This supplement to the DEFUSE 3 study contains the following items:

Protocol
- Original protocol (dated October 13, 2015)
- Final protocol (dated April 20, 2017)
- Summary of changes (Change Log).

Statistical Plan
- Original statistical analysis plan (Version 1.0)
- Final statistical analysis plan (Version 3.1)
- DSMB amendment July 19, 2017
- DSMB amendment July 24, 2017
- Summary of changes to SAP

PROTOCOL TITLE

DEFUSE 3:
Endovascular Therapy Following Imaging Evaluation for Ischemic Stroke 3

Protocol Version/Version Date

Version 1.3
October 14, 2015

Protocol Directors

Gregory Albers, MD
Michael Marks, MD
Maarten Lansberg, MD, PhD

Stanford University
Stanford Stroke Center
780 Welch Rd. Suite CJ350
Stanford, CA 94304

Supported by

The National Institute of Neurological
Disorders and Stroke (NINDS)
U01 NS092076

IDE Number

G150028

DEFUSE 3

AGREEMENT ON THE PROTOCOL

By signing below I confirm that:

1) I have read this protocol and it contains all necessary details for conducting this study

   AND

2) I agree to conduct the trial in compliance with this protocol and to adhere to all regulations that govern the conduct of the study.

_____ _____

Principal Investigator's Signature **Date**

_____

Principal Investigator's Name

_____

Site Name:

TABLE OF CONTENTS

1. SUMMARY OF TRIAL

DEFUSE 3 is a prospective, randomized, multi-center, Phase III, adaptive, blinded endpoint, controlled trial. A maximum of 476 patients will be randomized and treated between 6 and 16 hours of symptom onset. Subjects will be randomized 1:1 to endovascular therapy plus medical management or medical management alone. *Only the devices listed in this protocol application are approved for use in DEFUSE 3. The choice of device or devices employed is at the discretion of the clinical investigator.*

2. SCIENTIFIC BACKGROUND

Although stroke is the number one cause of adult disability in the United States[1], treatment options for stroke are limited. The only FDA approved treatment for stroke is administration of intravenous (iv) tissue plasminogen activator (tPA) within 3 hours after symptom onset. Nationwide, only about 4 percent of stroke patients receive this therapy.[2] The main reason for this low treatment rate is that most patients present to the hospital outside the time-window for tPA.[3, 4] Even when administered, tPA is often not effective because it either fails to recanalize the occluded artery[5-7] or because the brain is already irreversibly injured.[8] As a result, it is estimated that only 12-25% of treated patients benefit from tPA.[9] Thus, in order to improve outcomes from stroke we need better treatments that are available to a greater proportion of stroke patients.

2.1 State of the science on endovascular stroke therapy for acute stroke

Endovascular stroke therapy, the removal of blood clots with mechanical devices or thrombolytic drugs administered intra-arterially, is the most promising new treatment for patients who "fail" treatment with iv tPA or are not eligible for iv tPA. The main advantage of endovascular therapy is that it has a high rate of recanalization.[10, 11] Blood flow can be restored with a success rate of up to 82% with modern thrombectomy devices and 66% for intra-arterial thrombolysis.[10, 12, 13] This is approximately twice as effective as iv tPA which has a recanalization rate of 10-50% depending on the location of the blood clot.[6, 7] Despite higher rates of recanalization with endovascular therapy, two recent randomized controlled trials of endovascular therapy, IMS III and MR RESCUE, have failed to demonstrate a clinical benefit. Patient and treatment related factors likely both contributed to the neutral results of these trials (see below for details).

Patient-related factors: A central consideration in the optimization of patient selection for acute stroke therapies is the concept of the ischemic penumbra. Ischemic penumbra is defined as ischemic tissue that is potentially salvageable and is distinguished from the ischemic core that has already sustained irreversible injury. Clearly, the target of acute stroke therapies is salvage of the ischemic penumbra, preventing infarct growth and, most importantly, improved functional outcome. Acute stroke trials should therefore ideally be limited to patients with an ischemic penumbra. MRI-based studies, such as DEFUSE 1 and 2, indicate that MRI can be used to identify these patients.[8, 14, 15 16, 17]

Treatment-related factors: Recent studies have emphasized the importance of recanalization rates, demonstrating the influence on patient outcome of highly effective endovascular procedures that lead to complete or near-complete reperfusion.[18, 19] Currently, the most common metric for rating the quality of reperfusion is the modified Thrombolysis

In Cerebral Infarction (mTICI) scale, and a clear relationship exists between the degree of reperfusion on the mTICI scale and patient outcome.[20-22] Patients with >50% reperfusion (mTICI 2B-3) are much more likely to have a good outcome than patients with <50% reperfusion.

IMS III, the largest endovascular trial to date, did not use advanced imaging criteria to select patients.[23] Instead it used relatively strict time-criteria, anticipating that this would yield a high proportion of patients with a substantial penumbra. However, several categories of patients who likely did not have substantial penumbra were enrolled. Nearly one third of the patients did not have a vessel occlusion at angiography and 23% had distal MCA occlusions; both of these subgroups are unlikely to have substantial penumbral tissue. In addition, 42% of patients had some evidence of irreversible tissue injury (ASPECTS <8) on their baseline CT and 14% had evidence of extensive irreversible injury (ASPECTS <5).[24] Finally, the endovascular devices that were available during IMS III had relatively low rates of early reperfusion; mTICI 2B-3 was only 40% in IMS III.

MR RESCUE, a stroke trial that aimed to demonstrate benefit of endovascular therapy in patients with a penumbra based on MRI, had neutral results.[25] Several factors likely contributed to this. First, the rate of endovascular reperfusion was extremely low. Only 8 patients (24%) in the MR RESCUE penumbral group achieved TICI 2B-3 reperfusion during endovascular therapy. Second, patients in the penumbral group in MR RESCUE had larger baseline infarct core lesions (median volume 36 ml; IQR 24–51 ml) than the Target Mismatch patients in DEFUSE 2 (median volume 13 ml; IQR 5–26 ml). The combination of low rates of endovascular reperfusion and relatively large core lesions, both strong predictors of poor clinical outcome, likely explains the lack of a treatment effect in "penumbral patients" in MR RESCUE. Moreover, with only 8 endovascular patients in the penumbral cohort achieving TICI 2B-3 reperfusion, MR RESCUE was substantially underpowered.

New Generation Trials: Recently, a series of positive randomized studies of endovascular therapy with treatment initiated within 6 hours of stroke onset in the vast majority of patients were reported. This has prompted new guidelines endorsing endovascular therapy up to 6 hours after symptom onset. The American Heart Association is now calling for late window studies using advanced imaging for patient selection: ""Further randomized, controlled trials should be done to determine whether advanced imaging paradigms using CT perfusion and MRI perfusion, CTA, and diffusion imaging, including measures of infarct core, collateral flow status, and penumbra, are beneficial for selecting patients for acute reperfusion therapy who are beyond 6 hours from symptom onset. (New recommendation, 2015 AHA Guidelines).

DEFUSE 3 will address this new mandate. Enrollment is limited to patients with salvageable tissue (Target Mismatch patients) who are likely to respond favorably to endovascular reperfusion in the 6-16 hour window. Use of the latest generation thrombectomy devices, coupled with strict qualification and oversight criteria for the neuro-interventionalists, will result in high rates of reperfusion. Based on the compelling preliminary data from DEFUSE 2, the trial is adequately powered to demonstrate a clear treatment effect.

2.2. Prior studies and rationale for development

This application aims to shift the selection of patients for reperfusion therapy from a relatively arbitrary decision based on poorly validated clinical characteristics to an objective decision based on scientific evidence. Many factors affect the evolution of the ischemic penumbra into the ischemic core, and the rate of progression of irreversible injury is highly variable between individuals. This variability is likely mediated by the adequacy of collateral blood flow as well as the metabolic milieu of individual stroke patients. The individuality of penumbral evolution among stroke patients implies that identifying the extent of the ischemic core and penumbra is useful for making treatment decisions. Currently diffusion-weighted imaging (DWI) / perfusion-weighted imaging (PWI) magnetic resonance imaging (MRI) affords the best opportunity for approximating the ischemic core and penumbra in real time clinical practice.[27]

The DWI lesion provides a dependable estimation of the ischemic core and only very rarely shows permanent reversal following early reperfusion.[28, 29] PWI identifies hypoperfused ischemic tissue. Regions defined as abnormal on PWI that do not demonstrate a DWI abnormality, often referred to as the DWI/PWI mismatch, can estimate the ischemic penumbra.[30] It is critical that PWI utilizes an appropriately validated threshold parameter that excludes ischemic tissue with modest blood flow reduction (i.e. benign oligemia), because this tissue is unlikely to infarct even if reperfusion does not occur. Which PWI parameter is optimal, as well as what threshold to use to define critical hypoperfusion, has been the focus of multiple research efforts.[31] Prior work from our group and others supports the use of Tmax, thresholded at >6 seconds, as the optimal PWI parameter to identify ischemic tissue destined to become infarcted if timely reperfusion does not occur.[32] This Tmax threshold correlates well with the penumbral range of cerebral blood flow decline as determined by both positron emission tomography and Xenon CT.[33, 34]

Using a difference between the volume of the baseline PWI Tmax lesion and the DWI volume to identify mismatch, the DEFUSE and EPITHET studies found that most patients with a PWI/DWI mismatch responded favorably if reperfusion occurred following iv tPA treatment in the 3- to 6-hour time window. However, despite having a mismatch, patients with very large baseline DWI lesions (large core infarct volumes) had highly unfavorable outcomes following reperfusion. Patients with this MRI pattern, referred to as the Malignant profile, had a significantly higher rate of both parenchymal hemorrhage and severe disability/death if reperfusion occurred.[35] Mismatch patients who do not have the Malignant profile have been designated as having a Target Mismatch, and these patents respond extremely favorably to reperfusion following iv tPA therapy. In a pooled analysis of DEFUSE and EPITHET, Target Mismatch profile patients who experienced reperfusion had a 5-fold increase in favorable clinical response at 90 days and significantly less infarct growth when compared to those who did not reperfuse.[36] No association between reperfusion and favorable outcomes, or a reduction in infarct growth, was apparent for patients without the mismatch profile.

The DEFUSE 2 study utilized an automated mismatch analysis program (RAPID) to prospectively establish MRI profiles in a consecutive cohort of patients who then underwent endovascular therapy. DEFUSE 2 confirmed the concepts demonstrated in DEFUSE and EPITHET; Target Mismatch patients who achieve early reperfusion therapy have less infarct growth and more favorable clinical outcomes (8·8,95% CI 2·7–29·0).[37] No association between reperfusion and favorable outcomes or infarct growth was present in patients

without Target Mismatch. Furthermore, the positive association between reperfusion, favorable clinical response, and attenuation of infarct growth did not diminish in DEFUSE 2 patients with Target Mismatch who were treated relatively late (6-12 hours after symptom onset, see **Figure 1**).

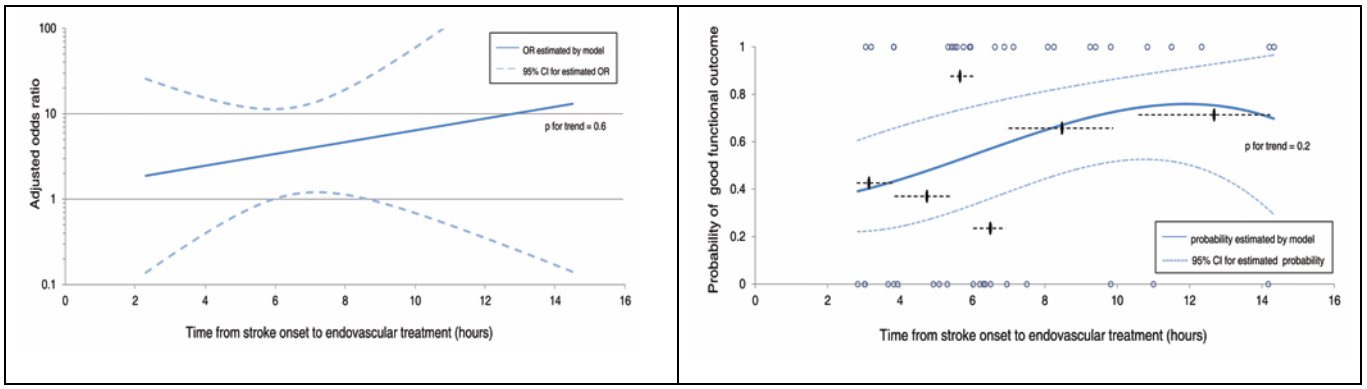

Figure 1: The effect of time on the association between reperfusion and good functional outcome (graph on the left) and the effect of time to treatment on the probability of good functional outcome patients with reperfusion (graph on the right) in Target mismatch patients. 95% CIs are indicated by dashed lines. Estimates are based on multivariate logistic regression, adjusted for age and baseline DWI volume. There is no significant effect of time in either model.[38]

This finding contrasts sharply with prior studies that did not use penumbral imaging to select patients and suggests that imaging findings may be of equal, or potentially even greater, importance than time from symptom onset for identification of optimal patients who might benefit from reperfusion therapy.

How could Target Mismatch patients who are treated late have outcomes that are as favorable as those of earlier treated patients? At later time points, the Target Mismatch profile identifies patients in whom the infarct is evolving at a relatively slow rate; a DWI lesion that is still considerably smaller than the PWI lesion reflects good collateral circulation. These collaterals typically allow prolonged, but not permanent, survival of the hypoperfused mismatch region. Evidence that the mismatch region is still at considerable risk for infarct expansion, even at later time points, was provided by the DEFUSE 2 finding that Target Mismatch patients imaged between 6 and 12 hours from symptom onset consistently demonstrated substantial infarct growth if reperfusion was not achieved.[38] Patients with slowly evolving infarct cores are ideal candidates for later time window reperfusion therapy, particularly endovascular therapies. One of the drawbacks of the endovascular approach is that the time between hospital arrival and achievement of endovascular reperfusion is typically at least 90 to 120 minutes. For patients with rapidly growing infarct cores (such as patients with the Malignant profile), substantial growth of the infarct core has been reported despite endovascular reperfusion. Therefore, removing this population, which represents about 20% of eligible patients, from a randomized endovascular trial, has important advantages.

DEFUSE 2 confirmed that early DWI lesions are an excellent surrogate for the ischemic core. Despite endovascular therapy, only 2 patients had a final infarct that was smaller than the baseline DWI lesion and the size and location of the early DWI lesion was a reliable predictor of the final infarct volume in patients with complete reperfusion.[39, 40] In DEFUSE 2, younger age and smaller DWI volume were significant independent predictors of favorable outcome. Subgroup analysis of EPITHET identified DWI lesion size ≤25 ml as a strong predictor of a

favorable response to reperfusion.[37, 41] These findings suggest that certain subgroups, in particular individuals with Target mismatch and small DWI lesions are most likely to benefit from reperfusion. The adaptive design of DEFUSE 3 (see below) has the potential to focus patient enrollment on a subgroup of patients (e.g. those with smaller DWI lesion volumes and/or shorter times from symptom onset to randomization) who respond most favorably to endovascular therapy. This will allow the study to identify the largest population that has a statistically reliable benefit of endovascular therapy.

New data (presented at the International Stroke Conference, February 2015) suggest that CT Perfusion studies, processed with the same software (RAPID) used in the studies described above, can identify the ischemic core with accuracy similar to MRI (Cereda, et al ISC 2015) and select patients who respond to endovascular reperfusion therapy in early time windows (Campbell, et al NEJM 2015, EXTEND-IA study, Saver, et al NEJM 2015, SWIFT PRIME study). Therefore, DEFUSE 3 will allow patient selection with both MRI and CT Perfusion.

3. INVESTIGATIONAL PLAN

3.1. Purpose

DEFUSE 3 is a prospective randomized Phase III multicenter controlled trial of patients with acute ischemic anterior circulation strokes due to large artery occlusion treated between 6-16 hours of stroke onset with endovascular thrombectomy therapy vs. control. The primary endpoint, the modified Rankin Score, will be assessed at 3 months. The patients' participation in the study concludes at that time (3 months from stroke onset). The study will randomize up to 476 patients over 4 years. The purpose of DEFUSE 3 is to assess the safety and efficacy of thrombectomy in carefully selected patients in an extended time window. Only the devices listed in this protocol will be used. Selection of the specific device (or devices) is determined by the individual endovascular therapist.

3.2. Protocol Design

DEFUSE 3 is a prospective randomized Phase III multicenter controlled trial of patients with acute ischemic anterior circulation strokes due to large artery occlusion treated between 6-16 hours of stroke onset. Patients who meet the inclusion criteria will undergo either CT Perfusion/CTA or MR DWI/PWI/MRA studies prior to randomization. Patients who have evidence of an ICA or MCA M1 occlusion and a Target Mismatch Profile will be randomized in a 1:1 ratio to treatment with one or more DEFUSE 3 approved thrombectomy devices (o*nly the devices listed in this protocol are approved for use in DEFUSE 3)* plus standard medical therapy versus standard medical therapy alone. Patients who are enrolled, but not randomized, will receive standard therapy according to local guidelines. Baseline data, and information about early stroke therapies, will be captured for this group of patients.

Randomization of a maximum of 476 patients is planned. A novel adaptive design (see below) will identify, at interim analyses, the group with the best prospect for showing benefit from endovascular treatment, based on baseline core lesion volumes and the times since stroke onset. Interim analyses will be conducted at 200 and 340 patients, at which time the study may stop for efficacy/futility, or the inclusion criteria may be adjusted in the case of futility.

Approximately 35 StrokeNet sites as well as 10 highly selected non-StrokeNet sites and 10

"back-up" sites will be chosen. Individual site selection will be based on a number of factors including endovascular volume, MRI and/or CT perfusion access, number of competing trials, clinical trial experience, and the diversity of their patient population. If a site does not enroll a patient within 4 months of activation, it will be placed on probation. If no enrollment occurs in the next 2 months, the site will be replaced with a "back-up" site.

3.3. Enrollment Criteria

3.3.1. Clinical Inclusion Criteria:

1. Signs and symptoms consistent with the diagnosis of an acute anterior circulation ischemic stroke
2. Age 18-85 years
3. Baseline NIHSSS is ≥ 6 and remains ≥6 immediately prior to randomization
4. Able to undergo an MRI or CT Perfusion within 90 minutes of arrival at the study site.
5. Endovascular treatment can be initiated (femoral puncture) between 6 and 16 hours of stroke onset. Stroke onset is defined as the time the patient was last known to be at their neurologic baseline (wake-up strokes are eligible if they meet the above time limits).
6. No significant pre-stroke disability (pre-baseline modified Rankin Scale score 0-2)
7. Patient willing/able to return for protocol required follow up visits
8. Patient/Legally Authorized Representative has signed the Informed Consent form.

3.3.2. Clinical Exclusion Criteria:

1. Other serious, advanced, or terminal illness (investigator judgment) or life expectancy is less than 6 months.
2. Pre-existing medical, neurological or psychiatric disease that would confound the neurological or functional evaluations or patient is already participating in another drug or device study
3. Pregnancy
4. Contraindication for MRI (e.g. pacemaker, severe claustrophobia)
5. Contraindications for both MRI and CT contrast precluding an MRI or CT contrast perfusion study. The hospital's local standard criteria should be applied to determine if contraindications exist.
6. Known allergy to iodine and previously refractory to pretreatment medications
7. Treated with tPA >4.5 hours after time last known well
8. Known hereditary or acquired hemorrhagic diathesis, coagulation factor deficiency; recent oral anticoagulant therapy with INR > 3 (recent use of one of the new oral anticoagulants is not an exclusion if eGFR > 30 ml/min).
9. Symptoms consistent with stroke in multiple locations.
10. Seizures at stroke onset if it makes the diagnosis of stroke doubtful and precludes obtaining an accurate baseline NIHSS assessment
11. Baseline blood glucose of <50mg/dL (2.78 mmol) or >400mg/dL (22.20 mmol)
12. Baseline platelet count < 50,000/uL
13. Severe, sustained hypertension (Systolic Blood Pressure >185 mmHg or Diastolic Blood Pressure >110 mmHg) not treatable with medications.
14. Current participation in another investigational drug or device study or registry
15. Presumed septic embolus; suspicion of bacterial endocarditis or cerebral vasculitis
16. Subjects who have had clot retrieval attempted using a neurothrombectomy device

prior to 6 hours from symptom onset

3.3.3. Neuroimaging Inclusion Criteria:

1. ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA or CTA

AND

2. Target Mismatch Profile on CT perfusion or MRI (ischemic core volume is < 70 ml, mismatch ratio is > 1.8 and mismatch volume* is ≥ 15 ml)

Notes: The mismatch volume is determined by the RAPID software in real time based on the difference between the ischemic core lesion volume and the Tmax>6s lesion volume. If both a CT perfusion and a multimodal MRI scan are performed prior to enrollment, the later of the 2 scans is assessed to determine eligibility. Only an intracranial MRA is required for patients screed with MRA; cervical MRA is not required. Cervical and intracranial CTA are typically obtained simultaneously in patients screened with CTA, but only the intracranial CTA is required for enrollment.

Alternative neuroimaging inclusion criteria (if perfusion imaging or CTA/MRA is technically inadequate):

A) If CTA (or MRA) is technically inadequate:

Tmax>6s perfusion deficit consistent with an ICA or MCA-M1 occlusion

AND

Target Mismatch Profile (ischemic core volume is < 70 ml, mismatch ratio is >1.8 and mismatch volume is >15 ml as determined by RAPID software)

B) If MRP is technically inadequate:

ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA (or CTA, if MRA is technically inadequate and a CTA was performed within 60 minutes prior to the MRI)

AND

DWI lesion volume < 25 ml

C) If CTP is technically inadequate:

Patient can be screened with MRI and enrolled if neuroimaging criteria are met.

3.3.4. Neuroimaging Exclusion Criteria:

1. ASPECTS score <6 on non-contrast CT (if baseline non-contrast CT was performed)
2. Evidence of intracranial tumor (except small meningioma) acute intracranial hemorrhage, neoplasm, or arteriovenous malformation
3. Significant mass effect with midline shift
4. Evidence of internal carotid artery flow limiting dissection or aortic dissection
5. Intracranial stent implanted in the same vascular territory that would preclude the safe deployment/removal of the neurothrombectomy device
6. Subjects with occlusions in multiple vascular territories (e.g., bilateral anterior circulation, or anterior/posterior circulation)

3.4. Enrollment and Randomization

3.4.1 Enrollment: All patients who meet the clinical criteria listed above are eligible for enrollment. This includes both patients who are directly admitted to the study site and

patients who are transferred from an outside hospital. ***The time of enrollment is the time when the informed consent is signed.*** After enrollment the RAPID output from a CT perfusion or multimodal MRI scan will be assessed. If the patient is confirmed to meet the Neuroimaging eligibility criteria listed above and not have any of the Neuroimaging exclusion criteria then the patient will be randomized. In general, patients will be enrolled prior to obtaining the RAPID output maps. In some situations the CT perfusion/ multimodal MRI may have been performed as part of standard care *prior* to the patient being assessed for study eligibility. Patients who are enrolled but do not meet the imaging criteria will not be randomized.

Determination of Target Mismatch and Large Artery Occlusion: At the conclusion of the MRI or CT Perfusion scan, the technologist sends the sequences from the console to RAPID with a single mouse click for automated processing. The RAPID software was developed based on data from DEFUSE 1 and was prospectively validated in DEFUSE 2. The system provides fully automated processing of brain images.

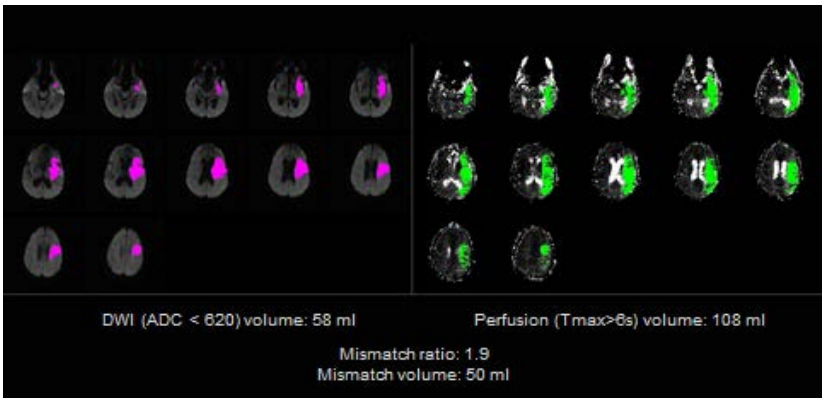

Figure 2. The RAPID mismatch summary allows investigators to quickly, accurately and easily determine if the patient meets the imaging criteria for enrollment. The patient above meets the Target Mismatch criteria: core volume is < 70 ml, mismatch ratio is ≥ 1.8 and mismatch volume is ≥ 15 ml.

The RAPID output maps, which identifies the volume and location of ischemic core and perfusion lesions, are emailed to investigators (protected health information is automatically removed) and auto-sent to PACS as well as to a secure email system for viewing within 5 minutes of completion of the scan. Immediately after the images are available, the investigator will review the results of the RAPID mismatch map (**Figure 2**) and the MRA/CTA to determine if the patient meets the imaging criteria (listed above). If a patient has undergone multiple imaging evaluations (both MRI and CT or multiple CTs or MRIs), the most recent imaging study will be used to determine if the patient meets the imaging criteria. The accuracy of the software for identifying the size and location of perfusion and diffusion lesions has been established by extensive validation and testing on blood flow phantoms; the software received FDA 510K clearance for clinical use in 2013. The agreement between local investigators and the Imaging Core Lab for identification of the mismatch profile in DEFUSE 2 was 97%, κ 0·92; 95% CI 0·83–1·

Baseline data, required for the stratification algorithm, will be captured for all enrolled patients, including for patients who do not meet the imaging criteria. Patients who do not meet the imaging eligibility criteria will receive standard care per local hospital practice and clinical data for the first 24 hours of their hospitalization will be captured on WebDCU™ (StrokeNet's web-based clinical trials management system).

3.4.2 Randomization using a Dynamic Stratification Algorithm: Once an enrolled patient

is determined to meet all Neuroimaging criteria, the patient ***will immediately be randomized*** on the WebDCU™ website. A dynamic stratification system will ensure well-balanced subgroups. The randomization algorithm, which will be programmed into the data capture system, will employ biased-coin minimization and the variance method with stratification weights.[44] The strategy is to balance treatment assignment along the marginal distribution of each stratification factor. The stratification factors used and their hierarchy will be: 1) core lesion volume, 2) age, 3) time from symptom onset to enrollment, 4) vessel occlusion (MCA vs. ICA) and 5) study site. When a new patient is enrolled, the site will enter the stratification factor values into the eCRF (electronic case report form) on WebDCU™. The dynamic randomization algorithm will determine an imbalance measure for each treatment group. The treatment group associated with the smallest imbalance measure will receive the largest probability of assignment in the biased-coin randomization. The optimal biased-coin acceptance region and stratification weights will be determined prior to study launch via simulations. The superior balancing characteristics of dynamic randomization over blocked randomization have been well established.

Patients will be assigned to either endovascular therapy plus medical therapy or to medical therapy alone (1:1 randomization). Crossover from medical to endovascular therapy is strictly prohibited; endovascular to medical therapy crossover is defined as a patient who is assigned to endovascular therapy but does not have a conventional angiogram performed. Endovascular to medical therapy crossover is only allowed if an endovascular contraindication arises after randomization. Sites will be closely monitored for crossovers (see site monitoring plan below).

3.5 Acute Treatment

3.5.1 Endovascular Therapy: In patients randomized to endovascular therapy, the goal for femoral artery puncture will be within 45 minutes of randomization; femoral artery puncture must occur within 60 minutes of randomization. Patients will be treated with thrombectomy devices (stent-retrievers) and/or suction thrombectomy systems currently cleared by the FDA for thrombus removal in patients experiencing an acute stroke within 8 hours of symptom onset following the published instructions for use for these devices. These devices will be used up to 16 hours following symptom onset in DEFUSE 3 based on an FDA IDE. The devices which will be used are the Trevo Retriever, the Solitaire device Revascularization Device and the Penumbra system thrombectomy system.
Standard medical therapy, based on current AHA guidelines, will also be provided for all patients. Individual investigators may use any of these devices or any combination of these devices to remove thrombus from the ICA, MCA M1 segment or, if needed, from M2 segments of the intracranial circulation. These are all approved anatomic locations for these devices. The use of thrombectomy devices should be performed in accordance with the indications for use. If there is a severe stenosis of the common carotid artery or the proximal internal carotid artery, investigators may also use other FDA devices approved for angioplasty or FDA devices approved for stenting of the carotid artery as deemed appropriate. The use of adjuvant intra-arterial (IA) thrombolytic medication is not currently approved by the FDA for stroke treatment and cannot be used in DEFUSE 3.

Sites will use local protocols for femoral access, sedation, heparin infusion, monitoring, etc. Sites will perform a cervical injection of the involved carotid circulation as a baseline

angiogram. At the conclusion of the procedure a post-treatment angiogram as a cervical injection of the involved carotid circulation will also be obtained. Imaging will cover the full region of the normal circulation in AP and lateral projections at 2-3 films per second through the entire venous phase. All baseline MRI and CT, as well as angiographic images obtained for the diagnostic and therapeutic portions of the procedure, will be transmitted to the core lab.

3.5.2 Medical Therapy: Patients randomized to medical therapy will receive standard medical therapy based on current AHA guidelines. Based on the time window for DEFUSE 3, it is anticipated that very few of the patients enrolled in DEFUSE 3 will have received iv tPA prior to randomization ("tPA failures"). For these patients, the sites' post-tPA protocol will be followed. Non-tPA treated patients randomized to medical therapy will be treated with aspirin, 325 mg on Day 1, and 81-325 mg/day (investigator's preference) on day 2-5, unless an indication for early anticoagulation is present (as determined by the patient's attending physician). All patients will receive standard DVT prevention therapy. Intravenous anticoagulants are prohibited (unless a clear indication for early anticoagulation is documented; dual antiplatelet therapy is prohibited unless carotid stenting was performed during the endovascular procedure or a clear indication for dual antiplatelet therapy is documented. At day 5 or discharge (whichever is sooner) subsequent antithrombotic therapy will be determined by the patient's attending physician.

3.6 Clinical and Imaging Evaluations

Follow-up (imaging and clinical): Randomized patients will be followed clinically for 90 days and will have an MRI/MRA/MR perfusion at 24 hours (range 18-30 hours) to assess infarct volume, recanalization, hemorrhage and reperfusion (**Table 1 below**).

Table 1 Schedule of Events

Evaluation	Baseline	24 hours after randomization	5 days or discharge	30 days	90 days
Informed Consent	✓				
History & Physical	✓			✓	✓
NIHSS Score	✓	✓	✓	✓	✓
Modified Rankin Scale+	✓		✓	✓	✓
TOAST subtype			✓	✓	
SSQOL				✓	✓
MRI or CTP scan	✓	✓++			
EKG / Laboratory Evaluation*	✓	✓			
Adverse Event Assessment		✓	✓	✓	✓

*Laboratory Evaluation includes CBC with Platelets, Creatinine, Glucose, INR, activated PTT, and Pregnancy test (if applicable). At 24 hour follow-up only creatinine is required. + Historical mRS at baseline, mRS to be performed by an mRS certified investigator who is blinded to treatment allocation at 30 and 90 days. ++ Patients will preferably undergo an MRI at 24 hours, if an MRI cannot be performed a CT/CTA/CTP can be substituted. For patients who are enrolled but not randomized the schedule of events is limited to all baseline data evaluations and a summary of stroke therapies received at 24-hrs.

3.6.1 Assessments and follow-up visits

Baseline visit: All items list in Table 1 above listed under "baseline" are to be performed

prior to randomization. The MRI or CT scan should be performed with the DEFUSE 3 (baseline) protocol, which will be installed at all study sites. In addition, the inclusion/exclusion page of the case report form must be completed to determine if the patient meets the eligibility requirements for the study. If the patient is eligible and the consent form is signed by the patient or authorized representative, then the randomization procedure should occur immediately.

24 hour visit (+/- 6 hours): The items listed for this visit in Table 1 should be performed between 18 and 30 hours from the time of randomization. The only laboratory value required at the 24 hour visit is a serum creatinine. If possible, the 24 hour follow-up imaging study should be performed with multimodal MRI, rather than CT perfusion. The MRI or CT scan performed at this time should be performed with the DEFUSE 3 protocol.

Day 5 (+/- 1 day) or Discharge visit: The items listed for this visit in Table 1 should be performed between day 4 and 6 from the time of randomization. If the patient is discharged prior to day 4, these items should be performed on the day of discharge.

30 and 90 day visits: The items listed for this visit in Table 1 should be performed on Day 30 (+/- 7) days and Day 90 (+/- 14) days. The mRS score must be performed by an mRS certified investigator who is blinded to treatment allocation at both the 30 and 90 day visits. If an in person visit is not possible, then the mRS should be performed by phone by an mRS certified investigator who is blinded to treatment allocation. If an in person visit is not possible, then the NIHSS score will be marked "not available" in the case report form.

3.6.2 Sources of Materials

Information on the clinical status of patients will be obtained from the patient's medical record. Study coordinators at the site will complete the DEFUSE 3 case report forms to collect basic demographic and medical information about the patients. Data will subsequently be entered into the StrokeNet's WebDCU electronic data capture system. MRI data will be electronically transmitted to the coordinating center at Stanford via RAPID (all patient identifiers are removed by the software prior to exporting the data outside of the site's firewall). All study sites will complete a stroke screening log in the StrokeNet's WebDCU electronic data capture system that documents all patients seen at their center, and reason for exclusion of patients not enrolled. Any patient who experiences a serious adverse effect (SAE) will be reported within 24 hours of the event on the StrokeNet's WebDCU electronic data capture system. The data collection process will include patient demographics, medical history, vital signs, laboratory assessments, NIHSS and mRS scores, and results of diagnostic studies performed to clarify stroke etiology.

3.7. Site Approval and Monitoring Plan

Site approval: Individual sites approved for participation in the study will be high-volume sites. Approval of neurointerventionalists at each site will follow a similar protocol to a previous endovascular trial.[45] A credentialing committee consisting of four experienced neurointerventionalists will determine if a site neurointerventionalist can participate. Approval will require a unanimous vote.

Selected sites will have access to emergent CT perfusion and/or MR imaging 24/7. Prior to activating a site, we will verify that RAPID is functional at the site. Together with the site's CT and/or MR technologists, we will install the DEFUSE 3 scan protocol on the local scanners

and perform a dummy-run to assess image quality and train the technologist in software handling and data sending. A site will be activated for enrollment after four test cases processed with RAPID have ensured good quality maps.

Monitoring for imaging quality: The Imaging Core Lab will monitor image quality throughout the study. If significant inadequacies or protocol errors are noted at a site, enrollment will be halted. Enrollment will resume after all imaging problems have been resolved and 2 repeat dummy runs have been obtained that demonstrate adequate image quality.

Table 2. Example Imaging Sequences for DEFUSE 3 scans

Sequence	Scan Parameters (3T)	Time
MRI		**6 min**
Localizer	128X256; 28 FOV;5/5mm, GRE	24 sec
Calibration		5 sec
DWI	128x128, 24 FOV, 5/0mm, 30 slices, 1 NEX, R=2; b=0 and 1000 s/mm^2 over 3 axes, TE/TR=min/7000ms.	25 sec
GRE	256x192; 24 FOV; 5/0 mm, 30 slices, TE/TR= 25/800ms, flip 20, interleaved EPI, 16 shots	27 sec
MRA intracranial	256x192, 1 mm; 4 slabs, 26 phase-encodes; 6 overlap, 22 FOV, 0.8 rFOV, fractional echo, ZIPx2, ZIPx512, minTE, flowcomp, TR=18ms, flip=18, inferior->superior ramppulse, R=2; 19 MIPS	143 sec
PWI	128x128; 24 FOV; 5/0 mm, 17 slices, TE/TR=35ms/1800ms, R=2 using 0.1mmol/kg Gadolinium @ 4ml/sec.	108 sec
CT (example below for GE VCT; comparable protocols will be used for other scanner models)		**5-6 min**
Non-con head	2.5 – 5mm, 40 slices, 120-140kV, 265-290mA	120-180 sec
CTA	0.625mm, 0.984:1/39.37cm, 120kV, 550mA , inject and observe for 15 sec until contrast concentration in ascending aorta reaches 80HU (smart prep) then the CT gantry moves along with the bolus of the contrast material from the aortic arch up to the apex of the brain in 5sec.	90 sec
CTP	22 FOV, 40mm, 8x5mm, 1.8sec time interval, 45 cycles, 80kV, 125mA; 2 runs	90 sec

Monitoring for bias: A detailed site-monitoring plan has been developed to detect bias. This plan will protect the study from enrollment, randomization, and treatment bias. The first component involves monitoring the percentage of each site's endovascular volume (6-16 hrs) that is enrolled in DEFUSE 3. Sites will report their volume of endovascular stroke procedures (6-16 hrs) each month on a screening log. If a DEFUSE 3 eligible patient is treated with endovascular therapy outside the DEFUSE 3 study, an explanation will be required detailing why the patient was not enrolled. Audits of the hospital database at each site will be performed every 6 months to verify that the endovascular volume matches the sites' screening log. The second component of the plan involves tracking of patients who are enrolled but not randomized. These patients will require an entry in the screening log with an explanation why the patient was not randomized as well as documentation whether endovascular therapy was performed outside of the study. A third component involves monitoring of crossover after randomization. Finally, it is expected that the majority of patients with be enrolled prior to obtaining or evaluation the results of the RAPID maps. The number of cases enrolled *after* the RAPID maps are reviewed will

be monitored. The Executive Committee will review the data described above for each site every 6 months. If evidence of enrollment bias is suspected, it will be investigated. If confirmed, the site will be placed on probation. If additional incidents of suspected bias are confirmed, the site will be withdrawn. Routine monitoring of the clinical sites for source to database verification will be performed by the StrokeNet Data Management Center.

3.8. Sample Size, Adaptive Design and Statistical Analysis

DEFUSE 3 will feature a novel adaptive trial design that will allow the study to focus on a subpopulation if interim or final analyses indicate futility in the overall population.[59] The adaptive design was developed specifically for DEFUSE 3. It is based on closed testing theory and the group sequential methods for the Generalized Likelihood Ratio (GLR) statistic developed by Lai and Shih.[60] The adaptive design was chosen because there is strong preliminary data that suggests that the effect of endovascular treatment is modified by two baseline variables: core lesion size and time-to-treatment. The way the adaptive design takes advantage of these biological assumptions (when they are true) is by reallocation of future accrual to the subgroup with the best prospects for showing efficacy. Specifically, if a subgroup is chosen at an interim analysis, subsequent enrollment is limited to patients in that subgroup. As a result, this subgroup will become larger than it would have been in the absence of the adaptive design. The criterion for deciding which subgroup has the best chance of showing a benefit from endovascular therapy combines both the estimated size of the effect in the subgroup and the sample size of the subgroup. The GLR statistic (Kullback-Leibler criterion) is used to identify this subgroup because it optimally balances those two criteria. It selects the subgroup that has the best chance of showing an effect because it has an apparently large effect and is also of substantial size (note there are 5 subgroups of increasingly larger size, **figure 3**). The adaptive design employs two biologically-based assumptions to limit the inflation of sample size; a monotonicity/contiguity assumption and a prior assumption that the effect is largest in the patients with the smallest DWI lesions and the shortest time to randomization (cell C_{11} in **figure 3**). The boundaries of the categories (cells) will be determined just prior to the 1st interim analysis based on the distribution of patients across these two dimensions (lesion volume and time-to-treatment).

		Time (hrs)	
		<10	10-16
Core lesion volume (ml)	≤20	C_{11}	C_{21}
	21-50	C_{12}	C_{22}
	51-70	C_{13}	C_{23}

Figure 3. The cohort is stratified according to core lesion volume and time to randomization. Exact boundaries of the stratification will be determined based on the distribution of patients at the time of the first interim analysis. After the 1st interim analysis, enrollment will continue in all 6 cells or will be limited to one of 5 sub-groups (C_{11}, C_{11+21}, $C_{11+21+12}$, $C_{11+21+12+22}$, or

Primary analysis: The primary endpoint is the distribution of scores on the modified Rankin Scale (mRS) at day 90. We will test the null hypothesis at the interim and final analysis using a normal approximation of the Wilcoxon-Mann-Whitney test (the generalized likelihood ratio [GLR] test). The primary analysis will be intention to treat, adjusted for design and not adjusted for covariates.

For each analysis, an efficacy bound will be set to control the overall (one-sided) Type I error rate at 2.5%. At each interim analysis a futility bound will be set to decide if the study should continue recruitment in the overall group, shift accrual and testing to a subgroup, or stop in its entirety. The futility boundary adapts when a subgroup is

selected to the fact that the maximum analyzed sample size is a random variable that is no larger than the fixed maximum number of patients randomized (n=476). Because subgroup selection reduces the maximum number of patients available for analysis at completion of the study, this method effectively allows an easier futility stop after subgroup selection. This setup replaces conditional power analyses with an automatic and more powerful adjustment of boundaries.

First interim analysis (n=200 enrolled): The null hypothesis is tested in the entire patient population:

1. If neither efficacy nor futility bound is crossed, the trial continues enrollment to the 2nd interim analysis.
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the overall population.
3. If the futility bound is crossed, the optimal subgroup is selected based on the Kullback-Leibler criterion and the null is tested in that subgroup. The futility bound is relaxed as described above, based on the expected maximum number of patients in the trial at completion (ie 476 minus the number of patients already enrolled in cells that will no longer be open for enrollment).
    3.1. If neither bound is crossed, the trial will continue with enrollment limited to the selected subgroup
    3.2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
    3.3. If the futility bound is crossed, the trial stops for futility.

Second interim analysis (n=340 enrolled): If, after the first interim analysis, the study proceeds with enrollment in the overall population (option 1 above), the testing at the 2nd interim analysis is identical to the first interim. If enrollment is limited to a selected subgroup (option 3.1), the null is tested in that subgroup:

1. If neither bound is crossed, the trial continues to the final analysis with enrollment of 136 additional patients limited to the selected subgroup
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
3. If the futility bound is crossed, the trial stops for futility.

Final analysis (n=476 enrolled): If, after the second interim analysis, the study proceeds with enrollment in the overall population the null is tested in the overall population:

1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in the overall population.
2. If the efficacy bound is not crossed, the optimal subgroup is selected and the null is tested in that group:
    2.1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in that subgroup
    2.2. If the efficacy bound is not crossed, endovascular therapy will be declared of no benefit.

If enrollment after one of the interim analyses is limited to a selected subgroup, the null will be tested in that subgroup only and efficacy or lack thereof will be declared as per options 2.1 and 2.2 above.

Power and sample size considerations: The projected overall effect of endovascular therapy is based on 1) the observed 90-day modified Rankin Scale outcomes in DEFUSE 2 of target mismatch patients treated >6hrs after symptom onset and 2) the assumption that early reperfusion will be achieved in 75% of the endovascular arm vs. 20% of the medical therapy arm.[20, 21, 61] Using these data, we projected the distributions on the mRS at 90 days in the endovascular and control arms of DEFUSE 3:

	mRS at day 90							
	0	1	2	3	4	5	6	Total
Endovascular group	18.0%	11.5%	19.6%	11.5%	16.4%	11.5%	11.5%	**100%**
Medical group	9.7%	7.9%	15.0%	17.7%	14.4%	17.7%	17.7%	**100%**

This distribution corresponds to a standardized effect of 0.36 for the primary analysis. Based on these data, the fixed sample size for a non-adaptive design requires a total of 376 patients (188/arm) to have 90% power at an alpha of 5% (Wilcoxon-Mann-Whitney test); 100 additional patients are added for the adaptive design to reach a maximum sample size of 476 for DEFUSE 3.

We ran simulations (n=5000) to compare the performance of a traditional fixed sample-size design (fixed n=476) to the adaptive design (max n=476) under various scenarios (see **Table 3**, top of next page). For the simulations the effect size is expressed as a standardized effect, where a standardized effect of 0.3 corresponds to a conservative projected effect of endovascular therapy (anticipated effect 0.36; see above).

Imaging outcomes: We hypothesize that endovascular treatment improves radiological outcomes in stroke patients with favorable clinical and imaging characteristics. DEFUSE 2 demonstrated a substantial reduction in infarct growth among Target mismatch patients treated in the 6-12 hour time-window who achieved early reperfusion: median growth 0.5 ml (IQR: -2 – 10) with reperfusion (n=23) vs. 39 ml (IQR: 18-121) without reperfusion (n=13), p<0.001. These data have been extrapolated to DEFUSE 3 using the same assumptions described above; anticipated an early reperfusion rate of 75% in the endovascular arm vs. 20% in the medical arm. This yields a sample size of 42 per group for 90% power. Therefore, DEFUSE 3 is highly powered to demonstrate differences lesion growth. Infarct volumes, ischemic lesion growth, and reperfusion rates at 24 hours will be compared between groups with the Mann-Whitney U test. The 24 hour endpoint is based on data demonstrating that assessment of infarct volume at 24 hours captures the effect of reperfusion therapies on infarct growth and predicts outcomes similarly to day 90 infarct volumes.[29, 62]

Scenario	Standardized effect in cells C11,	C12,	C21,	C22,	C31,	C32	Average standard. effect	Adaptive Design Average No. randomized	Power	Fixed Design Number randomized	Power
#0	0	0	0	0	0	0	0	361	2.2%	476	2.5%
#1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	354	80%	476	89%
#2	0.5	0.4	0.3	0	0	0	0.2	400	86%	476	55%
#3	0.5	0.5	0	0	0	0	0.17	403	87%	476	41%

Table 3. Under the null (Scenario #0), the adaptive design controls the total Type 1 error below 2.5%, stops early for futility 63% of the time, and the average number of randomizations is 361. If the effect is uniform across cells (scenario #1), the fixed-sample design is optimal, but the adaptive design results in only a small loss of power (from 89 to 80%). The adaptive design performs much better (higher power and smaller expected sample size) than the fixed sample, conventional trial when the effect size distribution across the subgroups is in accord with the biological assumptions (scenarios #2 and 3). If the effect is concentrated in two cells with small core volumes (scenario #3), the adaptive design maintains power (87%) while the conventional design collapses (41% power). The adaptive design also performs well compared to a non-adaptive, fixed sample that includes efficient multiple comparisons-adjusted testing for effect in subgroups at the end of the study. (see Lai et al[59])

Secondary analysis: Our secondary endpoint is the proportion of patients with mR S 0-2 at day 90 (indicating functional independence). The difference in the proportions of patients with mRS 0-2 between treatment arms will be assessed using logistic regression.

Subgroup analyses: Subgroup analyses of the effect of endovascular therapy on the primary and secondary endpoints will be performed. Subgroups will be defined based on the stratification variables, key demographic factors (such as race and ethnicity), tPA vs. no tPA, CTP vs. MRI selection, and witnessed vs. unwitnessed symptom onset, wake-up vs non-wake-up stroke, and TICI 0-2a vs. TICI 2b/3 results in cath lab. Homogeneity of treatment effects in these subgroups will be determined with the Breslow-Day test.

Missing data/lost to follow-up (LTFU): We will perform sensitivity analyses with standard methods for missing data (multiple imputation based on longitudinal models), but we do not expect this to alter the main study results given our estimated very low LTFU rate (<2%).

Coordination between MUSC and Stanford statistical teams: The primary statistician for the study is Dr. Phil Lavori at Stanford. He is voting member of the DEFUSE 3 Executive Committee (EC). Dr. Lavori will be blinded to all outcome data during the study. Dr. Lavori will become unblinded upon database lock and will conduct the final analyses. He will be responsible for developing and writing the statistical analysis plan (SAP) prior to the initiation of the study, and SAP amendments, if any, during the study. The statistical team at MUSC, led by Dr. Yeatts, will be unblinded throughout the study. The MUSC team will implement the adaptive design algorithm developed by Dr. Lavori; conduct and independently validate the interim analyses according to the SAP; generate Open and Closed Reports for the DSMB and interact with the DSMB in closed sessions; and collaborate with Dr. Lavori on validation of final analyses. After database lock, the MUSC statistical team will create the public use datasets (PUDS) and submit them to the NINDS.

DEFUSE 3 Timetable

Year 1	Year 2	Year 3	Year 4	Year 5
Install RAPID at all sites Begin enrollment (anticipated to begin mid-year)	Pt enrollment continues for a total of 4 yrs	1st interim analysis Potential modification of enrollment criteria based on adaptive design	2nd interim analysis Potential modification of enrollment criteria based on adaptive design	Finish enrollment Data Analyses Publication of results

3.9. Risk analysis

Description and analysis of all increased risks to the research subjects:

Potential complications of MRI scan include localized twitching sensation due to the magnetic field changes during the scan, anxiety due to claustrophobia and allergic reaction to the contrast agent. The allergic reaction may include headache, nausea, rash, hives, nasal congestion, sneezing, itching or swelling. If a severe reaction occurs, swelling of the throat, chest tightness, or a marked drop in blood pressure may occur. In addition, pain, bleeding, bruising, coldness or inflammation at the injection site may occur. Precautions will be taken for early detection and rapid treatment if such reactions occur.

Potential complications of CT scan include radiation exposure and allergic reaction to CT contrast agents.

Radiation doses:
Combined scanning with comprehensive stroke imaging, which includes a noncontrast head CT scan, perfusion imaging, and CT angiography of the cervicocranial vessels starting at the aortic arch results in a dose of approximately 7-10 mSv. (AJNR 2010 31: 1003-1009). According to the National Council on Radiation Protection and Measurement, the average annual radiation dose per person in the U.S. is 620 millirem (6.2 milliSieverts).

Reactions to contrast agents:

- **Mild**
  Nausea, vomiting**,** Headache**,** Cough**,** Nasal stuffiness, Altered taste**,** Flushing**,** Itching**,** Rash, hives**,** Sweats**,** Swelling of eyes or face
- **Moderate**
  Mild hypotension, Tachycardia or bradycardia, Bronchospasm, wheezing, Dyspnea, Laryngeal edema Generalized or diffuse erythema
- **Severe**
  Cardiopulmonary arrest, clinically manifest arrhythmias, Profound hypotension, Convulsions, Unresponsiveness, Respiratory, Laryngeal edema
  The rate of major reactions (e.g., anaphylaxis, death) is very low, estimated at one in 170,000 administrations.

Potential complications of endovascular therapy include stroke; new clot in an artery; total blockage of an artery; infection and pain in the region of insertion site; lack of blood

flow to the brain, rupture or puncture of an artery; significant tearing of the vessel wall; bleeding requiring blood transfusion; allergic reaction to contrast dye; abnormal low blood pressure requiring treatment; temporary closing of the artery (vessel spasm); formation of or dislodgments of clots which block the arteries (embolism). In rare circumstances, the procedure could result in death. At the puncture site in the groin, a blood clot or other blood vessel injury may occur and require blood transfusion or surgical repair. Infection may occur at the puncture site, this could cause pain and require additional medications. There is some chance of an allergic reaction to the x-ray contrast (dye) used during the angiogram procedure. Minor allergic reactions may include a rash or hives. There is also the possibility of a serious allergic reaction that could include shortness of breath and swelling, drop in blood pressure, and even death. Patients will be closely monitored for these reactions and receive prompt treatment to reverse any allergic reactions.

Safety of endovascular therapy beyond 8 hours

Mechanical thrombectomy devices have been used beyond 8 hours of stoke onset in a number of clinical trials and registries. No safety concerns have been associated with late window therapy. In DEFUSE 2 [37], patients were treated up to 12 hours after symptom onset and no safety concerns were identified in any time window. Based on both favorable safety data and encouraging efficacy data, Stryker Neurovascular has initiated the DAWN Study with a 24 hour treatment window. No safely issues have been identified to date in DAWN. The DAWN study is being run under an FDA IDE.

Methods to mitigate risks to subjects in the trial

Methods to mitigate risks to subjects in the trial include exclusion of subjects with bleeding disorders and selection of subjects via neuroimaging (infarct core lesions less than 70 ml) to minimize the risk of symptomatic intracranial hemorrhage. Computed Tomography (CT) scans will be performed for neurological deterioration (≥ 4 point increase in National Institutes of Health Stroke Scale (NIHSS) score) to identify new strokes, hemorrhage, or edema. Hospitals will follow their local standard of care safety procedures in order to reduce the risk of kidney dysfunction caused by contrast agents. Only investigators who are trained and experienced with use of the devices allowed within the trial are eligible to participate (see **Site Approval and Monitoring Plan**) above. The adaptive design will eliminate subgroups with an unfavorable therapeutic response.

Patients will be carefully screened for CT, MRI and endovascular treatment contraindications according to the inclusion/exclusion criteria and excluded from enrollment if any are present.

Radiation exposure: Radiation exposure during all tests will be minimized by optimizing the imaging protocols and by limiting fluoroscopy-time during the endovascular procedure. All CT sequences, including the CTP sequence, meet all FDA guidelines for radiation exposure.

Stopping rules or safety triggers for the study

Symptomatic ICH or death rates that exceed pre-specified thresholds with trigger a meeting of the DSMB to discuss the events and make a determination on the continuation of the trial. Below are the pre-specified triggers:

The DEFUSE 3 has established the following automatic stopping rules, based on identifying with 95% probability:

1) the rate of symptomatic ICH (NIHSS worsening of 4 or more points associated with ICH) in the endovascular group is exceeding 10%
2) the rate of 90 day mortality in the endovascular group is exceeding 20%.

If either threshold is crossed, the study will be automatically placed on hold until the investigators and the DSMB can conduct a review of events.

All adverse events (AEs) will be collected, recorded, and analyzed in accordance with Section 3.11 below. All AEs, regardless of severity, anticipation or relationship to the device or procedure will be recorded, collected, and reported.

Safety oversight for this study will be provided by both the DSMB and an independent Medical Safety Monitor. Please see Section 3.11 for details.

Patient Population: Four hundred and seventy six acute stroke patients meeting the pre-defined inclusion criteria will be enrolled in the trial. The mean age is anticipated to be 69 years of age. Our targeted planned enrollment breakdown is as follows:

Racial Categories	Ethnic Categories				Total
	Not hispanic or Latino		Hispanic or Latino		
	Female	Male	Female	Male	
American Indian/Alaska Native	2	2	1	1	6
Asian	13	13	1	1	28
Native Hawaiian or Other Pacific Islander	2	3	0	0	5
Black or African American	32	32	1	1	66
White	175	174	11	11	371
Racial Categories: Total of All Subjects	224	224	14	14	476

Imaging core lab: The Stanford imaging core lab has 15 years of experience with MR image storage and processing. It will perform the organization, archiving and blinded analysis of all imaging data collected in DEFUSE 3. They will be responsible for MRI and CT Perfusion image processing and artifact removal and will generate final lesion volumes for all MRI scans performed in the study.

DSA Angiograms sent to the core lab will include a baseline (pre-treatment) and a final angiogram for the territory of treatment. In addition, angiographic images from MRA or CTA will be sent to the core lab. The MRA or CTA will be used to assign a primary arterial lesion (AOL) from non-invasive imaging. The baseline DSA angiogram will also be used to assign a primary arterial occlusive lesion (AOL) and a pre-treatment mTICI score. The final angiogram will have a post-treatment mTICI score assigned.[22,52] This scoring system defines TICI 2A as partial perfusion of < 50% of the vascular distribution of the occluded artery and

2B as partial perfusion of > 50% of the vascular distribution. The scoring system will use these previously described definitions[22]:

- Grade 0 No perfusion
- Grade 1 Antegrade reperfusion past the initial occlusion, but limited distal branch filling with little or slow distal reperfusion
- Grade 2a Antegrade reperfusion of less than half of the occluded target artery previously ischemic territory (eg, in 1 major division of the MCA and its territory)
- Grade 2b Antegrade reperfusion of more than half of the previously occluded target artery ischemic territory (eg, in 2 major divisions of the MCA and their territories)
- Grade 3 Complete antegrade reperfusion of the previously occluded target artery ischemic territory, with absence of visualized occlusion in all distal branches

Two senior neurointerventionalists will perform the angiographic analysis, blinded to the clinical data, MR imaging and CT Perfusion results and the analysis of the other core lab reader. Any disagreements in the AOL interpretation or the mTICI scores will be adjudicated by common review of those cases and a consensus reading will be applied. The primary analysis of angiographic results will utilize the AOL determined from the non-invasive imaging studies. If the non-invasive imaging study is suboptimal the DSA AOL will be utilized. If there is a discrepancy between the Core lab and site investigator the core lab determination will be used for the analysis.

National Data Management Center: Data management and site monitoring will be performed by the StrokeNet National Data Management Center (NDMC) at Medical University of South Carolina (Director, Yuko Y. Palesch, PhD, see letter of support). The NDMC will create the database and set up the interface on the website (WebDCU™) where clinical site personnel will enter the data into the electronic CRF. Data quality assurance processes include: (1) logic and rule checks built into the database; (2) monitoring by the Data Manager at the NDMC; (3) risk-based monitoring by the statistical programmer at the NDMC; and (4) source verification monitoring by the Clinical Research Associates. The unblinded NDMC biostatistical team will generate periodic reports to the study team and the DSMB, perform interim statistical analyses, and interact with the DSMB in closed-session meetings. DEFUSE 3 data, including neuroimaging, will be shared in accordance with the StrokeNet data sharing policies. Anonymized neuroimaging will be stored on secure servers at the Stanford Stroke Center with nightly back-up. NINDS Common Data Elements will be used for both clinical and imaging data.

3.10. Description of devices

The following FDA approved thrombectomy devices will be included:

1) Trevo Retriever
2) Solitaire™ device Revascularization Device
3) Penumbra thrombectomy system including the following devices and pumps:
   Penumbra Aspiration Pump 115V
   Penumbra System Separator Flex [026, 032, 041 and 054]
   Penumbra System MAX Penumbra Pump MAX

3.11. Monitoring procedures

The coordination of the DEFUSE 3 Trial operations will be centralized through the following:
NIH StrokeNet National Coordinating Center (NCC)/ PI: Joseph Broderick, MD
University of Cincinnati
260 Stetson Street, Suite 2300
Cincinnati, Ohio 45267-0525

Leading the NCC team will be the Project Manager, Judy Spilker, who will be assigned to coordinate the following study oversight: trial communication required training activities, site assessment and or initiation visits, collection of trial related regulatory documents, recruitment performance tracking and site monitoring and performance analysis. Study oversight will be handled according to the Data Monitoring Standard Operating Procedure (SOP Number ADM 19).

DEFUSE 3 will have an independent Data Safety Monitoring Board (DSMB) appointed by the NIH to oversee study safety. Patients in both study arms will be assessed for the incidence of stroke-related mortality at 90 days, the incidence of symptomatic intracranial hemorrhage at 24 hours, and the incidence of significant neurologic deterioration at 5-7 days defined as ≥ 4 point increase in the NIHSS score. In the endovascular arm patients will be assessed for intra-procedural complications including: intra-procedural mortality, vessel perforation, arterial dissection, access site complication requiring surgical repair or blood transfusion, embolization and device failure. Any patient who experiences a SAE will be reported within 24 hours of the event.

The DSMB will meet in person or by teleconference, on a semi-annual basis, to monitor the cumulative safety data during participant follow-up. In no instance will more than 12 months elapse between DSMB reviews of cumulative safety data after the first participant has enrolled. The DSMB will monitor the study according to the guidelines specified in the study protocol and the operating procedures established at the initial meeting, unless the DSMB determines during the course of the trial that modification of the guidelines is in the best interest of the study and its participants.

Independent Medical Safety Monitor:
In addition to the DSMB, Dr. Andrew Demchuk has been appointed as the independent Medical Safety Monitor (MSM) for DEFUSE 3. Dr. Demchuk is not involved in the study and has no conflict of interest. He will be responsible for ongoing monitoring of reports of SAEs submitted by the clinical centers in real time to ensure good clinical practice and to identify safety concerns quickly. Dr. Demchuk may suggest protocol modifications to prevent the occurrence of particular AEs, e.g., modifying the protocol to require frequent measurement of laboratory values predictive of the event or to improve expeditious identification of SAEs. To minimize bias, he will usually evaluate SAEs blinded to treatment assignment, unless the DSMB approves partial or complete unblinding. Dr. Demchuk will prepare regular reports concerning SAEs (not segregated by treatment group) for submission to Dr. Albers, and subsequently to the DSMB. In the event of unexpected SAEs or an unduly high rate of SAEs, Dr. Demchuk will promptly contact Dr. Albers and the NINDS Program Official who will notify the DSMB Chair. In the event that he is unavailable for an extended period of time (i.e.,

extended vacation, sabbatical, illness, etc.) a back-up MSM will be nominated by the study PI and approved by the NINDS Program Official.

Adverse Event Reporting

Consideration of adverse events will hereafter consist of adverse events, serious adverse events, and adverse device effects, including anticipated adverse device effects and unanticipated adverse device effects.

- Adverse event (AE) is defined as any untoward/undesirable clinical occurrence in a clinical investigation of a subject which does not necessarily have a causal relationship with the treatment under investigation. An Adverse Event can therefore be any unfavorable and/or unintended sign, symptom, or disease temporally associated with the use of a device product, whether or not considered related to the device product. Only abnormal laboratory values that are deemed clinically significant by the investigator will be classified as adverse events.

- Serious adverse event (SAE) is defined as any untoward/undesirable adverse experience that results in any of the following outcomes: 1) death; 2) a life-threatening adverse experience; 3) inpatient hospitalization or prolongation of existing hospitalization; 4) a permanent/persistent or significant disability/incapacity or a congenital anomaly/birth defect; 5) important medical events that may not result in death, be life-threatening, or require hospitalization may be considered a serious adverse event when, based upon appropriate medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

- Anticipated adverse device effect (AADE) is defined as any adverse effect related to the device or procedure, which is identified in the protocol or the IFU for the device.

- Unanticipated Adverse Device Effects (UADEs) is defined as any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with a device, if that effect, problem, or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application, or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects.

Reports of UADEs will be made to the FDA within 10 days of receiving notification of the UADE (as required in 21 CFR 812 18p (b)(1)).

Safety Monitoring

The Principal Investigator and the MSM will monitor all AE reports to identify and trend all events that would require temporary discontinuation of study enrollment, to fully

characterize device safety, to modify the study protocol, or to terminate the study.

Reporting Procedures for All Adverse Events

After review with the subject by the study site personnel, all Adverse Events occurring during the study, whether or not attributed to the study and/or the devices, observed by the investigator or reported by the subject, will be documented on the appropriate case record form pages. The following attributes will be assigned by the reporting investigator:

1. Description of event
2. Date of onset
3. Date of resolution (if applicable)
4. Seriousness
5. Relationship to the study device and/or procedure(s)
6. Intensity
7. Action(s) taken
8. Outcome(s)

Intensity is defined as a measure of the severity of a reaction, effect or experience. The measurement(s) are described as mild, moderate or severe. The event itself, however, may be of relative minor medical significance. The intensity of Adverse Events is assessed as mild, moderate or severe according to the following index scale:

- **Mild**
  The Adverse Event is transient, requires no treatment, and does not interfere with the subject's daily activity.
- **Moderate**
  The Adverse Event introduces a low level of inconvenience or concern to the subject and may interfere with daily activities, but is usually ameliorated by simple therapeutic measures.
- **Severe**
  The Adverse Event interrupts the subject's usual daily activity and requires systematic therapy or other treatment.

The relationship of an AE to the study device or procedure will be graded as follows:

- **None** The Adverse Event is not associated with the study device use.*
- **Remote** The temporal association is such that the study device is not likely to have had an association with the observed Adverse Event.*
- **Possible. This causal relationship is assigned when the Adverse Event:**
  a) Follows a reasonable temporal sequence from device use, but
  b) Could have been produced by the subject's clinical state or other modes of therapy administered to the subject.

- **Probable. This causal relationship is assigned when the Adverse Event:**
  a) Follows a reasonable temporal sequence from device use;
  b) Abates upon discontinuation of the treatment;
  c) Cannot be reasonably explained by known characteristics of the subject's clinical state.
- **Definite. This causal relationship is assigned when the Adverse Event:**
  a) Follows a reasonable temporal sequence from device use;
  b) Abates upon discontinuation of the treatment; and
  c) Is confirmed by the reappearance of the Adverse Event on repeat exposure.

*For purposes of reporting UADEs, "None" and "Remote" will be considered as having no association with the device or treatment procedure

Serious Adverse Events All Serious Adverse Events including deaths will be reported to the MSM, the Central Institutional Review Board (IRB) and the FDA, as required.

4. INVESTIGATOR'S AGREEMENT

As this study will be carried out by the NIH StrokeNet, the names of the specific sites and investigators are not yet available. Investigator's agreements for key investigators at the coordinating site are included.

All investigators will be required to sign the following agreement**:**

INVESTIGATOR AGREEMENT
FOR THE CLINICAL INVESTIGATION OF THE DEFUSE 3 TRIAL

I_____ agree to participate as an Investigator on the DEFUSE 3 trial.

I have been provided a copy of the following Food and Drug Administration (FDA) regulations: 21 CFR Part 812, Investigational Device Exemptions; 21 CFR Part 50, Protection of Human Subjects; and 21 CFR Part 54, Financial Disclosure by Clinical Investigators.

I agree and/or certify that:

1. I will conduct the clinical investigation in accordance with this agreement, all requirements of the investigational plan, IDE regulations, other applicable regulations of the FDA, and any conditions of approval imposed by my reviewing Institutional Review Board (IRB) or FDA. I agree to abide by all of the responsibilities of Investigators addressed under 21 CFR Part 812, Subpart E and Subpart G, including but not limited to the following:

2. I will obtain written approval from the authorized IRB for the institution at which this investigation will be conducted.

3. I will ensure that Informed Consent is obtained from each subject participating in this clinical investigation in accordance with the informed consent regulation found in 21 CFR Part 50, and that a signed copy of the informed consent is available to the sponsor (sponsor-investigator) and the sponsor's (sponsor- investigator's) designated monitor.

4. I will ensure the accurate completion of protocol case report forms and, if I am not also the sponsor- investigator of the corresponding IDE application, I will submit completed protocol case report forms to the sponsor (sponsor-investigator) at the time frames specified in the Protocol and/or FDA regulations.

5. I have the appropriate, relevant qualification to conduct and to oversee the conduct of the clinical investigation as documented by the following: (***initial applicable statement***)
   ___ My relevant qualifications, including dates, location, extent and type of experience are listed in my most recent curriculum vitae (CV), which is attached to the Agreement and which will be maintained by the sponsor (sponsor-investigator) of the corresponding IDE application.
   ___ My curriculum vitae (CV) does not reflect my relevant qualifications, therefore attached to this Agreement is a statement of my relevant experience (including dates, location(s), extent and type of experience) which will be maintained by the sponsor (sponsor-investigator) of the corresponding IDE application.

6. There are no reasons to question my ability to oversee the appropriate conduct of this clinical investigation. (***Initial applicable statement***)
   ___ I have never participated in an investigation or other research activity which was terminated (disqualified) by the FDA, IRB (or equivalent), or sponsor of a study due to non-compliance issue.
   ___ I have participated in an investigation or other research activity which was terminated (disqualified) by the FDA, IRB (or equivalent), or sponsor of a study due to non-compliance issue. The specific circumstances leading to this termination and my role in the respective problems or issues and the resolution of these problems or issues are summarized in an attachment to this Agreement.

I further certify that I have not been debarred under the Generic Drug Enforcement Act of 1992, 21 USC §§ 335a and 335b. In the event that I become debarred or receive notice of an action or threat of action with respect to my debarment during the term of this Agreement, I agree to immediately notify the sponsor (sponsor-investigator) and the authorized IRB for my study site. If I am the sponsor-investigator of the corresponding IDE application I will notify the authorized IRB and the FDA.

As required by 21 CFR Part 54, Financial Disclosure by Clinical Investigator, I will disclose sufficient and accurate financial information to the sponsor (sponsor-investigator) by completing the Certification of Financial Interest Form (attached) and if applicable, the Disclosure of Financial Interest Form (attached). I will also notify the sponsor (sponsor-investigator) if my disclosed financial information changes at any time during the clinical investigation or up to one year following the closure of the study/

Site Name and Address:
_____
_____

_____
_____ _____
Investigator Signature Date

5. EXECUTIVE COMMITTEE / KEY PARTICIPATING INVESTIGATORS

The Executive committee, composed of experts in vascular neurology, endovascular therapy and neuroimaging, will provide the overall scientific guidance for the study. The committee will meet monthly by phone (1 hour/month) for the full duration of the study. Responsibilities include oversight of the overall conduct of the study with regard to protocol compliance and modifications/amendments, study progress, and problem-solving. Dr. Albers will chair the executive committee.

Key Participating Investigators at Coordinating Site

Gregory W. Albers, MD Principal Investigator Stanford Stroke Center 780 Welch Rd. Suite 350 Palo Alto, CA 94305 650-723-4448 galbers@stanfordmed.org	Michael Marks, MD Co-Principal Investigator Stanford University Medical Center 300 Pasteur Dr. Stanford, CA 94305-5105 650-723-6767 mmarks@stanford.edu	Maarten Lansberg, MD, PhD Protocol Director Stanford Stroke Center 780 Welch Rd. Suite 350 Palo Alto, CA 94305 650-723-4448 lansberg@stanford.edu

6. INSTITUTIONAL REVIEW BOARD

The University of Cincinnati Institutional Review Board will serve as the National Central Institutional Review Board for all participating sites. The **Central Institutional Review Board (CIRB)** for multicenter protocols is the single IRB of record. It has regulatory responsibility for assuring the protection of the rights and welfare of research participants in accordance with Standard Operating Procedure ADM 12; Central Institutional Review Board Reporting. The National Institute of Neurological Disorders and Stroke (NINDS) selected the University of Cincinnati Institutional Review Board (IRB) to serve as the CIRB for the NIH StrokeNet (StrokeNet).

University of Cincinnati IRB Registration # 00000180 FWA #: 00003152 Expiration Date: 6/27/2016

Michael Linke, PhD, CIP
Chairman CIRB
Michael.linke@va.gov
513-304-3540

7. COSTS

All of the eligible devices that will be used in this study are currently on the market. There will be no charges beyond the typical standard of care for use of these approved devices. These devices will be used and billed according to the standard of care for each institution.

8. REFERENCES

1. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: Systematic analysis of population health data. *The Lancet*. 2006;367:1747-1757
2. Kleindorfer D, Lindsell CJ, Brass L, Koroshetz W, Broderick JP. National us estimates of recombinant tissue plasminogen activator use: Icd-9 codes substantially underestimate. *Stroke*. 2008;39:924-928
3. Katzan IL, Hammer MD, Hixson ED, Furlan AJ, Abou-Chebl A, Nadzam DM. Utilization of intravenous tissue plasminogen activator for acute ischemic stroke. *Arch Neurol*. 2004;61:346-350
4. Qureshi AI, Kirmani JF, Sayed MA, Safdar A, Ahmed S, Ferguson R, et al. Time to hospital arrival, use of thrombolytics, and in-hospital outcomes in ischemic stroke. *Neurology*. 2005;64:2115-2120
5. Saqqur M, Uchino K, Demchuk AM, Molina CA, Garami Z, Calleja S, et al. Site of arterial occlusion identified by transcranial doppler predicts the response to intravenous thrombolysis for stroke. *Stroke*. 2007;38:948-954
6. Mori E, Yoneda Y, Tabuchi M, Yoshida T, Ohkawa S, Ohsumi Y, et al. Intravenous recombinant tissue plasminogen activator in acute carotid artery territory stroke. *Neurology*. 1992;42:976-982
7. del Zoppo GJ, Poeck K, Pessin MS, Wolpert SM, Furlan AJ, Ferbert A, et al. Recombinant tissue plasminogen activator in acute thrombotic and embolic stroke. *Ann Neurol*. 1992;32:78-86
8. Albers GW, Thijs VN, Wechsler L, Kemp S, Schlaug G, Skalabrin E, et al. Magnetic resonance imaging profiles predict clinical response to early reperfusion: The diffusion and perfusion imaging evaluation for understanding stroke evolution (defuse) study. *Annals of Neurology*. 2006;60:508-517
9. Lansberg MG, Schrooten M, Bluhmki E, Thijs VN, Saver JL. Treatment time-specific number needed to treat estimates for tissue plasminogen activator therapy in acute stroke based on shifts over the entire range of the modified rankin scale. *Stroke*. 2009;40:2079-2084
10. Fields JD, Lindsay K, Liu Kenneth C, Nesbit GM, Lutsep HL. Mechanical thrombectomy for the treatment of acute ischemic stroke. *Expert Review of Cardiovascular Therapy*. 2010;8:581-592
11. Lee M, Hong K-S, Saver JL. Efficacy of intra-arterial fibrinolysis for acute ischemic stroke: Meta-analysis of randomized controlled trials. *Stroke*. 2010;41:932-937
12. Smith WS, Sung G, Saver J, Budzik R, Duckwiler G, Liebeskind DS, et al. Mechanical thrombectomy for acute ischemic stroke: Final results of the multi merci trial. *Stroke*. 2008;39:1205-1212
13. Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, et al. Intra-arterial prourokinase for acute ischemic stroke. The proact ii study: A randomized controlled trial. Prolyse in acute cerebral thromboembolism. *Jama*. 1999;282:2003-2011
14. Thomalla G, Schwark C, Sobesky J, Bluhmki E, Fiebach JB, Fiehler J, et al. Outcome and symptomatic bleeding complications of intravenous thrombolysis within 6 hours in mri-selected stroke patients: Comparison of a german multicenter study with the pooled data of atlantis, ecass, and ninds tpa trials. *Stroke*. 2006;37:852-858
15. Hsia A, Kidwell C. Developments in neuroimaging for acute ischemic stroke: Diagnostic and clinical trial applications. *Current Atherosclerosis Reports*. 2008;10:339-346

16. Kakuda W, Lansberg MG, Thijs VN, Kemp SM, Bammer R, Wechsler LR, et al. Optimal definition for pwi/dwi mismatch in acute ischemic stroke patients. *J Cereb Blood Flow Metab*. 2008;28:887-891
17. Lansberg MG, Thijs V, Bammer R, Kakuda W, Hamilton S, Wechsler L, et al. Clinical and mri-based risk factors for symptomatic intracerebral hemorrhage following treatment with tissue plasminogen activator. *Stroke*. 2006;37:654
18. Saver JL, Jahan R, Levy EI, Jovin TG, Baxter B, Nogueira RG, et al. Solitaire flow restoration device versus the merci retriever in patients with acute ischaemic stroke (swift): A randomised, parallel-group, non-inferiority trial. *Lancet*. 2012;380:1241-1249
19. Nogueira RG, Lutsep HL, Gupta R, Jovin TG, Albers GW, Walker GA, et al. Trevo versus merci retrievers for thrombectomy revascularisation of large vessel occlusions in acute ischaemic stroke (trevo 2): A randomised trial. *Lancet*. 2012;380:1231-1240
20. Marks MP, Lansberg MG, Mlynash M, Kemp S, McTaggart RA, Zaharchuk G, et al. Angiographic outcome of endovascular stroke therapy correlated with mr findings, infarct growth, and clinical outcome in the defuse 2 trial. *Int J Stroke*. 2014
21. Yoo AJ, Simonsen CZ, Prabhakaran S, Chaudhry ZA, Issa MA, Fugate JE, et al. Refining angiographic biomarkers of revascularization: Improving outcome prediction after intra-arterial therapy. *Stroke*. 2013;44:2509-2512
22. Zaidat OO, Yoo AJ, Khatri P, Tomsick TA, von Kummer R, Saver JL, et al. Recommendations on angiographic revascularization grading standards for acute ischemic stroke: A consensus statement. *Stroke*. 2013;44:2650-2663
23. Broderick JP, Palesch YY, Demchuk AM, Yeatts SD, Khatri P, Hill MD, et al. Endovascular therapy after intravenous t-pa versus t-pa alone for stroke. *N Engl J Med*. 2013;368:893-903
24. Hill MD, Demchuk AM, Goyal M, Jovin TG, Foster LD, Tomsick TA, et al. Alberta stroke program early computed tomography score to select patients for endovascular treatment: Interventional management of stroke (ims)-iii trial. *Stroke*. 2014;45:444-449
25. Kidwell CS, Jahan R, Gornbein J, Alger JR, Nenov V, Ajani Z, et al. A trial of imaging selection and endovascular treatment for ischemic stroke. *N Engl J Med*. 2013;368:914-923
26. O.A. Berkhemer, P.S.S. Fransen, D. Beumer, et al. A Randomized Trial of Intraarterial Treatment for AcuteIschemic Stroke. *N Engl J Med* 2015;372:11-20.
27. Muir KW, Buchan A, von Kummer R, Rother J, Baron J-C. Imaging of acute stroke. *The Lancet Neurology*. 2006;5:755-768
28. Campbell B, Purushotham A, Christensen S, Desmond P, Nagakane Y, Parsons M, et al. The infarct core is well represented by the acute diffusion lesion: Sustained reversal is infrequent. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2012;32:50-56
29. Chemmanam T, Campbell BC, Christensen S, Nagakane Y, Desmond PM, Bladin CF, et al. Ischemic diffusion lesion reversal is uncommon and rarely alters perfusion-diffusion mismatch. *Neurology*. 2010;75:1040-1047
30. Baird AE, Warach S. Magnetic resonance imaging of acute stroke. *J Cereb Blood Flow Metab*. 1998;18:583-609
31. Dani KA, Thomas RG, Chappell FM, Shuler K, MacLeod MJ, Muir KW, et al. Computed tomography and magnetic resonance perfusion imaging in ischemic stroke: Definitions and thresholds. *Ann Neurol*. 2011;70:384-401

32. Olivot JM, Mlynash M, Thijs VN, Kemp S, Lansberg MG, Wechsler L, et al. Optimal tmax threshold for predicting penumbral tissue in acute stroke. *Stroke*. 2009;40:469-475
33. Zaro-Weber O, Moeller-Hartmann W, Heiss WD, Sobesky J. Maps of time to maximum and time to peak for mismatch definition in clinical stroke studies validated with positron emission tomography. *Stroke*. 2010;41:2817-2821
34. Olivot JM, Mlynash M, Zaharchuk G, Straka M, Bammer R, Schwartz N, et al. Perfusion mri (tmax and mtt) correlation with xenon ct cerebral blood flow in stroke patients. *Neurology*. 2009;72:1140-1145
35. Mlynash M, Lansberg MG, De Silva DA, Lee J, Christensen S, Straka M, et al. Refining the definition of the malignant profile: Insights from the defuse-epithet pooled data set. *Stroke; a journal of cerebral circulation*. 2011;42:1270-1275
36. Lansberg MG, Lee J, Christensen S, Straka M, De Silva DA, Mlynash M, et al. Rapid automated patient selection for reperfusion therapy: A pooled analysis of the echoplanar imaging thrombolytic evaluation trial (epithet) and the diffusion and perfusion imaging evaluation for understanding stroke evolution (defuse) study. *Stroke*. 2011;42:1608-1614
37. Lansberg MG, Straka M, Kemp S, Mlynash M, Wechsler LR, Jovin TG, et al. Mri profile and response to endovascular reperfusion after stroke (defuse 2): A prospective cohort study. *Lancet Neurology*. 2012;11:860-867
38. Lansberg MG, Cereda CW, Mlynash M, Mishra NK, Inoue M, Kemp S, et al. Response to reperfusion not time-dependent in patients with salvageable tissue. Neurology, in press 2015
39. Inoue M, Mlynash M, Christensen S, Wheeler HM, Straka M, Tipirneni A, et al. Early diffusion-weighted imaging reversal after endovascular reperfusion is typically transient in patients imaged 3 to 6 hours after onset. *Stroke*. 2014;45:1024-1028
40. Wheeler HM, Mlynash M, Inoue M, Tipirneni A, Liggins J, Zaharchuk G, et al. Early diffusion-weighted imaging and perfusion-weighted imaging lesion volumes forecast final infarct size in defuse 2. *Stroke*. 2013;44:681-685
41. Parsons MW, Christensen S, McElduff P, Levi CR, Butcher KS, De Silva DA, et al. Pretreatment diffusion- and perfusion-mr lesion volumes have a crucial influence on clinical response to stroke thrombolysis. *J Cereb Blood Flow Metab*. 2010;30:1214-1225
42. Marks MP, Olivot JM, Kemp S, Lansberg MG, Bammer R, Wechsler LR, et al. Patients with acute stroke treated with intravenous tpa 3-6 hours after stroke onset: Correlations between mr angiography findings and perfusion- and diffusion-weighted imaging in the defuse study. *Radiology*. 2008;249:614-623
43. Weisstanner C, Gratz PP, Schroth G, Verma RK, Kochl A, Jung S, et al. Thrombus imaging in acute stroke: Correlation of thrombus length on susceptibility-weighted imaging with endovascular reperfusion success. *Eur Radiol*. 2014
44. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*. 1975;31:103-115
45. Chimowitz MI, Lynn MJ, Turan TN, Fiorella D, Lane BF, Janis S, et al. Design of the stenting and aggressive medical management for preventing recurrent stroke in intracranial stenosis trial. *J Stroke Cerebrovasc Dis*. 2011;20:357-368
46. Pereira VM, Gralla J, Davalos A, Bonafe A, Castano C, Chapot R, et al. Prospective, multicenter, single-arm study of mechanical thrombectomy using solitaire flow restoration in acute ischemic stroke. *Stroke*. 2013;44:2802-2807

47. Molina CA, Montaner J, Abilleira S, Ibarra B, Romero F, Arenillas JF, et al. Timing of spontaneous recanalization and risk of hemorrhagic transformation in acute cardioembolic stroke. *Stroke*. 2001;32:1079-1084
48. Wolpert SM, Bruckmann H, Greenlee R, Wechsler L, Pessin MS, del Zoppo GJ. Neuroradiologic evaluation of patients with acute stroke treated with recombinant tissue plasminogen activator. The rt-pa acute stroke study group. *AJNR Am J Neuroradiol*. 1993;14:3-13
49. Alexandrov AV, Molina CA, Grotta JC, Garami Z, Ford SR, Alvarez-Sabin J, et al. Ultrasound-enhanced systemic thrombolysis for acute ischemic stroke. *N Engl J Med*. 2004;351:2170-2178
50. Rubin D. *Multiple imputation for nonresponse in surveys*. Wiley & Sons; 1987.
51. Campbell BC, Tu HT, Christensen S, Desmond PM, Levi CR, Bladin CF, et al. Assessing response to stroke thrombolysis: Validation of 24-hour multimodal magnetic resonance imaging. *Arch Neurol*. 2012;69:46-50
52. Tomsick T, Broderick J, Carrozella J, Khatri P, Hill M, Palesch Y, et al. Revascularization results in the interventional management of stroke ii trial. *AJNR Am J Neuroradiol*. 2008;29:582-587
53. Christou I, Alexandrov AV, Burgin WS, Wojner AW, Felberg RA, Malkoff M, et al. Timing of recanalization after tissue plasminogen activator therapy determined by transcranial doppler correlates with clinical recovery from ischemic stroke. *Stroke*. 2000;31:1812-1816
54. Nguyen TN, Malisch T, Castonguay AC, Gupta R, Sun CH, Martin CO, et al. Balloon guide catheter improves revascularization and clinical outcomes with the solitaire device: Analysis of the north american solitaire acute stroke registry. *Stroke*. 2014;45:141-145
55. San Roman L, Obach V, Blasco J, Macho J, Lopez A, Urra X, et al. Single-center experience of cerebral artery thrombectomy using the trevo device in 60 patients with acute ischemic stroke. *Stroke*. 2012;43:1657-1659
56. Turk AS, Frei D, Fiorella D, Mocco J, Baxter B, Siddiqui A, et al. Adapt fast study: A direct aspiration first pass technique for acute stroke thrombectomy. *J Neurointerv Surg*. 2014;6:260-264
57. Bang OY, Saver JL, Lee KH, Kim GM, Chung CS, Kim SJ, et al. Characteristics of patients with target magnetic resonance mismatch profile: Data from two geographically and racially distinct populations. *Cerebrovasc Dis*. 2010;29:87-94
58. Hacke W, Furlan AJ, Al-Rawi Y, Davalos A, Fiebach JB, Gruber F, et al. Intravenous desmoteplase in patients with acute ischaemic stroke selected by mri perfusion-diffusion weighted imaging or perfusion ct (dias-2): A prospective, randomised, double-blind, placebo-controlled study. *Lancet. Neurology*. 2009;8:141-150
59. Lansberg MG, Thijs VN, Bammer R, Olivot JM, Marks MP, Wechsler LR, et al. The mra-dwi mismatch identifies patients with stroke who are likely to benefit from reperfusion. *Stroke*. 2008;39:2491-2496
60. Mishra NK, Albers GW, Christensen S, Marks M, Hamilton S, Straka M, et al. Comparison of magnetic resonance imaging mismatch criteria to select patients for endovascular stroke therapy. *Stroke*. 2014;45:1369-1374

APPENDIX

I. DEFUSE 3 Patient Informed Consent Form

PROTOCOL TITLE

DEFUSE 3:
Endovascular Therapy Following Imaging Evaluation for Ischemic Stroke 3

Protocol Version/Version Date

Version 2.4

April 20, 2017

Protocol Directors

Gregory Albers, MD
Michael Marks, MD
Maarten Lansberg, MD, PhD

Stanford University
Stanford Stroke Center
780 Welch Rd. Suite CJ350
Stanford, CA 94304

Supported by

The National Institute of Neurological
Disorders and Stroke (NINDS)
U01 NS092076

IDE Number

G150028

DEFUSE 3

AGREEMENT ON THE PROTOCOL

By signing below I confirm that:

1) I have read this protocol and it contains all necessary details for conducting this study

   AND

2) I agree to conduct the trial in compliance with this protocol and to adhere to all regulations that govern the conduct of the study.

_____ _____

Principal Investigator's Signature **Date**

_____

Principal Investigator's Name

_____

Site Name:

TABLE OF CONTENTS

1. SUMMARY OF TRIAL

DEFUSE 3 is a prospective, randomized, multi-center, Phase III, adaptive, blinded endpoint, controlled trial. A maximum of 476 patients will be randomized and treated between 6 and 16 hours of symptom onset. Subjects will be randomized 1:1 to endovascular therapy plus medical management or medical management alone. *Only the devices listed in this protocol are approved for use in DEFUSE 3. The choice of device or devices employed is at the discretion of the clinical investigator.*

2. SCIENTIFIC BACKGROUND

Although stroke is the number one cause of adult disability in the United States[1], treatment options for stroke are limited. The only FDA approved treatment for stroke is administration of intravenous (iv) tissue plasminogen activator (tPA) within 3 hours after symptom onset. Nationwide, only about 4 percent of stroke patients receive this therapy.[2] The main reason for this low treatment rate is that most patients present to the hospital outside the time-window for tPA.[3, 4] Even when administered, tPA is often not effective because it either fails to recanalize the occluded artery[5-7] or because the brain is already irreversibly injured.[8] As a result, it is estimated that only 12-25% of treated patients benefit from tPA.[9] Thus, in order to improve outcomes from stroke, we need better treatments that are available to a greater proportion of stroke patients.

2.1 State of the science on endovascular stroke therapy for acute stroke

Endovascular stroke therapy, the removal of blood clots with mechanical devices or thrombolytic drugs administered intra-arterially, is the most promising new treatment for patients who "fail" treatment with iv tPA or are not eligible for iv tPA. The main advantage of endovascular therapy is that it has a high rate of recanalization.[10, 11] Blood flow can be restored with a success rate of up to 82% with modern thrombectomy devices and 66% for intra-arterial thrombolysis.[10, 12, 13] This is approximately twice as effective as iv tPA which has a recanalization rate of 10-50% depending on the location of the blood clot.[6, 7] Despite higher rates of recanalization with endovascular therapy, two recent randomized controlled trials of endovascular therapy, IMS III and MR RESCUE, have failed to demonstrate a clinical benefit. Patient and treatment related factors likely both contributed to the neutral results of these trials (see below for details).

Patient-related factors: A central consideration in the optimization of patient selection for acute stroke therapies is the concept of the ischemic penumbra. Ischemic penumbra is defined as ischemic tissue that is potentially salvageable and is distinguished from the ischemic core that has already sustained irreversible injury. Clearly, the target of acute stroke therapies is salvage of the ischemic penumbra, preventing infarct growth and, most importantly, improved functional outcome. Acute stroke trials should therefore ideally be limited to patients with an ischemic penumbra. MRI-based studies, such as DEFUSE 1 and 2, indicate that MRI can be used to identify these patients.[8, 14, 15 16, 17]

Treatment-related factors: Recent studies have emphasized the importance of recanalization rates, demonstrating the influence on patient outcome of highly effective endovascular procedures that lead to complete or near-complete reperfusion.[18, 19] Currently, the most common metric for rating the quality of reperfusion is the modified Thrombolysis

In Cerebral Infarction (mTICI) scale, and a clear relationship exists between the degree of reperfusion on the mTICI scale and patient outcome.[20-22] Patients with >50% reperfusion (mTICI 2B-3) are much more likely to have a good outcome than patients with <50% reperfusion.

IMS III, the largest endovascular trial to date, did not use advanced imaging criteria to select patients.[23] Instead it used relatively strict time-criteria, anticipating that this would yield a high proportion of patients with a substantial penumbra. However, several categories of patients who likely did not have substantial penumbra were enrolled. Nearly one third of the patients did not have a vessel occlusion at angiography and 23% had distal MCA occlusions; both of these subgroups are unlikely to have substantial penumbral tissue. In addition, 42% of patients had some evidence of irreversible tissue injury (ASPECTS <8) on their baseline CT and 14% had evidence of extensive irreversible injury (ASPECTS <5).[24] Finally, the endovascular devices that were available during IMS III had relatively low rates of early reperfusion; mTICI 2B-3 was only 40% in IMS III.

MR RESCUE, a stroke trial that aimed to demonstrate benefit of endovascular therapy in patients with a penumbra based on MRI, had neutral results.[25] Several factors likely contributed to this. First, the rate of endovascular reperfusion was extremely low. Only 8 patients (24%) in the MR RESCUE penumbral group achieved TICI 2B-3 reperfusion during endovascular therapy. Second, patients in the penumbral group in MR RESCUE had larger baseline infarct core lesions (median volume 36 ml; IQR 24–51 ml) than the Target Mismatch patients in DEFUSE 2 (median volume 13 ml; IQR 5–26 ml). The combination of low rates of endovascular reperfusion and relatively large core lesions, both strong predictors of poor clinical outcome, likely explains the lack of a treatment effect in "penumbral patients" in MR RESCUE. Moreover, with only 8 endovascular patients in the penumbral cohort achieving TICI 2B-3 reperfusion, MR RESCUE was substantially underpowered.

New Generation Trials: Recently, a series of positive randomized studies of endovascular therapy with treatment initiated within 6 hours of stroke onset in the vast majority of patients were reported. This has prompted new guidelines endorsing endovascular therapy up to 6 hours after symptom onset. The American Heart Association is now calling for late window studies using advanced imaging for patient selection: ""Further randomized, controlled trials should be done to determine whether advanced imaging paradigms using CT perfusion and MRI perfusion, CTA, and diffusion imaging, including measures of infarct core, collateral flow status, and penumbra, are beneficial for selecting patients for acute reperfusion therapy who are beyond 6 hours from symptom onset. (New recommendation, 2015 AHA Guidelines).

DEFUSE 3 will address this new mandate. Enrollment is limited to patients with salvageable tissue (Target Mismatch patients) who are likely to respond favorably to endovascular reperfusion in the 6-16 hour window. Use of the latest generation thrombectomy devices, coupled with strict qualification and oversight criteria for the neuro-interventionalists, will result in high rates of reperfusion. Based on the compelling preliminary data from DEFUSE 2, the trial is adequately powered to demonstrate a clear treatment effect.

2.2. Prior studies and rationale for development

This protocol aims to shift the selection of patients for reperfusion therapy from a relatively arbitrary decision based on poorly validated clinical characteristics to an objective decision based on scientific evidence. Many factors affect the evolution of the ischemic penumbra into the ischemic core, and the rate of progression of irreversible injury is highly variable between individuals. This variability is likely mediated by the adequacy of collateral blood flow as well as the metabolic milieu of individual stroke patients. The individuality of penumbral evolution among stroke patients implies that identifying the extent of the ischemic core and penumbra is useful for making treatment decisions. Currently diffusion-weighted imaging (DWI) / perfusion-weighted imaging (PWI) magnetic resonance imaging (MRI) affords the best opportunity for approximating the ischemic core and penumbra in real time clinical practice.[27]

The DWI lesion provides a dependable estimation of the ischemic core and only very rarely shows permanent reversal following early reperfusion.[28, 29] PWI identifies hypoperfused ischemic tissue. Regions defined as abnormal on PWI that do not demonstrate a DWI abnormality, often referred to as the DWI/PWI mismatch, can estimate the ischemic penumbra.[30] It is critical that PWI utilizes an appropriately validated threshold parameter that excludes ischemic tissue with modest blood flow reduction (i.e. benign oligemia), because this tissue is unlikely to infarct even if reperfusion does not occur. Which PWI parameter is optimal, as well as what threshold to use to define critical hypoperfusion, has been the focus of multiple research efforts.[31] Prior work from our group and others supports the use of Tmax, thresholded at >6 seconds, as the optimal PWI parameter to identify ischemic tissue destined to become infarcted if timely reperfusion does not occur.[32] This Tmax threshold correlates well with the penumbral range of cerebral blood flow decline as determined by both positron emission tomography and Xenon CT.[33, 34]

Using a difference between the volume of the baseline PWI Tmax lesion and the DWI volume to identify mismatch, the DEFUSE and EPITHET studies found that most patients with a PWI/DWI mismatch responded favorably if reperfusion occurred following iv tPA treatment in the 3- to 6-hour time window. However, despite having a mismatch, patients with very large baseline DWI lesions (large core infarct volumes) had highly unfavorable outcomes following reperfusion. Patients with this MRI pattern, referred to as the Malignant profile, had a significantly higher rate of both parenchymal hemorrhage and severe disability/death if reperfusion occurred.[35] Mismatch patients who do not have the Malignant profile have been designated as having a Target Mismatch, and these patents respond extremely favorably to reperfusion following iv tPA therapy. In a pooled analysis of DEFUSE and EPITHET, Target Mismatch profile patients who experienced reperfusion had a 5-fold increase in favorable clinical response at 90 days and significantly less infarct growth when compared to those who did not reperfuse.[36] No association between reperfusion and favorable outcomes, or a reduction in infarct growth, was apparent for patients without the mismatch profile.

The DEFUSE 2 study utilized an automated mismatch analysis program (RAPID) to prospectively establish MRI profiles in a consecutive cohort of patients who then underwent endovascular therapy. DEFUSE 2 confirmed the concepts demonstrated in DEFUSE and EPITHET; Target Mismatch patients who achieve early reperfusion therapy have less infarct growth and more favorable clinical outcomes (8·8,95% CI 2·7–29·0).[37] No association between reperfusion and favorable outcomes or infarct growth was present in patients

without Target Mismatch. Furthermore, the positive association between reperfusion, favorable clinical response, and attenuation of infarct growth did not diminish in DEFUSE 2 patients with Target Mismatch who were treated relatively late (6-12 hours after symptom onset, see **Figure 1**).

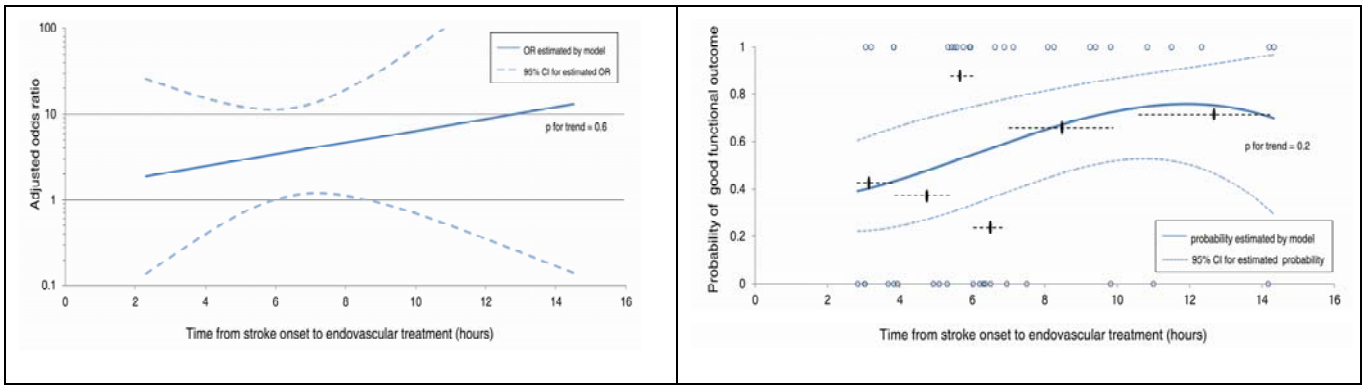

Figure 1: The effect of time on the association between reperfusion and good functional outcome (graph on the left) and the effect of time to treatment on the probability of good functional outcome patients with reperfusion (graph on the right) in Target mismatch patients. 95% CIs are indicated by dashed lines. Estimates are based on multivariate logistic regression, adjusted for age and baseline DWI volume. There is no significant effect of time in either model.[38]

This finding contrasts sharply with prior studies that did not use penumbral imaging to select patients and suggests that imaging findings may be of equal, or potentially even greater, importance than time from symptom onset for identification of optimal patients who might benefit from reperfusion therapy.

How could Target Mismatch patients who are treated late have outcomes that are as favorable as those of earlier treated patients? At later time points, the Target Mismatch profile identifies patients in whom the infarct is evolving at a relatively slow rate; a DWI lesion that is still considerably smaller than the PWI lesion reflects good collateral circulation. These collaterals typically allow prolonged, but not permanent, survival of the hypoperfused mismatch region. Evidence that the mismatch region is still at considerable risk for infarct expansion, even at later time points, was provided by the DEFUSE 2 finding that Target Mismatch patients imaged between 6 and 12 hours from symptom onset consistently demonstrated substantial infarct growth if reperfusion was not achieved.[38] Patients with slowly evolving infarct cores are ideal candidates for later time window reperfusion therapy, particularly endovascular therapies. One of the drawbacks of the endovascular approach is that the time between hospital arrival and achievement of endovascular reperfusion is typically at least 90 to 120 minutes. For patients with rapidly growing infarct cores (such as patients with the Malignant profile), substantial growth of the infarct core has been reported despite endovascular reperfusion. Therefore, removing this population, which represents about 20% of eligible patients, from a randomized endovascular trial, has important advantages.

DEFUSE 2 confirmed that early DWI lesions are an excellent surrogate for the ischemic core. Despite endovascular therapy, only 2 patients had a final infarct that was smaller than the baseline DWI lesion and the size and location of the early DWI lesion was a reliable predictor of the final infarct volume in patients with complete reperfusion.[39, 40] In DEFUSE 2, younger age and smaller DWI volume were significant independent predictors of favorable outcome. Subgroup analysis of EPITHET identified DWI lesion size ≤25 ml as a strong predictor of a

favorable response to reperfusion.[37, 41] These findings suggest that certain subgroups, in particular individuals with Target mismatch and small DWI lesions are most likely to benefit from reperfusion. The adaptive design of DEFUSE 3 (see below) has the potential to focus patient enrollment on a subgroup of patients (e.g. those with smaller DWI lesion volumes and/or shorter times from symptom onset to randomization) who respond most favorably to endovascular therapy. This will allow the study to identify the largest population that has a statistically reliable benefit of endovascular therapy.

New data (presented at the International Stroke Conference, February 2015) suggest that CT Perfusion studies, processed with the same software (RAPID) used in the studies described above, can identify the ischemic core with accuracy similar to MRI (Cereda, et al ISC 2015) and select patients who respond to endovascular reperfusion therapy in early time windows (Campbell, et al NEJM 2015, EXTEND-IA study, Saver, et al NEJM 2015, SWIFT PRIME study). Therefore, DEFUSE 3 will allow patient selection with both MRI and CT Perfusion.

3. INVESTIGATIONAL PLAN

3.1. Purpose

DEFUSE 3 is a prospective randomized Phase III multicenter controlled trial of patients with acute ischemic anterior circulation strokes due to large artery occlusion treated between 6-16 hours of stroke onset with endovascular thrombectomy therapy vs. control. The primary endpoint, the modified Rankin Scale, will be assessed at 3 months. The patients' participation in the study concludes at that time (3 months from stroke onset). The study will randomize up to 476 patients over 4 years. The purpose of DEFUSE 3 is to assess the safety and efficacy of thrombectomy in carefully selected patients in an extended time window. Only the devices listed in this protocol will be used. Selection of the specific device (or devices) is determined by the individual endovascular therapist.

3.2. Protocol Design

DEFUSE 3 is a prospective randomized Phase III multicenter controlled trial of patients with acute ischemic anterior circulation strokes due to large artery occlusion treated between 6-16 hours of stroke onset. Patients who meet the inclusion criteria will undergo either CT Perfusion/CTA or MR DWI/PWI/MRA studies prior to randomization. Patients who have evidence of an ICA or MCA M1 occlusion and a Target Mismatch Profile will be randomized in a 1:1 ratio to treatment with endovascular therapy (using one or more DEFUSE 3 approved thrombectomy devices) plus standard medical therapy versus standard medical therapy alone. Patients who are consented, but not randomized, will receive standard therapy according to local guidelines. Baseline data, and information about early stroke therapies, will be captured for this group of patients.

Randomization of a maximum of 476 patients is planned. At the first interim analysis when 200 subjects complete follow-up, if the overall analysis crosses the futility boundary, a novel adaptive design will identify, if it exists, a subgroup with the best prospect for showing benefit from endovascular treatment, based on baseline ischemic core lesion volumes and the time to treatment. The second interim analyses will be conducted at 340 patients at which time the study may stop for efficacy/futility, or the inclusion criteria may be adjusted in the case of futility.

Approximately 45 sites will be chosen. Individual site selection will be based on a number of factors including endovascular volume, MRI and/or CT perfusion access, number of competing trials, clinical trial experience, and the diversity of their patient population. If a site does not consent a patient within 4 months of activation, it will be placed on probation. If no patient consent occurs in the next 2 months, the site will be replaced with a "back-up" site.

3.3. Enrollment Criteria

3.3.1. Clinical Inclusion Criteria:

1. Signs and symptoms consistent with the diagnosis of an acute anterior circulation ischemic stroke
2. Age 18-90 years
3. Baseline NIHSSS is ≥ 6 and remains ≥6 immediately prior to randomization
4. Endovascular treatment can be initiated (femoral puncture) between 6 and 16 hours of stroke onset. Stroke onset is defined as the time the patient was last known to be at their neurologic baseline (wake-up strokes are eligible if they meet the above time limits).
5. modified Rankin Scale less than or equal to 2 prior to qualifying stroke (functionally independent for all ADLs)
6. Patient/Legally Authorized Representative has signed the Informed Consent form.

3.3.2. Clinical Exclusion Criteria:

1. Other serious, advanced, or terminal illness (investigator judgment) or life expectancy is less than 6 months.
2. Pre-existing medical, neurological or psychiatric disease that would confound the neurological or functional evaluations
3. Pregnant
4. Unable to undergo a contrast brain perfusion scan with either MRI or CT
5. Known allergy to iodine that precludes an endovascular procedure
6. Treated with tPA >4.5 hours after time last known well
7. Treated with tPA 3-4.5 hours after last known well AND any of the following: age >80, current anticoagulant use, history of diabetes AND prior stroke, NIHSS >25
8. Known hereditary or acquired hemorrhagic diathesis, coagulation factor deficiency; recent oral anticoagulant therapy with INR > 3 (recent use of one of the new oral anticoagulants is not an exclusion if estimated GFR > 30 ml/min).
9. Seizures at stroke onset if it precludes obtaining an accurate baseline NIHSS
10. Baseline blood glucose of <50mg/dL (2.78 mmol) or >400mg/dL (22.20 mmol)
11. Baseline platelet count < 50,000/uL
12. Severe, sustained hypertension (Systolic Blood Pressure >185 mmHg or Diastolic Blood Pressure >110 mmHg)
13. Current participation in another investigational drug or device study
14. Presumed septic embolus; suspicion of bacterial endocarditis
15. Clot retrieval attempted using a neurothrombectomy device prior to 6 hours from symptom onset
16. Any other condition that, in the opinion of the investigator, precludes an endovascular procedure or poses a significant hazard to the subject if an endovascular procedure was performed.

3.3.3. Neuroimaging Inclusion Criteria:

1. ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA or CTA

AND

2. Target Mismatch Profile on CT perfusion or MRI (ischemic core volume is < 70 ml, mismatch ratio is ≥ 1.8 and mismatch volume* is ≥ 15 ml)

Notes: The mismatch volume is determined by the RAPID software in real time based on the difference between the ischemic core lesion volume and the Tmax>6s lesion volume. If both a CT perfusion and a multimodal MRI scan are performed prior to enrollment, the later of the 2 scans is assessed to determine eligibility. Only an intracranial MRA is required for patients screened with MRA; cervical MRA is not required. Cervical and intracranial CTA are typically obtained simultaneously in patients screened with CTA, but only the intracranial CTA is required for enrollment.

Alternative neuroimaging inclusion criteria (if perfusion imaging or CTA/MRA is technically inadequate):

A) If CTA (or MRA) is technically inadequate:

Tmax>6s perfusion deficit consistent with an ICA or MCA-M1 occlusion

AND

Target Mismatch Profile (ischemic core volume is < 70 ml, mismatch ratio is ≥1.8 and mismatch volume is ≥15 ml as determined by RAPID software)

B) If MRP is technically inadequate:

ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA (or CTA, if MRA is technically inadequate and a CTA was performed within 60 minutes prior to the MRI)

AND

DWI lesion volume < 25 ml

C) If CTP is technically inadequate:

Patient can be screened with MRI and randomized if neuroimaging criteria are met.

3.3.4. Neuroimaging Exclusion Criteria:

1. ASPECT score <6 on non-contrast CT (if patient is enrolled based on CT perfusion criteria)
2. Evidence of intracranial tumor (except small meningioma) acute intracranial hemorrhage, neoplasm, or arteriovenous malformation
3. Significant mass effect with midline shift
4. Evidence of internal carotid artery dissection that is flow limiting or aortic dissection
5. Intracranial stent implanted in the same vascular territory that precludes the safe deployment/removal of the neurothrombectomy device
6. Acute symptomatic arterial occlusions in more than one vascular territory confirmed on CTA/MRA (e.g., bilateral MCA occlusions, or an MCA and a basilar artery occlusion).

3.4. Enrollment and Randomization

3.4.1 Enrollment: All patients who meet the clinical criteria listed above are eligible for

enrollment. This includes both patients who are directly admitted to the study site and patients who are transferred from an outside hospital. The time of enrollment is the time when the informed consent is signed. After obtaining consent the RAPID output from a CT perfusion or multimodal MRI scan will be assessed. If the patient is confirmed to meet the neuroimaging eligibility criteria listed above and not have any of the neuroimaging exclusion criteria, then the patient will be randomized. In general, patients will be consented prior to obtaining the RAPID output maps. In some situations the CT perfusion/ multimodal MRI may have been performed as part of standard care *prior* to the patient being assessed for study eligibility. Patients who are consented but do not meet the imaging criteria will not be randomized.

Determination of Target Mismatch and Large Artery Occlusion: At the conclusion of the MRI or CT Perfusion scan, the technologist sends the sequences from the console to RAPID with a single mouse click for automated processing. The RAPID software was developed based on data from DEFUSE 1 and was prospectively validated in DEFUSE 2. The system provides fully automated processing of brain images.

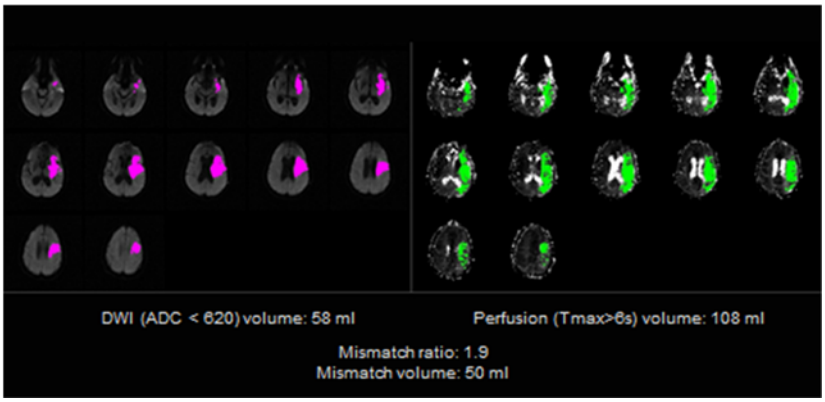

Figure 2. The RAPID mismatch summary allows investigators to quickly, accurately and easily determine if the patient meets the imaging criteria for enrollment. The patient above meets the Target Mismatch criteria: core volume is < 70 ml, mismatch ratio is ≥ 1.8 and mismatch volume is ≥ 15 ml.

The RAPID output maps, which identify the volume and location of ischemic core and perfusion lesions, are emailed to investigators (protected health information is automatically removed) and auto-sent to PACS as well as to a secure email system for viewing within 5 minutes of completion of the scan. Immediately after the images are available, the investigator will review the results of the RAPID mismatch map (**Figure 2**) and the MRA/CTA to determine if the patient meets the imaging criteria (listed above). If a patient has undergone multiple imaging evaluations (both MRI and CT or multiple CTs or MRIs), the most recent imaging study will be used to determine if the patient meets the imaging criteria. The accuracy of the software for identifying the size and location of perfusion and diffusion lesions has been established by extensive validation and testing on blood flow phantoms; the software received FDA 510K clearance for clinical use in 2013. The agreement between local investigators and the Imaging Core Lab for identification of the mismatch profile in DEFUSE 2 was 97%, κ 0·92; 95% CI 0·83–1·

Baseline data will be captured for all consented patients, including for patients who do not meet the imaging criteria. Patients who do not meet the imaging eligibility criteria will receive standard care per local hospital practice.

3.4.2 Randomization using a Dynamic Stratification Algorithm: Once a consented patient is determined to meet all Neuroimaging criteria, the patient ***will immediately be***

randomized on the WebDCU™ website. A dynamic stratification system will ensure well-balanced subgroups. The randomization algorithm, which will be programmed into the data capture system, will employ biased-coin minimization and the variance method with stratification weights.[44] The strategy is to balance treatment assignment along the marginal distribution of each stratification factor. The stratification factors used will be: 1) age, 2) core lesion volume, 3) time from symptom onset to enrollment, 4) baseline NIHSS, and 5) study site. When a new patient is enrolled, the site will enter the stratification factor values into the eCRF (electronic case report form) on WebDCU™. The dynamic randomization algorithm will determine an imbalance measure for each treatment group. The treatment group associated with the smallest imbalance measure will receive the largest probability of assignment in the biased-coin randomization. The biased-coin acceptance region and stratification weights are specified in the Randomization Plan. The superior balancing characteristics of dynamic randomization over blocked randomization have been well established.

Patients will be assigned to either endovascular therapy plus medical therapy or to medical therapy alone (1:1 randomization). Crossover from medical to endovascular therapy is strictly prohibited; endovascular to medical therapy crossover is defined as a patient who is assigned to endovascular therapy but does not have a conventional angiogram performed. Endovascular to medical therapy crossover is only allowed if an endovascular contraindication arises after randomization. Sites will be closely monitored for crossovers (see site monitoring plan below).

3.5 Acute Treatment

3.5.1 Endovascular Therapy: In patients randomized to endovascular therapy, the goal for femoral artery puncture will be within 45 minutes of randomization; femoral artery puncture must occur within 90 minutes of the completion of the qualifying imaging. Patients will be treated with thrombectomy devices (stent-retrievers) and/or suction thrombectomy systems currently cleared by the FDA for thrombus removal in patients experiencing an acute stroke within 8 hours of symptom onset following the published instructions for use for these devices. These devices will be used up to 16 hours following symptom onset in DEFUSE 3 based on an FDA IDE. The devices which will be used are the Trevo Retriever, the Solitaire Revascularization Device, Covidien MindFrame Capture Revascularization Device and the Penumbra thrombectomy system.

Standard medical therapy, based on current AHA guidelines, will also be provided for all patients. Individual investigators may use any of these devices or any combination of these devices to remove thrombus from the ICA, MCA M1 segment or, if needed, from M2 segments of the intracranial circulation. These are all approved anatomic locations for these devices. The use of thrombectomy devices should be performed in accordance with the indications for use. If there is a severe stenosis of the common carotid artery or the proximal internal carotid artery, investigators may also use other FDA devices approved for angioplasty or FDA devices approved for stenting of the carotid artery as deemed appropriate. The use of adjuvant intra-arterial (IA) thrombolytic medication is not currently approved by the FDA for stroke treatment and cannot be used in DEFUSE 3.

Sites will use local protocols for femoral access, sedation, heparin infusion, monitoring, etc. Sites will perform a cervical injection of the involved carotid circulation as a baseline

angiogram. At the conclusion of the procedure, a post-treatment angiogram as a cervical injection of the involved carotid circulation will also be obtained. Imaging will cover the full region of the normal circulation in AP and lateral projections at 2-3 films per second through the entire venous phase. All brain imaging from stroke onset through hospital discharge, including the baseline MRI and CT, as well as angiographic images obtained for the diagnostic and therapeutic portions of the procedure, will be transmitted to the core lab.

3.5.2 Medical Therapy: Patients randomized to medical therapy will receive standard medical therapy based on current AHA guidelines. Based on the time window for DEFUSE 3, it is anticipated that very few of the patients enrolled in DEFUSE 3 will have received iv tPA prior to randomization ("tPA failures"). For these patients, the sites' post-tPA protocol will be followed. Non-tPA treated patients randomized to medical therapy will be treated with aspirin, 325 mg on Day 1, and 81-325 mg/day (investigator's preference) on day 2-5, unless an indication for early anticoagulation is present (as determined by the patient's attending physician). All patients will receive standard DVT prevention therapy. Intravenous anticoagulants are prohibited (unless a clear indication for early anticoagulation is documented); dual antiplatelet therapy is prohibited unless carotid stenting was performed during the endovascular procedure or a clear indication for dual antiplatelet therapy is documented. Subsequent antithrombotic therapy will be determined by the patient's attending physician.

3.6 Clinical and Imaging Evaluations

Follow-up (imaging and clinical): Randomized patients will be followed clinically for 90 days and will have an MRI/MRA/MR perfusion at 24 hours (range 18-30 hours) to assess infarct volume, recanalization, hemorrhage and reperfusion (**Table 1 below**).

Table 1 Schedule of Events

	Screening	Enrollment	Baseline / Randomization	Endovascular Procedure	24 hours (+/-6 hrs)	Hospital Discharge	Day 30 (+/- 7 days)	Day 90 (+/- 14 days)
Screen Failure Log	X							
Informed Consent		X						
Subject Enrollment		X						
Inclusion and Exclusion Criteria			X					
MRI or CTP			X		X++			
Randomization			X					
Medical History			X					
Vital Signs			X					
NIH Stroke Scale			X		X	X	X	X
Modified Rankin Scale			X+			X	X	X
Baseline ASPECTS Score			X					
Baseline Labs*			X					
Endovascular Therapy				X				
24 Hour Labs					X			
Hospital Discharge						X		
Adverse Event Assessment				X	X	X	X	X
NeuroQOL								X

*Laboratory Evaluation includes CBC with Platelets, Creatinine, Glucose, INR, activated PTT, and Pregnancy test (if applicable). At 24 hour follow-up only creatinine is required. + Historical mRS at baseline, mRS/NIHSS to be performed by an NIHSS/mRS certified member of the research team who is blinded to treatment allocation at 30 and 90 days. ++ Patients will preferably undergo an MRI with MRA AND MR Perfusion at 24 hours; if an MR cannot be performed, a CT with CTA AND CTP can be substituted. For patients who are consented but not randomized, the schedule of events is limited to a summary of stroke therapies received within 24-hrs of stroke onset.

3.6.1 Assessments and follow-up visits

Baseline visit: All items in Table 1 above listed under "baseline" are to be performed prior to randomization. The MRI or CT scan should be performed with the DEFUSE 3 (baseline) protocol, which will be installed at all study sites. In addition, the inclusion/exclusion page of the case report form must be completed to determine if the patient meets the eligibility requirements for the study. If the patient is eligible and the consent form is signed by the patient or authorized representative, then the randomization procedure should occur immediately.

24 hour visit (+/- 6 hours): The items listed for this visit in Table 1 should be performed between 18 and 30 hours from the time of randomization. The only laboratory value required at the 24 hour visit is a serum creatinine. If possible, the 24 hour follow-up imaging study should be performed with multimodal MRI, rather than CT perfusion. The MRI or CT scan performed at this time should be performed with the DEFUSE 3 protocol.

Discharge visit: The items listed for this visit in Table 1 should be performed on the day of hospital discharge

30 and 90 day visits: The items listed for this visit in Table 1 should be performed on Day 30 (+/- 7) days and Day 90 (+/- 14) days. The mRS score must be performed by an mRS certified investigator who is blinded to treatment allocation at both the 30 and 90 day visits. If an in person visit is not possible, then the mRS should be performed by phone by an mRS certified investigator who is blinded to treatment allocation. If an in person visit is not possible, then the NIHSS score will be marked "not available" in the case report form.

Neurological worsening: If clinical worsening (defined as a $\geq$4 point increase on the NIHSS score) occurs prior to discharge, a CT scan or MRI should be obtained as soon as possible. Neurological worsening is a reportable adverse event.

3.6.2 Sources of Materials

Information on the clinical status of patients will be obtained from the patient's medical record. Study coordinators at the site will complete the DEFUSE 3 case report forms to collect basic demographic and medical information about the patients. Data will subsequently be entered into the StrokeNet's WebDCU electronic data capture system. Imaging data will be electronically transmitted to the coordinating center at Stanford via RAPID (all patient identifiers are removed by the software prior to exporting the data outside of the site's firewall). All study sites will complete a stroke screening log that documents all patients treated in the cath lab beyond 6 hours at their center, and reason for exclusion of patients not enrolled, in the StrokeNet's WebDCU electronic data capture system. Serious adverse events (SAEs) will be reported within 24 hours of the event in the StrokeNet's WebDCU electronic data capture system. The data collection process will include patient demographics, medical history, vital signs, laboratory assessments, NIHSS and mRS scores, and results of diagnostic studies performed to clarify stroke etiology.

3.7. Site Approval and Monitoring Plan

Site approval: Individual sites approved for participation in the study will be high-volume sites. Selected sites will have access to emergent CT perfusion and/or MR imaging 24/7. Prior to activating a site, we will verify that RAPID is functional at the site. Together with the

site's CT and/or MR technologists, we will install the DEFUSE 3 scan protocol on the local scanners and perform a dummy-run to assess image quality and train the technologists in software handling and data sending. A site will be activated for enrollment after test cases processed with RAPID have ensured good quality maps.

Monitoring for imaging quality: The Imaging Core Lab will monitor image quality throughout the study. If significant inadequacies or protocol errors are noted at a site, enrollment will be halted. Enrollment will resume after all imaging problems have been resolved and repeat dummy runs have been obtained that demonstrate adequate image quality.

Table 2. Example Imaging Sequences for DEFUSE 3 scans

Sequence	Scan Parameters (3T)	Time
MRI		**6 min**
Localizer	128X256; 28 FOV;5/5mm, GRE	24 sec
Calibration		5 sec
DWI	128x128, 24 FOV, 5/0mm, 30 slices, 1 NEX, R=2; b=0 and 1000 s/mm^2 over 3 axes, TE/TR=min/7000ms.	25 sec
GRE	256x192; 24 FOV; 5/0 mm, 30 slices, TE/TR= 25/800ms, flip 20, interleaved EPI, 16 shots	27 sec
MRA intracranial	256x192, 1 mm; 4 slabs, 26 phase-encodes; 6 overlap, 22 FOV, 0.8 rFOV, fractional echo, ZIPx2, ZIPx512, minTE, flowcomp, TR=18ms, flip=18, inferior->superior ramppulse, R=2; 19 MIPS	143 sec
PWI	128x128; 24 FOV; 5/0 mm, 17 slices, TE/TR=35ms/1800ms, R=2 using 0.1mmol/kg Gadolinium @ 4ml/sec.	108 sec
CT (example below for GE VCT; comparable protocols will be used for other scanner models)		**5-6 min**
Non-con head	2.5 – 5mm, 40 slices, 120-140kV, 265-290mA	120-180 sec
CTA	0.625mm, 0.984:1/39.37cm, 120kV, 550mA , inject and observe for 15 sec until contrast concentration in ascending aorta reaches 80HU (smart prep) then the CT gantry moves along with the bolus of the contrast material from the aortic arch up to the apex of the brain in 5sec.	90 sec
CTP	22 FOV, 40mm, 8x5mm, 1.8sec time interval, 45 cycles, 80kV, 125mA; 2 runs	90 sec

Monitoring for bias: A detailed site-monitoring plan has been developed to detect bias. This plan will protect the study from enrollment, randomization, and treatment bias. The first component involves monitoring the percentage of each site's endovascular volume (within 6-16 hrs) that is enrolled in DEFUSE 3. Sites will report their volume of endovascular stroke procedures (within 6-16 hrs) each month on a screening log. If a DEFUSE 3 eligible patient is treated with endovascular therapy outside the DEFUSE 3 study, an explanation will be required detailing why the patient was not enrolled. The second component of the plan involves tracking of patients who are consented but not randomized. These patients will require an entry in the WebDCU with an explanation why the patient was not randomized as well as documentation whether endovascular therapy was performed outside of the study. A third component involves monitoring of crossover after randomization. The Executive Committee will review the data described above for each site every 6 months. If evidence of enrollment bias is suspected, it will be investigated. If confirmed, the site will be placed on probation. If additional incidents of

suspected bias are confirmed, the site will be withdrawn. Routine monitoring of the clinical sites for source to database verification will be performed by the StrokeNet Data Management Center.

3.8. Sample Size, Adaptive Design and Statistical Analysis

DEFUSE 3 will feature a novel adaptive trial design that will allow the study to focus on a subpopulation if interim or final analyses indicate futility in the overall population.[59] The adaptive design was developed specifically for DEFUSE 3. It is based on closed testing theory and the group sequential methods for the Generalized Likelihood Ratio (GLR) statistic developed by Lai and Shih.[60] The adaptive design was chosen because there is strong preliminary data that suggests that the effect of endovascular treatment is modified by two baseline variables: core lesion size and time-to-treatment. The way the adaptive design takes advantage of these biological assumptions (when they are true) is by reallocation of future accrual to the subgroup with the best prospects for showing efficacy. Specifically, if a subgroup is chosen at an interim analysis, subsequent enrollment is limited to patients in that subgroup. As a result, this subgroup will become larger than it would have been in the absence of the adaptive design. The criterion for deciding which subgroup has the best chance of showing a benefit from endovascular therapy combines both the estimated size of the effect in the subgroup and the sample size of the subgroup. The GLR statistic (Kullback-Leibler criterion) is used to identify this subgroup because it optimally balances those two criteria. It selects the subgroup that has the best chance of showing an effect because it has an apparently large effect and is also of substantial size (note there are 5 subgroups of increasingly larger size, **figure 3**). The adaptive design employs two biologically-based assumptions to limit the inflation of sample size; a monotonicity/contiguity assumption and an a prior assumption that the effect is largest in the patients with the smallest DWI lesions and the shortest time to randomization (cell C_{11} in **figure 3**). The boundaries of the categories (cells) will be determined just prior to the 1st interim analysis based on the distribution of patients across these two dimensions (lesion volume and time-to-treatment).

		Time (hrs)	
		<10	10-16
Core lesion volume (ml)	≤20	C_{11}	C_{21}
	21-50	C_{12}	C_{22}
	51-70	C_{13}	C_{23}

Figure 3. The cohort is stratified according to core lesion volume and time to randomization. Exact boundaries of the stratification will be determined based on the distribution of patients just prior to the first interim analysis. Depending on the results of the 1st interim analysis, subsequent enrollment will continue in all 6 cells or will be limited to one of 5 sub-groups (C_{11}, C_{11+21}, $C_{11+21+12}$, $C_{11+21+12+22}$, or $C_{11+21+12+22+13}$).

Primary analysis: The primary endpoint is the distribution of scores on the modified Rankin Scale (mRS) at day 90. We will test the null hypothesis at the interim and final analysis using a normal approximation of the Wilcoxon-Mann-Whitney test (the generalized likelihood ratio [GLR] test). The primary analysis will be intention to treat, adjusted for design and not adjusted for covariates.

For each analysis, an efficacy bound will be set to control the overall (one-sided) Type I error rate at 2.5%. At each interim analysis a futility bound will be set to decide if the study should continue recruitment in the overall group, shift accrual and testing to a subgroup, or stop in its entirety. The futility boundary adapts when a subgroup is selected to the fact that the maximum analyzed sample size is a random variable that is no larger than the fixed maximum number of patients randomized (n=476). Because subgroup selection reduces the

maximum number of patients available for analysis at completion of the study, this method effectively allows an easier futility stop after subgroup selection. This setup replaces conditional power analyses with an automatic and more powerful adjustment of boundaries.

First interim analysis (n=200 randomized and completed 90 day follow-up): The null hypothesis is tested in the entire patient population:

1. If neither efficacy nor futility bound is crossed, the trial continues enrollment to the 2nd interim analysis.
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the overall population.
3. If the futility bound is crossed, the optimal subgroup is selected based on the Kullback-Leibler criterion and the null is tested in that subgroup. The futility bound is relaxed as described above, based on the expected maximum number of patients in the trial at completion (ie 476 minus the number of patients already enrolled in cells that will no longer be open for enrollment).
    3.1. If neither bound is crossed, the trial will continue with enrollment limited to the selected subgroup
    3.2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
    3.3. If the futility bound is crossed, the trial stops for futility.

Second interim analysis (n=340 randomized and completed 90 day follow-up): If, after the first interim analysis, the study proceeds with enrollment in the overall population (option 1 above), the testing at the 2nd interim analysis is identical to the first interim. If enrollment is limited to a selected subgroup (option 3.1), the null is tested in that subgroup:

1. If neither bound is crossed, the trial continues to the final analysis with enrollment of 136 additional patients limited to the selected subgroup
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
3. If the futility bound is crossed, the trial stops for futility.

Final analysis (n=476 randomized and completed 90 day follow-up): If, after the second interim analysis, the study proceeds with enrollment in the overall population, the null is tested in the overall population:

1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in the overall population.
2. If the efficacy bound is not crossed, the optimal subgroup is selected and the null is tested in that group:
    2.1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in that subgroup
    2.2. If the efficacy bound is not crossed, endovascular therapy will be declared of no benefit.

If enrollment after one of the interim analyses is limited to a selected subgroup, the null will be tested in that subgroup only and efficacy or lack thereof will be declared as per options 2.1 and 2.2 above.

Power and sample size considerations: The projected overall effect of endovascular therapy is based on 1) the observed 90-day modified Rankin Scale outcomes in DEFUSE 2 of target mismatch patients treated >6hrs after symptom onset and 2) the assumption that early reperfusion will be achieved in 75% of the endovascular arm vs. 20% of the medical therapy arm.[20, 21, 61] Using these data, we projected the distributions on the mRS at 90 days in the endovascular and control arms of DEFUSE 3:

	mRS at day 90							
	0	1	2	3	4	5	6	Total
Endovascular group	18.0%	11.5%	19.6%	11.5%	16.4%	11.5%	11.5%	**100%**
Medical group	9.7%	7.9%	15.0%	17.7%	14.4%	17.7%	17.7%	**100%**

This distribution corresponds to a standardized effect of 0.36 for the primary analysis. Based on these data, the fixed sample size for a non-adaptive design requires a total of 376 patients (188/arm) to have 90% power at an alpha of 5% (Wilcoxon-Mann-Whitney test); 100 additional patients are added for the adaptive design to reach a maximum sample size of 476 for DEFUSE 3.

We ran simulations (n=5000) to compare the performance of a traditional fixed sample-size design (fixed n=476) to the adaptive design (max n=476) under various scenarios (see **Table 3**, below). For the simulations the effect size is expressed as a standardized effect, where a standardized effect of 0.3 corresponds to a conservative projected effect of endovascular therapy (anticipated effect 0.36; see above).

Imaging outcomes: We hypothesize that endovascular treatment improves radiological outcomes in stroke patients with favorable clinical and imaging characteristics. DEFUSE 2 demonstrated a substantial reduction in infarct growth among Target mismatch patients treated in the 6-12 hour time-window who achieved early reperfusion: median growth 0.5 ml (IQR: -2 – 10) with reperfusion (n=23) vs. 39 ml (IQR: 18-121) without reperfusion (n=13), p<0.001. These data have been extrapolated to DEFUSE 3 using the same assumptions described above; anticipated an early reperfusion rate of 75% in the endovascular arm vs. 20% in the medical arm. This yields a sample size of 42 per group for 90% power. Therefore, DEFUSE 3 is highly powered to demonstrate differences in lesion growth. Infarct volumes, ischemic lesion growth, and reperfusion rates at 24 hours will be compared between groups with the Mann-Whitney U test. The 24 hour endpoint is based on data demonstrating that assessment of infarct volume at 24 hours captures the effect of reperfusion therapies on infarct growth and predicts outcomes similarly to day 90 infarct volumes.[29, 62] RAPID-assessed ischemic core volume at baseline will be correlated with 24h infarct volume (DWI volume) in subjects who achieve reperfusion without PH1 or PH2 intracranial hemorrhage. Pearson's correlation coefficient will be calculated and the median absolute error (ml) will be reported. Similarly, correlation of the baseline Tmax >6 volume and the 24h infarct volume in patients without PH1 or PH2 intracranial hemorrhage who have <10% reperfusion will be performed. Correlation of RAPID predicted infarct volume (coregistered baseline ischemic core and 24h Tmax >6 volume) with the actual 24h infarct volume will also be performed.

Scenario	Standardized effect in cells C11,	C12,	C21,	C22,	C31,	C32	Average standard. effect	Adaptive Design Average No. randomized	Power	Fixed Design Number randomized	Power
#0	0	0	0	0	0	0	0	361	2.2%	476	2.5%
#1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	354	80%	476	89%
#2	0.5	0.4	0.3	0	0	0	0.2	400	86%	476	55%
#3	0.5	0.5	0	0	0	0	0.17	403	87%	476	41%

Table 3. Under the null (Scenario #0), the adaptive design controls the total Type 1 error below 2.5%, stops early for futility 63% of the time, and the average number of randomizations is 361. If the effect is uniform across cells (scenario #1), the fixed-sample design is optimal, but the adaptive design results in only a small loss of power (from 89 to 80%). The adaptive design performs much better (higher power and smaller expected sample size) than the fixed sample, conventional trial when the effect size distribution across the subgroups is in accord with the biological assumptions (scenarios #2 and 3). If the effect is concentrated in two cells with small core volumes (scenario #3), the adaptive design maintains power (87%) while the conventional design collapses (41% power). The adaptive design also performs well compared to a non-adaptive, fixed sample that includes efficient multiple comparisons-adjusted testing for effect in subgroups at the end of the study. (see Lai et al[59])

Secondary analysis: Our secondary endpoint is the proportion of patients with mR S 0-2 at day 90 (indicating functional independence). The difference in the proportions of patients with mRS 0-2 between treatment arms will be assessed using logistic regression.

Subgroup analyses: Subgroup analyses of the effect of endovascular therapy on the primary and secondary endpoints will be performed. Subgroups will be defined based on the stratification variables, key demographic factors (such as race and ethnicity), tPA vs. no tPA, CTP vs. MRI selection, and witnessed vs. unwitnessed symptom onset, wake-up vs non-wake-up stroke, and TICI 0-2a vs. TICI 2b/3 results in cath lab.

Missing data/lost to follow-up (LTFU):

All effort is put forth to ensure near complete follow-up, in particular with the assessment of the primary outcome (mRS at 90 days), death (mRS=6), and stroke recurrence. If the primary outcome (mRS at 90 days) cannot be assessed in the clinic, it will instead be obtained by phone using a structured interview. If the subject's mRS cannot be obtained in clinic or by phone within the window of 60 to 120 days from randomization, then for primary analyses the day 30 mRS score will be used as the primary outcome (ie day 30 mRS carried forward). If neither the 30-day nor the 90-day mRS is available, then the mRS will be imputed. We do not expect this to alter the main study results given our estimated very low LTFU rate (<2%).

DEFUSE 3 Timetable

Year 1	Year 2	Year 3	Year 4	Year 5
Install RAPID at all sites Begin enrollment (anticipated to begin mid-year)	Pt enrollment continues for a total of 4 yrs	1st interim analysis Potential modification of enrollment criteria based on adaptive design	2nd interim analysis Potential modification of enrollment criteria based on adaptive design	Finish enrollment Data Analyses Publication of results

3.9. Risk analysis

Description and analysis of all increased risks to the research subjects:

Potential complications of MRI scan include localized twitching sensation due to the magnetic field changes during the scan, anxiety due to claustrophobia and allergic reaction to the contrast agent. The allergic reaction may include headache, nausea, rash, hives, nasal congestion, sneezing, itching or swelling. If a severe reaction occurs, swelling of the throat, chest tightness, or a marked drop in blood pressure may occur. In addition, pain, bleeding, bruising, coldness or inflammation at the injection site may occur. Precautions will be taken for early detection and rapid treatment if such reactions occur.

Potential complications of CT scan include radiation exposure and allergic reaction to CT contrast agents.

Radiation doses:
Combined scanning with comprehensive stroke imaging, which includes a noncontrast head CT scan, perfusion imaging, and CT angiography of the cervicocranial vessels starting at the aortic arch results in a dose of approximately 7-10 mSv. (AJNR 2010 31: 1003-1009). According to the National Council on Radiation Protection and Measurement, the average annual radiation dose per person in the U.S. is 620 millirem (6.2 milliSieverts).

Reactions to contrast agents:

- **Mild**
  Nausea, Vomiting**,** Headache**,** Cough**,** Nasal stuffiness, Altered taste**,** Flushing**,** Itching**,** Rash, Hives**,** Sweats**,** Swelling of eyes or face
- **Moderate**
  Mild hypotension, Tachycardia or Bradycardia, Bronchospasm, Wheezing, Dyspnea, Laryngeal edema, Generalized or diffuse erythema
- **Severe**
  Cardiopulmonary arrest, Clinically manifested arrhythmias, Profound hypotension, Convulsions, Unresponsiveness, Respiratory failure, Laryngeal edema
  The rate of major reactions (e.g., anaphylaxis, death) is very low, estimated at one in 170,000 administrations.

Potential complications of endovascular therapy include stroke; new clot in an artery; total blockage of an artery; infection and pain in the region of insertion site; lack of blood flow to the brain; rupture or puncture of an artery; significant tearing of the vessel wall; bleeding requiring blood transfusion; allergic reaction to contrast dye; abnormal low blood pressure requiring treatment; temporary closing of the artery (vessel spasm); formation of or dislodgments of clots which block the arteries (embolism). In rare circumstances, the procedure could result in death. At the puncture site in the groin, a blood clot or other blood vessel injury may occur and require blood transfusion or surgical repair. Infection may occur at the puncture site; this could cause pain and require additional medications. There is some chance of an allergic reaction to the x-ray contrast (dye) used during the angiogram procedure. Minor allergic reactions may include a rash or hives. There is also the possibility of a serious allergic reaction that could include shortness of breath and swelling, drop in

blood pressure, and even death. Patients will be closely monitored for these reactions and receive prompt treatment to reverse any allergic reactions.

Safety of endovascular therapy beyond 8 hours

Mechanical thrombectomy devices have been used beyond 8 hours of stoke onset in a number of clinical trials and registries. No safety concerns have been associated with late window therapy. In DEFUSE 2 [37], patients were treated up to 12 hours after symptom onset, and no safety concerns were identified in any time window. Based on both favorable safety data and encouraging efficacy data, Stryker Neurovascular has initiated the DAWN Study with a 24 hour treatment window. No safely issues have been identified to date in DAWN. The DAWN study is being run under an FDA IDE.

Methods to mitigate risks to subjects in the trial

Methods to mitigate risks to subjects in the trial include exclusion of subjects with bleeding disorders and selection of subjects via neuroimaging (infarct core lesions less than 70 ml) to minimize the risk of symptomatic intracranial hemorrhage. Computed Tomography (CT) scans will be performed for neurological deterioration (≥ 4 point increase in National Institutes of Health Stroke Scale (NIHSS) score) to identify new strokes, hemorrhage, or edema. Hospitals will follow their local standard of care safety procedures in order to reduce the risk of kidney dysfunction caused by contrast agents. Only investigators who are trained and experienced with use of the devices allowed within the trial are eligible to participate (see **Site Approval and Monitoring Plan**) above. The adaptive design will eliminate subgroups with an unfavorable therapeutic response.

Patients will be carefully screened for CT, MRI and endovascular treatment contraindications according to the inclusion/exclusion criteria and excluded from enrollment if any are present.

Radiation exposure: Radiation exposure during all tests will be minimized by optimizing the imaging protocols and by limiting fluoroscopy-time during the endovascular procedure. All CT sequences, including the CTP sequence, meet all FDA guidelines for radiation exposure.

Stopping rules or safety triggers for the study

Symptomatic ICH or death rates that exceed pre-specified thresholds will trigger a meeting of the DSMB to discuss the events and make a determination on the continuation of the trial. Below are the pre-specified triggers:

The DEFUSE 3 has established the following automatic stopping rules, based on identifying with 95% probability:

1) the rate of symptomatic ICH (NIHSS worsening of 4 or more points associated with ICH) in the endovascular group is exceeding 10%
2) the rate of 90 day mortality in the endovascular group is exceeding 20%.

If either threshold is crossed, the study will be automatically placed on hold until the investigators and the DSMB can conduct a review of events.

Adverse events (AEs) will be collected, recorded, and analyzed in accordance with Section 3.11 below.

Safety oversight for this study will be provided by both the DSMB and an independent Medical Safety Monitor. Please see Section 3.11 for details.

Patient Population: Four hundred and seventy six acute stroke patients meeting the pre-defined inclusion criteria will be enrolled in the trial. The mean age is anticipated to be 69 years of age. Our targeted planned enrollment breakdown is as follows:

Racial Categories	Ethnic Categories				Total
	Not hispanic or Latino		Hispanic or Latino		
	Female	Male	Female	Male	
American Indian/Alaska Native	2	2	1	1	6
Asian	13	13	1	1	28
Native Hawaiian or Other Pacific Islander	2	3	0	0	5
Black or African American	32	32	1	1	66
White	175	174	11	11	371
Racial Categories: Total of All Subjects	224	224	14	14	476

Imaging core lab: The Stanford imaging core lab has 15 years of experience with MR image storage and processing. It will perform the organization, archiving and blinded analysis of all imaging data collected in DEFUSE 3. They will be responsible for MRI and CT Perfusion image processing and artifact removal and will generate final lesion volumes for all MRI scans performed in the study.

DSA Angiograms sent to the core lab will include a baseline (pre-treatment) and a final angiogram for the territory of treatment. In addition, angiographic images from MRA or CTA will be sent to the core lab. The MRA or CTA will be used to assign a primary arterial lesion (AOL) from non-invasive imaging. The baseline DSA angiogram will also be used to assign a primary arterial occlusive lesion (AOL) and a pre-treatment mTICI score. The final angiogram will have a post-treatment mTICI score assigned.[22],[52] This scoring system defines TICI 2A as partial perfusion of < 50% of the vascular distribution of the occluded artery and 2B as partial perfusion of > 50% of the vascular distribution. The scoring system will use these previously described definitions[22]:

- Grade 0 No perfusion
- Grade 1 Antegrade reperfusion past the initial occlusion, but limited distal branch filling with little or slow distal reperfusion
- Grade 2a Antegrade reperfusion of less than half of the occluded target artery previously ischemic territory (eg, in 1 major division of the MCA and its territory)
- Grade 2b Antegrade reperfusion of more than half of the previously occluded target artery ischemic territory (eg, in 2 major divisions of the MCA and their territories)

- Grade 3 Complete antegrade reperfusion of the previously occluded target artery ischemic territory, with absence of visualized occlusion in all distal branches

Two senior neurointerventionalists will perform the angiographic analysis, blinded to the clinical data, MR imaging and CT Perfusion results, and the analysis of the other core lab reader. Any disagreements in the AOL interpretation or the mTICI scores will be adjudicated by common review of those cases and a consensus reading will be applied.

National Data Management Center: Data management and site monitoring will be performed by the StrokeNet National Data Management Center (NDMC) at Medical University of South Carolina (Director, Yuko Y. Palesch, PhD, see letter of support). The NDMC will create the database and set up the interface on the website (WebDCU™) where clinical site personnel will enter the data into the electronic CRF. Data quality assurance processes include: (1) logic and rule checks built into the database; (2) monitoring by the Data Manager at the NDMC; (3) central monitoring by the statistical programmer at the NDMC; and (4) risk-based source verification monitoring by the Clinical Research Associates. DEFUSE 3 data, including neuroimaging, will be shared in accordance with the StrokeNet data sharing policies. Anonymized neuroimaging will be stored on secure servers at the Stanford Stroke Center with nightly back-up. NINDS Common Data Elements will be used for both clinical and imaging data.

3.10. Description of devices

The following FDA approved thrombectomy devices will be included:

1) Trevo Retriever
2) Solitaire™ Revascularization Device
3) Covidien MindFrame Capture Revascularization Device
4) Penumbra thrombectomy system including the following devices and pumps:

Penumbra Aspiration Pump (1115V)	Penumbra System [026, 032, 041]
Penumbra System 054	Penumbra System Separator Flex [026, 032, 041, 054]
Penumbra System MAX	Penumbra Pump MAX
Penumbra System 110 Aspiration Tubing	Penumbra System Reperfusion Catheter ACE64 & ACE68

3.11. Monitoring procedures

The coordination of the DEFUSE 3 Trial operations will be centralized through the following:
NIH StrokeNet National Coordinating Center (NCC)/ PI: Joseph Broderick, MD
University of Cincinnati
260 Stetson Street, Suite 2300
Cincinnati, Ohio 45267-0525

Leading the NCC team will be the Project Manager, who will be assigned to coordinate the following study oversight: trial communication required training activities, site assessment and/or initiation visits, collection of trial related regulatory documents, recruitment performance tracking, site monitoring, and performance analysis. Study oversight will be handled according to the Data Monitoring Standard Operating Procedure (SOP Number ADM 19).

DEFUSE 3 will have an independent Data and Safety Monitoring Board (DSMB) appointed by the NIH to oversee study safety. Patients in both study arms will be assessed for the incidence of stroke-related mortality at 90 days, the incidence of symptomatic intracranial hemorrhage at 36 hours from symptom onset (defined as a ≥ 4 point worsening of immediate pre-deterioration NIHSS neurological status vs. post deterioration and associated with brain hemorrhage), and the incidence of significant neurologic deterioration prior to discharge (defined as ≥ 4 point worsening of the immediate pre-deterioration NIHSS neurological status vs. post deterioration and not attributed to sedation). In the endovascular arm patients will be assessed for intra-procedural complications including: intra-procedural mortality, vessel perforation, arterial dissection, access site complication requiring surgical repair or blood transfusion, embolization and device failure. SAEs will be reported within 24 hours of awareness of the event.

The DSMB will meet in person or by teleconference, on a semi-annual basis, to monitor the cumulative safety data during participant follow-up. In no instance will more than 12 months elapse between DSMB reviews of cumulative safety data after the first participant has been randomized. The DSMB will monitor the study according to the guidelines specified in the study protocol and the operating procedures established at the initial meeting, unless the DSMB determines during the course of the trial that modification of the guidelines is in the best interest of the study and its participants.

Independent Medical Safety Monitor:
In addition to the DSMB, Dr. Andrew Demchuk has been appointed as the independent Medical Safety Monitor (MSM) for DEFUSE 3. Dr. Demchuk is not involved in the study and has no conflict of interest. He will be responsible for ongoing monitoring of reports of SAEs submitted by the clinical centers in real time to ensure good clinical practice and to identify safety concerns quickly. Dr. Demchuk may suggest protocol modifications to prevent the occurrence of particular AEs, e.g., modifying the protocol to require frequent measurement of laboratory values predictive of the event or to improve expeditious identification of SAEs. In the event of unexpected SAEs or an unduly high rate of SAEs, Dr. Demchuk will promptly contact the DSMB Liaison who will notify the DSMB Chair. In the event that he is unavailable for an extended period of time (i.e., extended vacation, sabbatical, illness, etc.), a back-up MSM will be nominated by the study PI and approved by the DSMB.

Adverse Event Reporting
Consideration of adverse events will hereafter consist of adverse events, serious adverse events, and adverse device effects, including anticipated adverse device effects and unanticipated adverse device effects.

- Adverse event (AE) is defined as any untoward/undesirable clinical occurrence in a clinical investigation of a subject which does not necessarily have a causal relationship with the treatment under investigation. An Adverse Event can therefore be any unfavorable and/or unintended sign, symptom, or disease temporally associated with the use of a device product, whether or not considered related to the device product. Only abnormal laboratory values that are deemed clinically significant by the investigator will be classified as adverse events.

- Serious adverse event (SAE) is defined as any untoward/undesirable adverse experience that results in any of the following outcomes: 1) death; 2) a life-threatening adverse experience; 3) inpatient hospitalization or prolongation of existing hospitalization; 4) a permanent/persistent or significant disability/incapacity or a congenital anomaly/birth defect; 5) important medical events that may not result in death, be life-threatening, or require hospitalization may be considered a serious adverse event when, based upon appropriate medical judgment, they may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. This category includes the use of intra-arterial thrombolytics and/or intracranial stents.

- Anticipated adverse device effect (AADE) is defined as any adverse effect related to the device or procedure, which is identified in the protocol or the IFU for the device.

- Unanticipated Adverse Device Effects (UADEs) is defined as any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with a device, if that effect, problem, or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application, or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects.

Reports of UADEs will be made to the FDA within 10 days of receiving notification of the UADE (as required in 21 CFR 812 18p (b)(1)).

Safety Monitoring

The MSM will monitor all AE reports to identify and trend all events that would require temporary discontinuation of study enrollment, to fully characterize device safety, to modify the study protocol, or to terminate the study.

Reporting Procedures for All Adverse Events

All Adverse Events, whether or not attributed to the study and/or the devices, observed by the investigator or reported by the subject, will be recorded from the time of randomization through Day 5 or discharge, whichever is earlier. All SAEs will be recorded through Day 90.

The following attributes will be assigned by the reporting investigator:

1. Description of event
2. Date of onset
3. Date of resolution (if applicable)
4. Seriousness
5. Relationship to the study device and/or procedure(s)
6. Severity
7. Action(s) taken
8. Outcome(s)

Severity is defined as a measure of the intensity of a reaction, effect or experience. The measurement(s) are described as mild, moderate, severe, life threatening or death. The event itself, however, may be of relative minor medical significance. The severity of Adverse Events is assessed according to the following index scale:

- **Mild**
  asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated
- **Moderate**
  minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental Activities of Daily Living.
- **Severe**
  medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care Activities of Daily Living.
- **Life-threatening consequences;**
  urgent intervention indicated
- **Death** related to AE

The relationship of an AE to the study device or procedure will be graded as follows:

- Unrelated
- Unlikely
- Reasonable possibility
- Definitely

Serious Adverse Events All Serious Adverse Events including deaths will be reported to the MSM, the Central Institutional Review Board (IRB) and the FDA, as required.

4. INVESTIGATOR'S AGREEMENT

As this study will be carried out by the NIH StrokeNet, the names of the specific sites and investigators are not yet available. Investigator's agreements for key investigators at the coordinating site are included.

All investigators will be required to sign the following agreement:

INVESTIGATOR AGREEMENT
FOR THE CLINICAL INVESTIGATION OF THE DEFUSE 3 TRIAL

I_____ agree to participate as an Investigator on the DEFUSE 3 trial.

I have been provided a copy of the following Food and Drug Administration (FDA) regulations: 21 CFR Part 812, Investigational Device Exemptions; 21 CFR Part 50, Protection of Human Subjects; and 21 CFR Part 54, Financial Disclosure by Clinical Investigators.

I agree and/or certify that:

1. I will conduct the clinical investigation in accordance with this agreement, all requirements of the investigational plan, IDE regulations, other applicable regulations of the FDA, and any conditions of approval imposed by my reviewing Institutional Review Board (IRB) or FDA. I agree to abide by all of the responsibilities of Investigators addressed under 21 CFR Part 812, Subpart E and Subpart G, including but not limited to the following:

2. I will obtain written approval from the authorized IRB for the institution at which this investigation will be conducted.

3. I will ensure that Informed Consent is obtained from each subject participating in this clinical investigation in accordance with the informed consent regulation found in 21 CFR Part 50, and that a signed copy of the informed consent is available to the sponsor (sponsor-investigator) and the sponsor's (sponsor- investigator's) designated monitor.

4. I will ensure the accurate completion of protocol case report forms and, if I am not also the sponsor- investigator of the corresponding IDE application, I will submit completed protocol case report forms to the sponsor (sponsor-investigator) at the time frames specified in the Protocol and/or FDA regulations.

5. I have the appropriate, relevant qualification to conduct and to oversee the conduct of the clinical investigation as documented by the following: (***initial applicable statement***)
   ___ My relevant qualifications, including dates, location, extent and type of experience are listed in my most recent curriculum vitae (CV), which is attached to the Agreement and which will be maintained by the sponsor (sponsor-investigator) of the corresponding IDE application.
   ___ My curriculum vitae (CV) does not reflect my relevant qualifications, therefore attached to this Agreement is a statement of my relevant experience (including dates,

location(s), extent and type of experience) which will be maintained by the sponsor (sponsor-investigator) of the corresponding IDE application.

6. There are no reasons to question my ability to oversee the appropriate conduct of this clinical investigation. (***Initial applicable statement***)
   ___ I have never participated in an investigation or other research activity which was terminated (disqualified) by the FDA, IRB (or equivalent), or sponsor of a study due to non-compliance issue.
   ___ I have participated in an investigation or other research activity which was terminated (disqualified) by the FDA, IRB (or equivalent), or sponsor of a study due to non-compliance issue. The specific circumstances leading to this termination and my role in the respective problems or issues and the resolution of these problems or issues are summarized in an attachment to this Agreement.

I further certify that I have not been debarred under the Generic Drug Enforcement Act of 1992, 21 USC §§ 335a and 335b. In the event that I become debarred or receive notice of an action or threat of action with respect to my debarment during the term of this Agreement, I agree to immediately notify the sponsor (sponsor-investigator) and the authorized IRB for my study site. If I am the sponsor-investigator of the corresponding IDE application I will notify the authorized IRB and the FDA.

As required by 21 CFR Part 54, Financial Disclosure by Clinical Investigator, I will disclose sufficient and accurate financial information to the sponsor (sponsor-investigator) by completing the Certification of Financial Interest Form (attached) and if applicable, the Disclosure of Financial Interest Form (attached). I will also notify the sponsor (sponsor-investigator) if my disclosed financial information changes at any time during the clinical investigation or up to one year following the closure of the study.

Site Name and Address:
_____
_____
_____

_____ _____
Investigator Signature Date

5. EXECUTIVE COMMITTEE / KEY PARTICIPATING INVESTIGATORS

The Executive committee, composed of experts in vascular neurology, endovascular therapy and neuroimaging, will provide the overall scientific guidance for the study. The committee will typically meet monthly by phone (1 hour/month) for the full duration of the study. Responsibilities include oversight of the overall conduct of the study with regard to protocol compliance and modifications/amendments, study progress, and problem-solving. Dr. Albers will chair the executive committee.

Key Participating Investigators at Coordinating Site

Gregory W. Albers, MD Principal Investigator Stanford Stroke Center 780 Welch Rd. Suite 350 Palo Alto, CA 94305 650-723-4448 galbers@stanfordmed.org	Michael Marks, MD Co-Principal Investigator Stanford University Medical Center 300 Pasteur Dr. Stanford, CA 94305-5105 650-723-6767 mmarks@stanford.edu	Maarten Lansberg, MD, PhD Protocol Director Stanford Stroke Center 780 Welch Rd. Suite 350 Palo Alto, CA 94305 650-723-4448 lansberg@stanford.edu

6. INSTITUTIONAL REVIEW BOARD

The University of Cincinnati Institutional Review Board will serve as the National Central Institutional Review Board for all participating sites. The **Central Institutional Review Board (CIRB)** for multicenter protocols is the single IRB of record. It has regulatory responsibility for assuring the protection of the rights and welfare of research participants in accordance with Standard Operating Procedure ADM 12; Central Institutional Review Board Reporting. The National Institute of Neurological Disorders and Stroke (NINDS) selected the University of Cincinnati Institutional Review Board (IRB) to serve as the CIRB for the NIH StrokeNet (StrokeNet).

University of Cincinnati IRB Registration # 00000180 FWA #: 00003152 Expiration Date: 6/27/2016

Michael Linke, PhD, CIP
Chairman CIRB
Michael.linke@va.gov
513-304-3540

7. COSTS

All of the eligible devices that will be used in this study are currently on the market. There will be no charges beyond the typical standard of care for use of these approved devices. These devices will be used and billed according to the standard of care for each institution.

8. REFERENCES

1. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: Systematic analysis of population health data. *The Lancet*. 2006;367:1747-1757
2. Kleindorfer D, Lindsell CJ, Brass L, Koroshetz W, Broderick JP. National us estimates of recombinant tissue plasminogen activator use: Icd-9 codes substantially underestimate. *Stroke*. 2008;39:924-928
3. Katzan IL, Hammer MD, Hixson ED, Furlan AJ, Abou-Chebl A, Nadzam DM. Utilization of intravenous tissue plasminogen activator for acute ischemic stroke. *Arch Neurol*. 2004;61:346-350
4. Qureshi AI, Kirmani JF, Sayed MA, Safdar A, Ahmed S, Ferguson R, et al. Time to hospital arrival, use of thrombolytics, and in-hospital outcomes in ischemic stroke. *Neurology*. 2005;64:2115-2120
5. Saqqur M, Uchino K, Demchuk AM, Molina CA, Garami Z, Calleja S, et al. Site of arterial occlusion identified by transcranial doppler predicts the response to intravenous thrombolysis for stroke. *Stroke*. 2007;38:948-954
6. Mori E, Yoneda Y, Tabuchi M, Yoshida T, Ohkawa S, Ohsumi Y, et al. Intravenous recombinant tissue plasminogen activator in acute carotid artery territory stroke. *Neurology*. 1992;42:976-982
7. del Zoppo GJ, Poeck K, Pessin MS, Wolpert SM, Furlan AJ, Ferbert A, et al. Recombinant tissue plasminogen activator in acute thrombotic and embolic stroke. *Ann Neurol*. 1992;32:78-86
8. Albers GW, Thijs VN, Wechsler L, Kemp S, Schlaug G, Skalabrin E, et al. Magnetic resonance imaging profiles predict clinical response to early reperfusion: The diffusion and perfusion imaging evaluation for understanding stroke evolution (defuse) study. *Annals of Neurology*. 2006;60:508-517
9. Lansberg MG, Schrooten M, Bluhmki E, Thijs VN, Saver JL. Treatment time-specific number needed to treat estimates for tissue plasminogen activator therapy in acute stroke based on shifts over the entire range of the modified rankin scale. *Stroke*. 2009;40:2079-2084
10. Fields JD, Lindsay K, Liu Kenneth C, Nesbit GM, Lutsep HL. Mechanical thrombectomy for the treatment of acute ischemic stroke. *Expert Review of Cardiovascular Therapy*. 2010;8:581-592
11. Lee M, Hong K-S, Saver JL. Efficacy of intra-arterial fibrinolysis for acute ischemic stroke: Meta-analysis of randomized controlled trials. *Stroke*. 2010;41:932-937
12. Smith WS, Sung G, Saver J, Budzik R, Duckwiler G, Liebeskind DS, et al. Mechanical thrombectomy for acute ischemic stroke: Final results of the multi merci trial. *Stroke*. 2008;39:1205-1212
13. Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, et al. Intra-arterial prourokinase for acute ischemic stroke. The proact ii study: A randomized controlled trial. Prolyse in acute cerebral thromboembolism. *Jama*. 1999;282:2003-2011
14. Thomalla G, Schwark C, Sobesky J, Bluhmki E, Fiebach JB, Fiehler J, et al. Outcome and symptomatic bleeding complications of intravenous thrombolysis within 6 hours in mri-selected stroke patients: Comparison of a german multicenter study with the pooled data of atlantis, ecass, and ninds tpa trials. *Stroke*. 2006;37:852-858
15. Hsia A, Kidwell C. Developments in neuroimaging for acute ischemic stroke: Diagnostic and clinical trial applications. *Current Atherosclerosis Reports*. 2008;10:339-346

16. Kakuda W, Lansberg MG, Thijs VN, Kemp SM, Bammer R, Wechsler LR, et al. Optimal definition for pwi/dwi mismatch in acute ischemic stroke patients. *J Cereb Blood Flow Metab*. 2008;28:887-891
17. Lansberg MG, Thijs V, Bammer R, Kakuda W, Hamilton S, Wechsler L, et al. Clinical and mri-based risk factors for symptomatic intracerebral hemorrhage following treatment with tissue plasminogen activator. *Stroke*. 2006;37:654
18. Saver JL, Jahan R, Levy EI, Jovin TG, Baxter B, Nogueira RG, et al. Solitaire flow restoration device versus the merci retriever in patients with acute ischaemic stroke (swift): A randomised, parallel-group, non-inferiority trial. *Lancet*. 2012;380:1241-1249
19. Nogueira RG, Lutsep HL, Gupta R, Jovin TG, Albers GW, Walker GA, et al. Trevo versus merci retrievers for thrombectomy revascularisation of large vessel occlusions in acute ischaemic stroke (trevo 2): A randomised trial. *Lancet*. 2012;380:1231-1240
20. Marks MP, Lansberg MG, Mlynash M, Kemp S, McTaggart RA, Zaharchuk G, et al. Angiographic outcome of endovascular stroke therapy correlated with mr findings, infarct growth, and clinical outcome in the defuse 2 trial. *Int J Stroke*. 2014
21. Yoo AJ, Simonsen CZ, Prabhakaran S, Chaudhry ZA, Issa MA, Fugate JE, et al. Refining angiographic biomarkers of revascularization: Improving outcome prediction after intra-arterial therapy. *Stroke*. 2013;44:2509-2512
22. Zaidat OO, Yoo AJ, Khatri P, Tomsick TA, von Kummer R, Saver JL, et al. Recommendations on angiographic revascularization grading standards for acute ischemic stroke: A consensus statement. *Stroke*. 2013;44:2650-2663
23. Broderick JP, Palesch YY, Demchuk AM, Yeatts SD, Khatri P, Hill MD, et al. Endovascular therapy after intravenous t-pa versus t-pa alone for stroke. *N Engl J Med*. 2013;368:893-903
24. Hill MD, Demchuk AM, Goyal M, Jovin TG, Foster LD, Tomsick TA, et al. Alberta stroke program early computed tomography score to select patients for endovascular treatment: Interventional management of stroke (ims)-iii trial. *Stroke*. 2014;45:444-449
25. Kidwell CS, Jahan R, Gornbein J, Alger JR, Nenov V, Ajani Z, et al. A trial of imaging selection and endovascular treatment for ischemic stroke. *N Engl J Med*. 2013;368:914-923
26. O.A. Berkhemer, P.S.S. Fransen, D. Beumer, et al. A Randomized Trial of Intraarterial Treatment for AcuteIschemic Stroke. *N Engl J Med* 2015;372:11-20.
27. Muir KW, Buchan A, von Kummer R, Rother J, Baron J-C. Imaging of acute stroke. *The Lancet Neurology*. 2006;5:755-768
28. Campbell B, Purushotham A, Christensen S, Desmond P, Nagakane Y, Parsons M, et al. The infarct core is well represented by the acute diffusion lesion: Sustained reversal is infrequent. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2012;32:50-56
29. Chemmanam T, Campbell BC, Christensen S, Nagakane Y, Desmond PM, Bladin CF, et al. Ischemic diffusion lesion reversal is uncommon and rarely alters perfusion-diffusion mismatch. *Neurology*. 2010;75:1040-1047
30. Baird AE, Warach S. Magnetic resonance imaging of acute stroke. *J Cereb Blood Flow Metab*. 1998;18:583-609
31. Dani KA, Thomas RG, Chappell FM, Shuler K, MacLeod MJ, Muir KW, et al. Computed tomography and magnetic resonance perfusion imaging in ischemic stroke: Definitions and thresholds. *Ann Neurol*. 2011;70:384-401
32. Olivot JM, Mlynash M, Thijs VN, Kemp S, Lansberg MG, Wechsler L, et al. Optimal tmax threshold for predicting penumbral tissue in acute stroke. *Stroke*. 2009;40:469-475

33. Zaro-Weber O, Moeller-Hartmann W, Heiss WD, Sobesky J. Maps of time to maximum and time to peak for mismatch definition in clinical stroke studies validated with positron emission tomography. *Stroke*. 2010;41:2817-2821
34. Olivot JM, Mlynash M, Zaharchuk G, Straka M, Bammer R, Schwartz N, et al. Perfusion mri (tmax and mtt) correlation with xenon ct cerebral blood flow in stroke patients. *Neurology*. 2009;72:1140-1145
35. Mlynash M, Lansberg MG, De Silva DA, Lee J, Christensen S, Straka M, et al. Refining the definition of the malignant profile: Insights from the defuse-epithet pooled data set. *Stroke; a journal of cerebral circulation*. 2011;42:1270-1275
36. Lansberg MG, Lee J, Christensen S, Straka M, De Silva DA, Mlynash M, et al. Rapid automated patient selection for reperfusion therapy: A pooled analysis of the echoplanar imaging thrombolytic evaluation trial (epithet) and the diffusion and perfusion imaging evaluation for understanding stroke evolution (defuse) study. *Stroke*. 2011;42:1608-1614
37. Lansberg MG, Straka M, Kemp S, Mlynash M, Wechsler LR, Jovin TG, et al. Mri profile and response to endovascular reperfusion after stroke (defuse 2): A prospective cohort study. *Lancet Neurology*. 2012;11:860-867
38. Lansberg MG, Cereda CW, Mlynash M, Mishra NK, Inoue M, Kemp S, et al. Response to reperfusion not time-dependent in patients with salvageable tissue. Neurology, in press 2015
39. Inoue M, Mlynash M, Christensen S, Wheeler HM, Straka M, Tipirneni A, et al. Early diffusion-weighted imaging reversal after endovascular reperfusion is typically transient in patients imaged 3 to 6 hours after onset. *Stroke*. 2014;45:1024-1028
40. Wheeler HM, Mlynash M, Inoue M, Tipirneni A, Liggins J, Zaharchuk G, et al. Early diffusion-weighted imaging and perfusion-weighted imaging lesion volumes forecast final infarct size in defuse 2. *Stroke*. 2013;44:681-685
41. Parsons MW, Christensen S, McElduff P, Levi CR, Butcher KS, De Silva DA, et al. Pretreatment diffusion- and perfusion-mr lesion volumes have a crucial influence on clinical response to stroke thrombolysis. *J Cereb Blood Flow Metab*. 2010;30:1214-1225
42. Marks MP, Olivot JM, Kemp S, Lansberg MG, Bammer R, Wechsler LR, et al. Patients with acute stroke treated with intravenous tpa 3-6 hours after stroke onset: Correlations between mr angiography findings and perfusion- and diffusion-weighted imaging in the defuse study. *Radiology*. 2008;249:614-623
43. Weisstanner C, Gratz PP, Schroth G, Verma RK, Kochl A, Jung S, et al. Thrombus imaging in acute stroke: Correlation of thrombus length on susceptibility-weighted imaging with endovascular reperfusion success. *Eur Radiol*. 2014
44. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*. 1975;31:103-115
45. Chimowitz MI, Lynn MJ, Turan TN, Fiorella D, Lane BF, Janis S, et al. Design of the stenting and aggressive medical management for preventing recurrent stroke in intracranial stenosis trial. *J Stroke Cerebrovasc Dis*. 2011;20:357-368
46. Pereira VM, Gralla J, Davalos A, Bonafe A, Castano C, Chapot R, et al. Prospective, multicenter, single-arm study of mechanical thrombectomy using solitaire flow restoration in acute ischemic stroke. *Stroke*. 2013;44:2802-2807
47. Molina CA, Montaner J, Abilleira S, Ibarra B, Romero F, Arenillas JF, et al. Timing of spontaneous recanalization and risk of hemorrhagic transformation in acute cardioembolic stroke. *Stroke*. 2001;32:1079-1084

48. Wolpert SM, Bruckmann H, Greenlee R, Wechsler L, Pessin MS, del Zoppo GJ. Neuroradiologic evaluation of patients with acute stroke treated with recombinant tissue plasminogen activator. The rt-pa acute stroke study group. *AJNR Am J Neuroradiol*. 1993;14:3-13
49. Alexandrov AV, Molina CA, Grotta JC, Garami Z, Ford SR, Alvarez-Sabin J, et al. Ultrasound-enhanced systemic thrombolysis for acute ischemic stroke. *N Engl J Med*. 2004;351:2170-2178
50. Rubin D. *Multiple imputation for nonresponse in surveys*. Wiley & Sons; 1987.
51. Campbell BC, Tu HT, Christensen S, Desmond PM, Levi CR, Bladin CF, et al. Assessing response to stroke thrombolysis: Validation of 24-hour multimodal magnetic resonance imaging. *Arch Neurol*. 2012;69:46-50
52. Tomsick T, Broderick J, Carrozella J, Khatri P, Hill M, Palesch Y, et al. Revascularization results in the interventional management of stroke ii trial. *AJNR Am J Neuroradiol*. 2008;29:582-587
53. Christou I, Alexandrov AV, Burgin WS, Wojner AW, Felberg RA, Malkoff M, et al. Timing of recanalization after tissue plasminogen activator therapy determined by transcranial doppler correlates with clinical recovery from ischemic stroke. *Stroke*. 2000;31:1812-1816
54. Nguyen TN, Malisch T, Castonguay AC, Gupta R, Sun CH, Martin CO, et al. Balloon guide catheter improves revascularization and clinical outcomes with the solitaire device: Analysis of the north american solitaire acute stroke registry. *Stroke*. 2014;45:141-145
55. San Roman L, Obach V, Blasco J, Macho J, Lopez A, Urra X, et al. Single-center experience of cerebral artery thrombectomy using the trevo device in 60 patients with acute ischemic stroke. *Stroke*. 2012;43:1657-1659
56. Turk AS, Frei D, Fiorella D, Mocco J, Baxter B, Siddiqui A, et al. Adapt fast study: A direct aspiration first pass technique for acute stroke thrombectomy. *J Neurointerv Surg*. 2014;6:260-264
57. Bang OY, Saver JL, Lee KH, Kim GM, Chung CS, Kim SJ, et al. Characteristics of patients with target magnetic resonance mismatch profile: Data from two geographically and racially distinct populations. *Cerebrovasc Dis*. 2010;29:87-94
58. Hacke W, Furlan AJ, Al-Rawi Y, Davalos A, Fiebach JB, Gruber F, et al. Intravenous desmoteplase in patients with acute ischaemic stroke selected by mri perfusion-diffusion weighted imaging or perfusion ct (dias-2): A prospective, randomised, double-blind, placebo-controlled study. *Lancet. Neurology*. 2009;8:141-150
59. Lansberg MG, Thijs VN, Bammer R, Olivot JM, Marks MP, Wechsler LR, et al. The mra-dwi mismatch identifies patients with stroke who are likely to benefit from reperfusion. *Stroke*. 2008;39:2491-2496
60. Mishra NK, Albers GW, Christensen S, Marks M, Hamilton S, Straka M, et al. Comparison of magnetic resonance imaging mismatch criteria to select patients for endovascular stroke therapy. *Stroke*. 2014;45:1369-1374

APPENDIX

I. DEFUSE 3 Patient Informed Consent Form

§ (Section, page number and paragraph)	Original Document	Revised content	Rationale
	Amendment #1		
Throughout protocol; HEADER	**Version 1.3; 14OCT2015**	**Version 2.2; 19-NOV-2015**	
Pg 7	randomized in a 1:1 ratio to treatment with one or more DEFUSE 3 approved (only the devices listed in this IDE application are approved for use in DEFUSE 3) plus standard medical therapy versus standard medical therapy alone. Patients who are enrolled but not randomized, will receive…	randomized in a 1:1 ratio to treatment with **endovascular therapy** (using one or more DEFUSE 3 thrombectomy devices) plus standard medical therapy versus standard medical therapy alone. Patients who are **consented** but not randomized, will receive…	To clarify that patients are not randomized to the devices.
Pg 7	A novel adaptive design (see below) will identify, at interim analysis, the group with the best prospect for showing benefit from endovascular treatment, based on baseline core lesion volumes and the time since stroke onset. Interim analyses will be conducted at 300 and 400 patients	**At the first analysis when 200 subjects complete follow-up, if the overall analysis crosses the futility boundary**, a novel adaptive design will identify, **if it exists, a subgroup** with the best prospect for showing benefit from endovascular treatment, based on baseline **ischemic** core lesion volumes and the **time to treatment**. **The second** interim analysis will be conducted at **340** patients	Clarify timing of interim analyses
Pg 7; Clinical Inclusion Criteria 2	Age 18- 85 years	Age 18 – **90** years	based on the favorable response to endovascular therapy seen in recent trials (ESCAPE and MR CLEAN) in patients >80 years of age.
Pg 7; Clinical Inclusion	Able to undergo an MRI or CT Perfusion within 90 minutes of arrival at the study site.	Deleted	Some patients will have their stroke while hospitalized at the study site, making this inclusion rule irrelevant

Criteria 4			
Pg 7; clinical inclusion criteria	No significant pre-stroke disability (pre-baseline modified Rankin Scale score 0-2)	modified Rankin Scale less than or equal to 2 prior to qualifying stroke (functionally independent for all ADLs)	clarification
Pg 7; Clinical Inclusion Criteria 8	Patient willing/able to return for protocol required follow up visits	Deleted	Investigator can not know with confidence at the time of screening advance if patient willing/able to return for protocol required follow up visits
Pg 8; Clinical Exclusion Criteria 2	or patient is already participating in another drug or device study	Deleted	Deleted duplicate; Criteria is included in Exclusion criteria #14
Pg 8; Clinical Exclusion Criteria 4	Contraindication for MRI (e.g. pacemaker, severe claustrophobia)	Deleted	NA. if MRI is contraindicated, CT can be used
Pg 8; Clinical Exclusion Criteria 5	Contraindication for MRI contrast (renal failure with eGFR < 40ml/mim)	4. Unable to undergo a contrast brain perfusion scan with either MRI or CT	Clarification; if contrast is contraindication for both CT and MR
Pg 8; Clinical Exclusion Criteria 6	Known allergy to iodine and previously refractory to pretreatment medications	5. Known allergy to iodine that precludes an endovascular procedure	Clarification
Pg 7; Clinical Exclusion Criteria		7. Treated with tPA 3-4.5 hours after last known well AND any of the following; age >80, current anticoagulant use, history of diabetes or prior stroke, NIHSS >25	Criteria added per FDA request
Pg 9; Clinical Exclusion Criteria 9	Symptoms consistent with stroke in multiple locations	Deleted	Criteria further clarified in the Imaging Inc/Exc Criteria section
Pg 9; Clinical Exclusion Criteria 10	Seizures at stroke onset if it makes the diagnosis of stroke doubtful and precludes obtaining an accurate baseline NIHSS assessment	9. Seizures at stroke onset if it precludes obtaining an accurate baseline NIHSS	Clarification
Pg 7; Clinical	Severe, sustained hypertension (Systolic Blood Pressure >185 mmHg or Diastolic Blood	12. Severe, sustained hypertension (Systolic Blood Pressure >185 mmHg or Diastolic Blood	Clarification

Exclusion Criteria 13	Pressure >110 mmHg) not treatable with medications	Pressure >110 mmHg)	
Pg 7; Clinical Exclusion Criteria		15. Clot retrieval attempted using a neurothrombectomy device prior to 6 hours from symptom onset	Criteria is not new, but was moved from neuroimaging exclusion criteria
Pg 8; Neuro-imaging exc criteria 4	Evidence of internal carotid artery flow limiting dissection or aortic dissection	Evidence of internal carotid artery dissection that is flow limiting or aortic dissection	Wording clarification
Pg 8; Neuro-imaging exc criteria 5	Intracranial stent implanted in the same vascular territory that would preclude the safe deployment/removal of the neurothrombectomy device	Intracranial stent implanted in the same vascular territory that precludes the safe deployment/removal of the neurothrombectomy device	Wording clarification
Pg 8; Neuro-imaging exc criteria 6	Subjects with occlusions in multiple vascular territories (e.g., bilateral anterior circulation, or anterior/posterior circulation)	Occlusions in multiple vascular territories (e.g., bilateral anterior circulation, or anterior/posterior circulation)	Wording clarification
Pg 9	After enrollment the RAPID output from a CT perfusion or multimodal MRI scan will be assessed. If the patient is confirmed to meet the Neuroimaging eligibility criteria listed above and not have any of the Neuroimaging exclusion criteria, then the patient will be randomized. In general, patients will be enrolled prior to obtaining the RAPID output maps. In some situations the CT perfusion/ multimodal MRI may have been performed as part of standard care *prior* to the patient being assessed for study eligibility. Patients who are enrolled but do not meet the imaging criteria will not be randomized.	After **obtaining consent** the RAPID output from a CT perfusion or multimodal MRI scan will be assessed. If the patient is confirmed to meet the Neuroimaging eligibility criteria listed above and not have any of the Neuroimaging exclusion criteria, then the patient will be randomized. In general, patients will be **consented** prior to obtaining the RAPID output maps. In some situations the CT perfusion/ multimodal MRI may have been performed as part of standard care *prior* to the patient being assessed for study eligibility. Patients who are **consented** but do not meet the imaging criteria will not be randomized.	Wording changed for clarification
Pg 9	Baseline data required for the stratification	Baseline data will be captured for all	clarification

	algorithm, will be captured for all enrolled patients	**consented** patients	
Pg 9	and clinical data for the first 24 hours of their hospitalization will be captured on WebDCU™ (StrokeNet's web-based clinical trials management system).	Deleted	Wording change
Pg 10	The stratification factors used and their hierarchy will be: 1) core lesion volume, 2) age 3) time from symptom onset to enrollment, 4) vessel occlusion (MCA vs, ICA) and 5) study site.	The stratification factors used will be: 1) age, 2) core lesion volume, 3) time from symptom onset to enrollment, 4) **baseline NIHSS** and 5) study site.	Baseline NIHSS felt to be more important than vessel occlusion sites based on new data form the randomized endovascular trials
Pg 10	The optimal biased-coin acceptance region and stratification weights will be determined prior to study launch via simulations.	The biased-coin acceptance region and stratification weights are specified in the Randomization Plan.	clarification
Pg 10	The devices which will be used are the Trevo Retriever, the Solitaire Device Revascularization Device and the Penumbra thrombectomy system.	The devices which will be used are the Trevo Retriever, the Solitaire Revascularization Device, **Covidien MindFrame Capture Revascularization Device** and the Penumbra thrombectomy system.	Added device to include newly 510(k) cleared devices
Pg 11		All brain imaging from stroke onset through hospital discharge, including the	Added to clarify that all brain imaging through discharge will be collected
Pg 11; schedule of events		Updated throughout	Updated for clarification
Pg 11; schedule of events footer	Historical mRS at baseline, mRS to be performed by an mRS certified investigator who is blinded to treatment allocation at 30 and 90 days. For patients who are enrolled but not randomized, the schedule of events is limited to all baseline data evaluations and a summary of stroke therapies received at 24-hrs.	Historical mRS at baseline, mRS/**NIHSS** to be performed by an **NIHSS/**mRS certified **member of the research team** who is blinded to treatment allocation at 30 and 90 days. For patients who are **consented** but not randomized, the schedule of events is limited to a summary of stroke therapies received **within** 24-hrs **of stroke onset**.	clarification
Pg 12	Day 5 (+/- 1 day) or Discharge visit: The items listed for this visit in Table 1 should be performed between day 4 and 6 from the time of randomization. If the patient is discharged prior to day 4, these items should	**Discharge visit:** The items listed for this visit in Table 1 should be performed on the day of hospital discharge	Day 5 visit time point changed to Hospital Discharge visit

	be performed on the day of discharge.		
Pg 12		**Neurological worsening: If clinical worsening (defined as a ≥4 point increase on the NIHSS score) occurs prior to discharge, a CT scan or MRI should be obtained as soon as possible and the NIHSS score (obtained as soon as possible after the worsening was detected) should be documented in the case report form.**	Language added for clarification
Pg 12	All study sites will complete a stroke screening log in the StrokeNet's WebDCU electronic data capture system that documents all patients seen at their center, and reason for exclusion of patients not enrolled. Any patient who experiences a serious adverse effect (SAE) will be reported within 24 hours of the event on the StrokeNet's WebDCU electronic data capture system.	All study sites will complete a stroke screening log that documents all patients treated in the cath lab beyond 6 hours at their center, and reason for exclusion of patients not enrolled, in the StrokeNet's WebDCU electronic data capture system. Serious adverse events (SAEs) will be reported within 24 hours of the event in the StrokeNet's WebDCU electronic data capture system.	Wording changed to clarify the screening log procedure
Pg 113; Site Approval and Monitoring Plan	Approval of neurointerventionalists at each site will follow a similar protocol to a previous endovascular trial.45 A credentialing committee consisting of four experienced neurointerventionalists will determine if a site neurointerventionalist can participate. Approval will require a unanimous vote.	Deleted	Details of approval process for endovascular therapist to be determined by the endovascular committee
Pg 13; Monitoring for Bias	Audits of the hospital database at each site will be performed every 6 months to verify that the endovascular volume matches the sites' screening log. Finally, it is expected that the majority of patients with be enrolled prior to obtaining or evaluation the results of the RAPID maps. The number of cases enrolled *after* the RAPID maps are reviewed will be monitored.	**Deleted**	The plan for monitoring for bias has been simplified as deemed appropriate by the Protocol Director

Pg 15	First interim analysis (n=200 enrolled)	First interim analysis (n=200 **randomized and completed 90 day follow-up**)	Clarification of at what time point the analysis will take place
Pg 15	Second interim analysis (n=340 enrolled)	Second interim analysis (n=340 **randomized and completed 90 day follow-up)**	Clarification of at what time point the analysis will take place
Pg 15	Final analysis (n=476 enrolled)	Final analysis **(n=476 randomized and completed 90 day follow-up)**	Clarification of at what time point the analysis will take place
Pg 16		RAPID-assessed ischemic core volume at baseline will be correlated with 24h infarct volume (DWI volume) in subjects who achieve reperfusion without PH1 or PH2 intracranial hemorrhage. Pearson's correlation coefficient will be calculated and the median absolute error (ml) will be reported. Similarly, correlation of the baseline Tmax >6 volume and the 24h infarct volume in patients without PH1 or PH2 intracranial hemorrhage who have <10% reperfusion will be performed. Correlation of RAPID predicted infarct volume (coregistered baseline ischemic core and 24h Tmax >6 volume) with the actual 24h infarct volume will also be performed.	Further clarification of the statistical plan for handling imaging outcomes
Pg 17	Homogeneity of treatment effects in these subgroups will be determined with the Breslow-Day test.	**deleted**	
Pg 17 "Missing Data/Lost to F/U"	We will perform sensitivity analyses with standard methods for missing data (multiple imputation based on longitudinal models), but we do not expect this to alter the main study results given our estimated very low LTFU rate (<2%).	All effort is put forth to ensure near complete follow-up, in particular with the assessment of the primary outcome (mRS at 90 days), death (mRS=6), and stroke recurrence. If the primary outcome (mRS at 90 days) cannot be assessed in the clinic, it will instead be obtained by phone using a structured interview. If the subject's mRS cannot be obtained in clinic or by phone within the window of 60 to 120 days from randomization, then for primary analyses the day 30 mRS score will be used as the primary outcome (ie day 30 mRS carried	Clarification of how missing data will be handled

		forward). If neither the 30-day nor the 90-day mRS is available, then the mRS will be imputed. We do not expect this to alter the main study results given our estimated very low LTFU rate (<2%).	
Pg 17	**Coordination between MUSC and Stanford statistical teams**: The primary statistician for the study is Dr. Phil Lavori at Stanford. He is voting member of the DEFUSE 3 Executive Committee (EC). Dr. Lavori will be blinded to all outcome data during the study. Dr. Lavori will become unblinded upon database lock and will conduct the final analyses. He will be responsible for developing and writing the statistical analysis plan (SAP) prior to the initiation of the study, and SAP amendments, if any, during the study. The statistical team at MUSC, led by Dr. Yeatts, will be unblinded throughout the study. The MUSC team will implement the adaptive design algorithm developed by Dr. Lavori; conduct and independently validate the interim analyses according to the SAP; generate Open and Closed Reports for the DSMB and interact with the DSMB in closed sessions; and collaborate with Dr. Lavori on validation of final analyses. After database lock, the MUSC statistical team will create the public use datasets (PUDS) and submit them to the NINDS.	deleted	This was deleted from the protocol as it was felt unnecessary. The statistical plan has been developed and is in a separate document (the SAP)
Pg 20	All adverse events (AEs) will be collected, recorded, and analyzed in accordance with Section 3.11 below. All AEs, regardless of severity, anticipation or relationship to the device or procedure will be recorded, collected, and reported.	Adverse events (AEs) will be collected, recorded, and analyzed in accordance with Section 3.11 below.	Language changed for clarification purposes
Pg 21	The primary analysis of angiographic results will utilize the AOL determined from the non-invasive imaging studies. If the non-	deleted	Language removed from protocol as deemed unnecessary

	invasive imaging study is suboptimal the DSA AOL will be utilized. If there is a discrepancy between the Core lab and site investigator the core lab determination will be used for the analysis.		
Pg 21	The unblinded NDMC biostatistical team will generate periodic reports to the study team and the DSMB, perform interim statistical analyses, and interact with the DSMB in closed-session meetings.	deleted	Clarification
Pg 21		• **Penumbra System 110 Aspiration Tubing** • **Penumbra System Reperfusion Catheter ACE64& ACE 68** • **Mindframe Capture LP device**	Added device to include newly 510(k) cleared devices
Pg 21 "Monitoring Procedures"	"Judy Spilker"	**Deleted**	Personnel change
Pg 22 "Monitoring Procedures"	the incidence of symptomatic intracranial hemorrhage at 24 hours and the incidence of significant neurologic deterioration at 5-7 days defined as ≥ 4 point increase in the NIHSS score from the baseline score.	**the incidence of symptomatic intracranial hemorrhage at 36 hours (defined as a >/= 4 point worsening of immediate pre-deterioration NIHSS neurological status vs. post deterioration and associated with brain hemorrhage), and the incidence of significant neurologic deterioration prior to discharge (defined as ≥ 4 point worsening of the immediate pre-deterioration NIHSS neurological status vs. post deterioration and not attributed to sedation).**	Clarification in the definition of neurological worsening
Pg 22	Any patient who experiences an SAE will be reported within 24 hours of the event.	SAEs will be reported within 24 hours of awareness of the event.	To clarify the timeline for reporting SAEs
Pg 22	Dr. Demchuk will prepare regular reports concerning SAEs (not segregated by treatment group) for submission to Dr. Albers, and subsequently to the DSMB.	deleted	Deleted from protocol; information will be specified in the DSMB Charter

Pg 23; Reporting Procedures	After review with the subject by the study site personnel, all Adverse Events occurring during the study, whether or not attributed to the study and/or the devices, observed by the investigator or reported by the subject, will be documented on the appropriate case record form pages.	All Adverse Events, whether or not attributed to the study and/or the devices, observed by the investigator or reported by the subject, will be **recorded from the time of randomization through Day 5 or discharge, whichever is earlier. All SAEs will be recorded through Day 90.**	clarification
Pg 23	6. Intensity	6. Severity	clarification
Pg 24	Intensity is defined as a measure of the severity of a reaction, effect or experience. The measurement(s) are described as mild, moderate or severe. The event itself, however, may be of relative minor medical significance. The intensity of Adverse Events is assessed as mild, moderate or severe according to the following index scale: • *Mild* The Adverse Event is transient, requires no treatment, and does not interfere with the subject's daily activity. • *Moderate* The Adverse Event introduces a low level of inconvenience or concern to the subject and may interfere with daily activities, but is usually ameliorated by simple therapeutic measures. • *Severe* The Adverse Event interrupts the subject's usual daily activity and requires systematic therapy or other treatment.	**Severity** is defined as a measure of the **intensity** of a reaction, effect or experience. The measurement(s) are described as mild, moderate or severe. The event itself, however, may be of relative minor medical significance. The **severity** of Adverse Events is assessed according to the following index scale: • *Mild* *asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated* • *Moderate* minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental Activities of Daily Living. • *Severe* medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care Activities of Daily Living. • **Life-threatening consequences;** urgent intervention indicated • **Death** related to AE	Clarification and revised to match CDEs
	• **None** The Adverse Event is not associated with the study device use.*	• **Unrelated**	

Pg 24	• **<u>Remote</u>** The temporal association is such that the study device is not likely to have had an association with the observed Adverse Event.* • *<u>Possible.</u> This causal relationship is assigned when the Adverse Event:* a) Follows a reasonable temporal sequence from device use, but b) Could have been produced by the subject's clinical state or other modes of therapy administered to the subject. • *<u>Probable.</u> This causal relationship is assigned when the Adverse Event:* a) Follows a reasonable temporal sequence from device use; b) Abates upon discontinuation of the treatment; c) Cannot be reasonably explained by known characteristics of the subject's clinical state. • *<u>Definite.</u> This causal relationship is assigned when the Adverse Event:* a) Follows a reasonable temporal sequence from device use; b) Abates upon discontinuation of the treatment; and c) Is confirmed by the reappearance of the Adverse Event on repeat exposure.	• **Unlikely** • **Reasonable possibility** • **Definitely**	Revised to match CDEs

	*For purposes of reporting UADEs, "None" and "Remote" will be considered as having no association with the device or treatment procedure		
	Amendment #2		
Throughout protocol; HEADER	**Version 2.2; 19-NOV-2015**	**Version 2.3; October 10, 2016**	
Protocol; Page 7	Approximately 35 StrokeNet sites as well as 10 non-StrokeNet sites and 10 "back-up" sites	Approximately 45 StrokeNet sites will be chosen	To reflect change in NIH policy that every DEFUSE 3 site must be in the StrokeNet network as part of an RCC
Protocol; Page 7	If a site does not enroll a patient within 4 months of activation, it will be placed on probation. If no enrollment occurs in the next 2 months, the site will be replaced with a "back-up" site.	If a site does not **consent** a patient within 4 months of activation, it will be placed on probation. If no **patient consent** occurs in the next 2 months, the site will be replaced with a "back-up" site.	Clarifying language
Protocol; Page 7; 3.3.2: Exclusion Criteria # 7	Treated with tPA 3-4.5 hours after last known well AND any of the following; age >80, current anticoagulant use, history of diabetes or prior stroke, NIHSS >25	Treated with tPA 3-4.5 hours after last known well AND any of the following; age >80, current anticoagulant use, history of diabetes **AND** prior stroke, NIHSS >25	Correct typo – exclusion criteria
Protocol; Page 8; Exclusion criteria 16		16. Any other condition that, in the opinion of the investigator, precludes an endovascular procedure or poses a significant hazard to the subject if an endovascular procedure was performed.	Criteria added for patient safety. In the event that the patient meets all other criteria, but the investigator considers IA therapy too risky for the patient, this will give the option to not enroll.

Protocol page 8; Neuroimaging exclusion criteria 1	1. ASPECTS score < 6 on non-contrast CT (if baseline non-contrast CT was performed)	1. ASPECT score <6 on non-contrast CT (if patient is enrolled based on CT perfusion criteria)	Clarification that ASPECT score is only exclusionary if patient is screened for enrollment by CT imaging
Protocol; Page 8; exclusion criteria 6	6. Occlusions in multiple vascular territories (e.g., bilateral anterior circulation, or anterior/posterior circulation)	6. Acute symptomatic arterial occlusions in more than one vascular territory confirmed on CTA/MRA (e.g., bilateral MCA occlusions, or an MCA and a basilar artery occlusion).	clarification
Protocol; Page 10; section 3.5	the goal for femoral artery puncture will be within 45 minutes of randomization; femoral artery puncture must occur within 60 minutes of randomization.	the goal for femoral artery puncture will be within 45 minutes of randomization; femoral artery puncture must occur within **90 minutes of the completion of the qualifying imaging.**	Changed the timing **parameters.** Note the goal remains IA start to be within 45 mins of randomization
Protocol; Page 12	**Neurological worsening:** If clinical worsening (defined as a >4 point increase on the NIHSS score) occurs prior to discharge, a CT scan or MRI should be obtained as soon as possible and the NIHSS score (obtained as soon as possible after the worsening was detected) should be documented in the case report form.	**Neurological worsening:** If clinical worsening (defined as a ≥4 point increase on the NIHSS score) occurs prior to discharge, a CT scan or MRI should be obtained as soon as possible. Neurological worsening is a reportable adverse event.	change language to clarify that an NIHSSS is not documented in CRF for Neuro worsening
Protocol; Page 13; Section 3.7	A site will be activated for enrollment after **four** test cases processed with RAPID have ensured good quality maps.	A site will be activated for enrollment after test cases processed with RAPID have ensured good quality maps.	The number of required test cases differs among sites, and therefore the stipulation of "four" has been removed
Protocol; Page 13; Section 3.7	Enrollment will resume after all imaging problems have been resolved and **two** repeat dummy runs have been obtained that demonstrate adequate image quality.	Enrollment will resume after all imaging problems have been resolved and repeat dummy runs have been obtained that demonstrate adequate image quality.	The number of required dummy runs will differ among sites, and therefore the stipulation of "two" has been removed

Protocol; page 22	the incidence of symptomatic intracranial hemorrhage at 36 hours	the incidence of symptomatic intracranial hemorrhage at 36 hours **from symptom onset**	Added "from symptom onset" for clarification
Protocol; page 22	To minimize bias, he will usually evaluate SAEs blinded to treatment assignment, unless the DSMB approves partial or complete unblinding.		Deleted language to clarify that the safety monitor will not be blinded to treatment arm
Protocol; page 22	In the event of unexpected SAEs or an unduly high rate of SAEs, Dr. Demchuk will promptly contact Dr. Albers and the NINDS Program Official who will notify the DSMB Chair. In the event that he is unavailable for an extended period of time (i.e., extended vacation, sabbatical, illness, etc.), a back-up MSM will be nominated by the study PI and approved by the NINDS Program Official.	In the event of unexpected SAEs or an unduly high rate of SAEs, Dr. Demchuk will promptly contact **the DSMB Liaison** who will notify the DSMB Chair. In the event that he is unavailable for an extended period of time (i.e., extended vacation, sabbatical, illness, etc.), a back-up MSM will be nominated by the study PI and approved by **the DSMB**.	Language to clarify how the correspondence with the DSMB and MSM will be handled
Protocol; page 26	The committee will meet monthly by phone (1 hour/month) for the full duration of the study.	The committee will **typically** meet monthly by phone (1 hour/month) for the full duration of the study.	For clarification purposes
	Amendment #3		
Throughout protocol; HEADER	**Version 2.3; 10-OCT-2016**	**Version 2.4; April 20, 2017**	
Protocol; Page 12	Patients will preferably undergo an MRI/MRA/MR Perfusion at 24 hours; if an MR cannot be performed, a CT/CTA/CTP can be substituted	Patients will preferably undergo an **MRI with MRA AND MR** Perfusion at 24 hours; if an MR cannot be performed, a **CT with CTA AND CTP** can be substituted	To clarify each required image sequence

Protocol; Page 12	MRI data will be electronically transmitted to the coordinating center at Stanford via….	**Imaging** data will be electronically transmitted to the coordinating center at Stanford via …	To clarify that all imaging data, not only MRI data, will be transferred to Stanford
Protocol; Page 23		**This category includes the use of intra-arterial thrombolytics and/or intracranial stents.**	Added per CIRB request

DEFUSE 3

STATISTICAL ANALYSIS PLAN

Prepared by

Greg Albers, MD
Scott Hamilton, PhD
Maarten Lansberg, MD
Philip W. Lavori, PhD
Yuko Y. Palesch, Ph.D.
Sharon D. Yeatts, Ph.D.

Stanford University
and
DEFUSE 3 National Data Management Center
Data Coordination Unit
Department of Public Health Sciences
Medical University of South Carolina
Charleston, SC 29425

Version 1.0
December 29, 2015

Table of Contents

Delet
Delet
Delet
Delet
Delet
Delet

1. List of abbreviations

AE	adverse event
ASPECTS	Alberta Stroke Program Early CT Score
CRF	case report form
CT	computer tomography
CTA	computer tomography angiography
DCU	Data Coordination Unit at the Medical University of South Carolina
DCR	Data Clarification Request
DSMB	Data and Safety Monitoring Board
EC	Executive Committee
ICA	internal carotid artery
ICH	intracranial hemorrhage
IMM	independent medical monitor
ITT	intent-to-treat
IV	intravenous
LTFU	lost to follow up
MCA	middle cerebral artery
MRI	magnetic resonance imaging
MRA	magnetic resonance angiography
mRS	modified Rankin Scale
NDMC	National Data Management Center
NIHSSS	National Institutes of Health Stroke Scale score
NINDS	National Institute of Neurological Disorders and Stroke
OR	odds ratio
RR	relative risk
rt-PA	recombinant tissue plasminogen activator
SAE	serious adverse event
SAP	statistical analysis plan
sICH	symptomatic intracranial hemorrhage
TICI	thrombolysis in cerebral infarction

2. Statistical analysis plan and statistical reports

This document provides the details of the statistical analyses planned for the DEFUSE 3 Trial, including interim analyses for efficacy, futility, and subgroup selection. In addition, it discusses the statistical issues relevant to these analyses (e.g., sample data to be used, imputation of missing data, adjustments for multiplicity, etc.). Deviations to the SAP that are encountered while the trial progresses will be noted as they occur in the Addenda at the end of this document.

The NDMC generates DSMB Reports semiannually. Each semiannual report provides cumulative summary statistics on enrollment; subject status in the study (e.g., number completed 30 and 90 day assessments); baseline characteristics; protocol violations; safety data, including AEs and SAEs by AE code and relatedness to the study intervention; and data management/quality information (e.g., timeliness and completeness of data entry by the clinical centers via the StrokeNet WebDCU Website; number of DCRs generated and resolved). These statistics are reported by treatment group. If a semiannual report coincides in timing with one of the two planned interim analyses, the results are appended to the report.

3. Study Objectives

3.1. Efficacy

The primary objective of the DEFUSE 3 Trial is to determine if ischemic stroke subjects treated in the 6-16 hour time-window with endovascular therapy plus medical management have more favorable functional outcomes at 90 days, defined by mRS score, as compared to subjects treated with medical management alone.

For supportive evidence, the trial plans to evaluate the effectiveness of endovascular therapy plus medical management as compared to medical management alone by other clinical measures (e.g., NIHSSS) and imaging data (e.g., proportion with reperfusion and infarct growth at 24 hours).

3.2. Safety

The safety of endovascular therapy plus medical management as compared to medical management alone is monitored and evaluated by deaths and incidence of sICH, and other SAEs.

4. Study Design

The study has a two-arm parallel design. Eligible subjects are randomized in a 1:1 ratio to endovascular therapy plus medical management or medical management alone. Each subject is followed for 3 months from randomization.

5. Definition of eligible population and choice of analysis set

5.1. Eligible Population

At the outset of the DEFUSE 3 trial, an eligible patient has an acute ischemic stroke, is 18-90 years of age, has an NIHSSS of at least 6 and no more than 24, has no significant pre-stroke disability (pre-baseline mRS of 0-2), can undergo endovascular therapy between 6 and 16 hours of stroke onset, and has evidence of a large vessel occlusion and a large penumbra by neuro-imaging. At one of the two interim analyses, the study inclusion criteria may be altered by the adaptive design (see Section 9). The specific neuro-imaging criteria to qualify for randomization at the onset of the study are:

- ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA or CTA; AND
- Target Mismatch Profile on MRI or CT perfusion (ischemic core volume is <70 ml, mismatch ratio is ≥1.8 and mismatch volume is ≥15 ml as determined by RAPID software).

If perfusion imaging, the MRA, or the CTA is technically inadequate, alternative neuro-imaging inclusion criteria to qualify for randomization are:

A) If CTA (or MRA) is technically inadequate

- Tmax>6s perfusion deficit consistent with an ICA or MCA-M1 occlusion; AND
- Target Mismatch Profile (ischemic core volume is <70 ml, mismatch ratio is ≥1.8 and mismatch volume is ≥15 ml as determined by RAPID software).

B) If MRP is technically inadequate

- ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA (or CTA, if MRA is technically inadequate and a CTA was performed within 60 minutes prior to the MRI); AND
- DWI lesion volume <25 ml

C) If CTP is technically inadequate

- Patient can be screened with MRI and enrolled if neuroimaging criteria are met.

5.2. Adaptive design sample

The adaptive design sample is the group selected by the adaptive design. It includes all randomized subjects (all six cells; Figure 1) if no subgroup is selected; it includes a subset of all randomized subjects if the adaptive design results in the selection of one of five possible subgroups (see Figure 1 and Section 9). The primary efficacy analysis will be conducted in the adaptive design sample (see Section 9).

5.3. Safety sample

The safety sample includes all randomized subjects. Thus, the safety sample is the same, regardless of whether a subgroup is selected by the adaptive design. See Section 13 for details of the safety analyses.

6. Randomization

Randomization takes place centrally via the DEFUSE 3 Trial WebDCU™ website. The randomization scheme is the combination of minimization and the biased coin method and is never deterministic. A dynamic stratification system will ensure well-balanced subgroups. The randomization algorithm will employ biased-coin minimization and the variance method with stratification weights.[1] The strategy is to balance treatment assignment along the marginal distribution of each stratification factor. The stratification factors used and their hierarchy will be: 1) ischemic core volume, 2) age, 3) time from symptom onset to enrollment, 4) NIHSS score and 5) study site. When a new patient is enrolled, the site will enter the stratification factor values into the eCRF (electronic case report form) on WebDCU™. The dynamic randomization algorithm will determine an imbalance measure for each treatment group. The treatment group associated with the smallest imbalance measure will receive the largest probability of assignment in the biased-coin randomization. The randomization algorithm will be programmed into the WebDCU™ and validated using test samples by the NDMC. The detailed randomization scheme, including biased coin acceptance region and stratification weights, and source codes are provided in the Randomization Plan document. (Appendix 1)

7. Blinding

The acute treatment phase of the study is conducted in an open-label manner. However, study investigators who are not directly involved with acute treatment of the subject and who are blinded to treatment assignment will conduct all 30 and 90-day outcome assessments. To maintain blinding of the assessor, subjects are instructed not to discuss their initial hospitalization and treatment with the assessor.

In cases where an unblinded assessor performed the 30- or 90-day assessment (the CRF will capture if the assessor was blinded; this variable will be self-reported by the assessor), such occurrence will be marked as a protocol violation and presented in the final study report; nevertheless, the submitted data are used in the analysis.

8. Missing Data

Based on previous experiences with acute stroke trials, it is anticipated that there will be minimal loss to follow up for the 90-day assessment of the primary outcome. In the IMS I Trial, only 1 of 80 (1.25%) subjects were LTFU. In the IMS II Pilot Study, 2 of 73 (2.7%) were LTFU. In the IMS III Study, 27 of 656 (4.1%) subjects were LTFU. In SWIFT-PRIME, 4 of 191 (2.1%) were LTFU. In Fast-Mag, out of 1700 patients, 4 patients (0.2%) did not have at least day 30 follow-up.

All effort is put forth to ensure near complete follow-up, in particular with the assessment of the primary outcome (mRS at 90 days), death (mRS=6), and stroke recurrence. If the primary outcome (mRS at 90 days) cannot be assessed in the clinic, it will instead be obtained by phone using a structured interview. If the subject's mRS cannot be obtained in clinic or by phone within the window of 60 to 120 days from randomization, then for primary analyses the day 30 mRS score will be used as the primary outcome (ie, the day 30 mRS will be carried forward). If neither the 30-day nor the 90-day mRS is available, then the mRS will be imputed (ie single imputation; source for imputation algorithm to be decided).

9. Primary efficacy analysis

Regardless of whether or not a subgroup is selected, all efficacy outcome measures are primarily analyzed under the ITT principle. Under this principle, each subject is analyzed according to the treatment group to which they were randomly assigned. Definition of the sample included in the primary efficacy analysis (the adaptive design sample) is listed below in Section 9.3.

9.1. Primary outcome

The primary efficacy outcome measure is the mRS score at 90 days from randomization. Missing outcome is imputed according to Section 8.

9.2. Impact of adaptive design on the sample for the primary analysis

An adaptive trial design, developed for DEFUSE 3, will allow the study to test the primary efficacy hypothesis in a subpopulation (the adaptive design sample) if an interim or final analysis indicates futility in the overall population.[2] The design is based on closed testing theory and the group sequential methods for the Generalized Likelihood Ratio (GLR) statistic developed by Lai and Shih.[3] The adaptive design was chosen because there is strong preliminary data suggesting that the effect of endovascular treatment is modified by two baseline variables: ischemic core lesion size and time-to-treatment. The way the adaptive design takes advantage of these biological assumptions (when they are true) is by reallocation of future accrual to the subgroup with the best prospects for showing efficacy. Specifically, if a subgroup is chosen at an interim analysis, subsequent enrollment is limited to patients in that subgroup. As a result, this subgroup will become larger than it would have been in the absence of the adaptive design.

The criterion for deciding which subgroup has the best chance of showing a benefit from endovascular therapy combines both the estimated size of the effect in the subgroup and the sample size of the subgroup. The GLR statistic (Kullback-Leibler criterion) is used to identify this subgroup because it optimally balances those two criteria. It selects the subgroup that has the best chance of showing an effect because it has an apparently large effect and is also of substantial size (note there are 5 subgroups of increasingly larger size, **Figure 1**). The adaptive design employs two biologically-based assumptions to limit the inflation of sample size: a monotonicity / contiguity assumption and an assumption that the effect is largest in the patients with the smallest volumes and the shortest time to randomization (cell C_{11} in **Figure 1**). The cut-points of the categories (cells) will be determined just prior to the first interim analysis (n=200), blinded to subjects' treatment allocations and outcomes, based on the distribution of subjects across the two dimensions (lesion volume and time-to-treatment) to yield six categories (cells) of approximately equal number of subjects.

		Time (hrs)	
		<10	10-16
Core lesion volume (ml)	≤20	C_{11}	C_{21}
	21-50	C_{12}	C_{22}
	51-70	C_{13}	C_{23}

Figure 1. The cohort is stratified according to core lesion volume and time to randomization. Cut-points in the figure serve as examples. Exact cut-points of the stratification will be determined, blinded to treatment allocation and outcome, based on the distribution of subjects across the core and time variables, just prior to the first interim analysis. Based on the results of the 1st interim analysis, enrollment will continue in all 6 cells or the study entry criteria will adapt and enrollment will be limited to one of 5 sub-groups (C_{11}, C_{11+21}, $C_{11+21+12}$, $C_{11+21+12+22}$, or $C_{11+21+12+22+13}$).

9.3. Interim and final statistical analyses

The primary endpoint is the distribution of scores on the modified Rankin Scale (mRS) at day 90. We will test the primary efficacy and futility hypotheses at the interim and final analysis using the generalized likelihood ratio (GLR) test, based on whether the usual normal approximation to the Wilcoxon-Mann-Whitney test statistic crosses a futility or efficacy boundary at interim or final analysis. The primary analysis will be conducted in the adaptive design sample (see section 5.2 for definition), according to the intention to treat principle, adjusted for the adaptive design, and unadjusted for covariates. See Appendix 2 for detailed specification of the interim and final calculations.

The efficacy bounds at interim and final analysis are set to control the overall (one-sided) Type I error rate at 2.5%. There are three group sequential boundaries: an interim futility boundary b_f, an interim efficacy boundary b_e, and a final efficacy bound c, which are fixed before the first interim analysis. At each of the two interim analyses, the futility bound b_f is used to decide if the study should continue recruitment in the overall group, shift accrual and testing to a subgroup, or stop in its entirety. The futility boundary takes into account the fact that the maximum analyzed sample size is a random variable that is no larger than the fixed maximum number of subjects randomized (n=476). Because subgroup selection reduces the maximum number of subjects available for analysis at completion of the study, this method effectively allows an easier futility stop after subgroup selection. This setup replaces conditional power analyses with an automatic and more efficient adjustment of boundaries.

If the patients are equally distributed across the six cells, the efficacy boundary at the interim analyses will be 2.62, the efficacy boundary at the final analysis will be 2.61, and the futility boundary at the interim analyses will be -1.88. If the patients are not distributed equally, these boundaries will be adapted slightly.[2]

First interim analysis (based on primary outcome data obtained from the first 200 consecutively randomized subjects)

The null hypothesis is tested in the entire subject population, and, depending on the results:

1. If neither efficacy nor futility bound is crossed, the trial continues with enrollment in the overall population to the 2nd interim analysis.
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the overall population.
3. If the futility bound is crossed, the optimal subgroup is selected based on the Kullback-Leibler criterion and the null is tested in that subgroup.
    - 3.1. If neither bound is crossed, the trial will continue with enrollment limited to the selected subgroup.
    - 3.2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
    - 3.3. If the futility bound is crossed, the trial stops for futility.

Second interim analysis (after primary outcome data obtained from an additional 140 consecutively randomized subjects within the target population defined by the first interim analysis)

If, after the first interim analysis, the study proceeds with enrollment in the overall population (option 1, under first interim analysis), the testing at the 2nd interim analysis is identical to the first interim and the decisions to stop or proceed with enrollment are identical to those outlined above under the first interim analysis.

If, after the first interim analysis, enrollment is limited to a selected subgroup (option 3, under first interim analysis), the second interim analysis is based on a test of the null hypothesis in the selected subgroup only and, depending on the results:

1. If neither bound is crossed, the trial continues to the final analysis with enrollment of additional subjects limited to the selected subgroup
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
3. If the futility bound is crossed, the trial stops for futility. Note that there is no option for “second subgroup selection”.

Final analysis (after primary outcome data obtained from an additional 136 consecutively randomized subjects within the target population defined by the second interim analysis)

If, after the second interim analysis, the study proceeds with enrollment in the overall population the null is tested in the overall population, and, depending on the results:

1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in the overall population.
2. If the efficacy bound is not crossed, the optimal subgroup is selected and the null is tested in that group:
    - 2.1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in that subgroup
    - 2.2. If the efficacy bound is not crossed, endovascular therapy will be declared of no benefit.

If enrollment after one of the interim analyses is limited to a selected subgroup, then at the final analysis the null will be tested in that subgroup only and efficacy or lack thereof will be declared as per options 2.1 and 2.2 above.

The computation of the test statistics and the Kullback-Leibler selection criterion are specified in Appendix 2. Calculation of the test statistics will be carried out in SAS by the NDMC. A more detailed description of the adaptive design is provided in Appendix 3.

9.4. Reporting of primary results

The results of the study will be primarily expressed as whether or not an efficacy boundary was crossed at either one of the two interim analyses or at the final analysis. The efficacy boundary is set, a priori, to guarantee less than a 2.5% one-sided error rate. Crossing of the efficacy boundary will be considered evidence that endovascular therapy is beneficial, based on lower day-90 mRS scores in the endovascular group compared to controls.

9.5. Estimation of p-values, effect size estimates, and CIs

The treatment effect will be adjusted for study design and expressed as

- The Wilcoxon-Mann-Whitney measure of superiority (WMW), with its 95% confidence interval and p-value
- The average risk difference (ARD), defined as the probability that a randomly selected subject from the endovascular group has a better functional outcome than a randomly selected subject from the control group, with its 95% confidence interval, where ARD = WMW – 0.5.
- The average number needed to treat for benefit (NNT), with its 95% confidence interval, where NNT = 1 / ARD.
- The common odds ratio with its 95% confidence interval and p-value, calculated using a proportional odds model.

10. Sample size determination for primary efficacy analysis

The sample size determination begins with a preliminary estimate of the effect size that is plausible to expect and which is also clinically meaningful. Fixing standard operating characteristics at 5% two-sided Type 1 error and 10% Type 2 error (90% power) leads to a sample size lower bound for a hypothetical fixed-sample trial with no adaptation and no interim analysis. We then adjust the sample size for the group-sequential modification and the subgroup adaptive design.

The projected overall effect of endovascular therapy is based on 1) the observed 90-day modified Rankin Scale outcomes in DEFUSE 2 of target mismatch patients treated >6hrs after symptom onset and 2) the assumption that early reperfusion will be achieved in 75% of the endovascular arm vs. 20% of the medical therapy arm.[4-6] Using these data, we have projected the distributions on the mRS at 90 days for subjects in the endovascular and control arms of DEFUSE 3. (Table 1)

	mRS at day 90							
	0	1	2	3	4	5	6	Total
Endovascular group	18.0%	11.5%	19.6%	11.5%	16.4%	11.5%	11.5%	**100%**
Medical group	9.7%	7.9%	15.0%	17.7%	14.4%	17.7%	17.7%	**100%**

Table 1. Projected 90-day mRS distributions in DEFUSE 3

These distributions correspond to a standardized effect of 0.36 for the primary analysis. Based on these data, the fixed sample size for a non-adaptive design requires 376 patients (188 per arm) to have 90% power at an alpha of 5% (Wilcoxon-Mann-Whitney test); 100 patients are added for the adaptive design to reach a maximum sample size of 476 for DEFUSE 3. The size of this increase is based on simulations and is selected to preserve the desired operating characteristics, while allowing shrinkage in effect size to 0.30, since the above estimate of 0.36 may be optimistic. The sample size of 476 is also the largest sample that can be accrued within budget and time limitations.

Simulations (n=5000) are used to compare the performance of a traditional fixed sample-size design (fixed n=476) to the adaptive design (max randomized n=476) under the null and various alternative scenarios (**Table 2**). For the simulations the effect size is expressed as a standardized effect *in the disjoint cells*, which are cumulated to form the subgroups as described above, where a standardized effect of 0.3 corresponds to a conservative projected effect of endovascular therapy (anticipated effect 0.36; see above).

Sim.	Standardized effect in cells* C_{11},	C_{21},	C_{12},	C_{22},	C_{13},	C_{23}	Average standard. effect	**Adaptive Design** Average No. randomized	Power	**Fixed Design** Number randomized	Power
#0	0	0	0	0	0	0	0	361	2.2%	476	2.5%
#1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	354	80%	476	89%
#2	0.5	0.4	0.3	0	0	0	0.2	400	86%	476	55%
#3	0.5	0.5	0	0	0	0	0.17	403	87%	476	41%

Table 2. **Simulations comparing the adaptive and the fixed trial designs**. *Cells are defined in figure 1. Under the null (scenario #0), the adaptive design controls the total Type 1 error below 2.5%, stops early for futility 63% of the time, and the average number of randomizations is 361. If the effect is uniform across cells - and therefore also over cumulated subgroups - (scenario #2), the fixed-sample design is optimal, but the adaptive design results in only a small loss of power (from 89 to 80%). When the effect size distribution across the subgroups is in accord with the biological assumptions (scenarios #2 and 3), so that the effect in the cumulated subgroups declines as more cells with a null effect are added in, the adaptive design performs much better (higher power and smaller expected sample size) than the fixed sample, conventional trial. The adaptive design also performs well compared to a non-adaptive, fixed sample that includes efficient multiple comparisons-adjusted testing for effect in subgroups at the end of the study.[2] Note that in all of the tabulated examples (other than the first, which is the null case), the overall population null is false, because the effect is positive in some cell or cells, and the effect is null in the others, so that the effect never becomes negative. If we allow negative effects in any of the cells, then the advantages of the adaptive design over the conventional design grow even larger.

11. Exploratory analyses of the primary outcome after trial completion

11.1. Analysis in Subgroups Defined by Sex, Race, and Ethnicity

The primary efficacy analysis will be repeated in subgroups defined by sex, race (White, African-American, Asian, Other), and ethnicity (Hispanic, Non-Hispanic). The following treatment effects and corresponding 95% confidence intervals will be estimated, as in Section 9.5:

- The Wilcoxon-Mann-Whitney measure of superiority (WMW)
- The average risk difference (ARD), defined as the probability that a randomly selected patient from the endovascular group has a better functional outcome than a randomly selected patient from the control group, with its 95% confidence interval, where ARD = WMW – 0.5.
- The average number needed to treat for benefit (NNT)

11.2.Analysis adjusting for covariates

The secondary analysis of the primary endpoint will be the same rank-based analysis comparing the distribution of the 90-day Rankin scores between treatment groups while stratifying for prognostically important covariates. The Generalized Cochran-Mantel-Haenszel test (CMH) is the analogue to the WMH, designed to test against the alternative that there is a uniform shift of size "delta" in the Rankin score distribution from one group to the other after stratification. The CMH test will be stratified by age, baseline NIHSS Score, Baseline Ischemic Core Volume, and Time from symptom onset to randomization. A rich body of clinical research has established the prognostic importance of these covariates. Cut-points will generally follow those used in the randomization, but may be altered depending on the numbers of patients enrolled into each category.

To assess the relationship between important covariates and the size of the treatment effect, a model based regression analysis will be performed. Specifically, we will create a multivariable ordinal regression model with the 90-day Rankin outcome as the dependent variable and use a literature based model-building process. The model building process will include preliminary variable selection, a final model selection, and a final model assessment. We will evaluate each candidate predictor for proportionality and linearity (model assumptions). The correlation between candidate continuous variables will be analyzed. Variables with a rho coefficient higher than 0.7 will not be jointly entered into the final model building process - only the variable with the lowest p-value if it is less than 0.1. The base model will include treatment group. During the preliminary variable selection, potential predictors will be added to the base model one at a time to obtain a p-value for each. Variables considered candidates for the final model will have a p-value ≤ 0.1. At the final model selection, variables will be added sequentially starting with the variable with the lowest p-value from the group of candidate predictors. The criteria for keeping a variable in the final model will be a p-value ≤ 0.05. Each time a predictor is kept in the final model all previously added variables will be re-assessed. Any previously entered variables whose p-value has increased above 0.05 will be dropped from the model. All qualifying variables from the preliminary selection phase will be considered along with their two way interaction with site of arterial occlusion. A statistical assessment of the model will be examined by the shrinkage statistic. A shrinkage statistic below 0.85 will indicate the model is overfitting the data, and the number of predictors should be reduced. The following variables will be considered for inclusion in the adjusted model:

- Age
- Baseline NIHSS score
- Baseline ischemic core volume
- Time from symptom onset to randomization
- Sex
- Admission SBP
- Baseline glucose

11.3.Analysis under the as-treated principle

Due to the nature of the study, some subjects in the endovascular treatment plus medical management group may not receive endovascular therapy. One instance when this may occur is if symptoms resolve spontaneously between randomization and start of the endovascular procedure. Another potential reason is the absence of an arterial occlusion on the baseline angiogram (due to spontaneous recanalization). Although it would be rare, we

may also see some subjects who are randomized to the medical management only group, but are treated with endovascular therapy. Therefore, after completion of the study, we will repeat the primary analysis (the Wilcoxon-Mann-Whitney measure of superiority) within the adaptive design sample, but with patients categorized under the as-treated principle.

11.3.1. Group assignment under the "as-treated" principle

- **Endovascular therapy**: Patients who present to the endovascular suite (cath lab) and undergo a femoral puncture within 24 hours after time of onset of the qualifying stroke are assigned to the endovascular treatment arm under the "as-treated" principle.
- **Medical therapy**: Patients who do not meet the above criteria for endovascular treatment are assigned to the medical treatment arm under the "as-treated" principle.

12. Analyses of secondary efficacy outcomes

12.1.Analyses of secondary clinical efficacy outcomes

The effect of endovascular treatment will be assessed on two secondary clinical efficacy outcomes: 1) functional independence at 90 days, defined as an mRS score ≤2 at day 90; and 2) change in NIHSS score between baseline and 24 hours of >8 points or a 24-hour NIHSSS of 0-1. Results for both outcomes will be expressed as an unadjusted risk ratio with its 95% confidence interval and p-value.

12.2.Analyses of functional independence in subgroups

The unadjusted effect of endovascular treatment on the secondary efficacy outcome "functional independence" (mRS ≤2) will be analyzed in the following subgroups, assuming sufficient numbers of subjects are enrolled in each subgroup. The effect of endovascular therapy will be estimated in subgroups defined by sex, race, and ethnicity, regardless of the number of subjects enrolled in each subgroup, as suggested by NIH guidance. Results will be expressed as unadjusted risk ratios with their 95% confidence intervals and p-values. We will assess for differences in the treatment effect between groups (eg, age <70 vs $\geq$70) using a Breslow-Day test for binary covariates, or by including an interaction term into a logistic regression model for covariates with greater than 2 categories.

- Time from symptom onset to randomization (using cutpoint from adaptive design)
- Baseline ischemic core lesion volume (using cutpoints from adaptive design)
- Age at randomization (<70 and $\geq$70 years old)
- Baseline NIHSSS (<18 and ≥18)
- Baseline ASPECTS (<8 and $\geq$8)
- Primary occlusion site (M1 and ICA)
- IV rt-PA treatment (yes and no)
- Sex
- Race (White, African American, Asian, Other)
- Ethnicity (Hispanic and Non-Hispanic)
- Baseline atrial fibrillation (yes and no)

12.3.Analyses of imaging efficacy outcomes

We hypothesize that endovascular treatment improves radiological outcomes. We will compare four imaging outcomes between treatment groups:

- The proportions of successful reperfusion, where successful reperfusion is defined as a >90% reduction in the volume of brain tissue with critical hypoperfusion (Tmax>6sec) between baseline and 24 hours. Results will be expressed as an unadjusted risk ratio with its 95% confidence interval and p-value.
- The proportion of subjects with recanalization of the primary arterial occlusive lesion (ICA or M1) at the time of the 24-hour follow-up scan is compared between the two treatment arms. Results will be expressed as an unadjusted risk ratio with its 95% confidence interval and p-value.
- Infarct volumes at 24 hours, defined as the lesion volume as outlined on the 24-hour DWI. The 24-hour endpoint is based on data demonstrating that assessment of infarct volume at 24 hours captures the effect of reperfusion therapies on infarct growth and predicts outcomes similarly to day 90 infarct volumes.[7, 8]
- Ischemic lesion growth between baseline and 24 hours defined as the difference between the baseline ischemic core lesion volume and the 24-hour DWI lesion volume. Absolute lesion growth will be calculated by subtracting the baseline ischemic core lesion volume from the 24-hour DWI lesion volume. Relative lesion growth will be calculated by dividing the 24-hour DWI lesion volume by the baseline ischemic core lesion volume. DEFUSE 2 demonstrated a substantial reduction in infarct growth among Target mismatch patients treated in the 6-12 hour time-window who achieved early reperfusion: median growth 0.5 ml (IQR: -2 – 10) with reperfusion (n=23) vs. 39 ml (IQR: 18-121) without reperfusion (n=13), p<0.001.

  Cube root and log transformations of the absolute lesion growth and relative lesion growth observations will be applied to achieve approximate Gaussian distributions. The transformation that maximizes the Gaussian likelihood multiplied by the Jacobian will be used and the resulting set of observations will be tested for Gaussian properties by standard techniques. If neither transformation adequately achieves a Gaussian distribution, the Wilcoxon rank-sum will be used to compare the treatment groups. Stratification of the test by age, baseline NIHSS, and time to randomization categories will be considered. If the transformations are successful, the transformed absolute and relative lesion growth outcomes will be analyzed by linear regression with the treatment group and independently predictive baseline covariates. Potential baseline covariates (age, NIHSS, time to randomization, etc.) will be evaluated for inclusion in the linear regression models using standard model building techniques.

We also hypothesize that final infarct volumes (assessed on 24-hour DWI) can be predicted based on baseline ischemic core and critically hypoperfused tissue volumes. To analyze this relationship, subjects will be divided into two groups: 1) patients with reperfusion, defined as >90% reduction in the Tmax>6-second lesion volume between baseline and 24 hours (or TICI 2b–3 at end of procedure if perfusion imaging inadequate) and (2) a “no reperfusion” group, defined as <10% reperfusion at 24 hours. Spearman’s rho will be used to assess correlations between: 1) baseline ischemic core volume and 24-hour DWI lesion volume in patients with reperfusion; and 2) the union of baseline ischemic core volume and 24-hour critically hypoperfused lesion volume (Tmax6) and 24-hour DWI lesion volume in patients without reperfusion. Baseline demographics and disease characteristics along with treatment group will be evaluated for inclusion in the model. If necessary, transformation of the volume measures will be used to achieve a more Gaussian distribution.

13. Safety analyses

All safety outcome measures are analyzed under the as-treated principle. Under this principle, each subject is analyzed according to the treatment that the subject received. The external Medical Safety Monitor and the DSMB will monitor safety variables throughout the study at frequent intervals. Details of this process will be specified in the Safety Monitoring Plan.

13.1. Monitoring of deaths and sICH

Deaths and sICH rates will be monitored quarterly. Symptomatic ICH or death rates that exceed pre-specified thresholds will trigger a meeting of the DSMB. The DEFUSE 3 study will be placed on hold if it is determined with 95% probability that either:

1) the symptomatic ICH rate (NIHSS worsening of 4 or more points associated with ICH) exceeds 10% in the endovascular group; OR
2) the 90-day mortality rate exceeds 20% in the endovascular group.

The study will remain on hold until the investigators and the DSMB can conduct a review of events and make a determination on the continuation of the trial.

13.2. Summary of Adverse Events (AEs) and Serious Adverse Events (SAEs)

All AEs and SAEs are summarized by MedDRA preferred term (as coded based on the AE CRF) and by treatment group in terms of frequency of the event, number of subjects having the event, timing relative to randomization, and relatedness to the intervention.

For the following specific events, the proportions by treatment group, as well as the RR and corresponding 95% confidence interval, are provided:

- Stroke related mortality within 90 days of randomization
- sICH within 36 hours of randomization
- Significant neurological deterioration prior to discharge (defined as an increase of 4 points or more on the NIHSS)
- PH1 and PH 2 rates on the 24 hr scan

At the end of the study, the cumulative incidences of these events are compared between the two treatment groups using Fisher's exact test.

All brain scans obtained prior to discharge will also be assessed for PH by the core lab.

14. Coordination between Stanford and MUSC statistical teams

The primary statistician for the study is Dr. Phil Lavori at Stanford. He is a voting member of the DEFUSE 3 Executive Committee (EC). Dr. Lavori will be blinded to all outcome data during the study. Dr. Lavori will become unblinded upon database lock and will conduct the final analyses. He will be responsible for developing and writing the statistical analysis plan (SAP) prior to the initiation of the study, and SAP amendments, if any, during the study. The statistical team at MUSC, led by Dr. Yeatts, will be unblinded throughout the study. The MUSC team will implement the adaptive design algorithm developed and written by Dr. Lavori; conduct and independently validate the interim analyses according to the SAP; generate Open and Closed Reports for the DSMB and interact with the DSMB in closed sessions; and collaborate with Dr.

Lavori on validation of final analyses. After database lock, the MUSC statistical team will create the public use datasets (PUDS) and submit them to the NINDS.

15. References

1. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*. 1975;31:103-115
2. Lai TL, Lavori PW, Liao OY. Adaptive choice of patient subgroup for comparing two treatments. *Contemp Clin Trials*. 2014;39:191-200
3. Lai TL, Shih M-C. Power, sample size and adaptation considerations in the design of group sequential clinical trials. *Biometrika*. 2004;91:507-528
4. Nogueira RG, Lutsep HL, Gupta R, Jovin TG, Albers GW, Walker GA, et al. Trevo versus merci retrievers for thrombectomy revascularisation of large vessel occlusions in acute ischaemic stroke (trevo 2): A randomised trial. *Lancet*. 2012;380:1231-1240
5. Saver JL, Jahan R, Levy EI, Jovin TG, Baxter B, Nogueira RG, et al. Solitaire flow restoration device versus the merci retriever in patients with acute ischaemic stroke (swift): A randomised, parallel-group, non-inferiority trial. *Lancet*. 2012;380:1241-1249
6. Pereira VM, Gralla J, Davalos A, Bonafe A, Castano C, Chapot R, et al. Prospective, multicenter, single-arm study of mechanical thrombectomy using solitaire flow restoration in acute ischemic stroke. *Stroke*. 2013;44:2802-2807
7. Campbell B, Purushotham A, Christensen S, Desmond P, Nagakane Y, Parsons M, et al. The infarct core is well represented by the acute diffusion lesion: Sustained reversal is infrequent. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2012;32:50-56
8. Campbell BC, Tu HT, Christensen S, Desmond PM, Levi CR, Bladin CF, et al. Assessing response to stroke thrombolysis: Validation of 24-hour multimodal magnetic resonance imaging. *Arch Neurol*. 2012;69:46-50

DEFUSE 3

STATISTICAL ANALYSIS PLAN

Prepared by

Greg Albers, MD
Scott Hamilton, PhD
Maarten Lansberg, MD
Philip W. Lavori, PhD
Yuko Y. Palesch, Ph.D.
Sharon D. Yeatts, Ph.D.

Stanford University
and
DEFUSE 3 National Data Management Center
Data Coordination Unit
Department of Public Health Sciences
Medical University of South Carolina
Charleston, SC 29425

Version 2.0, June 3, 2017 (Added changes to address early analysis as described in Section 2)

Version 3.0, August 24, 2017 (Added clarifications for programming purposes prior to unblinding of Stanford Study Team)

Version 3.1, September 8, 2017 (Added clarifications for programming purposes prior to unblinding of Stanford Study Team)

Table of Contents

1. List of abbreviations

AE	adverse event
ASPECTS	Alberta Stroke Program Early CT Score
CRF	case report form
CT	computer tomography
CTA	computer tomography angiography
DCU	Data Coordination Unit at the Medical University of South Carolina
DCR	Data Clarification Request
DSMB	Data and Safety Monitoring Board
EC	Executive Committee
ICA	internal carotid artery
ICH	intracranial hemorrhage
IMM	independent medical monitor
ITT	intent-to-treat
IV	intravenous
LTFU	lost to follow up
MCA	middle cerebral artery
MRI	magnetic resonance imaging
MRA	magnetic resonance angiography
mRS	modified Rankin Scale
NDMC	National Data Management Center
NIHSSS	National Institutes of Health Stroke Scale score
NINDS	National Institute of Neurological Disorders and Stroke
OR	odds ratio
RR	relative risk
rt-PA	recombinant tissue plasminogen activator
SAE	serious adverse event
SAP	statistical analysis plan
sICH	symptomatic intracranial hemorrhage
TICI	thrombolysis in cerebral infarction

2. Statistical analysis plan, statistical reports and early interim analysis

This statistical analysis plan (SAP) was modified on June 3, 2017 to accommodate significant external events that occurred in May 2017. On May 16, 2017, the results of DAWN, a clinical trial that enrolled similar patients and studied a similar intervention as DEFUSE 3, were presented at an international meeting. The results of this study demonstrated a substantial clinical benefit of endovascular therapy over medical therapy. Based on the results of the DAWN study, on May 24, 2017, the DEFUSE 3 Central IRB requested that enrollment in DEFUSE 3 be halted; at that time, 182 patients had been randomized. On May 26, 2017, the DSMB and the DEFUSE 3 Executive Committee both recommended that the trial be halted. The NINDS instead requested an early data analysis and revision of the DEFUSE 3 Statistical Analysis Plan (SAP) to provide for the option of continuing the study if the early interim analysis did not cross the stopping boundary for efficacy. Based on this request, the SAP has been modified to provide for this unplanned early analysis. This analysis will be conducted as follows:

Upon database freeze, the DEFUSE 3 unblinded Statistician (Dr. Sharon Yeatts) will perform a test of the overall null hypothesis for the primary endpoint (mRS shift analysis) in the full DEFUSE 3 sample at the one-sided Type 1 error probability of 0.023 (additional details of how this analysis will be performed are presented in section 9.4). For the early interim analysis, we will impute the

90-day mRS for subjects who have not yet reached the end of the study protocol with their respective 30-day mRS. Subjects with no 30-day or 90-day mRS scores will be excluded. The DSMB will be provided with the results of this analysis and would likely recommend stopping the study if the result is significant at the one-sided 0.023 level. Subsequently, upon submission of the last-randomized subject's 90-day outcome data into the WebDCU™ (anticipated in mid-August 2017) and the database lock, the analysis will be repeated at the one-sided alpha level of 0.023. The manuscript for the DEFUSE 3 Trial results will reflect this second analysis.

If the initial analysis result (i.e., based on the frozen, not locked, database) is not significant at the one-sided alpha level of 0.023, the DSMB will review the result and may request any additional analyses in any subgroups that they desire in order to make recommendations to NINDS. Following the DSMB review and recommendation, the NINDS may make the decision to continue the study with recruitment of the DEFUSE 3 eligible patients per the current eligibility criteria, and the study may continue per the current protocol. In this case, the final primary outcome analysis will be tested at the one-sided alpha=0.001 (which will be referred to as "alpha2"). If the decision is made to terminate the study, upon submission of the last-randomized subject's 90-day outcome data into the WebDCU™ (anticipated in mid-August 2017) and the database lock, the analysis will be repeated at the one-sided alpha level of 0.023.

The NINDS may also direct the study team to continue the study in a subset of the current DEFUSE 3 eligible patient population (such as DAWN-ineligible patients). If this is deemed feasible, then the final analysis will include all patients enrolled (both before and after the subset selection was imposed) and will be evaluated with a one-sided alpha of 0.001 (which will be referred to as alpha3).

The rationale for recommending the extreme split of the alpha is that:
(a) given the conditions currently imposed by the CIRB (i.e., recruit only DAWN-ineligible patients), continuing the DEFUSE 3 study with the current eligibility criteria is unlikely to be feasible; and
(b) the sites have reported significant concerns regarding recruiting the DAWN-ineligible patients due to its complex definition.

To date, the DEFUSE 3 Executive Committee has proposed and the DSMB had effectively recommended alpha2 + alpha3 =0, since they do not envision any circumstance where continuation of DEFUSE 3 in its current form is feasible. If a continuation is requested by the NINDS, the sample size will need to be reassessed based on the desired effect in the modified study population and the one-sided alpha2/alpha3 required for efficacy. This evaluation will likely result in increased total sample size for the study.

This proposed plan supersedes the prior SAP. No further interim analyses for adaptation will be conducted, and Sections 9.2-9.4 of the prior SAP are now deleted. Other changes include: Subgroup analysis related to time from symptom onset to randomization and baseline core volume (Section 12.2 of the prior SAP) define subgroups based on medians, rather than "cutpoint from adaptive design". An additional pre-specified exploratory analysis has been included that compares the primary and secondary efficacy and safety endpoints in DAWN-eligible vs DAWN-ineligible patients (see section 11.1). The remainder of the SAP will be followed as pre-specified prior to this revision.

This document provides the details of the statistical analyses planned for the DEFUSE 3 Trial, including the original interim analyses for efficacy, futility, and subgroup selection, and the revision to reflect the early interim analysis for overwhelming efficacy. In addition, it discusses

the statistical issues relevant to these analyses (e.g., sample data to be used, imputation of missing data, adjustments for multiplicity, etc.).

The NDMC generates DSMB Reports semiannually. Each semiannual report provides cumulative summary statistics on enrollment; subject status in the study (e.g., number completed 30 and 90 day assessments); baseline characteristics; protocol violations; safety data, including AEs and SAEs by AE code and relatedness to the study intervention; and data management/quality information (e.g., timeliness and completeness of data entry by the clinical centers via the StrokeNet WebDCU™ Website; number of DCRs generated and resolved). These statistics are reported by treatment group.

3. Study Objectives

3.1. Efficacy

The primary objective of the DEFUSE 3 Trial is to determine if ischemic stroke subjects treated in the 6-16 hour time-window with endovascular therapy plus medical management have more favorable functional outcomes at 90 days, defined by mRS score, as compared to subjects treated with medical management alone.

For supportive evidence, the trial plans to evaluate the effectiveness of endovascular therapy plus medical management as compared to medical management alone by other clinical measures (e.g., mRS 0-2 outcomes at 90 days) and imaging data (e.g., proportion with reperfusion and infarct growth at 24 hours).

3.2. Safety

The safety of endovascular therapy plus medical management as compared to medical management alone is monitored and evaluated by deaths and incidence of sICH, and other SAEs.

4. Study Design

The study has a two-arm parallel design. Eligible subjects are randomized in a 1:1 ratio to endovascular therapy plus medical management or medical management alone. Each subject is followed for 3 months from randomization.

5. Definition of eligible population and choice of analysis set

5.1. Eligible Population

At the outset of the DEFUSE 3 trial, an eligible patient has an acute ischemic stroke, is 18-90 years of age, has an NIHSSS of at least 6 and no more than 24, has no significant pre-stroke disability (pre-baseline mRS of 0-2), can undergo endovascular therapy between 6 and 16 hours of stroke onset, and has evidence of a large vessel occlusion and a large penumbra by neuro-imaging. ~~At one of the two interim analyses, the study inclusion criteria may be altered by the adaptive design (see Section 9).~~ The specific neuro-imaging criteria to qualify for randomization at the onset of the study are:

- ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA or CTA; AND
- Target Mismatch Profile on MRI or CT perfusion (ischemic core volume is <70 ml, mismatch ratio is ≥1.8 and mismatch volume is ≥15 ml as determined by RAPID software).

If perfusion imaging, the MRA, or the CTA is technically inadequate, alternative neuro-imaging inclusion criteria to qualify for randomization are:

A) If CTA (or MRA) is technically inadequate
- Tmax>6s perfusion deficit consistent with an ICA or MCA-M1 occlusion; AND
- Target Mismatch Profile (ischemic core volume is <70 ml, mismatch ratio is ≥1.8 and mismatch volume is ≥15 ml as determined by RAPID software).

B) If MRP is technically inadequate
- ICA or MCA-M1 occlusion (carotid occlusions can be cervical or intracranial; with or without tandem MCA lesions) by MRA (or CTA, if MRA is technically inadequate and a CTA was performed within 60 minutes prior to the MRI); AND
- DWI lesion volume <25 ml

C) If CTP is technically inadequate
- Patient can be screened with MRI and enrolled if neuroimaging criteria are met.

5.2. ~~Adaptive design sample~~Analysis Sample

~~The adaptive design sample is the group selected by the adaptive design. It includes all randomized subjects (all six cells; Figure 1) if no subgroup is selected; it includes a subset of all randomized subjects if the adaptive design results in the selection of one of five possible subgroups (see Figure 1 and Section 9). The primary efficacy analysis will be conducted in the adaptive design sample (see Section 9).~~ The analysis sample consists of all randomized patients and the primary analysis will be conducted per the ITT principle.

5.3. Safety sample

The safety sample includes all randomized subjects. Thus, the safety sample is the same, regardless of whether a subgroup is selected to continue enrollment.~~by the adaptive design.~~ See Section 13 for details of the safety analyses.

6. Randomization

Randomization takes place centrally via the DEFUSE 3 Trial WebDCU™ website. The randomization scheme is the combination of minimization and the biased coin method and is never deterministic. A dynamic stratification system will ensure well-balanced subgroups. The randomization algorithm will employ biased-coin minimization and the variance method with stratification weights.[1] The strategy is to balance treatment assignment along the marginal distribution of each stratification factor. The stratification factors used and their hierarchy will be: 1) ischemic core volume, 2) age, 3) time from symptom onset to enrollment, 4) NIHSS score and 5) study site. When a new patient is enrolled, the site will enter the stratification factor values into the eCRF (electronic case report form) on WebDCU™. The dynamic randomization

algorithm will determine an imbalance measure for each treatment group. The treatment group associated with the smallest imbalance measure will receive the largest probability of assignment in the biased-coin randomization. The randomization algorithm will be programmed into the WebDCU™ and validated using test samples by the NDMC. The detailed randomization scheme, including biased coin acceptance region and stratification weights, and source codes are provided in the Randomization Plan document. (Appendix 1)

7. Blinding

The acute treatment phase of the study is conducted in an open-label manner. However, study investigators who are not directly involved with acute treatment of the subject and who are blinded to treatment assignment will conduct all 30 and 90-day outcome assessments. To maintain blinding of the assessor, subjects are instructed not to discuss their initial hospitalization and treatment with the assessor.

In cases where an unblinded assessor performed the 30- or 90-day assessment (the CRF will capture if the assessor was blinded; this variable will be self-reported by the assessor), such occurrence will be marked as a protocol violation and presented in the final study report; nevertheless, the submitted data are used in the analysis.

8. Handling of Missing Data at the Final Analysis

Based on previous experiences with acute stroke trials, it is anticipated that there will be minimal loss to follow up for the 90-day assessment of the primary outcome. In the IMS I Trial, only 1 of 80 (1.25%) subjects were LTFU. In the IMS II Pilot Study, 2 of 73 (2.7%) were LTFU. In the IMS III Study, 27 of 656 (4.1%) subjects were LTFU. In SWIFT-PRIME, 4 of 191 (2.1%) were LTFU. In Fast-Mag, out of 1700 patients, 4 patients (0.2%) did not have at least day 30 follow-up.

~~In DEFUSE 3, at each analysis stage, the definitive sample for the boundary crossing analysis will consist of the first N consecutively recruited subjects, where N is the design-specified sample size at each interim stage or the total sample at the final stage, who are in the selected subgroup if one has been chosen.~~ All effort is put forth to ensure near complete follow-up, in particular with the assessment of the primary outcome (mRS at 90 days), death (mRS=6), and stroke recurrence. If the primary outcome (mRS at 90 days) cannot be assessed in the clinic, it will instead be obtained by phone using a structured interview. The final analysis will be conducted when all subjects in the analysis sample have reached the upper limit of the window for the 90-day outcome (120 days after randomization), and when the coordinating center believes it has exhausted all reasonable efforts to obtain outcome data collected within the window but delayed in data entry.

Missing 90-day mRS data (no mRS available within a 60-120 day window) will be handled by a hybrid approach: if the 30-day mRS is available, it will be “carried forward”, if not, the 90-day mRS will be “multiply imputed” (multiple imputation) using 'age' and 'NIHSS score at hospital discharge' as predictor variables.

All final analyses described in the SAP will use the multiple imputation data if applicable. ~~Specifically, the standardized, multiple imputation adjusted Wilcoxon rank sum statistic will be calculated.The Kullback-Leibler score, used in the~~

~~adaptive design for subgroup selection, will also be estimated using multiple imputation data. (see section 9.2)~~

The multiple imputation model will be based on all data accumulated in the study, and all missing data will be imputed based on the final model. ~~If, at any stage, the adaptive design specifies testing of the null in a subgroup, the standardized, multiple imputation adjusted Wilcoxon-Mann-Whitney test statistic will be calculated in that subgroup without refitting the multiple imputation model or changing the imputations.~~

9. Primary efficacy analysis

~~Regardless of whether or not a subgroup is selected~~. All efficacy outcome measures are analyzed under the ITT principle. Under this principle, each subject is analyzed according to the treatment group to which they were randomly assigned. ~~Definition of the sample included in the primary efficacy analysis (the adaptive design sample) is listed below in Section 9.3.~~

9.1. Primary outcome

The primary efficacy outcome measure is the mRS score at 90 days from randomization. Missing outcome is imputed according to Section 8.

9.2. ~~Impact of adaptive design on the sample for the primary analysis~~

~~An adaptive trial design, developed for DEFUSE 3, will allow the study to test the primary efficacy hypothesis in a subpopulation (the adaptive design sample) if an interim or final analysis indicates futility in the overall population.[2] The design is based on closed testing theory and the group sequential methods for the Generalized Likelihood Ratio (GLR) statistic developed by Lai and Shih.[3] The adaptive design was chosen because there is strong preliminary data suggesting that the effect of endovascular treatment is modified by two baseline variables: ischemic core lesion size and time-to-treatment. The way the adaptive design takes advantage of these biological assumptions (when they are true) is by reallocation of future accrual to the subgroup with the best prospects for showing efficacy. Specifically, if a subgroup is chosen at an interim analysis, subsequent enrollment is limited to patients in that subgroup. As a result, this subgroup will become larger than it would have been in the absence of the adaptive design.~~

The criterion for deciding which subgroup has the best chance of showing a benefit from endovascular therapy combines both the estimated size of the effect in the subgroup and the sample size of the subgroup. The GLR statistic (Kullback-Leibler criterion) is used to identify this subgroup because it optimally balances those two criteria. It selects the subgroup that has the best chance of showing an effect because it has an apparently large effect and is also of substantial size (note there are 5 subgroups of increasingly larger size, **Figure 1**). The adaptive design employs two biologically-based assumptions to limit the inflation of sample size: a monotonicity / contiguity assumption and an assumption that the effect is largest in the patients with the smallest volumes and the shortest time to randomization (cell C_{11} in **Figure 1**). The cut-points of the categories (cells) will be determined just prior to the first interim analysis (n=200), blinded to subjects' treatment allocations and outcomes, based on the distribution of subjects across the two dimensions (lesion volume and time-to-treatment) to yield six categories (cells) of approximately equal number of subjects.

		Time (hrs)	
		<10	10-16
Core lesion volume (ml)	≤20	C_{11}	C_{21}
	21-50	C_{12}	C_{22}
	51-70	C_{13}	C_{23}

Figure 1. The cohort is stratified according to core lesion volume and time to randomization. Cut-points in the figure serve as examples. Exact cut-points of the stratification will be determined, blinded to treatment allocation and outcome, based on the distribution of subjects across the core and time variables, just prior to the first interim analysis. Based on the results of the 1st interim analysis, enrollment will continue in all 6 cells or the study entry criteria will adapt and enrollment will be limited to one of 5 sub-groups (C_{11}, C_{11+21}, $C_{11+21+12}$, $C_{11+21+12+22}$, or $C_{11+21+12+22+13}$).

9.3. Interim and final statistical analyses

The primary endpoint is the distribution of scores on the modified Rankin Scale (mRS) at day 90. We will test the primary efficacy and futility hypotheses at the interim and final analysis using the generalized likelihood ratio (GLR) test, based on whether the usual normal approximation to the Wilcoxon-Mann-Whitney test statistic crosses a futility or efficacy boundary at interim or final analysis. The primary analysis will be conducted in the adaptive design sample (see section 5.2 for definition), according to the intention to treat principle, adjusted for the adaptive design, and unadjusted for covariates. See Appendix 2 for detailed specification of the interim and final calculations.

The efficacy bounds at interim and final analysis are set to control the overall (one-sided) Type I error rate at 2.5%. There are three group sequential boundaries: an interim futility boundary b_f, an interim efficacy boundary b_e, and a final efficacy bound c, which are fixed before the first interim analysis. At each of the two interim analyses, the futility bound b_f is used to decide if the study should continue recruitment in the overall group, shift accrual and testing to a subgroup, or stop in its entirety. The futility boundary takes into account the fact that the maximum analyzed sample size is a random variable that is no larger than the fixed maximum number of subjects randomized (n=476). Because subgroup selection reduces the maximum number of subjects available for analysis at completion of the study, this method effectively allows an easier futility stop after subgroup selection. This setup replaces conditional power analyses with an automatic and more efficient adjustment of boundaries.

If the patients are equally distributed across the six cells, the efficacy boundary at the interim analyses will be 2.62, the efficacy boundary at the final analysis will be 2.61, and the futility boundary at the interim analyses will be -1.88. If the patients are not distributed equally, these boundaries will be adapted slightly.[2]

First interim analysis (based on primary outcome data obtained from the first 200 consecutively randomized subjects)

The null hypothesis is tested in the entire subject population, and, depending on the results:

1. If neither efficacy nor futility bound is crossed, the trial continues with enrollment in the overall population to the 2nd interim analysis.
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the overall population.
3. If the futility bound is crossed, the optimal subgroup is selected based on the Kullback-Leibler criterion and the null is tested in that subgroup.
    - 3.1. If neither bound is crossed, the trial will continue with enrollment limited to the selected subgroup.
    - 3.2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
    - 3.3. If the futility bound is crossed, the trial stops for futility.

Second interim analysis (after primary outcome data obtained from an additional 140 consecutively randomized subjects within the target population defined by the first interim analysis)

If, after the first interim analysis, the study proceeds with enrollment in the overall population (option 1, under first interim analysis), the testing at the 2nd interim analysis is identical to the first interim and the decisions to stop or proceed with enrollment are identical to those outlined above under the first interim analysis.

If, after the first interim analysis, enrollment is limited to a selected subgroup (option 3, under first interim analysis), the second interim analysis is based on a test of the null hypothesis in the selected subgroup only and, depending on the results:

1. If neither bound is crossed, the trial continues to the final analysis with enrollment of additional subjects limited to the selected subgroup
2. If the efficacy bound is crossed, the trial stops and efficacy is declared in the selected subgroup
3. If the futility bound is crossed, the trial stops for futility. Note that there is no option for "second subgroup selection".

Final analysis (after primary outcome data obtained from an additional 136 consecutively randomized subjects within the target population defined by the second interim analysis)

If, after the second interim analysis, the study proceeds with enrollment in the overall population the null is tested in the overall population, and, depending on the results:

1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in the overall population.
2. If the efficacy bound is not crossed, the optimal subgroup is selected and the null is tested in that group:
    - 2.1. If the efficacy bound is crossed, endovascular therapy is declared efficacious in that subgroup
    - 2.2. If the efficacy bound is not crossed, endovascular therapy will be declared of no benefit.

~~If enrollment after one of the interim analyses is limited to a selected subgroup, then at the final analysis the null will be tested in that subgroup only and efficacy or lack thereof will be declared as per options 2.1 and 2.2 above.~~

~~The computation of the test statistics and the Kullback-Leibler selection criterion are specified in Appendix 2. Calculation of the test statistics will be carried out in SAS by the NDMC. A more detailed description of the adaptive design is provided in Appendix 3.~~

9.4. Interim and Final statistical analysis

The overall null hypothesis for the primary endpoint will be tested using the Wilcoxon rank sum statistic and asymptotic pvalue generated via the NPAR1WAY procedure in SAS. Tied values will be assigned the average rank for the corresponding value, as is the default. The continuity correction will not be applied.

Based on external information generated from the DAWN trial, an early interim analysis was requested by NINDS in order to make a determination as to whether DEFUSE should continue in the original population or a subpopulation or be terminated early. For the early interim analysis, the efficacy hypothesis will be tested at the one-sided Type 1 error probability of 0.023. The 90-day mRS for subjects who have not yet reached the end of the study protocol will be imputed with their respective 30-day mRS. Subjects with no 30-day or 90-day mRS scores will be excluded. The DSMB will be provided with the results of this analysis and would likely recommend stopping the study if the result is significant at the one-sided 0.023 level. Subsequently, upon submission of the last-randomized subject's 90-day outcome data into the WebDCU™ (anticipated in mid-August 2017) and the database lock, the analysis will be repeated at the one-sided alpha level of 0.023. The manuscript for the DEFUSE 3 Trial results will ~~reflect~~ contain results from this second analysis executed on the final locked database.

If the interim analysis is not significant at the one-sided alpha level of 0.023, it is anticipated that the DSMB will review the results and may request any additional subgroup analyses in order to make recommendations to NINDS. Such requests will be documented in an addendum to the SAP. Following the DSMB review and recommendation, the NINDS may make the decision to continue the study with recruitment of the DEFUSE 3 eligible patients per the current eligibility criteria, and the study may continue per the current protocol. In this case, the final primary outcome analysis will be tested at the one-sided alpha=0.001 (which will be referred to as "alpha2"). If the decision is made to terminate the study, upon submission of the last-randomized subject's 90-day outcome data into the WebDCU™ (anticipated in mid-August 2017) and the database locked, the analysis will be repeated at the one-sided alpha level of 0.023.

The NINDS may also direct the study team to continue the study in a subset of the current DEFUSE 3 eligible patient population (such as DAWN-ineligible patients). If this is deemed feasible, then the final analysis will include all patients enrolled (both before and after the subset selection was imposed) and will be evaluated with a one-sided alpha of 0.001 (which will be referred to as alpha3).

The rationale for recommending the extreme split of the alpha is that:
(a) given the conditions currently imposed by the CIRB (i.e., recruit only DAWN-ineligible patients), continuing the DEFUSE 3 study with the current eligibility criteria is unlikely to be

feasible; and (b) the sites have reported significant concerns regarding recruiting the DAWN-ineligible patients due to its complex definition.

To date, the DEFUSE 3 Executive Committee has proposed and DSMB had effectively recommended alpha2 + alpha3 =0, since they do not envision any circumstance where continuation of DEFUSE 3 in its current form is feasible. If a continuation is requested by the NINDS, a sample size will need to be reassessed based on the desired effect in the modified study population and the one-sided alpha2/alpha3 required for efficacy. This evaluation will likely result in increased total sample size for the study.

9.5. Reporting of primary results

The results of the study will be primarily expressed as whether or not an efficacy boundary was crossed at ~~either one of the two interim analyses or at~~ the final analysis. ~~The efficacy boundary is set, a priori, to guarantee less than a 2.5% one-sided error rate.~~ Crossing of the efficacy boundary will be considered evidence that endovascular therapy is beneficial, based on lower day-90 mRS scores in the endovascular group compared to controls.

9.6. Estimation of p-values, effect size estimates, and CIs after trial completion

The treatment effect will be ~~adjusted for study design and~~ expressed as

- The common odds ratio with its 95% confidence interval and p-value, calculated using a proportional odds model.
- The average number needed to treat for benefit (NNT), with its 95% confidence interval, where NNT = 1 / (P_EndovascularSuperior – P_MedicalSuperior).[4]

10. Sample size determination for primary efficacy analysis

The sample size determination begins with a preliminary estimate of the effect size that is plausible to expect and which is also clinically meaningful. Fixing standard operating characteristics at 5% two-sided Type 1 error and 10% Type 2 error (90% power) leads to a sample size lower bound for a hypothetical fixed-sample trial with no adaptation and no interim analysis. We then adjust the sample size for the group-sequential modification and the subgroup adaptive design.

The projected overall effect of endovascular therapy is based on 1) the observed 90-day modified Rankin Scale outcomes in DEFUSE 2 of target mismatch patients treated >6hrs after symptom onset and 2) the assumption that early reperfusion will be achieved in 75% of the endovascular arm vs. 20% of the medical therapy arm.[5-7] Using these data, we have projected the distributions on the mRS at 90 days for subjects in the endovascular and control arms of DEFUSE 3. (Table 1)

	mRS at day 90							
	0	1	2	3	4	5	6	Total
Endovascular group	18.0%	11.5%	19.6%	11.5%	16.4%	11.5%	11.5%	**100%**
Medical group	9.7%	7.9%	15.0%	17.7%	14.4%	17.7%	17.7%	**100%**

Table 1. **Projected 90-day mRS distributions in DEFUSE 3**

These distributions correspond to a standardized effect of 0.36 for the primary analysis. Based on these data, the fixed sample size for a non-adaptive design requires 376 patients (188 per arm) to have 90% power at an alpha of 5% (Wilcoxon-Mann-Whitney test); 100 patients were

added for the adaptive design to reach a maximum sample size of 476 for DEFUSE 3. The size of this increase is based on simulations and is selected to preserve the desired operating characteristics, while allowing shrinkage in effect size to 0.30, since the above estimate of 0.36 may be optimistic. The sample size of 476 is also the largest sample that can be accrued within budget and time limitations.

~~Simulations (n=5000) are used to compare the performance of a traditional fixed sample-size design (fixed n=476) to the adaptive design (max randomized n=476) under the null and various alternative scenarios (**Table 2**). For the simulations the effect size is expressed as a standardized effect *in the disjoint cells*, which are cumulated to form the subgroups as described above, where a standardized effect of 0.3 corresponds to a conservative projected effect of endovascular therapy (anticipated effect 0.36; see above).~~

~~Sim.~~	~~Standardized effect in cells*~~ C_{11}, C_{21}, C_{12}, C_{22}, C_{13}, C_{23}	~~Average standard. effect~~	~~Adaptive Design~~		~~Fixed Design~~	
			~~Average No. randomized~~	~~Power~~	~~Number randomized~~	~~Power~~
~~**#0**~~	~~0 0 0 0 0 0~~	~~0~~	~~361~~	~~2.2%~~	~~476~~	~~2.5%~~
~~**#1**~~	~~0.3 0.3 0.3 0.3 0.3 0.3~~	~~0.3~~	~~354~~	~~80%~~	~~476~~	~~89%~~
~~**#2**~~	~~0.5 0.4 0.3 0 0 0~~	~~0.2~~	~~400~~	~~86%~~	~~476~~	~~55%~~
~~**#3**~~	~~0.5 0.5 0 0 0 0~~	~~0.17~~	~~403~~	~~87%~~	~~476~~	~~41%~~

~~**Table 2. Simulations comparing the adaptive and the fixed trial designs. *Cells are defined in figure 1. Under the null (scenario #0), the adaptive design controls the total Type 1 error below 2.5%, stops early for futility 63% of the time, and the average number of randomizations is 361. If the effect is uniform across cells - and therefore also over cumulated subgroups - (scenario #2), the fixed-sample design is optimal, but the adaptive design results in only a small loss of power (from 89 to 80%). When the effect size distribution across the subgroups is in accord with the biological assumptions (scenarios #2 and 3), so that the effect in the cumulated subgroups declines as more cells with a null effect are added in, the adaptive design performs much better (higher power and smaller expected sample size) than the fixed sample, conventional trial. The adaptive design also performs well compared to a non-adaptive, fixed sample that includes efficient multiple comparisons-adjusted testing for effect in subgroups at the end of the study.[2] Note that in all of the tabulated examples (other than the first, which is the null case), the overall population null is false, because the effect is positive in some cell or cells, and the effect is null in the others, so that the effect never becomes negative. If we allow negative effects in any of the cells, then the advantages of the adaptive design over the conventional design grow even larger.**~~

11. Exploratory analyses of the primary outcome after trial completion

11.1. DAWN eligible vs DAWN ineligible

The primary efficacy and safety analyses (mRS shift at 90 days, mRS 0-2 at 90 days, death rate and SICH rate) will be performed in "DAWN eligible" vs. "DAWN-ineligible" subgroups. The DAWN-ineligible subgroup will be defined as any patient who has one or more of the following characteristics:

- Pre-stroke mRS >1
- Baseline NIHSS <10
- Baseline ischemic core > 51 ml (based on Central Lab reading)
- Age >80 with baseline ischemic core > 21ml (based on Central Lab reading)
- Age < 80 and NIHSS < 20 and Baseline ischemic core >31 ml (based on Central Lab reading)

All other patients will be considered "DAWN eligible".

11.2.Analysis in Subgroups Defined by Sex, Race, and Ethnicity

The primary efficacy analysis will be repeated in subgroups defined by sex, race (White, African-American, Asian, Other), and ethnicity (Hispanic, Non-Hispanic). The treatment effect and corresponding 95% confidence intervals will be estimated, as described in Section 12.2.

11.3.Analysis adjusting for covariates

Please note: if the final analysis is performed with 200 patients or less, we will still perform the analysis of covariates as described below, but acknowledge that the value of this analysis may be limited because of the small sample size.

An additional ~~The secondary~~ analysis of the primary endpoint will be the same rank-based analysis comparing the distribution of the 90-day Rankin scores between treatment groups while stratifying for prognostically important covariates. The Generalized Cochran-Mantel-Haenszel test (CMH) is the analogue to the WMW, designed to test against the alternative that there is a uniform shift of size "delta" in the Rankin score distribution from one group to the other after stratification. The CMH test will be stratified by age, baseline NIHSS Score, Baseline Ischemic Core Volume (based on Central Lab reading), and Time from symptom onset to randomization (using the same categories used for the stratified randomization). A rich body of clinical research has established the prognostic importance of these covariates. Cut-points will generally follow those used in the randomization, but may be altered depending on the numbers of patients enrolled into each category.

To assess the relationship between important covariates and the size of the treatment effect, a model based regression analysis will be performed. Specifically, we will create a multivariable ordinal regression model with the 90-day Rankin outcome as the dependent variable and use a literature based model-building process. The model building process will include preliminary variable selection, a final model selection, and a final model assessment. We will evaluate each candidate predictor for proportionality and linearity (model assumptions). The correlation between candidate continuous variables will be analyzed. Variables with a rho coefficient higher than 0.7 will not be jointly entered into the final model building process - only the variable with the lowest p-value if it is less than 0.1. The base model will include treatment group. During the preliminary variable selection, potential predictors will be added to the base model one at a time to obtain a p-value for each. Variables considered candidates for the final model will have a p-value ≤ 0.1. At the final model selection, variables will be added sequentially starting with the variable with the lowest p-value from the group of candidate predictors. The criteria for keeping a variable in the final model will be a p-value ≤ 0.05. Each time a predictor is kept in the final model all previously added variables will be re-assessed. Any previously entered variables whose p-value has increased above 0.05 will be dropped from the model. All qualifying variables from the preliminary selection phase will be considered along with their two-way interaction with treatment assignment ~~site of arterial occlusion~~. A statistical assessment of the model will be examined by the shrinkage statistic. A shrinkage statistic below 0.85 will indicate the model is overfitting the data, and the number of predictors should be reduced. The following variables will be considered for inclusion in the adjusted model:

- Age
- Baseline NIHSS score

- Baseline ischemic core volume
- Time from symptom onset to randomization
- Sex
- Admission SBP
- Baseline glucose

11.4.Analysis under the as-treated principle

Due to the nature of the study, some subjects in the endovascular treatment plus medical management group may not receive endovascular therapy. One instance when this may occur is if symptoms resolve spontaneously between randomization and start of the endovascular procedure. Another potential reason is the absence of an arterial occlusion on the baseline angiogram (due to spontaneous recanalization). Although it would be rare, we may also see some subjects who are randomized to the medical management only group, but are treated with endovascular therapy. Therefore, after completion of the study and if there is a crossover population, we will repeat the primary analysis (the Wilcoxon-Mann-Whitney measure of superiority) ~~within the adaptive design sample~~, but with patients categorized under the as-treated principle.

11.4.1. Group assignment under the "as-treated" principle

- **Endovascular therapy**: Patients who present to the endovascular suite (cath lab) and undergo a femoral puncture within 24 hours after time of onset of the qualifying stroke are assigned to the endovascular treatment arm under the "as-treated" principle.
- **Medical therapy**: Patients who do not meet the above criteria for endovascular treatment are assigned to the medical treatment arm under the "as-treated" principle.

12. Analyses of secondary efficacy outcomes

12.1.Analyses of secondary clinical efficacy outcome: functional independence

The effect of endovascular treatment will be assessed on a secondary clinical efficacy outcome: functional independence at 90 days, defined as an mRS score ≤2 at day 90. Results will be expressed as an unadjusted risk ratio with its 95% confidence interval and p-value.

12.2.Analyses of functional independence in subgroups

The unadjusted effect of endovascular treatment on the secondary efficacy outcome "functional independence" (mRS ≤2) will be analyzed in the following subgroups, assuming sufficient numbers of subjects are enrolled in each subgroup (minimum required is 10% of the total sample size in each subgroup). Results will be expressed as unadjusted risk ratios with their 95% confidence intervals and p-values. We will assess for differences in the treatment effect between subgroups (eg, age <70 vs ≥70) using a Breslow-Day test for binary covariates, or by including an interaction term into a logistic regression model for covariates with greater than 2 categories.

- Time from symptom onset to randomization (using the categories from the dynamic stratification: <9, 9-12, >12 hrs~~cutpoint from adaptive design)~~

- Baseline ischemic core lesion volume based on Central Lab reading (~~using cutpoints from adaptive design~~using the three categories from the dynamic stratification: <10.0, 10.0-25.0, > 25.0 mL)
- Age at randomization (<70 and ≥70 years old)
- Baseline NIHSSS using the categories from the dynamic randomization (<13, 13 - 18 and >18)
- Baseline ASPECTS based on Central Lab reading (<8 and ≥8)
- Primary occlusion site based on Central Lab reading (M1 and ICA)
- IV rt-PA treatment (yes and no)
- Patient selection (CTP versus MRI)
- ~~Collateral grade (0-1, 2, 3)~~
- Method by which symptom onset time was determined (last known well vs exact time of symptom onset)
- Sex
- Race (White, Non-white)
- Ethnicity (Hispanic and Non-Hispanic)
- Baseline atrial fibrillation (yes and no)

12.3.Analyses of imaging efficacy outcomes

We hypothesize that endovascular treatment improves radiological outcomes. We will compare four imaging outcomes between treatment groups:

- The proportions of successful reperfusion, where successful reperfusion is defined as a >90% reduction in the volume of brain tissue with critical hypoperfusion (Tmax>6sec after artifact removal by core lab) between baseline and 24 hours. Results will be expressed as an unadjusted risk ratio with its 95% confidence interval and p-value.
- The proportion of subjects with recanalization of the primary arterial occlusive lesion (ICA or M1) at the time of the 24-hour follow-up scan is compared between the two treatment arms. Results will be expressed as an unadjusted risk ratio with its 95% confidence interval and p-value.
- Infarct volumes at 24 hours, defined as the lesion volume as outlined on the 24-hour DWI (or CT if DWI not performed). The 24-hour endpoint is based on data demonstrating that assessment of infarct volume at 24 hours captures the effect of reperfusion therapies on infarct growth and predicts outcomes similarly to day 90 infarct volumes.[8, 9] The Wilcoxon rank-sum will be used to compare the treatment groups.
- Ischemic lesion growth between baseline and 24 hours defined as the difference between the baseline ischemic core lesion volume and the 24-hour DWI lesion volume (or CT if DWI not performed). Absolute lesion growth will be calculated by subtracting the baseline ischemic core lesion volume from the 24-hour DWI lesion volume. DEFUSE 2 demonstrated a substantial reduction in infarct growth among Target mismatch patients treated in the 6-12 hour time-window who achieved early reperfusion: median growth 0.5 ml (IQR: -2 – 10) with reperfusion (n=23) vs. 39 ml (IQR: 18-121) without reperfusion (n=13), p<0.001. The Wilcoxon rank-sum will be used to compare the treatment groups.

We also hypothesize that final infarct volumes (assessed on 24-hour DWI) can be predicted based on baseline ischemic core and critically hypoperfused tissue volumes. To analyze this relationship, subjects will be divided into two groups: 1) patients with reperfusion, defined as >90% reduction in the Tmax>6-second lesion volume between baseline and 24 hours (or TICI 2b–3 at end of procedure if perfusion imaging inadequate) and (2) a “no reperfusion” group, defined as <10% reperfusion at 24 hours. Spearman’s rho will be used

to assess correlations between: 1) baseline ischemic core volume and 24-hour DWI lesion volume in patients with reperfusion; and 2) the union of baseline ischemic core volume and 24-hour critically hypoperfused lesion volume (Tmax6) and 24-hour DWI lesion volume in patients without reperfusion.

12.4.Additional pre-planned analyses

- Assessment of the effect of endovascular therapy on the following endpoints will be assessed: (1) >8 point improvement on the NIHSS between baseline and 24 hrs or an NIHSS score of 0-1 at 24 hrs; (2) discharge destination (home vs other); (3) number of days spent at home during first 90 days after stroke; and (4) mRS score at day 30 (both with an ordinal analysis and dichotomous at 0-2).
- Assessment of the interaction between time-to-~~treatment~~randomization and the effect of endovascular therapy (expressed as the common odds ratio for shift on the 90-day mRS)
- Assessment of imaging predictors of growth of the ischemic lesion between baseline and 24 hours stratified by treatment allocation and reperfusion status. Specific baseline imaging predictors that will be evaluated include the ratio of the Tmax 10 / Tmax 6 lesion volumes and the CBV value within the Tmax 6 lesion volume.

13. Safety analyses

All safety outcome measures are analyzed under the as-treated principle. Under this principle, each subject is analyzed according to the treatment that the subject received. The external Medical Safety Monitor and the DSMB will monitor safety variables throughout the study at frequent intervals. Details of this process are specified below.

13.1.Monitoring of deaths and sICH

The primary safety endpoints are deaths and sICH rates; these will be monitored quarterly. Symptomatic ICH or death rates that exceed pre-specified thresholds will trigger a meeting of the DSMB. The DEFUSE 3 study will be placed on hold if it is determined with 95% probability that either:

1) the symptomatic ICH rate (NIHSS worsening of 4 or more points associated with ICH) within 36 hours of randomization exceeds 10% in the endovascular group; OR
2) the 90-day mortality rate exceeds 20% in the endovascular group.

The study will remain on hold until the investigators and the DSMB can conduct a review of events and make a determination on the continuation of the trial.

13.2.Summary of Adverse Events (AEs) and Serious Adverse Events (SAEs)

All AEs and SAEs are summarized by MedDRA preferred term (as coded based on the AE CRF) and by treatment group in terms of frequency of the event, number of subjects having the event, timing relative to randomization, and relatedness to the intervention.

For the following specific events, the proportions by treatment group, as well as the RR and corresponding 95% confidence interval, are provided:

- Stroke related mortality within 90 days of randomization

- sICH within 36 hours of randomization
- Significant neurological deterioration prior to discharge (defined as an increase of 4 points or more on the NIHSS)
- PH1 and PH 2 rates on the 24 hr scan

At the end of the study, the cumulative incidences of these events are compared between the two treatment groups using Pearson Chi-Square. Fisher's exact test will be used if any cells are 5 or less.

All brain scans obtained prior to discharge will also be assessed for PH by the core lab.

14. Coordination between Stanford and MUSC statistical teams

The primary statistician for the study is Dr. Phil Lavori at Stanford. He is a voting member of the DEFUSE 3 Executive Committee (EC). Dr. Lavori will be blinded to all outcome data during the study. ~~Dr. Lavori will become unblinded~~ Upon database lock of the clinical database, Dr Scott Hamilton will receive unblinded data and will conduct the final analyses. Drs. Lavori and Hamilton will be responsible for developing and writing the statistical analysis plan (SAP) prior to the initiation of the study, and SAP amendments, if any, during the study. The statistical team at MUSC, led by Dr. Yeatts, will be unblinded throughout the study. The MUSC team will ~~implement the adaptive design algorithm developed and written by Dr. Lavori~~; conduct and independently validate the interim analyses according to the SAP; generate Open and Closed Reports for the DSMB and interact with the DSMB in closed sessions; and collaborate with Dr. ~~Lavori~~ Hamilton on validation of final analyses. After database lock, the MUSC statistical team will create the public use datasets (PUDS) and submit them to the NINDS.

15. References

1. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*. 1975;31:103-115
2. Lai TL, Lavori PW, Liao OY. Adaptive choice of patient subgroup for comparing two treatments. *Contemp Clin Trials*. 2014;39:191-200
3. Lai TL, Shih M-C. Power, sample size and adaptation considerations in the design of group sequential clinical trials. *Biometrika*. 2004;91:507-528
4. Howard G, Waller JL, Voeks JH, Howard VJ, Jauch EC, Lees KR, et al. A simple, assumption-free, and clinically interpretable approach for analysis of modified rankin outcomes. *Stroke*. 2012;43:664-669
5. Nogueira RG, Lutsep HL, Gupta R, Jovin TG, Albers GW, Walker GA, et al. Trevo versus merci retrievers for thrombectomy revascularisation of large vessel occlusions in acute ischaemic stroke (trevo 2): A randomised trial. *Lancet*. 2012;380:1231-1240
6. Saver JL, Jahan R, Levy EI, Jovin TG, Baxter B, Nogueira RG, et al. Solitaire flow restoration device versus the merci retriever in patients with acute ischaemic stroke (swift): A randomised, parallel-group, non-inferiority trial. *Lancet*. 2012;380:1241-1249
7. Pereira VM, Gralla J, Davalos A, Bonafe A, Castano C, Chapot R, et al. Prospective, multicenter, single-arm study of mechanical thrombectomy using solitaire flow restoration in acute ischemic stroke. *Stroke*. 2013;44:2802-2807
8. Campbell B, Purushotham A, Christensen S, Desmond P, Nagakane Y, Parsons M, et al. The infarct core is well represented by the acute diffusion lesion: Sustained reversal is infrequent. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*. 2012;32:50-56

9. Campbell BC, Tu HT, Christensen S, Desmond PM, Levi CR, Bladin CF, et al. Assessing response to stroke thrombolysis: Validation of 24-hour multimodal magnetic resonance imaging. *Arch Neurol*. 2012;69:46-50

Memorandum

Date: July 19, 2017

To: The DEFUSE 3 SAP

From: Sharon Yeatts
Jyoti Arora

RE: Interim Analysis Plan, as requested by the DEFUSE 3 DSMB

An early interim analysis was requested by NINDS, in response to external information generated from the DAWN trial, in order to make a determination as to whether DEFUSE 3 should continue in the original population or a subpopulation, or be terminated early.

The details of this proposed analysis were put forth in Version 2.0 of the DEFUSE 3 SAP. On June 23, via email communication from Peter Gilbert, a revision to that analysis plan was requested. The details of the plan, as requested by the DSMB, are outlined below; we have noted details which deviate from the plan outlined in the SAP. This information is documented in a memo to the SAP, rather than a revision to the SAP, in order to keep the study team blinded.

For this early interim analysis, the overall null hypothesis for the primary endpoint will be tested using the Wilcoxon rank sum statistic and asymptotic pvalue, generated via the PROC NPAR1WAY procedure in SAS. The 90-day mRS for subjects who have not yet reached the end of the study protocol will be imputed with their respective 30-day mRS. Subjects with no 30-day or 90-day mRS scores will be excluded. This is as specified in Version 2.0 of the DEFUSE 3 SAP. The SAP stipulates a one-sided level of significance of 0.023 for this analysis; however, the DSMB will meet to decide on decision rules prior to reviewing the results, and their decision rules are presumed to override the level of significance stated in the SAP.

The DSMB requested that the above analysis be repeated in the subgroup of DEFUSE 3 subjects who are considered DAWN-eligible. The DSMB will decide whether to review the subgroup analysis at the time of the meeting.

Per DSMB request, the power of these analyses at the available sample size will be calculated, as will the nominal threshold for declaring efficacy or harm if this was a typical unplanned interim analysis.

The DSMB requested an additional evaluation of harm associated with the endovascular therapy. This analysis plan is based on telephone consultation with Michael Parides (on July 7) and email communication from Peter Gilbert (on July 13). Briefly, the probability of a 'harmful' event under each treatment group ($\pi_{endovascular}$ and $\pi_{medical\ management}$) will be considered as random variables with prior probability distribution Beta (1,1), representing a uniform distribution. The posterior distributions are also represented by Beta, where the parameters are defined according to the number of subjects (and the number of harmful events) in each treatment arm. The distribution of the difference between $\pi_{endovascular}$ and $\pi_{medical\ management}$ will then be estimated via simulation. This process will be repeated for two definitions of harmful event: mRS 4-6 and mRS 3-6. The analysis will be conducted in the full cohort as well as in the DAWN-eligible subgroup.

The DSMB further requested that the absolute treatment effect, according to dichotomy defined by mRS 0-2 vs 3-6, and its corresponding 95% confidence intervals, also be presented. Again, the intervals will be constructed for the full cohort as well as the DAWN-eligible subgroup.

The semi-annual report for the full DEFUSE 3 cohort was emailed to Peter Gilbert on July 19. The semi-annual report based on the DAWN-eligible subgroup, as well as the requested analyses (in the overall cohort and in the DAWN-eligible subgroup) , will be brought to the meeting by the unblinded study statisticians. The individual analysis documents will be held in separate, signed and sealed envelopes; after agreeing on the decision rules for stopping the study, the unblinded statisticians will provide the DSMB with the relevant results for the full cohort. The DSMB will determine which of the ancillary analyses it wishes to see at that time.

Memorandum

Date: July 24, 2017

To: The DEFUSE 3 SAP

From: Sharon Yeatts
Jyoti Arora

RE: the DSMB review of interim analysis results

An early interim analysis was requested by NINDS, in response to external information generated from the DAWN trial, in order to make a determination as to whether DEFUSE 3 should continue in the original population or a subpopulation, or be terminated early.

The details of this proposed analysis were put forth in Version 2.0 of the DEFUSE 3 SAP. On June 23, via email communication from Peter Gilbert, a revision to that analysis plan was requested. The details of the plan, as requested by the DSMB, are outlined below; we have noted details which deviate from the plan outlined in the SAP. This information is documented in a memo to the SAP, rather than a revision to the SAP, in order to keep the study team blinded.

For this early interim analysis, the overall null hypothesis for the primary endpoint was tested using the Wilcoxon rank sum statistic and asymptotic pvalue, generated via the PROC NPAR1WAY procedure in SAS. The 90-day mRS for subjects who have not yet reached the end of the study protocol was imputed with their respective 30-day mRS. Subjects with no 30-day or 90-day mRS scores were excluded. This is as specified in Version 2.0 of the DEFUSE 3 SAP. The SAP stipulated a one-sided level of significance of 0.023 for this analysis; however, the DSMB met to decide on decision rules prior to reviewing the results, and their decision rules override the level of significance stated in the SAP.

Per DSMB request, the nominal threshold for declaring efficacy or harm if this was a typical unplanned interim analysis was provided. Two boundaries were calculated under a general design, according to a maximum sample size of 476 subjects (as the original maximum for the adaptive design) and according to a maximum sample size of 376 subjects (the sample size required for the non-adaptive design). The DSMB decided to proceed under the latter (boundary provided in the table below), and that the analysis should consider a two-sided alternative. Dr. Yeatts then emailed the board members the password for the corresponding document.

Look	Information Fraction	2-sided Efficacy Boundary
		Z-scale
Interim	0.48	±3.018
Final	1.0	±1.967

The DSMB then requested to see the two-sided ordinal analysis results for the DAWN-eligible subgroup, and Dr. Yeatts emailed the board members the password for the corresponding document.

The DSMB then reviewed the closed semi-annual report for the full cohort.

Summary of changes to the Statistical Analysis Plan for the DEFUSE 3 study

The statistical analysis plan was modified to accommodate significant external events that occurred in May 2017.

May 16, 2017: The results of the DAWN trial, that enrolled similar patients and studied a similar intervention as DEFUSE 3, were presented at the European Stroke Organization Conference. The results of this study demonstrated a substantial clinical benefit of endovascular thrombectomy over medical therapy.

May 24, 2017: The DEFUSE 3 CIRB requested that enrollment in DEFUSE 3 be held for patients who met the eligibility criteria for the DAWN study. At that time, 182 patients had been randomized in DEFUSE 3.

May 26, 2017: After consultation with the DSMB, the DEFUSE 3 sponsor (NINDS) placed enrollment on hold for the entire study, and requested a revision of the DEFUSE 3 Statistical Analysis Plan to allow an early interim analysis which would include a pre-specified subgroup analysis of the primary and secondary efficacy endpoints in both "DAWN eligible" and "DAWN-ineligible" subgroups.

The main changes to the SAP included:

1. Text and analyses that related to the - originally planned – group sequential adaptive enrichment design were removed
2. New decision rules for the early interim (and final) analysis were added (section 9.4)
3. Additional subgroup analysis in the 'DAWN eligble' vs 'DAWN ineligible' subgroups was added (section 11.1)

July 19, 2017: A memo to the SAP is filed that specifies that ‘the DSMB will meet to decide on decision rules prior to reviewing the results, and their decision rules are presumed to override the level of significance stated in the SAP.’

July 24, 2017: A memo to the SAP is filed specifying that the DSMB elected to not adapt the stopping rules as specified in the SAP, but instead used as their interim stopping rule a 2-sided efficacy boundary (Z-score of 3.018) according to a maximum sample size of 376 and the O’Brien-Fleming spending function.