Results of the MRI Substudy of the Intravenous Magnesium Efficacy in Stroke Trial

Chelsea S. Kidwell, MD; Kennedy R. Lees, MD; Keith W. Muir, MD; Christopher Chen, MD; Stephen M. Davis, MD; Deidre A. De Silva, MD; Christopher J. Weir, PhD; Sidney Starkman, MD; Jeffry R. Alger, PhD; Jeffrey L. Saver, MD; for the MR IMAGES Investigators

Background and Purpose—Although magnesium is neuroprotective in animal stroke models, no clinical benefit was confirmed in the Intravenous Magnesium Efficacy in Stroke (IMAGES) trial of acute stroke patients. The Magnetic Resonance in IMAGES (MR IMAGES) substudy investigated the effects of magnesium on the imaging surrogate outcome of infarct growth.

Methods—IMAGES trial patients in participating centers were randomized to receive either intravenous magnesium or placebo within 12 hours of stroke onset. Infarct growth was defined as volume difference between baseline diffusion-weighted imaging and day 90 fluid-attenuated inversion recovery image lesions. Patients who died were imputed the largest infarct growth observed.

Results—Among the 90 patients included in the primary analysis, there was no difference in infarct growth (median absolute growth, \( P=0.639 \); median percentage growth, \( P=0.616 \); proportion with any growth, \( P=0.212 \)) between the 46 treated with magnesium and 44 with placebo. Infarct growth correlated with NIHSS score change from baseline to day 90. There was a trend showing baseline serum glucose correlated with infarct growth with magnesium treatment, but not in the placebo group. The mismatch frequency was reduced from 73% to 47% by increasing the mismatch threshold from >20% to >100% of core volume.

Conclusions—Infarct growth, confirmed here as a surrogate for clinical progression, was similar between magnesium and placebo treatment, paralleling the main IMAGES trial clinical outcomes. Glucose was a covariate for infarct growth with magnesium treatment. A more stringent mismatch threshold to define penumbra more appropriately would have excluded half of the patients in this 12-hour time window stroke study. (Stroke. 2009;40:1704-1709.)

Key Words: diffusion-weighted imaging ■ glucose ■ magnesium ■ magnetic resonance imaging ■ surrogate endpoint

Neuroprotective therapies for acute ischemic stroke are postulated to work by salvaging the ischemic penumbra that surrounds the infarct core. More than 100 randomized, clinical trials of neuroprotective therapies have been performed in humans in the acute stroke setting. Despite promising results in the laboratory setting, no neuroprotective agent to date has been definitively shown to be of clinical benefit in humans with acute ischemic stroke. The desired outcome and hence gold-standard measurement of efficacy in human clinical stroke trials is neurological and functional improvement. One barrier to demonstrating the efficacy of neuroprotective agents in human phase III clinical trials is the need to evaluate large numbers of patients to demonstrate their modest clinical effect. There is a pressing need for biologically based surrogate endpoints that can discern treatment effects using smaller sample sizes. Such intermediate surrogate endpoints could be used in phase II clinical trials to rapidly identify and screen promising agents in a cost-effective manner. These surrogate measures could also be helpful in phase III clinical trials by providing physiological evidence and offering biological plausibility for efficacy of neuroprotective agents.

MRI offers promise as a surrogate outcome measure in acute stroke trials. MR diffusion-weighted imaging (DWI) and perfusion weighted imaging (PWI) visualize brain regions of bioenergetic and blood flow compromise almost immediately after ischemia onset. DWI combined with PWI can provide an approximation of the extent of the ischemic...
penumbra. Therapeutic salvage of the penumbral region has been demonstrated on MRI studies in humans with acute stroke. Serial imaging studies allow quantification of infarct growth over time and, therefore, an assessment of whether a therapeutic agent prevents progression of infarction in the penumbral region. The approach of measuring infarct growth closely parallels the pathological ascertainment of infarct volume as a standard gauge of therapeutic efficacy in animal stroke models. Moreover, in humans, infarct growth after ischemic stroke has been shown to closely correlate with poor clinical outcome. However, attenuation of ischemic DWI lesion growth with neuroprotective agents has to date only been definitively demonstrated in animal models.

The IntraVenous Magnesium Efficacy in Stroke (IMAGES) trial was designed to test whether intravenous magnesium sulfate, administered within 12 hours of stroke onset, reduces death or disability at 90 days. Preclinical trials, as well as pilot trials, in humans suggested that magnesium has a neuroprotective effect in ischemic stroke. However, in the main IMAGES trial, clinical outcome at day 90 with magnesium did not differ significantly from placebo. The MRI substudy was undertaken after the main trial had begun to provide an auxiliary biological measure of drug effect. The primary hypothesis tested in this substudy was that magnesium sulfate attenuates infarct growth, as measured from baseline DWI volume to day 90 infarct volume, among enrolled ischemic stroke patients.

**Patients and Methods**

**Patient Selection**

MR IMAGES was a multicenter, double-blind, randomized, placebo-controlled clinical trial designed as a substudy within the larger IMAGES trial. Between January 2000 and October 2003, 8 centers in 5 countries participated in the MRI substudy. Randomization in the MRI substudy was allowed to continue beyond the end of recruitment to the main trial, up to the time of unblinding of IMAGES data. At that point, the Steering Committee and Data Safety and Monitoring Board opted to close enrollment in the MRI substudy, even though the target sample size had not been reached. Inclusion and exclusion criteria for the MRI substudy were identical to those of the main IMAGES trial, except for additional imaging-related criteria. A minimum DWI lesion volume was prespecified for enrollment, initially ≥5 mL, which was altered to ≥3 mL during recruitment. Patients were excluded for the following reasons: pacemaker or other metal implant precluding MRI study; history of second- or third-degree heart block; recent stroke within previous 30 days or rapidly resolving deficit; and known allergic reaction to MRI contrast (gadolinium).

Enrollment into the MRI substudy occurred via the interactive voice–response system used for the main IMAGES trial. Study drug administration also proceeded as described in the main IMAGES trial manuscript. Serum glucose was measured at baseline before drug administration. Clinical outcome measures collected included NIHSS at baseline, 48 hours, and day 90, as well as the Barthel index and modified Rankin score at day 90. Patient who died were allocated a day 90 NIHSS score of 42. Poor clinical outcome was defined by global score of Barthel index <95 or modified Rankin score >1; Barthel index <95; Barthel index <60; modified Rankin score >1; and death.

The substudy was overseen by the IMAGES trial Data Safety and Monitoring Board until enrollment was closed in the main trial. After that time, a substudy-specific medical monitor was appointed (K.M.) and an independent Data Safety and Monitoring Board was established. The substudy was approved by the institutional review board at each site and informed consent was obtained for all patients enrolled.

**Imaging Methods and Outcome Measures**

All substudy sites were required to have MRI scanners capable of performing DWI. MRI was performed at screening and day 90. The screening MRI included the following sequences: a T1-weighted sagittal scout, DWI, and, when available, PWI (T2*-weighted imaging for bolus tracking perfusion images). The day 90 study included scout, DWI, and fluid-attenuated inversion recovery sequences. Recommended sequence parameters were provided to sites to standardize imaging methodology. Images were sent to the UCLA core imaging laboratory for processing and analysis. Imaging analyses were performed by a single reader (C.S.K.) blinded to treatment assignment and clinical information.

The following regions of interest were outlined using a semiautomated segmentation technique: baseline DWI lesion volume, baseline perfusion lesion volume defined as the region within the involved vascular territory with a Tmax (time-to-peak of the residue function) value >2 seconds, and day 90 infarct volume from fluid-attenuated inversion recovery sequence (if fluid-attenuated inversion recovery was not obtained, a T2-weighted sequence was used). For each patient, infarct growth was measured by absolute lesion growth (day 90 lesion−baseline DWI lesion volume); percentage lesion growth (day 90 lesion−baseline DWI lesion volume)/baseline DWI lesion volume×100%; and growth dichotomized as >0% and ≤0% percentage lesion growth. A conservative adjustment measure was used to correct for possible bias effects from patients who left the study because of death—these patients were imputed to have the largest infarct growth among subjects with day 90 imaging. Seventeen patients with initial DWI lesions less than the entry criteria (<3 mL) were enrolled, and they were retained in the primary intention-to-treat analysis. Separate analyses were performed for subgroups of patients with clinically defined lacunar stroke using the Oxfordshire Community Stroke Project classification system; imaging-defined lacunar stroke (small, deep infarct <1.5 cm in diameter without evidence of large-vessel territory involvement on DWI or PWI sequences); baseline DWI lesion volume ≥3 mL; diffusion–perfusion mismatch (|PWI-DWI|/DWI lesion volume) >20%; both baseline DWI lesion volume ≥3 mL as well as >20% diffusion–perfusion mismatch; and treatment within 6 hours of symptom onset. After a publication based on the Diffusion and Perfusion Imaging Evaluation for Understanding Stroke Evolution (DEFUSE) study cohort that suggested that the optimal mismatch threshold should be raised >20%, we also studied a subgroup with diffusion–perfusion mismatch >100%. This cut-off was chosen as a compromise between the cut-off of 160% (which optimally predicted favorable clinical outcome with early perfusion but severely restricted the number of patients) and 80% (which was suggested as less restrictive).

**Statistical Analysis**

The prespecified primary study end point was infarct growth from baseline DWI to day 90 final lesion volume in the intent-to-treat population. Group differences were analyzed using the Mann-Whitney U test for continuous variables and the χ² test or Fisher exact test for binary variables. Statistical significance of possible interactions was studied using logistic regression for categorical outcomes and 2-way ANOVA testing for continuous outcomes. The target sample size of 150 patients was calculated based on the following assumptions: (1) 85% survival at day 90 in the placebo group and (2) magnesium effect estimated as 25% reduction of infarct growth, conservatively derived from animal studies that show magnesium reduces infarct volume by 20% to 65%. The sample size necessary to detect a 25% difference in infarct growth using a 2-tail test with alpha=0.05 and power=80% was calculated to be 72 in each treatment arm.

**Results**

A total of 104 patients were enrolled in the MR IMAGES study (Figure). Pretreatment imaging data could not be...
analyzed in 6 patients. Of the 98 patients with baseline imaging data available, 74 underwent a day 90 MRI. Among the patients alive at day 90 but without outcome MRI (8 placebo, 8 magnesium), the final Barthel and modified Rankin score scores did not differ from those with final outcome scanning. As prespecified, maximum lesion growth (183% for percentage growth; 214 mL for absolute growth) was assigned to patients who died before their day 90 scan. Therefore, there were 90 patients in the primary outcome analysis (44 placebo, 46 magnesium). Among patients who were alive at day 90, Barthel index was ascertainment for 70 patients (34 placebo, 36 magnesium) and modified Rankin score in 86 (42 placebo, 44 magnesium).

Mean time to treatment across the 90 patients was 7 hours 52 minutes (SD, 2 hours 30 minutes; range, 2 hours 15 minutes to 11 hours 50 minutes). Only 2 patients were enrolled <3 hours and 26 patients <6 hours from onset. Median NIHSS score was 12 (range, 2–28). Median baseline DWI lesion volume was 21 mL (range, 0–303 mL). Seventy-one patients had a baseline DWI lesion >3 mL. Of the 60 patients with baseline perfusion imaging available for analysis, median baseline PWI lesion volume was 79 mL (range, 0–438 mL) and median baseline DWI volume was 24 mL (range, 0–290 mL). The proportion of patients with perfusion–diffusion mismatch decreased from 73% (44/60), using a threshold of >20%, to 47% (28/60), using a threshold of >100%. Median day 90 infarct volume was 14 mL (range, 0–531 mL). Fifty-one patients (57% of 90) had percentage infarct growth >0% from the baseline to the day 90 scan.

Among the 90 patients with outcome data, baseline DWI lesion volume correlated with baseline NIHSS (Spearman correlation coefficient = 0.669; P < 0.001) and final infarct size (Spearman correlation coefficient = 0.948; P < 0.001). Absolute lesion growth (Spearman correlation coefficient = 0.654; P < 0.001), as well as percentage lesion growth (Spearman correlation coefficient = 0.581; P < 0.001) both correlated with change of the NIHSS score from baseline to day 90. Among the 60 patients with baseline PWI performed, baseline PWI lesion volume correlated with baseline NIHSS (Spearman correlation coefficient = 0.499; P < 0.001) and final infarct volume (Spearman correlation coefficient = 0.805; P < 0.001). There were also correlations between mismatch volume with final infarct volume (Spearman correlation coefficient = 0.382; P = 0.003) and infarct growth volume (Spearman correlation coefficient = 0.234; P = 0.027).

For the 90 patients contributing to the primary analysis, there were no significant differences between magnesium and placebo groups in baseline characteristics (Table 1). There were 18 patient with clinically defined lacunar stroke (9 placebo, 9 magnesium) and 15 patients with imaging-defined lacunar stroke (9 placebo, 6 magnesium).

Among the 90 intention-to-treat patients, there were no significant differences between magnesium and placebo treatment in the primary outcome measure of infarct growth using various measurements (Table 2). There were also no differences in clinical outcomes between treatment groups. For the groups of patients who fulfilled the prespecified criterion of DWI lesion ≥3 mL, there was no difference between magnesium and placebo treatment in terms of absolute growth, percentage growth, and proportion with growth >0% (Table 3).

Exploratory analyses failed to demonstrate any significant difference for infarct growth between placebo and magne-
sium treatment for the following subgroups: clinically defined lacunar stroke, imaging-defined lacunar stroke, patients with diffusion–perfusion mismatch >20%, patients with both baseline DWI lesion volume ≥3 mL as well as diffusion–perfusion mismatch, and patients treated within 6 hours of symptom onset (supplemental Table I, available online at http://stroke.ahajournals.org). Using the prespecified global clinical outcome measure, percentage infarct growth (P = 0.015) and absolute infarct growth (P = 0.004) were greater among patients with poor global clinical outcome. The proportion with infarct growth >0% was also higher among patients with poor global clinical outcome (4% vs 42%; P = 0.084).

Baseline glucose levels were available for 88 of the 90 intention-to-treat patients. The median baseline serum glucose level was 7.1 mmol/L (128 mg/dL), with a range of 3.4 to 22 mmol/L (62–396 mg/dL). Among all patients, serum glucose correlated significantly with absolute (P = 0.028) and percentage (P = 0.024) growth; however, the higher serum glucose among patients with growth >0% was not statistically significant (P = 0.103; Table 4). Serum glucose also correlated with absolute growth (P = 0.012) and percentage growth (P = 0.048) for the magnesium-treated group. In the magnesium-treated group, serum glucose trended higher among patients with growth >0% compared to patients with growth ≤0% (P = 0.064).

However, in the placebo group, serum glucose did not correlate as strongly with absolute (P = 0.498) or percentage (P = 0.233) infarct growth. Among placebo-treated patients, there was no significant difference between patients with growth >0% and growth ≤0% (P = 0.552). Using logistic regression, the interaction of baseline glucose level and treatment allocation tended to predict infarct growth >0% (P = 0.071), independent of baseline glucose level and treatment allocation alone.

**Discussion**

The results of the MR IMAGES substudy are consistent with the findings of the main IMAGES clinical trial: no significant difference was demonstrated in infarct growth between placebo and magnesium-treated groups. Although the sample target size was not reached, the current data do not suggest that the results would have differed for the primary analysis with recruitment of additional patients.

Our study provides further evidence that imaging lesion volumes are good surrogates for clinical outcome. As previously shown, both acute DWI and PWI lesion volumes correlated well with baseline neurological status. Infarct growth was confirmed as a good surrogate for clinical neurological progression, here measured as a change in NIHSS from baseline to day 90, consistent with published literature. The weaker correlations of final infarct volume with PWI and mismatch lesion volumes, compared to DWI lesion volume, support previously published opinions that mismatch does not optimally define the penumbra. A perfusion deficit with Tmax ≥2 seconds has been shown to include areas of benign oligemia, which are not at risk for proceeding to infarction.

Most studies using mismatch as an inclusion criteria, including the Echoplanar Imaging Thrombolytic Evaluation Trial (EPITHET), the DEFUSE study, Desmoteplase in Acute Ischemic Stroke (DIAS) trial, and the Dose Escalation of Desmoteplase for Acute Ischemic Stroke (DEDAS) trial, used a mismatch threshold of >20%. Post hoc analyses of the DEFUSE study population found that a higher mismatch threshold of >80% more than doubled the odds of a favorable clinical response with early reperfusion. The
reduction of the included group from 73% to 47% by raising the mismatch threshold from >20% to >100% suggests that more than half of the patients in MR IMAGES population may not have had an appropriate penumbral pattern with potentially salvageable tissue. This is an important finding which may explain the negative results of studies with mismatch inclusion of >20% investigating the effect of intravenous tissue plasminogen activator in the 3- to 6-hour window (EPITHET) and intravenous desmoteplase up to 9 hours after symptom onset (DIAS and DEDAS). Future studies should consider alternative approaches to defining the penumbra, pending further post hoc analysis from the EPITHET, DEFUSE, and DEDAS trials.

Proof-of-concept trials using surrogate imaging outcomes may be useful in determining the feasibility of large phase III trials with clinical endpoints, testing the physiological effects of a drug with fewer patients and identifying subgroups of patients or possible covariates. Whereas studies of reperfusion therapies have demonstrated attenuated infarct growth and that this attenuation correlates with improved clinical outcome, there has been such positive trials testing neuroprotective agents. A phase II trial testing the neuroprotective effect of magnesium infusion in acute stroke using imaging surrogate outcomes would have likely yielded similar results as found in this study. Thus, a phase II proof-of-concept study designed similarly to the MR IMAGES substudy would have likely demonstrated no difference between magnesium and placebo groups in imaging surrogate outcome of infarct growth. Such a proof-of-concept study would have provided important insights and may have influenced the decision to proceed with the main IMAGES study.

A published further analysis of the IMAGES trial found that magnesium improved clinical outcome among patients with lacunar clinical syndromes. In this MRI substudy, we did not find any significant treatment-related differences in imaging outcome measures among patients with clinically defined lacunar syndrome and imaging-defined lacunar infarction. Possible reasons for these differing results are the very small number of patients with lacunar stroke and the known propensity for errors in infarct growth measurement for small lesions. None of the other subgroup analyses revealed any difference in infarct growth among treatment groups.

The deleterious effect of hyperglycemia on infarct growth in the overall group concurs with published literature showing that hyperglycemia is associated with poorer imaging and clinical outcomes in acute stroke. There was a trend toward greater negative influence of higher serum glucose on infarct growth with magnesium treatment compared to its effects with spontaneous infarct growth in the placebo group. Although not statistically significant, this finding is intriguing. The neuroprotective effect of magnesium chloride was attenuated by hyperglycemia in an animal model, as demonstrated by compensation with insulin coadministration. No effect of magnesium sulfate on glucose levels has been noted in animal models, multiple clinical trials in pregnancy, myocardial ischemia, and brain injury, or in clinical practice.

The limitations of this investigation include the shortcomings of studies with imaging outcomes such as corrupted/lost data and need to use imputed volumes for patients who died before day 90. In addition, there is no established best measure for infarct growth, hence we used 3 measures in this substudy, absolute growth in mL, percentage growth, and growth dichotomized for the presence of any growth >0%. Third, we did not reach the planned sample size. Fourth, 19% of patients enrolled did not fulfill the prespecified criterion of DWI lesion >3 mL. Finally, we had only 1 measurement of glucose. Persistent hyperglycemia after stroke has been shown to be a better indicator of infarct evolution and clinical outcome compared to any isolated measurement.

Summary
In conclusion, the lack of treatment effect on infarct growth in this MR substudy concurs with the clinical outcome finding in the main IMAGES study. These data, together with the confirmation that infarct growth correlates with clinical progression, support use of imaging surrogates in proof-of-concept studies before larger-scale clinical trials. An important secondary finding from this randomized imaging-based substudy is that a more appropriate mismatch threshold would exclude a sizeable proportion of patients in prior acute stroke trials using mismatch selection.

Appendix

Study Sites and Investigators
University of California Los Angeles (UCLA), USA: Chelsea S. Kidwell, Jeffrey L. Saver, Sidney Starkman, Bruce Ovbiagele, Abdul Rahman, Jeffrey R. Alger, Mark Pregenzer.
UCLA Olive View Medical Center, USA: Chelsea S. Kidwell, Jeffrey L. Saver, Sidney Starkman, Bruce Ovbiagele, Abdul Rahman, Jeffrey R. Alger, Mark Pregenzer.
Singapore General Hospital, Singapore: Christopher Chen, Hui-Meng Chang, Meng-Cheong Wong, Rhomy Esagunde, Aprilynn Reyes, Huey-Bee Pey, Francis Goh, Catherine Hong.
Southern General Hospital, Scotland: Keith W. Muir, Donald Hadley, Aileen Rowe, Isobel MacDonald, Barrie Condon, Michael McCormick.
Western Infirmary, Scotland: Kennedy R. Lees, Matthew R. Walters, Nigel McMillan, John Foster, Elizabeth Colquhoun, Belinda Manak, Muriel Coburn.
Royal Melbourne Hospital, Australia: Stephen M. Davis, Peter J. Hand, Bernard Yan, Patricia M. Desmond.
AZ. Sint-Jan AV, Brugge, Belgium: Geert Vanhooren, Valérie Schotte, Marc D’Hooghe, Ida Dehaene, Michel Van Zandijcke, Marcel Marchau, Kristof Verhoeven, Ludo Vanopdenbosch, Pieter Jan Simons, Elisabeth Foncke, Luc Régal, Jan Casselman.
University of South Carolina, USA: Te Long Hwang, Alex Rice.

Acknowledgments
The authors thank the members of the Data Safety and Monitoring Board for their support of the study: Bruce Dobkin, MD; Steven Warach, MD, PhD, and Michael Hill, MD.

Sources of Funding
This study was supported by a grant from the National Institutes of Neurological Disorders and Stroke (NINDS)/National Institutes of Health (NIH), K23 NS 02088 (to C.S.K., C.C., J.L.S.), by the Medical Research Council of the United Kingdom, and by the Chief Scientist Office of Scotland.
Disclosures

Stephen M. Davis is a member of the Speaker’s Bureau for Sanofi Aventis, Bristol Myers Squibb, Pfizer, and Boehringer Ingelheim; received honoraria from Sanofi Aventis, Bristol Myers Squibb, Pfizer, and Boehringer Ingelheim; and is a member of the Advisory Board for Novo Nordisk and Astra Zeneca. Jeffrey L. Saver received a research grant from National Institutes of Neurological Disorders and Stroke (NINDS)/National Institutes of Health (NIH), K23 NS 02088. There are no other conflicts to report.

References


Results of the MRI Substudy of the Intravenous Magnesium Efficacy in Stroke Trial
Chelsea S. Kidwell, Kennedy R. Lees, Keith W. Muir, Christopher Chen, Stephen M. Davis, Deidre A. De Silva, Christopher J. Weir, Sidney Starkman, Jeffry R. Alger and Jeffrey L. Saver for the MR IMAGES Investigators

Stroke. 2009;40:1704-1709; originally published online March 19, 2009;
doi: 10.1161/STROKEAHA.108.537613

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/40/5/1704

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org//subscriptions/