(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 April 2006 (06.04.2006)

PCT

(10) International Publication Number
WO 2006/036220 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01) *G01N 33/48* (2006.01)
G06F 19/00 (2006.01) *G01N 33/50* (2006.01)

(21) International Application Number:
PCT/US2005/017274

(22) International Filing Date: 18 May 2005 (18.05.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/948,834 22 September 2004 (22.09.2004) US
11/046,593 29 January 2005 (29.01.2005) US

(71) Applicant *(for all designated States except US)*: **PREDICTION SCIENCES LLC** [—/US]; A California Limited Liability Corporation, 9404 Genesee Avenue, Suite 210, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CASTELLANO, Mar** [ES/ES]; Department of Neurology, Hospital Universitario de Girona, Doctor Josep Trueta Avide Franca, E-17007 Girona (ES). **DAVALOS, Antoni** [ES/ES]; Department of Neurology, Hospital Universitario de Girona, Doctor Josep Trueta Avide Franca, E-17007 Girona (ES). **DIAMOND, Cornelius, Allen** [US/US]; Prediction Sciences LLC, 9404 Genesee Avenue Suite 210, La Jolla, CA 92037 (US).

(74) Agent: **FUESS, William, Craig**; Fuess & Davidenas, 10951 Sorrento Valley Road, Suite 2G, San Diego, CA 92121-1613 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2006/036220 A2

(54) Title: CELLULAR FIBRONECTIN AS A DIAGNOSTIC MARKER IN STROKE AND METHODS OF USE THEREOF

(57) Abstract: The present invention relates to methods for the diagnosis and evaluation of stroke and stroke sub-type. A variety of bio-markers are disclosed for assembling a panel for such diagnosis and evaluation. Methods are disclosed for selecting markers and correlating their combined levels with a clinical outcome of interest. In various aspects: the invention provides methods for early detection and differentiation of stroke subtypes, for determining the prognosis of a patient presenting with stroke symptoms, and identifying a patient at risk for hemorrhagic transformation after thrombolyic therapy. Methods are disclosed that provide rapid, sensitive and specific assays to greatly increase the number of patients that can receive beneficial stroke treatment and therapy, and reduce the costs associated with incorrect stroke diagnosis.

Cellular fibronectin as a diagnostic marker in stroke and methods of use thereof

RELATED APPLICATIONS

[0001] The present application is descended from, and claims benefit of priority of, U.S. provisional patent application number 60/505,606, the contents of which of which are hereby incorporated herein in their entirety, including all tables, figures, and claims. The present application is a continuation-in-part of U.S. utility patent application number 10/948,834, which application is itself descended from U.S. provisional patent applications 60/505,606 and 60/556,411

FIELD OF THE INVENTION

[0002] The present invention generally relates to the identification and use of diagnostic markers for cardiovascular disease and cerebral injury. In a various aspects, the present invention particularly relates to methods for (1) the early detection and differentiation of cardiovascular events stroke and transient ischemic attacks, and (2) the identification of individuals at risk for hemorrhagic transformation after both presentation with stroke symptoms and subsequent administration of tissue plasminogen activator (tPA) therapy.

BACKGROUND OF THE INVENTION

[0003] The following discussion of the background of the invention is merely provided to aid the reader in understanding the invention and is not admitted to describe or constitute prior art to the present invention.

[0004] A stroke is a sudden interruption in the blood supply of the brain. Most strokes are caused by an abrupt blockage of arteries leading to the brain (ischemic stroke). Other strokes are caused by bleeding into brain tissue when a blood vessel bursts (hemorrhagic stroke). Because stroke occurs rapidly and requires immediate treatment, stroke is also called a brain attack. When the symptoms of a stroke last only a short time (less than an hour), this is called a

transient ischemic attack (HT) or mini-stroke. Stroke has many consequences.

[0005] The effects of a stroke depend on which part of the brain is injured, and how severely it is injured. A stroke may cause sudden weakness, loss of sensation, or difficulty with speaking, seeing, or walking. Since different parts of the brain control different areas and functions, it is usually the area immediately surrounding the stroke that is affected. Sometimes people with stroke have a headache, but stroke can also be completely painless. It is very important to recognize the warning signs of stroke and to get immediate medical attention if they occur.

[0006] Stroke or brain attack is a sudden problem affecting the blood vessels of the brain. There are several types of stroke, and each type has different causes. The three main types of stroke are listed below.

[0007] Ischemic stroke is the most common type of stroke -- accounting for almost 80% of strokes – an is caused by a clot or other blockage within an artery leading to the brain.

[0008] Intracerebral hemorrhage is a type stroke caused by the sudden rupture of an artery within the brain. Blood is then released into the brain, compressing brain structures.

[0009] Subarachnoid hemorrhage is also a type of stroke caused by the sudden rupture of an artery. A subarachnoid hemorrhage differs from an intracerebral hemorrhage in that the location of the rupture leads to blood filling the space surrounding the brain rather than inside of it.

[0010] Ischemic stroke occurs when an artery to the brain is blocked. The brain depends on its arteries to bring fresh blood from the heart and lungs. The blood carries oxygen and nutrients to the brain, and takes away carbon dioxide and cellular waste. If an artery is blocked, the brain cells (neurons) cannot make enough energy and will eventually stop working. If the artery remains blocked for more than a few minutes, the brain cells may die. This is why immediate medical treatment is absolutely critical.

[0011] Ischemic stroke can be caused by several different kinds of diseases. The most common problem is narrowing of the arteries in the neck or head. This is most often caused

atherosclerosis, or gradual cholesterol deposition. If the arteries become too narrow, blood cells may collect and form blood clots. These blood clots can block the artery where they are formed (thrombosis), or can dislodge and become trapped in arteries closer to the brain (embolism). Another cause of stroke is blood clots in the heart, which can occur as a result of irregular heartbeat (for example, atrial fibrillation), heart attack, or abnormalities of the heart valves. While these are the most common causes of ischemic stroke, there are many other possible causes. Examples include use of street drugs, traumatic injury to the blood vessels of the neck, or disorders of blood clotting.

[0012] Ischemic stroke can further be divided into two main types: thrombotic and embolic.

[0013] A thrombotic stroke occurs when diseased or damaged cerebral arteries become blocked by the formation of a blood clot within the brain. Clinically referred to as cerebral thrombosis or cerebral infarction, this type of event is responsible for almost 50% of all strokes. Cerebral thrombosis can also be divided into an additional two categories that correlate to the location of the blockage within the brain: large-vessel thrombosis and small-vessel thrombosis. Large-vessel thrombosis is the term used when the blockage is in one of the brain's larger blood-supplying arteries such as the carotid or middle cerebral, while small-vessel thrombosis involves one (or more) of the brain's smaller, yet deeper penetrating arteries. This latter type of stroke is also called a lacuner stroke.

[0014] An embolic stroke is also caused by a clot within an artery, but in this case the clot (or emboli) was formed somewhere other than in the brain itself. Often from the heart, these emboli will travel the bloodstream until they become lodged and cannot travel any further. This naturally restricts the flow of blood to the brain and results in almost immediate physical and neurological deficits.

[0015] Thrombolytic therapy has been proven to be effective for the treatment of acute ischemic stroke, but the increased risk of tissue plasminogen activator (tPA) is still of great clinical concern (see for instance The National Institutes of Neurological Disorders, and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. *New England Journal of Medicine* 1995;333:1581-7).

[0016] As it is critical to restore proper blood flow to the brain as soon as possible to prevent tissue damage, rapid diagnosis of stroke is critical to the survival of the patient and the minimization of any effects of the stroke to the patient. If caught from three to six hours after occurrence most stroke patients can expect full or partial recovery

[0017] Current state of the art diagnosis of stroke involves a physical examination and imaging procedures such as computed tomography (CT) scan, angiogram, electrocardiogram, magnetic resonance imaging (MRI), Single photon emission computed tomography (SPECT) and positron emission tomography (PET).

[0018] While physical examination is rapid, it only can detect large strokes (defined to be significant impairment of symptoms on the National Institutes of Health Stroke Scale, (NIHSS) of greater than 12). In addition, prior studies have found that the accuracy of stroke identification by medical personnel is modest and variable from one community to another. Sensitivity for stroke recognition by prehospital personnel has ranged widely, and positive predictive values have remained between 64% and 77% (see for instance Zweifler RM, York D, U TT, Mendizabal JE, Rothrock JF. Accuracy of paramedic diagnosis of stroke. *J Stroke Cerebrovasc Dis*. 1998;7:446–448.). These studies have consistently suggested a tendency for prehospital personnel to overdiagnose stroke by not recognizing stroke mimics, such as patients with alcohol and drug intoxication, postictal hemiparesis, hypoglycemia or other metabolic encephalopathies, and other nonstroke causes of acute neurological deficits. Finally, any clinical neurological screening test will be limited by the training and experience of the examiner. This suggests the need for an adjunctive clinical test that can provide diagnostic information above and beyond screening clinical exams.

[0019] CT scan produces x-ray images of the brain and is used to determine the location and extent of hemorrhagic stroke. It has widespread availability. CT scan usually cannot produce images showing signs of ischemic stroke until 48 hours after onset. This insensitivity to acute stroke limits its use to post-stroke damage assessment.

[0020] SPECT and PET involve injecting a radioactive substance into the bloodstream and monitoring it as it travels through blood vessels in the brain. These tests allow physicians to

detect damaged regions of the brain resulting from reduced blood flow. However, this takes several hours, and thus is not used for rapid diagnosis of stroke.

[0021] MRI with magnetic resonance angiography (MRA) uses a magnetic field to produce detailed images of brain tissue and arteries in the neck and brain, allowing physicians to detect small-vessel infarct (i.e., stroke in small blood vessels deep in brain tissue). However, as a practical issue, most hospitals do not have these specialized and highly expensive MRI services available in the acute setting. Thus, without a practical and widely available radiological test, the diagnosis of stroke remains largely a clinical decision.

[0022] Thrombolytic therapy has been proven to be effective for the treatment of acute ischemic stroke, but the increased risk of hemorrhagic transformation (HT) associated with tissue plasminogen activator (tPA) administration is still of great clinical concern. HT after cerebral ischemia seems to be related to the disruption of the vascular endothelium (see for instance Hamann GF, Okada Y, del Zoppo GJ. Hemorrhagic transformation and microvascular integrity during focal cerebral ischemia/reperfusion. *J Cereb Blood Flow Metab.* 1996;16:1373–1378.). In patients who receive tPA treatment, endothelial injury may be the result of free radical generation secondary to thrombolytic-induced reperfusion (see for instance), as well as of the upregulation of matrix metalloproteinases (MMPs) (see for instance Lapchak PA, Chapman DF, Zivin JA. Pharmacological effects of the spin trap agents *N*-t-butyl-phenylnitrone (PBN) and 2,2,6,6-tetramethylpiperidine-*N*-oxyl (TEMPO) in a rabbit thromboembolic stroke model: combination studies with the thrombolytic tissue plasminogen activator. *Stroke.* 2001;32:147–152.), a group of enzymes that are able to degrade the basal membrane components. The association between high levels of MMP-9 and the risk of HT in patients with acute ischemic stroke who have and have not received tPA have been previously reported (see for instance Sumii T, Lo EH. Involvement of matrix metalloproteinase in thrombolysis-associated hemorrhagic transformation after embolic focal ischemia in rats. *Stroke.* 2002;33:831–836.; Montaner J, Molina CA, Monasterio J, Abilleira S, Arenillas JF, Ribo´ M, Quintana M, Alvarez-Sabin J. Matrix metalloproteinase-9 pretreatment level predicts intracranial hemorrhagic complications after thrombolysis in human stroke. *Circulation.* 2003;107:598–603.). However, despite the available data, the underlying molecular mechanisms related to HT after thrombolytic treatment have yet to be fully elucidated.

[0023] Accordingly, there is a present need in the art for a rapid, sensitive and specific differential diagnostic assay for stroke, stroke subtype, and stroke mimic that can also identify those individuals at risk for hemorrhagic transformation after presentation with stroke symptoms and subsequent administration of tPA therapy. Such a diagnostic assay would greatly increase the number of patients that can receive beneficial stroke treatment and therapy and in so doing reduce the costs associated with incorrect stroke diagnosis. Some content of this patent application was first published in the journal Stroke in its May 27, 2004, electronic issue, but the exact thresholds of markers crucial to the teaching of the invention are disclosed in the instant invention and prior related applications.

BRIEF SUMMARY OF THE INVENTION

[0024] The present invention relates to the identification and use of diagnostic markers for stroke, endothelial damage and hemorrhagic transformation (HT) after thrombolytic therapy. The methods and compositions described herein can meet a need in the healing arts for rapid, sensitive and specific diagnostic assay to be used in the diagnosis and differentiation of various cardiac events. Moreover, the methods and compositions of the present invention can also be used to facilitate the treatment of stroke patients and the development of additional diagnostic and/or prognostic indicators.

[0025] In various aspects, the present invention relates to (1) materials and procedures for identifying markers that are associated with the diagnosis, prognosis, or differentiation of stroke and/or determination of HT in a patient; (2) using such markers in diagnosing and treating a patient and/or monitoring the course of a treatment regimen; (3) using such markers to identify subjects at risk for one or more adverse outcomes related to stroke and/or determination of HT; and (4) using at one of such markers an outcome marker for screening compounds and pharmaceutical compositions that might provide a benefit in treating or preventing such conditions.

[0026] In one of its aspects, the invention discloses methods for determining a diagnosis or prognosis related to a cardiac event such as stroke, or for differentiating between stroke sub-type and/or determination of HT. The preferred method includes analyzing a fluid sample

obtained from a person who has an unknown diagnosis for the levels of one or more markers specific to the damage caused by said cardiac event. In the case of stroke, these markers would be drawn from the group consisting of markers relating to vascular damage, glial activation, inflammatory mediation, thrombosis, cellular injury, apoptosis, myelin breakdown, and specific and non-specific markers of cerebral injury. The analysis of the preferred method thus more precisely includes identifying one or more markers the presence or amount of which is associated with the diagnosis, prognosis, or differentiation of stroke and/or determination of HT. Once such marker(s) are identified, the next in the method the level of such marker(s) in a sample obtained from a subject of interest can be measured. In certain embodiments of the preferred method, these markers can be compared to a level that is associated with the diagnosis, prognosis, or differentiation of stroke and/or determination of HT. By correlating the subject's marker level(s) to the diagnostic marker level(s), the presence or absence of stroke, and also the probability of future adverse outcomes, etc., in a patient may be rapidly and accurately determined.

[0027] In another of its aspects, the instant invention is embodied in methods for choosing one or more marker(s) for differentiation of stroke and/or determination of HT that together, and as a group, have maximal sensitivity, specificity, and predictive power. Said maximal sensitivity, specificity, and predictive power is in particular realized by choosing one or more markers as constitute a group by process of plotting receiver operator characteristic (ROC) curves for (1) the sensitivity of a particular combination of markers versus (2) specificity for said combination at various cutoff threshold levels. In addition, the instant invention further discloses methods to interpolate the nonlinear correlative effects of one or more markers chosen by any methodology to such that the interaction between markers of said combination of one or more markers promotes maximal sensitivity, specificity, and predictive accuracy in the prediction of any of the occurrence of stroke, identification of stroke subtype, or likelihood of HT.

[0028] For purposes of the following discussion, the methods described as applicable to the diagnosis and prognosis of stroke generally may be considered applicable to the diagnosis and prognosis of other cardiac events.

[0029] The term "marker" as used herein refers to proteins or polypeptides to be used as targets

for screening test samples obtained from subjects. "Proteins or polypeptides" used as markers in the present invention are contemplated to include any fragments thereof, in particular, immunologically detectable fragments. One of skill in the art would recognize that proteins which are released by cells of the central nervous system which become damaged during a cerebral attack could become degraded or cleaved into such fragments. Additionally, certain markers are synthesized in an inactive form, which may be subsequently activated, e.g., by proteolysis. Examples of such markers are described hereinafter. The term "related marker" as used herein refers to one or more fragments of a particular marker that may be detected as a surrogate for the marker itself. These related markers may be, for example, "pre," "pro," or "prepro" forms of markers, or the "pre," "pro," or "prepro" fragment removed to form the mature marker. Exemplary markers that are synthesized as pre, pro, and prepro forms are described hereinafter. In preferred embodiments, these "pre," "pro," or "prepro" forms or the removed "pre," "pro," or "prepro" fragments are used in an equivalent fashion to the mature markers in the methods described herein.

[0030] Preferred markers of the invention can (1) differentiate between ischemic stroke, hemorrhagic stroke, and stroke mimics, and also (2) predict HT in subjects. Preferred markers are drawn from the group including c-Fn, MMP-9, myelin basic protein, IL-1, IL-1r□, IL-1□, IL-8, IL-10, NCAM, VCAM, ICAM, S100b, HSP60, BDNF, D-Dimer, TGF-□, NT-3, VEGF, CK-BB, caspase 3, MCP-1 Calbindin-D, thrombin-antithrombin III complex, tissue factor, GFAP, NSE-γγ, vWF, VEGF, FPA, and NR2A. Each of these terms are defined hereinafter or in referenced related applications and are known to those of ordinary skill in the art. Particularly preferred markers from this group are ones that have proven highly predictive of hemorrhagic transformation: namely, cellular fibronectin (c-Fn) and matrix metalloprotein-9 (MMP-9).

[0031] Those of ordinary skill in the art know that marker levels vary at certain time points; for example, the level of a marker may be at one level at three hours post-stroke event, and another level at nine hours post-stroke event. Thus when using multiple markers together which may or may not be correlated with each other it is necessary to provide interpretation through an algorithm that relates all markers together. This algorithm in current state of the art is a simple threshold level above which a marker is said to be indicative of an adverse event in the human body. A particular diagnosis and/or prognosis of said adverse event may depend upon the

comparison of each marker to this value; alternatively, if only a subset of markers are outside of a normal range, then this subset may be indicative of a said adverse event.

[0032] Thus, in certain embodiments of the methods of the present invention, a plurality of markers are combined using an algorithm to increase the predictive value of the analysis in comparison to that obtained from the markers taken individually or in smaller groups. Most preferably, one or more markers for vascular damage, glial activation, inflammatory mediation, thrombosis, cellular injury, apoptosis, myelin breakdown, and specific and non-specific markers of cerebral injury are combined in a single assay to enhance the predictive value of the described methods. This assay is usefully predictive of multiple outcomes, for instance: determining whether or not a stroke occurred, then determining the sub-type of stroke, then further predicting stroke prognosis. Moreover, different marker combinations in the assay may be used for different indications. Correspondingly, different algorithms interpret the marker levels as indicated on the same assay for different indications.

[0033] Preferred panels comprise markers for the following purposes: (1) diagnosis of stroke; (2) diagnosis of stroke mimics; (3) diagnosis of stroke and indication if an acute stroke has occurred; (4) diagnosis of stroke and indication if an non-acute stroke has occurred; (5) diagnosis of stroke, indication if an acute stroke has occurred, and indication if an non-acute stroke has occurred; (6) diagnosis of stroke and indication if an ischemic stroke has occurred; (7) diagnosis of stroke and indication if a hemorrhagic stroke has occurred; (8) diagnosis of stroke, indication if an ischemic stroke has occurred, and indication if a hemorrhagic stroke has occurred; (9) diagnosis of stroke and prognosis of a subsequent adverse outcome; (10) diagnosis of stroke and prognosis of a subsequent hemorrhagic transformation; and (11) diagnosis of stroke, indication if a hemorrhagic stroke has occurred, and further diagnosis of whether a subarachnoid hemorrhagic stroke has occurred.

[0034] In preferred embodiments, particular thresholds for one or more markers in a panel are not relied upon to determine if a profile of marker levels obtained from a subject are indicative of a particular diagnosis/prognosis. Rather, in accordance with the present invention, an evaluation of the entire profile is made by (1) first training an algorithm with marker information from samples from a test population and a disease population to which the clinical

outcome of interest has occurred to determine weighting factors for each marker, and (2) then evaluating that result on a previously unseen population. Certain persons skilled in bioinformatics will recognize this procedure to be tanatamount to the construction, and to the training, of a machine learning algorithm. The evaluation is determined by maximizing the numerical area under the ROC curve for the sensitivity of a particular panel of markers versus specificity for said panel at various individual marker levels. From this number, the skilled artisan can then predict a probability that a subject's current marker levels in said combination is indicative of the clinical marker of interest. For example, (1) the test population might consist solely of samples from a group of subjects who have had ischemic stroke and no other comorbid disease conditions, while (2) the disease population might consist solely of samples from a group of subjects who have had hemorrhagic stroke and no other comorbid disease conditions. A third, "normal" population might also be used to establish baseline levels of markers as well in a non-diseased population.

[0035] In preferred embodiments of the marker, and marker panel, selection methods of the present invention, the aforementioned weighting factors are multiplicative of marker levels in a nonlinear fashion. Each weighting factor is a function of other marker levels in the panel combination, and consists of terms that relate individual contributions, or independent and correlative, or dependent, terms. In the case of a marker having no interaction with other markers in regards to then clinical outcome of interest, then the specific value of the dependent terms would be zero.

[0036] The term "test sample" as used in this specification refers to a sample of bodily fluid obtained for the purpose of diagnosis, prognosis, or evaluation of a subject of interest, such as a patient. In certain embodiments, such a sample may be obtained for the purpose of determining the outcome of an ongoing condition or the effect of a treatment regimen on a condition. Preferred test samples include blood, serum, plasma, cerebrospinal fluid, urine and saliva. In addition, one of skill in the art would realize that some test samples would be more readily analyzed following a fractionation or purification procedure, for example, separation of whole blood into serum or plasma components.

[0037] The term "markers of glial activation" as used in this specification refers to markers that

indicate glial cell function. Glia mediate neuroendocrine and neuroimmune functions and are also important in synaptic remodeling and the loss of synaptic connections that occur during aging. These functions are carried out by changes in glia, including changes in shape, interactions with neurons and other glia, and gene expression. The predominant change that occurs in glia during aging is glial activation, which can progress to reactive gliosis in response to neurodegeneration. Markers distinguish normal and reactive glia. During aging, astrocytes hypertrophy and exhibit signs of metabolic activation, and astrocytic processes surround neurons. Microglia also become activated and subsets of activated microglial increase in number and may enter the phagocytic or reactive stage. Yet glial cells are intimately involved in the biochemical metabolic and neurotrophic support of the function of neurons, and glial actions at the synapses are crucial to normal neuronal transmission. Glia take up excess glutamate (which can be neurotoxic) and produce neurotrophic factors which keep cells alive, as well as interacting with other systems in transmitter-like actions. Thus, a loss of normal glial function could have dramatic impacts on normal neuronal function. Such specific markers of glial activation include, but are not limited to, GFAP, S100B, Mac-1, TLR4, TGF-β1 and CD14.

[0038] The term "markers of vascular damage" as used in this specification refers to markers that indicate endothelial damage. When the endothelium is damaged or becomes dysfunctional, a cascade leading to atherogenesis is precipitated, initiating a cycle of injury, immunologic induction, and amplification. Dysfunctional endothelium leads to increased permeability to lipoproteins and up-regulation of leukocyte and endothelial adhesion molecules. In response to the presence of certain activating substances, including oxidized LDL, monocyte chemotactic protein 1, interleukin (IL)-8, and platelet-derived growth factor (PDGF), leukocytes migrate into the wall of the artery. Such specific markers of vascular damage include, but are not limited to, endothelin-1 (ET-1), von Willebrand factor (vWf), and soluble (S-) adhesion molecules E-selectin, intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule-1 (VCAM-1), plasma indexes of endothelial damage/dysfunction and soluble thrombomodulin (sTM).

[0038] The term "markers of inflammatory mediation" as used in this specification refers to markers that indicate an inflammatory response to a cerebral injury. Inflammatory responses are initiated and perpetuated by the interaction of immune cells with cells of the affected vessel

wall. This is directed by a network of chemical messengers, which, in a state of vascular health, exist as balanced but opposing forces. These markers include various cytokines, proteases, adhesion molecules, and acute phase proteins as participants in the generation of vascular inflammation. Such specific markers of vascular damage include, but are not limited to, Cellular adhesion molecules such as Intracellular adhesion molecule-1, Vascular cellular adhesion molecule-1, NCAM and Selectins such as E-Selectin; Chemokines such as monocyte chemoattractant protein-1; Cytokines such as Interleukins 1, 1β, 1 receptor antagonist, 6, 8, 10, 18, transforming growth factor β, and Tumor necrosis factor-α; Proteases such as the matrix metalloproteinases MMP-9, MMP-3, and MMP-2; Accessory signaling markers such as CD40/CD40L; and acute phase proteins such as C-reactive protein, vascular endothelial growth factor, ceruloplasmin, fibrinogen, α 1-acid glycoprotein, α 1-antitrypsin, and haptoglobin.

[0039] The term "markers of thrombosis" as used in this specification refers to markers that indicate an coagulation event in ischaemic stroke. The blood clotting system is activated when blood vessels are damaged, exposing collagen, the major protein that connective tissue is made from. Platelets circulating in the blood adhere to exposed collagen on the cell wall of the blood vessel and secrete chemicals that start the clotting process as follows: Platelet aggregators cause platelets to clump together (aggregate). They also cause the blood vessels to contract (vasoconstrict), which reduces blood loss. Platelet aggregators include adenosine diphosphate (ADP), thromboxane A2, and serotonin (5-HT). Coagulants such as fibrin then bind the platelets together to form a permanent plug (clot) that seals the leak.

[0040] Fibrin is formed from fibrinogen in a complex series of reactions called the coagulation cascade. The enzymes that comprise the coagulation system are called coagulation factors, which are numbered in the order in which they were discovered. They include factor XII, factor XI, factor IX, factor X, factor VII, and factor V. The activation of the coagulation factors results in the formation of thrombin, which acts as a cofactor for the conversion of fibrinogen into fibrin. After the leak has been sealed with a blood clot, the body responds with another set of chemical messengers that oppose the actions of these chemicals. These include: Platelet aggregation inhibitors and vasodilators, such as nitric oxide and prostacyclin, which is also known as prostaglandin I2 (PGI2) Plasminogen activators that promote the breakdown of fibrin, such as tissue plasminogen activator (t-PA) Anticoagulants that inhibit enzymes in the

coagulation cascade, such as antithrombin III (activated by heparin) and proteins C and S.

[0040] Such specific markers of thrombosis include, but are not limited to, von Willebrand factor, thrombin-antithrombin III complex, proteins C and S, tissue factor, fibrinopeptide A, plasmin-α-2-antiplasmin complex, prothrombin fragment 1+2, D-dimer, platelet factor 4, and β-thromboglobulin.

[0041] The term "marker of cellular injury and myelin breakdown" as used in this specification refers to markers associated with damage to the structural and functional molecules of the cell. Although any biologically important molecule in a cell can be the target of injury producing stress, four biochemical systems are particularly vulnerable: (1) the cell membrane, (2) energy metabolism, (3) protein synthesis, and (4) genes. Because many of the biochemical systems of the cell are inter-dependent, injury at one site typically leads to secondary injury to other cellular processes.

[0042] Myelin is the outer lipid rich (fatty) layer that covers nerves and nervous system pathways in the brain and spinal cord. The myelin sheath, a lipid-rich multilamellar membrane of relative stability, both insulates and enhances conduction in nerve axons. A notable feature of myelin-specific proteins, in particular myelin basic protein, is their susceptibility to proteolytic activity and their encephalitogenicity, which induces inflammatory demyelination in the CNS. The final common pathway of myelin breakdown in vivo is well documented and there is evidence that myelin disruption can be mediated directly by soluble (circulating) factors and for following receptor-driven phagocytosis by macrophages. However the exact mechanism(s) of demyelination in ischemic attack is still unresolved, both antigen-specific and--non-specific events having the potential to generate the myelinolytic process.

[0043] Cerebral injury leads to breakdown of the blood–brain barrier (BBB), exposing CNS antigens to the peripheral circulation and allowing the peripheral circulation access to the brain. The breakdown of the BBB leads to rapid acquisition of MBP-reactive T cell clones and Igs in stroke patients. but does not lead to autoimmune encephalitis. The degradation of myelin basic protein (MBP) by proteinase yields encephalitogenic peptides and its loss has been found to cause structural alteration of the myelin sheath. This suggests that MBP degradation is an initial

step in the breakdown of myelin in demyelinating diseases. A calcium-activated neutral proteinase (calpain), which degrades MBP, was found to increase in activity in MS tissue and cerebrospinal fluid (CSF), and its presence in myelin suggests that myelin may be autodigested in demyelinating disease. The source of increased proteinase activity has been indicated as macrophages, lymphocytes, and proliferative astrocytes (reactive cells). Increased proteinase activity is found in Schwann cells in Wallerian degeneration, and the presence of calpain in myelin-forming oligodendrocytes and Schwann cells suggests that these cells are likely sources of degradative enzymes.

[0044] Such specific markers of cellular injury and myelin breakdown include, but are not limited to, creatinine phosphokinase brain band, tissue factor, Proteolipid protein, RU Malendialdehyde, calpain, and myelin basic protein.

[0045] The term "marker of apoptosis or growth factors" as used in this specification refers to markers involved in neuronal cell death. Numerous studies in experimental models of ischemia have now reported that apoptosis contributes to neuronal death (reviewed by Chalmers-Redman et al Mechanisms of nerve cell death: apoptosis or necrosis after cerebral ischemia. In: Green AR, Cross AJ, eds. *Neuroprotective Agents and Cerebral Ischemia* . San Diego, Calif: Academic Press; 1997:1–25.). Apoptosis requires the activation of a "cell death" gene program, and many of the extracellular signals that regulate apoptosis have been identified. For example, interaction between the Fas/APO-1 molecule, a cell surface protein, with its ligand (Fas-L) leads to programmed cell death. Soluble (s) Fas/APO-1, a molecule lacking the transmembrane domain of Fas/APO-1, blocks apoptosis by inhibiting interaction between Fas/APO-1 and Fas-L on the cell surface (see for instance Cheng J et al., Protection from Fas-mediated apoptosis by a soluble form of the Fas molecule. *Science.* 1994;263:1759–1762.). Fas expression has been detected on B and T cells and on neutrophils. It has been suggested that the Fas/Fas-L pathway is one of the major mechanisms for T-cell–mediated cytotoxicity. It has also been demonstrated by in situ hybridization that the expression of Fas/APO-1 was induced in murine brain after transient global cerebral ischemia. Another gene product, Bcl-2, has been shown to suppress apoptosis and to protect primary neuronal cell cultures from apoptosis induced by nerve growth factor depletion.

[0046] Macrophages and T lymphocytes kill target cells by inducing apoptosis, one of the potential mechanisms whereby the inflammatory cells invading the infarcted brain area participate in neuronal cell death. Stroke patients displayed an intrathecal production of proinflammatory cytokines, such as interleukin (IL)-1ß, IL-6, IL-8, and granulocyte-macrophage colony-stimulating factor (GM-CSF), and of the anti-inflammatory cytokine IL-10 within the first 24 hours after the onset of symptoms, supporting the notion of localized immune response to the acute brain lesion in humans. Some of these cytokines (eg, IL-1ß and IL-8) stimulate influx of leukocytes to the infarcted brain, a prerequisite for Fas/APO-1– and bcl-2–mediated apoptosis. TNF- α, a powerful cytokine inducing apoptosis in the extraneural compartment of the body, has been demonstrated to protect rat hippocampal, septal, and cortical cells against metabolic-excitotoxic insults and to facilitate regeneration of injured axons. More importantly, TNF-α and -ß protect neurons against amyloid ß-protein–triggered toxicity.

[0047] Other evidence demonstrates that apoptosis involves the activation of caspases, a unique family of structurally related, highly conserved, aspartate-specific, cysteine proteases that are necessary to carry out the signal for apoptotic cell death. Two members of the caspase family, caspase-1 and caspase-3, are known to cleave the most abundant caspase target substrate, actin. The 45-kDa actin is cleaved by caspase activation between Asp11 and Asn12 and between Asp244 and Gly245 to produce N-terminal 32-kDa fragments and C-terminal 15-kDa fragments. A polyclonal antibody to the last 5 amino acids of the C-terminus of the 32-kDa fragment of actin generated by caspase cleavage of intact actin has been developed and named "fractin" for "fragment of actin." Fractin labeling provides indirect evidence of caspase activation and demonstrates initiation of an apoptotic pathway, but does not rule out secondary necrosis. Other markers for apoptosis include biochemical evidence of oligointernucleosomal DNA fragmentation into approximately 180-bp multiples resulting from endonuclease activation that can be demonstrated with a typical "laddering" appearance on agarose gel electrophoresis. In addition, the terminal deoxynucleotidyl transferase–mediated dUTP-biotin nick end labeling (TUNEL) technique, which identifies 3'-OH ends of DNA-strand breaks, has been widely used as a marker of DNA damage or repair. However, the lack of specificity of TUNEL in detecting oligointernucleosomal DNA fragmentation precludes its use as a defining feature of apoptosis.

[0048] Such specific markers of apoptosis and growth factors include, but are not limited to, Brain natriuretic peptide, caspase 3, calbindin-D, heat shock protein 60 and 70, c-fos, c-jun, ubiquitin, and cytochrome C.

[0049] The term "specific marker of cerebral injury" as used in this specification refers to proteins or polypeptides that are associated with brain tissue and neural cells, and which can be correlated with a cerebral injury, but are not correlated with other types of injury. Such specific markers of cerebral injury include, but are not limited to, adenylate kinase, brain-derived neurotrophic factor, calbindin-D, lactate dehydrogenase, myelin basic protein, neural cell adhesion molecule, neuron-specific enolase, neurokinin A, neurokinin B, neurotensin, neurotrophin-3, neurotrophin-4/5, neuropeptide Y, proteolipid protein, substance P, thrombomodulin, and protein kinase C gamma.

[0050] The term "non-specific marker of cerebral injury" as used in this specification refers to proteins or polypeptides that are elevated in the event of cerebral injury, but may also be elevated due to non-cerebral events. Non-specific markers include, but are not limited to, ApoC-I and ApoC-II, A-type natriuretic peptide, B-type natriuretic peptide, C-type natriuretic peptide,adrenomedullin, β - thromboglobulin, C-reactive protein, Cardiac Troponin I and Troponin T, Creatine kinase MB, D-dimer, E-selectin, endothelin-1, endothelin-2, and endothelin-3, A- , F- , and H- Fatty acid binding protein, fibrinopeptide A, hemoglobin α_2, chain head activator, insulin-like growth factor-1, MMP-3, plasmin-α-2-antiplasmin complex, platelet factor 4, 8- epi PGF sub(2α), PGI2, PGE2, prothrombin fragment 1+2, thrombin-antithrombin III complex, tissue factor, transforming growth factor β and von Willebrand factor.

[0051] The term “diagnosis”, as used in this specification refers to predict the type of disease or condition from a set of marker values and/or patient symptoms. This is in contrast to disease prediction, which is to predict the occurrence of disease before it occurs, and the term “prognosis”, which is to predict disease progression at a future point in time from one or more indicator value(s) at a previous point in time.

[0052] The term "correlating," as used in this specification refers to a process in which a set of examples of clinical inputs from subjects, such as marker levels, and their corresponding outputs,

such as whether a subject suffered from a specific type of stroke, are related to each other. This relationship can be determined by comparing such examples to examples from a control and/or disease-free population at a later point in time, and selecting those indicators which can differentiate between the two disease states as a function of time alone or in combination at a certain probability level. The selection process is described herein. The selected markers, each at a certain level range which might be a simple threshold, are said to be correlative or associative with one of the disease states. Said correlated markers can be then be used for disease detection, diagnosis, prognosis and/or treatment outcome. Preferred methods of correlating markers is by performing marker selection by a feature selection algorithm and classification by mapping functions described herein. A preferred probability level is a 3% chance, 5% chance, a 7% chance, a 10% chance, a 15% chance, a 20% chance, a 25% chance, a 30% chance, a 35% chance, a 40% chance, a 45% chance, a 50% chance, a 55% chance, a 60% chance, a 65% chance, a 70% chance, a 75% chance, a 80% chance, a 85% chance, a 90% chance, a 95% chance, and a 100% chance. Each of these values of probability is plus or minus 2% or less. A preferred threshold level for markers of the present invention is about 25 pg/mL, about 50 pg/mL, about 60 pg/mL, about 75 pg/mL, about 100 pg/mL, about 150 pg/mL, about 200 pg/mL, about 300 pg/mL, about 400 pg/mL, about 500 pg/mL, about 600 pg/mL, about 750 pg/mL, about 1000 pg/mL, and about 2500 pg/mL. The term "about" in this context refers to +/- 10%.

[0053] In yet another of its aspects, the present invention is embodied in methods for determining a treatment regimen for use in a patient diagnosed with stroke. The methods preferably comprise determining a level of one or more diagnostic or prognostic markers as described herein, and using the markers to determine a diagnosis for a patient. For example, a prognosis might include the development or predisposition to delayed neurologic deficits after stroke onset. One or more treatment regimens that improve the patient's prognosis by reducing the increased disposition for an adverse outcome associated with the diagnosis can then be used to treat the patient. Such methods may also be used to screen pharmacological compounds for agents capable of improving the patient's prognosis as above.

[0054] In yet another of its aspect, the present invention relates to methods of identifying a patient at risk for hemorrhagic transformation after thrombolytic therapy. Such methods

preferably comprise comparing an amount of a marker predictive of a subsequent hemorrhagic transformation, said marker selected from the group consisting of cellular fibronectin (c-Fn), and matrix metalloprotease-9 (MMP-9), in a test sample from a patient diagnosed with an acute ischemic stroke to a predictive level of said marker, wherein said patient is identified as being at risk for hemorrhagic transformation by a level of said marker equal to or greater than said predictive level.

[0055] In yet another of its aspects, the present invention is embodied in methods of differentiating ischemic stroke from hemorrhagic stroke using such marker combination panels.

[0056] In yet another of its aspects, the present invention is embodied in kits for determining the diagnosis or prognosis of a patient. These kits preferably comprise devices, software and reagents for measuring one or more marker levels in a patient sample, and instructions for performing the assay. Additionally, the kits contain a computer software program to be run on a computer or other means for converting marker level(s) to a prognosis. Such kits preferably contain sufficient reagents to perform one or more such determinations, and are standardized to run on an instrument used to analyze blood samples, such as Abbott Laboratories' AxSYM®, Roche Diagnostics' Cardiac Reader®, or Dade Behring's Stratus® CS Analyzer.

BRIEF DESCRIPTION OF THE DRAWINGS

[0057] In the following, the invention will be explained in further detail with reference to the drawings, in which:

[0058] FIG. 1 is a table detailing demographics of the study population;

[0059] FIG.s 2A and 2B are graphs showing baseline levels of c-Fn at two time periods for patients who experienced HT and those who did, subclassified by severity;

[0060] FIG.s 3A and 3B are graphs showing baseline levels of MMP-9 at two time periods for patients who experienced HT and those who did, subclassified by severity;

[0061] FIG. 4 is a chart showing the univariate statistically significance for each of the markers studied with respect to HT outcome;

[0062] FIG. 5 is a Receiver operator curve for discrimination of patients whom will undergo a relevant hemorrhagic transformation following administration of TPA based on c-Fn;

[0063] FIG. 6 is a Receiver operator curve for discrimination of patients whom will undergo a relevant hemorrhagic transformation following administration of TPA based on MMP-9;

[0064] FIG. 7 is a graph of the joint distribution of c-Fn and MMP-9; and

[0065] FIG. 8 is a graph of relationship between MMP-9 and c-Fn for subjects with an observed HT.

DETAILED DESCRIPTION OF THE INVENTION

[0066] In accordance with the present invention, there are provided methods and apparatus for the identification and use of a panel of markers for the diagnosis of hemorrhage, particularly hemorrhage following tPA therapy, and differentation of such patients from patients who have stroke-like symptoms, e.g. stroke mimics.

[0067] Fibronectins are adhesive dimeric glycoproteins that promote cell–cell and cell–matrix interactions (see for instance Hynes RO. Fibronectins. *Sci Am.* 1986;254:42–51.). Plasma fibronectin (p-Fn) is primarily produced by hepatocytes, but plasma also contains small quantities of cellular fibronectin (c-Fn), which is mainly synthesized by endothelial cells (see for instance Peters JH, Sporn LA, Ginsberg MH, Wagner DD. Human endothelial cells synthesize, process, and secrete fibronectin molecules bearing an alternatively spliced type II homology (ED1). *Blood.* 1990;75:1801–1808.). Because c-Fn is largely confined to the vascular endothelium, high plasma levels of this molecule might be indicative of endothelial damage. In fact, plasma c-Fn levels have been reported to be increased in patients with vascular injury secondary to vasculitis, sepsis, acute major trauma, and diabetes, (see for instance Peters JH, Maunder RJ, Woolf AD, Cochrane GH, Ginsberg MH. Elevated plasma levels of ED1 ("cellular") fibronectin in patients with vascular injury. *J Lab Clin Med.* 1989;113:586–597; Kanters SD, Banga JD, Algra A, Frijns RC, Beutler JJ, Fijnheer R. Plasma levels of cellular fibronectin in diabetes. *Diabetes Care.* 2000;24:323–327.). Since HT after cerebral ischemia seems to be the result of the continuous disappearance of basal membrane components (see for instance Hamann GF, Okada Y, del Zoppo GJ. Hemorrhagic transformation and microvascular integrity during focal cerebral ischemia/reperfusion. *J Cereb Blood Flow Metab.* 1996;16:1373–1378.), in the instant invention we show high levels of plasma c-Fn are associated with HT in patients who received thrombolytic treatment with tPA. We also show elevation in patients with acute ischemic stroke for the first time.

[0068] Because thrombolytic therapy is the only treatment for ischemic stroke proven to be effective, the investigation of the underlying mechanisms responsible for HT, the most feared complication associated with this therapy, as well as the identification of factors that can improve the benefit/risk ratio of tPA administration is of critical importance. The instant

invention demonstrates that plasma c-Fn levels are significantly higher in patients in whom HT develops after tPA administration and teaches that c-Fn levels >3.6 μg/mL can predict the development of HI-2 and PH after tPA administration with a sensitivity and negative predictive value of 100%. Therefore, c-Fn is a useful marker of those patients who are at greatest risk for HT after the administration of thrombolytic treatment.

[0069] The loss of microvascular integrity secondary to the continuous disappearance of the antigens of the endothelial components has been reported as being responsible for HT after ischemic injury.3 Among these antigens, c-Fn is especially important because it mediates the interaction between the endothelium and blood cells as well as other blood components.8 Moreover, Fn plays an important role in blood clot formation by mediating the adhesion of platelets to fibrin (see for instance Hynes RO. Fibronectins. *Sci Am.* 1986;254:42–51.), so the disappearance of the c-Fn of the vascular endothelium secondary to ischemia might damage this clotting mechanism, facilitating HT development. Although high c-Fn levels have been previously reported in patients with ischemic stroke, no previous data are available on the association between c-Fn levels in patients with acute ischemic stroke.

[0070] The increase of vascular permeability and subsequent extravasation of serum components leading to HT after tPA administration may be the result of several mechanisms including the activation of MMPs, which is secondary to ischemia, and the administration of tPA (see for instance Sumii T, Lo EH. Involvement of matrix metalloproteinase in thrombolysis-associated hemorrhagic transformation after embolic focal ischemia in rats. *Stroke*. 2002;33:831–836.). The instant invention also details the significant association between MMP-9 levels and HT in patients who received tPA and in a nonselected series of ischemic stroke patients. However, the fact that c-Fn is almost exclusively located at the endothelium suggests that this molecule could be a more specific marker of a high risk for HT. This hypothesis is supported by our finding that c-Fn levels, but not MMP-9 levels, remained independently associated with HT in the logistic regression analysis. Moreover, the predictive capacity of plasma c-Fn levels for the development of HI-2 and PH was higher than the predictive capacity of MMP-9 levels. However, although the difference did not reach statistical significance, probably because of the small sample size, there was a clear trend for the levels of c-Fn to be higher in patients with symptomatic HT, whereas MMP levels were similar in

symptomatic and asymptomatic bleedings. Because neurological deterioration usually occurs in patients with more severe HT, c-Fn levels probably reflect not only endothelial damage but also the degree of endothelial damage. In agreement with this hypothesis, we have observed a positive correlation between c-Fn levels and hypodensity volume at 24 to 36 hours of evolution of the ischemia, which probably reflects the relationship between endothelial and brain injuries. This positive correlation could lead us to argue that c-Fn levels are just an epiphenomenon of the extent of brain damage. However, we find that it is the plasma c-Fn concentrations rather that infarct volume that independently predicted HT.

[0071] The basal lumina disruption and the subsequent release of c-Fn after brain ischemic injury into the plasma, as well as accelerated Fn synthesis by endothelial cells and other cells such as polymorphonuclear leukocytes arriving at the ischemic tissue as part of the ischemic inflammatory cascade, could be among the participating mechanisms. Interleukins and transforming growth factor, whose expression is increased as a result of ischemia (see for instance Feuerstein GZ, Wang X, Barone FC. Inflammatory mediators and brain injury: the role of cytokines and chemokines in stroke and CNS diseases. In: Ginsberg MD, Bogousslavsky J, eds. *Cerebrovascular Disease:Pathophysiology, Diagnosis, and Management.* Boston, Mass: Blackwell Science; 1998:507–531.), have been shown to stimulate Fn synthesis (see for instance Roberts CJ, Birkenmeier TM, McQuillar JJ, Akiyama SK, Yamada SS, Chen WT, Yamada KM, McDonald JA. Transforming growth factor beta stimulates the expression of fibronectin and of both subunits of the human fibronectin receptor by cultured human lung fibroblast. *J Biol Chem.*1988;263:4586–4592.). Increased c-Fn synthesis could be an attempt to decrease endothelial destruction by MMPs, which might explain the positive correlation between c-Fn and MMP-9 in the instant invention.

[0072] Recently, many researchers have investigated the possibility of blood-borne markers of stroke and its subtypes. This approach is well established in the clinical setting of suspected myocardial ischemia. In acute coronary syndromes, the myocardial isoform of creatinine phosphokinase and troponin play an important role both in treatment decisions and clinical research. Similarly, B-type natriuretic peptide has become a routine part of the assessment of patients with congestive heart failure and dyspnea. However, the ischemic cascade of glial activation and ischemic neuronal injury in stroke is far more complex than myocardial ischemia

and less amenable to the use of a single biochemical marker. Indeed, the authors of the instant invention know of no individual biochemical marker has been demonstrated to possess the requisite sensitivity and specificity to allow it to function independently as a clinically useful diagnostic marker.

[0073] Thus a panel of markers was envisioned to overcome this deficiency in 1998 or earlier for detecting stroke (see for instance Misz M, Olah L, Kappelmayer J, Blasko G, Udvardy M, Fekete I, Csepany T, Ajzner E, Csiba L. Hemostatic abnormalities in ischemic stroke, Orv Hetil. 1998 Oct 18;139(42):2503-7; Tarkowski E, Rosengren L, Blomstrand C, Jensen C, Ekholm S, Tarkowski A. Intrathecal expression of proteins regulating apoptosis in acute stroke. Stroke. 1999 Feb;30(2):321-7; Stevens H, Jakobs C, de Jager AE, Cunningham RT, Korf J. Neurone-specific enolase and N-acetyl-aspartate as potential peripheral markers of ischaemic stroke. Eur J Clin Invest. 1999 Jan;29(1):6-11.) or its sub-types (see for instance Soderberg S, Ahren B, Stegmayr B, Johnson O, Wiklund PG, Weinehall L, Hallmans G, Olsson T. Leptin is a risk marker for first-ever hemorrhagic stroke in a population-based cohort. Stroke. 1999 Feb;30(2):328-37).

[0074] In many studies since this time, many blood-borne proteomic markers have been shown to be associated with stroke and its sub-types. For example, acute stroke has been associated with serum elevations of numerous inflammatory and anti-inflammatory mediators such as interleukin 6 (IL-6) and matrix metalloproteinase-9 (MMP-9) (see for instance Kim JS, Yoon SS, Kim YH, Ryu JS. Serial measurement of interleukin-6, transforming growth factor-beta, and S-100 protein in patients with acute stroke. *Stroke*. 1996;27:1553–1557.; Dziedzic T, Bartus S, Klimkowicz A, Motyl M, Slowik A, Szczudlik A. Intracerebral hemorrhage triggers interleukin-6 and interleukin-10 release in blood. *Stroke*. 2002;33:2334–2335.; Beamer NB, Coull BM, Clark WM, Hazel JS, Silberger JR. Interleukin-6 and interleukin-1 receptor antagonist in acute stroke. *Ann Neurol*. 1995; 37:800–805.; Montaner J, Alvarez-Sabin J, Molina C, et al. Matrix metalloproteinase expression after human cardioembolic stroke: temporal profile and relation to neurological impairment. *Stroke*. 2001;32:1759–1766.; Perini F, Morra M, Alecci M, Galloni E, Marchi M, Toso V. Temporal profile of serum anti-inflammatory and pro-inflammatory interleukins in acute ischemic stroke patients. *Neurol Sci*. 2001;22:289–296.; Vila N, Castillo J, Davalos A, Chamorro A. Proinflammatory cytokines and early neurological worsening in

ischemic stroke. *Stroke*. 2000;31: 2325–2329), markers of impaired hemostasis and thrombosis (see for instance Fon EA, Mackey A, Cote R, et al. Hemostatic markers in acute transient ischemic attacks. *Stroke*. 1994;25:282–286.; Takano K, Yamaguchi T, Uchida K. Markers of a hypercoagulable state following acute ischemic stroke. *Stroke*. 1992;23:194–198.), and markers of glial activation such as S100b (see for instance Buttner T, Weyers S, Postert T, Sprengelmeyer R, Kuhn W. S-100 protein: serum marker of focal brain damage after ischemic territorial MCA infarction. *Stroke*. 1997;28:1961–1965.; Martens P, Raabe A, Johnsson P. Serum S-100 and neuron-specific enolase for prediction of regaining consciousness after global cerebral ischemia. *Stroke*. 1998;29:2363–2366.). Several of these mediators, including IL-6, have been shown to be elevated within hours after ischemia and correlate with infarct volume (see for instance Fassbender K, Rossol S, Kammer T, et al. Proinflammatory cytokines in serum of patients with acute cerebral ischemia: kinetics of secretion and relation to the extent of brain damage and outcome of disease. *J Neurol Sci*. 1994;122:135–139.; Tarkowski E, Rosengren L, Blomstrand C, et al. Early intrathecal production of interleukin-6 predicts the size of brain lesion in stroke. *Stroke*. 1995;26:1393–1398).

[0075] Other authors have looked at the differentiation between TIA and stroke (see for instance Dambinova SA, Khounteev GA, Skoromets AA. Multiple panel of biomarkers for TIA/stroke evaluation. *Stroke*. 2002;33:1181–1182.) or type of hemorrhage (see for instance McGirt MJ, Lynch JR, Blessing R, Warner DS, Friedman AH, Laskowitz DT. Serum von Willebrand factor, matrix metalloproteinase-9, and vascular endothelial growth factor levels predict the onset of cerebral vasospasm after aneurysmal subarachnoid hemorrhage. *Neurosurgery*. 2002;51:1128–1134).

[0076] To this date, most of these studies have been in small number of patients and while have individual markers in common, the panels proposed in each have not been replicated. This is due to the fact that many reported panels merely linearly add the effects of multiple markers, or perform simple logistic regression to get correlative effects of a panel. One such example of the current state of the art is that of Reynolds *et al.* (Mark A. Reynolds, Howard J. Kirchick, Jeffrey R. Dahlen, Joseph M. Anderberg, Paul H. McPherson, Kevin K. Nakamura, Daniel T. Laskowitz, Gunars E. Valkirs, and Kenneth F. Buechler, Early biomarkers of stroke, *Clinical Chemistry* 49:10 1733–1739, 2003). In this paper, a five marker panel consisting of S-100β, B-

type neurotrophic growth factor, von Willebrand factor, matrix metalloproteinase-9, and monocyte chemotactic protein-1 was disclosed as suggested blood-borne panel to diagnosis acute ischemic stroke. In this analysis, univariate analysis was used to select an initial pool of candidate markers, and then multivariate analysis was used to achieve the final panel. However, as shown in the instant invention, this methodology is flawed. The result of this paper was tested on data used to train such, a typical mistake which usually leads to an irreproducible result.

[0077] Another example of the state of the art is U.S. Patent application 20040121343 and/or U.S. patent Ser. No. 10/225,082. In these application, a variety of markers for the diagnosis of stroke are envisioned, the mere presence or absence of such markers in the blood being indicative of disease. This methodology is fatally flawed, however, since it does not indicate how to relate the collective nonlinear effects of all markers to the outcome of interest, i.e. specify an algorithm to select among such markers and another to classify such markers as related to outcome. Instead, the application anticipates using the thresholded values of such markers as an indicator, giving a simple binary response of each as a value. As such markers are all treated as independent variables, there is no interaction between them, another fatal flaw.

[0078] Most existing statistical and computational methods for biomarker feature selection, such as U.S. Patent applications 20040126767 and/or U.S. patent application 20040219509, have focused on differential expression of markers between diseased and control data sets. This metric is tested by simple calculation of fold changes, by t-test, and/or F test. These are based on variations of linear discriminant analysis (i.e., calculating some or the entire covariance matrix between features).

[0079] However, the majority of these data analysis methods are not effective for biomarker identification and disease diagnosis for the following reasons. First, although the calculation of fold changes or t-test and F-test can identify highly differentially expressed biomarkers, the classification accuracy of identified biomarkers by these methods, is, in general, not very high. This is because linear transforms typically extract information from only the second-order correlations in the data (the covariance matrix) and ignore higher-order correlations in the data. We have shown that proteomic datasets are inherently non-symmetric (unpublished data). For

such cases, nonlinear transforms are necessary. Second, most scoring methods do not use classification accuracy to measure a biomarker's ability to discriminate between classes. Therefore, biomarkers that are ranked according to these scores may not achieve the highest classification accuracy among biomarkers in the experiments. Even if some scoring methods, which are based on classification methods, are able to identify biomarkers with high classification accuracy among all biomarkers in the experiments, the classification accuracy of a single marker cannot achieve the required accuracy in clinical diagnosis. Third, a simple combination of highly ranked markers according to their scores or discrimination ability is usually not be efficient for classification, as shown in the instant invention. If there is high mutual correlation between markers, then complexity increases without much gain.

[0080] Accordingly, the instant invention provides a methodology that can be used for biomarker feature selection and classification, and is applied in the instant application to detection of stroke and its subtypes.

[0081] Exemplary Biomarkers related to detection of hemorrhage.

[0082] A comprehensive methodology for identification of one or more markers for the prognosis, diagnosis, and detection of disease has been described previously. Suitable methods for identifying such diagnostic, prognostic, or disease-detecting markers are described in detail in U.S. patent application Ser. No. 11/046,592, entitled 'Cellular fibronectin as a diagnostic marker in stroke and methods of use thereof', filed January 29, 2005, which is hereby incorporated by reference in its entirety, including all tables, figures, and claims. Briefly, our method of predicting relevant markers given an individual's test sample is an automated technique of constructing an optimal algorithmic mapping between a given set of input marker data and a given clinical variable of interest. We illustrate this method further in said U.S. patent application in the section entitled "Methodology of Marker Selection, Analysis, and Classification"

[0083] We first obtain patient test samples of some bodily fluid, such as blood, cerebrospinal fluid, or urine from two or more groups of patients. Preferred fluid is blood. The patients are those exhibiting symptoms of a disease event, say stroke, which is determined at a later time, and those not exhibiting the same disease event, which are viewed as controls, though these patients might

have another disease event distinct from the first. Samples from these patients are taken at various time periods after the event has occurred, and assayed for various markers as described within. Clinical information, such as sex, age, time from onset of symptoms to treatment, NIHSS score, biochemistry and vital signs at admission, and neuroimaging findings are collected at various time periods. Preferred time periods for the instant invention include 0, 3 hours, 6 hours, 9 hours, 12 hours, 15 hours, 18 hours, 24 hours, 36 hours, 48 hours, 72 hours, 1 week, 2 weeks, 3 weeks, 4 weeks, 3 months and 6 months. Time is measured either from onset of symptoms or admission into a clinical setting where the patient receives care. This marker and clinical information form a set of examples of clinical inputs and their corresponding outputs, the outputs being the clinical outcome of interest, for instance stroke and stroke subtype occurrence or non-occurrence, hemorrhagic transformation, or stroke mimic subtype. These quantities are as described in the Introduction.

[0084] We then use an algorithm to select the most relevant clinical inputs that correspond to the outcome for each time period. This process is also known as feature selection. In this process, the minimum number of relevant clinical inputs that are needed to fully differentiate and/or predict disease prognosis, diagnosis, or detection with the highest sensitivity and specificity are selected for each time period. The feature selection is done with an algorithm that selects markers that differentiate between patient disease groups, say hemorrhagic versus ischemic. The relevant clinical input combinations might change at different time periods, and might be different for different clinical outcomes of interest.

[0085] We then train a classifier to map the selected relevant clinical inputs to the outputs. A classifier assigns relative weightings to individual marker values. We note that the construct of a classifier is not crucial to our method. Any mapping procedure between inputs and outputs that produces a measure of goodness of fit, for example, maximizing the area under the receiver operator curve of sensitivity versus 1-specificity, for the training data and maximizes it with a standard optimization routine on a series of validation sets would also suffice.

[0086] Once the classifier is trained, it is ready for use by a clinician. The clinician enters the same classifier inputs used during training of the network by assaying the selected markers and collecting relevant clinical information for a new patient, and the trained classifier outputs a maximum likelihood estimator for the value of the output given the inputs for the current patient.

The clinician or patient can then act on this value. We note that a straightforward extension of our technique could produce an optimum range of output values given the patient's inputs as well as specific threshold values for inputs.

[0087] One versed in the ordinary state of the art knows that many other markers in the literature once measured from the blood in a diseased and healthy patient, selected through use of an feature selection algorithm might be diagnostic of cardiovascular illness if measured in combination with others and evaluated together with a nonlinear classification algorithm. These markers have been previously described in U.S. patent application Ser. No. 11/046,592, entitled 'Cellular fibronectin as a diagnostic marker in stroke and methods of use thereof', filed January 29, 2005, and have been previously considered for diagnosis or prognosis of cardiovascular illness and thus are not novel in themselves. The list detailed in said patent application is meant to serve as illustrative and not meant to be exhaustive. Selected marker descriptions in the following list are similar to U.S. patent application 20040126767 and/or U.S. patent application 20040219509, both of which are noted as prior art. However, the instant invention goes beyond what is taught or anticipated in these applications, providing a rigorous methodology of discovering which representative markers are best suited to building a predictive model for determining a clinical outcome and building a model for interpolating between such markers to determine clinical outcome, while the methodology described in U.S. Patent application 20040126767 and/or U.S. patent application 20040219509 rely on simple linear relationships between markers and linear optimization techniques to find them. As also previously discussed in the instant invention, neither the general markers used, the idea of combinations of such markers, nor techniques used to analyze them are novel.

[0088] Markers specifically related to hemorrhagic transformation

[0089] Cellular Fibronectin, or ED1+. is an adhesive glycoprotein, is a fibronectin synthesized in endothelial cells. It contains an extra Type III domain (ED1, or EDA/EIIIA), as a result of alternative mRNA splicing. It circulates in the blood in small quantities. Endothelial cells do not express the ED1 domain under normal circumstances, but the ED1 domain is included in fibronectin molecules in pathological conditions (see for instance Dubin D, Peters JH, Brown LF, Logan B, Kent KC, Berse B, Berven S, Cercek B, Sharifi BG, Pratt RE: Balloon

catheterization induced arterial expression of embryonic fibronectins. Arterioscler Thromb Vasc Biol. 15:1958 1967, 1995.) Because ED1-fn is not stored in cellular granules, concentration increases indicate increased synthesis (26). Because c-Fn is largely confined to the vascular endothelium, high plasma lvels of this molecule might be indicative of endothelial damage. Plasma c-Fn levels have been reported to be increased in patients with vascular injury secondary to vasculitiis, sepsis, acute major trauma, diabetes, and patients with ischemic stroke (see for instance Peters et al. Elevated plasma levels of ED1+ 'cellular fibronectin' in patients with vascular injury J Lab Clin Med. 1989. 113:586-597).

[0090] Atherosclerotic plaque rupture is part of a dynamic inflammatory process of atherosclerotic vascular disease which starts from inception and continues through plaque growth, rupture and ultimately thrombosis. Suggested markers of atherosclerotic plaque rupture that would be suitable for inclusion in a stroke or stroke sub-type diagnostic include human neutrophil elastase, inducible nitric oxide synthase, lysophosphatidic acid, malondialdehyde-modified low density lipoprotein, and various members of the matrix metalloproteinase (MMP) family, including MMP-1, -2, -3, 7, 9, 12, and -19, the main family members of which we discuss below.

[0091] Matrix metalloproteinases (MMPs) are a family of zinc-binding proteolytic enzymes that normally remodel the extracellular matrix and pathologically attack substrates as part of the neuroinflammatory response. MMP-2 (72 kDa, gelatinase A) and MMP-9 (92 kDa, gelatinase B) specifically attack type IV collagen, laminin, and fibronectin, which are the major components of the basal lamina around cerebral blood vessels. Proenzyme activation and enzyme activities are tightly regulated by tissue inhibitors of MMPs (TIMPs) and interactions with surrounding extracellular matrix molecules. Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) play a significant role in regulating angiogenesis, the process of new blood vessel formation. Interstitial collagenase (MMP-1), 72 kDa gelatinase A/type IV collagenase (MMP-2), and 92 kDa gelatinase B/type IV collagenase (MMP-9) dissolve extracellular matrix (ECM) and may initiate and promote angiogenesis. TIMP-1, TIMP-2, TIMP-3, and possibly, TIMP-4 inhibit neovascularization. A new paradigm is emerging that matrilysin (MMP-7), MMP-9, and metalloelastase (MMP-12) may block angiogenesis by converting plasminogen to angiostatin, which is one of the most potent

angiogenesis antagonists. MMPs and TIMPs play a complex role in regulating angiogenesis. MMP-9 has been implicated as a marker of stoke severity(Montaner *et al.*, Matrix Metalloproteinase Expression After Human Cardioembolic Stroke, *Stroke.* 2001;32:1759.)

[0092] How to measure various markers

[0093] One of ordinary skill in the art know several methods and devices for the detection and analysis of the markers of the instant invention. With regard to polypeptides or proteins in patient test samples, immunoassay devices and methods are often used. These devices and methods can utilize labeled molecules in various sandwich, competitive, or non-competitive assay formats, to generate a signal that is related to the presence or amount of an analyte of interest. Additionally, certain methods and devices, such as biosensors and optical immunoassays, may be employed to determine the presence or amount of analytes without the need for a labeled molecule.

[0094] Preferably the markers are analyzed using an immunoassay, although other methods are well known to those skilled in the art (for example, the measurement of marker RNA levels). The presence or amount of a marker is generally determined using antibodies specific for each marker and detecting specific binding. Any suitable immunoassay may be utilized, for example, enzyme-linked immunoassays (ELISA), radioimmunoassay (RIAs), competitive binding assays, and the like. Specific immunological binding of the antibody to the marker can be detected directly or indirectly. Direct labels include fluorescent or luminescent tags, metals, dyes, radionuclides, and the like, attached to the antibody. Indirect labels include various enzymes well known in the art, such as alkaline phosphatase, horseradish peroxidase and the like. For an example of how this procedure is carried out on a machine, one can use the RAMP Biomedical device, called the Clinical Reader sup.TM., which uses the fluoresent tag method, though the skilled artisan will know of many different machines and manual protocols to perform the same assay. Diluted whole blood is applied to the sample well. The red blood cells are retained in the sample pad, and the separated plasma migrates along the strip. Fluorescent dyed latex particles bind to the analyte and are immobilized at the detection zone. Additional particles are immobilized at the internal control zone. The fluorescence of the detection and internal control zones are measured on the RAMP Clinical Reader sup.TM., and the ratio

between these values is calculated. This ratio is used to determine the analyte concentration by interpolation from a lot-specific standard curve supplied by the manufacturer in each test kit for each assay.

[0095] The use of immobilized antibodies specific for the markers is also contemplated by the present invention and is well known by one of ordinary skill in the art. The antibodies could be immobilized onto a variety of solid supports, such as magnetic or chromatographic matrix particles, the surface of an assay place (such as microtiter wells), pieces of a solid substrate material (such as plastic, nylon, paper), and the like. An assay strip could be prepared by coating the antibody or a plurality of antibodies in an array on solid support. This strip could then be dipped into the test sample and then processed quickly through washes and detection steps to generate a measurable signal, such as a colored spot.

[0096] The analysis of a plurality of markers may be carried out separately or simultaneously with one test sample. Several markers may be combined into one test for efficient processing of a multiple of samples. In addition, one skilled in the art would recognize the value of testing multiple samples (for example, at successive time points) from the same individual. Such testing of serial samples will allow the identification of changes in marker levels over time. Increases or decreases in marker levels, as well as the absence of change in marker levels, would provide useful information about the disease status that includes, but is not limited to identifying the approximate time from onset of the event, the presence and amount of salvagable tissue, the appropriateness of drug therapies, the effectiveness of various therapies, identification of the severity of the event, identification of the disease severity, and identification of the patient's outcome, including risk of future events.

[0097] An assay consisting of a combination of the markers referenced in the instant invention may be constructed to provide relevant information related to differential diagnosis. Such a panel may be constucted using 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more or individual markers. The analysis of a single marker or subsets of markers comprising a larger panel of markers could be carried out methods described within the instant invention to optimize clinical sensitivity or specificity in various clinical settings. The clinical sensitivity of an assay is defined as the percentage of those with the disease that the assay correctly predicts, and the

specificity of an assay is defined as the percentage of those without the disease that the assay correctly predicts (Tietz Textbook of Clinical Chemistry, 2.sup.nd edition, Carl Burtis and Edward Ashwood eds., W. B. Saunders and Company, p. 496).

[0098] The analysis of markers could be carried out in a variety of physical formats as well. For example, the use of microtiter plates or automation could be used to facilitate the processing of large numbers of test samples. Alternatively, single sample formats could be developed to facilitate immediate treatment and diagnosis in a timely fashion, for example, in ambulatory transport or emergency room settings. Particularly useful physical formats comprise surfaces having a plurality of discrete, addressable locations for the detection of a plurality of different analytes. Such formats include protein microarrays, or "protein chips" (see, e.g., Ng and Ilag, J. Cell Mol. Med. 6: 329-340 (2002)) and capillary devices.

[0099] In another embodiment, the present invention provides a kit for the analysis of markers. Such a kit preferably comprises devises and reagents for the analysis of at least one test sample and instructions for performing the assay. Optionally the kits may contain one or more means for using information obtained from immunoassays performed for a marker panel to rule in or out certain diagnoses. Marker antibodies or antigens may be incorporated into immunoassay diagnostic kits depending upon which marker autoantibodies or antigens are being measured. A first container may include a composition comprising an antigen or antibody preparation. Both antibody and antigen preparations should preferably be provided in a suitable titrated form, with antigen concentrations and/or antibody titers given for easy reference in quantitative applications.

[0100] The kits may also include an immunodetection reagent or label for the detection of specific immunoreaction between the provided antigen and/or antibody, as the case may be, and the diagnostic sample. Suitable detection reagents are well known in the art as exemplified by radioactive, enzymatic or otherwise chromogenic ligands, which are typically employed in association with the antigen and/or antibody, or in association with a second antibody having specificity for first antibody. Thus, the reaction is detected or quantified by means of detecting or quantifying the label. Immunodetection reagents and processes suitable for application in connection with the novel methods of the present invention are generally well known in the art.

[0101] The reagents may also include ancillary agents such as buffering agents and protein stabilizing agents, e.g., polysaccharides and the like. The diagnostic kit may further include where necessary agents for reducing background interference in a test, agents for increasing signal, software and algorithms for combining and interpolating marker values to produce a prediction of clinical outcome of interest, apparatus for conducting a test, calibration curves and charts, standardization curves and charts, and the like.

[0102] In a more particular aspect the invention relates to a rapid multiple marker panel containing antibodies to selected markers that employs latex agglutination. Thus, in one embodiment the invention provides a kit for diagnosing stroke or stroke sub-type comprising: (1) an agglutinating immunosorbent for said selected markers, and (2) a control such as saline or a known concentration of said selected markers.

[0103] In another embodiment the invention relates to a kit for detecting various markers indicative of stroke or stroke subtype diagnosis comprising: (1) an immunosorbent for selected markers indicative of stroke or stroke subtype diagnosis, and (2) an indicator reagent comprising secondary antibodies attached to a signal generating compound for each individual marker. The secondary antibodies can be specific for each individual marker or for the primary antibodies in the immunosorbent. In a preferred embodiment the kits further comprise an immunosorbent for glutamate or polyglutamate, and/or an immunosorbent for homocysteine or polyhomocysteine, and secondary antibodies against the glutamate and/or homocysteine, or to the primary antibodies on the immunosorbents against the glutamate or homocysteine. The immunosorbent preferably comprises anti-antibodies for the biomarkers bound to a solid support.

[0104] In another aspect the present invention relates to a test-kit that relies upon PCR amplification for measuring selected markers indicative of stroke or stroke subtype diagnosis. Thus, in another embodiment the invention provides a kit comprising: (a) one or more oligonucleotide primers attached to a solid phase, (b) indicator reagent attached to a signal-generating compound capable of generating a detectable signal from oligonucleotides, and (c) a control sample (i.e. template cDNA). The reagents may also include ancillary agents such as

buffering agents, polymerase agents, and the like. The diagnostic kit may further include, where necessary, other members of the signal-producing system of which system the detectable group is a member (e.g., enzyme and non-enzyme substrates), agents for reducing background interference in a test, agents for increasing the signal, apparatus for conducting a test, and the like.

[0105] In another embodiment of test-kit comprises (a) a solid phase to which biological fluids for receiving total DNA including selected marker cDNA indicative of stroke or stroke subtype diagnosis could be attached, (b) oligonucleotide primers, preferably in a ready-to-use PCR buffer, and (c) a control sample (i.e. template cDNA). Ancillary agents as described above may similarly be included.

[0106] In another embodiment the invention provides a diagnostic kit for detecting selected markers indicative of stroke or stroke subtype diagnosis autoantibodies comprising (a) a polypeptide of the selected markers indicative of stroke or stroke subtype diagnosis, fragment thereof, or analog or derivative thereof, (b) an indicator reagent comprising a secondary antibody specific for the autoantibody or the polypeptide attached to a signal-generating compound; and (c) a control sample, such as a known concentration of said selected markers indicative of stroke or stroke subtype diagnosis polyclonal antibodies. The reagents may also include ancillary agents such as buffering agents and protein stabilizing agents, e.g., polysaccharides and the like. The diagnostic kit may further include, where necessary, other members of the signal-producing system of which system the detectable group is a member (e.g., enzyme and non-enzyme substrates), agents for reducing background interference in a test, agents to increase the signal, apparatus for conducting a test, calibration and standardization information or instructions, and the like.

[0107] Methodology of Marker Selection, Analysis, and Classification

[0108] Non-linear techniques for data analysis and information extraction are important for identifying complex interactions between markers that contribute to overall presentation of the clinical outcome. However, due to the many features involved in association studies such as the one proposed, the construction of these in-silico predictors is a complex process. Often one

must consider more markers to test than samples, missing values, poor generalization of results, selection of free parameters in predictor models, confidence in finding a sub-optimal solution and others. Thus, the process for building a predictor is as important as designing the protocol for the association studies. Errors at each step can propagate downstream, affecting the generalizability of the final result.

[0109] An overview of our process of model development, describing the five main steps and some techniques that the instant invention will use to build an optimal biomarker panel of response for each clinical outcome, is described in detail in U.S. provisional patent application Ser. No. 60/505,606, filed Sept. 23, 2003, now U.S. patent application Ser. No. 10/948,834, entitled 'DIAGNOSTIC MARKERS OF CARDIOVASCULAR ILLNESS AND METHODS OF USE THEREOF, which is hereby incorporated by reference in its entirety, including all tables, figures, and claims. One of ordinary skill in the art will know that it is best to use a 'toolbox' approach to the various steps, trying several different algorithms at each step, and even combining several as in Step Five. Since one does not know *a priori* the distribution of the true solution space, trying several methods allows a thorough search of the solution space of the observed data in order to find the most optimal solutions (i.e. those best able to generalize to unseen data). One also can give more confidence to predictions if several independent techniques converge to a similar solution.

[0117] Method for Defining Panels of Markers

[0118] In practice, data may be obtained from a group of subjects. The subjects may be patients who have been tested for the presence or level of certain markers. Such markers and methods of patient extraction are well known to those skilled in the art. A particular set of markers may be relevant to a particular condition or disease. The method is not dependent on the actual markers. The markers discussed in this document are included only for illustration and are not intended to limit the scope of the invention. Examples of such markers and panels of markers are described in the instant invention and the incorporated references.

[0119] Well-known to one of ordinary skill in the art is the collection of patient samples. A preferred embodiment of the instant invention is that the samples come from two or more

different sets of patients, one a disease group of interest and the other(s) a control group, which may be healthy or diseased in a different indication than the disease group of interest. For instance, one might want to look at the difference in blood-borne markers between patients who have had stroke and those who had stroke mimic to differentiate between the two populations.

[0120] The blood samples are assayed, and the resulting set of values are put into a database, along with outcome, also called phenotype, information detailing the illness type, for instance stroke mimic, once this is known. Additional clinical details such as time from onset of symptoms and patient physiological, medical, and demographics, the sum total called patient characteristics, are put into the database. The time from onset is important to know as initial marker values from onset of symptoms can change significantly over time on a timeframe of tens of minutes. Thus, a marker may be significant at one point in the patient history and not at another in predicting diagnosis or prognosis of cardiovascular disease, damage or injury. The database can be simple as a spreadsheet, i.e. a two-dimensional table of values, with rows being patients and columns being filled with patient marker and other characteristic values.

[0121] From this database, a computerized algorithm can first perform pre-processing of the data values. This involves normalization of the values across the dataset and/or transformation into a different representation for further processing. The dataset is then analyzed for missing values. Missing values are either replaced using an imputation algorithm, in a preferred embodiment using KNN or MVC algorithms, or the patient attached to the missing value is exised from the database. If greater than 50% of the other patients have the same missing value then value can be ignored.

[0122] Once all missing values have been accounted for, the dataset is split up into three parts: a training set comprising 33-80% of the patients and their associated values, a testing set comprising 10-50% of the patients and their associated values, and a validation set comprising 1-50% of the patients and their associated values. These datasets can be further sub-divided or combined according to algorithmic accuracy. A feature selection algorithm is applied to the training dataset. This feature selection algorithm selects the most relevant marker values and/or patient characteristics. Preferred feature selection algorithms include, but are not limited to, Forward or Backward Floating, SVMs, Markov Blankets, Tree Based Methods with node

discarding, Genetic Algorithms, Regression-based methods, kernel-based methods, and filter-based methods.

[0123] Feature selection is done in a cross-validated fashion, preferably in a naïve or k-fold fashion, as to not induce bias in the results and is tested with the testing dataset. Cross-validation is one of several approaches to estimating how well the features selected from some training data is going to perform on future as-yet-unseen data and is well-known to the skilled artisan. Cross validation is a model evaluation method that is better than residuals. The problem with residual evaluations is that they do not give an indication of how well the learner will do when it is asked to make new predictions for data it has not already seen. One way to overcome this problem is to not use the entire data set when training a learner. Some of the data is removed before training begins. Then when training is done, the data that was removed can be used to test the performance of the learned model on ``new'' data.

[0124] Once the algorithm has returned a list of selected markers, one can optimize these selected markers by applying a classifer to the training dataset to predict clinical outcome. A cost function that the classifier optimizes is specified according to outcome desired, for instance an area under receiver-operator curve maximizing the product of sensitivity and specificity of the selected markers, or positive or negative predictive accuracy. Testing of the classifier is done on the testing dataset in a cross-validated fashion, preferably naïve or k-fold cross-validation. Further detail is given in U.S. patent application 09/611,220, incorporated by reference. Classifiers map input variables, in this case patient marker values, to outcomes of interest, for instance, prediction of stroke sub-type. Preferred classifiers include, but are not limited to, neural networks, Decision Trees, genetic algorithms, SVMs, Regression Trees, Cascade Correlation, Group Method Data Handling (GMDH), Multivariate Adaptive Regression Splines (MARS), Multilinear Interpolation, Radial Basis Functions, Robust Regression, Cascade Correlation + Projection Pursuit, linear regression, Non-linear regression, Polynomial Regression, Regression Trees, Multilinear Interpolation, MARS, Bayes classifiers and networks, and Markov Models, and Kernel Methods.

[0125] The classification model is then optimized by for instance combining the model with other models in an ensemble fashion. Preferred methods for classifier optimization include, but

are not limited to, boosting, bagging, entropy-based, and voting networks. This classifier is now known as the final predictive model. The predictive model is tested on the validation data set, not used in either feature selection or classification, to obtain an estimate of performance in a similar population.

[0126] The predictive model can be translated into a decision tree format for subdividing the patient population and making the decision output of the model easy to understand for the clinician. The marker input values might include a time since symptom onset value and/or a threshold value. Using these marker inputs, the predictive model delivers diagnositic or prognostic output value along with associated error. The instant invention anticipates a kit comprised of reagents, devices and instructions for performing the assays, and a computer software program comprised of the predictive model that interprets the assay values when entered into the predictive model run on a computer. The predictive model receives the marker values via the computer that it resides upon.

[0127] Once patients are exhibiting symptoms of cardiovascular illness, for instance stroke, a blood sample is drawn from the patient using standard techniques well known to those of ordinary skill in the art and assayed for various blood-borne markers of cardiovascular illness. Assays can be preformed through immunoassays or through any of the other techniques well known to the skilled artisan. In a preferred embodiment, the assay is in a format that permits multiple markers to be tested from one sample, such as the Luminex platform.TM., and/or in a rapid fashion, defined to be under 30 minutes and in the most preferred enablement of the instant invention, under 15 minutes. The values of the markers in the samples are inputed into the trained, tested, and validated algorithm residing on a computer, which outputs to the user on a display and/or in printed format on paper and/or transmits the information to another display source the result of the algorithm calculations in numerical form, a probability estimate of the clinical diagnosis of the patient. There is an error given to the probability estimate, in a preferred embodiment this error level is a confidence level. The medical worker can then use this diagnosis to help guide treatment of the patient.

EXAMPLE I

[0128] Prospectively studied patients (n= 87, 59% men; mean age 67±12 years) received intravenous tPA following the European Cooperative Acute Stroke Study (ECASS) II criteria.12 Thrombolytic therapy was administered within 6 hours from the beginning of the symptoms at a dose of 0.9 mg/kg body weight, with an upper dose limit of 90 mg per patient. Ten percent of the total dose was given as a bolus over 1 to 2 minutes, followed by a 60-minute infusion of the remaining dose. The mean time to the infusion of the drug was 160±46 minutes. Seventy-one patients received the treatment within 3 hours from onset of symptoms, whereas 16 patients received tPA between 3 and 6 hours within onset of symptoms. Thirty healthy control subjects matched by age and sex (male: 57%; mean age: 63±9 years) and without history of neurological disorders or vascular risk factors were also included in the study. To determine the effect of stroke on the levels of the molecules, plasma c-Fn and MMP-9 concentrations were also determined in 100 patients with acute ischemic stroke who did not receive tPA treatment and in whom HT did not develop (male: 59%; mean age: 67±6 years; mean time to inclusion: 7.2_3.9 hours). Neither the patients nor the controls had inflammatory, hematological, or infectious diseases, cancer, or severe renal or liver failure. The ethics committee approved the protocol in each center, and informed consent was obtained from patients or their relatives. Medical history recording potential stroke risk factors, clinical examination, blood and coagulation tests, 12-lead electrocardiogram, chest radiography, and noncontrast cranial computed tomography (CT) scan were performed at admission. Stroke subtype was classified according to the Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria. Stroke severity was assessed by a certified neurologist using the National Institutes of Health Stroke Scale (NIHSS) at admission and at 24 to 36 hours. Neurological deterioration was defined as death or an increase of ≥4 points in the NIHSS score between the 2 examinations.

[0129] Early CT signs of infarction were evaluated in the first radiological examination. The volume of hypodensity and the presence of HT were evaluated on a second cranial CT, which was performed 24 to 36 hours after treatment. Hypodensity volume was determined by using the formula 0.5 x *a* x *b* x *c*, where *a* and *b* are the largest perpendicular diameters measured on CT and *c* is the slice thickness. The HT type was classified according to the ECASS II criteria. Hemorrhagic infarction type 1 (HI-1) was defined as small petechiae along the margins of the infarct, and HI type 2 (HI-2) was defined as more confluent petechiae within the infarct area but without a space-occupying effect. Parenchymal hemorrhage type 1 (PH-1) was defined as blood clots in ≤30% of the infarcted area with some slight space-occupying effect, and PH type 2 (PH-

2) as blood clots in ≥30% of the infarcted area with substantial space-occupying effect. All CT examinations were performed by 1 investigator in each center blinded to the clinical and analytical data. Symptomatic HT was considered as being associated with neurological deterioration.

[0130] Laboratory Tests

[0131] Blood samples were taken from all patients at admission before tPA administration. Samples were collected in glass test tubes containing EDTA. Suspension of plasma was centrifuged at 3000*g* for 5 minutes and immediately frozen and stored at -80°C. Plasma MMP-9 and c-Fn levels were measured with commercially available quantitative sandwich enzyme-linked immunoabsorbent assay kits obtained from Biotrack, Amersham Pharmacia UK, and Adeza Biomedical, respectively. Determinations were performed in an independent laboratory blinded to clinical and radiological data. The intra-assay and interassay coefficients of variation were <5% for MMP-9 and c-Fn determinations.

[0132] Proportions between groups were compared using the χ^2 test. Continuous variables are expressed as mean±SD and were compared using the Student *t* test. Given that MMP-9 and c-Fn concentrations are not normally distributed, their levels were expressed as median (quartiles), and comparisons were made using the Mann–Whitney test or Kruskal–Wallis test as appropriate. The association between c-Fn levels and baseline continuous variables was assessed by calculating the Spearman correlation coefficient.

[0133] Statistical Analysis

[0134] We used cutoff values, as described by Robert et al,(Robert C, Vermont J, Bosson JL. Formulas for threshold computations. *Comput Biomed Res.* 1991;24:514–519.) to estimate the sensitivity, specificity, and predictive values of a specific concentration of plasma MMP-9 and c-Fn for HT. The importance of MMP-9 and c-Fn in the development of HT after tPA administration was determined by logistic regression analysis after adjusting for those variables evaluated at admission that were related to HT in the univariate analysis. Because plasma levels of c-Fn have been reported to increase with age and in patients with diabetes, these 2 variables were forced into the analysis. To test whether the odds of HT for c-Fn was modified by the volume of hypodensity, a second analysis was performed including this factor into the model.

Plasma MMP-9 and c-Fn were included as continuous variables because the cutoff values meant that there was a linearity of the odds ratios.

[0135] Twenty-six (30%) of the 87 patients included in the study had HT. Fifteen patients (17.2%) had HI-1, 7 (8%) had HI-2, 2 (2.3%) had PH-1, and 2 (2.3%) had PH-2. Figure 1 shows the main characteristics of patients with and without HT. The severity of neurological deficit at admission evaluated by the NIHSS score was significantly higher in patients with HT, who also displayed significantly greater volumes of hypodensity on the second cranial CT. Both clinical groups presented with similar systolic and diastolic blood pressures and glucose levels before tPA administration. Neurological deterioration was observed in 15 patients (17.2%). In 8 patients, the neurological worsening was associated with HT: 2 patients had HI-1, 3 displayed HI-2, and 3 had PH.

[0136] Results

[0137] Plasma c-Fn concentrations before tPA administration were significantly higher in patients with HT (4.8 [3.4,5.9] μg/mL) than in those without HT (1.7 [1.4, 2.5] μg/mL) and both the healthy subjects (1.3 [0.9, 1.6] μg/mL) and the patients not treated with tPA (1.4 [1.1, 1.8] μg/mL) (all $P<0.001$). Moreover, we found that the greater the severity of the bleeding the higher the levels of c-Fn (Figure 2A). A similar effect was found in those patients who were treated within 3 hours of onset of symptoms (n=71) (Figure 2B). The levels of c-Fn were not statistically different in patients with symptomatic and asymptomatic HT, although there was a clear trend for the levels to be higher in patients with symptomatic HT (5.8 [4.0, 6.9] μg/mL versus 4.5 [2.7, 5.4] μg/mL; $P<0.054$).

[0138] Plasma MMP-9 concentrations before tPA administration were also significantly higher in those patients with HT (170.3 [101.4, 196.2] ng/mL) than in those without HT (87.2 [54.8, 115.1] ng/mL) and in both the healthy subjects (53.7 [39.5, 79.4] ng/mL) and the patients not treated with tPA (62 [40, 93.8] ng/mL) (all $P<0.001$). As observed with c-Fn levels, the greater the severity of the bleeding, the higher the levels of MMP-9, both in patients treated within 6 (Figure 3A) and 3 hours (Figure 3B). No differences were found in MMP-9 levels between symptomatic and asymptomatic HT (170 [121, 214] ng/mL versus 170 [87, 194] ng/mL; $P<0.338$).

[0139] Plasma c-Fn concentrations were significantly higher in patients with early signs of ischemia (n=37) on cranial CT (2.8 [1.7, 5.0] μg/mL) than in those without (1.9 [1.4, 2.9] μg/mL) (P<0.012). Plasma c-Fn correlated positively with MMP-9 levels (r=0.671, P<0.001) and the hypodensity volume (r=0.364, P<0.001). No correlation was found between plasma c-Fn levels and other variables related to HT such as serum glucose concentrations, blood pressure levels, or the severity of neurological deficit at admission.

[0140] As shown in Figure 4, only plasma c-Fn levels remained independently associated with HT after adjustment for age, history of diabetes, baseline NIHSS score, and plasma MMP-9 levels. The odds of c-Fn levels for HT did not substantially change after the inclusion of the volume of hypodensity into the analysis (OR, 2.1; 95% CI, 1.3 to 3.4; P<0.002). Plasma c-Fn level was also the only factor independently associated with HT after adjustment for potential confounders when the analysis was limited to the 71 patients treated within 3 hours (OR, 1.9; 95% CI, 1.2 to 3.3; P<0.006).

[0141] Because it has been reported that HI-2 and PH occur more often in patients who receive tPA treatment, and also based on the observation of the authors results in Figure 1 that clearly demonstrate higher levels of c-Fn in patients with HI-2 and PH, we calculated the c-Fn and MMP-9 cutoff values with the highest sensitivity and specificity for these types of HT. Receiver operating curves for c-Fn and MMP-9 are given in Figures 5 and 6, respectively. This post-hoc explanatory analysis showed that plasma MMP-9 concentrations ≥140 ng/mL predicted the development of HI-2 and PH with a sensitivity of 81%, specificity of 88%, positive predictive value of 41%, and negative predictive value of 98%. More interestingly, the sensitivity, specificity, and positive and negative predictive values of plasma c-Fn ≥3.6 μg/mL for the prediction of HI-2 and PH were 100%, 96%, 44%, and 100%, respectively.

[0142] Marker threshold Determination

[0143] The subjects with HT that are relevant, as defined by either HI-2, or PT-1, or PT-2, have univariate normal distributions for their c-Fn and MMP-9 assayed values.

[0144] From one perspective, it may be assumed that the data [c-Fn, MMP-9] were sampled at random from a **joint** normal distribution. Bounds may be determined such that the probability that a sample will fall outside the bounds is less than a predefined quantity. For the selected

probabilities of 5%, 1% and 0.5%, these bounds were plotted based on the assumptions of IID data samples and an underlying **joint** normal distribution. This plot is shown in Figure 7. The minimum values for c-Fn and MMP-9 were identified for each of the three bounds. This is shown in Table 1.

[0145] **Table 1: Confidence values for various thresholds of c-Fn and MMP-9**

Probability of new sample contained in bounds	Minimum MMP-9 value	Minimum c-Fn value
95%	91.0000	4.3270
99%	61.0000	3.7069
99.5%	50.0000	3.4734

[0146] Alternatively, the threshold for c-Fn corresponding to the probability that a patient with relevant hemorrhagic transformation will have an assayed result below the threshold may be determined on a univariate basis, again with the assumption that c-Fn is normally distributed. Table 2 below shows a series of selected sensitivities for identifying a RHT using c-Fn alone and the corresponding c-Fn threshold. If a subject has a c-Fn greater than the threshold, then the test result is positive for RHT.

[0147] **Table 2: Confidence values for various c-Fn thresholds**

Sensitivity	c-Fn Threshold
95%	4.637
99%	3.957
99.5%	3.707
99.9%	3.1876

[0148] However, there is also a worst-case perspective. There is a correlation between the c-Fn and MMP-9 values that is observed for patients with a HT. It may be assumed that the RHT is a subclass of the HT, rather than a totally different class. The relationship between MMP-9 and c-Fn for subjects with an observed HT is shown in Figure 8. It may be assumed that we have observed a sparse sampling of the distribution for c-Fn as a function of MMP-9. It may also be noticed that the majority of RHT observations are above the reference line, and that those RHT below the reference line are very close to it with respect to their c-Fn and MMP-9 deviations. Subsequently, we project the minumum expected values for MMP-9 to derive corresponding minimum approximate expected values for c-Fn.

[0149] Assuming the reference line then corresponds to the approximate lower minimum for c-Fn as a function of MMP-9 in the case of RHT, the minimum expected values for MMP-9 and the corresponding predicted approximate minimum values for c-Fn are given below in Table 3. The corresponding false positive percentage of Non-HT subjects that would be classified as Rel HT is also provided.

[0150] **Table 3: C-Fn threshold values and expected false-positive rates**

Sensitivity for RHT	MMP-9 Minimum Expected Value	CFN threshold	%FP
95%	105.5	3.42965	4.21%
99%	72.83	2.668439	13.68%
99.5%	60.88	2.390004	18.95%
99.9%	36.21	1.815193	33.68%

[0151] Note that by using both MMP-9 and c-Fn, the false positive RHT diagnostic test rate may be decreased while maintaining the same sensitivity for identifying RHT subjects. In this case we set the sensitivity in the worst-case scenario to 99%, for example, and the MMP-9 and c-Fn thresholds at 72.83 and 2.67, respectively. The test must then have values exceeding both of these levels to indicate a positive RHT. Considering the first scenario where we have a representative sampling, a threshold value of approximately 3.2 would provide a very high degree of sensitivity for RHT (99.9%) with a specificity of approximately 95%.

[0152] However, if we do not have a reasonable sampling for the RHT subjects to directly estimate the minimum c-Fn threshold for the diagnostic, as in the second scenario, but there is a relationship between MMP-9 and c-Fn for HT in general, then we may exploit this relationship to better estimate the lower bounds for c-Fn for use in a diagnostic. In this case, a c-Fn threshold with high sensitivity of 99 or 99.5% (values of 2.67 or 2.39, respectively) could be selected with expected specificities of approximately, 85% and 80% respectively. While not ideal, the very high expected sensitivity, even given a potentially weak original sample distribution for the RHT subjects), will allow a firm basis for excluding patients from receiving therapy that is expected to result in a RHT, while the diagnostic specificity of 80-85% will allow many patients to receive therapy with high confidence.

[0153] One skilled in the art readily appreciates that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The examples provided herein are representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Modifications therein and other uses will occur to those skilled in the art. These modifications are encompassed within the spirit of the invention and are defined by the scope of the claims.

[0154] It will be readily apparent to a person skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

[0155] All patents and publications mentioned in the specification are indicative of the levels of those of ordinary skill in the art to which the invention pertains. All patents and publications are herein incorporated by reference to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference.

[0156] The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

[0157] Other embodiments are set forth within the following claims.

CLAIMS

We claim:

1. A method of determining presence or risk of hemorrhage in a human subject CHARACTERIZED IN THAT

a test sample obtained from a human subject is analyzed for presence or amount of (1) cellular fibronectin and (2) one or more additional markers both proteomic and non-proteomic for, or mass spectrometry peak levels of, any of apoptosis, cellular adhesion, cellular injury, coagulation, glial activation, inflammatory mediation, myelin breakdown, thrombosis, vascular damage, and specific and non-specific markers of cerebral injury; and then

(1) the presence or amount of said cellular fibronectin and said more additional markers or peak levels, is correlated with (2) clinical patient information, other than the cellular fibronectin and additional makers or peak levels for cerebral injury, in order to deduce a probability of present or future risk of a hemorrhage for the subject.

2. The method according to claim 1 FURTHER CHARACTERIZED IN THAT the correlating transpires by

determining the expression levels or mass spectrometry peak levels of one or more proteomic marker(s) or mass-to-charge ratio(s) and the numerical quantity of one or more non-proteomic marker(s) or mass-to-charge ratio(s) from humans suspected or known to have some form of hemorrhage;

comparing said determined levels and numerical values to humans known to have said matched type of hemorrhage; and

training an algorithm to identify patterns of differences in said humans which correlate with the prescience or absence of said matched type of hemorrhage, respectively.

3. The method according to claim 2 FURTHER CHARACTERIZED IN THAT the training of the algorithm on characteristic protein patterns includes the steps of

obtaining numerous examples of (i) said proteomic and non-proteomic data, and (ii) historical clinical results corresponding to this proteomic and non-proteomic data;

constructing an algorithm suitable to map (i) said protein expression levels or mass spectrometry peak mass-to-charge ratio(s) and said non-proteomic values as inputs to the

algorithm, to (ii) the historical clinical results as outputs of the algorithm;

exercising the constructed algorithm to so map (i) the said protein expression levels or mass spectrometry peak mass-to-charge ratio(s) and said non-proteomic values as inputs to (ii) the historical clinical results as outputs; and

conducting an automated procedure to vary the mapping function, inputs to outputs, of the constructed and exercised algorithm in order that, by minimizing an error measure of the mapping function, a more optimal algorithm mapping architecture is realized;

wherein realization of the more optimal algorithm mapping architecture, also known as feature selection, means that any irrelevant inputs are effectively excised, meaning that the more optimally mapping algorithm will substantially ignore said protein expression levels or mass spectrometry peak mass-to-charge ratio(s) and said non-proteomic values that are irrelevant to output clinical results; and

wherein realization of the more optimal algorithm mapping architecture, also known as feature selection, also means that any relevant inputs are effectively identified, making that the more optimally mapping algorithm will serve to identify, and use, those input protein expression levels or mass spectrometry peak mass-to-charge ratio(s) and said non-proteomic values that are relevant, in combination, to output clinical results that would result in a clinical detection of disease, disease diagnosis, disease prognosis, or treatment outcome or a combination of any two, three or four of these actions.

4. The method according to claim 3 FURTHER CHARACTERIZED IN THAT the constructed algorithm is drawn from the group consisting essentially of: linear or nonlinear regression algorithms; linear or nonlinear classification algorithms; ANOVA; neural network algorithms; genetic algorithms; support vector machines algorithms; hierarchical analysis or clustering algorithms; hierarchical algorithms using decision trees; kernel based machine algorithms such as kernel partial least squares algorithms, kernel matching pursuit algorithms, kernel fisher discriminate analysis algorithms, or kernel principal components analysis algorithms; Bayesian probability function algorithms; Markov Blanket algorithms; a plurality of algorithms arranged in a committee network; and forward floating search or backward floating search algorithms.

5. The method according to claim 3 FURTHER CHARACTERIZED IN THAT the feature selection process employs an algorithm drawn from the group consisting essentially of: linear or

nonlinear regression algorithms; linear or nonlinear classification algorithms; ANOVA; neural network algorithms; genetic algorithms; support vector machines algorithms; hierarchical analysis or clustering algorithms; hierarchical algorithms using decision trees; kernel based machine algorithms such as kernel partial least squares algorithms, kernel matching pursuit algorithms, kernel fisher discriminate analysis algorithms, or kernel principal components analysis algorithms; Bayesian probability function algorithms; Markov Blanket algorithms; recursive feature elimination or entropy-based recursive feature elimination algorithms; a plurality of algorithms arranged in a committee network; and forward floating search or backward floating search algorithms.

6. The method according to claim 3 FURTHER CHARACTERIZED IN THAT a tree algorithm is trained to reproduce the performance of another machine-learning classifier or regressor by enumerating the input space of said classifier or regressor to form a plurality of training examples sufficient (1) to span the input space of said classifier or regressor and (2) train the tree to emulate the performance of said classifier or regressor.

7. The method according to claim 1 FURTHER CHARACTERIZED IN THAT the risk of hemorrhage is risk of hemorrhage following thrombolytic therapy.

8. The method of claim 7 FURTHER CHARACTERIZED IN THAT the diagnosing the risk of hemorrhage following thrombolytic therapy further includes

measuring the level of cellular fibronectin (c-Fn) alone;

evaluating the patient's risk of hemorrhage following thrombolytic therapy from said measured level of c-Fn; and

administering stroke therapy as appropriate to the evaluated risk of hemorrhage.

9. The method according to claim 7 FURTHER CHARACTERIZED IN THAT the one or more additional markers includes, in addition to cellular fibronectin (c-Fn), the proteomic markers MMP-9, TAFI, and PAI-1.

10. The method according to claim 7 FURTHER CHARACTERIZED IN THAT the one or more additional markers includes, in addition to cellular fibronectin (c-Fn), a proteomic marker of endothelial injury.

11. The method of claim 1 FURTHER CHARACTERIZED IN THAT the markers in addition to cellular-fibronectin, are selected from the group consisting of two or more of the following: Glial fibrillary acidic protein, apolipoprotein CI (ApoC-I), apolipoprotein CIII (ApoC-III), serum amyloid A (SAA), Platelet factor 4 (PF4), platelet-derived growth factor, antithrombin-III fragment (AT-III fragment), bradykinin, renin, haptoglobin, Creatine kinase brain band (CK-BB), adenylate kinase, lactate dehydrogenase, troponin I, troponin T, Brain Derived Neurotrophic Factor, CPK, LDH Isoenzymes, Thrombin-Antithrombin III, calcitonin, procalcitonin, c-tau, Protein C, Protein S, fibrinogen, Factor VIII, activated Protein C resistance, E-selectin, P-selectin, von Willebrand factor (vWF), platelet-derived microvesicles (PDM), plasminogen activator inhibitor-1 (PAI-1), angiotensin I, angiotensin II, angiotensin III, annexin V, arginine vasopressin, B-type natriuretic peptide (BNP), pro-BNP, atrial natriuretic peptide (ANP), N-terminal pro-ANP, pro- ANP, C-type natriuretic peptide, (CNP), c-fos, c-jun, ubiquitin, cytochrome C, beta-enolase, cardiac troponin I, cardiac troponin T, urotensin II, creatine kinase-MB, glycogen phosphorylase-BB, KL-6, endothelin-1, endothelin-2, and endothelin-3, A- , F- , and H- Fatty acid binding protein (A-, F-, H-FABP), phosphoglyceric acid mutase-MB, aldosterone, S-100beta (S100β), myelin basic protein, NR2A or NR2B NMDA receptor or fragment thereof (a subtype of N-methyl-D-aspartate (NMDA) receptors), Intracellular adhesion molecule (ICAM or CD54), Neuronal cell adhesion molecule, (NCAM or CD56), C-reactive protein, caspase-3, cathepsin D, hemoglobin alpha.sub.2, human lipocalin-type prostaglandin D synthase, interleukin-1 beta, interleukin-1 receptor angonist, interleukin 2, interleukin 2 receptor, interleukin-6, IL-1, IL-8, IL-10, monocyte chemotactic protein-1, soluble intercellular adhesion molecule-1, soluble vascular cell adhesion molecule-1, MMP-2, MMP-3, MMP-9, MMP-12, MMP-9, tissue factor (TF), NDKA, RAGE, RNA-BP, TRAIL, TWEAK, UFD1, fibrin D-dimer (D-dimer), total sialic acid (TSA), TpP, heat shock protein 60, heat shock protein 70, tumor necrosis factor alpha, tumor necrosis factor receptors 1 and 2, VEGF, Calbindin-D, Proteolipid protein RU Malendialdehyde, neuron-specific enolase gamma gamma isoform (NSE γγ isoform), thrombus precursor protein, Chimerin, Fibrinopeptide A (FPA), plasmin-α 2AP complex (PAP), plasmin inhibitory complex (PIC), beta-thromboglobulin (βTG), Prothrombin fragment 1+2, PGI2, Creatinine phosphokinase brain band, neurotrophin-3 (NT-3), neurotrophin-4/5 (NT-4/5), neurokinin A, neurokinin B, neurotensin, neuropeptide Y, Lactate dehydrogenase (LDH), soluble thrombomodulin (sTM), Insulin-like growth factor-1 (IGF-1), protein kinase C gamma (PKC-γ), Secretagogin, PGE2, 8-epi PGF.sub.2alpha and

Transforming growth factor βeta (TGF-β) or markers related thereto.

12. The method of claim 11 FURTHER CHARACTERIZED IN THAT the determination of diagnostic or prognostic outcome is made in accordance with an algorithm drawn from the group consisting essentially of: linear or nonlinear regression algorithms; linear or nonlinear classification algorithms; ANOVA; neural network algorithms; genetic algorithms; support vector machines algorithms; hierarchical analysis or clustering algorithms; hierarchical algorithms using decision trees; kernel based machine algorithms such as kernel partial least squares algorithms, kernel matching pursuit algorithms, kernel fisher discriminate analysis algorithms, or kernel principal components analysis algorithms; Bayesian probability function algorithms; Markov Blanket algorithms; recursive feature elimination or entropy-based recursive feature elimination algorithms; a plurality of algorithms arranged in a committee network; and forward floating search or backward floating search algorithms.

13. The method of claim 12 w FURTHER CHARACTERIZED IN THAT the diagnostic outcome, in addition to risk of hemorrhage, is expanded to differentiate between stroke and stroke mimic.

14. The method of claim 1 FURTHER CHARACTERIZED IN THAT when clinical patient information is selected from a group consisting of Complete blood count (CBC), Coagulation test, Blood chemistry (glucose, serum electrolytes {Na, Ca, K}), Leukocyte and Neutrophil counts, platelet count, and Blood lipids tests.

15. The method of claim 1 FURTHER CHARACTERIZED IN THAT the clinical patient information is selected from a group consisting of age, weight, height, body mass index, computed tomography scan information, Magnet Resonance Image scan information, gender, time from onset of stroke-like symptoms, time to recanalization, ethnicity, heart rate, blood pressure, respiration rate, blood oxygenation, previous personal and/or familial history of cardiac events, recent cranial trauma and unequal eye dilation.

16. The method of claim 1 FURTHER CHARACTERIZED IN THAT in determination of diagnostic or prognostic outcome where both proteomic markers and non-proteomic markers are used.

17. The method of claim 1 FURTHER CHARACTERIZED IN THAT the test sample obtained from the subject is selected from the group consisting of blood, urine, blood plasma, blood serum, cerebrospinal fluid, saliva, perspiration or brain tissue, or a derivative thereof.

18. A method according to claim 1 FURTHER CHARACTERIZED IN THAT the obtaining of the test sample from the subject is within a specific time window from onset of symptoms; and

wherein the correlating is between (1) proteomic and non-marker marker values, and (2) the probability of present or future risk of a hemorrhage for the subject, for said specific time window from onset of symptoms.

19. The method of claim 18 FURTHER CHARACTERIZED IN THAT the correlating is in accordance with an algorithm drawn from the group consisting essentially of: linear or nonlinear regression algorithms; linear or nonlinear classification algorithms; ANOVA; neural network algorithms; genetic algorithms; support vector machines algorithms; hierarchical analysis or clustering algorithms; hierarchical algorithms using decision trees; kernel based machine algorithms such as kernel partial least squares algorithms, kernel matching pursuit algorithms, kernel fisher discriminate analysis algorithms, or kernel principal components analysis algorithms; Bayesian probability function algorithms; Markov Blanket algorithms; recursive feature elimination or entropy-based recursive feature elimination algorithms; a plurality of algorithms arranged in a committee network; and forward floating search or backward floating search algorithms.

20. A kit for determining presence or risk of hemorrhage in a human subject CHARACTERIZED IN THAT

a device having reagents at each of a plurality of discrete locations, each reagent and corresponding location configured and arranged to immobilize for detection one of said plurality of subject-derived markers, supports the analysis of cellular fibronectin and additional markers; and

a computer algorithm residing on a computer calculates in consideration of analyzed cellular fibronectin and additional markers a probability of present or future risk of hemorrhage for the subject.

21. The kit according to claim 20 FURTHER CHARACTERIZED IN THAT the device at

each of said plurality of discrete locations includes a poly- or monoclonal antibody specific for an individual protein or protein fragment and that binds one of said plurality of subject-derived markers.

22. The kit according to claim 20 FURTHER CHARACTERIZED IN THAT the device is an immunoassay.

23. The kit according to claim 20 FURTHER CHARACTERIZED IN THAT the device is capable of determining the levels of cellular fibronectin and said additional markers within 20 minutes or less.

24. A method of determining presence or risk of hemorrhage in a human subject, said method CHARACTERIZED IN THAT

the presence or amount of cellular fibronectin in the human subject is determined; and

the risk of hemorrhage following thrombolyic therapy is predicted in accordance with the determined presence or amount.

25. The method according to claim 24 FURTHER CHARACTERIZED IN THAT the predicting is that hemorrhage is more probable than not when the determined presence is of an amount of cellular fibronectin greater than or equal to 2.8 μg/mL or greater than or equal to 3.6 μg/mL in serum.

26. The method according to claim 24 FURTHER CHARACTERIZED IN THAT the predicting is that hemorrhage is more probable than not when the determined presence is of an amount of cellular fibronectin greater than or equal to 4.64 μg/mL, greater than or equal to 3.96 μg/mL, greater than or equal to 3.71 μg/mL,or greater than or equal to 3.19 μg/mL in serum.

27. The method according to claim 24 FURTHER CHARACTERIZED IN THAT the predicting is that hemorrhage is more probable than not when the determined presence is of an amount of cellular fibronectin greater than or equal to 3.47 μg/mL and an amount of MMP-9 greater than or equal to 50 ng/ml, an amount of cellular fibronectin greater than or equal to 3.71 μg/mL and an amount of MMP-9 greater than or equal to 61 ng/ml, or an amount of cellular fibronectin greater than or equal to 4.33 μg/mL and an amount of MMP-9 greater than or equal to 91 ng/ml in serum.

28. The method according to claim 24 FURTHER CHARACTERIZED IN THAT the predicting is that hemorrhage is more probable than not when the determined presence is of an amount of cellular fibronectin greater than or equal to 1.81 μg/mL and an amount of MMP-9 greater than or equal to 36.2 ng/ml, an amount of cellular fibronectin greater than or equal to 2.39 μg/mL and an amount of MMP-9 greater than or equal to 60.9 ng/ml, an amount of cellular fibronectin greater than or equal to 2.67 μg/mL and an amount of MMP-9 greater than or equal to 72.8 ng/ml, or an amount of cellular fibronectin greater than or equal to 3.43 μg/mL and an amount of MMP-9 greater than or equal to 105.5 ng/ml in serum.

29. A panel comprising a plurality of markers selected to selectively identify the occurrence or nonoccurrence of a stroke in a subject exhibiting one or more symptoms associated with the diagnosis of stroke CHARACTERIZED IN THAT one marker(s) selected to distinguish the occurrence of a stroke in said subject from one or more stroke mimic conditions includes cellular fibronectin.

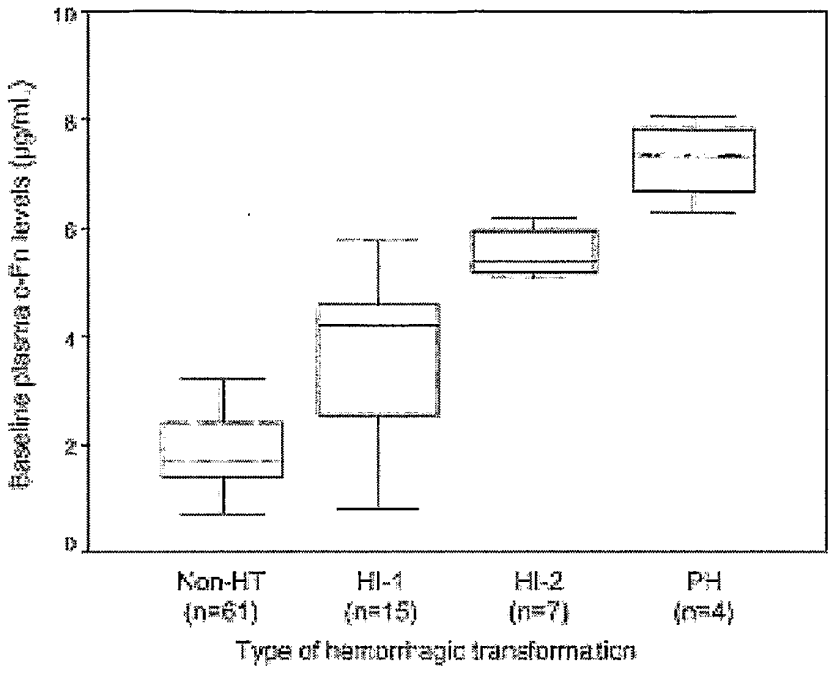
FIGURE 1

	Hemorrhagic Transformation (n=26)	Nonhemorrhagic Transformation (n=61)	*P*
Male, n (%)	11 (46.2)	37 (63.9)	0.156
Age, y	69.6±10.2	66.5±12.5	0.232
Time from stroke onset to treatment, min	158.6±32.4	159.8±50.4	0.897
Clinical characteristics			
NIHSS score at admission	15.9±4.1	12.4±5.4	0.003
Stroke subtype, n (%)			0.331
Large-artery atherosclerosis	8 (30.8)	16 (26.2)	
Cardioembolism	13 (50)	27 (44.3)	
Small-vessel disease	0	3 (4.9)	
Undetermined cause	4 (15.4)	15 (24.6)	
Others	1 (3.8)	0	
Previous treatment with aspirin, n (%)	5 (19.2)	7 (11.5)	0.332
Biochemistry and vital signs at admission			
Plasma glucose, mg/dL	147±41	151±60	0.736
Systolic blood pressure, mm Hg	157±24	163±25	0.274
Diastolic blood pressure, mm Hg	85±15	86±76	0.691
INR	1.1±0.1	1.1±0.1	0.670
Platelet count, 10^5/mm^3	218±60	217±59	0.976
Neuroimaging findings			
Early signs of infarction, n (%)	16 (61.5)	21 (34.4)	0.059
Volume of hypodensity at 24–36 h, mL	107±163	29±60	0.002

Continuous variables are expressed as mean±SD.
INR indicates International Normalized Ratio.

FIGURE 2

A


Baseline plasma c-Fn levels (µg/mL)
0
2
4
6
8
10
Non-HT (n=61)
HI-1 (n=15)
HI-2 (n=7)
PH (n=4)
Type of hemorrhagic transformation

B

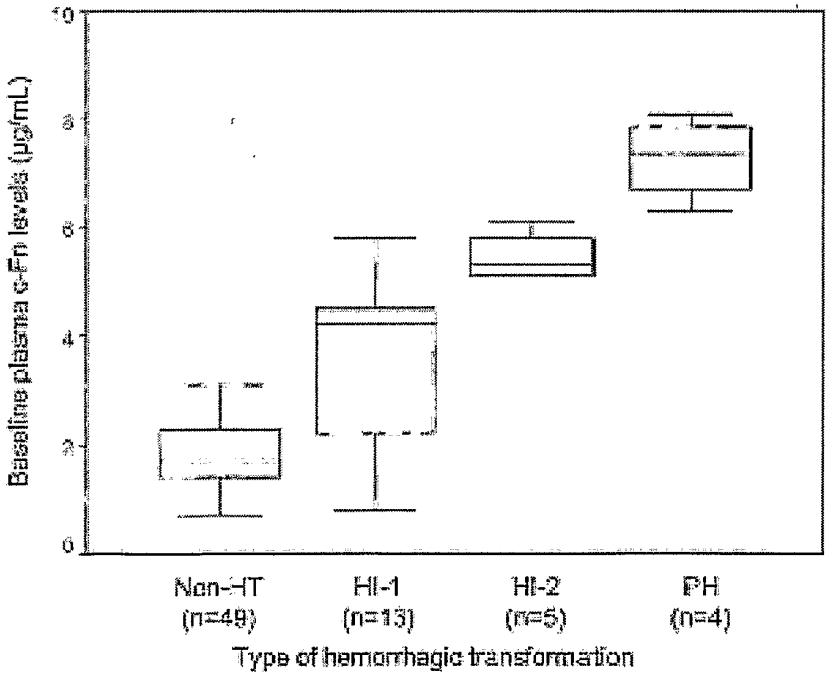

Baseline plasma c-Fn levels (µg/mL)
0
2
4
6
8
10
Non-HT (n=49)
HI-1 (n=13)
HI-2 (n=5)
PH (n=4)
Type of hemorrhagic transformation

FIGURE 3

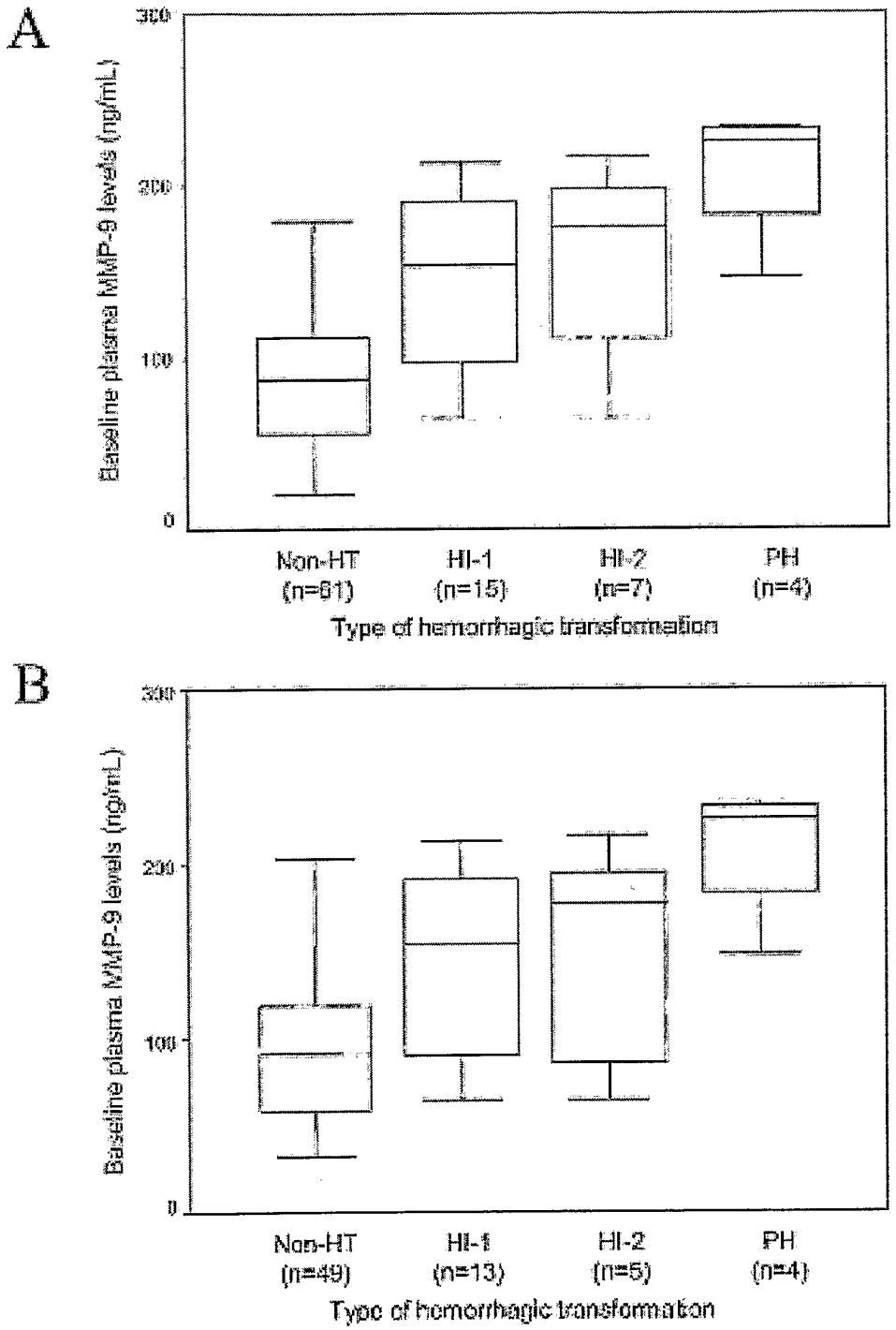

A
300
200
100
0
Baseline plasma MMP-9 levels (ng/mL)
Non-HT (n=61)
HI-1 (n=15)
HI-2 (n=7)
PH (n=4)
Type of hemorrhagic transformation
B
300
200
100
0
Baseline plasma MMP-9 levels (ng/mL)
Non-HT (n=49)
HI-1 (n=13)
HI-2 (n=5)
PH (n=4)
Type of hemorrhagic transformation

FIGURE 4

Cross validated receiver operator curve (ROC) time from onset subset comparison

	IL-6	TNF	MMP-9	FBN (CFN)
N	80	80	87	86
N REL_HT	10	10	10	9
p	0.0064	0.0088	4.30E-04	8.94E-07

Kruskalwallis ANOVA summary. Complete marker data for each marker only. N: Total number of cases with complete marker data. N REL_HT Number of relevant HT cases. p: Significance.

FIGURE 5

Receiver operator curve (ROC) for discrimination of patients whom will undergo a relevant hemorrhagic transformation following administration of TPA based on c-Fn

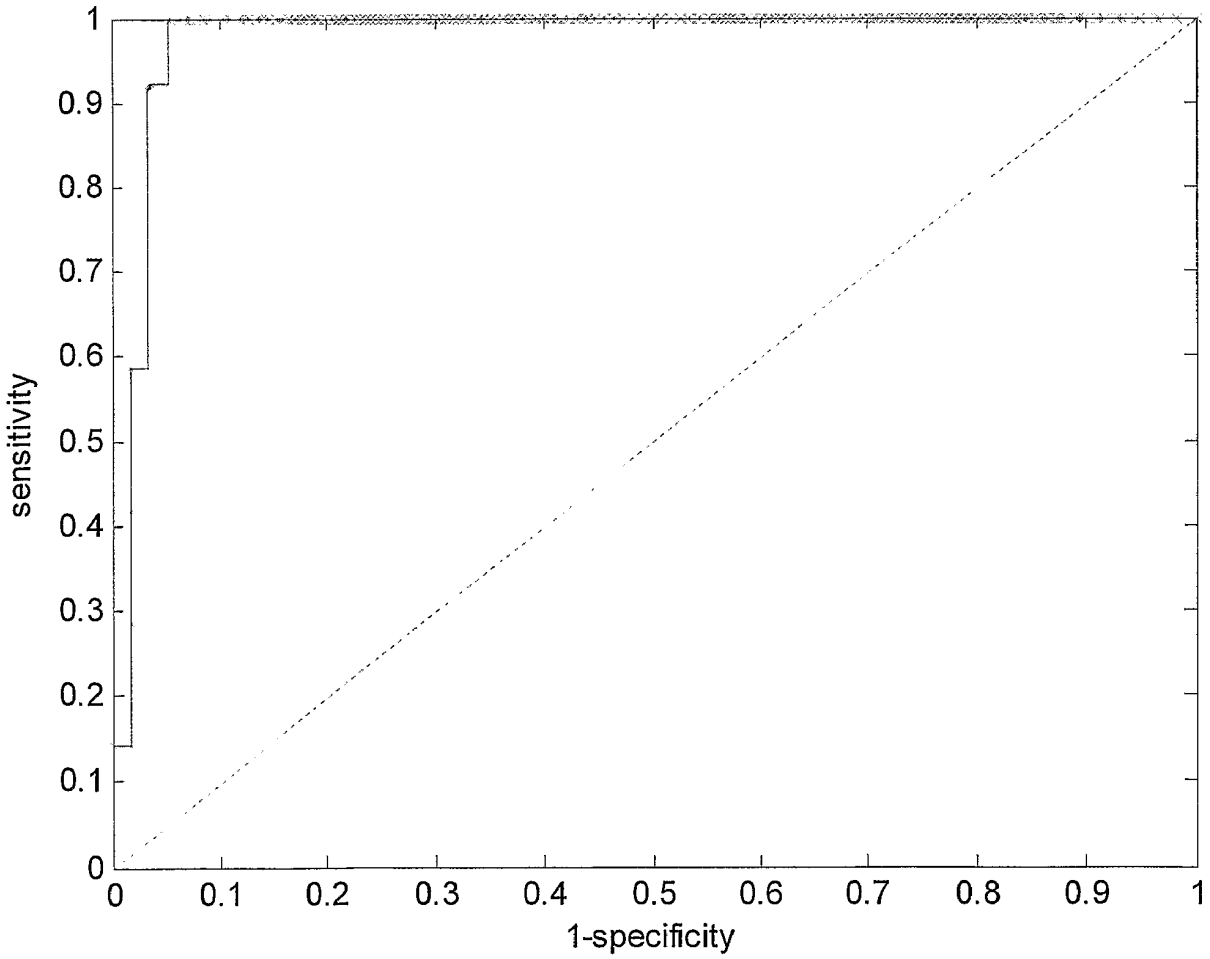


FIGURE 6

Receiver operator curve (ROC) for discrimination of patients whom will undergo a relevant hemorrhagic transformation following administration of TPA based on MMP-9

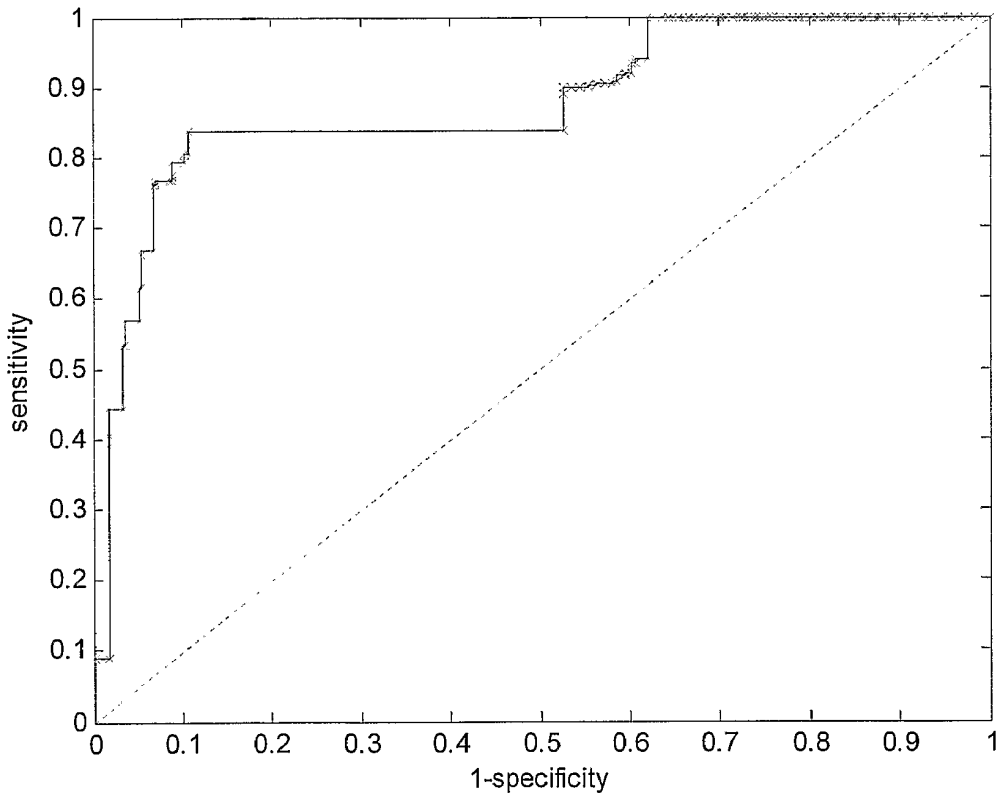


FIGURE 7

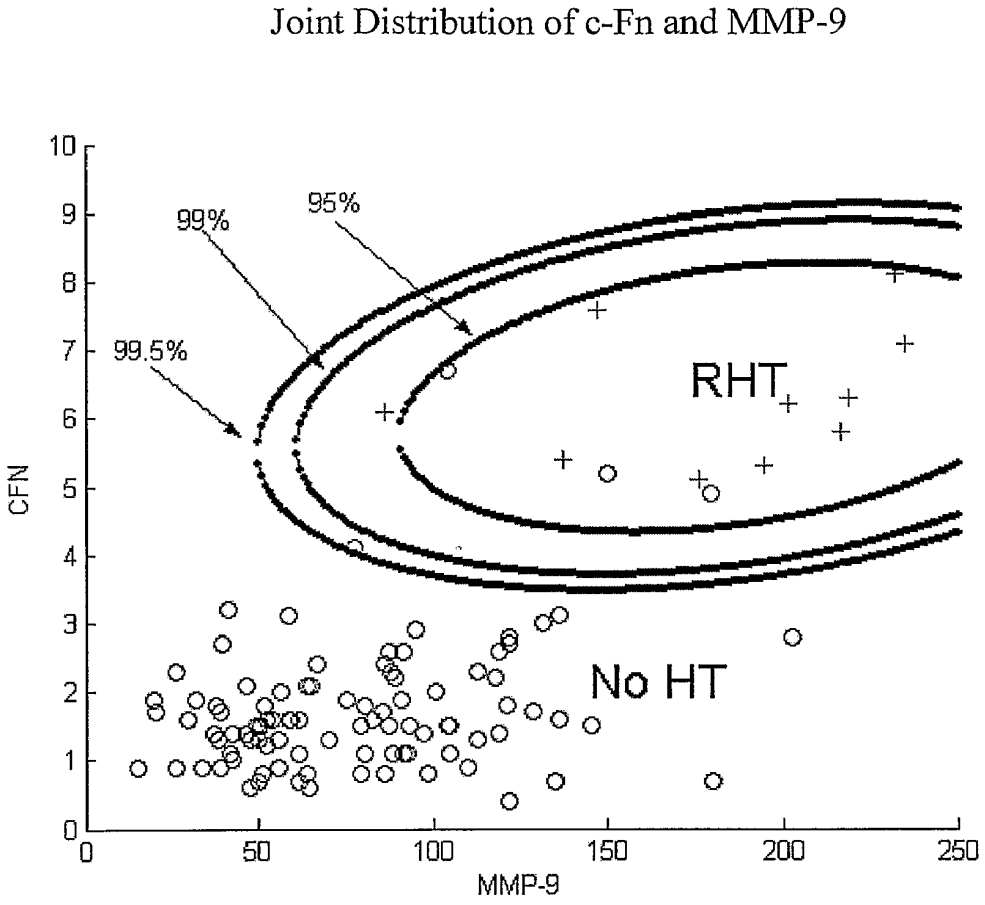
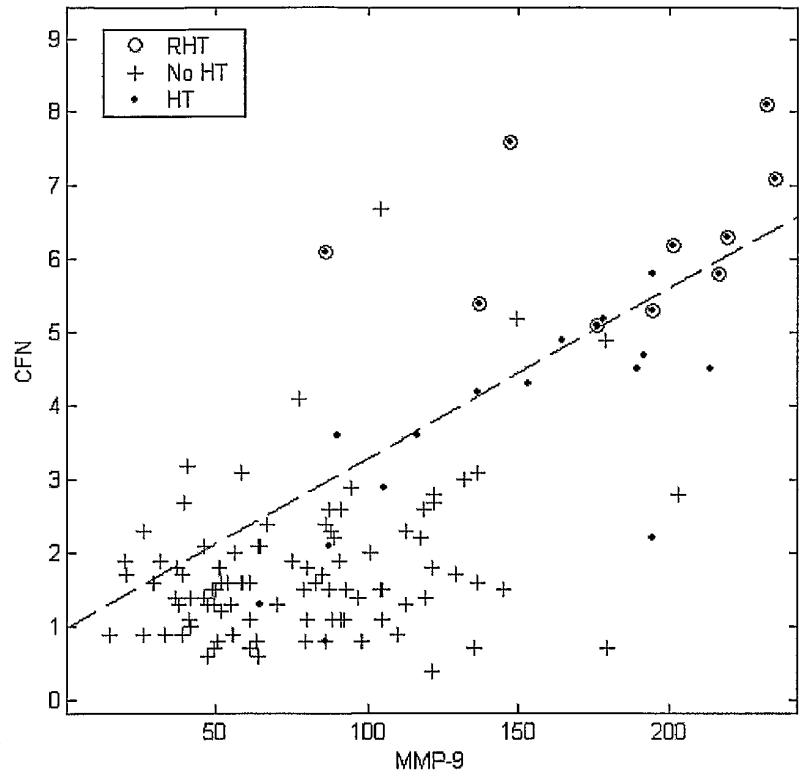

Joint Distribution of c-Fn and MMP-9
99%
95%
99.5%
RHT+
No HT
CFN
MMP-9
0
1
2
3
4
5
6
7
8
9
10
0
50
100
150
200
250

FIGURE 8

Graph of relationship between MMP-9 and CFN for subjects with an observed HT



专利名称(译)	细胞纤连蛋白作为中风的诊断标志物及其使用方法		
公开(公告)号	EP1792178A4	公开(公告)日	2007-10-03
申请号	EP2005778919	申请日	2005-05-18
[标]申请(专利权)人(译)	SCI预测		
发明人	CASTELLANO, MAR, HOSPITAL UNI. DE GIRONA DAVALOS, ANTONI, HOSPITAL UNI. DE GIRONA DIAMOND, CORNELIUS, ALLEN, PREDICTION SCIENCES LLC		
IPC分类号	G01N33/48 G01N33/50 G01N33/53 G06F19/00 G01N33/68 G06F19/18 G06F19/24		
CPC分类号	G01N33/6893 G01N33/6848 G01N2333/78 G01N2800/2871 G16B20/00 G16B40/00 Y02A90/26		
优先权	11/046592 2005-01-29 US 10/948834 2004-09-22 US		
其他公开文献	EP1792178A2		
外部链接	Espacenet		

摘要(译)

本发明涉及用于诊断和评估中风和中风亚型的方法。公开了多种生物标记物，用于组装用于这种诊断和评估的面板。公开了用于选择标记物并将它们的组合水平与感兴趣的临床结果相关联的方法。在各个方面：本发明提供了用于早期检测和分化中风亚型的方法，用于确定患有中风症状的患者的预后，以及鉴定在溶栓治疗后具有出血性转化风险的患者。公开了提供快速，灵敏和特异性测定的方法，以大大增加可以接受有益中风治疗和治疗的患者数量，并降低与错误中风诊断相关的成本。

Sensitivity for RHT	MMP-9 Minimum Expected Value	CFN threshold	%FP
95%	105.5	3.42965	4.21%
99%	72.83	2.668439	13.68%
99.5%	60.88	2.390004	18.95%
99.9%	36.21	1.815193	33.68%