td>
<td>FOR: AAGTGACCCTCAGGTGGATATT</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>REV: GTGGTTTTTGTGAGCCAAATCCT</td>
<td></td>
</tr>
<tr>
<td>E2M</td>
<td>NM_004048.2</td>
<td>FOR: GAGGCTATCCAGGCTACTCCA</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>XM_006725182.2</td>
<td>REV: CGGCAGGCATACTCTCATC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>XM_005254549.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTB</td>
<td>NM_001110.3</td>
<td>FOR: CATGTACGGCCTTACTCCAGGC</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>XM_006715764.1</td>
<td>REV: CTCCTATTGTCACCGCAGAT</td>
<td></td>
</tr>
</tbody>
</table>

1 Listed by NCBI accession number  
2 All targets were amplified for 40 cycles of 95°C (15s) / 60°C (60s)
Appendix.

Associated manuscripts (3)
Machine learning approach identifies a pattern of gene expression in peripheral blood which can accurately detect ischemic stroke.

Grant C. O'Connell, MS\textsuperscript{1,2,\#}; Ashley B. Petrone, PhD\textsuperscript{1}; Madison B. Treadway, BSc\textsuperscript{3}; Connie S. Tennant, RN\textsuperscript{1}; Noelle Lucke-Wold, RN\textsuperscript{1}; Paul D. Chantler, PhD\textsuperscript{4,5}; Taura L. Barr, RN, PhD\textsuperscript{6,\#}

\textsuperscript{1}Center for Basic and Translational Stroke Research, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia
\textsuperscript{2}Department of Pharmaceutical Sciences, School of Pharmacy, West Virginia University, Morgantown, WV
\textsuperscript{3}Department of Biology, Eberly College of Arts and Sciences, West Virginia University, Morgantown, WV
\textsuperscript{4}Center for Cardiovascular and Respiratory Sciences, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia
\textsuperscript{5}Division of Exercise Physiology, School of Medicine, West Virginia University, Morgantown, West Virginia
\textsuperscript{6}CereDx Incorporated, Morgantown, WV

\#Corresponding authors:

Taura L. Barr
CereDx Incorporated
Blanchette Rockefeller Neurosciences Institute
Eight Medical Center Drive
Morgantown, WV 26505
Phone: 304-825-3131
Email: tbarr@ceredx.com

Grant C. O’Connell
West Virginia University
Robert C. Byrd Health Sciences Center
One Medical Center Drive
Morgantown, WV 26505
Phone: 920-707-0458
Email: goconnell.wvu@gmail.com

Word Count (with references): 7,396
Tables: 2
Figures: 5
Supplemental Materials: 4

Running Head: Machine learning for stroke biomarker discovery.

Keywords: Biomarkers; Cerebrovascular Disease; Cardiovascular Disease; Pattern Recognition; Immune System, Systems Biology
Abstract:

Early and accurate diagnosis of stroke improves the probability of positive outcome. The objective of this study was to identify a pattern of gene expression in peripheral blood which could potentially be optimized to expedite the diagnosis of acute ischemic stroke (AIS). A discovery cohort was recruited comprised of 39 AIS patients and 24 neurologically asymptomatic controls. Peripheral blood was sampled at emergency department admission, and genome-wide expression profiling was performed via microarray. A machine learning technique known genetic algorithm k-nearest neighbors (GA/kNN) was then used to identify a pattern of gene expression which could optimally discriminate between groups. This pattern of expression was then assessed via qRT-PCR in an independent validation cohort, where it was evaluated for its ability to discriminate between an additional 39 AIS patients and 30 neurologically asymptomatic controls, as well as 20 acute stroke mimics. GA/kNN identified ten genes (ANTXR2, STK3, PDK4, CD163, MAL, GRAP, ID3, CTSZ, KIF1B, and PLXDC2) whose coordinate pattern of expression was able to identify 98.4% of discovery cohort subjects correctly (97.4% sensitive, 100% specific). In the validation cohort, the expression levels of the same ten genes were able to identify 95.6% of subjects correctly when comparing AIS patients to asymptomatic controls (92.3% sensitive, 100% specific), and 94.9% of subjects correctly when comparing AIS patients to stroke mimics (97.4% sensitive, 90.0% specific). The transcriptional pattern identified in this study shows strong diagnostic potential, and warrants further evaluation to determine its true clinical efficacy.
Introduction:

Stroke is currently the leading cause of disability and fifth leading cause of death in the United States. It is well established that early and accurate diagnosis improves outcome by increasing the probability of successful intervention, however, the diagnostic tools currently available to clinicians for the identification of stroke have significant limitations.

While neuroradiological imaging is the gold standard for diagnosis of stroke, it is inaccessible in the field and at the initial point of contact in emergency departments. Furthermore, such imaging techniques are often not immediately available in hospitals without dedicated stroke centers, such as smaller facilities and those which serve rural areas. As a result, crucial decisions regarding the triage of potential strokes by emergency department staff and emergency medical technicians are based on the assessment of overt patient symptoms using stroke recognition and severity scales such as the Cincinnati prehospital stroke scale (CPSS) and National Institutes of Health stroke scale (NIHSS). In the hospital setting, the ability to identify stroke with such assessments is highly inconsistent, with an estimated sensitivity ranging from 44-85%, and specificity ranging between 64-98%. The sensitivity and specificity of these assessments is even lower in the pre-hospital setting, where the ability to quickly identify stroke facilitates the transfer of patients to stroke ready hospitals, increasing the chances of appropriate treatment and positive outcome. Due to these current limitations, a rapidly measurable blood-based biomarker panel could be invaluable in informing pre-hospital and in-hospital decisions early in the acute phase of care, and could ultimately expedite access to interventional treatment.

As a result, there has been a substantial push for the identification of stroke-associated peripheral blood biomarkers. The earliest stroke biomarker studies focused on the peripheral blood proteome, and countless protein-based biomarker panels have been evaluated to date. While a handful of these protein-based panels have demonstrated a strong ability to differentiate between stroke patients and healthy controls lacking the presence of cardiovascular disease (CVD) risk factors, a majority have failed to achieve specificities and sensitivities approaching 90% when tested against clinically relevant control groups. More recently, the peripheral blood transcriptome has emerged as a potential source of stroke biomarkers, as preliminary reports have suggested that gene expression in the peripheral immune system is highly responsive to ischemic brain injury. Most notably, Tang et al. identified a panel of 18 genes whose expression levels demonstrated the ability to discriminate between acute ischemic stroke (AIS) patients and healthy controls with 93.5% sensitivity and 89.5% specificity using combined expression data generated from three blood draws obtained over the first 24 hours of hospitalization. While the necessity to obtain multiple blood samples limited this biomarker panel with regards to acute stroke triage, this work provided proof of principle that stroke-induced transcriptional changes in the peripheral immune system could be used to identify stroke with relatively high levels of accuracy. Thus, it is plausible that implementation of a robust biomarker discovery approach could identify transcriptional stroke markers with the potential to be diagnostically useful during the acute phase of care.
Analysis of high-dimensional gene expression data using a pattern recognition approach known as genetic algorithm k-nearest neighbors (GA/kNN) has been successfully used in a small number of cancer studies to identify diagnostically relevant biomarker panels with strong discriminatory ability. The GA/kNN approach combines a powerful search heuristic, GA, with a non-parametric classification method, kNN. In GA/kNN analysis, a small combination of genes (referred to as a chromosome) is generated by random selection from the total pool of gene expression data (Supplemental Figure 1A). The ability of this randomly generated chromosome to discriminate between sample classes is then evaluated using kNN. In this evaluation, each sample is plotted as a vector in a multi-dimensional feature space where the coordinates of the vector are comprised of the expression levels of the genes of the chromosome. The class of each sample is then predicted based on the majority class of the nearest neighbors, or other samples which lie closest in Euclidian distance within the feature space (Supplemental Figure 1B). The ability of the chromosome to discriminate between classes is quantified as a fitness score, or the proportion of samples which the chromosome is correctly able to classify. A termination cutoff (minimum proportion of correct classifications) determines the level of fitness required to pass evaluation. A chromosome which passes kNN evaluation is labeled as a near-optimal solution and recorded, while a chromosome which fails undergoes repeated cycles of mutation and re-evaluation until a near optimal solution is reached (Supplemental Figure 1A). This entire search paradigm is performed multiple times (typically hundreds of thousands) to generate a heterogeneous pool of near-optimal solutions (Supplemental Figure 1C). The discriminatory ability of each gene is then ranked according to the number of times it appears in the near-optimal solution pool (Supplemental Figure 1D), and the collective discriminatory ability of the top ranked genes can then be tested via kNN in a leave one out cross validation (Supplemental Figure 1E). This approach has been utilized to generate biomarker panels capable of optimally discriminating between cancerous and non-cancerous colon biopsies, primary and metastatic melanoma tumors, as well as between B-cell lymphoma sub-types, all with accuracies ranging between 95 and 100%.

While GA/kNN has proven robust in several applications in the field of cancer, it has yet to be utilized for biomarker discovery in the realm of cardiovascular disease (CVD). In this study, we applied the GA/kNN approach to analyze peripheral blood gene expression data generated via microarray to identify transcriptional patterns which could potentially be optimized for the detection of AIS in the acute phase of care.
Results:

Discovery Cohort:

In order to identify potential transcriptional biomarkers for the identification of AIS, we first recruited a discovery cohort comprised of 39 AIS patients and 24 neurologically asymptomatic controls. In terms of demographic and clinical characteristics, AIS patients were older than controls, and displayed a higher prevalence of CVD risk factors such as hypertension and dyslipidemia (Table 1). Furthermore, AIS patients displayed a more substantial history of cardiac conditions such as myocardial infarction and atrial fibrillation, and higher proportion of AIS patients reported as currently taking antihypertensives and anticoagulants.

Peripheral whole blood was sampled from patients at emergency department admission, and genome-wide expression profiling was performed via microarray. Gene expression data were subjected to GA/kNN analysis, and genes were ranked based on the ability of their expression levels to discriminate between AIS patients and controls, according to the number of times they were selected as part of a near-optimal solution (Figure 1A). The expression levels of top 50 genes identified by GA/kNN displayed a strong ability to discriminate between groups using kNN in leave one out cross validation; a combination of just the top ten ranking genes (ANTXR2, STK3, PDK4, CD163, MAL, GRAP, ID3, CTSZ, KIF1B, and PLXDC2) were able to classify 98.4% of subjects in the discovery cohort correctly with a sensitivity of 97.4% and specificity of 100% (Figure 1B).

In order to evaluate the robustness of our GA/kNN analysis in terms of its ability to select optimally discriminative genes, we compared the ability of the expression levels of top 50 genes selected by GA/kNN to differentiate between stroke patients and controls to that of genes selected at random. Specifically, we compared the accuracy of GA/kNN-selected genes to the accuracy of 50 sets of 50 genes randomly generated from the total pool of gene expression data, as well to the accuracy of 50 sets of 50 genes randomly selected from a sub-pool of genes which displayed greater than 1.7 fold differential regulation between groups. The top genes selected by GA/kNN performed significantly better than genes selected at random genome-wide, as well as significantly better than genes selected at random from those which were differentially regulated greater than 1.7 fold (Figure 1C). Collectively, the results of this analysis, in combination with the levels of accuracy observed, suggest that our biomarker discovery strategy was effective at selecting genes with optimal diagnostic potential in terms of the subjects of the discovery cohort. Because the use of genes beyond the top ten did not appear to improve overall accuracy (Figure 1B), and displayed diminishing diagnostic robustness relative to genes selected at random (Figure 1C), we chose to focus on only the top ten genes for the remainder of our analysis.

When comparing the peripheral blood expression levels of the top ten genes between AIS patients and controls, the magnitude of differential expression was modest in terms of fold change in the case of most genes, however differences in expression levels between groups were highly consistent across all subjects, which was reflected by high levels of statistical significance in parametric statistical testing (Figure 2A). The combined discriminatory power of the top ten genes was evident when their coordinate expression levels were plotted on
a continuum for each individual subject; the overall pattern of expression was strikingly different between AIS
patients and controls, and it was clear that the overall pattern of expression was more diagnostically powerful
than the expression levels of any given gene on its own (Figure 2B).

In order to more intuitively explore the relationship between the pattern of gene expression observed
across the top ten genes and relevant clinical characteristics, we first used principle components analysis to
describe the expression levels of the top ten genes as single composite RNA expression variable. The
expression levels of the top ten genes were highly correlated, and a single principle component was able to
describe 70% of the collective variance in expression (Supplemental Table 1A). The result component scores
(composite RNA expression) were strongly correlated with the expression levels of each of the individual
candidate genes (Supplemental Table 1B), and visually appeared to summarize the gene expression pattern
well (Figure 2C).

We first used this composite RNA expression variable to examine the influence of potentially confounding
intergroup differences in clinical and demographic characteristics on the expression levels of the top ten genes.
Stroke, age, anticoagulant status, hypertension, antihypertensive status, dyslipidemia, history of myocardial
infarction, and history of atrial fibrillation were regressed against the composite RNA expression levels of the top
ten genes using multiple regression. We then performed variance decomposition via the LMG method to estimate
the relative contributions of each regressor to the total variance in composite RNA expression explained by the
resultant regression model.21 Stroke remained significantly associated with the composite RNA expression levels
of the top ten genes after accounting for all potentially confounding factors included in the model (Figure 3A),
and was responsible for a majority of the explained variance (77.9%, Figure 3B). In terms of potentially
confounding factors, both anti-hypertensive status and anticoagulant status were significantly associated with
the composite RNA expression levels of the top ten genes after accounting for all other regressors (Figure 3A),
however these associations only accounted for a small amount of the variance in composite RNA expression
explained by the model (6.5% and 4.5% respectively, Figure 3B). Results of this multiple regression analysis
were supported by the results of a more traditional logistic regression analysis in which the composite RNA
expression levels of the top ten genes were identified as the only significant predictor of stroke when considering
the same potentially confounding covariates (Supplemental Table 2). Taken as a whole, these findings suggest
that the pattern of differential expression observed across the top ten genes between groups is highly associated
with stroke independently of the assessed potential confounding factors. While these findings do suggest that
antihypertensive status and anticoagulant status may influence the expression levels of the top ten genes, the
effect of this influence on expression levels are likely minimal relative to the effect of stroke, and intergroup
differences in these factors were likely not significant drivers of the selection of these genes by GA/kNN.

We next used this used this composite RNA expression variable to examine the potential influence of
stroke severity and time to blood draw on the pattern of gene expression observed across the top ten genes.
The composite RNA expression levels of the top ten genes displayed a significant positive association with stroke
severity as assessed by the NIHSS (Figure 4A), suggesting that the expression levels of the top ten genes are likely directly responsive to stroke pathology. We observed a weak nonsignificant negative relationship between the composite RNA expression levels of the top ten genes and the time from symptom onset to blood draw (Figure 4B). However, this negative relationship was likely driven by the influence of stroke severity, given that the composite expression levels of these genes were positively associated with stroke severity, and patients undergoing more severe strokes generally presented to the emergency department earlier than patients undergoing less severe strokes (Figure 4B). Collectively, these observations suggest that the stroke-induced differential expression of the top ten genes may have additional utility for the stratification of stroke severity, and is relatively temporally stable during the acute phase of care.

Validation Cohort:

We then tested the diagnostic ability of gene expression pattern identified in the discovery cohort in an independent validation cohort enrolled via a second geographically and socioeconomically distinct clinical site (see methods section). This validation cohort was comprised of an additional 39 AIS patients along with two different control groups, one consisting of 30 neurologically asymptomatic controls, and the other consisting of 20 acute stroke mimics. Like in the discovery cohort, AIS patients were older than neurologically asymptomatic controls, however, AIS patients and asymptomatic controls were better matched in terms of the prevalence of co-morbidities and CVD risk factors (Table 2). AIS patients were also significantly older than stroke mimics, however, extremely well matched in terms of all other clinical and demographic characteristics (Table 2).

Peripheral blood samples were once again obtained from patients at emergency department admission, and the expression levels of the top ten genes identified by GA/KNN in the discovery cohort were measured via qRT-PCR. The overall pattern of differential expression between AIS patients and asymptomatic controls observed across the top 10 genes in the discovery cohort was also seen when comparing AIS patients and asymptomatic controls in the validation cohort (Figure 5A). The strong ability of the top ten genes to differentiate between stroke patients and asymptomatic controls in the discovery cohort using kNN was also recapitulated in the validation cohort; the expression levels of the top ten genes used in combination were able to classify 95.6% of subjects correctly with a sensitivity of 92.3% and a specificity of 100% (Figure 5B).

When comparing AIS patients to stroke mimics, the overall pattern of differential expression observed across the top ten genes was identical to that observed when comparing AIS patients to asymptomatic controls, however the magnitude of these expression differences was smaller in the case of several genes (Figure 5C). Despite this reduction in the magnitude of differential expression, the expression levels of the top ten genes used in combination were still able to accurately discriminate between AIS patients and stroke mimics, classifying 94.9% of subjects correctly with a sensitivity of 97.4% and a specificity of 90.0% (Figure 5D). However, it is important to note that it was evident that all ten genes were required to achieve high levels of diagnostic accuracy when comparing AIS patients to stroke mimics (Figure 5D), whereas similar levels of accuracy could be achieved with as few as the top four markers when comparing AIS patients to neurologically asymptomatic controls in both
the discovery cohort (Figure 1B) and the validation cohort (Figure 5B). Despite this, the collective validation cohort results supported those of the discovery cohort, and provide further evidence that the top ten markers selected by GA/kNN have high potential performance for identification of AIS.
Discussion:

The primary objective of this study was to apply the GA/kNN approach to identify a pattern of gene expression in peripheral blood which could potentially be optimized to identify AIS in the acute phase of care. The ten transcriptional markers identified by GA/kNN in our analysis proved robust in their combined ability to differentiate between AIS patients and controls in both the discovery cohort and the independent validation cohort; not only did these markers display levels of diagnostic accuracy which exceed those reported in a majority of previous stroke biomarker studies, they also demonstrated characteristics which suggest they have the potential to be clinically useful. In addition to having diagnostic utility, some of the markers identified in this study may represent viable therapeutic targets in the context of stroke immunopathology.

With regards to the countless number peripheral blood biomarker explorations which have been performed to date, to our knowledge, only one prior investigation has reported similar levels of diagnostic accuracy to those which we observed in this study in terms discriminating between stroke patients and clinically relevant control populations. Dambinova et al.\textsuperscript{22} recently reported that plasma levels of brain-derived NR2 peptide, a degradation product of \textit{N-methyl-D-aspartate} (NMDA) receptor cleavage, could be used to differentiate between stroke patients and a combination of acute stroke mimics and neurologically asymptomatic controls with 92\% sensitivity and 96\% specificity. However, a majority of blood samples in this prior study were obtained between 24 and 72 hours post symptom onset, and it is currently unknown if NR2 peptide would exhibit an equivalent level of diagnostic performance early in the acute phase of care. The ten marker panel identified in our analysis was tested earlier in the progression of pathology, and thus exhibits an obvious advantage in that it has the potential to provide actionable diagnostic information at an early enough time point to influence critical triage decisions which impact outcome.

The ten marker panel identified in our analysis displayed several favorable characteristics which could make it well suited for identification of ischemic stroke in the acute care setting. Most notably, the pattern of differential expression we observed between AIS patients and controls appeared to be relatively temporally stable. This is of clinical relevance from the standpoint that it is well established that acute stroke patients tend to arrive to the emergency department in two waves, the first within four hours from symptom onset (typically patients with more severe overt symptoms), and the second more than eight hours from symptom onset (typically patients with milder symptoms).\textsuperscript{23} For this reason, a potential diagnostic for identification of acute stroke needs to be diagnostically robust across a wide time window with regards to the progression of stroke pathology. Another diagnostically beneficial characteristic we observed was that the stroke-associated pattern of expression across these ten markers was positively correlated with the NIHSS. Thus, these markers may have utility in stratifying injury severity, information which is commonly considered when making decisions regarding the prescription of interventional treatment.\textsuperscript{4} These characteristics, along with the fact that we observed levels of sensitivity and specificity which well exceed those achievable via the tools currently available to clinicians for the
identification of stroke during acute triage, suggest that the ten marker panel identified in our analysis has legitimate potential for future clinical implementation.

In addition to having diagnostic utility, some of the markers identified in this study may represent potential therapeutic targets in the context of stroke immunopathology. Perhaps the most interesting of these markers from this standpoint is $CD163$. It is well established that stroke induces a state of peripheral adaptive immune suppression characterized by a limited capacity of lymphoid cells to respond to antigen.\(^{24,25}\) This suppressed adaptive immune state leaves patients highly susceptible to post-stroke infection,\(^{26}\) which is the leading cause of death in the post-acute phase of care.\(^{27}\) $CD163$ encodes for a protein known as cluster of differentiation 163 (CD163), a membrane-bound scavenger receptor for extracellular hemoglobin which is predominantly expressed on immune populations of myeloid lineage.\(^{28,29}\) Mature CD163 is known to undergo ectodomain shedding to generate a soluble truncated peptide (sCD163) which has been shown in multiple studies to directly interact with lymphocytes and inhibit antigen mediated activation.\(^{30-32}\) Interestingly, we observed elevated RNA expression levels of CD163 in the peripheral blood of AIS patients; it is possible that CD163 expression is increased in the innate peripheral immune system in response to stroke-induced increases in circulating free hemoglobin,\(^{33}\) subsequently driving an increase in levels of circulating sCD163 which act to suppress lymphocyte activation. In support of this hypothesis, unpublished preliminary data from our laboratory suggest that plasma levels of sCD163 are elevated in AIS patients during the acute phase of care, and are positively correlated with RNA expression levels of CD163 in whole blood. Ongoing work in our laboratory is aimed at characterizing the relationship between peripheral blood sCD163 levels and stroke-induced adaptive immune dysfunction, as CD163 may be therapeutically targetable as a means of rescuing adaptive immune responsiveness following stroke.

In addition to $CD163$, the markers identified in this study included several other genes which may be pathologically relevant within the context of the stroke-induced peripheral immune response. We observed downregulated expression levels of $MAL$ and $GRAP$ in the peripheral blood of AIS patients; both genes encode proteins which are critically involved in T-cell receptor activation and signal transduction.\(^{34,35}\) Furthermore, AIS patients exhibited elevated expression levels of $STK3$, a gene encoding a serine threonine kinase involved in pro-apoptotic signal transduction\(^{36,37}\) and suppression of lymphocyte proliferation.\(^{38}\) Taken as a whole, the differential regulation we observed across these genes is consistent with suppressed adaptive immune state induced in response to stroke, and may be mechanistically involved in blunting the responsiveness of the adaptive immune system following ischemic brain injury. Conversely, two of the markers identified as being upregulated in the peripheral blood of AIS patients in this study, $KIF1B$ and $ANTXR2$, may be mechanistically involved in the innate immune response to ischemic insult. It is well established that stroke induces robust recruitment of myeloid-derived innate immune populations such as neutrophils and monocytes from the peripheral blood into the brain parenchyma\(^{39,40}\); both genes encode proteins which have been shown to play a role in cellular adhesion and migration,\(^{41-44}\) and thus may be mechanistically involved in this process.
Collectively, the findings reported here are exciting, however, it is important to note that this study was not without limitations. Perhaps most notably was the fact that AIS patients and neurologically asymptomatic controls in our discovery cohort were not well matched with regards to several clinical and demographic characteristics; thus, intergroup differences in these factors had the potential to confound the selection of stroke specific genes in our GA/kNN analysis. To account for this possible limitation, we utilized a relatively high termination cut-off for optimal solution selection; under these conditions, a confounding factor would have to be almost ubiquitously present in one group, and nearly ubiquitously absent in the other, for it to influence the selection of candidate genes. The results of our multiple regression analysis suggest that this strategy was largely successful, however, they did infer that medication status may influence the expression of the candidate genes. Despite this, the ten candidate genes were still able to demonstrate high levels of diagnostic accuracy when discriminating between groups which were better matched in terms of these factors in the validation cohort.

Taken as a whole, the results of this preliminary study demonstrate that a highly accurate RNA-based companion diagnostic for AIS is plausible using a relatively small number of markers, and also highlight the potential power of machine learning approaches for biomarker discovery in the realm of CVD. The ten transcriptional biomarkers identified in this study displayed levels of diagnostic performance which well exceed those reported in a majority of previous stroke biomarker investigations, as well as several characteristics which suggest they may have true clinical utility for identification of ischemic stroke during the acute phase of care. Furthermore, future exploration of these markers may reveal novel mechanisms which underlie the peripheral immune response to stroke, and lead to novel therapeutic targets in the context of stroke-induced immunopathology. Due to the robust results of this preliminary analysis, the ten transcriptional biomarkers identified in this study warrant further evaluation to determine their true clinical efficacy.
Methods:

Discovery cohort patients:

Acute ischemic stroke patients and neurologically asymptomatic controls were recruited at Suburban Hospital, Bethesda, MD, which serves an upper-class metro area bordering Washington DC. AIS cases were of mixed etiology, and diagnosis was confirmed by MRI according to the established criteria for diagnosis of acute ischemic cerebrovascular syndrome (AICS). The median time from symptom onset to blood draw was 5.3 hours, as determined by the time the patient was last known to be free of AIS symptoms. In the case of patients who received thrombolytic therapy, blood samples were collected prior to the administration of recombinant tissue plasminogen activator (rtPA). Injury severity was determined according to NIHSS at the time of blood draw. Control subjects were deemed neurologically normal by a trained neurologist at the time of enrolment. Demographic information was collected from either the subject or significant other by a trained clinician. All procedures were approved by the institutional review boards of the National Institute of Neurological Disorders/National Institute on Aging at National Institutes of Health and Suburban Hospital. Written informed consent was obtained from all subjects or their authorized representatives prior to any study procedures.

Blood collection and RNA extraction:

Peripheral whole blood samples were collected via PAXgene RNA tubes (Qiagen, Valencia, CA) and stored at -80°C until RNA extraction. Total RNA was extracted via PreAnalytiX PAXgene blood RNA kit (Qiagen) and automated using the QIAcube system (Qiagen). Quantity and purity of isolated RNA was determined via spectrophotometry (NanoDrop, Thermo Scientific, Waltham, MA). Quality of RNA was confirmed by chip capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA).

RNA amplification and microarray:

RNA was amplified and biotinylated using the TotalPrep RNA amplification kit (Applied Biosystems, Grand Island, NY). Samples were hybridized to HumanRef-8 expression bead chips (Illumina, San Diego, CA) containing 25,000 unique probes and scanned using the Illumina BeadStation. Raw probe intensities were background subtracted, quantile normalized, and then summarized at the gene level using Illumina GenomeStudio. Sample labeling, hybridization, and scanning were performed per standard Illumina protocols. Raw data are assessable through the National Center for Biotechnology Information Gene Expression Omnibus via accession number GSE16561.

GA/kNN analysis:

Normalized microarray data were filtered based on absolute fold difference between stroke and control; genes exhibiting a greater than 1.7 absolute fold difference in expression between AIS and control were retained for analysis. Filtered gene expression data were z-transformed and GA/kNN analysis was performed using C source code developed by Li et al., compiled in Linux Mint. Two-thousand near-optimal solutions were collected.
per sample using five nearest neighbors, majority rule, a chromosome length of five, and a termination cutoff of 0.97. Leave one out cross validation was performed using the top 50 ranked genes. The top 50 genes were tested against random gene combinations which were selected using the R sample() function (R 2.14, R Project for Statistical Computing).

**Validation cohort patients:**

AIS patients, acute stroke mimics, and neurologically asymptomatic controls were recruited at Ruby Memorial Hospital, Morgantown, WV, which serves an impoverished rural region of West Virginia which displays some of the highest CVD rates in the nation. As with the discovery cohort, AIS cases were of mixed etiology, and diagnosis was confirmed via neuroradiological imaging. Patients admitted to the emergency department as suspected strokes based on the overt presentation of stroke-like symptoms, but receiving a negative diagnosis for stroke upon imaging according to the established AICS diagnostic criteria were identified as acute stroke mimics. Discharge diagnoses of stroke mimics included cases of seizures, complex migraines, and other conditions which induce neurological symptoms such as hypertensive encephalopathy. The median time from symptom onset to blood draw was 4.6 hours and all blood was sampled before the administration of rtPA. Assessment of injury severity, screening of neurologically asymptomatic controls, and collection of demographic information was performed in an identical manner. All procedures were approved by the institutional review boards of West Virginia University and Ruby Memorial Hospital. Written informed consent was obtained from all subjects or their authorized representatives prior to study procedures.

**Quantitative reverse transcription PCR:**

cDNA was generated from purified RNA using the Applied Biosystems high capacity reverse transcription kit. For qPCR, target sequences were amplified from 10 ng of cDNA input using sequence specific primers (Supplemental Table 3) and detected via SYBR green (PowerSYBR, Thermo-Fisher) on the RotorGeneQ (Qiagen). Raw amplification plots were background corrected and CT values were generated via the RotorGeneQ software package. All reactions were performed in triplicate. Transcripts of B2M, PPIB, and ACTB were amplified as references and normalization was performed using the NORMAgene data-driven normalization algorithm.

**Statistical analysis:**

Parametric statistical analysis was performed using SPSS (IBM, Chicago, Ill) in combination with R 2.14 via the SPSS R integration plug-in. Chi-squared tests were used for comparison of dichotomous variables, while student t-tests were used for comparison of continuous variables. Spearman’s rho was used to assess the strength of correlational relationships. For multiple regression analysis, variance decomposition was performed using the relaimpo R package. Penalized logistic regression was performed using the logistf R package. The level of significance was established at 0.05 for all parametric statistical testing. In the cases of multiple comparisons, p-values were adjusted using Holm’s Bonferonni method.
Acknowledgments:

The authors would foremost like to thank the subjects and their families, as this work was truly made possible by their selfless contribution. We also would like to thank the stroke team Ruby Memorial Hospital and the NIH stroke team at Suburban Hospital for supporting this research effort.

Contributions:

Work was conceptualized by GCO and TLB. Procedures for collection of clinical samples and recruitment of human subjects were overseen by TLB and PDC. Recruitment of subjects and collection of samples were performed by GCO, ABP, NL, and CST. Experiments were designed by GCO and performed by GCO and MBT. Data were analyzed by GCO. Manuscript was written by GCO with contributions from TLB, ABP, NL, CST, and PDC.

Funding:

Work was partially funded via a Robert Wood Johnson Foundation Nurse Faculty Scholar award to TLB (70319) and a National Institutes of Health CoBRE sub-award to TLB (P20 GM109098).

Competing interests:

GCO and TLB have a patent pending re: genomic patterns of expression for stroke diagnosis. TLB serves as chief scientific officer for CereDx Incorporated, a biotech firm which develops diagnostics for brain injury. The remaining authors report no potential conflicts of interest.
References:


Figure legends:

Table 1. Discovery cohort clinical and demographic characteristics.

Table 2. Validation cohort clinical and demographic characteristics.

Figure 1. Top 50 genes selected by GA/kNN for identification of AIS.

(A) The top 50 peripheral blood transcripts ranked by GA/kNN based on their ability to discriminate between AIS patients and neurologically asymptomatic controls in the discovery cohort. (B) Combined ability of the expression levels of top 50 genes selected by GA/kNN to discriminate between AIS patients and neurologically asymptomatic controls in the discovery cohort using kNN. (C) Ability of the expression levels of the top 50 genes selected by GA/kNN to discriminate between neurologically asymptomatic controls and AIS patients via kNN compared to the expression levels of genes selected at random. The accuracy of the top 10 genes selected by GA/kNN was specifically tested against the accuracy of randomly selected genes using single sample two-way t-test.

Figure 2. Differential expression of top ranked genes within the discovery cohort.

(A) Peripheral blood differential expression of the top ten genes selected by GA/kNN in discovery cohort neurologically asymptomatic controls and AIS patients, with fold changes reported relative to control. Statistical significance of intergroup differences in gene expression was determined via two-sample two-way t-test, and p-values were corrected to account for multiple comparisons via Holm’s Bonferroni method. (B) Coordinate pattern of peripheral blood expression across the top ten genes plotted for individual subjects in both experimental groups. (C) Composite RNA expression levels of the top ten genes generated via principle components analysis.

Figure 3. Influence of potentially confounding clinical and demographic characteristics on the expression levels of the top ten genes.

(A) Multiple regression model generated by regressing potentially confounding clinical and demographic characteristics against the composite RNA expression levels of the top ten genes selected by GA/kNN in the discovery cohort. (B) Graphical representation of the relative contribution of each regressor towards the total variance in composite RNA expression explained by the model.

Figure 4. Influence of stroke severity and time to blood draw on the coordinate expression levels of the top ranked genes in discovery cohort AIS patients.

(A) Relationship between stroke severity, as assessed by NIHSS, and composite RNA expression levels of the top ten genes in discovery cohort AIS patients. (B) Relationship between time from symptom onset to blood draw and composite RNA expression levels of the top 10 genes in discovery cohort AIS patients, with indication of stroke severity. Strength of correlations were tested via Spearman’s rho.
Figure 5. Differential expression and discriminatory ability of top ranked genes within the validation cohort.

(A) Peripheral blood differential expression of the top ten genes between validation cohort neurologically asymptomatic controls and AIS patients. (B) Combined ability of the expression levels of the top ten genes to discriminate between neurologically asymptomatic controls and AIS patients. (C) Peripheral blood differential expression of the top ten genes between acute stroke mimics and AIS patients. (D) Combined ability of the expression levels of the top ten genes to discriminate between acute stroke mimics and AIS patients. All gene expression values are reported as fold change relative to control. Statistical significance of intergroup differences in gene expression was determined via two-sample two-way t-test, and p-values were corrected to account for multiple comparisons via Holm’s Bonferroni method.

Supplemental Figure 1. GA/kNN paradigm used for the selection of transcriptional biomarkers from high-dimensional expression data.

(A) A first generation chromosome is generated via random selection from the total pool of gene expression data. (B) The ability of the first generation chromosome to discriminate between groups is evaluated using kNN. (A) If the chromosome passes evaluation, it is added to the pool of near-optimal solutions. If the chromosome fails evaluation, it undergoes mutation and re-evaluation until a near-optimal solution is found. (C) This search is repeated to generate a pool containing hundreds of thousands of near-optimal solutions. (D) Genes are then ranked based on the number of times they appear in the near-optimal solution pool. (E) The combined discriminatory ability of the top ranked genes is then tested in leave one out cross validation via kNN.

Supplemental Table 1. Use of principal components analysis for generation of a composite RNA expression variable summarizing the coordinate expression levels of the top ten genes.

(A) Correlation matrix depicting the correlations between the expression levels of the top ten genes identified by GA/kNN in the discovery cohort. Strength of correlations were tested via Spearman’s rho. (B) Principal component loadings and coefficients used for generation of the composite RNA expression variable.

Supplemental Table 2. Influence of potentially confounding clinical and demographic characteristics on the predictive power of the top ten genes identified by GA/kNN.

Firth’s penalized logistic regression was performed with stroke as the dependent variable and composite RNA expression, age, anticoagulant status, hypertension, antihypertension medication status, dyslipidemia, history of myocardial infarction, and history of atrial fibrillation as regressors. In the resultant model, composite RNA expression remained a highly significant predictor of stroke after considering all other regressors. Furthermore, none of the potentially confounding factors were significant predictors of stroke in the model.

Supplemental Table 3. Primers and thermocycling conditions used for qRT-PCR.
Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic (n=24)</th>
<th>Ischemic stroke (n=39)</th>
<th>Stat (df)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>59.9 ± 9.7</td>
<td>73.1 ± 14.0</td>
<td>t = -4.40 (61)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>14 (58.3)</td>
<td>22 (56.4)</td>
<td>χ² = 0.12 (1)</td>
<td>0.731</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>0 ± 0.0</td>
<td>5.3 ± 6.4</td>
<td>t = 5.17 (38)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>4 (16.7)</td>
<td>15 (38.5)</td>
<td>χ² = 7.02 (1)</td>
<td>0.008*</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>7 (29.2)</td>
<td>25 (64.1)</td>
<td>χ² = 11.2 (1)</td>
<td>0.001*</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>0 (0.00)</td>
<td>18 (46.2)</td>
<td>χ² = 15.5 (1)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>2 (8.30)</td>
<td>11 (28.2)</td>
<td>χ² = 3.58 (1)</td>
<td>0.056</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>2 (8.30)</td>
<td>8 (15.4)</td>
<td>χ² = 0.67 (1)</td>
<td>0.414</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>0 (0.00)</td>
<td>6 (15.4)</td>
<td>χ² = 4.08 (1)</td>
<td>0.043*</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>0 (0.00)</td>
<td>6 (15.4)</td>
<td>χ² = 4.08 (1)</td>
<td>0.043*</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>8 (33.3)</td>
<td>29 (74.4)</td>
<td>χ² = 10.3 (1)</td>
<td>0.001*</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>1 (4.20)</td>
<td>7 (17.9)</td>
<td>χ² = 2.55 (1)</td>
<td>0.111</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>5 (20.8)</td>
<td>17 (43.6)</td>
<td>χ² = 3.39 (1)</td>
<td>0.066</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>1 (4.20)</td>
<td>20 (51.3)</td>
<td>χ² = 14.9 (1)</td>
<td>&gt;0.001*</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>0 (0.00)</td>
<td>9 (23.1)</td>
<td>χ² = 6.46 (1)</td>
<td>0.011*</td>
</tr>
<tr>
<td>rtPA n (%)</td>
<td>2 (8.30)</td>
<td>2 (5.13)</td>
<td>χ² = 0.26 (1)</td>
<td>0.612</td>
</tr>
</tbody>
</table>
Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic v Acute Ischemic Stroke</th>
<th>Stroke Mimic v Acute Ischemic Stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Asym (n=30)</td>
<td>AIS (n=39)</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>51.5 ± 14.3</td>
<td>73.1 ± 13.3</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>25 (83.3)</td>
<td>25 (64.1)</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>0.0 ± 0.0</td>
<td>8.6 ± 7.5</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>16 (53.3)</td>
<td>15 (38.5)</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>17 (56.7)</td>
<td>32 (82.1)</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>11 (36.7)</td>
<td>16 (41.0)</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>2 (6.70)</td>
<td>8 (20.5)</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>1 (3.30)</td>
<td>7 (17.9)</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>0 (0.00)</td>
<td>13 (33.3)</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>0 (0.00)</td>
<td>11 (28.2)</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>15 (50.0)</td>
<td>27 (69.2)</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>2 (6.70)</td>
<td>8 (20.5)</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>7 (23.3)</td>
<td>14 (35.9)</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>1 (3.30)</td>
<td>23 (58.9)</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>0 (0.00)</td>
<td>13 (33.3)</td>
</tr>
<tr>
<td>tPA n (%)</td>
<td>1 (3.30)</td>
<td>9 (23.1)</td>
</tr>
</tbody>
</table>
Figure 1.
Figure 3.

A

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Std Error</th>
<th>p</th>
<th>R² Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.176</td>
<td>0.313</td>
<td>4E-04*</td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>1.874</td>
<td>0.151</td>
<td>1E-13*</td>
<td>0.661 (77.9%)</td>
</tr>
<tr>
<td>Hypertension Medication</td>
<td>0.454</td>
<td>0.149</td>
<td>0.004*</td>
<td>0.055 (6.5%)</td>
</tr>
<tr>
<td>Anticoagulant/Antiplatelet</td>
<td>-0.358</td>
<td>0.138</td>
<td>0.012*</td>
<td>0.038 (4.5%)</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>-0.176</td>
<td>0.154</td>
<td>0.259</td>
<td>0.031 (3.6%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.012</td>
<td>0.145</td>
<td>0.934</td>
<td>0.029 (3.4%)</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>-0.251</td>
<td>0.217</td>
<td>0.234</td>
<td>0.005 (0.6%)</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>0.175</td>
<td>0.243</td>
<td>0.475</td>
<td>0.011 (1.3%)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.001</td>
<td>0.005</td>
<td>0.824</td>
<td>0.019 (2.2%)</td>
</tr>
</tbody>
</table>

B

- Stroke
- Hypertension Medication
- Anticoagulant/Antiplatelet
- Dyslipidemia
- Hypertension
- Age
- Atrial Fibrillation
- Myocardial Infarction

Model: R²=0.848, p=1E-12*
Figure 4.

A

B

STROKE SEVERITY:
- MILD (NIHSS<5)
- MODERATE (≤ NIHSS<10)
- SEVERE (NIHSS≥10)

STROKE SEVERITY (NIHSS)

TIME TO BLOOD DRAW (MINUTES)
Figure 5.

A

<table>
<thead>
<tr>
<th>GENE</th>
<th>ASYMPOMATIC CONTROL</th>
<th>ACUTE ISCHEMIC STROKE</th>
<th>FOLD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTXR2</td>
<td></td>
<td></td>
<td>1.2</td>
<td>0.004*</td>
</tr>
<tr>
<td>STK3</td>
<td></td>
<td></td>
<td>1.5</td>
<td>3E-10*</td>
</tr>
<tr>
<td>PDK4</td>
<td></td>
<td></td>
<td>1.7</td>
<td>1E-08*</td>
</tr>
<tr>
<td>CD163</td>
<td></td>
<td></td>
<td>1.8</td>
<td>4E-08*</td>
</tr>
<tr>
<td>MAL</td>
<td></td>
<td></td>
<td>-1.4</td>
<td>0.004*</td>
</tr>
<tr>
<td>GRAP</td>
<td></td>
<td></td>
<td>-1.4</td>
<td>5E-04*</td>
</tr>
<tr>
<td>ID3</td>
<td></td>
<td></td>
<td>-1.6</td>
<td>2E-04*</td>
</tr>
<tr>
<td>CTSZ</td>
<td></td>
<td></td>
<td>1.3</td>
<td>3E-06*</td>
</tr>
<tr>
<td>KIF1B</td>
<td></td>
<td></td>
<td>1.6</td>
<td>5E-08*</td>
</tr>
<tr>
<td>PLXDC2</td>
<td></td>
<td></td>
<td>1.7</td>
<td>5E-09*</td>
</tr>
</tbody>
</table>

B

PERCENT (%) vs NUMBER OF TOP RANKED GENES

C

<table>
<thead>
<tr>
<th>GENE</th>
<th>STROKE MIMIC</th>
<th>ACUTE ISCHEMIC STROKE</th>
<th>FOLD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTXR2</td>
<td></td>
<td></td>
<td>1.2</td>
<td>0.020*</td>
</tr>
<tr>
<td>STK3</td>
<td></td>
<td></td>
<td>1.5</td>
<td>3E-06*</td>
</tr>
<tr>
<td>PDK4</td>
<td></td>
<td></td>
<td>1.8</td>
<td>2E-06*</td>
</tr>
<tr>
<td>CD163</td>
<td></td>
<td></td>
<td>1.8</td>
<td>1E-05*</td>
</tr>
<tr>
<td>MAL</td>
<td></td>
<td></td>
<td>-1.3</td>
<td>0.032*</td>
</tr>
<tr>
<td>GRAP</td>
<td></td>
<td></td>
<td>-1.3</td>
<td>0.033*</td>
</tr>
<tr>
<td>ID3</td>
<td></td>
<td></td>
<td>-1.5</td>
<td>0.038*</td>
</tr>
<tr>
<td>CTSZ</td>
<td></td>
<td></td>
<td>1.3</td>
<td>1E-05*</td>
</tr>
<tr>
<td>KIF1B</td>
<td></td>
<td></td>
<td>1.5</td>
<td>3E-05*</td>
</tr>
<tr>
<td>PLXDC2</td>
<td></td>
<td></td>
<td>1.5</td>
<td>5E-04*</td>
</tr>
</tbody>
</table>

D

PERCENT (%) vs NUMBER OF TOP RANKED GENES
Supplemental Figure 1.

A: Random formation of first generation chromosome

- Gene 4, Gene 6, Gene 2, Gene 3, Gene 1

- Fitness evaluation by 5-dimensional kNN

  - Pass
  - Fail

- Added to pool of near-optimal solutions

B: Mutation to form second generation chromosome

- Gene 4, Gene 6, Gene 2, Gene 8, Gene 1

- Fitness evaluation by 5-dimensional kNN

  - Pass
  - Fail

- Added to pool of near-optimal solutions

C: Further rounds of mutation (maximum 50 generations)

- Repeated to generate pool containing thousands of near-optimal solutions

D: Leave-one-out cross validation by kNN to determine combined discriminatory ability

E: Ranking of genes by discriminatory ability

- Discriminatory rank: Gene 1, Gene 2, Gene 3, Gene 4, Gene 5, Gene 6, Gene 7, Gene 8, Gene 9, Gene 10
- Selection count: 4, 3, 2, 2, 2, 1, 1, 0
Supplemental Table 1.

### A

<table>
<thead>
<tr>
<th></th>
<th>MAL</th>
<th>GRAP</th>
<th>ID3</th>
<th>PLXDC2</th>
<th>KIF1B</th>
<th>CD163</th>
<th>CTSZ</th>
<th>PDK4</th>
<th>STK3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRAP</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.85</td>
<td>0.88</td>
<td>-0.72</td>
<td>-0.80</td>
<td>-0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID3</td>
<td></td>
<td></td>
<td></td>
<td>-0.72</td>
<td>-0.80</td>
<td>-0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLXDC2</td>
<td>-0.72</td>
<td></td>
<td></td>
<td>-0.72</td>
<td>-0.76</td>
<td>-0.79</td>
<td>0.83</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF1B</td>
<td>-0.65</td>
<td>-0.71</td>
<td>0.69</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD163</td>
<td>-0.76</td>
<td>-0.76</td>
<td>-0.79</td>
<td>0.83</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTSZ</td>
<td>-0.43</td>
<td>-0.46</td>
<td>-0.46</td>
<td>0.57</td>
<td>0.54</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDK4</td>
<td>-0.59</td>
<td>-0.93</td>
<td>-0.52</td>
<td>0.51</td>
<td>0.54</td>
<td>0.59</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STK3</td>
<td>-0.57</td>
<td>-0.72</td>
<td>-0.64</td>
<td>0.74</td>
<td>0.76</td>
<td>0.79</td>
<td>0.59</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANTXR2</td>
<td>-0.51</td>
<td>-0.56</td>
<td>-0.50</td>
<td>0.77</td>
<td>0.78</td>
<td>0.83</td>
<td>0.56</td>
<td>0.41</td>
<td>0.58</td>
</tr>
</tbody>
</table>

KMO: 0.88  Bartlett's Sphericity: p<0.001

### B

<table>
<thead>
<tr>
<th></th>
<th>MAL</th>
<th>GRAP</th>
<th>ID3</th>
<th>PLXDC2</th>
<th>KIF1B</th>
<th>CD163</th>
<th>CTSZ</th>
<th>PDK4</th>
<th>STK3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>0.86</td>
<td>-0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC2</td>
<td></td>
<td>-0.91</td>
<td>-0.71</td>
<td>0.92</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.88</td>
<td>-0.88</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Composite RNA Expression

PC1: 70% Variance
A=0.88
Supplemental Table 2.

Regression Summary: Likelihood Ratio Test\(=60.0\), df=8, \(p=5\times10^{-10}\)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Std Error</th>
<th>(\chi^2)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.23</td>
<td>3.39</td>
<td>0.24</td>
<td>0.63</td>
</tr>
<tr>
<td>Composite RNA Expression</td>
<td>3.02</td>
<td>0.85</td>
<td>24.83</td>
<td>6E-7*</td>
</tr>
<tr>
<td>Age</td>
<td>0.05</td>
<td>0.06</td>
<td>0.45</td>
<td>0.51</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.28</td>
<td>1.62</td>
<td>0.01</td>
<td>0.91</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>-1.31</td>
<td>2.42</td>
<td>0.21</td>
<td>0.65</td>
</tr>
<tr>
<td>Myocardial Infarction</td>
<td>-1.27</td>
<td>2.13</td>
<td>0.20</td>
<td>0.66</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>0.90</td>
<td>1.81</td>
<td>0.14</td>
<td>0.71</td>
</tr>
<tr>
<td>Hypertension Medication</td>
<td>-1.17</td>
<td>1.55</td>
<td>0.43</td>
<td>0.51</td>
</tr>
<tr>
<td>Anticoagulant/Antiplatelet</td>
<td>1.69</td>
<td>1.37</td>
<td>0.88</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Supplemental Table 3.

<table>
<thead>
<tr>
<th>GENE</th>
<th>TRANSCRIPTS¹</th>
<th>PRIMERS (5’ to 3’)²</th>
<th>PRODUCT (bp)</th>
</tr>
</thead>
</table>
| ANTXR2 | NM_058172.5, NM_001145794.1 | FOR: GATCTCTACTTCGTCTGGGACA  
REV: AAATCTCTCGCAAGTTGCTG | 90 |
| STK3 | NM_006281.3, XM_011517258.1, XM_011517255.1, XM_011517254.1, XM_011517253.1, XM_011517252.1, XM_011517251.1, XM_011517250.1, XM_011517249.1, XM_011517247.1, NM_001256312.1, NM_001256313.1 | FOR: CGATGTGAGATCCACCTTGG  
REV: GTCTTTGTACTTGTGGTGAGGTT | 105 |
| PDK4 | NM_002612.3 | FOR: GCCCDAGTCACCAATCAAAATCT  
REV: GTTCATCATGACATCAGGAGTGA | 82 |
| CD163 | NM_004244.5, XM_005253529.3, XM_005253528.3 | FOR: GCCCTCTTTTACCTAGCG  
REV: GCAATGGTTTTTCATGCTAGTGCCT | 89 |
| MAL | NM_002371.3, NM_022439.2 | FOR: GCCCTCTTTTACCTACGC | 95 |
| GRAP | NM_006613.3 | FOR: AGCCCTTGCTCAAGTCACC  
REV: CTAACCTACCAGAGACAGC | 180 |
| ID3 | NM_002167.4 | FOR: GAGAGGACTGACTGTTAGCC  
REV: TCCTTTGTGCTTTGGAGATGAC | 170 |
| CTSZ | NM_001336.3 | FOR: GACGGGACTGCTGGCAAGAG  
REV: CGATGACGTTCTCAGCAGGA | 198 |
| PLXDC2 | NM_032812.8, XM_011519750.1 | FOR: ACTCAGATCCGAGGAGTACACGA  
REV: CCGGCTGGCAGAATCAGATG | 75 |
| KIF1B | NM_015074.3, XM_003416.3 | FOR: AAACAGGTAATTCCGCTGTGC  
REV: GTAAGCTCCAACTTGGAGACAT | 78 |
| PPID | NM_000942.4 | FOR: AACATCGGCCTAGGTATTTC  
REV: TGGCTTTTTGTGAGCAATCTC | 153 |
| B2M | NM_004048.2, XM_006725182.2, XM_005254549.2 | FOR: GAGGCTATCCACGTACTCCA  
REV: CGGCAGGCTACCTCATTCTT | 248 |
| ACTB | NM_001101.3, XM_006715764.1 | FOR: CATGTACCTTGTCTATCCAGCC  
REV: CTCCTTAATGTCACGCGACGAT | 250 |

¹Listed by NCBI accession number  
²All targets were amplified for 40 cycles of 95°C (15s) / 60°C (60s)
Previously identified transcriptional biomarkers of ischemic stroke are associated with peripheral immune status and injury severity.

Grant C. O’Connell, MS1,2,#; Madison B. Treadway, BSc3; Ashley B. Petrone, PhD1; Connie S. Tennant, RN1; Noelle Lucke-Wold, RN1; Yasser Kabbani, MD4; Abdul R. Tarabishy, MD4; Paul D. Chantler, PhD5,6; Taura L. Barr, RN, PhD7,#

1Center for Basic and Translational Stroke Research, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia
2Department of Pharmaceutical Sciences, School of Pharmacy, West Virginia University, Morgantown, WV
3Department of Biology, Eberly College of Arts and Sciences, West Virginia University, Morgantown, WV
4Department of Neuroradiology, Ruby Memorial Hospital, Morgantown, WV
5Center for Cardiovascular and Respiratory Sciences, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia
6Division of Exercise Physiology, School of Medicine, West Virginia University, Morgantown, West Virginia
7CereDx, Morgantown, WV

#Co-corresponding authors:

Taura L. Barr
CereDx Incorporated
Blanchette Rockefeller Neurosciences Institute
Eight Medical Center Drive
Morgantown, WV 26505
Phone: 304-825-3131
Email: tbarr@ceredx.com

Grant C. O’Connell
West Virginia University
Robert C. Byrd Health Sciences Center
One Medical Center Drive
Morgantown, WV 26505
Phone: 920-707-0458
Email: goconnell.wvu@gmail.com

Word Count (with references): 4,447
Tables: 2
Figures: 4
Supplemental Materials: 1

Running Head: Transcriptional stroke biomarkers are associated with immune status.

Keywords: NLR; neutrophil to lymphocyte ratio; cerebrovascular disease; microarray; genomics
Abstract:

The peripheral immune response to stroke plays a central role in pathology. Our group recently identified a panel of ten genes whose whole blood expression levels can be used to identify acute ischemic stroke (AIS) with high levels of accuracy. The purpose of this study was to determine the role of these transcriptional biomarkers with regards to the stroke-induced peripheral immune response. We first aimed to identify where these genes are expressed in the peripheral immune system by measuring their expression levels on six leukocyte sub-populations isolated from the peripheral blood of healthy donors. All genes previously reported as being upregulated in whole blood in response to stroke were predominantly expressed on populations of myeloid origin, while all genes previously reported as being downregulated in stroke were predominantly expressed on populations of lymphoid origin. We then examined the relationship between the expression levels of these genes and the neutrophil-monocyte to lymphocyte ratio (NMLR) in whole blood obtained from 36 AIS patients at emergency department admission. All genes previously reported as being upregulated in whole blood in response to stroke were positively associated with NMLR, while all genes previously reported as being upregulated in stroke were negatively associated with NMLR. We then demonstrated that these transcriptional stroke biomarkers can be used as a surrogate measure of white blood cell differential that could potentially be used to monitor stroke severity. Taken as a whole, these results suggest that these transcriptional stroke biomarkers are differentially regulated in whole blood as markers of stroke-induced changes in white blood cell count, and not changes at the level of transcription. Our findings not only provide insight into the mechanism by which these transcripts are stroke biomarkers, but also into the results of other prior studies which have performed whole blood expression profiling in stroke pathology.
Introduction:

There are currently close to 800,000 strokes reported annually in the United States, and stroke is responsible for 1 in every 19 mortalities.\textsuperscript{1} For those who survive, stroke constitutes a life altering event, with over 50% of patients requiring significant impatient rehabilitation or skilled nursing assistance upon discharge.\textsuperscript{1} Despite stroke’s prevalence and societal impact, therapeutic strategies aimed to improve stroke outcome remain limited. It is becoming increasingly evident that the peripheral immune system plays a critical role in stroke pathology, and may be therapeutically targetable as means of limiting secondary tissue damage and other detrimental post-stroke complications.\textsuperscript{2-4}

Stroke triggers an immediate systemic inflammatory response which results in a dramatic shift in the phenotype of the peripheral immune system. The innate arm of the peripheral immune system undergoes rapid expansion, as monocyte and neutrophil counts increase significantly in the hours following stroke onset.\textsuperscript{5,6} Simultaneously, the adaptive arm of the immune system shifts into a state of suppression characterized by lymphopenia and limited responsiveness.\textsuperscript{6-8} It has been hypothesized that this phenotypical shift may be an evolutionarily conserved mechanism intended to provide an innate immune population to infiltrate the ischemic region to aid in repair while simultaneously limiting the possibility of an adaptive autoimmune response directed against central nervous system antigens.\textsuperscript{9} However, if this immune response becomes exaggerated or prolonged, it constitutes a state of immune dysfunction that can severely impact stroke outcome.

If the innate immune system becomes excessively activated during the post-stroke immune response, it can result in prolonged inflammation in and around the infarcted region, damage healthy tissues, and enhance the injury.\textsuperscript{10,11} Furthermore, hyper-activation of innate immune system also increases the risk of developing post-stroke BBB complications such as edema and hemorrhage,\textsuperscript{12} adverse events which are associated with poor clinical prognosis.\textsuperscript{13-15} In terms of the adaptive immune system, if adaptive immune suppression becomes prolonged, it leaves patients highly susceptible to post-stroke infection,\textsuperscript{5,6,16,17} which is the leading cause of death in the post-acute phase of care.\textsuperscript{18} For these reasons, the presence of post-stroke peripheral immune dysfunction is becoming increasingly linked to poor prognosis.

The status can of the peripheral immune system can be monitored via white blood cell differential in the acute phase of care, and can be used predict stroke outcome. Elevated neutrophil to lymphocyte ratio (NLR) at hospital admission has recently been linked to longer hospitalizations,\textsuperscript{19} and increased risk of mortality following ischemic stroke.\textsuperscript{20-22} Similarly, heightened monocyte counts in the acute phase of care have been linked to increased mortality risk after intracerebral hemorrhage.\textsuperscript{23} Despite the known associations between post-stroke peripheral immune dysfunction and poor prognosis, the mechanisms which drive the peripheral immune changes which occur in response to stroke remain unclear. Better understanding these mechanisms will undoubtedly lead to future development of immune-based interventions aimed to improve stroke outcome.
Our group recently identified a panel of ten genes whose expression levels are differentially regulated in peripheral whole blood during stroke pathology which can be used to identify acute ischemic stroke (AIS) with high levels of sensitivity and specificity. The expression levels of seven of these genes were demonstrated to be upregulated in peripheral whole blood in response to ischemic stroke (ANTXR2, STK3, PDK4, CD163, CTSZ, KIF1B, and PLXDC2), while the expression levels of the remaining 3 genes were demonstrated to be downregulated (MAL, GRAP, and ID3). It is plausible that these genes play a mechanistic role in modulating the peripheral immune changes which occur in response to stroke, however the role of these genes in post-stroke immune pathology has yet to be explored. Thus, the purpose of this study was to gain a better understanding of the potential role of these genes with regards to the stroke-induced peripheral immune response. We first aimed to identify where these genes are expressed in the peripheral immune system, and then to determine the relationship between the expression levels of these genes and stroke-induced peripheral immune dysfunction and injury severity.
Methods:

Experimental Design:

In order to determine where the previously identified stroke-associated genes are expressed in peripheral blood, leukocyte sub-populations including neutrophils, monocytes, CD4+ T-lymphocytes, CD8+ T-lymphocytes, B-lymphocytes, and NK-cells were isolated from peripheral blood samples obtained from three healthy donors. qRT-PCR was then used measure expression levels of target genes in RNA extracted from the isolated sub-populations.

In order to determine the relationship between the peripheral blood expression levels of the previously identified genes and stroke-induced peripheral immune dysfunction and injury severity, blood was sampled from 36 AIS patients and 15 acute stroke mimics at emergency department admission. Peripheral blood expression levels of target genes were measured using qRT-PCR, peripheral immune status was determined via neutrophil-monocyte to lymphocyte ratio (NMLR), and injury severity was assessed via NIHSS and infarct volume. The relationship between the expression levels of target genes, neutrophil-monocyte to lymphocyte ratio, NIHSS, and infarct volume were assessed within the AIS group. Additionally, expression levels of target genes and NMLR were compared between AIS patients and stroke mimics.

Participants:

Healthy donors, AIS patients, and stroke mimics were recruited from 2011 to 2015 at Ruby Memorial Hospital, Morgantown, WV. For AIS patients, diagnosis was confirmed by MRI or CT. Patients admitted to the emergency department with stroke-like systems but receiving a negative diagnosis for stroke upon imaging were identified as stroke mimics. Samples were collected at ED admission, and always within 24 hours of symptom onset, as determined by the time the patient was last known to be free of stroke symptoms. Injury severity was determined according to the NIH stroke scale (NIHSS) at the time of blood draw. Demographic information was collected from either the subject or significant other by a trained clinician. All procedures were approved by the institutional review boards of West Virginia University and Ruby Memorial Hospital. Written informed consent was obtained from all subjects or their authorized representatives prior to study procedures.

Blood collection:

Parallel peripheral intravenous whole blood samples were collected from subjects via PAXgene RNA tubes (Qiagen, Valencia, CA) and K2EDTA Vacutainers (Becton Dickenson, Franklin Lakes, NJ). PAXgene RNA tubes were frozen immediately and stored at -80ºC until RNA extraction, while K2EDTA tubes were stored at room temperature until white blood cell differential (less than 30 minutes) or leukocyte isolation (performed immediately).
Isolation of leukocyte subpopulations:

Leukocytes were isolated from EDTA-treated blood via immuno-magnetic negative selection (EasySep Direct, StemCell Technologies). Neutrophils, monocytes, CD4+ T-lymphocytes, CD8+ T-lymphocytes, B-lymphocytes, and NK-cells were isolated from 4 mLs of blood per population according to the manufacture recommended protocol. Isolated cells were rinsed once in PBS and lysed in Qiagen buffer RLT containing BME, flash frozen in liquid nitrogen, and stored at -80 until RNA extraction.

RNA extraction and qRT-PCR:

Whole blood RNA was extracted from PAXgene tubes via the PreAnalytiX PAXgene blood RNA kit (Qiagen) and automated using the QIAcube system (Qiagen). RNA was isolated from leukocyte lysates via the RNeasy Micro kit (Qiagen). Quantity and purity of isolated RNA was determined via spectrophotometry (NanoDrop, Thermo Scientific, Waltham, MA). cDNA was generated from purified RNA via Applied Biosystems high capacity reverse transcription kit. For qPCR, target sequences were amplified from 10 ng of cDNA input using previously published sequence specific primers (O’Connell et al. 2016, in press, npj Genomic Medicine) and detected via SYBR green (PowerSYBR, Thermo-Fisher) on the RotorGeneQ (Qiagen). Raw amplification plots were background corrected and CT values were generated via the RotorGeneQ software package. All reactions were performed in triplicate. B2M, PPIB, and ACTB were amplified as reference transcripts and normalization was performed using the NORMAgene data-driven normalization algorithm.

White blood cell differential and NMLR:

Complete blood count was obtained from EDTA-treated blood via optical flow cytometry and cellular impedance using an automated clinical hematology analyzer (Cell-Dyn, Abbott Diagnostics, Santa Clara, CA). NMLR was calculated as the geometric mean of absolute neutrophil and absolute monocyte count divided by absolute lymphocyte count (NMLR = \( \sqrt[4]{\text{neutrophil count} \times \text{monocyte count}} / \text{lymphocyte count} \)).

Neuroradiological imaging:

Neuroradiological imaging was performed using either MRI or CT within 24 hours of symptom onset. The Brainlab iPlan software package was used to calculate infarct volume via manual tracing, and all infarct volume calculations were verified by a neuroradiologist.

Statistics:

All statistics were performed using the GraphPad Prism statistical software package. Chi-squared test was used for groupwise comparison of dichotomous variables. T-test or one-way ANOVA was used for groupwise comparisons where appropriate. Spearman’s rho was used to test the strength of correlational relationships. The null hypothesis was rejected when p<0.05.
Results:

Transcriptional stroke biomarkers exhibit lineage specific expression in peripheral blood:

In order to get a better understanding regarding the potential role of the previously identified stroke-associated genes in stroke-induced immune response, we first looked to identify where these stroke associated genes are expressed in the peripheral immune system. Interestingly, almost every gene exhibited some degree of lineage-specific expression. Each of the seven genes which have previously been shown to be up-regulated in stroke were all predominantly expressed in cell populations of myeloid origin (Figure 1A, 1B, 1C, 1D, 1H, 1I, 1J). Conversely, all of genes which have been previously identified as being downregulated in the peripheral blood following stroke were predominantly expressed in cells of lymphoid origin (Figure 1E, 1F, 1G). In some cases, the difference in expression between myeloid and lymphoid populations was quite substantial, in excess of hundred fold.

Expression levels of transcriptional stroke biomarkers are highly correlated with peripheral immune status:

We next explored the relationship between the expression levels of the previously identified stroke-associated genes and white blood cell differential in the peripheral blood of AIS patients at emergency department admission. The whole blood expression levels of each of the seven genes which have previously been shown to be up-regulated in stroke were to some degree positively associated with NMLR. Conversely, the whole blood expression levels of all three genes which have previously been shown to be down-regulated in stroke were negatively associated with NMLR.

Transcriptional stroke biomarkers can be used as a surrogate for white blood cell differential:

Because of the strong pattern of lineage-specific expression and the relationships we observed with NMLR, we hypothesized that the changes in expression levels previously reported in whole blood are reflective of stroke-induced changes in white blood cell count, and not by changes at the level of transcription. Thus, the whole blood expression levels of these genes may serve as a surrogate for white blood cell differential. To examine this possibility, we calculated a transcriptional-NMLR (tNMLR) by dividing the geometric mean of the whole blood expression levels of the seven myeloid expressed genes by the whole blood expression levels of the three lymphoid expressed genes (tNMLR = \(7^{th}\text{ root}\left[\text{ANTXR2*STK3*PDK4*CD163*CTSZ*KIF1B*PLXDC2}\right] / \left[3^{rd}\text{ root}\left[\text{MAL*GRAP*ID3}\right]\right]\)). We then examined the relationship tNMLR and white blood cell differential obtained NMLR in the AIS group. Unsurprisingly, we observed a strong positive correlation between tNMLR and NMLR in the peripheral blood of AIS patients at emergency department admission. These results support the idea that the expression levels of these genes are a surrogate indicator of white blood cell differential.
Transcriptional stroke biomarkers are associated with injury severity in a similar manner as white blood cell differential:

Because NMLR has been shown to have diagnostic utility in stroke, we next aimed to determine whether tNMLR could provide similar diagnostic information. Consistent with prior reports, NMLR was higher in AIS patients than stroke mimics at emergency department admission. tNMLR was also elevated in stroke patients, however to a greater degree of statistical significance. In terms of injury severity, both NMLR and tNMLR exhibited positive associations with NIHSS in AIS patients at emergency department admission, however neither relationship was statistically significant. NMLR and tNMLR were also both positively associated with infarct volume, however only tNMLR exhibited a statistically significant relationship. Collectively, these results suggest that a surrogate white blood cell differential derived from the expression levels of the previously identified transcriptional stroke biomarkers can provide similar diagnostic information in the acute phase of care as NMLR.
Discussion:

The primary objective of this study was to determine the potential role of ten transcriptional stroke biomarkers previously identified by our group in stroke immunopathology. Our findings suggest that these ten transcriptional biomarkers are most likely differentially regulated in whole blood as markers of stroke-induced changes in white blood cell differential, and are likely not altered at the level of transcription in stroke pathology. These results not only provide insight into our previous findings, but also into the findings of studies by other groups which have tried to elucidate mechanisms of stroke immunopathology using expression profiling data generated from whole blood. Furthermore, our results also reaffirm the potential diagnostic usefulness of white blood cell differential in stroke, and suggest that transcriptional biomarkers could be used as a surrogate measure in cases where cytometry is not available, such as in pre-clinical settings.

It is well established that neutrophil and monocyte counts increases early in stroke pathology, while lymphocyte counts drop. All seven genes which we previously identified as being upregulated in whole blood in response to stroke were predominantly expressed in myeloid cell populations and positively associated with NMLR. All three genes which we previously identified as being down regulated in AIS were predominantly expressed in lymphoid cell populations and negatively associated with NMLR (summarized in Table 2). Taken as a whole, this suggests that these genes are likely not differentially regulated at the level of transcription in response to stroke, and likely exhibit altered expression in whole blood simply as markers of changes in white blood cell differential. While this does not diminish the diagnostic potential of these genes as biomarkers, it suggests they may not play a substantial mechanistic role in driving stroke immune pathology.

Two previous studies have used microarray for expression profiling of whole blood in an attempt identify mechanisms which drive stroke-induced immunopathology. Our findings suggest that it is plausible that many of the genes identified as being differentially regulated in these studies exhibit altered expression in stroke as markers of changes in white blood cell differential as well. This would suggest that the biological mechanisms implied by these studies may in some ways be flawed. To assess this possibility, we measured the leukocyte sub-population expression levels of three genes identified these studies which have been suggested to play a mechanistic role in driving post stroke immune suppression. Expression levels of ARG1 and NPL were demonstrated to be upregulated in whole blood in in response to stroke in both studies, expectedly, we observed hundreds to thousands fold higher expression levels of these genes on cells of myeloid origin than on cells of lymphoid origin (Supplemental Figure 1A, 1B). These studies identified expression levels of CCR7 as being downregulated in whole blood in response to stroke; expectedly, we observed hundreds fold higher expression levels on cells of lymphoid origin than on cells of myeloid origin (Supplemental Figure 1C). While not definitively, these observations provide reason to believe that at least some of the genes suggested by these whole blood expression profiling studies as being mechanistic drivers of stroke immunopathology exhibit altered expression in whole blood simply as markers of changes in WBC differential, and may not play as mechanistic of a role as previously suggested. Genes which play a central mechanistic role in stroke immunopathology are likely ones
which are differentially regulated at the level of transcription. Our collective observations suggest that that changes in white blood cell differential which occur in response to stroke makes reliably identifying such genes in whole blood unrealistic. Future studies which aim to use expression profiling of the peripheral immune system should take care to avoid this confound by measuring stroke induced changes in isolated leukocyte sub-populations or on single cells.

Our finding supports prior work which suggest white blood cell differential could be diagnostically useful in the acute phase of care. Prior studies have focused on the use of neutrophil to lymphocyte ratio (NLR) as a diagnostic marker in stroke, while to our knowledge, this study was the first to examine NMLR. Several of these studies have demonstrated that NLR is elevated in AIS patients early in pathology and is associated with injury severity. Similarly, in this study, we observed that NMLR was significantly higher in AIS patients at emergency department admission as opposed to stroke mimics, and is to some degree positively associated with injury severity. Interestingly, we observed more distinct differences in NMLR between AIS patients and stroke mimics compared to NLR as well as stronger associations between injury severity and NMLR than NLR (data not shown). These observations suggest that NMLR may be more diagnostically robust than NLR, however future study in a larger sample size would have to be used to definitely draw such a conclusion. Furthermore, our results suggest that the ten transcriptional stroke biomarkers identified by our previous work could be used as a surrogate measure for NMLR.

Having a rapidly assessable surrogate measure of white blood cell differential could be diagnostically useful in the prehospital setting. Assessment of white blood cell differential by cytometry is cumbersome and expensive, limiting its accessibility in the field. However, recent advances in nucleic acid detection suggest that cost-effective rapid quantification of nucleic acids via portable devices will soon be possible. Thus, the transcriptional biomarkers explored in this study may be able to serve as a surrogate measure of WBC differential that is field assessable to first responders. Such an assessment could be valuable in making triage decisions with regards to potential strokes. Because it has been established that alterations in white blood cell differential are associated with injury severity and predictive of mortality, a field assessment which could provide similar diagnostic information could be used to identify high risk patients early and allow for transport to more stroke-ready facilities. Additionally, such a diagnostic assessment may be of value in resource-poor world regions where the prevalence of stroke is dramatically increasing, however access to modern medical technologies such as cytometry or neuroradiological imaging are limited.

We have already established that the transcriptional biomarkers explored in this study can be used to identify AIS with high levels of sensitivity and specificity; our results suggest that their expression levels are altered in peripheral whole blood as markers of stroke-induced changes in white blood cell count. These findings not only provide insight into the mechanism by which these transcripts are stroke biomarkers, but also into the results of other prior studies which have performed whole blood expression profiling in stroke pathology. Additionally, our results also reaffirm the potential diagnostic usefulness of white blood cell differential in stroke,
and suggest that transcriptional biomarkers have the potential to be used as a surrogate measure in the pre-clinical setting.
Acknowledgments:

The authors would foremost like to thank the subjects and their families, as this work was truly made possible by their selfless contribution. We also would like to thank the stroke team at Ruby Memorial Hospital for supporting this research effort.

Contributions:

Work was conceptualized by TLB and GCO. Experiments were designed by GCO and performed by GCO and MBT. Measurement of infarct volume was performed by YK and ART. Data were analyzed by GCO. Patient recruitment and collection of samples was performed by ABP, NL, CST and PDC. Manuscript was written by GCO with contributions from TLB, ABP, NL, CST, and PDC.

Funding:

Work was funded via a Robert Wood Johnson Foundation Nurse Faculty Scholar award to TLB (70319) and a National Institutes of Health CoBRE sub-award to TLB (P20 GM109098).

Disclosures:

GCO and TLB have a patent pending re: genomic patterns of expression for stroke diagnosis. TLB serves as chief scientific officer for CereDx, a biotech firm which develops diagnostics for brain injury. The remaining authors report no potential conflicts of interest.
References:


Figure Legends:

**Table 1. Clinical and demographic characteristics**

**Table 2. Summary**

**Figure 1. Expression levels of transcriptional stroke biomarkers on isolated leukocyte sub-populations.**

Expression levels of (A) ANTXR2, (B) STK3, (C) PDK4, (D) CD163, (E) MAL, (F) GRAP, (G) ID3, (H) CTSZ, (I) KIF1B, and (J) PLXDC2 in peripheral whole blood and isolated leukocyte populations obtained from one male and two female healthy donors. Expression levels are presented as fold change relative to the population of lowest expression.

**Figure 2. Relationships between whole blood expression levels of transcriptional stroke biomarkers and peripheral immune status in AIS patients.**

Relationship between peripheral whole blood expression levels of (A) ANTXR2, (B) STK3, (C) PDK4, (D) CD163, (E) MAL, (F) GRAP, (G) ID3, (H) CTSZ, (I) KIF1B, and (J) PLXDC2 and NMLR in AIS patients at emergency department admission.

**Figure 3. Relationship between NMLR and tNMLR**

Relationship between NMLR and tNMLR in peripheral blood of AIS patients at emergency department admission.

**Figure 4. Relationship between tNMLR and measures of injury severity.**

(A) NMLR and (B) tNMLR in AIS patients and stroke mimics at emergency department admission. Relationships between both (C) NMLR and (D) tNMLR and NIHSS in AIS patients at emergency department admission. Relationships between both (E) NMLR and (F) tNMLR and infarct volume in AIS patients at emergency department admission. Infarct volume categories comprised of three groups of twelve patients each.

**Supplemental Figure 1. Cell population-specific expression levels of genes highly cited as being potential drivers of stroke immunopathology based on the finding of prior stroke microarray investigations performed in whole blood.**

Expression levels of (A) NPL, (B) ARG1, and (C) CCR7 in peripheral whole blood and on isolated leukocyte populations obtained from one male and two female healthy donors. Expression levels are presented as fold change relative to the population of lowest expression, and were statistically compared via one-way ANOVA.
Table 1.

<table>
<thead>
<tr>
<th></th>
<th>MIMIC (n=16)</th>
<th>STROKE (n=38)</th>
<th>STAT (df)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>60.2 ± 17.2</td>
<td>73.1 ± 12.85</td>
<td>t = -2.95 (49)</td>
<td>0.005 *</td>
</tr>
<tr>
<td>Male n (%)</td>
<td>7 (46.7)</td>
<td>14 (38.9)</td>
<td>$\chi^2 = 0.26$ (1)</td>
<td>0.607</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>8 (53.3)</td>
<td>22 (61.1)</td>
<td>$\chi^2 = 0.26$ (1)</td>
<td>0.607</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>5.0 ± 4.5</td>
<td>9.0 ± 7.6</td>
<td>t = -1.98 (49)</td>
<td>0.064</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>4 (26.7)</td>
<td>14 (38.9)</td>
<td>$\chi^2 = 0.69$ (1)</td>
<td>0.405</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>13 (86.7)</td>
<td>30 (83.3)</td>
<td>$\chi^2 = 0.09$ (1)</td>
<td>0.765</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>10 (66.7)</td>
<td>16 (44.4)</td>
<td>$\chi^2 = 2.09$ (1)</td>
<td>0.148</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>5 (33.3)</td>
<td>7 (19.4)</td>
<td>$\chi^2 = 1.14$ (1)</td>
<td>0.286</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>4 (26.7)</td>
<td>7 (19.4)</td>
<td>$\chi^2 = 0.33$ (1)</td>
<td>0.567</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>3 (20.0)</td>
<td>12 (33.3)</td>
<td>$\chi^2 = 0.91$ (1)</td>
<td>0.341</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>5 (33.3)</td>
<td>10 (27.7)</td>
<td>$\chi^2 = 0.16$ (1)</td>
<td>0.692</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>13 (86.7)</td>
<td>25 (69.4)</td>
<td>$\chi^2 = 1.65$ (1)</td>
<td>0.199</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>5 (33.3)</td>
<td>7 (19.4)</td>
<td>$\chi^2 = 1.14$ (1)</td>
<td>0.287</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>9 (60.0)</td>
<td>13 (36.1)</td>
<td>$\chi^2 = 2.46$ (1)</td>
<td>0.116</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>11 (73.3)</td>
<td>22 (61.1)</td>
<td>$\chi^2 = 0.69$ (1)</td>
<td>0.405</td>
</tr>
<tr>
<td>rtPA n (%)</td>
<td>0 (0.00)</td>
<td>12 (33.3)</td>
<td>$\chi^2 = 6.54$ (1)</td>
<td>0.012 *</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>1 (6.70)</td>
<td>8 (22.2)</td>
<td>$\chi^2 = 1.76$ (1)</td>
<td>0.184</td>
</tr>
<tr>
<td>GENE</td>
<td>REGULATION IN AIS</td>
<td>CORRELATION WITH NMLR</td>
<td>PREDOMINANT EXPRESSION</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>ANTXR2</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
<tr>
<td>STK3</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
<tr>
<td>PDK4</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
<tr>
<td>CD163</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
<tr>
<td>MAL</td>
<td>↓</td>
<td>−</td>
<td>LYMPHOID</td>
<td></td>
</tr>
<tr>
<td>GRAP</td>
<td>↓</td>
<td>−</td>
<td>LYMPHOID</td>
<td></td>
</tr>
<tr>
<td>ID3</td>
<td>↓</td>
<td>−</td>
<td>LYMPHOID</td>
<td></td>
</tr>
<tr>
<td>CTSZ</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
<tr>
<td>KIF1B</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
<tr>
<td>PLXDC2</td>
<td>↑</td>
<td>+</td>
<td>MYELOID</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1.
Figure 2.
Figure 3.
Figure 4.
Supplemental Figure 1.
ADAM17-dependent generation of soluble CD163 by monocytes perpetuates post-stroke suppression of the peripheral adaptive immune system.

Grant C. O’Connell, MS\textsuperscript{1,2,\#}, Ashley B. Petrone, PhD\textsuperscript{1}; Connie S. Tennant, RN\textsuperscript{1}; Noelle Lucke-Wold, RN\textsuperscript{1}; Paul D. Chantler, PhD\textsuperscript{3,4}; Taura L. Barr, RN, PhD\textsuperscript{5,\#}

\textsuperscript{1}Center for Basic and Translational Stroke Research, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia

\textsuperscript{2}Department of Pharmaceutical Sciences, School of Pharmacy, West Virginia University, Morgantown, WV

\textsuperscript{3}Center for Cardiovascular and Respiratory Sciences, Robert C. Byrd Health Sciences Center, West Virginia University, Morgantown, West Virginia

\textsuperscript{4}Division of Exercise Physiology, School of Medicine, West Virginia University, Morgantown, West Virginia

\textsuperscript{5}CereDx, Morgantown, WV

\#Co-corresponding authors:

Taura L. Barr
CereDx Incorporated
Blanchette Rockefeller Neurosciences Institute
Eight Medical Center Drive
Morgantown, WV 26505
Phone: 304-825-3131
Email: tbarr@ceredx.com

Grant C. O’Connell
West Virginia University
Robert C. Byrd Health Sciences Center
One Medical Center Drive
Morgantown, WV 26505
Phone: 920-707-0458
Email: goconnell.wvu@gmail.com

Word Count (with references): 7,822
Tables: 2
Figures: 5
Supplemental Materials: 8

Cover Title: sCD163 modulates stroke-induced adaptive immune suppression.

Keywords: TACE, Immune therapy, Autoimmunity, Cerebrovascular disease, Cardiovascular disease
Abstract:

The peripheral adaptive immune system responds to stroke by shifting into a state of suppression. While it is becoming increasingly evident that the duration and magnitude of this response significantly influences patient outcome, its underlying mechanisms are not yet fully understood. CD163 is a myeloid-expressed scavenger receptor which can be shed from the plasma membrane via the metalloprotease ADAM17 to generate a soluble peptide which is known to directly interact with lymphocytes and exert inhibitory effects. Based on prior observations of elevated transcriptional levels of CD163 in the peripheral blood of ischemic stroke patients, we hypothesized that CD163 may be a mediator of stroke-induced adaptive immune suppression. To address this hypothesis, we first assessed cellular ADAM17 activity and soluble CD163 levels in the peripheral blood of a modestly-sized cohort of ischemic stroke patients (n=39), neurologically asymptomatic controls (n=20), and acute stroke mimics (n=20). ADAM17 activity and soluble CD163 levels were elevated in ischemic stroke patients relative to control groups, and negatively associated with post-stroke lymphocyte counts. Subsequent in vitro experiments suggested that this stroke-induced elevation in circulating soluble CD163 likely originates from monocytes, as serum from ischemic stroke patients stimulated ADAM17-dependant CD163 shedding in primary monocyte cultures generated from the peripheral blood of healthy individuals. Additional in vitro experiments further suggested that post-stroke elevations in circulating soluble CD163 have the capacity to elicit direct suppressive effects on the adaptive immune system, as serum from ischemic stroke patients inhibited the proliferation of healthy donor-derived lymphocytes, an effect which was partially ameliorated following serum CD163 depletion. Collectively, these observations novelly suggest that CD163 plays a mechanistic role in modulation of peripheral immunity following ischemic stroke, as a factor directly linking innate immune system activation and adaptive immune suppression.
Introduction:

Stroke triggers a systemic inflammatory response which results in a dramatic shift in the phenotype of the peripheral immune system. The innate arm of the peripheral immune system undergoes rapid activation, a phenomenon which results in significant elevations in circulating neutrophil and monocyte counts in the hours following onset.\textsuperscript{1,2} Conversely, the adaptive arm of the peripheral immune system shifts into a state of suppression, often characterized by a prolonged period of lymphopenia and limited antigen responsiveness.\textsuperscript{2–4} While the pathophysiological role of the innate immune response to stroke has been long appreciated, a growing body of evidence suggests that the adaptive immune response to stroke also critically influences patient outcome.

It is likely that the state of adaptive immune suppression which develops in response to stroke largely serves to limit the possibility of an autoimmune response as the blood brain barrier becomes disrupted and the peripheral immune system becomes exposed to unfamiliar central nervous system (CNS) antigens.\textsuperscript{5} Expectedly, heightened adaptive immune activity immediately following stroke has been shown to prolong the inflammatory state and promote secondary tissue damage, factors which can ultimately contribute to poor prognosis.\textsuperscript{6} While the state of adaptive immune suppression is acutely protective, if prolonged, it can constitute a significant detriment to recovery.\textsuperscript{7–11} Patients which display prolonged adaptive immune suppression following stroke become highly susceptible to post-stroke infection,\textsuperscript{12–15} which is one of the leading causes of death in the post-acute phase of care.\textsuperscript{8} Despite the fact that is becoming increasingly evident that this delicate balancing act between maintenance of self-tolerance and pathogen susceptibility plays a significant role in recovery, the mechanisms which drive the peripheral adaptive immune response to stroke are not yet fully understood.

CD163 is a membrane-bound scavenger receptor for extracellular hemoglobin which is believed to be expressed exclusively within the innate immune system,\textsuperscript{16} where it is predominantly found on monocytes and macrophages,\textsuperscript{17} and to a lesser extent on neutrophils.\textsuperscript{18} Various stimuli can trigger CD163 ectodomain shedding via cleavage by the metalloprotease ADAM metallopeptidase domain 17 (ADAM17),\textsuperscript{19,20} resulting in generation of a soluble truncated peptide (sCD163) which has been shown in multiple studies to directly interact with lymphocytes and inhibit antigen-induced proliferation.\textsuperscript{21–23} Interestingly, a genome-wide transcriptomic screen performed by our group recently identified CD163 as being robustly up-regulated in the peripheral blood of ischemic stroke patients within hours of system onset (O’Connell et al. 2016, in-press, \textit{npj Genomic Medicine}). Furthermore, other groups have reported heightened ADAM17 activity in animal models of ischemic brain injury,\textsuperscript{24–26} as well as stroke-induced increases in peripherally circulating levels of various ADAM17 substrates such as tumor necrosis factor alpha (TNFα) in human subjects.\textsuperscript{27,28} Therefore, it is possible that stroke induces a rise in peripherally circulating levels of sCD163 via coordinate increases in CD163 expression and ADAM17 activity; such a rise in sCD163 levels could subsequently contribute to suppression of the adaptive immune system via the inhibitory effects of sCD163 on lymphocyte activity. Thus, the primary objective of this study was to investigate sCD163 as a potential modulator of post-stroke adaptive immune suppression.
Results:

In order to determine whether stroke drives an increase in peripheral sCD163 production in vivo, we first recruited 39 acute ischemic stroke patients (AIS) along with two control groups, the first consisting of 20 neurologically asymptomatic subjects, and the second consisting of 20 acute stroke mimics (Table 1). Peripheral venous blood was sampled within 24 hours of symptom onset, and expression levels of ADAM17 and CD163 were assessed in terms of RNA and protein using a combination of qRT-PCR, enzyme activity assay, and ELISA.

Consistent with our previous reports, transcriptional levels of both ADAM17 and sCD163 were elevated in the whole blood of AIS patients relative to that of control groups (Figure 1A, 1B). Protein assays supported these observations, as AIS patients displayed heightened cellular ADAM17 activity in total hemocyte fractions generated from peripheral blood (Figure 1C), as well as elevated plasma levels of sCD163 (Figure 1D). Furthermore, both cellular ADAM17 activity and plasma sCD163 levels displayed positive associations with stroke severity in terms of NIHSS (Figure 1E, 1F) and infarct volume (Supplemental Figure 1A, 1B); while our relatively small sample size and the variability inherent to such measures of injury status kept these associations from reaching statistical significance, these relationships tentatively inferred that both factors are directly responsive to stroke pathology. Taken as a whole, these observations provided associative evidence supporting our hypothesis that stroke induces a peripheral rise in sCD163 levels via increased ADAM17 activity.

To explore whether the elevations in sCD163 levels which we observed in AIS play a role in modulation of the stroke-induced peripheral immune response, we next assessed the relationship between circulating sCD163 levels and white blood cell counts within the AIS group. Plasma sCD163 levels were significantly negatively associated with post-stroke absolute lymphocyte counts (Figure 2A), as well as significantly positively associated with both absolute monocyte counts (Figure 2B) and absolute neutrophil counts (Figure 2C). Collectively, these results provided in vivo associative evidence which supported our hypothesis that stroke-induced increases in myeloid-derived circulating sCD163 may contribute to post-stroke suppression of the adaptive immune system.

Next, we wanted to determine the direct effects of the post-stroke peripheral inflammatory milieu on ADAM17-mediated sCD163 shedding by peripheral blood innate immune cell populations. To do this, primary neutrophil and monocyte cultures generated from the peripheral blood of twelve healthy donors (Supplemental Table 1) were treated with pooled serum samples derived from a subset of ten AIS patients, ten neurologically asymptomatic controls, and ten acute stroke mimics which were relatively well matched in terms of clinical and demographic characteristics (Table 2). Serum incubation was performed in both the presence and absence of the ADAM17 inhibitors TAPI-0 and Marimastat, and cellular ADAM17 activity and sCD14 production were assessed following treatment.

Neutrophil cultures across all treatment conditions appeared largely phenotypically identical when visually inspected following treatment (Figure 3A). ADAM17 activity did appear to be elevated in neutrophil
cultures treated with serum obtained from AIS patients relative to those treated with serum obtained from both neurologically asymptomatic controls and acute stroke mimics, however this increase was not statistically significant (Figure 3B). Overall, neutrophil cultures appeared to generate relatively little sCD163, although a limited number of cultures treated with AIS serum did appear to exhibit increased sCD163 production in response to treatment. This effect was not widespread however, and no statistically significant differences were observed in sCD163 production between cultures treated with AIS serum and those treated with serum from control groups (Figure 3C). A second independent experiment using pooled serum samples derived from a separate sub-set of subjects (Supplemental Table 2) yielded similar results (Supplemental Figure 2).

Counter to what was observed in terms of neutrophil cultures, there were striking visual differences between monocyte cultures treated with serum derived from AIS patients and those treated with serum derived from control groups. Cultures treated with control serum generally displayed a mix of adherent and semi-adherent cells, however several cultures incubated with AIS serum contained large patches of cells which appeared highly adherent with prominent cytoplasmic vacuoles characteristic of a shift towards a macrophage phenotype (Figure 4A). ADAM17 activity was significantly elevated in monocytes treated with AIS serum relative to those treated with serum from both control groups (Figure 4B). Furthermore, supernatants recovered from cultures treated with AIS serum contained significantly higher concentrations of monocyte-derived sCD163 than those recovered from cultures treated with serum from either control group, and this effect was largely ablated by inhibition of ADAM17 (Figure 4C). Once again, similar results were observed in a second independent experiment (Supplemental Figure 3) using pooled serum samples derived from a separate sub-set of subjects (Supplemental Table 2). Collectively, these observations demonstrate that soluble factors present in peripheral circulation following ischemic stroke have the capacity to trigger ADAM17-dependant sCD163 shedding from peripheral blood monocytes, however likely not from neutrophils.

Lastly, we wanted to determine the direct effect of post-stroke-elevations in circulating sCD163 levels on the capacity of the peripheral blood to support lymphocyte proliferation. To do this, PHA-stimulated proliferation of healthy-donor derived lymphocytes was assessed via BrdU incorporation in the presence of pooled serum obtained from control groups, pooled serum from AIS patients, and pooled serum obtained from AIS patients which had been depleted of sCD163 by immunoprecipitation (Supplemental Figure 4). Consistent with prior reports, visual observations suggested decreased proliferative activity in lymphocyte cultures treated with unmanipulated AIS serum relative to those treated with serum from control groups, as they appeared to contain smaller and more diffuse clusters of blasting lymphocytes (Figure 5A). However, visual indications suggested that the inhibitory effect of AIS serum was at least moderately counteracted by the depletion of sCD163 (Figure 5A). In agreement with these visual observations, we observed significantly lower levels of BrDU incorporation in lymphocytes stimulated in the presence of AIS serum relative to those which were stimulated in the presence of serum derived from control groups, an effect which was partially ablated as a result of sCD163 depletion (Figure 5B). A second independent experiment using pooled serum samples derived from a separate sub-set of subjects (Supplemental Table 2) yielded similar results (Supplemental Figure 5). Taken together, these results
novelly suggest that post-stroke elevations in circulating levels of sCD163 can have direct inhibitory effects on
the proliferative capacity of lymphocytes.
Discussion:

The primary objective of this work was to determine the role of CD163 within the context of stroke immunopathology. We hypothesized that coordinate elevations in CD163 expression and ADAM17 activity induced by stroke could drive increases in circulating levels of sCD163, which could ultimately contribute to suppression of the peripheral adaptive immune system via the known inhibitory effects of sCD163 on lymphocyte activity. Our collective findings supported our hypothesis, and suggest that sCD163 plays a unique role as a factor which synergistically links stroke-induced activation of the innate immune system and suppression of the adaptive immune system. Such a mechanism illustrates a departure from the often compartmentalized view of stroke immunopathology, and highlights the necessity to approach the peripheral immune system from a global perspective in future work, especially with regards to immunotherapeutic design.

Observations of elevated levels of sCD163 in the peripheral blood of AIS patients, along with the fact that these elevations were negatively associated with lymphocyte counts, provided in vivo associative evidence supporting our hypothesis that CD163 plays a role in modulation of the adaptive immune system following stroke. Furthermore, our in vitro observations suggested that this stroke-driven increase in circulating sCD163 most likely originates from peripheral blood monocytes, as serum from AIS patients stimulated ADAM17-dependant production shedding in monocyte cultures generated from the peripheral blood of healthy individuals. Our in vitro observations further suggested that stroke-induced elevations in circulating sCD163 have the capacity to elicit direct suppressive effects on the adaptive immune system, as depletion of sCD163 from AIS patient-derived serum was able to partially rescue its capacity to support lymphocyte proliferation. Collectively, our findings infer a mechanism in which peripherally circulating factors induced by stroke trigger ADAM17-dependant CD163 shedding by monocytes, driving an increase in sCD163 levels which act to suppress peripheral lymphocyte activity, likely as means of mitigating the risk of an adaptive immune response directed against CNS antigens (Supplemental Figure 6).

To our knowledge, our results are the first to demonstrate a direct role for the innate immune system in mediating stroke-induced suppression of the peripheral adaptive immune system. While our results highlighted sCD163 as a means of such cross-communication, it is likely other innate-derived factors function to mediate the adaptive immune response to stroke via similar mechanisms. This notion is accentuated by the fact that complete ablation of sCD163 from stroke-derived serum was not adequate to fully rescue its capacity to support lymphocyte proliferation, inferring that there are a multitude of other soluble factors present in post-stroke peripheral circulation which have similar inhibitory effects. While the origins of these factors are inevitably diverse, it is likely that some arise via a similar mechanism as sCD163, as biological mechanisms are often redundant.30 Interesting from this regard is cluster of differentiation 14 (CD14), a membrane-bound pattern recognition receptor which is largely exclusive to the innate immune system.31,32 Much like CD163, CD14 can be shed via an MMP-dependent mechanism to generate a soluble protein (sCD14) which has been shown to directly interact with lymphocytes and exert suppressive effects.33,34 Preliminary results from our laboratory suggest that
circulating levels of sCD14 are also elevated in response to stroke, and further exploration may reveal that sCD14 plays similar role as sCD163 in modulation of the peripheral adaptive immune response.

While this study identified a clear role for sCD163 in stroke immunopathology in terms of its end-action, our experiments did not look to identify the soluble factors present in peripheral circulation following stroke which are responsible for triggering the shedding of CD163 from monocytes. It is plausible to believe that this phenomenon is likely driven by a synergistic combination of stroke-induced damage associated molecules (DAMs), cytokines, and hormones, as members of the aforementioned families of factors are known to stimulate CD163 shedding. Notable from this standpoint is extracellular hemoglobin, a DAM which has been previously demonstrated as being acutely elevated in peripheral circulation following stroke as a result of intravascular and extravascular hemolysis. Free hemoglobin has been characterized as a potent inducer of CD163 shedding, and thus, may help drive production of sCD163 in response to stroke. Along similar lines, levels of interleukin-6, cortisol, and reactive oxygen species, are all known to be elevated acutely in stroke pathology, and all have been shown to promote sCD163 production. Further exploration into these and other potential molecular signals which may trigger sCD163 generation within the context of stroke could not only provide a more complete picture of the mechanisms identified in this study, but a better understanding of post-stroke adaptive immune suppression as a whole, as the factors which drive post-stroke sCD163 production likely exert suppressive effects on the peripheral immune system via multiple parallel mechanisms.

While our findings are exciting, it is important to note that potentially confounding the interpretation of our results is the fact that elevations in sCD163 have been previously reported in atherosclerosis as a result of excessive intravascular macrophage activity within atherosclerotic plaques. As atherosclerosis is closely linked with stroke as one of its primary pathogenic drivers, it leaves open the possibility that our results were atherosclerosis driven, and not by the acute event of stroke itself. However, we find this scenario unlikely, as subjects in our control groups were well matched with AIS patients in terms of the prevalence of cardiovascular disease risk factors associated with atherosclerosis. Supporting our notion that the elevations in sCD163 which we observed in AIS were a direct result of the stroke-induced neurological insult is the fact that it has recently been reported that levels of macrophage-derived sCD163 are elevated in the cerebrospinal fluid of pediatric traumatic brain injury patients, providing concurrent evidence that cerebral trauma has the capacity to directly trigger increased sCD163 production. Furthermore, our findings, taken with theirs, suggest the possibility that stroke-induced elevations in sCD163 may not only modulate adaptive immune activity in the periphery, but could also serve a similar function within the lesion in terms of adaptive cellular infiltrates.

From a broader perspective, our findings are intriguing as they demonstrate a departure from the compartmentalized view of stroke immunopathology. While there are some exceptions, research regarding stroke immunopathology has long and too often treated the adaptive and innate immune responses as independent entities, however our results and those of select others suggest that these responses are likely synergistically linked via direct mechanisms. This realization is of utmost importance with regards to the current
push towards the development of novel stroke therapeutics which target the peripheral immune system. Currently, experimental immune therapies are being developed which largely aim to improve stroke outcome via one of two often discrete avenues: either by inhibiting the acute innate immune response to stroke as means of limiting excessive tissue damage and mitigating the risk of negative secondary cerebrovascular events such as edema and hemorrhagic transformation,\textsuperscript{44–46} or by post-acutely stimulating the peripheral adaptive immune system as a means of limiting the risk of complications induced by post-stroke infection.\textsuperscript{47–50} However, immunotherapeutics proposed along this line of thinking often fail to account the possible interplay between the adaptive and innate immune systems, and the potential repercussions that could stem from modulating one response without consideration for the other.

A perfect example in this regard is the fact that ADAM17 inhibitors are often cited as a potential stroke therapeutic under the premise that they could ultimately limit excessive innate-driven inflammation and its associated complications via a reduction in TNFα production.\textsuperscript{51–54} However, our results clearly demonstrate that inhibition of ADAM17 within the innate immune system could directly interfere with mechanisms which act to suppress the adaptive immune system, an unintended effect which could put patients at elevated risk for the development of long-term autoimmune complications associated with loss of CNS self-tolerance. Such an example highlights the necessity for a global view regarding the peripheral immune system in terms of the development of novel stroke immunotherapeutics. Unfortunately, therapeutic design from such a perspective is currently unrealistic, as the mechanisms which drive the stroke-induced peripheral immune response, particularly that of the adaptive immune system, are still largely unknown. This current gap in knowledge may underlie the reality that stroke immunotherapeutics have thus-far been largely unsuccessful.\textsuperscript{55} Further future work which aims to characterize the peripheral immune response to stroke and its underlying mechanisms from a panoptic systems biology approach, preferably within human-centric models, would likely generate the foundational basis for the development of stroke immunotherapeutics with higher chances for clinical success.

Collectively, our results demonstrate a novel role for CD163 as factor which mechanistically couples stroke induced-activation of the peripheral innate immune system and suppression of the peripheral adaptive immune system. Our results provide a unique example which suggests that the innate immune system employs protective mechanisms aimed at mitigating the risk of post-stroke autoimmune complications driven by adaptive immune system overactivation. Further work which looks to characterize similar pathways of cross-communication between the innate and adaptive immune systems within the context of stroke would likely provide a better understanding of stroke pathology as a whole, and allow for more informed development of stroke-targeted immunotherapeutics.
Methods:

Patients:

AIS patients, acute stroke mimics, and neurologically asymptomatic controls were recruited at Ruby Memorial Hospital, Morgantown, WV. AIS cases were of mixed etiology, and diagnosis was confirmed by neuroradiological imaging according to the established criteria for diagnosis of acute ischemic cerebrovascular syndrome. Patients admitted to the emergency department as suspected strokes based on the overt presentation of stroke-like symptoms, but receiving a negative diagnosis for stroke upon imaging were identified as acute stroke mimics. Discharge diagnoses of stroke mimics included cases of seizures, complex migraines, and other conditions which induce neurological symptoms such as hypertensive encephalopathy. Blood was sampled within 24 hours of onset, as determined by the time the patient was last known to be free of AIS symptoms. In the case of patients who received thrombolytic therapy, blood samples were collected prior to the administration of recombinant tissue plasminogen activator (rtPA). Injury severity was determined according to NIHSS at the time of blood draw. Control subjects were deemed neurologically normal by a trained neurologist at the time of enrolment. Demographic information was collected from either the subject or significant other by a trained clinician. All procedures were approved by the institutional review boards of West Virginia University and Ruby Memorial Hospital. Written informed consent was obtained from all subjects or their authorized representatives prior to study procedures.

Neuroradiological imaging:

Neuroradiological imaging was performed using either MRI or CT within 24 hours of symptom onset. CT imaging was performed on either a Toshiba Aquilion 64 or Toshiba Aquilion 320 scanner (5 mm slices acquired at 120 KVp, 250 mA). MRI imaging was performed using either a Siemens Aera or Siemens Verio scanner (5 mm slices, 5 mm inter-slice gap, acquired via diffusion-weighted echoplanar imaging with a B-factor of 1000 s/mm²). The BrainLAB iPlan Neuroradiology software package (BrainLAB, Westchester, Ill) was used to calculate infarct volume via manual tracing as previously described, and all infarct volume calculations were verified by an experienced neurologist.

Blood collection:

Parallel peripheral intravenous whole blood samples were collected from subjects via PAXgene RNA tubes (Qiagen, Valencia, CA), K₂EDTA Vacutainers (Becton Dickenson, Franklin Lakes, NJ), and serum separator tubes (Becton Dickenson). PAXgene RNA tubes were frozen immediately and stored at -80°C until RNA extraction. K₂EDTA tubes were stored at room temperature until white blood cell differential (less than 30 minutes), plasma isolation (less than 30 minutes), or leukocyte isolation (performed immediately). Serum separator tubes were stored at room temperature until serum isolation (less than 30 minutes).
Isolation of serum and plasma:

For plasma isolation, EDTA-treated blood was spun at 2,000*g for 10 minutes to sediment hemocytes. Plasma was collected for ELISA, and the total hemocyte fraction was retained for subsequent protein extraction. Plasma was additionally spun at 10,000*g for 10 minutes to remove any residual blood cells or debris. For serum isolation, serum separator tubes were incubated at room temperature for a minimum of 15 minutes to allow for clot formation, and subsequently spun at 2,000*g for 10 minutes to sediment the resulting thrombus. Resultant samples were stored at -80ºC until analysis.

RNA extraction:

Whole blood RNA was extracted from PAXgene tubes via the PreAnalytiX PAXgene blood RNA kit (Qiagen) and automated using the QIAcube system (Qiagen). RNA was isolated from leukocyte lysates via the RNeasy Micro kit (Qiagen). Quantity and purity of isolated RNA was determined via spectrophotometry (NanoDrop, Thermo Scientific, Waltham, MA). Quality of RNA was confirmed by chip capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA).

Quantitative reverse transcription PCR:

cDNA was generated from purified RNA using the Applied Biosystems high capacity reverse transcription kit. For qPCR, target sequences were amplified from 10 ng of cDNA input using sequence specific primers (Supplemental Table 3) and detected via SYBR green (PowerSYBR, Thermo-Fisher) on the RotorGeneQ (Qiagen). Raw amplification plots were background corrected and CT values were generated via the RotorGeneQ software package. All reactions were performed in triplicate. Transcripts of B2M, PPIB, and ACTB were amplified as references and normalization was performed using the NORMAgene data-driven normalization algorithm.58

White blood cell differential:

Complete blood count was obtained from EDTA-treated blood via combined optical flow cytometry and cellular impedance using an automated clinical hematology analyzer (Cell-Dyn, Abbott Diagnostics, Santa Clara, CA).

Isolation of primary leukocytes for cell culture:

Starting with 30 mLs of healthy donor-derived EDTA-treated blood, peripheral blood mononuclear cells (PBMCs) were separated from polymorphonuclear cells (PMNs) and RBCs via centrifugation over a polysaccharide-sodium diatrizoate density gradient (Lymphoprep, 1.077 g/mL, StemCell Technologies) for 30 minutes at 400*g. Resultant interphase PBMC fractions were collected, resuspended in RPMI 1640 (Life Technologies, Grand Island, NY) containing 10% autologous serum, separated into two aliquots, and set aside at room temperature. Pelleted PMN/RBCs were immediately placed on ice for PMN enrichment via RBC lysis induced by two 5 minute incubations with a 1:10 volume of ice cold ammonium chloride (ACK) buffer.
resultant PMNs were rinsed in ice cold PBS, counted via automated cytometer (Cellmeter X1, Nexcelom Bioscience, Lawrence, MA), and immediately plated for experiments.

Following PNM isolation, monocytes were further enriched from one aliquot of PBMCs via a second round of density gradient centrifugation as described by Menck et al. Briefly, PBMC suspensions were layered over an iso-osmotic 46% Percoll (1.131 g/mL, Sigma Aldrich) solution prepared in RPMI 1640 containing 10% autologous serum and centrifuged at 500 g for 30 minutes. Monocytes were collected from the resultant interphase, rinsed in PBS, counted, and immediately plated for experiments.

Following PMN and monocyte isolation, the remaining aliquot of PBMCs was rinsed in PBS and seeded in T75 flasks for lymphocyte expansion. Lymphocytes were expanded for 72 hours under standard mammalian cell culture conditions in a growth media comprised of RPMI 1640 containing 2% PHA-M (Gibco, Grand Island, NY), 2 mM L-Alanyl-Glutamine (GlutaMAX, Gibco), 25 mM HEPES (Gibco), 50 uM BME (Gibco), and 1% antibiotic antimycotic (Gibco), supplemented with 10% autologous serum. Following expansion, actively proliferating lymphocytes were rinsed in PBS, counted, and plated for experiments.

Neutrophil culture:

Neutrophils from healthy donors were seeded in 12-well plates at a density of 7.5*10^5 cells per well in 250 uL of RPMI 1640 containing 10% patient serum along with either Marimastat (10 uM, Sigma Aldrich, St-Louis, MO), TAPI-0 (10 uM, Sigma Aldrich), or vehicle (DMSO, Fisher Scientific). Following three hours of incubation under standard mammalian culture conditions, cells were lysed in NP40 lysis buffer (Life Technologies) and cellular ADAM17 activity was assessed via enzyme activity assay. Cell culture supernatants were collected and concentrations of sCD163 were quantified via ELISA; levels of neutrophil-derived sCD163 were determined by subtracting the concentration of sCD163 in media incubated without cells from the concentrations of sCD163 observed in cell culture supernatants.

Monocyte culture:

Monocytes from healthy donors were seeded in 12-well plates at a density of 1*10^6 cells per well in 250 uL of RPMI 1640 containing 20% patient serum along with either Marimastat (10 uM), TAPI-0 (10 uM), or vehicle (DMSO). Following three hours of incubation under standard mammalian culture conditions, cells were lysed in NP40 lysis buffer and cellular ADAM17 activity was assessed via enzyme activity assay. Cell culture supernatants were collected and concentrations of sCD163 were quantified via ELISA; levels of monocytes-derived sCD163 were determined by subtracting the concentration of sCD163 in media incubated without cells from the concentrations of sCD163 observed in cell culture supernatants.

sCD163 depletion from human serum:

sCD163 was depleted from pooled serum samples via immunoprecipitation using biotinylated goat polyclonal antibody raised against the extracellular region of CD163 (BAF1607, R&D Systems, Minneapolis,
MN). Biotinylated normal goat IgG (BAF108, R&D Systems) was used as an immunoprecipitation control. Immunoglobulins were conjugated to polymer-coated superparamagnetic beads (Dynabeads Streptavidin T1, Thermo Fisher, Waltham MA) at a ratio of 40 ug of immunoglobulins per 1 mg of beads. For immunoprecipitation, 5 mL of pooled serum was precleared with 2.5 mg of unconjugated beads, and then incubated with 2.5 mg of conjugated beads for 12 hours at 4°C. Beads were subsequently separated from serum samples via multiple rounds of magnetic separation. Following immunoprecipitation, pooled serum was filtered at 14 microns to remove potential microbial contaminants prior to cell culture. Depletion of sCD163 was subsequently confirmed via ELISA (Supplemental Figure 4).

**Lymphocyte proliferation assay:**

Expanded lymphocytes from healthy donors were seeded in 48 well plates at 5*10^4 cell per well in 200 uL growth media supplemented 20% pooled patient serum. Following 48 hours of culture, BrdU (Roche Life Sciences, Indianapolis, IN) was added to cultures at a final concentration of 20 uM. After allowing for 24 hours for BrdU incorporation, lymphocytes were adhered via centrifugation and fixed for BrdU quantification. Incorporated BrdU was quantified via colorimetric ELISA (Roche Life Sciences) per manufacture instructions.

**sCD163 ELISA:**

sCD163 was measured in plasma samples and cell culture supernatants using a commercially available ELISA assay (RAB0082, Sigma-Aldrich). Plasma samples were diluted 1:50, neutrophil cell culture supernatants were diluted 1:5, and monocyte cultures were diluted 1:10 prior to analysis. Due to the potential confound of hemolysis on concentrations of leukocyte derived analytes, plasma samples were screened for hemolysis prior to ELISA. Plasma absorbance was measured at 385 and 414 nm via spectrophotometry (NanoDrop, Thermo Scientific, Waltham, MA) and used to calculate a hemolysis score as described by Appierto et al.\(^6\) Hemoglobin-depleted plasma (HemogloBind, Biotech Support Group, Monmouth Junction, NJ) spiked with serial dilutions of sonicated red blood cells were used as a positive control. Plasma samples with detectable hemolysis were excluded from analysis.

**ADAM17 activity assay:**

Total hemocyte fractions were thawed, mixed with NP-40 lysis buffer (Thermo Fisher) at a 1 to 1 ratio, and depleted of hemoglobin via the HemogloBind hemoglobin removal kit (Biotech Support Group, Monmouth Junction, NJ) as described by Park et al.\(^6\) Protein concentrations of hemoglobin-depleted total hemocyte lysates and cell culture lysates were determined via DC-protein assay (Bio-Rad, Hercules, CA). ADAM17 activity was measured in 25 ug of total hemocyte lysate, or 10 ug of cell culture lysate, via the InnoZyme TACE activity kit (EMD Millipore, Temecula, CA) per manufacture instructions. Due to the reversible nature of ADAM17 inhibition by both marimastat and TAPI-0, experimental concentrations of drug were maintained in assay buffers when testing samples from inhibitor-treated conditions.
Phase Contrast Microscopy:

Phase contrast images of primary leukocyte cultures were obtained from an AxioObserver inverted microscope (Ziess, Thornwood, NY) equipped with a AxioCam mR5 digital camera (Ziess) using the AxioVision imaging software suite (Ziess). Scale bars were generated via imaging of a 0.07-1.50 mm scale calibration slide (Motic, Richmond, BC, CA).

Statistical analysis:

Parametric statistical analysis was performed using SPSS (IBM, Chicago, Ill) in combination with R 2.14 via the SPSS R integration plug-in. Chi-squared tests were used for comparison of dichotomous variables. Student t-tests or one-way ANOVA were used for comparison of continuous variables. Spearman's rho was used to assess the strength of correlational relationships. The level of significance was established at 0.05 for all parametric statistical testing. In the cases of multiple comparisons, p-values were adjusted using the Bonferonni method.
Acknowledgments:

The authors would foremost like to thank the subjects and their families, as this work was truly made possible by their selfless contribution. We also would like to thank the stroke team Ruby Memorial Hospital for supporting this research effort.

Contributions:

Work was conceptualized by GCO and TLB. Procedures for collection of clinical samples and recruitment of human subjects were overseen by TLB and PDC. Recruitment of subjects and collection of samples were performed by GCO, ABP, NL, and CST. Experiments were designed and performed by GCO. Data were analyzed by GCO. Manuscript was written by GCO with contributions from TLB, ABP, NL, CST, and PDC.

Funding:

Work was partially funded via a Robert Wood Johnson Foundation Nurse Faculty Scholar award to TLB (70319) and a National Institutes of Health CoBRE sub-award to TLB (P20 GM109098).

Competing interests:

GCO and TLB have a patent pending re: genomic patterns of expression for stroke diagnosis. TLB serves as chief scientific officer for CereDx Incorporated, a biotech firm which develops diagnostics for brain injury. The remaining authors report no potential conflicts of interest.
References:


37. Olsson T, Astrom M, Eriksson S, Forssell A. Hypercortisolism revealed by the dexamethasone


47. Meisel C, Meisel A. clinical implications of basic research Suppressing Immunosuppression after Stroke. 2015.


Figure Legends:

Table 1. Clinical and demographic characteristics.

Table 2. Clinical and demographic characteristics of subject sub-populations used for in vitro experiments.

Figure 1. ADAM17 and CD163 expression in the peripheral blood following ischemic stroke.

mRNA expression of (A) ADAM17 and (B) CD163 in peripheral whole blood obtained from AIS patients and controls, with expression levels presented as fold difference relative to asymptomatic control group. (C) Cellular ADAM17 activity in the peripheral whole blood of AIS patients and controls. (D) Plasma concentrations of sCD163 in AIS patients and controls. (E) Relationship between peripheral blood cellular ADAM17 activity and NIHSS within the AIS group. (F) Relationship between plasma sCD163 levels and NIHSS within the AIS group. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferonni post-hoc, while strength of correlational relationships were tested via Spearman’s rho.

Figure 2. Relationship between plasma sCD163 levels and post-stroke peripheral immune status.

Relationships between plasma sCD163 levels and (A) absolute lymphocyte counts, (B) absolute monocyte counts, and (C) absolute neutrophil counts in the peripheral blood of AIS patients. strength of correlational relationships were tested via Spearman’s rho.

Figure 3. Effects of ischemic stroke serum on neutrophil ADAM17-dependant sCD163 production.

(A) Morphology of healthy donor-derived neutrophils following three hours incubation with 10% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. (B) ADAM17 activity in neutrophil lysates collected following treatment. (C) Neutrophil-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferonni post-hoc.

Figure 4. Effects of ischemic stroke serum on monocyte ADAM17-dependant sCD163 production.

(A) Morphology of healthy donor-derived monocytes following three hours incubation with 20% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. White arrowheads indicate cytoplasmic vacuoles. (B) ADAM17 activity in monocyte lysates collected following treatment. (C) Monocyte-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferonni post-hoc.
Figure 5. Influence of post-stroke peripheral blood sCD163 levels on the capacity to support lymphocyte proliferation.

(A) Morphology of healthy donor-derived lymphocytes following 72 hours of PHA-stimulated proliferation in the presence of 20% serum obtained from AIS patients and controls, which was either unmanipulated in terms of sCD163 levels, or depleted of sCD163 via immunoprecipitation using anti-CD163 monoclonal antibody. White arrowheads indicate clusters of blasting lymphocytes. (B) BrdU incorporation over the final 24 hours of treatment. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferroni post-hoc.

Supplemental Table 1. Demographic characteristics of healthy leukocyte donors.

Supplemental Table 2. Clinical and demographic characteristics of subject subpopulations used for independent replicate in vitro experiments.

Supplemental Figure 1. Relationships between post-stroke peripheral blood cellular ADAM17 activity and plasma sCD163 levels and infarct volume.

(A) Relationship between peripheral blood cellular ADAM17 activity and infarct volume within the validation cohort AIS group. (B) Relationship between plasma sCD163 levels and infarct volume within the validation cohort AIS group. Strength of correlational relationships were tested via Spearman’s rho.

Supplemental Figure 2. Effects of ischemic stroke serum on neutrophil ADAM17-dependant sCD163 production in an independent replicate experiment.

(A) ADAM17 activity in healthy-donor neutrophil lysates collected following three hours incubation with 10% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. (B) Neutrophil-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferroni post-hoc.

Supplemental Figure 3. Effects of ischemic stroke serum on monocyte ADAM17-dependant sCD163 production in an independent replicate experiment.

(A) ADAM17 activity in healthy-donor monocyte lysates collected following three hours incubation with 10% serum obtained from AIS patients and control subjects, in either the presence and absence of ADAM17 inhibitors. (B) Monocyte-derived sCD163 levels in cell culture supernatants collected following treatment, presented as the difference in sCD163 levels observed between cell culture supernatants and serum-supplemented media incubated in the absence of cells. Intergroup comparisons were made via one-way ANOVA with subsequent Bonferroni post-hoc.
**Supplemental Figure 4.** Confirmation of serum sCD163 depletion using immunoprecipitation.

(A) Concentrations of sCD163 in pooled serum samples obtained from validation cohort AIS patients and controls measured with ELISA following immunoprecipitation via either anti-CD163 polyclonal antibody or isotype control immunoglobulin.

**Supplemental Figure 5.** Influence of post-stroke peripheral blood sCD163 levels on the capacity to support lymphocyte proliferation in an independent replicate experiment.

(A) BrdU incorporation in healthy donor-derived lymphocytes over the final 24 of 72 hours PHA-stimulated proliferation in the presence of 20% serum obtained from AIS patients and controls, which was either unmanipulated in terms of sCD163 levels, or depleted of sCD163 via immunoprecipitation using anti-CD163 monoclonal antibody.

**Supplemental Figure 6.** Proposed role for CD163 in modulation of the stroke-induced peripheral immune response.

Stroke-driven elevations in circulating levels of damage associated molecules (DAMs), proinflammatory cytokines, and hormones coordinately trigger innate immune system activation and ADAM17-dependant CD163 shedding by monocytes, resulting in increased generation of sCD163. Heightened circulating sCD163 levels subsequently act to suppress lymphocyte activity, likely as a means of limiting the risk of a CNS-directed autoimmune response resulting from adaptive immune system activation.
Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic (n=20)</th>
<th>Stroke mimic (n=20)</th>
<th>Ischemic stroke (n=39)</th>
<th>Stat (df)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>58.6 ± 11.1</td>
<td>58.0 ± 17.0</td>
<td>73.1 ± 13.3</td>
<td>F= 11.2 (2, 76)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>15 (75.0)</td>
<td>9 (45.0)</td>
<td>25 (64.1)</td>
<td>χ²= 3.96 (2)</td>
<td>0.138</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>0.0 ± 0.0</td>
<td>4.7 ± 4.9</td>
<td>8.6 ± 7.5</td>
<td>F= 14.6 (2, 76)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>11 (55.0)</td>
<td>5 (25.0)</td>
<td>15 (38.5)</td>
<td>χ²= 3.79 (2)</td>
<td>0.151</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>15 (75.0)</td>
<td>17 (85.0)</td>
<td>32 (82.1)</td>
<td>χ²= 0.71 (2)</td>
<td>0.703</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>8 (40.0)</td>
<td>13 (65.0)</td>
<td>16 (41.0)</td>
<td>χ²= 3.55 (2)</td>
<td>0.169</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>2 (10.0)</td>
<td>7 (35.0)</td>
<td>8 (20.5)</td>
<td>χ²= 3.75 (2)</td>
<td>0.153</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>1 (5.00)</td>
<td>5 (25.0)</td>
<td>7 (17.9)</td>
<td>χ²= 3.03 (2)</td>
<td>0.219</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>0 (00.0)</td>
<td>3 (15.0)</td>
<td>18 (33.3)</td>
<td>χ²= 9.55 (2)</td>
<td>0.008*</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>0 (00.0)</td>
<td>6 (30.0)</td>
<td>11 (28.2)</td>
<td>χ²= 7.38 (2)</td>
<td>0.025*</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>14 (70.0)</td>
<td>16 (80.0)</td>
<td>27 (69.2)</td>
<td>χ²= 0.82 (2)</td>
<td>0.662</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>2 (10.0)</td>
<td>6 (30.0)</td>
<td>8 (20.5)</td>
<td>χ²= 2.48 (2)</td>
<td>0.299</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>6 (30.0)</td>
<td>12 (60.0)</td>
<td>14 (35.9)</td>
<td>χ²= 4.41 (2)</td>
<td>0.111</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>1 (5.00)</td>
<td>12 (60.0)</td>
<td>23 (59.0)</td>
<td>χ²= 17.8 (2)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>1 (5.00)</td>
<td>0 (00.0)</td>
<td>13 (33.3)</td>
<td>χ²= 13.0 (2)</td>
<td>0.001*</td>
</tr>
<tr>
<td>rtPA n (%)</td>
<td>0 (00.0)</td>
<td>2 (10.0)</td>
<td>9 (23.1)</td>
<td>χ²= 8.21 (2)</td>
<td>0.044*</td>
</tr>
<tr>
<td></td>
<td>Asymptomatic (n=10)</td>
<td>Stroke mimic (n=10)</td>
<td>Ischemic stroke (n=10)</td>
<td>Stat (df)</td>
<td>P</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>------------------------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>57.5 ± 11.1</td>
<td>56.4 ± 17.1</td>
<td>63.5 ± 14.3</td>
<td>f= 0.63 (2, 27)</td>
<td>0.536</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>8 (80.0)</td>
<td>5 (50.0)</td>
<td>5 (50.0)</td>
<td>$x^2$= 2.50 (2)</td>
<td>0.286</td>
</tr>
<tr>
<td>NIHSS (mean ± SD)</td>
<td>0.0 ± 0.0</td>
<td>1.6 ± 2.4</td>
<td>6.0 ± 6.9</td>
<td>f= 5.96 (2, 27)</td>
<td>0.007*</td>
</tr>
<tr>
<td>Family history of stroke n (%)</td>
<td>5 (50.0)</td>
<td>2 (20.0)</td>
<td>4 (40.0)</td>
<td>$x^2$= 2.01 (2)</td>
<td>0.366</td>
</tr>
<tr>
<td>Hypertension n (%)</td>
<td>9 (90.0)</td>
<td>6 (60.0)</td>
<td>6 (60.0)</td>
<td>$x^2$= 0.48 (2)</td>
<td>0.786</td>
</tr>
<tr>
<td>Dyslipidemia n (%)</td>
<td>5 (50.0)</td>
<td>7 (70.0)</td>
<td>4 (40.0)</td>
<td>$x^2$= 1.87 (2)</td>
<td>0.391</td>
</tr>
<tr>
<td>Diabetes n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>$x^2$= 0.00 (2)</td>
<td>1.000</td>
</tr>
<tr>
<td>Previous stroke n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>$x^2$= 0.57 (2)</td>
<td>0.749</td>
</tr>
<tr>
<td>Atrial fibrillation n (%)</td>
<td>0 (00.0)</td>
<td>2 (20.0)</td>
<td>3 (30.0)</td>
<td>$x^2$= 3.36 (2)</td>
<td>0.166</td>
</tr>
<tr>
<td>Myocardial infarction n (%)</td>
<td>0 (00.0)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>$x^2$= 2.22 (2)</td>
<td>0.329</td>
</tr>
<tr>
<td>Hypertension medication n (%)</td>
<td>9 (90.0)</td>
<td>8 (80.0)</td>
<td>7 (70.0)</td>
<td>$x^2$= 1.25 (2)</td>
<td>0.535</td>
</tr>
<tr>
<td>Diabetes medication n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>$x^2$= 0.00 (2)</td>
<td>1.000</td>
</tr>
<tr>
<td>Cholesterol medication n (%)</td>
<td>2 (20.0)</td>
<td>7 (70.0)</td>
<td>3 (30.0)</td>
<td>$x^2$= 5.83 (2)</td>
<td>0.054</td>
</tr>
<tr>
<td>Anticoagulant or antiplatelet n (%)</td>
<td>1 (10.0)</td>
<td>5 (50.0)</td>
<td>6 (60.0)</td>
<td>$x^2$= 5.83 (2)</td>
<td>0.054</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>2 (20.0)</td>
<td>$x^2$= 0.57 (2)</td>
<td>0.749</td>
</tr>
<tr>
<td>rtPA n (%)</td>
<td>0 (00.0)</td>
<td>0 (00.0)</td>
<td>3 (30.0)</td>
<td>$x^2$= 6.87 (2)</td>
<td>0.035*</td>
</tr>
</tbody>
</table>
Figure 2.

A

LYMPHOCYTE COUNT
(10^6 CELLS/mL)

sCD163 (ng/mL)

B

MONOCYTE COUNT
(10^6 CELLS/mL)

sCD163 (ng/mL)

C

NEUTROPHIL COUNT
(10^6 CELLS/mL)

sCD163 (ng/mL)

r=-0.334
p=0.045

r=0.352
p=0.035

r=0.032
p=0.046
Figure 3.
Figure 4.
Figure 5.
Supplemental Table 1.

<table>
<thead>
<tr>
<th></th>
<th>n=12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>25.5 ± 3.9</td>
</tr>
<tr>
<td>Female n (%)</td>
<td>6 (50.0)</td>
</tr>
<tr>
<td>History of metabolic syndrome n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>History of cardiovascular disease n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>History of autoimmune disease n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>History of cancer n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Prior or current pregnancy n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>History of head injury n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>History of neurological disease n (%)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Current smoker n (%)</td>
<td>0 (0.00)</td>
</tr>
</tbody>
</table>
Supplemental Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic (n=10)</th>
<th>Stroke mimic (n=10)</th>
<th>Ischemic stroke (n=10)</th>
<th>Stat (df)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (mean ± SD)</strong></td>
<td>59.7 ± 10.5</td>
<td>59.6 ± 15.9</td>
<td>64.8 ± 8.8</td>
<td>t = 0.54 (2, 27)</td>
<td>0.589</td>
</tr>
<tr>
<td><strong>Female n (%)</strong></td>
<td>7 (70.0)</td>
<td>4 (40.0)</td>
<td>4 (40.0)</td>
<td>χ² = 2.40 (2)</td>
<td>0.301</td>
</tr>
<tr>
<td><strong>NIHSS (mean ± SD)</strong></td>
<td>0.0 ± 0.0</td>
<td>6.9 ± 4.9</td>
<td>5.9 ± 3.7</td>
<td>t = 0.66 (2, 27)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td><strong>Family history of stroke n (%)</strong></td>
<td>6 (60.0)</td>
<td>3 (30.0)</td>
<td>6 (60.0)</td>
<td>χ² = 2.40 (2)</td>
<td>0.301</td>
</tr>
<tr>
<td><strong>Hypertension n (%)</strong></td>
<td>6 (60.0)</td>
<td>9 (90.0)</td>
<td>8 (80.0)</td>
<td>χ² = 2.61 (2)</td>
<td>0.271</td>
</tr>
<tr>
<td><strong>Dyslipidemia n (%)</strong></td>
<td>3 (30.0)</td>
<td>6 (60.0)</td>
<td>5 (50.0)</td>
<td>χ² = 1.87 (2)</td>
<td>0.391</td>
</tr>
<tr>
<td><strong>Diabetes n (%)</strong></td>
<td>2 (20.0)</td>
<td>6 (60.0)</td>
<td>5 (50.0)</td>
<td>χ² = 3.53 (2)</td>
<td>0.171</td>
</tr>
<tr>
<td><strong>Previous stroke n (%)</strong></td>
<td>0 (00.0)</td>
<td>4 (40.0)</td>
<td>0 (00.0)</td>
<td>χ² = 9.23 (2)</td>
<td>0.009*</td>
</tr>
<tr>
<td><strong>Atrial fibrillation n (%)</strong></td>
<td>0 (00.0)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>χ² = 1.07 (2)</td>
<td>0.585</td>
</tr>
<tr>
<td><strong>Myocardial infarction n (%)</strong></td>
<td>0 (00.0)</td>
<td>5 (50.0)</td>
<td>4 (40.0)</td>
<td>χ² = 6.67 (2)</td>
<td>0.035*</td>
</tr>
<tr>
<td><strong>Hypertension medication n (%)</strong></td>
<td>5 (50.0)</td>
<td>8 (80.0)</td>
<td>6 (60.0)</td>
<td>χ² = 2.01 (2)</td>
<td>0.366</td>
</tr>
<tr>
<td><strong>Diabetes medication n (%)</strong></td>
<td>1 (10.0)</td>
<td>5 (50.0)</td>
<td>5 (50.0)</td>
<td>χ² = 4.59 (2)</td>
<td>0.101</td>
</tr>
<tr>
<td><strong>Cholesterol medication n (%)</strong></td>
<td>4 (40.0)</td>
<td>5 (50.0)</td>
<td>5 (50.0)</td>
<td>χ² = 0.26 (2)</td>
<td>0.874</td>
</tr>
<tr>
<td><strong>Anticoagulant or antiplatelet n (%)</strong></td>
<td>0 (00.0)</td>
<td>7 (70.0)</td>
<td>5 (50.0)</td>
<td>χ² = 10.8 (2)</td>
<td>0.004*</td>
</tr>
<tr>
<td><strong>Current smoker n (%)</strong></td>
<td>0 (00.0)</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>χ² = 1.07 (2)</td>
<td>0.585</td>
</tr>
<tr>
<td><strong>rtPA n (%)</strong></td>
<td>0 (00.0)</td>
<td>0 (00.0)</td>
<td>5 (50.0)</td>
<td>χ² = 12.0 (2)</td>
<td>0.002*</td>
</tr>
</tbody>
</table>
Supplemental Figure 1.

A

INJURED VOLUME (mL)

300
250
200
150
100
50
0

ADAM17 ACTIVITY (RFU/mg)

1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000

B

INJURED VOLUME (mL)

300
250
200
150
100
50
0

sCD163 (ng/mL)

50 100 150 200 250 300 350 400 450 500 550 600 650

r = 0.037
p = 0.820

r = 0.234
p = 0.234
Supplemental Figure 2.
Supplemental Figure 3.
Supplemental Figure 4.
Supplemental Figure 6.