STROKE

Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia

Gemma Llovera,[1,2] Kerstin Hofmann,[1,2] Stefan Roth,[1,2] Angelica Salas-Pérdomo,[3,4] Maura Ferrer-Ferrer,[3,4] Carlo Perego,[5] Elisa R. Zanier,[5] Uta Mamrak,[1,2] Andre Rex,[6] Hélène Party,[7] Véronique Agin,[7] Claudine Fauchon,[8] Cyrille Orset,[7,8] Benoît Haelewyn,[7,8] Maria-Grazia De Simoni,[5] Ulrich Dirnagl,[6] Ulrike Grittner,[9] Anna M. Planas,[3,4] Nikolaus Plesnila,[1,2] Denis Vivien,[7,8] Arthur Liesz[1,2]*

Numerous treatments have been reported to provide a beneficial outcome in experimental animal stroke models; however, these treatments (with the exception of tissue plasminogen activator) have failed in clinical trials. To improve the translation of treatment efficacy from bench to bedside, we have performed a preclinical randomized controlled multicenter trial (pRCT) to test a potential stroke therapy under circumstances closer to the design and rigor of a clinical randomized control trial. Anti-CD49d antibodies, which inhibit the migration of leukocytes into the brain, were previously investigated in experimental stroke models by individual laboratories. Despite the conflicting results from four positive and one inconclusive preclinical studies, a clinical trial was initiated. To confirm the preclinical results and to test the feasibility of conducting a pRCT, six independent European research centers investigated the efficacy of anti-CD49d antibodies in two distinct mouse models of stroke in a centrally coordinated, randomized, and blinded approach. The results pooled from all research centers revealed that treatment with CD49d-specific antibodies significantly reduced both leukocyte invasion and infarct volume after the permanent distal occlusion of the middle cerebral artery, which causes a small cortical infarction. In contrast, anti-CD49d treatment did not reduce lesion size or affect leukocyte invasion after transient proximal occlusion of the middle cerebral artery, which induces large lesions. These results suggest that the benefits of immune-targeted approaches may depend on infarct severity and localization. This study supports the feasibility of performing pRCTs.

INTRODUCTION

Ischemic stroke, a leading cause of death and disability worldwide (*1*), induces the rapid loss of specific brain functions, a result of insufficient blood flow to specific brain regions. Together with other downstream effects, stroke triggers an acute inflammatory response in the brain, which activates harmful signaling cascades that contribute to secondary brain damage (*2*). Although stroke places enormous medical and economic burdens on society, thrombolysis with tissue plasminogen activator (tPA) is currently the only clinically approved therapy for ischemic stroke. Safety considerations require, however, that tPA must be administered within a very narrow time window after the onset of symptoms, and tPA treatment has important contraindications such as hemorrhage or oral anticoagulation (*3*, *4*). Because of these constraints, tPA is only provided to a small percentage of all stroke patients in industrialized countries; the overall therapeutic impact of tPA is negligible when developing countries are also considered.

In recent decades, several experimental therapies have been developed and tested for their ability to mitigate stroke-related brain injury, which can include postischemic, excitotoxic neuronal damage, spreading depolarization, apoptosis, and/or inflammation (*5*, *6*). Unfortunately, none of the tested drugs that show promise in animal models of stroke have so far been translated into clinical use for stroke patients (*7*, *8*), and most major pharmaceutical companies have stopped research and development in this area. The cause of this failure in clinical trials is multifactorial and may include poorly designed preclinical and clinical studies, biased selection of substances for clinical testing, and underpowered clinical trials with overambitious and pathophysiologically irrelevant therapeutic windows (*5*, *9*). Indeed, academic and industrial researchers, as well as funding agencies and journals, now recognize the existence of a "reproducibility crisis": the results of preclinical studies in many research fields, including stroke, lack robustness, and only a small fraction of these studies can be replicated (*10*).

To overcome the current limitations of preclinical in vivo studies, strict operational and statistical guidelines have been developed for the data reporting, and suggestions have been implied for design and performance of preclinical studies (*11*). In addition, pRCTs have been proposed as a way to help bridge the gap between experimental laboratory research and clinical trials (*12*–*14*). Such studies would be primarily confirmatory, designed to test efficacy of previously defined therapeutic concepts in a study design with the highest standards for statistics, analysis, and reporting (*15*). Finally, such a study will have many of the

[1]Institute for Stroke and Dementia Research, Klinikum der Universität München, Feodor-Lynen-Straße 17, 81377 Munich, Germany. [2]Munich Cluster for Systems Neurology (SyNergy), 81377 Munich, Germany. [3]Department of Brain Ischemia and Neurodegeneration, Institut d'Investigacions Biomèdiques de Barcelona (IIBB), Consejo Superior de Investigaciones Científicas (CSIC), 08036 Barcelona, Spain. [4]Àrea de Neurociències, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), 08036 Barcelona, Spain. [5]Neuroscience Department, IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, 20156 Milan, Italy. [6]Department of Experimental Neurology and Center for Stroke Research Berlin (CSB), Charité–Universitätsmedizin Berlin; German Center for Neurodegenerative Diseases (DZNE) and German Center for Cardiovascular Diseases (DZHK), Berlin sites; Excellence Cluster NeuroCure, 10117 Berlin, Germany. [7]INSERM, UMR-S U919, Université de Caen Basse-Normandie, team Serine Proteases and Pathophysiology of the neurovascular Unit, GIP Cyceron, F-14074 Caen Cedex, France. [8]Experimental Stroke Research Platform (ESRP), IBiSA platform, Centre Universitaire de Ressources Biologiques (CURB), Université de Caen Basse-Normandie, F-14074 Caen Cedex, France. [9]Department of Biostatistics and Clinical Epidemiology, Charité–Universitätsmedizin Berlin, 12203 Berlin, Germany.
*Corresponding author. E-mail: arthur.liesz@med.uni-muenchen.de

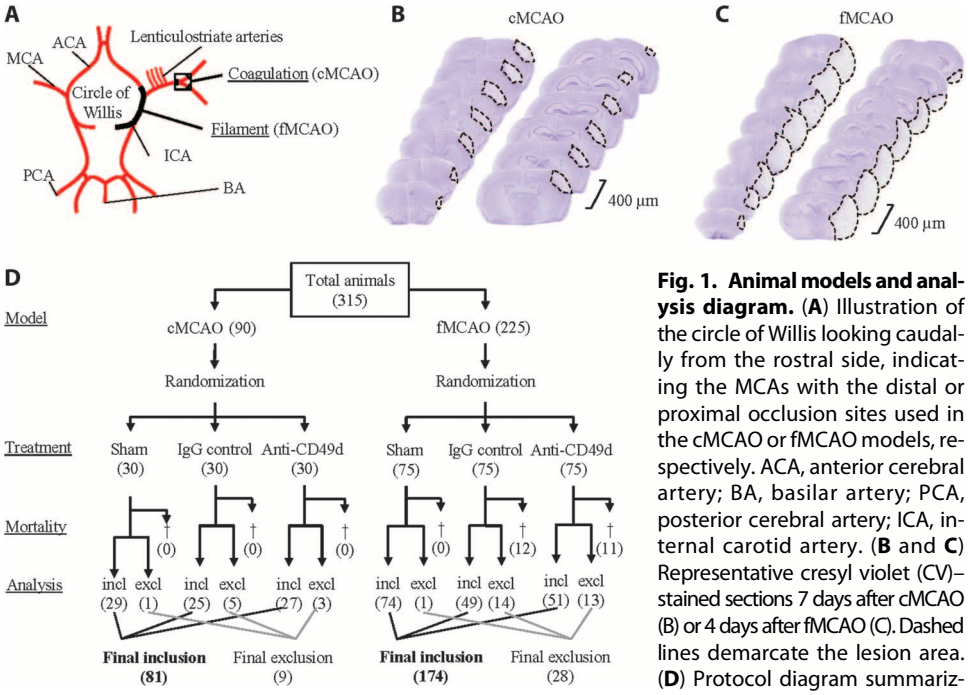


Fig. 1. Animal models and analysis diagram. (**A**) Illustration of the circle of Willis looking caudally from the rostral side, indicating the MCAs with the distal or proximal occlusion sites used in the cMCAO or fMCAO models, respectively. ACA, anterior cerebral artery; BA, basilar artery; PCA, posterior cerebral artery; ICA, internal carotid artery. (**B** and **C**) Representative cresyl violet (CV)–stained sections 7 days after cMCAO (B) or 4 days after fMCAO (C). Dashed lines demarcate the lesion area. (**D**) Protocol diagram summarizing the number of animals included (incl) (315 mice), with exclusion (excl) per group and included animal numbers for final analysis. Exclusion criteria were defined as described in Materials and Methods. "†" indicates dead animals.

characteristics of adequately powered clinical trials. Several international consortia have been established to perform such studies, and multicenter investigations have been reported, but no pRCT has been published to date. Therefore, our objective was to implement a preclinical multicenter trial by applying rigorous design, experimental, and analysis standards to test the efficacy of an anti-inflammatory therapeutic intervention (that is, blocking leukocyte migration with anti-CD49d antibodies) in two experimental models of stroke.

Brain leukocyte invasion is a key mechanism that mediates secondary neuronal injury after an acute brain lesion (*2*, *16*). Transendothelial migration is tightly regulated by adhesion molecules, specifically the interaction between the molecules VLA-4 (leukocyte very late antigen-4) and VCAM-1 (vascular cell adhesion molecule-1) (*17*, *18*). Antibodies against the α chain of VLA-4 (anti-CD49d antibodies) have shown efficacy in several models of autoimmune diseases, and the humanized antibody natalizumab is currently one of the most effective therapies for patients with multiple sclerosis (*19*). Four single-center studies reported that anti-CD49d antibodies confer cerebroprotective properties in mouse models of brain ischemia by blocking the postischemic invasion of leukocytes into the brain (*20*–*23*). However, one of these studies reported different efficacies between different stroke models (*20*). The neuroprotective potential of anti-CD49d treatment was challenged further by a subsequent study that reported no treatment effect (*24*). A phase 2 clinical trial is currently testing the effect of natalizumab in patients after acute ischemic stroke (ClinicalTrials.gov identifier: NCT01955707).

The availability of reliable preclinical data is essential for selecting the drug candidates with the highest clinical potential and for designing the most effective clinical study, thereby minimizing unnecessary, costly, and even harmful clinical trials. We therefore tested anti-CD49d treatment in a multicenter, randomized preclinical trial. The predetermined primary end point for the trial was infarct volume measured in two ischemic stroke models. The secondary end points were functional outcome and the invasion of leukocytes into the brain.

RESULTS

For this experimental trial, we used two different models of focal brain ischemia (Fig. 1, A to C) to address potential pathophysiological differences among commonly used stroke models (*25*, *26*). Small lesions confined to the cortex were induced by coagulation of the distal middle cerebral artery (cMCAO), whereas lesions in the cortex and subcortical structures were induced by transiently occluding the MCA at its origin for 60 min with an endovascular filament (fMCAO). A total of 315 male C57BL/6J mice from the following five independent European research centers were randomized in this trial by the author indicated by their initials [center 1: Caen (D.V.), center 2: Milan (M.-G.D.S.), center 3: Barcelona (A.M.P.), center 4: Munich (N.P.), and center 5: Berlin (U.D.)], whereas the sixth center centrally coordinated the study and analyzed the data (Munich, A.L.). A total of 81 and 174 mice were included in the final analyses of the cMCAO and fMCAO groups, respectively (Fig. 1D). The animals' physical features, mortality incidences, and reasons for exclusion are summarized in Table 1. The body temperature of the mice at center 3 differed significantly from the other centers, and the body weight of the mice at center 4 differed significantly from the other centers; however, all values were within their respective normal physiological range for mice.

Anti-CD49d treatment reduces infarct volume after cortical stroke

Previous studies performed in individual laboratories with relatively small sample sizes suggested that anti-CD49d antibodies might improve outcome after stroke (*20*–*23*); in contrast, another study reported no significant benefit of this treatment (*24*). We therefore examined the efficacy of anti-CD49d treatment in experimental ischemic stroke by performing a pRCT, using pooled infarct volume of all centers per each stroke model as the primary end points. The anti-CD49d antibody or a control immunoglobulin G (IgG) antibody (300 µg each) was administered intraperitoneally in a randomized and blinded fashion 3 hours after stroke induction. Analysis of the pooled data from all centers revealed that treatment with anti-CD49d antibodies resulted in 19% smaller infarct volume 7 days after cMCAO compared to administration of control IgG ($P < 0.05$) (Fig. 2A). In contrast, 4 days after inducing fMCAO, infarct volume did not differ between control IgG–treated and anti-CD49d–treated mice, regardless of whether the data were analyzed for each individual center or as a pooled data set (Fig. 2B). To additionally show the effect sizes (ESs, standardized mean difference) of

Table 1. Animal characteristics.

Center	Stroke model	Time after lesion (days)	Total *n*	Mortality	Exclusion: No infarct	Exclusion: Other criteria*	Final *n*	Mean body temperature†	Mean body weight
1	Filament	4	45	1	9	1	34	37.43	25.02
2	Filament	4	45	9	6	—	30	37.88	23.13
3	Filament	4	45	6	3	—	36	38.77‡	23.35
4	Filament	4	45	—	—	—	45§	—	20.60‡
5	Filament	4	45	7	2	7	29	37.95	24.81
1	Coagulation	7	30	—	—	—	30	37.51	24.62
2	Coagulation	7	30	—	2	2	26	37.74	23.39
3	Coagulation	7	30	—	5	—	25	38.35‡	23.88

*Other exclusion criteria: Neuroscore <8 at 24 hours after fMCAO, overt mechanical damage, or disruption of the brain sample or sham animals with lesion. †Baseline values before anesthesia induction. ‡$P < 0.05$ [analysis of variance (ANOVA)]. §$P < 0.05$ (χ^2).

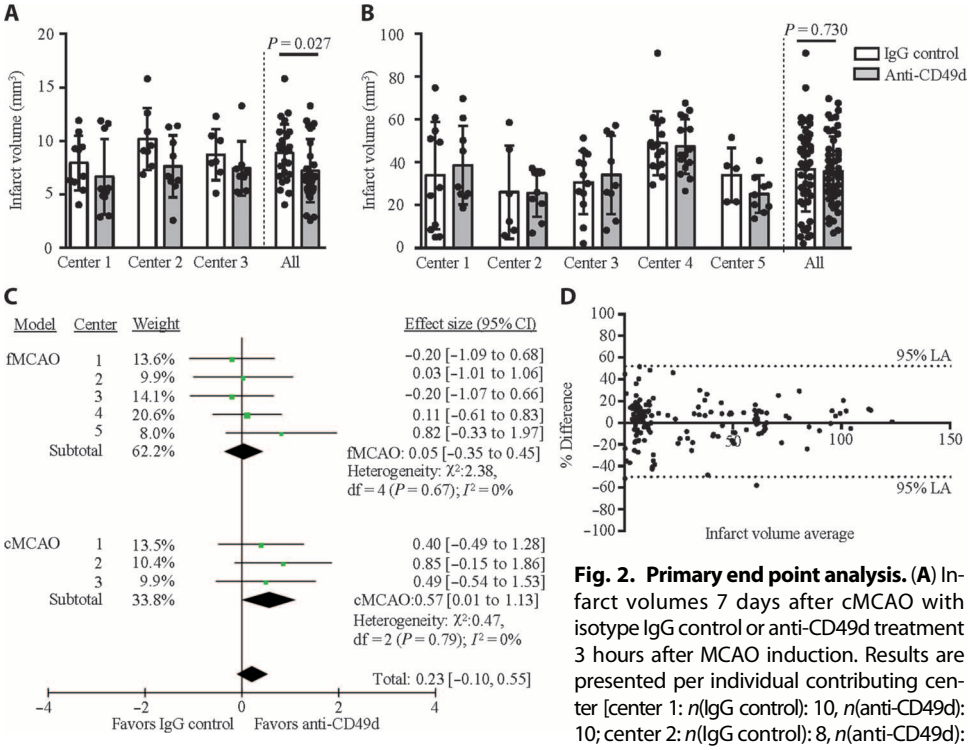


Fig. 2. Primary end point analysis. (**A**) Infarct volumes 7 days after cMCAO with isotype IgG control or anti-CD49d treatment 3 hours after MCAO induction. Results are presented per individual contributing center [center 1: *n*(IgG control): 10, *n*(anti-CD49d): 10; center 2: *n*(IgG control): 8, *n*(anti-CD49d): 9; and center 3: *n*(IgG control): 7, *n*(anti-CD49d): 8] and as pooled samples from all centers [*N*(IgG control): 25, *N*(anti-CD49d): 27]. Mixed model analysis revealed significantly (β = 1.8, SE = 0.8, *P* = 0.027) lower infarct volumes in the anti-CD49d treatment group. (**B**) Corresponding results for infarct volumes 4 days after fMCAO (β = −0.3, SE = 3.3, *P* = 0.947). Results are presented per individual contributing center [center 1: *n*(IgG control): 11, *n*(anti-CD49d): 9; center 2: *n*(IgG control): 6, *n*(anti-CD49d): 9; center 3: *n*(IgG control): 12, *n*(anti-CD49d): 9; center 4: *n*(IgG control): 15, *n*(anti-CD49d): 15; and center 5: *n*(IgG control): 5, *n*(anti-CD49d): 9] and as pooled samples from all centers [*N*(IgG control): 49, *N*(anti-CD49d): 51]. Center labels for contributing centers are consistent between the two stroke models. Data are presented as means ± SD. (**C**) Forest plot of ES estimation by the inverse-variance method in a random-effects model in individual centers and in all-center pooled data, corresponding to centers depicted in (A) and (B). (**D**) Bland-Altman plot for interrater concordance of infarct volume measurement for the two independent and blinded raters at the central study center. LA, limit of agreement.

antibody treatment in the two stroke models, we performed a meta-analysis using inverse-variance weighting (random-effects model) from all centers; these results are illustrated as a forest plot (Fig. 2C). Consistent with the results from the mixed model analysis used to test the primary end point (Fig. 2, A and B), the result from the cMCAO model showed no significant effect of treatment if analyzed separately for each center, whereas a subtotal analysis (pooled for cMCAO) revealed a significant treatment effect [ES, 0.57; 95% confidence interval (CI), 0.01 to 1.13]. As in the mixed model primary analysis, no significant effect was found in the analysis for pooled infarct volume after fMCAO. When the data were pooled from all treatment centers and both models, the ES no longer favored treatment (ES, 0.23; 95% CI, −0.1 to 0.55).

Next, we investigated the accuracy of our assessment of the primary end point. Infarct volume was independently measured by two researchers (G.L. and A.L.) at the central study center. These researchers were blinded with respect to the groups, and concordance of their results was confirmed using Bland-Altman analysis and calculation of intraclass correlation coefficient (ICC), which revealed excellent interrater reliability (ICC, 0.99; 95% CI, 0.990 to 0.994; Fig. 2D and fig. S1A). Moreover, each study center reevaluated the infarct volumes of the samples, confirming reliability between the central analysis and the analyses from each center (ICC range, 0.88 to 0.99; fig. S1B). Together, these results support the hypothesis that anti-CD49d treatment reduces infarct volume in cortical lesions after permanent MCA occlusion; moreover, our results suggest that treatment efficacy might depend on infarct type and severity, because anti-CD49d treatment was not effective after transient occlusion of the proximal MCA.

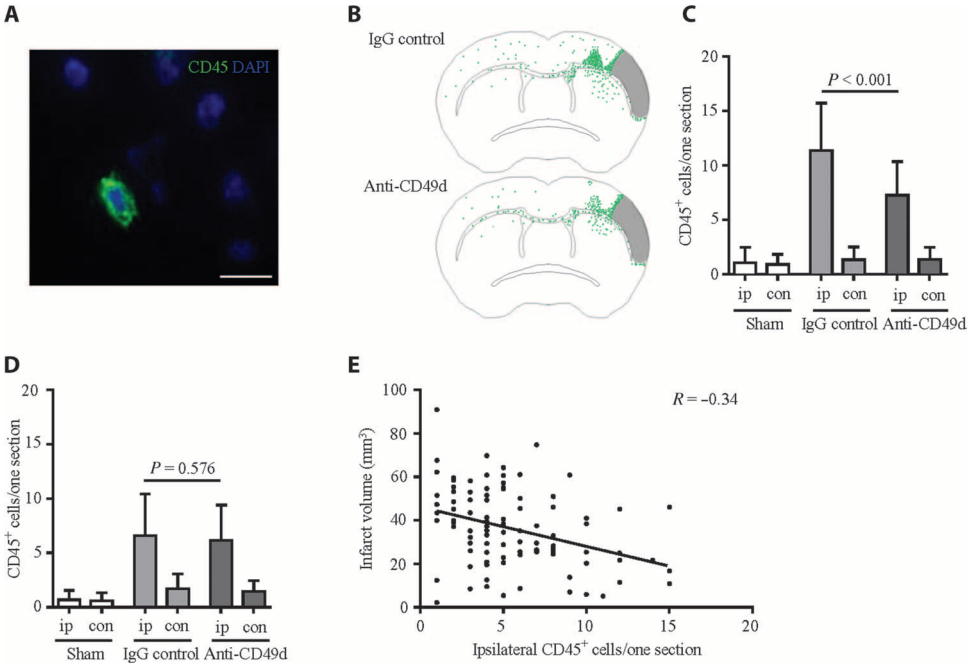


Fig. 3. Leukocyte brain infiltration after cMCAO and fMCAO. (**A**) Representative image of CD45 immunohistological staining in the peri-infarct area 7 days after cMCAO. Green, CD45–Alexa Fluor 488; blue, 4′,6-diamidino-2-phenylindole (DAPI). Scale bar, 10 μm. (**B**) Cumulative topographic representation of each single $CD45^+$ leukocyte detected on one coronal section per brain, excluding the infarct core area. Topographic maps for each section were superimposed for both treatment groups [*n*(IgG control): 25, *n*(anti-CD49d): 27]. Total $CD45^+$ count in ipsilateral hemispheres was 284 cells in the IgG group and 196 cells in the anti-CD49d group. (**C**) Quantitative analysis of $CD45^+$ cell count per one coronal section in ipsilateral (ip) and contralateral (con) hemispheres of sham-operated animals ($n = 30$) and anti-CD49d–treated ($n = 25$) or IgG control–treated ($n = 27$) mice 7 days after cMCAO. (**D**) Corresponding quantification of $CD45^+$ cells 4 days after fMCAO, sham-operated animals ($n = 74$), and anti-CD49d–treated ($n = 49$) or IgG control–treated ($n = 51$) mice. Data in (C) and (D) are presented as means ± SD. (**E**) Linear regression analysis of $CD45^+$ cell counts per ipsilateral hemisphere of both treatment groups (*x* axis) and infarct volume (*y* axis) 4 days after fMCAO. *r*, Pearson correlation coefficient.

Anti-CD49d treatment reduces the invasion of leukocytes into the brain after cortical stroke

Previous studies (*20*, *24*) suggested that anti-CD49d treatment exerts its primary therapeutic effect on brain ischemia by inhibiting leukocyte invasion of the brain. To investigate this process, we prepared brain sections from mice after MCAO or sham surgery; we then immunostained the sections with antibodies against CD45, a marker for leukocytes (Fig. 3A). The stained sections were analyzed by researchers who were blinded with respect to the surgical and treatment groups. Cumulative topographic maps of leukocyte invasion after cMCAO revealed that leukocytes infiltrated the region surrounding the infarct core, with predominant localization in the peri-infarct cortical area (Fig. 3B). Anti-CD49d treatment significantly ($P < 0.001$) reduced the number of leukocytes in the infarct area after cMCAO (Fig. 3, B and C), consistent with our finding of reduced lesion volume in anti-CD49d–treated mice after cMCAO. Mice that underwent fMCAO had significantly fewer cerebral leukocytes than those that underwent MCAOcMCAO [mean (SD), 5.4 (3.6) versus 11.4 (4.4) cells per section, respectively; $P < 0.001$], and anti-CD49d treatment did not affect the invasion of leukocytes after fMCAO (Fig. 3D). Because of the high variability in lesion size and location in the fMCAO model, we were unable to generate topographic maps; however, we measured a weak yet significant inverse correlation between infarct volume and cerebral leukocyte count in the fMCAO-induced group (Fig. 3E), suggesting a relationship between lesion size and cellular poststroke neuroinflammation.

Other secondary outcome parameters were not affected by anti-CD49d treatment

In addition to the infiltration of leukocytes into the brain, we also measured the following outcomes as predefined secondary end points: mortality, physiological parameters, and behavioral deficits. Our analysis revealed that none of these outcomes was significantly affected by anti-CD49d treatment. For the cMCAO model, we used the rotarod test and the adhesive removal test, two well-established functional tests for this stroke model. An analysis of functional outcome with the per-protocol analysis revealed that sensorimotor deficits in both tests (at three time points after stroke) did not differ significantly between the sham surgery group and either the control-treated or anti-CD49d–treated cMCAO groups. Hence, we were unable to discriminate between cMCAO and sham surgery, indicating that these commonly used tests lacked sufficient sensitivity to detect the limited neurological deficits in the cMCAO model (fig. S2, A to C and E). Additionally, the mean absolute values for both tests differed greatly between individual centers (fig. S2, A and B), indicating high intercenter variability in performing these tests, despite the use of harmonized procedures. To eliminate intercenter variability as a source for increased variability for this parameter, we performed a secondary analysis in which the poststroke deficits were normalized to their respective baseline values (fig. S2F). This analysis revealed a cMCAO-induced deficit in the adhesive removal test; however, treatment with the anti-CD49d antibody had no significant effect on this deficit. Because of the substantial behavioral deficits induced by fMCAO, the tests used for cMCAO could not be applied in this model. Therefore, we tested deficits in the fMCAO model by a composite Neuroscore, with no difference apparent between treatment groups (fig. S3A); this lack of treatment effect is consistent with the lack of treatment effect on infarct volume in this model. The composite Neuroscore test robustly detected significant deficits in the animals (compared to sham treatment) in all five research centers 4 days after stroke. A correlation analysis of this large cohort of animals ($n = 174$ fMCAO-induced mice) revealed moderate correlation between infarct volume and Neuroscore outcome ($r = 0.76$ and $r = 0.77$, 2 and 4 days after stroke, respectively; fig. S3B). Although cMCAO did not cause mortality (within 7 days after stroke), fMCAO resulted in similar mortality rates in the control IgG–treated (16.0%) and the anti-CD49d–treated (14.6%) groups (fig. S4). Similarly, the physiological parameters (body temperature and body weight) were affected by fMCAO but not by cMCAO (fig. S5). Specifically, fMCAO-induced brain

damage led to pronounced hypothermia and weight loss, whereas sham surgery did not; however, these parameters were not affected by treatment with the anti-CD49d antibody.

DISCUSSION

Here, we report the results of an international multicenter, randomized, controlled, and blinded experiment for the preclinical testing of drug efficacy. The design of this study was modeled on phase 3 clinical randomized controlled trials, which are the current standard in clinical drug development and are required to license a new drug. Our study was performed in response to the "replication crisis" (*10*), the exceedingly low reproducibility, and lack of robustness of preclinical results in academic research (*27*–*29*). In addition, our approach may help to resolve the translational roadblock, in which promising preclinical approaches fail to be translated to clinically effective therapies (*30*–*32*). Thus, international research consortia—which provide larger sample sizes with sufficient statistical power and a study design that reduces the confounding effects of bias—have been proposed as a way to improve the robustness and translational predictability of preclinical research (*13*, *14*, *33*–*35*).

To improve the quality and reliability of preclinical research, several international research consortia have been established, including the European Union (EU)–funded Multi-PART (Multicentre Preclinical Animal Research Team) consortium for stroke research and the National Institutes of Health–funded CAESAR (Consortium for preclinicAl assESsment of cARdioprotective therapies) consortium for cardiovascular research. Systematic efforts by these consortia and others have revealed that inadequate reporting of data, statistical flaws, and missing cross-validation of data from independent study centers are major flaws in preclinical translational research (*5*, *13*, *33*, *36*, *37*). Virtually all preclinical studies performed by individual research groups are underpowered and often fail to adequately control for bias (for example, using randomization and blinding) (*38*). For example, the average group size in more than 2000 different experiments involving more than 35,000 rodents (in which hundreds of different treatments were tested in ischemic stroke models) is only 8.6 animals per group (median, 8; range, 1 to 54) (*39*). Ioannidis *et al.* (*34*) examined a random sample of in vivo animal studies published since 2006 and found that fewer than 30% of studies used randomization and fewer than 10% performed a blinded assessment of outcome; with respect to experimental stroke research, these percentages are even lower (*40*). A meta-analysis of 49 published preclinical meta-analyses (comprising 730 primary individual studies in the field of neuroscience) revealed that the median statistical power was only 21% (*41*). This means that of 100 studies that are conducted to investigate a genuine treatment effect, only 21 will actually be powered in such a way as to be able to demonstrate the effect, a result of small sample sizes. Until as recently as several decades ago, clinical trials were challenged by the same issues that face preclinical biomedical research today, including low internal validity, low statistical power, and a high rate of false positives. By developing and implementing rigorous standards for the design, execution, analysis, and reporting of studies, today's clinical trials can deliver robust and relevant results. Learning—and adopting key measures—from current clinical trial design (for example, randomization, blinding, and a priori power analysis), as in our pRCT, may help to overcome the current crisis of preclinical translational medicine. Such an approach has clear advantages over a meta-analysis, which pools results from studies with different methodological and statistical standards, potentially differing protocols, and lack of monitoring. In contrast, as explained in detail below, our pRCT approach aimed to harmonize protocols and warranted a prospective, blinded, and randomized study design, central monitoring of data quality, and centralized analysis and data deposition.

Here, we used the two most common experimental models of ischemic stroke to test a drug candidate that is already being tested in a phase 2 clinical trial after only a few, small-scale, single-center studies in mice. In all but one case (*8*), this approach in stroke research has led until now to costly failures and the exposure of patients to ineffective—or potentially harmful—drugs (*42*). Many other medical fields share this frustrating experience. The primary objective in the design of our study was to achieve maximum harmonization of the methods used by the various participating study centers. The basic protocol, as well as the primary and secondary end points, was predefined by coordination between the six laboratories. In clinical trials, patient care among various trial sites is relatively uniform and performed in accordance with national and international guidelines. In our study, achieving harmonization of the materials and standard operating procedures was a challenge in the initial phase of our study because the various laboratories generally use different strategies for testing drugs in the stroke models. We achieved the best possible harmonization of the surgical procedures and behavioral tests by using same-age and same-gender mice obtained from the same commercial breeder, centrally supplying all critical materials (for example, the filaments used to induce transient stroke), and defining the methods used for the surgical procedures, behavioral tests, and data acquisition. Several factors could not be harmonized fully, including the use of specific anesthetics and analgesics (a result of differences in local regulations regarding the conduct of animal experiments), the expertise of the surgeons, postsurgical care, body temperature maintenance, and monitoring of cerebral blood flow after vessel occlusion; all of these factors can potentially influence stroke outcome. Nevertheless, comparable differences in equipment, legal requirements, and the skills of the medical personnel also confound most clinical RCTs. Other variables included between-site differences in the use of sedation protocols, in postoperative anesthesia and analgesic drugs, and in mouse microbiota.

These remaining methodological differences might actually increase the robustness of our obtained results with respect to clinical translation. Indeed, we were able to demonstrate that drug efficacy could be demonstrated despite the above-mentioned procedural differences. Furthermore, increasing the sample size by pooling the efforts of multiple laboratories reduced statistical variability and increased sensitivity of outcome measures such as the infarct volume in the cMCAO model, cerebral leukocyte counts, and the Neuroscore.

Another aspect of the study that was handled individually by each study site was the approval of the animal experiments. Because of differing regulatory processes in Spain, France, Italy, and Germany, a uniform approval for all European partner sites was not achievable despite the European directive 2010/63/EU. However, future pRCTs could get approval more easily or via a single-point process for all partner sites if pRCTs showed more robust results than a single study and were regarded as a part of the preclinical evaluation of drug development.

Natalizumab, a humanized anti-CD49d antibody, is currently one of the most effective disease-modifying drugs for multiple sclerosis (*19*). With more than 7 years of clinical use, the efficacy and safety profiles for natalizumab are well established. The major risk associated with this

treatment approach (the induction of progressive multifocal leukoencephalopathy) usually emerges after long-term treatment, but not after a single dose. Recently, several research groups established that lymphocyte invasion plays a role in stroke pathophysiology (*20*, *43*–*46*). For example, lymphocyte-deficient mice have substantially smaller lesion volumes compared to wild-type mice (*43*). Therefore, targeting the adaptive immune system is a promising therapeutic strategy for stroke. However, the contribution of secondary neuroinflammation to poststroke pathophysiology is far less explored, and the translatability of murine mechanisms to the human situation regarding stroke immunology is still under discussion (*47*, *48*). On the basis of previous experimental studies with anti-CD49d antibodies in stroke models (*20*–*22*) and the safety data on natalizumab use in patients with multiple sclerosis, a phase 2 clinical trial was initiated recently to test the effect of natalizumab in patients with acute ischemic stroke [ACTION (Effect of Natalizumab on Infarct Volume in Acute Ischemic Stroke) trial, ClinicalTrials.gov identifier: NCT01955707]. During the ACTION trial, an animal study was published that reported no effect of natalizumab on stroke (*24*), raising concerns about the efficacy of this approach (*24*, *49*). In contrast, in our multicenter study, consistent with previous results (*20*), we found that anti-CD49d treatment exerted a modest yet significant neuroprotective effect in the cMCAO model, but not in the fMCAO model of transient occlusion, which causes extensive damage. Notably, this protective effect was statistically significant when the cMCAO results were pooled from all centers but not when analyzed in each individual center. When we pooled results from both stroke models, no significant treatment effect was observed.

The difference in the neuroprotective efficacy of anti-CD49d treatment between the two models might be attributed to biological differences in the underlying pathophysiology and/or insufficient statistical power as a result of the unexpectedly high variability in the fMCAO model. In a head-to-head comparison of the fMCAO and cMCAO models, the authors found large differences in neuroinflammatory markers between the two models (*26*). Unexpectedly, moderate-sized cortical lesions, which appeared after permanent occlusion in the cMCAO model, induced higher levels of leukocyte brain invasion, microgliosis, and proinflammatory cytokine release than did the extensive hemispheric lesions of the transient ischemia model (*26*). Consistent with these results, we also found about twofold higher cerebral leukocyte counts after cMCAO than fMCAO. However, it should be noted that different time points (4 days after fMCAO versus 7 days after cMCAO) were used in the two models because of the high mortality rate in the fMCAO-treated mice after 4 to 5 days; moreover, the cMCAO model requires craniectomy and penetration of the dura. However, a recent study by Chu *et al.* reported twofold higher cerebral leukocyte cell counts after permanent occlusion by a filament in the fMCAO model compared to transient occlusion in the same stroke model, indicating that there may be pronounced leukocyte brain invasion in situations in which reperfusion is not established, regardless of the stroke model used (*50*). In our study, the cerebral leukocyte counts in the anti-CD49d–treated cMCAO group (the permanent occlusion model) were considerably higher than in the fMCAO group, and we found no treatment effect of anti-CD49d on the relatively low leukocyte numbers in the fMCAO model. Therefore, inhibition of the migration of leukocytes into the brain with anti-CD49d antibodies may be more effective in strokes that trigger a more robust inflammatory reaction. This possibility has been generally overlooked in preclinical testing of immunotherapeutics for stroke, and it has also not been considered in the design of clinical studies.

Because of its unique design (pooling results from five centers and the unprecedented large sample size in a single preclinical experiment), our study provides a critical view of commonly used methods in experimental stroke research. The variability of outcomes differed substantially between the two models; specifically, the overall SD in the control-treated cMCAO group was 30% [mean (SD), 8.9 (2.7); $n = 25$ mice], which is considerably lower than the surprisingly high SD in the control-treated fMCAO group [53%; mean (SD), 36.4 (19.9); $n = 49$]. Furthermore, our study also revealed an unexpectedly weak sensitivity of some of the most widely used behavioral tests, particularly in the cMCAO model, which induced only subtle neurological deficits. The rotarod and adhesive tape removal tests are two of the most common tests performed in experimental stroke research (*51*), and they have been successfully used by each laboratory in our study. Nevertheless, the rotarod test failed to detect poststroke deficits compared to sham surgery, even after the test results were normalized to baseline values, thus excluding intercenter differences. Moreover, the adhesive removal test results showed wide variability, even after normalization, and therefore lacked statistical power to detect treatment effects; nevertheless, this test was able to detect significant deficits after cMCAO. In addition, the Neuroscore, which was used to quantify deficits after fMCAO, also had high variability and revealed only a moderate correlation between test results and lesion volume. It was also apparent that the laboratory with previous experience using the Neuroscore (center 2) obtained more consistent results with this test than did laboratories that used this test for the first time (fig. S3). This finding might reflect a more general challenge with respect to harmonizing procedures, which was a necessary step in designing our pRCT. For example, deciding a priori to use mice of a specific age or deciding to centrally distribute the surgical material might necessitate changes in local standard operating procedures, which, in turn, could increase variability. Nevertheless, we reasoned that harmonizing the basic procedures and using similar materials were essential for achieving a multicenter, randomized, blinded pRCT, thereby minimizing intercenter effects and facilitating the analysis of pooled data based on group means.

On the basis of the previous considerations and new methodological insights gained in our study, we were able to define several points of potential improvement for future pRCTs: an obvious limitation for our analysis strategy was the constraint to a per-protocol analysis, lacking an additional intention-to-treat analysis. This is due to a substantial amount of missing data for the primary outcome measure (infarct volume) because of mortality or no infarct demarcation. This might be circumvented in future studies by using inclusion criteria similar to clinical trials such as imaging modalities for confirmation of an infarct, a defined neurological deficit range as an inclusion criterion, and use of a behavioral readout marker or mortality as primary end point. In analogy to clinical trials, which pool stroke patients with differing stroke etiology, future pRCTs might also consider pooling outcome of different models for testing drug efficacy.

One surprise in our study was the low performance of the behavioral tests, as well as exceedingly high variability of the fMCAO model. It would be desirable to evaluate such methodological characteristics in "pretrials" before future pRCTs, with the aim to validate the used methods and grade of harmonization in-between study centers before performing the actual pRCT. Such a pretrial would also give the opportunity to include a dose-escalation study for the investigated drug before defining the dose for the pRCT. Finally, central assessment of behavioral outcomes based on video recordings might reduce variability and improve test sensitivity, which should be evaluated in a pretrial for a future pRCT.

Given the lack of funding for international collaborative confirmatory research, the apparent lack of incentives for individual researchers and laboratories, and the competitiveness associated with preclinical biomedicine, so-called pRCT studies—although deemed necessary—are often believed to be impossible to implement. Here, we report the feasibility of this approach, despite the existence of several obstacles, and we hope that this study will encourage researchers in all biomedical research fields to consider forming similar consortia to perform essential preclinical trials before advancing to clinical development; this approach is particularly relevant to biomedical research fields that have traditionally been hampered by lack of reproducibility and translational roadblock. We estimated the total full cost of our study at about €165,000 ($ 180,000), including more than €30,000 direct drug costs (table S1). Because the use of pRCTs may avoid unnecessary clinical trials and improve our ability to predict successful translation, these costs for a confirmatory study seem reasonable (*52*).

Our study demonstrates that poststroke treatment with anti-CD49d antibodies confers a neuroprotective benefit in a specific mouse model of stroke. We found differences between the results from two ischemic stroke models with respect to cerebral leukocyte invasion and the efficacy of anti-CD49d treatment; therefore, future clinical trials testing immunotherapeutic drugs for stroke will need to ensure that the included study population feature a substantial neuroinflammatory reaction to the brain injury, which may improve its potential to profit from such therapeutic approaches. Finally, the ability of a pRCT to advance translational research and increase the reliability of preclinical findings should be tested by comparing the results of clinical trials with the results of their preceding pRCTs.

MATERIALS AND METHODS

Study design

A key objective of this study was to implement and to test the feasibility of performing an international pRCT in experimental biomedicine. Another objective was to use this pRCT approach to robustly test the neuroprotective effect of CD49d-specific antibodies in two murine models of experimental stroke. This study was performed from June 2013 (initiation) through October 2014 (unblinding) by an international consortium consisting of six independent research groups. The infarct volume of the pooled samples from all centers and from each model was a predetermined primary end point of the study. The secondary end points were functional outcome and the migration of leukocytes into the brain. The Munich study center (A.L.) initiated the trial, coordinated the study design, and performed central data analysis, whereas the other five centers [Caen (D.V.), Milan (M.-G.D.S.), Barcelona (A.M.P.), Munich (N.P.), and Berlin (U.D. and A.R.)] contributed to study design, performed the experimental part of the study, and validated centrally determined infarct volumes. In total, 315 male C57BL/6J mice (8 to 10 weeks of age) were used [90 mice for cMCAOMCAO model (30 mice per treated group) and 225 mice for fMCAOMCAO (75 mice per treated group)]. Overall sample size per stroke model was determined a priori by performing a power calculation with G*Power (version 3.1) software using the two-tailed Wilcoxon-Mann-Whitney test for two groups. For both models, an α level of 0.05 and a power of 0.9 were required. For the cMCAOMCAO model, a Cohen's *d* ES of 0.9 was used [a conservative assumption based on our previous results demonstrating an ES *d* of 1.2 (*20*)]. In contrast, previous studies using the fMCAO model in mice did not detect a significant treatment effect that we could use to calculate ES. Therefore, we assumed an ES *d* of 0.6 to be biologically relevant, a highly conservative assumption based on our previous experience, which revealed greater variability of infarct volume after fMCAO compared to cMCAO. In addition, a 4-day mortality rate of 20% was assumed for the fMCAO group. This yielded a final sample size of 29 mice per treatment group for cMCAO and 63 mice per treatment group for fMCAO. Coded samples of anti-CD49d and control antibodies, as well as some additional items (for example, the filaments used to occlude the MCA and the behavioral test equipment), were distributed to each study center. The operational exclusion criteria were predefined (see below), and analytical exclusion criteria (see specifications below) were defined during data acquisition. The mice were allocated to the treatment and surgery groups using randomization lists generated with an online randomizer tool (www.randomizer.org) provided by the coordinating center. Treatment, surgery, analysis of infarct volume, and secondary outcome measures were performed by researchers who were blinded with respect to the treatment groups. Unblinding was performed after the statistical analyses were completed. Digital scans of the brain sections and all analysis files were stored on a central database that was accessible to all study contributors. The complete set of data is reported, including outliers. The full data set obtained from this study is publicly available at the figshare repository (http://dx.doi.org/10.6084/m9.figshare.1289824). The study protocol is available in the Supplementary Materials as fig. S6.

Animals

This study was conducted in accordance with the respective national guidelines regarding the use of experimental animals, and all procedures were approved by the respective government and institutional committees for the individual research groups [Munich: Regierungspräsidium Oberbayern; Barcelona: Ethics Committee (CEEA) of the University of Barcelona and the Departament d'Agricultura, Ramaderia, Pesca, Alimentació i Medi Natural de la Generalitat de Catalunya; Milan: Institutional guidelines and authorization by the Italian Ministry of Health; Berlin: Landesamt für Gesundheit und Soziales, Berlin; and Caen: Regional ethics committee of Lower Normandy]. Approval for the performed experiments was reached at each study site by different means, such as including the experiments in a 3-year laboratory program, as an addendum to an existing license, or applying for a new license as a "proof-of-principle" study. In total, 315 male C57BL/6J mice (8 to 10 weeks of age) were supplied to the research groups from Charles River Laboratories (stock #664). The animals were housed in a controlled temperature (22 ± 2°C) with a 12/12-hour light/dark cycle and access to pelleted food and water ad libitum. The number of mice per cage, the use of environmental enrichment, and food type were all determined individually by each research group in accordance with locally approved standard procedures. Where applicable, body weight (measured daily) and body temperature (measured on the surgery day and 1, 3, and 7 days after surgery) were measured for each mouse. All procedures regarding the study design, animal experiments, statistical analysis, and data reporting fulfill the criteria of the ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines (see checklist in the Supplementary Materials as fig. S7).

Antibody treatment

Animals received an intraperitoneal injection of antibodies 3 hours after stroke was induced (by coagulation occlusion or filament insertion); 300 μg of monoclonal mouse anti-CD49d antibody (clone R1-2, eBioscience) or rat IgG2b isotype control antibody (clone LTF-2, Bio

X Cell) was diluted to 1 mg/ml in phosphate-buffered saline (PBS) and prewarmed to 37°C before injection. The antibody concentration was chosen on the basis of information in two previous reports that used the same antibody clone in experimental stroke models (*20*, *24*).

Transient MCAO (fMCAO) model

Animals were anesthetized with isoflurane in a 30 to 70% mixture of O_2 and N_2O, and the temporal bone was exposed with an incision between the ear and the eye. A laser Doppler probe (at centers 1, 3, and 4) was affixed to the skull above the MCA territory. The animal was then placed in the supine position. Using a midline neck incision, the common carotid artery and the external carotid artery (right side, groups 1 to 3; left side, groups 4 and 5) were isolated and ligated; a 2-mm silicon-coated filament (#701912PK5Re, Doccol) was threaded into the internal carotid artery, and MCA occlusion was confirmed by a reduction in the corresponding laser Doppler flow. After 60 min of occlusion, the animals were reanesthetized, and the filament was removed. For the postsurgical survival period, the animals were housed in their home cages with access to water and food. Sham-operated control mice received the same surgical procedure, except the filament was not inserted. Body temperature was maintained during surgery using a feedback-controlled heating pad. Analgesia was provided during surgery and in the postoperative phase in accordance with each center's local regulations.

Surgical exclusion criteria: Centers 1, 3, and 4 used laser Doppler flow during the operation to confirm sufficient MCA occlusion (a reduction in blood flow to <20% of the baseline value). In addition, at all five contributing centers, mice that did not develop sufficient neurological deficits (that is, Neuroscore <8) 24 or 48 hours after MCAO were excluded. Mice that died during the observation period were excluded from all analyses, but differences in the mortality rate between the treatment groups during the observation period were determined.

Permanent MCAO (cMCAO) model

Centers 1, 2, and 3 performed the cMCAO procedure as described (*53*). In brief, the animals were anesthetized with isoflurane in a 30 to 70% mixture of O_2 and N_2O and placed in a lateral position. The skin was incised between the eye and the ear, the temporal muscle was removed, and the MCA was identified. A burr hole was drilled over the MCA, and the dura mater was removed carefully. The MCA was permanently occluded with either bipolar electrocoagulation forceps or an electrocauterizer. Permanent occlusion of the MCA was confirmed visually before closing the wound with sutures. During surgery, body temperature was maintained using a feedback-controlled heating pad. Sham-operated mice received the same surgical procedure without MCA coagulation.

Surgical exclusion criteria: Mice that developed a subarachnoid hemorrhage during surgery were excluded from the analysis. In addition, mice that died during the observation period were excluded from analysis.

Functional outcome tests

The rotarod and adhesive removal tests were performed 1, 3, and 7 days after cMCAO. These tests were chosen because they are the most commonly used tests for measuring the effect of cMCAO, which results in only minor behavioral deficits, and because these tests were used previously by all three centers that performed the cMCAO model.

The rotarod test was used to measure coordination and sensorimotor performance before and after cMCAO (*54*). Mice were trained daily for 3 days before MCAO or sham surgery, and baseline performance was recorded the day before cMCAO using the following strategy: the rod accelerated from 8 to 40 rpm over 240 s, with a mean of three consecutive trials per mouse and time points. The latency to fall off the rod was recorded, and postsurgical performance was calculated by dividing the postsurgical values by the individual animal's baseline performance.

The adhesive removal test was used to evaluate sensory and motor deficits (*55*). A round (4-mm-diameter) adhesive sticker was placed on the palmar side of the forepaw; the same pressure was applied for each adhesive application. The day before surgery (and at the indicated times after cMCAO or sham surgery), three trials per test were performed. The latency to contact the paw with the adhesive and the latency to remove the adhesive were recorded. Motor performance was expressed as the difference between the latency to contact the paw and the latency to remove the adhesive.

The Neuroscore was performed before surgery and 2 and 4 days after fMCAO or sham surgery; this test was used to evaluate the general status and focal neurologic dysfunction after fMCAO, which induces substantial deficits in contrast to the cMCAO model, and was performed as described (*56*). The score ranges from 0 (no deficits) to 56 (representing the poorest performance in all items) and is calculated as the sum of the general and focal deficits. The Neuroscore results were expressed as a composite neurological score, which included the following general deficits (scores): fur (0 to 2), ears (0 to 2), eyes (0 to 4), posture (0 to 4), spontaneous activity (0 to 4), and epileptic behavior (0 to 12); and the following focal deficits: body asymmetry (0 to 4), gait (0 to 4), climbing on a surface inclined at 45° (0 to 4), circling behavior (0 to 4), front-limb symmetry (0 to 4), circling behavior (0 to 4), and whisker response to light touch (0 to 4).

Tissue sampling and processing

Mice were deeply anesthetized 7 days after cMCAO, 4 days after fMCAO, and at the respective times for mice in the corresponding sham surgery groups. The brain was removed, frozen immediately on powdered dry ice, and stored at −20°C. All biological samples were shipped on dry ice to the central analysis center in Munich, where the samples were stored at −80°C until further analysis and for central archiving. All brains were cryosectioned by researchers who were blinded with respect to the surgery and treatment groups. Coronal brain sections (20 μm thick) were prepared at 400-μm intervals and used to analyze infarct volume. In addition, 12-μm-thick coronal sections were obtained at the level of the anterior commissure and were used for immunohistochemistry. Sections were mounted on Superfrost Plus slides (Thermo Scientific) and stored at −80°C.

Infarct volumetry

Infarct volume was measured in CV-stained sections as described (*53*). In brief, one series of serial (20-μm-thick) sections from each animal was air-dried at room temperature for 30 min. The slides were stained with a 0.1% CV solution at 60°C for 10 min and then rinsed twice in distilled water for 1 min. After washing and dehydration, the slides were covered with Roti-Histokit mounting medium (Roth). CV-stained sections were scanned at 600 dpi on a flatbed scanner (LiDE 210, Canon). For the cMCAO model, the direct infarct demarcation on CV-stained sections (unstained area) was measured after confirming the absence of edema at the investigated time point (7 days after cMCAO or sham surgery). For the fMCAO model, we corrected the infarct volume for edema with the following equation: (Ischemic area) = (Direct lesion volume) − [(Ipsilateral hemisphere) − (Contralateral hemisphere)]. In both models, total infarct volume was determined by integrating the measured areas and distances over the sections. To validate the sections in the MCAO mice with missing lesion demarcation on CV staining,

neuronal loss was measured with the terminal deoxynucleotidyl transferase–mediated deoxyuridine triphosphate nick end labeling (TUNEL) method (TUNEL Apoptosis Detection Kit, Millipore). All sections were analyzed independently by two central investigators (G.L. and A.L.), both of whom were blinded with respect to the treatment groups; the mean of these two measurements was used for subsequent analyses. In addition, the original CV-stained slides were returned to the respective originating center for further validation. Interrater reliability between the two central investigators and the individual participating centers was measured and is shown in fig. S1.

Analytical exclusion criteria: The following criteria were used to exclude samples from the analysis: (i) missing biological samples due to the death of an animal before the study's end point (day 4 for fMCAO and day 7 for cMCAO; $n = 23$ mice excluded); (ii) overt mechanical damage or disruption when preparing the brain sections, rendering the sample unusable ($n = 2$ mice excluded); (iii) no ischemic brain lesion detectable in the MCAO group (detected as demarcation on CV staining or no neuronal apoptosis measured using TUNEL staining) ($n = 27$ mice excluded); (iv) demarcation of an ischemic lesion on CV-stained sections in sham-operated mice ($n = 2$ mice excluded); and (v) Neuroscore <8 points 24 hours after fMCAO ($n = 6$ mice excluded). The individual animals that were excluded on the basis of the aforementioned criteria were also excluded from the analysis of behavioral data sets and physiological parameters.

Immunofluorescence

We performed immunofluorescence staining for CD45 to identify leukocytes that had infiltrated the brain. Coronal sections (12 μm thick) were prepared at the level of the anterior commissure and dried at room temperature for 1 hour. After rinsing in PBS (pH 7.4), the slides were fixed with acetone at −20°C, rinsed in PBS, and incubated in blocking buffer (BB) containing 0.1% Triton, 0.05% Tween 20, 10% fetal calf serum, and 1% bovine serum albumin (w/v) in PBS at room temperature for 1 hour. The slides were then incubated overnight at 4°C in anti-CD45 antibody (1:100, clone 104-2, Abcam) in BB. The slides were then rinsed in PBS and incubated for 1 hour at room temperature in Alexa Fluor 488 goat anti-mouse IgG (H+L) (1:100, Invitrogen). Finally, the slides were rinsed in PBS, counterstained for 2 min with DAPI (1:4000, Invitrogen), and mounted using Fluoromount medium (Sigma). The slides were analyzed using an epifluorescence microscope (Axiovert 200M, Zeiss). CD45-positive cells were counted in one 12-μm section per brain at the position of the anterior commissure (about 0.1 mm rostral to bregma). The infarct core was identified with corresponding CV-stained sections, and this area was excluded from the localization analysis and quantification of CD45-positive cells. For the cMCAO model, the location of each CD45-positive cell was marked on a topographic map according to the mouse brain atlas, yielding a cumulative localization map of leukocyte invasion for each treatment group.

Statistical analysis

This study was designed as a prospective, multicenter, randomized controlled trial of experimental ischemic stroke induced in mice via two distinct models. The above-mentioned operational and analytical exclusion criteria were applied. If a mouse met one or more of the exclusion criteria, data and samples were excluded from all analyses. Sufficient normal distribution of all data sets was verified by checking histograms (unimodal distribution) and skewness ($|\text{skewness}| < 1$). Physiological parameters and cerebral leukocyte counts were analyzed by ANOVA followed by Tukey's multiple-comparison test (GraphPad Prism 6.0). The characteristics of the samples obtained from the individual centers (Table 1) were tested with Fisher's exact test, the χ^2 test, or ANOVA (where indicated), and mortality was analyzed by comparing the survival curves with the log-rank (that is, Mantel-Cox) test (GraphPad Prism 6.0). Infarct volume and behavioral deficits were tested using a linear mixed-effects model to account for heterogeneity between centers (SPSS version 22, IBM). We used random intercept models with treatment group as the independent variable and infarct volume as the dependent variable. ES estimates (standardized mean difference) were calculated using inverse-variance weighting in a random-effects model using RevMan version 5.3; RevMan 5.3 was also used to create the forest plot in Fig. 2C. Correlation analysis for infarct volume and behavioral deficits was performed using linear regression (GraphPad Prism). Concordance analysis for interrater reliability was done using the Bland-Altman method (GraphPad Prism) and calculation of ICCs (SPSS). A two-sided significance level of $\alpha = 0.05$ for both primary hypotheses was applied. *P* values for secondary hypotheses should be interpreted cautiously as exploratory analyses. No adjustment for duality of the two stroke models was applied. An overview of means and 95% CI is available for all primary and secondary end points in table S2.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/7/299/299ra121/DC1
Fig. S1. Analyses for interrater concordance.
Fig. S2. Behavioral tests before and after cMCAO.
Fig. S3. Composite Neuroscore for functional deficits after fMCAO.
Fig. S4. Mortality after cMCAO and fMCAO.
Fig. S5. Physiological parameter analysis after cMCAO and fMCAO.
Fig. S6. Study protocol.
Fig. S7. The ARRIVE guidelines checklist.
Table S1. Study costs (in euro).
Table S2. fMCAO and cMCAO means and CIs.

REFERENCES AND NOTES

1. G. A. Donnan, M. Fisher, M. Macleod, S. M. Davis, Stroke. *Lancet* **371**, 1612–1623 (2008).
2. Á. Chamorro, A. Meisel, A. M. Planas, X. Urra, D. van de Beek, R. Veltkamp, The immunology of acute stroke. *Nat. Rev. Neurol.* **8**, 401–410 (2012).
3. W. Hacke, M. Kaste, E. Bluhmki, M. Brozman, A. Dávalos, D. Guidetti, V. Larrue, K. R. Lees, Z. Medeghri, T. Machnig, D. Schneider, R. von Kummer, N. Wahlgren, D. Toni; ECASS Investigators, Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. *N. Engl. J. Med.* **359**, 1317–1329 (2008).
4. E. C. Jauch, J. L. Saver, H. P. Adams Jr., A. Bruno, J. J. Connors, B. M. Demaerschalk, P. Khatri, P. W. McMullan Jr., A. I. Qureshi, K. Rosenfield, P. A. Scott, D. R. Summers, D. Z. Wang, M. Wintermark, H. Yonas; American Heart Association Stroke Council; Council on Cardiovascular Nursing; Council on Peripheral Vascular Disease, Council on Clinical Cardiology, Guidelines for the early management of patients with acute ischemic stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **44**, 870–947 (2013).
5. V. E. O'Collins, M. R. Macleod, G. A. Donnan, L. L. Horky, B. H. van der Worp, D. W. Howells, 1,026 experimental treatments in acute stroke. *Ann. Neurol.* **59**, 467–477 (2006).
6. C. Iadecola, J. Anrather, Stroke research at a crossroad: Asking the brain for directions. *Nat. Neurosci.* **14**, 1363–1368 (2011).
7. M. A. Moskowitz, Brain protection: Maybe yes, maybe no. *Stroke* **41**, S85–S86 (2010).
8. J. A. Zivin, M. Fisher, U. DeGirolami, C. C. Hemenway, J. A. Stashak, Tissue plasminogen activator reduces neurological damage after cerebral embolism. *Science* **230**, 1289–1292 (1985).
9. K. A. Hossmann, The two pathophysiologies of focal brain ischemia: Implications for translational stroke research. *J. Cereb. Blood Flow Metab.* **32**, 1310–1316 (2012).
10. M. McNutt, Journals unite for reproducibility. *Science* **346**, 679 (2014).
11. C. Kilkenny, W. J. Browne, I. C. Cuthill, M. Emerson, D. G. Altman, Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLOS Biol.* **8**, e1000412 (2010).
12. D. J. Lefer, R. Bolli, Development of an NIH consortium for preclinicAl AssESsment of CARdioprotective therapies (CAESAR): A paradigm shift in studies of infarct size limitation. *J. Cardiovasc. Pharmacol. Ther.* **16**, 332–339 (2011).

13. U. Dirnagl, A. Hakim, M. Macleod, M. Fisher, D. Howells, S. M. Alan, G. Steinberg, A. Planas, J. Boltze, S. Savitz, C. Iadecola, S. Meairs, A concerted appeal for international cooperation in preclinical stroke research. *Stroke* **44**, 1754–1760 (2013).
14. P. M. Bath, M. R. Macleod, A. R. Green, Emulating multicentre clinical stroke trials: A new paradigm for studying novel interventions in experimental models of stroke. *Int. J. Stroke* **4**, 471–479 (2009).
15. J. Kimmelman, J. S. Mogil, U. Dirnagl, Distinguishing between exploratory and confirmatory preclinical research will improve translation. *PLOS Biol.* **12**, e1001863 (2014).
16. Q. Wang, X. N. Tang, M. A. Yenari, The inflammatory response in stroke. *J. Neuroimmunol.* **184**, 53–68 (2006).
17. B. Engelhardt, L. Sorokin, The blood–brain and the blood–cerebrospinal fluid barriers: Function and dysfunction. *Semin. Immunopathol.* **31**, 497–511 (2009).
18. B. Engelhardt, R. M. Ransohoff, The ins and outs of T-lymphocyte trafficking to the CNS: Anatomical sites and molecular mechanisms. *Trends Immunol.* **26**, 485–495 (2005).
19. R. Rudick, C. Polman, D. Clifford, D. Miller, L. Steinman, Natalizumab: Bench to bedside and beyond. *JAMA Neurol.* **70**, 172–182 (2013).
20. A. Liesz, W. Zhou, É. Mracskó, S. Karcher, H. Bauer, S. Schwarting, L. Sun, D. Bruder, S. Stegemann, A. Cerwenka, C. Sommer, A. H. Dalpke, R. Veltkamp, Inhibition of lymphocyte trafficking shields the brain against deleterious neuroinflammation after stroke. *Brain* **134**, 704–720 (2011).
21. K. Becker, D. Kindrick, J. Relton, J. Harlan, R. Winn, Antibody to the α4 integrin decreases infarct size in transient focal cerebral ischemia in rats. *Stroke* **32**, 206–211 (2001).
22. J. K. Relton, K. E. Sloan, E. M. Frew, E. T. Whalley, S. P. Adams, R. R. Lobb, Inhibition of α4 integrin protects against transient focal cerebral ischemia in normotensive and hypertensive rats. *Stroke* **32**, 199–205 (2001).
23. J. Neumann, M. Riek-Burchardt, J. Herz, T. R. Doeppner, R. König, H. Hütten, E. Etemire, L. Männ, A. Klingberg, T. Fischer, M. W. Görtler, H. J. Heinze, P. Reichardt, B. Schraven, D. M. Hermann, K. G. Reymann, M. Gunzer, Very-late-antigen-4 (VLA-4)-mediated brain invasion by neutrophils leads to interactions with microglia, increased ischemic injury and impaired behavior in experimental stroke. *Acta Neuropathol.* **129**, 259–277 (2015).
24. F. Langhauser, P. Kraft, E. Göb, J. Leinweber, M. K. Schuhmann, K. Lorenz, M. Gelderblom, S. Bittner, S. G. Meuth, H. Wiendl, T. Magnus, C. Kleinschnitz, Blocking of α4 integrin does not protect from acute ischemic stroke in mice. *Stroke* **45**, 1799–1806 (2014).
25. S. T. Carmichael, Rodent models of focal stroke: Size, mechanism, and purpose. *NeuroRx* **2**, 396–409 (2005).
26. W. Zhou, A. Liesz, H. Bauer, C. Sommer, B. Lahrmann, N. Valous, N. Grabe, R. Veltkamp, Postischemic brain infiltration of leukocyte subpopulations differs among murine permanent and transient focal cerebral ischemia models. *Brain Pathol.* **23**, 34–44 (2013).
27. F. Prinz, T. Schlange, K. Asadullah, Believe it or not: How much can we rely on published data on potential drug targets? *Nat. Rev. Drug Discov.* **10**, 712 (2011).
28. C. G. Begley, L. M. Ellis, Drug development: Raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
29. S. Perrin, Preclinical research: Make mouse studies work. *Nature* **507**, 423–425 (2014).
30. G. N. Duda, D. W. Grainger, M. L. Frisk, L. Bruckner-Tuderman, A. Carr, U. Dirnagl, K. M. Einhäupl, S. Gottschalk, E. Gruskin, C. Huber, C. H. June, D. J. Mooney, E. T. Rietschel, G. Schütte, W. Seeger, M. M. Stevens, R. Urban, A. Veldman, G. Wess, H. D. Volk, Changing the mindset in life sciences toward translation: A consensus. *Sci. Transl. Med.* **6**, 264cm12 (2014).
31. M. Endres, B. Engelhardt, J. Koistinaho, O. Lindvall, S. Meairs, J. P. Mohr, A. Planas, N. Rothwell, M. Schwaninger, M. E. Schwab, D. Vivien, T. Wieloch, U. Dirnagl, Improving outcome after stroke: Overcoming the translational roadblock. *Cerebrovasc. Dis.* **25**, 268–278 (2008).
32. S. R. Bornstein, J. Licinio, Improving the efficacy of translational medicine by optimally integrating health care, academia and industry. *Nat. Med.* **17**, 1567–1569 (2011).
33. U. Dirnagl, M. Fisher, International, multicenter randomized preclinical trials in translational stroke research: It's time to act. *J. Cereb. Blood Flow Metab.* **32**, 933–935 (2012).
34. J. P. Ioannidis, S. Greenland, M. A. Hlatky, M. J. Khoury, M. R. Macleod, D. Moher, K. F. Schulz, R. Tibshirani, Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* **383**, 166–175 (2014).
35. J. P. Ioannidis, Why most published research findings are false. *PLOS Med.* **2**, e124 (2005).
36. N. A. Crossley, E. Sena, J. Goehler, J. Horn, B. van der Worp, P. M. Bath, M. Macleod, U. Dirnagl, Empirical evidence of bias in the design of experimental stroke studies: A metaepidemiologic approach. *Stroke* **39**, 929–934 (2008).
37. D. W. Howells, E. S. Sena, M. R. Macleod, Bringing rigour to translational medicine. *Nat. Rev. Neurol.* **10**, 37–43 (2014).
38. U. Dirnagl, Bench to bedside: The quest for quality in experimental stroke research. *J. Cereb. Blood Flow Metab.* **26**, 1465–1478 (2006).
39. CAMARADES (2014).
40. K. K. Tsilidis, O. A. Panagiotou, E. S. Sena, E. Aretouli, E. Evangelou, D. W. Howells, R. Al-Shahi Salman, M. R. Macleod, J. P. Ioannidis, Evaluation of excess significance bias in animal studies of neurological diseases. *PLOS Biol.* **11**, e1001609 (2013).
41. K. S. Button, J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, M. R. Munafò, Power failure: Why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14**, 365–376 (2013).
42. Enlimomab Acute Stroke Trial Investigators, Use of anti-ICAM-1 therapy in ischemic stroke: Results of the Enlimomab Acute Stroke Trial. *Neurology* **57**, 1428–1434 (2001).
43. C. Kleinschnitz, N. Schwab, P. Kraft, I. Hagedorn, A. Dreykluft, T. Schwarz, M. Austinat, B. Nieswandt, H. Wiendl, G. Stoll, Early detrimental T-cell effects in experimental cerebral ischemia are neither related to adaptive immunity nor thrombus formation. *Blood* **115**, 3835–3842 (2010).
44. T. Shichita, Y. Sugiyama, H. Ooboshi, H. Sugimori, R. Nakagawa, I. Takada, T. Iwaki, Y. Okada, M. Iida, D. J. Cua, Y. Iwakura, A. Yoshimura, Pivotal role of cerebral interleukin-17–producing γδT cells in the delayed phase of ischemic brain injury. *Nat. Med.* **15**, 946–950 (2009).
45. A. E. Baird, The forgotten lymphocyte: Immunity and stroke. *Circulation* **113**, 2035–2036 (2006).
46. G. Yilmaz, T. V. Arumugam, K. Y. Stokes, D. N. Granger, Role of T lymphocytes and interferon-γ in ischemic stroke. *Circulation* **113**, 2105–2112 (2006).
47. J. Seok, H. S. Warren, A. G. Cuenca, M. N. Mindrinos, H. V. Baker, W. Xu, D. R. Richards, G. P. McDonald-Smith, H. Gao, L. Hennessy, C. C. Finnerty, C. M. López, S. Honari, E. E. Moore, J. P. Minei, J. Cuschieri, P. E. Bankey, J. L. Johnson, J. Sperry, A. B. Nathens, T. R. Billiar, M. A. West, M. G. Jeschke, M. B. Klein, R. L. Gamelli, N. S. Gibran, B. H. Brownstein, C. Miller-Graziano, S. E. Calvano, P. H. Mason, J. P. Cobb, L. G. Rahme, S. F. Lowry, R. V. Maier, L. L. Moldawer, D. N. Herndon, R. W. Davis, W. Xiao, R. G. Tompkins; Inflammation and Host Response to Injury, Large Scale Collaborative Research Program, Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3507–3512 (2013).
48. J. Mestas, C. C. Hughes, Of mice and not men: Differences between mouse and human immunology. *J. Immunol.* **172**, 2731–2738 (2004).
49. W. R. Schabitz, U. Dirnagl, Are we ready to translate T-cell transmigration in stroke? *Stroke* **45**, 1610–1611 (2014).
50. H. X. Chu, H. A. Kim, S. Lee, J. P. Moore, C. T. Chan, A. Vinh, M. Gelderblom, T. V. Arumugam, B. R. Broughton, G. R. Drummond, C. G. Sobey, Immune cell infiltration in malignant middle cerebral artery infarction: Comparison with transient cerebral ischemia. *J. Cereb. Blood Flow Metab.* **34**, 450–459 (2014).
51. M. Balkaya, J. M. Kröber, A. Rex, M. Endres, Assessing post-stroke behavior in mouse models of focal ischemia. *J. Cereb. Blood Flow Metab.* **33**, 330–338 (2013).
52. D. W. Howells, E. S. Sena, V. O'Collins, M. R. Macleod, Improving the efficiency of the development of drugs for stroke. *Int. J. Stroke* **7**, 371–377 (2012).
53. G. Llovera, S. Roth, N. Plesnila, R. Veltkamp, A. Liesz, Modeling stroke in mice: Permanent coagulation of the distal middle cerebral artery. *J. Vis. Exp.* e51729 (2014).
54. B. J. Jones, D. J. Roberts, The quantitative measurement of motor inco-ordination in naive mice using an accelerating rotarod. *J. Pharm. Pharmacol.* **20**, 302–304 (1968).
55. V. Bouet, M. Boulouard, J. Toutain, D. Divoux, M. Bernaudin, P. Schumann-Bard, T. Freret, The adhesive removal test: A sensitive method to assess sensorimotor deficits in mice. *Nat. Protoc.* **4**, 1560–1564 (2009).
56. F. Orsini, P. Villa, S. Parrella, R. Zangari, E. R. Zanier, R. Gesuete, M. Stravalaci, S. Fumagalli, R. Ottria, J. J. Reina, A. Paladini, E. Micotti, R. Ribeiro-Viana, J. Rojo, V. I. Pavlov, G. L. Stahl, A. Bernardi, M. Gobbi, M. G. De Simoni, Targeting mannose-binding lectin confers long-lasting protection with a surprisingly wide therapeutic window in cerebral ischemia. *Circulation* **126**, 1484–1494 (2012).

Acknowledgments: We thank F. Marchesi and D. Olivari for their skilled technical assistance. **Funding:** This work was supported by the excellence cluster of the German research foundation "Munich Cluster for Systems Neurology (SyNergy)" to A.L. and N.P. and the Daimler and Benz Foundation (32-05/12) to A.L.; grants from the Spanish Ministry of Economy (MINECO; SAF2011-30492 and SAF2014-56279) to A.M.P.; grants from the INSERM (French National Institute for Health and Medical Research), the University of Caen Basse-Normandie, and the GIS IBiSA to D.V.; the Italian Ministry of Health Young Investigators Award #GR-2008-1136044 to E.R.Z.; and grants from Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) (01EO0801, Center for Stroke Research Berlin) and the Hermann and Lilly Schilling Foundation to U.D. **Author contributions:** A.L. initiated this trial; A.L., U.D., A.M.P., N.P., M.-G.D.S., and D.V. designed the study protocol, supervised the study, and reviewed original data sets; G.L. performed the central analysis of all biological samples and data sets; V.A., C.F., M.F.-F., B.H., K.H., G.L., U.M., C.O., H.P., C.P., A.R., S.R., A.S.-P., and E.R.Z. performed experiments; U.G. performed central statistical analysis; and U.D., A.L., G.L., A.M.P., N.P., M.-G.D.S., and D.V. wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The full data set obtained from this study is publicly available at the figshare repository (http://dx.doi.org/10.6084/m9.figshare.1289824).

Submitted 24 February 2015
Accepted 23 June 2015
Published 5 August 2015
10.1126/scitranslmed.aaa9853

Citation: G. Llovera, K. Hofmann, S. Roth, A. Salas-Pérdomo, M. Ferrer-Ferrer, C. Perego, E. R. Zanier, U. Mamrak, A. Rex, H. Party, V. Agin, C. Fauchon, C. Orset, B. Haelewyn, M.-G. De Simoni, U. Dirnagl, U. Grittner, A. M. Planas, N. Plesnila, D. Vivien, A. Liesz, Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia. *Sci. Transl. Med.* **7**, 299ra121 (2015).

Science Translational Medicine
AAAS

Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia

Gemma Llovera, Kerstin Hofmann, Stefan Roth, Angelica Salas-Pérdomo, Maura Ferrer-Ferrer, Carlo Perego, Elisa R. Zanier, Uta Mamrak, Andre Rex, Hélène Party, Véronique Agin, Claudine Fauchon, Cyrille Orset, Benoît Haelewyn, Maria-Grazia De Simoni, Ulrich Dirnagl, Ulrike Grittner, Anna M. Planas, Nikolaus Plesnila, Denis Vivien and Arthur Liesz (August 5, 2015)
Science Translational Medicine **7** (299), 299ra121. [doi: 10.1126/scitranslmed.aaa9853]

Editor's Summary

Tested, just like a human

Over the last few decades, clinical trial design and analysis have become increasingly stringent. These refinements––designed to ensure valid conclusions for formal drug approval––have improved clinical trial reliability. Now, Llovera *et al*. have applied the principles of the gold-standard randomized, controlled clinical trial to a preclinical investigation. They tested an antibody to CD49d, which inhibits leukocyte migration into the brain, in two mouse models of stroke. Data from their six-center, preclinical, randomized controlled trial in mice show that the antibody significantly reduced both leukocyte invasion and infarct volume after a small cortical stroke but that it did not have any effect in the other model, in which the animal suffered a larger injury. The authors outline the many lessons learned from their experience for further application of preclinical randomized controlled trials to translational research.

Science Translational Medicine (print ISSN 1946-6234; online ISSN 1946-6242) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005.

STROKE

Neuroprotective therapies: Preclinical reproducibility is only part of the problem

Michael Tymianski[1,2,3,4]

Among the many unknowns in the translational path to developing drugs for acute stroke, addressing the reproducibility of preclinical data is only one piece of a multifaceted and incomplete puzzle (Llovera *et al.*, this issue).

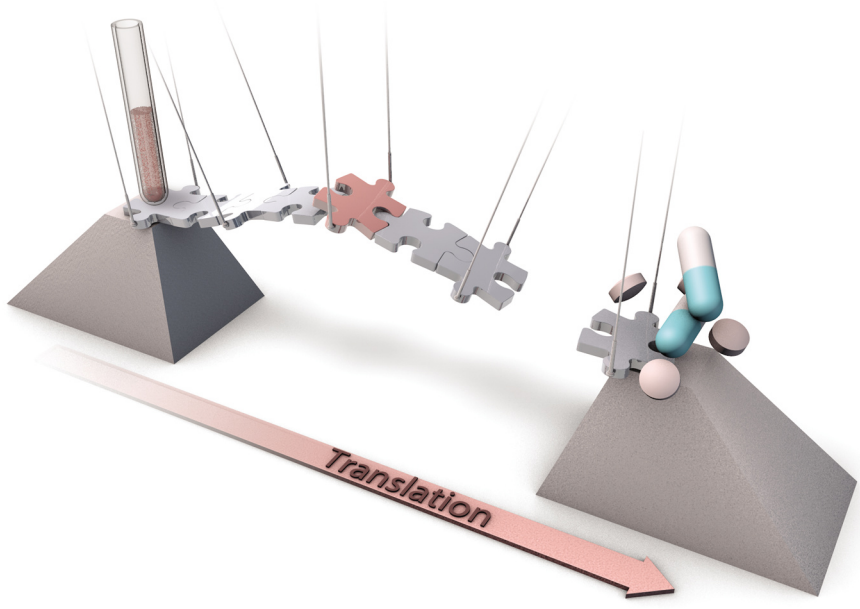
At this moment in time, there is no conclusive evidence that neuroprotection can be of clinical benefit to humans suffering from acute stroke. Despite the absence of such evidence, stroke neuroprotection research continues on the basis of the belief—derived from decades of preclinical data in which multiple neuroprotectants were tested with promising results—that clinically beneficial neuroprotection is possible in people. This unsettling situation leaves the field without an evidence-based roadmap with which to conduct translational research efficiently and with a high chance of success. In this issue of *Science Translational Medicine*, Llovera *et al.* (*1*) investigate one preclinical portion of the translational path that has puzzled stroke researchers—verifying the efficacy of a neuroprotectant.

The two key unanswered questions about neuroprotection in humans are (i) "Is it possible?" and (ii) "Is it practicable?" Both must be answered before a proposed treatment can be proven to be clinically beneficial. The former is a matter of biology, assessing whether the therapeutic has the capability to reduce ischemic brain damage. The latter is a matter of implementation—whether it is feasible to test the therapy in a clinical trial capable of showing a meaningful clinical benefit.

Answers to the first question require an understanding of pathophysiology and possible therapeutic targets, the mechanism of action of an agent, and the suitability of experimental stroke models. This research must be followed by exploratory preclinical studies of a therapeutic effect and then verification of this effect in larger, more definitive studies, such as those suggested by Llovera *et al.* The second question, that of practicability, cannot be addressed until the first is resolved. Once researchers settle on a biologically efficacious neuroprotectant, the drug must be tested in preclinical safety and pharmacology studies conducted to regulatory standards and manufactured at high purity, potency, and stability. This work is then followed by the demonstration of clinical safety (typically in a phase 1 trial) and the design and conduct of phase 2 and 3 trials that meet regulatory guidelines. Although this process applies to all therapeutics, it has proven to be particularly difficult for acute stroke neuroprotectants. The past failures to translate neuroprotectants to clinical utility have triggered a "nuclear winter" of pessimism, limited funding, and near-absent interest from pharmaceutical companies that previously were partners in at least the latter aspects of this journey. Thus, the future depends on academic researchers to chart the roadmap from bench to clinic.

In the new work, Llovera *et al.* (*1*) tackle a critical aspect of the translational path of a stroke neuroprotectant: the verifiability of biological efficacy. They propose a solution to the "reproducibility crisis" currently plaguing preclinical data in the life sciences (*2*–*4*) by conducting an animal study designed with standards analogous to those of a randomized, placebo-controlled multicenter clinical trial. In essence, their paper demonstrates the feasibility of conducting such a study at several individual laboratories while coordinating it centrally, and generating data that can be pooled and analyzed under a single protocol. In doing so, they aim to overcome several causes of poor reproducibility, including small sample size, uninformed data handling, a pressure to publish, bias against reporting negative results, and deficiencies in methods reporting (*4*).

However, over the past six decades, more than 1000 experimental stroke treatments have failed to be translated to clinical utility (*5*), and it is difficult to believe that all of these failures resulted only from irreproducible or even false results. Rather, additional factors might be responsible for the



Puzzling over neuroprotection. A series of complex quandaries have marred the translational path of stroke therapies—from basic research into disease biology and target discovery; to target validation and preclinical safety, e□cacy, and reproducibility studies; to human clinical trials. These pieces must be aligned to successfully develop a medication for for patients su□ering from acute stroke.

CREDIT: C. BICKEL/*SCIENCE TRANSLATIONAL MEDICINE*

[1]Toronto Western Hospital Research Institute, Toronto, Ontario M5T 2S8, Canada. [2]Department of Physiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [3]Institute of Medical Science, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [4]Department of Surgery, University of Toronto, Toronto, Ontario M5S 1A8, Canada.
E-mail: mike.tymianski@uhn.ca

inability to translate seemingly promising preclinical stroke therapies. Therefore, like a puzzle with an unknown number of pieces, a translational roadmap for stroke must be constructed systematically by using the best tools that science has to offer, including adherence to rigorous standards of scientific inquiry, and minimizing wasted time and resources. In this context, the new work (*1*) addresses these important needs but constitutes only one piece of the puzzle whose scope has not yet been fully defined.

ARE WE BARKING UP THE RIGHT TREE?

It seems self-evident that a drug acts on a target, whether or not it is known, and that the target's pathobiology should be understood as it pertains to the disease being treated. Nevertheless, many stroke targets have been identified over decades of research, but most were advanced to preclinical testing of efficacy before gaining a comprehensive understanding of the target biology. Not surprisingly, target validity is questionable in many such cases, and this gap may contribute to the inability to reproduce research results in patients. In the case of antibodies to the α chain of the leukocyte integrin VLA-4 [antibodies to CD49d, such as the one used by Llovera *et al.* (*1*)], the target is seemingly clear. However, the rationale for inhibiting the α chain of VLA-4 in human stroke is predicated on whether leukocyte invasion of the brain—that is, inflammation—is causally related to the tissue injury of stroke and to the accompanying functional disability in humans and whether this injury can be reduced with anti-CD49d therapy.

Such an approach also presupposes that in stroke victims, there is still enough brain left to save by the time the inflammatory response kicks in. Although available preclinical literature links stroke-induced inflammation to the progression of damage, it is less clear whether limiting the inflammatory response—which peaks later after stroke onset than do other injurious pathways—is both necessary and sufficient to improve stroke outcome in a clinical setting. By the time destructive inflammation occurs, various other pathways that also lead to cell and tissue destruction have been triggered, and the inflammatory response occurs after other "white elephants," such as excitotoxicity, have entered the room. The impact of various concomitant pathways might be large enough to explain much of the initial damage of stroke, in which case treating inflammation could be moot. A further confounder is that the relative contributions of various stroke mechanisms to ischemic brain damage might differ among species, and it remains unclear how much inflammation contributes to the clinical stroke phenotype in humans. If the biology remains controversial, then even the most rigorous preclinical mouse study, including that by Llovera and colleagues, risks producing results that may not be translatable.

The gap between target identification and validation, which often results from an insufficient understanding of the underlying disease biology, is not particular to stroke. Target validation projects, typically conducted by pharmaceutical companies to increase confidence in the biology of the targets before embarking on expensive clinical development programs, frequently fail to reproduce research findings.

For example, attempts to replicate data published in leading cancer journals failed up to 89% of the time (*3*). As stated by Prinz (*2*), "the reproducibility of published data did not significantly correlate with journal impact factors, the number of publications on the respective target, or the number of independent groups that authored the publications." In many instances, preclinical research publications generate many secondary publications, spawning entire directions of research even when results cannot be reproduced (*3*). Bridging the gap between target identification and validation that includes an accurate understanding of disease and target biology is thus an essential initial step that should precede extensive animal studies, such as the one reported by Llovera *et al.* (*1*). This sequence of events should increase the efficiency of translation, reduce the volume of contradictory publications about the value of a given therapeutic, and lessen the high failure rates observed in subsequent development.

LEAPS OF FAITH ACROSS THE EVOLUTIONARY GAP

Llovera and colleagues (*1*) conducted their study in mice, a popular species in stroke research owing to its capacity for facile genetic manipulation. However, when a study does not capitalize on genetic modifications to test a biological question, mouse stroke models might not offer any translational value over existing models in other species, such as rats, in which the stroke models are well established. Indeed, the authors found "unexpectedly high variability" in the filamentous middle cerebral artery occlusion (fMCAO) mouse model (*1*). According to Carmichael (*6*), "The exquisite sensitivity of the mouse to infarct extension over minutes of MCAO may account for the variation in infarct size in this model: Published studies of infarct size within the same strain, using the same duration of MCAO and the same survival period, produce measures of infarct volume that range over a fivefold difference." Also, MCAO in the mouse of 60 min or more might produce a malignant stroke (*6*) that is more difficult to neuroprotect. In (*1*), poststroke treatment with antibodies to CD49d conferred a neuroprotective benefit in one, but not another, mouse model of stroke. Thus, it is possible that the choice of species, one so prone to variability, prevented even this comprehensive multicenter study from drawing definitive conclusions about the suitability of antibodies to CD49d for further translation.

The value of a preclinical model for predicting clinical success cannot be validated until a clinical success occurs. Until then, no model can claim to validly represent the human situation. The possibility that preclinical animal stroke studies might not predict efficacy in humans is now accepted. However, the corollary that predictions of efficacy based on animal data represent a leap of faith often elicits strong emotional responses. Those who advocate one preclinical model over another argue in favor of the close similarities in brain biology between mice and men or lay the blame on the poor methodological quality of past studies rather than on the model system itself. Although all perspectives have merit, they remain unvalidated, and thus the knowledge gap must be navigated cautiously—especially so if animal studies produce negative results. It is tempting to discontinue development of stroke therapies that do not work in one model or another, often without a thorough appraisal of the scientific rigor of the study. However, it is poor science to do so. If a failed agent has a clearly understood and validated biological target and a validated mechanism of action, more may be gained from analyzing the failure than from moving to the next agent.

For example, failure may result from technical error, inadequate dosing, unfavorable pharmacokinetics, malignant strokes that cannot be fixed, or interspecies differences. The efficacy of anti-CD49d treatment in one, but not in a second, mouse stroke model (*1*) occurred in a test system that was not designed to interpret such a result.

Although the reasons for this discrepancy remain unclear, understanding them might advance the field more than would arbitrarily abandoning the project and proceeding with a multicenter preclinical study of a different therapy. The preclinical research environment should encourage not only the funding and reporting of properly controlled negative studies but also the funding of failure analysis—a process that is common in manufacturing and various engineering industries but largely absent from the biomedical literature. Failure analysis might prevent the repetition of previous mistakes, making the process of translation more streamlined, less expensive, and more likely to succeed.

HINDSIGHT IS NOT ALWAYS 20/20

In cynomolgus macaques, stroke damage after MCAO reaches a plateau after 3 to 4 hours (*7*). Evidence from human studies similarly suggests that on average, the time during which treatment initiation can be beneficial is short, on the order of 3 hours and certainly no longer than 4.5 hours after symptom onset (*7*). No drug can be clinically beneficial if it is administered too late. Yet, virtually every major in-hospital clinical trial of a neuroprotectant ever conducted enrolled patients in a window of 4 or more hours after stroke symptoms began—and often much longer. The overwhelming preclinical evidence from experiments in rats, cats, and primates supports neuroprotection as a viable strategy when treatment is administered less than 3 hours after the experimental stroke is initiated (*7*). In contrast, there is a relative paucity of solid preclinical evidence to support neuroprotection beyond this window, especially in higher species.

Although the mantra "time is brain" resounds at every stroke treatment center, the design of past clinical trials, especially insofar as they pertain to the therapeutic window, has rarely heeded the critical need to be aligned with that of the preclinical studies that preceded it. One exception was a preclinical study in which cynomolgus macaques were subjected to small embolic strokes induced by intracarotid injections of polystyrene microemboli, effectively producing a permanent brain arterial occlusion. Administration of the neuroprotectant NA-1, a PSD95 inhibitor, at 1 hour after stroke onset reduced the number and volume of the resulting infarcts (*8*). These results were used to design the human ENACT trial (clinicaltrials.gov NCT00728182) in which NA-1 was given to patients undergoing brain endovascular procedures in an average window of 2 hours after the onset of ischemia; this trial revealed promising results in reducing stroke burden (*9*). ENACT provides evidence that bridging the gap in study conduct by aligning the designs of preclinical and clinical studies can lead to scientific progress.

USE OF FORESIGHT: A LESSON FROM INDUSTRY

A tool often used by industry is the Target Product Profile (TPP). This key strategic document provides a summary of the product under development, its desired characteristics, the studies and activities that must be completed to demonstrate the product's desirable features (including safety and efficacy), and a conceptualization of how the product would be used once it is clinically approved. According to the U.S. Food and Drug Administration (FDA), "The TPP embodies the notion of beginning with the goal in mind" (*10*). All too often, a neuroprotectant that works in a laboratory setting—for example, in mice—does not possess the characteristics that can carry it through to the clinic. For example, each of the following might impose substantial impediments to translation even if the agent is verified to reduce stroke damage in a preclinical model: Lack of target conservation in humans, incompatibility with tissue plasminogen activator (tPA) treatment, lack of adequate bioanalytical assays to detect drug levels, unfavorable pharmacokinetics, an onerous dosing regimen, extensive contraindications, known reactions with commonly taken drugs, a drug that is difficult or expensive to manufacture or that has limited stability, or lack of adequate patent protection. Such impediments may not be deal killers individually, but they should be considered before embarking on an extensive development program, including definitive preclinical efficacy studies. Academic research may be streamlined by bridging the gap between the hypothesis-driven research that typically drives scientific progress and the practical requirements of translational science more often found in industry.

The approach of conducting multicenter preclinical trials of a therapy sets a new standard for life sciences research, which is currently plagued by irreproducibility of a substantial proportion of preclinical data (*2*–*5*). It is an important piece of the larger unsolved translational puzzle in stroke research, but there are other key pieces that must be defined before the roadmap to success is clear. Whether this new standard should be a requirement or be regarded as just a bonus to a research program can only be determined in hindsight, once we can look back on the roadmap to success.

REFERENCES AND NOTES

1. G. Llovera, K. Hoffmann, S. Roth, A. Salas-Pérdomo, M. Ferrer-Ferrer, C. Perego, E. R. Zanier, U. Marmak, A. Rex, H. Party, V. Agin, C. Fauchon, C. Orset, B. Haelewyn, M.-G. De Simoni, U. Dirnagl, U. Grittner, A. M. Planas, N. Plesnila, D. Vivien, A. Liesz, Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia. *Sci. Transl. Med.* **7**, 299ra121 (2015).
2. F. Prinz, T. Schlange, K. Asadullah, Believe it or not: How much can we rely on published data on potential drug targets? *Nat. Rev. Drug Discov.* **10**, 712 (2011).
3. C. G. Begley, L. M. Ellis, Drug development: Raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
4. S. C. Landis, S. G. Amara, K. Asadullah, C. P. Austin, R. Blumenstein, E. W. Bradley, R. G. Crystal, R. B. Darnell, R. J. Ferrante, H. Fillit, R. Finkelstein, M. Fisher, H. E. Gendelman, R. M. Golub, J. L. Goudreau, R. A. Gross, A. K. Gubitz, S. E. Hesterlee, D. W. Howells, J. Huguenard, K. Kelner, W. Koroshetz, D. Krainc, S. E. Lazic, M. S. Levine, M. R. Macleod, J. M. McCall, R. T. Moxley 3rd, K. Narasimhan, L. J. Noble, S. Perrin, J. D. Porter, O. Steward, E. Unger, U. Utz, S. D. Silberberg, A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* **490**, 187–191 (2012).
5. V. E. O'Collins, M. R. Macleod, G. A. Donnan, L. L. Horky, B. H. van der Worp, D. W. Howells, 1,026 experimental treatments in acute stroke. *Ann. Neurol.* **59**, 467–477 (2006).
6. S. T. Carmichael, Rodent models of focal stroke: Size, mechanism, and purpose. *NeuroRx* **2**, 396–409 (2005).
7. M. Tymianski, Novel approaches to neuroprotection trials in acute ischemic stroke. *Stroke* **44**, 2942–2950 (2013).
8. D. J. Cook, L. Teves, M. Tymianski, A translational paradigm for the preclinical evaluation of the stroke neuroprotectant Tat-NR2B9c in gyrencephalic nonhuman primates. *Sci. Transl. Med.* **4**, 154ra133 (2012).
9. M. D. Hill, R. H. Martin, D. Mikulis, J. H. Wong, F. L. Silver, K. G. Terbrugge, G. Milot, W. M. Clark, R. L. Macdonald, M. E. Kelly, M. Boulton, I. Fleetwood, C. McDougall, T. Gunnarsson, M. Chow, C. Lum, R. Dodd, J. Poublanc, T. Krings, A. M. Demchuk, M. Goyal, R. Anderson, J. Bishop, D. Garman, M. Tymianski, ENACT trial investigators, Safety and efficacy of NA-1 in patients with iatrogenic stroke after endovascular aneurysm repair (ENACT): A phase 2, randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* **11**, 942–950 (2012).
10. U.S. Department of Health and Human Services, FDA, *Guidance for Industry and Review Staff. Target Product Profile—A Strategic Development Process Tool* (2007); available at www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm080593.pdf

Acknowledgments: M.T. is a Canada Research Chair (Tier 1) in Translational Stroke Research. **Competing financial interest:** M.T. is president and CEO of NoNO Inc., a biotechnology company founded to translate PSD95 inhibitors discovered in his academic laboratory to human trials.

10.1126/scitranslmed.aac9412

Citation: M. Tymianski, Neuroprotective therapies: Preclinical reproducibility is only part of the problem. *Sci. Transl. Med.* **7**, 299fs32 (2015).

Neuroprotective therapies: Preclinical reproducibility is only part of the problem

Michael Tymianski (August 5, 2015)

Science Translational Medicine **7** (299), 299fs32. [doi: 10.1126/scitranslmed.aac9412]

Editor's Summary