Principles of precision medicine in stroke

Jason D Hinman,1 Natalia S Rost,2 Thomas W Leung,3 Joan Montaner,4 Keith W Muir,5 Scott Brown,6 Juan F Arenillas,7 Edward Feldmann,8 David S Liebeskind1

ABSTRACT

The era of precision medicine has arrived and conveys tremendous potential, particularly for stroke neurology. The diagnosis of stroke, its underlying aetiology, theranostic strategies, recurrence risk and path to recovery are populated by a series of highly individualised questions. Moreover, the phenotypic complexity of a clinical diagnosis of stroke makes a simple genetic risk assessment only partially informative on an individual basis. The guiding principles of precision medicine in stroke underscore the need to identify, value, organise and analyse the multitude of variables obtained from each individual to generate a precise approach to optimise cerebrovascular health. Existing data may be leveraged with novel technologies, informatics and practical clinical paradigms to apply these principles in stroke and realise the promise of precision medicine. Importantly, precision medicine in stroke will only be realised once efforts to collect, value and synthesise the wealth of data collected in clinical trials and routine care starts. Stroke theranostics, the ultimate vision of synchronising tailored therapeutic strategies based on specific diagnostic data, demand cerebrovascular expertise on big data approaches to clinically relevant paradigms. This review considers such challenges and delineates the principles on a roadmap for rational application of precision medicine to stroke and cerebrovascular health.

Precision medicine is broadly defined as tailored diagnosis and therapy for each individual patient.1 Driven by an explosion in genomic sequencing capabilities together with simultaneous revolutions in computing capabilities, recent years have seen the initial application of this approach take root in the practice of clinical medicine, though almost exclusively in the field of oncology. Recent nationwide efforts2 to expand precision medicine to many other areas of medicine have been promoted, including stroke.3–8 The meaning of the application of precision medicine to a phenotypically diverse disease such as stroke, however, is far more complicated than the comparatively simple genomic identification of the disease-causing mutation in a patient with cancer. Stroke is a basket term that historically includes multiple distinct disorders that acutely alter neurovascular tree, unique pathophysiology and likewise span both acute and chronic time points. For example, stroke due to cerebral microvascular disease demonstrates considerable overlap with neurocognitive disorders.7,8 Such phenotypic variation poses several challenges to the application of precision stroke medicine. Each of these phenotypic variations is associated with different types of input data that need to be appropriately and intelligently integrated to carry value in tailoring therapy during the acute, recovery and prevention phases of cerebrovascular disease. These challenges underscore the value and importance of standardising clinical, biological (genomic, serological and novel) and imaging data collection, much of which is already collected during focused, hypothesis-driven group-wise clinical trials yet not readily used. Moreover, true stroke precision medicine will require experienced and rational valuation of these multiple data points and appropriate statistical methodology to harness their collective potential for individualised decision-making. Stroke theranostics, the ultimate vision of synchronising tailored therapeutic strategies based on specific diagnostic data, demand cerebrovascular expertise on big data approaches to clinically relevant paradigms. This review considers such challenges and delineates the principles on a roadmap for rational application of precision medicine to stroke and cerebrovascular health.

CLINICAL DATA

Standardised methods

Precision diagnosis and treatment require big data in numerous dimensions, considerable depth, extensive volume and high quality. Past clinical trials, however, collected a negligible fraction of the vast array of variables established in the NIH-defined common data elements (CDE). The development of CDE for inclusion in a given clinical trial and collection for future study provides a standard language to define what each variable reflects.9 Precision stroke medicine requires rigorous standardisation, definition of variables, consistent data collection, reliable central adjudication and, where appropriate, quantification.10 Clinical variables require particularly meticulous definition to their preserve future value, as they may be identified by a wide spectrum of criteria, from spoken history to complex physiological measurements. Standardisation of clinical data elements permits data analysis across trials; the larger sample size creates power to study subgroups, which cannot be explored in single trials.

Challenges in data collection

A precision approach requires that data acquisition and analysis be made a central feature of trial design, requiring dedicated infrastructure independent of the focus on the diagnostic method or treatment being studied. A potential foundation for such a sustainable and effective research model would include StrokeNet, the NINDS stroke trials...
The breadth of data elements to be potentially collected in trial design has exploded, from a basic set of clinical variables to addition of imaging, physiological and genomic data. Even when specific markers or detailed data sets are of unproven value, source data collection must be structured within the framework of a trial and made available for future study. An unfortunate example of missed opportunity might include stroke imaging studies which failed to discern accurate stroke subtypes or ischaemic infarct patterns, thresholds and volumes associated with distinct mechanisms.

The location or environment where data are collected is as essential a consideration as the standardisation of the data. Local collection of data elements in real time differs greatly from central adjudication of those variables. Consider an example from acute stroke imaging: the extent of early ischaemic changes on CT scan or the degree of reperfusion may vary considerably between local and central readings. Local, real-time readings by treating physicians enhance the generalisability of the data. Central adjudication by an imaging core laboratory provides reliability. Generalisability may also be increased by collecting information regarding excluded participants via screening logs and consideration of dropouts and withdrawal of care.

Temporal aspects of data ascertainment are also relevant. Specific variables may have different meaning depending on whether they were measured in the acute period of stroke onset or in the follow-up of an acute stroke case and timing must be precisely identified. Clinical trials may refer to ‘admission’ values for National Institutes of Health Stroke Scale (NIHSS), yet it is not always reliably documented when this takes place. Since such measurements change considerably over time, detecting any fluctuation in these parameters may be critical in a patient who suffered a stroke. Moreover, collection of data in prior trials presbytesophly centred around the timing of the investigational treatment, while ignoring prehospital, intensive care unit (ICU) or postdischarge periods up until the standard day 90 evaluation. This has limited the study of potentially critical treatment variables such as early haemodynamic management, ICU treatment or rehabilitation therapies that influence clinical outcomes in stroke. The use of serial, periprocedural multimodal imaging will increase, resulting in a huge number of variables. Long-term data storage also plays an essential role. Data archiving ensures maximal study interpretation and financial investment. Digital Imaging and Communications in Medicine imaging data sets and biological specimens, for example, could be tapped for decades. Collection of source imaging data sets allow for repeated analyses and has proven invaluable. Raw data captured today can be stored and reanalysed tomorrow given new software developments. A clinical trial network could archive routinely collected imaging data to keep pace with evolving imaging technology. Postprocessing techniques such as computational fluid dynamics are novel, have been successfully applied to older data sets, but are continuously improved on and replaced. Archived data have many lives. Consider the potential of a neuroimaging library, functioning as a repository of lesion atasling. Novel software can be tested in larger data sets and remote real-time viewing and analysis can be developed. Similarly, collection of biological samples such as blood constituents do not require knowledge of every molecular signature, as storage of samples may be tapped decades later to identify novel findings.

Clinical utility

The WASID trial of intracranial atherosclerosis illustrates the potential of a ‘data stored for future study’ approach. Posthoc analyses of collaterals on conventional angiography were superior to the degree of arterial stenosis for predicting patient outcomes. Subsequent analyses also demonstrated that fractional flow measures were useful in predicting outcomes. In SAMMPRIS, the centrally collected imaging data sets provided the basis to validate the role of collaterals pioneered in the WASID trial. Thus, previous negative trial data in stroke provided fertile ground to examine subgroups from an imaging perspective, since the data were collected and stored. The role of breast-cancer gene-expression products in prognosis developed as a result of studies of preserved biological tissues. Later retrospective studies of randomised clinical trials suggested little benefit of chemotherapy in women having low recurrence risk. Now, prospective randomised trials study the effects of adjuvant chemotherapy stratified by predicted risk.

BIOLOGICAL DATA

Biological data—genomics

Genetics and genomics fuelled the personalised medicine concept. For the first time in decades since the seminal deciphering of DNA structure, the reality of gene therapy and genome editing for monogenic disorders is closer than ever. However, stroke as a phenotype is considered to be a complex genetic trait, that is, influenced collectively by numerous genetic variants with individually small effect sizes and the search for specific stroke genes remains a work in progress. There are very few stroke syndromes that are considered ‘genetic’; in fact, while stroke is known to run in families, a component of shared environment is likely to modify the heritability contribution seen in twin and family history studies.

Genetic contribution is probably stronger in younger individuals with stroke, as well as in those with haemorrhagic stroke subtypes and non-cardioembolic (non-CE) ischaemic stroke; furthermore, there is also suggested shared genetic contribution between large artery atherosclerosis (LAA) and small vessel ischaemic stroke subtypes as shown recently in large-scale genetic studies of patients with ischaemic stroke.

Diagnosis of stroke often involves a combination of clinical formulation (based on the semiology and timeline of a presenting syndrome), results of neurological examination and neuroimaging. Current challenges to the role of genomics in clinical stroke diagnoses are imposed by the rapid workflow of acute stroke therapy and the untouched, yet unmet need for personalised secondary stroke prevention. Gene expression profiling, a genomic approach based on interrogation of RNA molecules circulating in peripheral blood has demonstrated early promise in differentiation of ischaemic stroke from acute non-stroke pathologies (such as myocardial infarction and seizures), as well as intracranial haemorrhage and transient ischaemic attack (TIA). Furthermore, specific genome-wide RNA signatures might exist for individual ischaemic stroke subtypes such as CE versus LAA, and lacunar versus non-lacunar strokes. While gene expression profiling has not yet been incorporated into current clinical routine in the era of hyperacute diagnosis, the rationale for using rapid, accurate and inexpensive point-of-care testing to improve precision of stroke subtyping is compelling. Beyond the acute phase, accurate determination of stroke aetiology is an important step toward therapeutic management, personalised risk stratification and formulation of a secondary stroke prevention strategy for each individual patient.

Once a diagnosis of acute stroke is established, treatment options and outcome prediction become a priority. For stroke syndromes associated with well-established underlying genetic disorders such as cerebral autosomal dominant arteriopathy...
with subcortical infarcts and leukoencephalopathy (CADASIL), mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) or Fabry disease (FD), targeted genetic counselling (CADASIL), disease-specific interventions (MELAS) or enzyme-replacement therapy (FD) now exist.\textsuperscript{15} \textsuperscript{16} \textsuperscript{19} In patients with lobar intracranial haemorrhage, presence of APOE gene mutation (ε2 or ε4 allele) suggests an increased risk of recurrent haematoma and may guide personalised selection of antithrombotics.\textsuperscript{16} For individuals with Marfan-like features, stroke and elevated homocysteine, mutational screening in the CBS gene may help identify homocystinuria and guide therapeutic approaches.\textsuperscript{15} 

Beyond monogenic syndromes, there is a growing field of pharmacogenetics, that is, the study of genetic variation in drug response, which is applicable to large populations of patients who suffered a stroke. In an era of targeted management of vascular risk factors, individual response to antithrombotics, statins and even drugs commonly used to control blood pressure and diabetes mellitus is often tested through trial and error, rather than theranostics.\textsuperscript{21} There is evidence that a large proportion in response variability to warfarin may be explained by common genetic variants in VKORC1 and CYP2C9 genes, prompting specific Food and Drug Administration (FDA) dose recommendations. Similarly, the CYP2C19*2 variant was associated with reduced concentrations of clopidogrel in blood and increased risk of cardiovascular events in its users, prompting the FDA request for a black-box warning regarding reduced effectiveness in poor metabolisers. For the direct thrombin inhibitor dabigatran, a polymorphism CES1 (rs2244613) was noted to reduce drug concentration by 15% and decrease risk of bleeding by 33%.\textsuperscript{22} A number of genetic variants have also been identified in association with individual patient response to other stroke-specific classes of drugs (such as statins or antihypertensive agents), supporting the need for precision stroke medicine and theranostic development.\textsuperscript{21} 

In acute stroke, evidence of genetic influence on individual response to alteplase is emerging, with variants in A2M (rs669) and F12 (rs1801020) being linked to haemorrhagic transformation and in-hospital death.\textsuperscript{23} Currently, genetic architecture of post-stroke outcomes is poorly understood; however, translational studies of genetic variability in cerebral murine collateral circulation may prove important in understanding pathophysiology of collateral response in acute ischaemic stroke.\textsuperscript{24} Similarly, the ongoing investigations of leukoaraiosis, a highly-heritable neuroimaging trait associated with increased risk of stroke and poor post-stroke outcomes hold promise for insight into mechanisms of brain susceptibility to acute ischaemia.\textsuperscript{25} 

Pursuing complex genetic underpinnings of stroke across the spectrum of clinical applications in precision medicine (ie, diagnosis, treatment, prognosis and prevention) will require large-scale, high-throughput studies that are based on thorough phenotype ascertainment and cutting-edge methodology including genome-wide genotyping, whole-genome (and/or exome) sequencing and gene expression studies. In addition, systematic strategies to translate genetic discovery into clinically relevant applications are urgently needed. If fully integrated, a personalised approach in stroke incorporating genetics and genomics may improve precision diagnosis; streamline treatment options that are safe and effective; and boost the predictive ability of risk and outcome models in stroke.

**Biological data—serological**

Blood biomarkers have been used for many years in other vascular conditions, such as troponins for myocardial infarction diagnosis, glycosylated haemoglobin for diabetes monitoring or low-density lipoprotein cholesterol level for starting lipid-lowering therapies, yet no biomarker is routinely used in stroke management either for diagnostic, prognostic or risk-stratification purposes. Recently, there has been great interest in the potential of serological markers of neurovascular disorders, from predicting stroke risk among healthy individuals to establishing a biochemical diagnosis of stroke or in using biomarkers as tools for outcome prediction.

Most surrogate biomarkers in stroke are focused on underlying pathophysiology, such as inflammation and immune system response, apoptosis, coagulation and fibrinolysis, tissue remodeling or heart damage. Recently, enthusiasm has mounted for high-throughput techniques based on a large screening process, avoiding selection bias and generating extensive lists of molecules to be evaluated. Among the new -omics techniques, proteomic studies (the study of the entire set of proteins expressed by a genome under a particular situation) are still the most promising tools for discovering new biomarkers, as proteins are the most widely used biomarkers and easily detectable in quick point-of-care devices. In fact, due to the complexity of the nervous system, proteomic tools have been invaluable in terms of biomarker and disease target discovery directly in the brain.\textsuperscript{26} However, taking into account that blood is employed as the most suitable biological sample, a whole proteome approach could be addressed with mass spectrometry-based techniques thus covering a wide number of proteins directly in the serum or plasma. Alternative proteomic techniques are required for the discovery of small specific indicators of disease, as identification of clinically significant biomarkers is challenged by the presence of tens of thousands of circulating proteins and differences in protein abundance ranging up to 11 orders of magnitude. Antibody libraries and protein arrays are promising but still expensive to conduct in large studies.\textsuperscript{27} 

Precision stroke medicine may be enhanced with blood biomarkers for predicting stroke risk, complementing subtype diagnosis and refining prognosis. Key priorities include prediction of stroke risk among healthy people, among those with vascular risk factors and identification of silent ischaemic brain lesions; diagnosis of stroke and TIA; differentiation between stroke subtypes (ischaemic vs haemorrhagic), between ischaemic stroke aetiologies; evaluation of response to reperfusion therapies; prediction of the global prognosis of stroke and of specific complications of stroke; and evaluation of secondary prevention strategies (see http://stroke-biomarkers.com/).

FDA has approved the use of lipoprotein-associated phospholipase A2 (PLAC test) to predict risk of stroke developing among healthy individuals. Lp-PLA 2 is a serine lipase, which circulates mainly linked to low-density lipoproteins known to be a good predictor of both first-ever and recurrent strokes in population-based studies. Such strategies will allow increasing vascular prevention strategies among the highest risk people.

Rapid biochemical diagnosis that would help distinguish stroke mimics may streamline triage of the right patients to stroke centres. The idea of identifying the ‘troponins’ of stroke has encouraged several groups and biotechnology companies to initiate such discovery programmes, yet a robust, sensitive and rapid blood test is still lacking. Similarly, a prehospital or bedside rapid diagnostic test to differentiate ischaemic and haemorrhagic stroke is desirable. Rapid differentiation may enable initiation of tailored prehospital therapies, such as blood pressure lowering in intracranial haemorrhage (ICH), rapid reversal of anticoagulation in the presence of warfarin-associated ICH or delivering tPA to patients who suffered an ischaemic stroke in the field. Although
prehospital acute stroke treatment is becoming reality.\textsuperscript{28} Implementation of resource-intensive approaches with a portable CT scanner are difficult and blood biomarkers may play a role, especially in low-income countries. Examples of such biomarkers include glial fibrillary acidic protein (GFAP), or RBP4 that is as a promising biomarker to distinguish ischemia from hemorrhage when combined with GFAP and others.\textsuperscript{29}

Serological biomarkers may refine determination of stroke aetiology and classification, as many patients currently receive a diagnosis of undetermined stroke. For example, increased blood levels of natriuretic peptides (B-type natriuretic peptide or N-terminal pro-BNP) have been repeatedly associated with CE stroke.\textsuperscript{30} In the future, the use of blood biomarkers will allow clinicians to rapidly guide other diagnostic tests and accelerate initiation of tailored secondary prevention.

Blood biomarkers may clarify extreme variations in stroke prognosis, commonly influenced by numerous factors such as age, stroke severity, comorbidities and complications. Early prediction of complications, such as cerebral oedema, haemorrhagic transformation or post-stroke infection represents a unique opportunity to improve stroke outcome. Early detection seems feasible by combining clinical, imaging and biological data (such as interleukin6 (IL6), copeptin, endostatin) but advancing such algorithms in routine clinical practice requires proper data integration strategies and further research. Furthermore, blood biomarker levels should be linked not just with general measures of prognosis such as disability or death, but with specific causes of worsening that culminate in poor outcome. Several studies have focused on biomarkers of stroke-associated infections such as IL-6, the Th1/Th2 profile, copeptin, procalcitonin or the recently described serum amyloid A.\textsuperscript{31} However, the accuracy, sensitivity and specificity of all these biomarkers are still not adequate for clinical use. Moreover, lack of common criteria for the outcome end point (eg, post-stroke pneumonia), further thwarts the ability to identify proper blood biomarkers.\textsuperscript{32} Even before clinical implementation, semi-interventional trials are necessary to link a biomarker with a specific intervention, such as the STROKE Adverse outcome is associated With NoScomial Infections trial,\textsuperscript{33} which is evaluating the use of a biomarker- guided (ultrasensitive procalcitonin) strategy in the indication of preventive antibiotic therapy in patients with severe stroke.

Theranostics of blood biomarkers beyond genetic propensity noted above may be used to refine treatment response, as patients respond differently to thrombolysis and mechanical reperfusion strategies. Serological biomarkers may selectively evaluate tPA response, prediction of haemorrhagic transformation and identify patients with unsuccessful tPA-induced rencanalisation. In terms of haemorrhage, fibronectin, MMP-9 and VAP-1 are considered potential biomarkers,\textsuperscript{34,35} yet a large cohort study that combines such measures is lacking.

Implementation of serological biomarkers in routine clinical practice remains a major challenge. A limiting factor is the extended period required for the validation and commercialisation of a discovered biomarker. The search for novel stroke biomarkers should therefore leverage other data types to accelerate development.

**Novel biological data**

Beyond the obvious potential for added diagnostic and prognostic value afforded by genomic and serological data, additional biological specimens collected from patients may add precision to the diagnosis of individual cases. These range from routine analysis of cerebrospinal fluid and brain biopsy samples in selected cases\textsuperscript{36} to samples uniquely available from patients who suffered a stroke such as histological and proteomic analysis of clots retrieved during thrombectomy\textsuperscript{37} and cerebrovascular endothelial cells that can be captured from stent retrievers after use in thrombectomy.\textsuperscript{38} Though presently these additive biological variables are not routinely used, they hold the potential to unlock additional diagnostic, prognostic and theranostic value.

**IMAGING DATA**

Neuroimaging of the brain and related vascular structures remains the foundation for the determination of stroke aetiology. Imaging data are routinely acquired in every potential patient who suffered a stroke, providing a wealth of data on cerebrovascular health and evolution of active pathophysiology. The phenotype of stroke, reflecting contributions from multiple genes and additional influences such as lifestyle and environmental factors, cannot be ascertained without imaging. While studies on stroke genomics are still at a rudimentary stage of searching for susceptible loci or genes associated with each subtype of stroke, one pre-requisite for the success of these genetic studies is a ‘precise’ stroke aetiology or subtype. Single nucleotide polymorphisms or variants that each contribute modestly to the phenotype would remain obscure if the stroke subtype classification is rough or imprecise. In fact, imprecise stroke subtyping has already led to a large number-needed-to-treat in conventional stroke treatment studies when a minority of study participants drove the cohort effect. LAA, the most common stroke aetiology around the world, provides such an example. Based on TOAST classification, cerebral ischaemia attributed to atherosclerosis of either the extracranial internal carotid artery or intracranial arteries have been collectively grouped as LAA. However, LAA does not focus on the location and functional impact of atherosclerosis, a critical factor in governing prognosis and related treatments. In fact, based on the site of symptomatic lesion, LAA strokes broadly fall into two relatively distinct phenotypes, extracranial carotid artery disease (ECAD) and intracranial atherosclerotic disease (ICAD). Compared with ECAD, ICAD has a significantly higher rate of early relapse. Moreover, ICAD and ECAD respond differently to the same treatment strategy. For instance, angioplasty/stenting proven effective for ECAD was found even inferior to medical treatment for ICAD. Epidemiologically, the diverging prevalence of ECAD versus ICAD in Caucasians, Asians, Hispanics and Blacks strongly suggests a genetic contribution and a distinction between these entities currently lumped as LAA. If a genomic investigation on LAA indiscriminately recruits both patients with ICAD and ECAD based on TOAST classification, it will dilute any potential power in detecting an associated variant. Hence, TOAST is not ‘precise’ enough for current clinical or research purposes. However, would the Causative Classification of Stroke (CCS), an improved version of TOAST classification, be precise enough for risk prediction or treatment allocation? CCS now permits distinguishing patients with symptomatic ICAD from patients with symptomatic ECAD. Nevertheless, ICAD can cause ischaemic strokes through multiple stroke mechanisms: artery-to-artery thromboembolism from unstable plaques; occlusion of penetrating artery ostium by a clot or atheroma; Venturi effect causing hypoperfusion of perforators; high-grade steno-occlusion leading to cortical ischaemia; or a combination of these (eg, thromboembolism with impaired washout at border zones). Without understanding the ‘precise’ stroke mechanism in each individual patient with ICAD, treatments that have been designed with a ‘one-size-fits-all’ approach may be useful for some patients but not helpful or even harmful for others. For example, an patient with ICAD with ipsilateral hemispheric...
Cerebrovascular disease

hypoperfusion may benefit from angioplasty of middle cerebral artery stenosis; but the same endovascular procedure may pose a significant risk of perforator stroke if one does not appreciate the concurrent presence of branch atheromatous disease adjacent to the stenosis.

The dynamic nature of either intracranial or extracranial atherosclerotic plaques further complicates precise strategies. An apparent high-grade stenosis at the acute stage may regress significantly after stringent control of cardiovascular risks; yet some lesions may contrarily continue to progress, despite treatment. Hence, the current approach of relying on a single scan or imaging snapshot of a vascular lesion is inadequate to portray phenotype. Serial imaging data, or repeated evaluations over time, that reflect the temporal nature of cerebrovascular disorders may elucidate the homeodynamics that govern the evolution of collaterals, cerebral autoregulation and infarct expansion. Similar ‘imprecision’ or ambiguity in diagnosis may also occur in other TOAST subtypes. For example, a small artery-to-artery thromboembolic infarct from LAA may be classified as a lacunar infarct or small vessel disease based on TOAST if vascular imaging is inadequate.

Brain imaging has the capacity to identify viable tissue and thus select patients most likely to benefit from interventions, especially reperfusion. The concept was demonstrated using positron emission tomography in the 1990s, when clinically similar acute patients were found to fall into three groups based on imaging findings. One group with severely reduced blood flow accompanied by failure of use of oxygen had uniformly poor outcome; one with preserved blood flow and use of oxygen had uniformly good outcome; and the third group, with variable outcome, had hypoperfusion but preserved metabolic capacity. This last group was believed to demonstrate imaging features of the ischaemic penumbra, and to represent a suitable target group for interventions such as reperfusion therapies. Subsequent developments of technology have translated this concept into more usable technologies, primarily CT perfusion (CTP) and MRI.

It is generally accepted that tissues with probability of infarction regardless of reperfusion, the ischaemic core, are denoted by the diffusion-weighted MRI (DWI) lesion, and by significantly reduced cerebral blood flow (CBF) or cerebral blood volume (CBV) on CT perfusion. The core is likely also represented by visible hypodensity on non-contrast CT (NCCT), although with much lower sensitivity compared to DWI or CTP. Potentially viable penumbra is defined by hypoperfusion, identified as tissue with prolonged contrast transit (but normal DWI signal on MRI, ‘diffusion-perfusion mismatch,’ or normal CBF or CBV on CTP), various definitions employing different time-based metrics including mean transit time, time to peak, delay time or time to the maximal tissue contrast peak.

Patient selection by imaging increases clinical trial efficiency since it increases the likelihood of selecting a responder population (and is rooted in pathophysiology), excluding those with either no need for treatment or no likelihood of benefit. Calculation of hypothetical sample size for clinical trials shows substantial reductions in sample size with an imaging-selection approach. The penalty is of fewer eligible patients and slower recruitment. Observational data support the concept that reperfusion of patients with imaging selection is beneficial whereas reperfusion of patients lacking an imaging target is not; data also suggest possible increased risk of haemorrhagic transformation when reperfusion occurs in the face of extensive irreversible ischaemia (the ‘large core’ group with >70 mL core tissue).

Recent clinical trials deploying imaging selection for reperfusion treatment have shown larger benefits than similar trials without tissue imaging selection. Examples include tissue selection in a tenecteplase trial, showing superiority over alteplase, while a trial without tissue selection criteria did not demonstrate benefit, and the endovascular thrombectomy trials deploying perfusion selection having a larger treatment effect than those without. A strategy for penumbral imaging selection to maximise clinical trial efficiency needs to be distinguished from evidence of any given treatment being ineffective in those being excluded; indeed, the exclusion of patients from clinical trials may impede our understanding of clinical application of an intervention. Proportional benefits from intravenous thrombolysis or endovascular thrombectomy do not appear to differ among those with more extensive NCCT hypodensity or CTP-defined large core, within the constraints of fulfilling other trial eligibility criteria and small numbers being recruited at the severe end of the spectrum. Absolute benefits are lower and therefore there may be either a clinical or health economic case for limiting treatment in such patients. There are nonetheless further issues that should be addressed before basing a precision medicine approach on penumbral imaging. While observational studies of intravenous thrombolysis note better clinical outcomes among those with a ‘target mismatch’ profile on CTP compared with historical control participants treated on the basis of non-contrast CT alone, two pertinent factors require consideration. First, there is substantial disagreement between clinical interpretation of what constitutes an appropriate penumbral pattern and automated software analysis.

Software analyses are not standardised and competing systems are used in different studies. Second, the additional time incurred in acquisition, processing and review of additional perfusion imaging may delay treatment and since there is rapidly diminishing benefit from both intravenous thrombolysis and endovascular thrombectomy over short time intervals, such delay may be detrimental. The use of advanced multimodal imaging at local sites with central telestroke triage offers another example of how a precision medicine approach to imaging acquisition, data transfer, automated postprocessing and clinical decision-making tools may be used in the future.

STRATEGIES FOR DATA INTEGRATION

Data selected for precision medicine need not be strictly molecular or imaging based. Epidemiological precision medicine uses large data sets, such as administrative claims. The Observational Health Data Sciences and Informatics collaborative uses electronic health records and health insurance claims data. Diverse data sources and large sample sizes inherent in this approach have the potential to improve the derived prognostic models.

Using clinical expertise to value data

Success in precision stroke medicine requires that investigators advance from identifying risk factors that are merely associated with outcomes to clinical features rooted in pathophysiology that actually cause outcomes. For example, baseline stroke severity is a prominent determinant of outcome, while collateral circulation determines stroke severity and is based in the pathophysiology that determines the spectrum of stroke severity outcomes. It is not surprising that imaging collateral flow before treatment of ischaemic stroke has become increasingly important. Similarly, the spot sign, reflecting potential contrast extravasation on CT angiography (CTA), predicts early intracerebral haematoma growth and reflects underlying pathophysiology. Perfusion delays are increasingly analysed in acute stroke, but the context in which they occur likely affects their...
interpretation. In a population of patients who suffered acute stroke with intracranial atherosclerosis, perfusion delays identified on CTP or perfusion-weighted imaging are likely chronic, thus limiting their value.4

Presence versus degree
Statistical analyses and clinical interpretability often require a sacrifice of precision in that a data parameter be changed from a continuous variable and dichotomised into ‘present or absent’ or ‘high or low’.12 In acute stroke, despite decades of research, the concept of ischaemic penumbra continues to be clarified. The methods of data definition, collection, storage and analysis inherent in precision medicine must lead to large, high-quality data sets and new analytic methods if this tension is to be relieved.

Reference standards
Cardiologists put fractional flow reserve (FFR) on the clinical map with invasive, pressure sensitive wires. Over time, they advanced to CTA-based computational fluid dynamics to measure the same parameters. Neurologists need to be equally flexible in accepting non-invasive methods of data acquisition and advanced software methods in order to develop more practical approaches to investigation and new reference standards.1

If large data sets originating in multiple locations are to be gathered in precision medicine, it is clear that owing to variability in availability of multimodal CT and MRI, cross-modality interpretation of various neurovascular imaging features will also be important.10

Statistics, incremental value and shared decision-making
While data are increasingly collected in precision medicine with new technologies, their introduction into the clinical arena still follows a common process. Investigators still ask whether a test reliably measures what it seeks to measure. Analyses of large data sets or randomised trials are performed to characterise clinical utility. Then, new tests are scrutinised for their incremental value: what do they add to previously used decision tools? This explosion of data will not be easy for scientists or patients to absorb and use. New approaches to data presentation, risk quantification and communication of uncertainty are urgently needed.12 It is becoming increasingly clear that the patient’s perspective must be accounted for if the promise of precision medicine is to bear practical fruit.

Embedding data collection into randomised controlled trial design
Integrating the principles of precision medicine with the logistics of clinical trials for stroke will require a shift in philosophy from the broad—not to the narrow, but to the precise. Large trials in stroke, especially those conducted with the objective of regulatory approval, typically focus on comparisons of performance between treatments (or between treatment and no treatment), while including the widest possible population in which an advantage is expected to be observed with the investigational therapy. Such trials often enrol a heterogeneous population in which the observed treatment effect is really not a single effect but an average across patients who may have benefitted greatly and others who benefitted less, or not at all, from the investigational therapy. This paradigm is understandable and sometimes necessary—for instance, to justify the expense of a trial conducted by a commercial entity, the greatest return on investment will be sought and the broadest plausible study population will therefore be enrolled. However, this runs counter to the objective of personalising medicine. In a large trial, subgroup analyses may be feasible to evaluate risk and benefit, but even when the sample size is sufficient to permit meaningful subgroups, differences in response among those subgroups are often not of interest to the trial sponsor. Indeed, in a commercial setting the sponsor is actually discouraged from thoroughly seeking such differences lest a subgroup with poor response be found: the result of such a finding is likely to be a narrowing of the resulting regulatory label. Few studies are therefore appropriately statistically powered to investigate populations below the mass of ‘all enrolled subjects’. The most useful change in focus is not simply to narrow the population under study. A clinical trial focused on a smaller, more homogeneous population may be able to answer specific questions about treatment outcomes impossible to find in broader investigations, but it cannot effectively extrapolate beyond its inclusion criteria. For instance, a trial enrolling only those patients for whom tPA is unavailable as an option, or only those patients who are treated particularly quickly after stroke onset, by definition cannot offer reliable evidence about outcomes in other cohorts of patients. From the perspective of precision medicine, one type of desirable study design is therefore one which plans enrollment of a broad population, and also collects data which clearly identify potentially useful phenotypes and which prospectively mandates analyses designed to distinguish outcomes in one subgroup of patients from another. Such a trial must necessarily be exploratory in nature, since if only thoroughly vetted hypotheses are tested, then truly novel, unplanned findings cannot arise.

Analysis techniques: statistics for precision stroke medicine
In order to embrace the exploratory nature of such trials, methods of analysis that make the greatest use of such data must be considered. As noted above, statistical methods in stroke trials commonly focus on distinguishing differences in outcome between alternative treatments, with the p value comparing groups deemed paramount. In the past, analysis has often started with dichotomisation of multivariate or continuous end points, such as the compression of the seven-point modified Rankin scale to a binary outcome of success or failure based on a cutpoint. Such approaches are intuitively appealing and comparatively easy to summarise, but they discard information in favour of simplicity. Precision medicine will instead require analysis methods designed to retain maximum information and flexibility in pattern recognition. Initial moves towards the comprehensive over the intuitive have already been seen, including the use of ordinal or ‘shift’ analyses incorporating the full distribution of Rankin scores in recent trials of mechanical thrombectomy. To uncover new relationships, the data will have to speak for itself, and learning must also accumulate over time. For this purpose, the Bayesian analysis paradigm should be emphasised. While Bayesian methods are not new to clinical trials, including stroke trials commonly focus on distinguishing differences in outcome between alternative treatments, with the p value comparing groups deemed paramount. In the past, analysis has often started with dichotomisation of multivariate or continuous end points, such as the compression of the seven-point modified Rankin scale to a binary outcome of success or failure based on a cutpoint. Such approaches are intuitively appealing and comparatively easy to summarise, but they discard information in favour of simplicity. Precision medicine will instead require analysis methods designed to retain maximum information and flexibility in pattern recognition. Initial moves towards the comprehensive over the intuitive have already been seen, including the use of ordinal or ‘shift’ analyses incorporating the full distribution of Rankin scores in recent trials of mechanical thrombectomy. To uncover new relationships, the data will have to speak for itself, and learning must also accumulate over time. For this purpose, the Bayesian analysis paradigm should be emphasised. While Bayesian methods are not new to clinical trials, including stroke, they have often been used merely as a technique for reducing sample size, thereby reducing the burden of trial management. The true value of Bayesian analysis is in its ability to continuously adapt by incorporating new information, as opposed to the traditional, frequentist approach to data analysis, in which trials start without inferential information. Fully Bayesian methods also avoid the use of p values entirely, which is appealing in an environment in which hypotheses may initially be vague and where promising patterns of information should not be discarded because of a failure to achieve traditional statistical significance. Similarly, statistical methods which prioritise pattern recognition over formal hypothesis testing will play a prominent role: methods such as non-parametric analysis,
classification and regression trees and discriminant analysis can all be expected to appear more commonly than in past analysis approaches.

Pathophysiology as target
The overwhelming focus on treatment effect must shift to studying the influence of baseline pathophysiology in the precision medicine world. The results of MR-RESCUE show that outcomes may be better predicted by baseline imaging of pathophysiological states. Patients with a favourable penumbral pattern had improved outcomes, smaller infarct volumes and attenuated infarct growth, as compared with patients with a non-penumbral pattern, regardless of treatment assignment. Anatomical selection criteria have enjoyed a long history in stroke trials, especially the measurement of per cent stenosis. Emphasis on function rather than structure is likely to be more fruitful. A recent study of cerebrovascular FFR in WASID/SONIA revealed that an entire group of patients excluded from SAMMPRIS, those with 40–69% anatomic stenosis, have a higher stroke risk when baseline physiology rather than anatomy is used for patient selection. Yet, these patients have not been enrolled in aggressive clinical trials. Thus, novel imaging approaches may extend useful treatments to wider populations of patients. Anatomical selection may identify select patient groups believed to harbour similar risk but which turn out to have highly variable risk when selected by more precise functional measures available from imaging studies. Too often, imaging data are merely inclusion criteria—if all patients are enrolled because they harbour a specific imaging finding, which the trial cannot determine whether the imaging approach that detected that finding was beneficial, as there is no comparison group. The efficiency of a clinical trial can be increased if patients are randomised to different imaging approaches or results, prior to randomisation to treatment, as in MR-RESCUE or the cardiology studies of FFR: FAME and FAME II. Perhaps various therapies may be combined and studied in a clinical trial where the main axis is defined by the underlying pathophysiology. Redesigning clinical trials around baseline pathophysiology using an adaptive trial model may improve our knowledge of natural history and hence individualised treatment effect, rather than the tradition of testing one therapy across heterogeneous patterns of influential pathophysiology.

EXAMPLES OF FUTURE APPLICATIONS
Recent advances in three-dimensional printing and biofabrication techniques may make it theoretically possible to develop precise, imaging-derived representations of cerebral aneurysms, intracranial stenoses or the microvascular circulation ex vivo. This type of platform could easily be used to study how individualised genomic, serological and imaging variables described above might contribute to cerebrovascular occlusion on an individual level. With ready availability of such models and the increasingly faster ability to produce precise models that reflect individual patient characteristics, stroke precision medicine may become a reality. Imagine when a neurointerventionalist can approach each patient’s aneurysm with not only a forethought treatment plan, but a trial and error tested approach; or a stroke neurologist can advise a patient with an intracranial stenosis or diffuse white matter microvascular disease, more precisely what their risk of recurrence is on a specific drug regimen. By wisely incorporating the principles of precision medicine together with the already accumulated and yet to be realised patient variables, true stroke precision medicine becomes imminently achievable and is likely to reduce cost long-term.

Box 1 Bullet points
- Precision medicine in stroke has the potential to reach well beyond genomics to incorporate the multitude of clinical, biological and imaging data points now routinely collected but rarely analysed from each patient.
- Stroke precision medicine will require the establishment of data repositories to which clinical trial and community generated data points can be synthesised and available for individualised outcome analyses.
- Multidisciplinary approaches to the valuation of various data points, the refinement of statistical approaches to handle disparate weighting of variables, the incorporation of new data variables and the synthesis of individualised data into clinically usable end points are needed.
- Continued support for new technologies including computational flow dynamics, three-dimensional printing and direct from patient data modelling hold the potential to develop methodology for the individualised prediction of recurrence risk, medication effects, procedural outcomes and recovery of stroke.
- Adherence to these principles in precision medicine in stroke can help to guide the precision medicine revolution in the practice of stroke medicine.

CONCLUSION
Relatively concurrent explosions in data collection, genomics, computing and a multitude of advances in stroke research have prefaced the realisation of precision medicine in stroke. Implementing this concept in practice, however, will require a multidisciplinary effort to bring together experts in traditionally disparate disciplines within the stroke field. True stroke precision medicine will require accurate clinical phenotyping, genomic and proteomic expertise, spatial and temporal imaging analysis, incorporation of new biological variables and the statistical and experiential knowledge to help assign the appropriate weight to each variable (box 1). Though daunting, patients already think of medical decision-making in personal terms that increasingly demand individualised precision. There is nothing more powerful than being able to convey to a patient a precise diagnosis and treatment plan, never more so than in stroke when the outcome in question is not only mortality but physical and mental disability. Proper consideration of how one might apply the principles of precision medicine to stroke presages a new era of individualised therapy in stroke.

Author affiliations
1Department of Neurology, Neurovascular Imaging Research Core and the UCLA Stroke Center, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, USA
2Department of Neurology, Philip Kistler Stroke Research Center, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA
3Division of Neurology, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, Hong Kong
4Neurovascular Research Laboratory, Vall d’Hebron Research Institute (VHIR), Barcelona & IBIS Stroke Programme, Hospital Virgen Macarena-Rocio, Sevilla, Spain
5Institute of Neuroscience & Psychology, Glasgow, UK
6Altair Biostatistics, St. Louis Park, Minnesota, USA
7Stroke Unit, Department of Neurology and Medicine, Hospital Clínico Universitario, Universidad de Valladolid, Valladolid, Spain
8Baystate Medical Center, Springfield, Massachusetts, USA
