Design and rationale for examining neuroimaging genetics in ischemic stroke

The MRI-GENIE study

OPEN

Anne-Katrin Giese, MD
Markus D. Schirmer, PhD
Kathleen L. Donahue, BS
Lisa Cloonan, BA
Robert Irie, PhD
Stefan Winzeck, MS
Mark J.R.J. Bouts, PhD
Elissa C. McIntosh, MA
Steven J. Mocking, MSc
Adrian V. Dalca, PhD
Ramesh Sridharan, PhD
Huichun Xu, MD, PhD
Petrea Frid, MD
Eva Giralt-Steinhauer, MD, PhD
Lukas Holmegaard, MD
Jaume Roquer, MD, PhD
Johan Wasselius, MD, PhD
John W. Cole, MD
Patrick F. McArdle, PhD
Joseph P. Broderick, MD
Jordi Jimenez-Conde, MD, PhD
Christina Jern, MD, PhD
Brett M. Kissela, MD, MS
Dawn O. Kleindorfer, MD
Robin Lemmens, MD, PhD
Arne Lindgren, MD, PhD
James F. Meschia, MD
Tatjana Rundek, MD, PhD
Ralph L. Sacco, MD, MS
Reinhold Schmidt, MD
Pankaj Sharma, MD, PhD
Agnieszka Slowik, MD, PhD
Vincent Thijs, MD, PhD
Daniel Woo, MD, MS
Bradford B. Worrall, MD, MSc
Steven J. Kittner, MD, MPH
Braxton D. Mitchell, PhD, MPH

Author list continued on next page

ABSTRACT

Objective: To describe the design and rationale for the genetic analysis of acute and chronic cerebrovascular neuroimaging phenotypes detected on clinical MRI in patients with acute ischemic stroke (AIS) within the scope of the MRI-GENetics Interface Exploration (MRI-GENIE) study.

Methods: MRI-GENIE capitalizes on the existing infrastructure of the Stroke Genetics Network (SiGN). In total, 12 international SiGN sites contributed MRIs of 3,301 patients with AIS. Detailed clinical phenotyping with the web-based Causative Classification of Stroke (CCS) system and genome-wide genotyping data were available for all participants. Neuroimaging analyses include the manual and automated assessments of established MRI markers. A high-throughput MRI analysis pipeline for the automated assessment of cerebrovascular lesions on clinical scans will be developed in a subset of scans for both acute and chronic lesions, validated against gold standard, and applied to all available scans. The extracted neuroimaging phenotypes will improve characterization of acute and chronic cerebrovascular lesions in ischemic stroke, including CCS subtypes, and their effect on functional outcomes after stroke. Moreover, genetic testing will uncover variants associated with acute and chronic MRI manifestations of cerebrovascular disease.

Conclusions: The MRI-GENIE study aims to develop, validate, and distribute the MRI analysis platform for scans acquired as part of clinical care for patients with AIS, which will lead to (1) novel genetic discoveries in ischemic stroke, (2) strategies for personalized stroke risk assessment, and (3) personalized stroke outcome assessment. *Neurol Genet* 2017;3:e180; doi: 10.1212/NXG.0000000000000180

GLOSSARY

ADC = apparent diffusion coefficient; **AIS** = acute ischemic stroke; **CE** = cardioembolic; **CCS** = Causative Classification of Stroke; **CCSc** = causative CCS; **DICOM** = Digital Imaging and Communications in Medicine; **DWI** = diffusion-weighted imaging; **DWIv** = DWI volume; **FLAIR** = fluid-attenuated inversion recovery; **GISCOME** = Genetics of Ischemic Stroke Functional Outcome; **GWAS** = genome-wide association studies; **ICC** = intraclass correlation coefficient; **LAA** = large artery atherosclerosis; **MGH** = Massachusetts General Hospital; **MRI-GENIE** = MRI-GENetics Interface Exploration; **mRS** = modified Rankin Scale; **PHI** = protected health information; **QC** = quality control; **SAO** = small artery occlusion; **SiGN** = Stroke Genetics Network; **SNP** = single nucleotide polymorphism; **SWI** = susceptibility-weighted imaging; **TOAST** = Trial of Org 10172 Acute Stroke Treatment; **VLSM** = voxel-based lesion–symptom mapping; **WMHv** = white matter hyperintensity volume; **XNAT** = eXtensible Neuroimaging Archive Toolkit.

Genome-wide association studies (GWAS) have been instrumental in elucidating the genetics of complex vascular traits (ischemic stroke[1,2] and coronary artery disease[3,4]) and their risk factors (blood pressure,[5] atrial fibrillation,[6] hyperlipidemia,[7] and diabetes mellitus[8]). Despite recent advances in prevention and treatment, stroke remains a leading cause of adult neurologic disability and death in the United States and worldwide.[9] Recent GWAS have uncovered several risk loci for ischemic stroke and its subtypes,[10,11] specifically *PITX2* and *ZFHX3* for cardioembolic (CE) stroke,[11,12] *HDAC9*[11,12] and *TSPAN2*[11] for large artery stroke, and *ALDH2*[11] for small artery stroke. These results highlight the necessity for large-scale collaborations such as

*These authors contributed equally to the manuscript.

Author affiliations are provided at the end of the article.

Funding information and disclosures are provided at the end of the article. Go to Neurology.org/ng for full disclosure forms. The Article Processing Charge was funded by the NIH.

Jonathan Rosand, MD, MSc
Polina Golland, PhD
Ona Wu, PhD*
Natalia S. Rost, MD, MPH*

Correspondence to Dr. Rost: nrost@partners.org

Supplemental data at Neurology.org/ng

METASTROKE and Stroke Genetics Network (SiGN) to identify risk loci for these complex diseases.

Improving our understanding of the mechanisms underlying stroke is crucial, as treatment options remain limited for this debilitating disease. Unlike clinical diagnosis of stroke, cerebrovascular phenotypes detected on brain MRI can be characterized with high validity and quantified with good precision using advanced neuroimaging analysis techniques.[13] Targeting these specific endophenotypes will lead to substantial pathophysiologic insight into stroke mechanisms and foster future advances in individualized therapy and prevention.

The MRI–Genetics Interface Exploration (MRI-GENIE) study aims to bridge current knowledge gaps by facilitating genetic discovery and developing novel therapeutic and preventive strategies in stroke through automated, multimodal MRI analysis. Leveraging the existing infrastructure of SiGN and harnessing its expertise, MRI-GENIE focuses on the subset of richly phenotyped and genotyped participants for whom clinical MRIs have been obtained. This article outlines the premises, methodology, and aims of MRI-GENIE.

METHODS

MRI-GENIE capitalizes on SiGN, an ongoing multicenter, NIH-funded collaboration within the community of stroke neurologists, geneticists, and neuroimaging analysts, which enabled the initial development of the SiGN Imaging Platform.[14] We have amassed the largest-to-date collection of ischemic stroke cases with comprehensively ascertained cerebrovascular phenotypes and genome-wide data. The project is funded by the NIH-NINDS (R01NS086905, N.R. Rost—PI) to undertake the first major study to jointly model MRI-derived traits obtained during acute ischemic stroke (AIS) evaluation, causative and phenotypic stroke subtypes, and traditional vascular risk factors to accelerate the pace of genetic discoveries and advance clinical applications in risk and outcome prediction in ischemic stroke.

Structure of MRI-GENIE. ***Participating study sites.*** MRI-GENIE is founded on the existing collaborations between members of the multidisciplinary clinical stroke research team at Massachusetts General Hospital (MGH), Massachusetts Institute of Technology (MIT), and the NINDS SiGN investigators. To date, 12 sites from the initial SiGN study have contributed phenotypes, images, and genotypic data of 3,301 participants to MRI-GENIE (For a summary, see table and figure; detailed study descriptions have been published previously[11]). Of those sites, 7 are European centers (BASICMAR—Spain, BRAINS—United Kingdom, GRAZ—Austria, KRAKOW—Poland, LEUVEN—Belgium, LUND STROKE REGISTER—Sweden, and SAHLSIS—Sweden) and 5 are based in the United States (GASROS, GCNKSS, GEOS, ISGS, and MIAMISR). Informed consent of study participants to data sharing was mandatory for all sites. Shared data include basic demographics, vascular risk factors and detailed Causative Classification of Stroke (CCS) phenotyping, genotypic data, and clinical MRIs.

Study oversight. The primary aims and progress of the MRI-GENIE study are overseen by a Scientific Steering Committee. In conjunction with the SiGN Publication Committee, the MRI-GENIE Steering Committee critically reviews project proposals by collaborators to avoid potential overlap with existing SiGN projects and to assess feasibility of the proposed projects. This effort is supported by the Phenotyping Committee, which is in charge of the data access to phenotypes previously obtained through SiGN, current data acquisition for neuroimaging markers, and quality control (QC) of new neuroimaging phenotypes. In addition, it is responsible for the oversight of statistical analysis of MRI-derived phenotypes and functional outcomes related to specific stroke subtypes. The Neuroimaging Analysis Committee is in charge of designing, validating, and implementing the MRI pipeline to automatically assess acute and chronic neuroimaging markers. Moreover, it facilitates and monitors the assessment of manually obtained neuroimaging markers. The Genetic Analysis Committee conducts the primary genetic analyses for the MRI-GENIE study to identify genetic variants associated with acute and chronic MRI-based manifestations of cerebrovascular disease. It is also essential in conducting secondary analyses and additional projects as proposed by collaborators (detailed listings of committee memberships are available in coinvestigator appendix e-1 at Neurology.org/ng).

Imaging platform. An integral part of the MRI-GENIE study is the centrally maintained imaging platform hosting deidentified acute or subacute brain MRIs obtained within 48 hours of symptom onset from all contributing SiGN sites. The MRI-GENIE Imaging Platform is maintained centrally at MGH and has been described previously.[14] Initially developed in the scope of an NIH-funded project to create a centralized system to share canonical human stroke data (R01 NS063925-01A1, O. Wu/Sorensen—PI), the underlying technology for the imaging platform integrates the eXtensible Neuroimaging Archive Toolkit (XNAT)[15] as the back-end data repository with a flexible, open-source content management system with user-friendly features (conglomeration of Plone,[16] Deliverance,[17] and NGINX[18]) as the front end for the users. An example of such features is the ability to search the imaging repository by stroke-specific clinical phenotypic variables (e.g., age, sex, or infarct location). Images can be viewed for semiquantitative analysis via a web-based XNAT plugin.[15]

Upon receipt of MRIs from the individual sites in the "Digital Imaging and Communications in Medicine" (DICOM) format, all images were deidentified to remove protected health information (PHI) potentially embedded into the DICOM headers. In addition, DICOM files with image type indicative of screen shots were removed to ensure elimination of files with potential PHI that may be "burned" into the screen shots. Each site provided phenotypic data (e.g., SiGN ID, sex, race, ethnicity, age, and infarct location). Age and sex were cross-referenced with phenotypes documented in the SiGN Phenotype Database, to flag potential discrepancies between databases.

The imaging platform is open to collaborators for the exploration of phenotypic and genetic underpinnings of AIS. Individual investigators receive access in a project-based manner, after the MRI-GENIE Scientific Steering Committee and the SiGN Publication Committee have reviewed the project proposal. Incorporating the computational workflow for automated segmentation of acute and chronic cerebrovascular phenotypes with

Table Basic demographic data for the MRI-GENIE study sites (n = 12)

Study name	Center	Total scans	Mean age (SD)	Sex (% female)	Race (% Caucasian)
BASICMAR	IMIM-Hospital del Mar, Spain	124	69.8 ± 11.0	37.1	94.4
BRAINS	Imperial College—London, UK	70	63.2 ± 16.4	47.1	94.3
GASROS	Mass General Hospital	457	64.9 ± 14.5	35.4	93.2
GCNKSS	U Cincinnati	245	64.3 ± 14.3	49.0	72.7
GEOS	U Maryland	76	41.8 ± 6.5	26.3	52.6
SAHLSIS	U of Gothenburg—Sweden	401	52.4 ± 11.7	38.7	100
GRAZ	Medical University—Graz, Austria	373	63.3 ± 13.7	30.0	100
ISGS	Mayo Clinic—Florida	425	65.1 ± 14.7	40.9	84.2
KRAKOW	Jagiellonian University—Poland	224	60.5 ± 13.9	46.4	100
LEUVEN	U Hospitals—Leuven, Belgium	448	66.9 ± 14.7	42.0	99.6
LUND STROKE REGISTER	Lund University Hospital, Sweden	196	63.4 ± 12.8	39.3	100
MIAMISR	U Miami	262	62.1 ± 13.8	37.0	59.5

Abbreviation: MRI-GENIE = MRI-GENetics Interface Exploration.

the imaging platform will complete the process for centralized data abstraction, collection, and sharing for future genetic studies.

MRI-derived phenotypes. The analysis of MRI-derived phenotypes will include both automated, volumetric analyses as well as manual, semiquantitative analyses of acute and chronic cerebrovascular phenotypes.

Image preprocessing. Prior to phenotypic analyses, imaging data from some sites required a manual review. As part of the de-identification process, some sites removed all metadata information from the DICOM headers, including series description, prior to transmission of images. This necessitated a visual review of cases by a trained operator for sites for which no series descriptions were available. For these sites, sequence labels for T2 fluid-attenuated inversion recovery (T2 FLAIR), diffusion-weighted imaging (DWI), susceptibility-weighted imaging (SWI), and magnetic resonance angiography were ascertained. Although the quality of specific sequences may preclude future analysis of specific phenotypes, no participants were excluded for image quality reasons (e.g., motion artifact and low-quality images) or incomplete imaging data (e.g., missing FLAIR sequence) at this stage to maximize the overall sample size in analysis of individual phenotypes. For manual semiquantitative analyses, useful information can often be extracted (e.g., location of acute infarct) by expert image analysts despite artifacts (e.g., motion).

Image QC. Image sequences will be excluded from automated and semiquantitative analysis based on visual inspection due to excessive motion artifact, incomplete sequence acquisition, or severe bilateral brain pathology that precludes accurate assessment. Details of the MRI data acquisition that are embedded in the DICOM headers (e.g., MRI manufacturer, model name, field strength, dimensions, echo time, and repetition time) will be retained to allow for subset analysis of phenotypic information as a function of data acquisition variability.

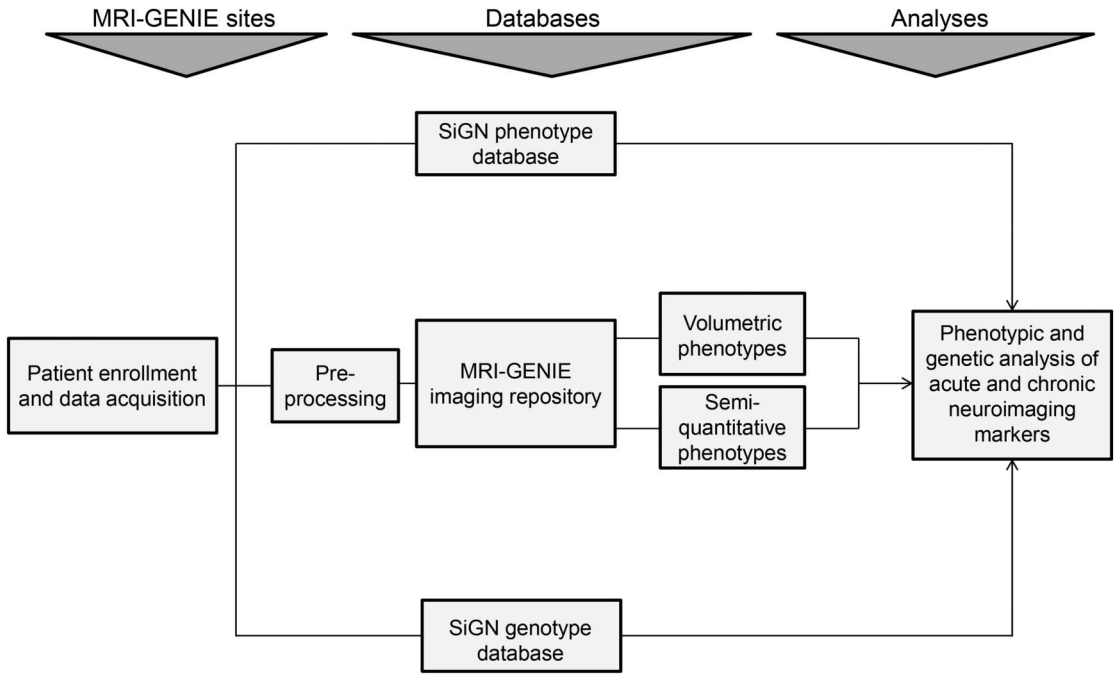
Quantitative analyses of acute and chronic neuroimaging markers. The acute infarct volume measured on DWI will be analyzed through an automated segmentation algorithm.[19,20] The algorithm will be trained on an independent cohort of acute stroke data and then applied to MRI-GENIE participants. The results will be validated across sites against a random sample set of images selected from each site for which manual lesion volumes will be drawn. The automatically determined acute DWI volume (DWIv) will be used to evaluate the effect of stroke lesion volume and topography on poststroke outcomes. Voxel-based lesion–symptom mapping (VLSM) has been used previously to investigate the relationship between ischemic stroke lesion location and stroke severity and outcome, but so far has been limited to the middle cerebral artery territory. The sample size of MRI-GENIE will allow for investigating all vascular territories with VLSM.[21]

The pre-existing burden of cerebrovascular white matter disease, measured as white matter hyperintensity volume (WMHv), will be computed by a separate automatic analysis algorithm based on T2-FLAIR images.[22] Key components of the WMH analysis will be the registration of all T2-FLAIR images to a common atlas space and segmentation of WMH based on the voxel intensity information and spatial priors that pattern WMH in the brain.[23] The algorithm will be trained on manually outlined WMH segmentations. Validation across sites will be performed for each site on a random sample of scans representing the entire severity spectrum of WMH.

QC of automated, quantitative phenotypes. The results of automatic acute DWI and WMH lesion analysis will be cross-validated with manually drawn lesion volumes obtained on a random subset of scans from each site representative of the disease spectrum. Intraclass correlation coefficient (ICC) will be used to assess the agreement between the manual and the automated volumes. Further evaluation of the automated algorithms will involve voxel-based comparison of the automated masks with manual outlines and classifying true positives, true negatives, false positives, and false negatives. Performance of the automated algorithms will be assessed by analysis of sensitivity, specificity, and Dice similarity coefficient.[24] Subset analysis of algorithm accuracy as a function of MRI data acquisition parameters will be performed to assess for bias as a function of scanner data quality.

Semiquantitative analysis of acute and chronic neuroimaging markers. MRIs will be systematically reviewed for acute and chronic markers of cerebrovascular disease to facilitate topography-based and stroke subtype–specific analyses. Specific data sets may be excluded on a case-by-case basis if artifacts preclude an accurate reading. DWI will be used to ascertain acute infarct location (vascular territories: middle cerebral artery, anterior cerebral artery, posterior cerebral artery, vertebrobasilar artery, and multiple vascular territories),

Figure **Organizational structure of MRI-GENIE**



MRI-GENIE = MRI-GENetics Interface Exploration; SiGN = Stroke Genetics Network.

number of acute infarct lesions, and DWI-based stroke subtypes (cortical, subcortical, and watershed infarct). When available, maps of the apparent diffusion coefficient (ADC) will be cross-referenced to minimize inclusion of lesions that are not pertinent to the index stroke. This is necessary because in the subacute to chronic stage, within hyperintense DWI lesions, the corresponding ADC values may be pseudonormal or elevated as a result of vasogenic edema.[25] T2-FLAIR sequences will be used to screen for subacute and chronic infarcts, as well as WMH severity using the Fazekas scale.[26] Cerebral microbleed (CMB, ≤ 10 mm) count and location, macrohemorrhages (>10 mm),[27] and hemorrhagic infarct transformation will be rated on 2D T2* gradient echo or 3D SWI sequences. In addition, the location of arterial occlusion, collateral circulation grade, and evidence of significant extracranial or intracranial large artery stenosis will be evaluated on MR angiography.

QC of semiquantitative phenotypes. Readers will be systematically trained on a standardized training set of AIS MRIs to attain the independent rater status (e.g., Fazekas score or CMBs). Furthermore, agreement between raters will be evaluated with the ICC for ordinal and continuous data, and a Cohen kappa will be used to assess interrater agreement for categorical data.

AIS subtyping. All patients with AIS in SiGN underwent extensive phenotyping through the web-based, standardized algorithm "Causative Classification of Ischemic Stroke" (details on CCS phenotyping in SiGN have been previously published).[14] In brief, CCS incorporates multiple clinical symptoms, clinical examinations, and testing results obtained throughout the clinical stroke evaluation and assigns both a phenotypic CCS subtype based on abnormal testing results at the time of stroke and causative CCS (CCSc) subtyping based on prior medical history. If challenged with multiple potential causes of ischemic stroke, CCSc assigns the most likely cause of stroke based on clinical data. CCSc allocates one of 5 different causative and phenotypic subtypes based on symptoms, vascular risk factors, and diagnostic tests. The 5 major categories are large artery atherosclerosis (LAA), small artery occlusion (SAO), CE stroke, other, and undetermined causes of stroke. Known rare causes of ischemic stroke were excluded from MRI-GENIE. The exact workup was performed as clinically directed by individual study centers. This includes patients with rare monogenic causes of stroke (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes, and sickle cell syndrome), infectious causes (infective endocarditis, meningitis, and primary infection of the arterial wall), hypercoagulability (acute disseminated intravascular coagulation and heparin-induced thrombocytopenia type II), distinct vascular and cardiac disorders (acute arterial dissection, dilated cardiomyopathy, papillary fibroelastoma, left atrial myxoma, and cerebral venous thrombosis), as well as migraine-related, drug-induced, or iatrogenic causes of ischemic stroke.

In addition to CCS, Trial of Org 10172 Acute Stroke Treatment (TOAST)[28] subtyping was conducted by the sites if required by the individual study protocols. Ischemic stroke cases are assigned to the most likely cause of stroke: LAA, SAO, CE, other, or undetermined. In SiGN, CCS and TOAST agree moderately (agreement rate 70%), varying by site and stroke subtype with LAA having the highest agreement across CCS and TOAST and SAO having the lowest agreement rate.[29] TOAST subtyping will be valuable in instances if future replication cohorts will not have appropriate CCS subtyping.

Functional outcomes and clinical characteristics. Baseline stroke severity was assessed using the NIH Stroke Scale,[30,31] and the modified Rankin Scale (mRS)[32] was used to assess functional outcomes at 3–6 months after stroke in a subset of the SiGN sites. These data will be available through collaboration with the Genetics of Ischemic Stroke Functional Outcome (GISCOME) study. The mRS measures the degree of dependence and disability after neurologic injury, ranging from

0 (asymptomatic) to 6 (death). Other clinical baseline characteristics were collected per protocol of the individual studies.

Genotyping and analysis strategy. Stroke cases were either genotyped previously and genotypes submitted to SiGN, or they were genotyped as part of SiGN, at the Center for Inherited Disease Research on the Illumina HumanOmni5Exome-4v1 array.[14] This platform includes 4.5 million single nucleotide polymorphisms (SNPs) genotyped across the genome, resulting in excellent coverage of both common and infrequent variants (>1%). In addition, 240,000 rare polymorphic variants (≤1%) were genotyped. The majority of cases are of European ancestry (table), but a small number of African Americans (n = 249, self-reported) and Hispanics (n = 153, self-reported) are available for analysis. The SiGN Data Management Core has undertaken the data cleaning and QC procedures for the primary analysis in SiGN. QC procedures included data cleaning by subjects with removal of samples with (1) a poor genotyping rate (<98%), (2) identity problems (sex mismatch, unexpected duplicates, and cryptic relatedness), (3) chromosomal anomalies, (4) batch effects, and (5) ethnic outliers. Poor-quality SNPs were identified on the basis of high levels of missingness. The cleaned data are maintained at the SiGN Data Management Core. The analysis plan for the MRI-GENIE portion of SiGN includes a principal component analysis to identify and account for population stratification in subsequent genome-wide association testing. Genotype imputation has been performed on the University of Michigan Imputation Server (imputationserver.sph.umich.edu/index.html)[33] using the Haplotype Reference Consortium panel.[34] The primary genetic analysis will be to test for the association of SNPs with the DWI and WMH volumes. The association analyses will be performed under a linear regression model with allelic dose (0, 1, or 2 copies of the reference allele) as the independent variable adjusted for age, sex, and population stratification as calculated by principal component analysis. DWI and WMH volumes will be adjusted for average head size and natural log transformed to facilitate modeling with linear regression because of nonlinear volume distributions. The current sample size of 3,301 participants will provide 80% power to detect variants accounting for as little as 1.2% of the variation in DWI or WMH at genome-wide levels of statistical significance (i.e., $p < 5 \times 10^{-8}$). Secondary analyses will include the modeling of CCS subtypes and ethnic group–specific analyses. SNPs found to be associated with WMHv and/or DWIv will be tested for association with 90-day mRS data available in the GISCOME study. Additional analyses will be performed to examine stroke subtype–specific genetic effects on DWIv and WMHv as well as the genetic underpinnings of additional neuroimaging markers.

Incoming sites. In the spirit of open collaboration, the MRI-GENIE Imaging Platform is available to collaborators. We are currently incorporating 8 new sites (estimated additional total n = 3,890) including (1) Secondary Prevention of Small Subcortical Strokes, n ≈ 1,000, (2) Siblings With Ischemic Stroke Study, n ≈ 300, and (3) Washington State University–St. Louis stroke patient collection, n ≈ 640, (4) Helsinki-2000 study, n ≈ 300, (5) Australian Stroke Genetics Cohort, n ≈ 100, (6) Stroke in Young Fabry Patients, n ≈ 800, (7) University of Campinas stroke patient collection, n ≈ 150, and (8) Follow-up of Transient Ischemic Attack and Unelucidated Risk Factor Evaluation Study/Observational Dutch Young Symptomatic StrokE studY, n ≈ 600. These collaborations will lead to the development of one of the largest (n ≈ 7,000) databases of patients with ischemic stroke with MRI and genome-wide genotyping available to date.

CONCLUSIONS Quantitative neuroimaging has recently been used to gain further insight into physiology and anatomy in both healthy participants (e.g., intracranial volume[35]) and clinical cohorts (e.g., hippocampal volumes[36] and structural neuroimaging biomarkers[37] in Alzheimer disease). In this article, we aim to analyze acute (e.g., cerebral infarct volume) and chronic (e.g., WMH volume and cerebral microbleeds) neuroimaging phenotypes in patients with AIS. While genetic loci associated with WMH have previously been reported in healthy populations,[38,39] no loci associated with WMH have yet been identified in patients with AIS, despite efforts with large sample sizes.[40] Similarly, no studies have linked specific genetic loci with cerebral microbleeds or stroke lesion volume. To facilitate these genetic studies, MRI-GENIE will develop, validate, and disseminate an automated analysis pipeline for large-scale phenotypic analysis of clinical brain MRI, as part of the future advances in personalized prediction modeling of stroke risk and outcomes.

AUTHOR AFFILIATIONS

From the Department of Neurology (A.-K.G., M.D.S, K.L.D., L.C., J.R., N.S.R.), Massachusetts General Hospital, Harvard Medical School, Boston; Program in Medical and Population Genetics (A.-K.G., J.R.), Broad Institute of MIT and Harvard; Computer Science and Artificial Intelligence Laboratory (M.D.S., A.V.D., R.S., P.G.), MIT, Cambridge, MA; Department of Radiology (R.I., S.W., M.J.R.J.B., E.C.M., S.J.M., A.V.D., O.W.), Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital (MGH), Harvard Medical School, Charlestown; Division of Endocrinology, Diabetes and Nutrition (H.X., P.F.M., B.D.M.), Department of Medicine, University of Maryland School of Medicine, Baltimore; Department of Clinical Sciences Lund, Neurology (P.F., A.L.) and Department of Clinical Sciences Lund, Radiology (J.W.), Lund University, Sweden; Department of Neurology and Rehabilitation Medicine, Neurology (P.F.) and Department of Radiology, Neuroradiology (J.W.), Skåne University Hospital, Malmö, Sweden; Department of Neurology (E.G.-S., J.R., J.J.-C.), Neurovascular Research Group (NEUVAS), IMIM-Hospital del Mar (Institut Hospital del Mar d'Investigacions Mèdiques), Universitat Autonoma de Barcelona, Spain; Institute of Neuroscience and Physiology (L.H.) and Institute of Biomedicine (C.J.), Sahlgrenska Academy, University of Gothenburg, Sweden; Department of Neurology (J.W.C., S.J.K.), University of Maryland School of Medicine and Veterans Affairs Maryland Health Care System, Baltimore; Department of Neurology and Rehabilitation Medicine (J.P.B., B.M.K., D.O.K., D.W.), University of Cincinnati College of Medicine, OH; Department of Neurosciences (R.L.), Experimental Neurology, KU Leuven—University of Leuven; VIB—Center for Brain & Disease Research (R.L.); Department of Neurology (R.L.), University Hospitals Leuven, Belgium; Department of Neurology and Rehabilitation Medicine, Neurology (A.L.), Skåne University Hospital, Lund, Sweden; Department of Neurology (J.F.M.), Mayo Clinic, Jacksonville; Department of Neurology (T.R.), Miller School of Medicine, University of Miami, FL; Department of Neurology (R.S.), Clinical Division of Neurogeriatrics, Medical University Graz, Austria; Institute of Cardiovascular Research (P.S.), Royal Holloway University of London (ICR2UL), Egham; Ashford and St Peter's Hospital (P.S.), UK; Department of Neurology (A.S.), Jagiellonian University Medical College, Krakow, Poland; Stroke Division (V.T.), Florey Institute of Neuroscience and Mental Health; Department of Neurology (V.T.), Austin Health, Heidelberg, Australia; Departments of Neurology and Public Health Sciences (B.B.W.), University of Virginia, Charlottesville; Geriatric Research and Education Clinical Center (B.D.M.), Veterans Administration Medical Center, Baltimore, MD; and Center for Human Genetic Research (J.R.), Massachusetts General Hospital, Boston.

AUTHOR CONTRIBUTIONS

Anne-Katrin Giese: review of imaging data, data analysis, drafting and revising the manuscript. Markus D. Schirmer: design of white matter hyperintensity lesion outlining algorithm and critical revision for intellectual content. Kathleen L. Donahue and Lisa Cloonan: maintenance and review of imaging data and critical revision for intellectual content. Robert Irie: design, development, and maintenance of imaging database and critical revision for intellectual content. Stefan Winzeck: design of acute lesion outlining algorithm and critical revision for intellectual content. Mark J.R.J. Bouts: maintenance of imaging database and critical revision for intellectual content. Elissa C. McIntosh: maintenance and review of imaging data and critical revision for intellectual content. Steven J. Mocking: maintenance of imaging database, design of acute lesion outlining algorithm, and critical revision for intellectual content. Adrian V. Dalca and Ramesh Sridharan: design of white matter hyperintensity lesion outlining algorithm and critical revision for intellectual content. Huichun Xu: analysis of genetic data and critical revision for intellectual content. Petrea Frid and Eva Giralt-Steinhauer: acquisition of imaging data and critical revision for intellectual content. Lukas Holmegaard: acquisition of data and critical revision for intellectual content. Jaume Roquer and Johan Wasselius: acquisition of imaging data and critical revision for intellectual content. John W. Cole, Patrick F. McArdle, Joseph P. Broderick, Jordi Jimenez-Conde, Christina Jern, Brett M. Kissela, James F. Meschia, Tatjana Rundek, Ralph L. Sacco, Reinhold Schmidt, Pankaj Sharma, Agnieszka Slowik, Vincent Thijs, Daniel Woo, and Bradford B. Worrall: acquisition of data and critical revision for intellectual content. Steven J. Kittner: study concept and design and critical revision for intellectual content. Braxton D. Mitchell: study concept and design, genetic analyses, and critical revision for intellectual content. Jonathan Rosand: study concept and design and critical revision for intellectual content. Polina Golland: study concept and design, design of white matter hyperintensity lesion outlining algorithm, and critical revision for intellectual content. Ona Wu: study concept and design, design and development of imaging database, design of acute lesion outlining algorithm, and critical revision for intellectual content. Natalia S. Rost: study concept and design, study supervision, drafting the manuscript, and critical revision for intellectual content.

ACKNOWLEDGMENT

On behalf of the BASICMAR project, the authors thank Carolina Soriano-Tarraga and Alejandra Gomez for their critical support of the study. R. Lemmens is a senior clinical investigator of FWO Flanders.

STUDY FUNDING

NIH-NINDS (MRI-GENIE: R01NS086905—PI N.S. Rost; K23NS064052, R01NS082285—N.S. Rost; SiGN: U01 NS069208—J. Rosand and S.J. Kittner; and R01NS059775, R01NS063925, R01NS082285, P50NS051343, R01NS086905, and U01 NS069208—O. Wu), NIH NIBIB (P41EB015902—P. Golland; U01NS030678—B.M. Kissela and D.O. Kleindorfer; and EB015325—O. Wu), Swedish Heart and Lung Foundation, the Swedish Stroke Association—A. Lindgren, Swedish Research Council and the Swedish Heart and Lung Foundation—C. Jern, Spanish Ministry of Science and Innovation, Instituto de Salud Carlos III (Funding for Research in Health [PI051737], [PI10/02064], [PI12/01238], and [PI15/00451—J. Jimenez-Conde]), Fondos FEDER/EDRF Red de Enfermedades Vasculares Cerebrales (RD16/0019/0002—J. Roquer), Fundació la Marató TV3 (76/C/2011—J. Jimenez-Conde) and Recercaixa'13 (JJ086116—J. Jimenez-Conde), Wistron Corporation (P. Golland).

DISCLOSURE

A.-K. Giese has received research support from the NIH/NINDS. M.D. Schirmer reports no disclosures. K.L. Donahue has received research support from the NIH/NINDS. L. Cloonan has been an employee of Decision Resources Group. R. Irie has received travel funding from the Department of Defense. S. Winzeck reports no disclosures. M.J.R.J. Bouts has received research support from Leiden University. E.C. McIntosh reports no disclosures. S.J. Mocking has been an employee of Google. A.V. Dalca has served on the editorial board of *Imaging Genetics*; has received publishing royalties from *Imaging Genetics* and Elsevier; has received research support from Massachusetts Institute of Technology; and has been an employee of Massachusetts Institute of Technology and Massachusetts General Hospital (part of Partners HealthCare, a nonprofit); in addition, his wife is employed by Partners HealthCare. R. Sridharan reports no disclosures. H. Xu holds a patent for Methods for diagnosing etiologies of ischemic stroke by detecting expression levels of genes and has received research support from the NIH, University of Maryland Claude D. Pepper Older Americans Independence Center, and AHA. P. Frid, E. Giralt-Steinhauer, and L. Holmegaard report no disclosures. J. Roquer has received research support from Spain's Ministry of Health. J. Wasselius reports no disclosures. J.W. Cole has received publishing royalties from Springer. P.F. McArdle has received research support from the NIH. J.P. Broderick has served on a scientific advisory board for Astra Zeneca; has received travel funding from universities as invited visiting professor over the past 2 years (approximately 4–5 times per year); has received honoraria for speaking placed in faculty academic fund within department; holds a patent for Method for Controlling the Lysis of Coagulated Blood with Apolipoprotein e4 Phenotype; has received publishing royalties from Elsevier; has been a consultant for Pfizer; and has received research support from Genentech and the NIH/NINDS. J. Jimenez-Conde has been an employee of IMIM-Hospital del Mar and has received research support from the NIH/NINDS, ISCIII/Spanish Government Grants, and La Marató TV3. C. Jern has received research support from the Sahlgrenska Academy and the Sahlgrenska University, the Swedish Medical Research Council, the Swedish Heart Lung Foundation, and the Swedish Stroke Foundation. B.M. Kissela has served on the scientific advisory board of Ipsen Pharmaceuticals; has received compensation to perform event adjudication for clinical trials sponsored by Janssen and AbbVie; and has received research support from the NIH. D.O. Kleindorfer has received travel funding from Genentech; has served on the speakers bureau of Genentech; and has received research support from the NIH/NINDS. R. Lemmens has received research support from FWO Flanders. A. Lindgren has served on scientific advisory boards for Astra Zeneca, Bristol-Myers Squibb, Boehringer Ingelheim, Sanofi, Pfizer, Bayer, and ReNeuron; has received travel funding/speaker honoraria from Astra Zeneca, Bristol-Myers Squibb, Boehringer Ingelheim, Sanofi, Pfizer, and Bayer; and has served on the editorial board of *Neurologi i Sverige*. J.F. Meschia has served on a scientific advisory board for the NINDS NeuroNEXT clinical trial network; has served on the editorial boards of *Mayo Clinic Proceedings* and the *European Journal of Neurology*; and has received research support from the NIH/NINDS. T. Rundek has served on the editorial boards of *Stroke*, *Neurology*, *Cerebrovascular Disease*, and the *Journal of Ultrasound in Medicine*; and has received research support from the NIH/NINDS. R.L. Sacco has served on scientific advisory boards for SOCRATES Trial through UCSF (indirect sponsor Astra Zeneca) and EUCLID trial through Duke Clinical Research Institute (indirect sponsor Astra Zeneca); has served on the editorial boards of *Stroke* and *Neuroepidemiology*; has been a consultant for Boehringer Ingelheim; and has received research support from Boehringer Ingelheim, the NIH/NINDS, the Evelyn McKnight Brain Institute, and the American Heart Association Bugher Foundation. R. Schmidt has served on scientific advisory boards for Rivastigmine Patch; Novartis JSW Life Sciences, and Dementia Trial Axon Neuroscience; has received travel expenses and/or honoraria for lectures from Pfizer, Novartis, Merz Austria, and Lundbeck, and for educational activity from Takeda and Axon Neuroscience; has served on the editorial board of *Clinical Neurology and Neurosurgery*; has been an employee of Pfizer, Novartis, Merz Austria, Lundbeck, and Takeda; has been a consultant for QPS and Austria Axon Neuroscience; and has received research support from Austrian Science Fund. P. Sharma has served on scientific advisory boards for Bohringer and Bayer and has served on the editorial board of the *Journal of Royal Society of Medicine Cardiovascular Disease*. A. Slowik reports no disclosures. V. Thijs has served on scientific advisory boards for Bayer, Boehringer Ingelheim, Amgen, Pfizer/BMS, Daiichi Sankyo, and Medtronic; has served on the editorial boards of *Acta Neurologica Belgica*, *Stroke*, *International Journal of Stroke*, *European Stroke Journal*, and *European Journal of Emergency Medicine*; has been a consultant for Bayer, Boehringer Ingelheim, Amgen, Pfizer/BMS, Daiichi Sankyo, and Medtronic; and has received research support from the NIH/NINDS and VIB Flanders. D. Woo has served on the editorial board of *Stroke Research and Treatment* and has received research support from the

NIH. B.B. Worrall has served on the editorial board of *Neurology*; receives publishing royalties from Merritt's Neurology; and has received research support from the NIH/NINDS and the University of Virginia. S.J. Kittner has served on the editorial board of *Neuroepidemiology*; has received research support from the NIH/NINDS and has participated in legal proceedings regarding Atlantic States Insurance (review of medical case). B.D. Mitchell has served on the editorial board of *Diabetes* and has received research support from the NIH/NINDS. J. Rosand has served on scientific advisory boards for Pfizer and DSMB; has served on the editorial boards of *Lancet Neurology* and *Stroke*; and has received research support from the NIH/NINDS. P. Golland has received travel funding/speaker honoraria from the University of Utah School of Computing Visiting Committee; has served on the editorial board of *Journal of Medical Image Analysis*; and has received research support from Wistron Corporation and the NIH. O. Wu has received travel funding/speaker honoraria from RSNA and the NIH; holds patents for Delay-compensated calculation of tissue blood flow and Method for evaluating novel stroke treatments using a tissue risk map; receives publishing royalties from Nobel Prize Internet Archive; has been a consultant for Penumbra; has received research support from Genentech and the NIH/NINDS; and receives royalty payments from General Electric, Olea Medical, and Imaging Biometrics. N.S. Rost has served on scientific advisory boards for Genzyme Co., Merck, Boston Biomedical Associates, Omniox, and Bio-Telemetry; has served on the editorial boards of *Stroke* and *Current Treatment Options in Cardiovascular Medicine*; receives publishing royalties from UpToDate; has been a consultant for Genzyme Co., Merck, Boston Biomedical Associates, and Omniox; and has received research support from the NIH/NINDS. Go to Neurology.org/ng for full disclosure forms.

Received March 9, 2017. Accepted in final form June 30, 2017.

REFERENCES

1. Markus HS, Bevan S. Mechanisms and treatment of ischaemic stroke—insights from genetic associations. Nat Rev Neurol 2014;10:723–730.
2. Lindgren A. Stroke genetics: a review and update. J Stroke 2014;16:114–123.
3. Pjanic M, Miller CL, Wirka R, Kim JB, DiRenzo DM, Quertermous T. Genetics and genomics of coronary artery disease. Curr Cardiol Rep 2016;18:102.
4. McPherson R, Tybjaerg-Hansen A. Genetics of coronary artery disease. Circ Res 2016;118:564–578.
5. Geller DS. New developments in the genetics of hypertension: what should clinicians know? Curr Cardiol Rep 2015;17:122.
6. Tucker NR, Clauss S, Ellinor PT. Common variation in atrial fibrillation: navigating the path from genetic association to mechanism. Cardiovasc Res 2016;109:493–501.
7. Hegele RA, Ginsberg HN, Chapman MJ, et al. The polygenic nature of hypertriglyceridaemia: implications for definition, diagnosis, and management. Lancet Diabetes Endocrinol 2014;2:655–666.
8. Karaderi T, Drong AW, Lindgren CM. Insights into the genetic susceptibility to type 2 diabetes from genome-wide association studies of obesity-related traits. Curr Diab Rep 2015;15:83.
9. Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying cause of death 1999-2014 on CDC WONDER online database, released 2015. Data are from the multiple cause of death files, 1999-2014, as compiled from data provided by the 57 vital statistics jurisdictions through the vital statistics Cooperative Program. Available at: wonder.cdc.gov/ucd-icd10.html. Accessed November 21, 2016.
10. Malik R, Traylor M, Pulit SL, et al. Low-frequency and common genetic variation in ischemic stroke the METASTROKE collaboration. Neurology 2016;86:1217–1226.
11. NINDS Stroke Genetics Network (SiGN), International Stroke Genetics Consortium. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurol 2016;15:174–184.
12. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol 2012;11:951–962.
13. Masdeu JC, Irimia P, Asenbaum S, et al. EFNS guideline on neuroimaging in acute stroke. Report of an EFNS task force. Eur J Neurol 2006;13:1271–1283.
14. Meschia JF, Arnett DK, Ay H, et al. Stroke genetics network (SiGN) study design and rationale for a genome-wide association study of ischemic stroke subtypes. Stroke 2013;44:2694–2702.
15. Marcus DS, Olsen TR, Ramaratnam M, Buckner RL. The extensible neuroimaging archive toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. Neuroinformatics 2007;5:11–33.
16. The Plone Foundation. Plone CMS: open source content management. Available at: plone.org/homepage. Accessed March 1, 2017.
17. The open planning project deliverance. Available at: pythonhosted.org/Deliverance/. Accessed March 1, 2017.
18. NGINX Inc. NGINX: high performance load balancer, web server, & reverse proxy. Available at: nginx.com/. Accessed March 1, 2017.
19. Winzeck S, Bouts M, McIntosh E, et al. Early identification of tissue at risk of infarction after Acute Ischemic Stroke. 2015 NIPS Workshop on Machine Learning for Healthcare; December 11, 2015; Montreal, Canada.
20. Wu O, Winzeck S, Mocking SJT, et al. 19th Workshop of the International Stroke Genetics Consortium, April 28–29, 2016, Boston, Massachusetts, USA: 2016.002 Modeling acute lesion volume and topography for genetic analysis (MRI-GENIE study). Neurol Genet 2017;3 (1 suppl 1):S2–S11. doi: 10.1212/NXG.0000000000000100.
21. Wu O, Cloonan L, Mocking SJ, et al. Role of acute lesion topography in initial ischemic stroke severity and long-term functional outcomes. Stroke 2015;46:2438–2444.
22. Schirmer MD, Dalca A, Sridharan R, et al. 19th Workshop of the International Stroke Genetics Consortium, April 28–29, 2016, Boston, Massachusetts, USA: 2016.008 The automated pipeline for white matter hyperintensity detection: the MRI-GENIE study. Neurol Genet 2017;3(1 suppl 1):S2–S11. doi: 10.1212/NXG.0000000000000100.
23. Dalca AV, Sridharan R, Cloonan L, et al. Segmentation of cerebrovascular pathologies in stroke patients with spatial and shape priors. Med Image Comput Comput Assist Interv 2014;17:773–780.
24. Mocking S, Rost NS, Fitzpatrick KM, et al. Automatic Segmentation of Diffusion MRI from the Genes Associated with Stroke Risk and Outcomes Study. ISMRM 21st Annual Meeting & Exhibition; April 20–26, 2013; Salt Lake City, UT.
25. Copen WA, Schwamm LH, González RG, et al. Ischemic stroke: effects of etiology and patient age on the time course of the core apparent diffusion coefficient. Radiology 2001;221:27–34.
26. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. AJR Am J Roentgenol 1987;149:351–356.

27. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol 2013;12:822–838.
28. Adams HP, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 1993;24:35–41.
29. McArdle PF, Kittner SJ, Ay H, et al. Agreement between TOAST and CCS ischemic stroke classification the NINDS SiGN Study. Neurology 2014;83:1653–1660.
30. Brott T, Adams HP, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke 1989;20:864–870.
31. Lyden P, Brott T, Tilley B, et al. Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. Stroke 1994;25:2220–2226.
32. Bonita R, Beaglehole R. Recovery of motor function after stroke. Stroke 1988;19:1497–1500.
33. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287.
34. Loh P-R, Danecek P, Palamara PF, et al. Reference-based phasing using the haplotype reference consortium panel. Nat Genet 2016;48:1443–1448.
35. Adams HHH, Hibar DP, Chouraki V, et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nat Neurosci 2016;19:1569–1582.
36. Song A, Yan J, Kim S, et al. Network-based analysis of genetic variants associated with hippocampal volume in Alzheimer's disease: a study of ADNI cohorts. BioData Min 2016;9:3.
37. Moon SW, Dinov ID, Kim J, et al. Structural neuroimaging genetics interactions in Alzheimer's disease. J Alzheimers Dis 2015;48:1051–1063.
38. Fornage M, Debette S, Bis JC, et al. Genome-wide association studies of cerebral white matter lesion burden: the CHARGE consortium. Ann Neurol 2011;69:928–939.
39. Verhaaren BFJ, Debette S, Bis JC, et al. Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. Circ Cardiovasc Genet 2015;8:398–409.
40. Traylor M, Zhang CR, Adib-Samii P, et al. Genome-wide meta-analysis of cerebral white matter hyperintensities in patients with stroke. Neurology 2016;86:146–153.