Edaravone improves functional and structural outcomes in animal models of focal cerebral ischemia: A systematic review

Simiao Wu1,2, Emily Sena1,3, Kieren Egan1, Malcolm Macleod1, and Gillian Mead1

Edaravone has been used in patients with acute ischemic stroke in Japan for over 10 years but does not have marketing authorization in Europe or America. Either patients in Europe and America are not receiving an effective treatment, or those in Asia are being given a treatment which is not effective. Finding out which of these is true will require further clinical trials, and a better understanding of its efficacy in animal models may help inform the design of those trials so that it might be tested under conditions where there is the greatest prospect of success. We systematically reviewed the efficacy of edaravone in animal models of focal ischemia and summarized data using weighted mean difference DerSimonian and Laird random-effects modeling. We used stratified meta-analysis and metaregression to assess the influence of study design and methodological quality. We identified 49 experiments describing outcome in 814 animals; 30 experiments (519 animals) reported functional and 35 experiments (503 animals) reported structural outcome. Edaravone improved functional and structural outcome by 30·3% (95% confidence interval 23·4–37·2%) and 25·5% (95% confidence interval, 21·1–29·9%), respectively. For functional outcome, there was an inverse relationship between study quality and effect size (P < 0·0017). Effect sizes were larger in studies where randomization or blinded assessment was not reported. There was no evidence of publication bias. Edaravone is a promising treatment for stroke. However, because of the methodological weakness in current animal studies, no sufficient preclinical evidence is available to optimize the study design of clinical trials. Higher quality animal studies are expected to inform further clinical study.

Key words: acute stroke therapy, cerebral infarction, ischemic stroke, neuroprotection, oxygen toxicity, treatment

Introduction

Acute ischemic stroke is a major cause of mortality and morbidity worldwide (1), and although thousands of candidate drugs have been tested in animal models of stroke, few have been demonstrated to be effective in humans (2). Thus, the development of new treatments for stroke remains challenging. In contrast to thrombolytic treatment which aims to restore blood flow, it is proposed that neuroprotective agents might prevent neuronal death in the face of an ischemic injury by interfering with pathological cellular processes. One group of putative neuroprotective compounds are antioxidants, which reduce oxidative stress either by inhibiting free radical production or by enhancing their elimination.

Edaravone is a free radical scavenger which has protective effects on neurons and cerebrovascular endothelial cells in vitro; it is thought that this occurs by protecting cell membranes against oxidative stress (3–6). Edaravone also has beneficial effects in animal models of several neurodegenerative diseases characterized by oxidative stress including neonatal hypoxic-ischemic encephalopathy (7), intracerebral (8) or subarachnoid hemorrhage (9), and traumatic brain injury (10). A benefit in animal models of focal cerebral ischemia was first reported in Japan (11–13), where following clinical testing (14) edaravone was approved in 2001 as a neuroprotective agent for acute ischemic stroke. Subsequent phase III clinical trials were conducted (15,16). A Cochrane systematic review (17) included data from three trials (496 patients) and reported a pooled risk ratio of 1·99 [95% confidence interval (CI), 1·60–2·49] for improvement in neurological status but noted a moderate risk of bias in included studies and small numbers of patients included; it concluded that further clinical trials were required. The clinical use of edaravone appears largely to be restricted to Asian countries (18).

Therefore, although edaravone is a promising drug for stroke, further clinical trials are needed. Given that there are animal data which might inform decisions about time windows for treatment or efficacy in the face of comorbidities, a robust and systematic summary of those data may assist in the design of those clinical trials. We have therefore investigated the efficacy of edaravone in animal models of ischemic stroke and explored the impact of study design (including dose of drug) and study quality on reported outcome.

Methods

Search strategy and study selection

In December 2011, we searched three electronic databases (Pubmed: U.S. National Library of Medicine, Bethesda, MD, USA Embase: Elsevier, Amsterdam, NL ISI Web of Science: Thomson Reuters, New York, USA) using keywords [(cerebral ischemia) OR (ischemic stroke)] AND [edaravone OR MCI-186 OR (MCI 186) OR radicut OR 3-methyl-1-phenyl-2-pyrazolin-5-one] with the search limited to ‘Animals’. One investigator (S. W.) screened all titles and available abstracts for eligibility. Full texts of potentially eligible studies were obtained, and two investigators (S. W., K. E.) independently applied inclusion and exclusion criteria to each paper; disagreement was resolved by either discussion or arbitration by a third investigator (E. S.).
Inclusion and exclusion criteria
We included controlled studies describing the effect of edaravone in animal models of focal cerebral ischemia where outcome was reported as a change in structural (infarct volume) or functional (neurobehavioral score) outcome. Publications in any language, published in either full text or abstract, using temporary or permanent occlusion, or thromboembolic models of ischemia and involving any species, age, and gender were eligible for inclusion. Studies with fewer than three animals in a treatment or control group were excluded.

Data extraction
Outcome measures, data on characteristics of study design, and study quality were extracted and recorded in the Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Stroke (CAMARADES) database. One investigator (S. W.) extracted the data and entered them into the database, and another investigator (K. E.) checked these data against the original publications. Values of graphical data were requested from the authors. When these data were not available, we used a digital ruler software to measure them from the graphs. We contacted authors to obtain missing or unpublished data.

We recorded the number of animals in each group. For each treatment comparison (a given dose, route, and timing of edaravone treatment with outcome compared with that in an appropriate control group), we extracted mean values and standard error or standard deviation of neurobehavioral scores (at the last time of assessment) or infarct volume. In experiments where outcome data for sham animals had not been reported, we inferred sham neurobehavioral scores for unimpaired animals and sham infarct volumes of zero.

We extracted data for study design including species, gender, age and body weight, types of ischemic models, duration of ischemia, drug regimen (dose, route, time of administration), time-point of outcome assessment, and methods of measurement. The total dose of edaravone (mg) in the first 24 h was standardized to units of mg/(kg.24 h). We recorded the number of CAMARADES quality checklist items scored as described previously (19).

Statistical analysis
To account for different brain sizes and neurobehavioral scales, we calculated a normalized mean difference effect size, being the improvement in treatment group expressed as a percentage of the observed outcome in the control group. The standard error for each comparison was similarly expressed. We used DerSimonian and Laird random-effects meta-analysis (20) to derive pooled estimates of efficacy. As our primary analysis, we prespecified subgroup analyses for categorical variables of study design (use of animals with comorbidities or not, types of ischemia, and dose of drug) and study quality (randomization, blinded assessment, and the number of quality checklist items scored), used the chi-squared statistic (χ²) to assess the significance of differences between studies [with n-1 degrees of freedom (df)], and used the Bonferroni correction to calculate significance levels adjusted for multiple testing (giving critical P values of 0.0017 for neurobehavioral outcome and 0.0015 for infarct volume). For continuous variables (doses of edaravone, duration of ischemia, time of administration, and time of assessment), we also used the more conservative metaregression as a secondary analysis using STATA/IC 10 (StataCorp, College Station, Texas, USA) with linear function metareg. We used funnel plotting (21), Egger regression (22), and ‘trim and fill’ (23) to seek evidence of publication bias.

The study protocol is available at http://www.camarades.info/index_files/protocols.html.

Results
We retrieved 328 citations from the electronic search (Fig. 1). Of these, 30 met our inclusion criteria, but of these, seven were excluded at the stage of data extraction because they reported qualitative data only (24) or did not report the number of animals (25–30). In response to reviewer’s comments, we added additional drug terms ([BRN 0609575] OR Norphenazone OR UNII-S798V6YJR OR (Bi Cun) OR (Yi Da Sheng) OR (You min) OR (Edvo) OR (Nuravon]) in a search conducted in February 2013 and limited to papers published before December 2011. This identified nine publications, of which eight had been identified in the origin search. The other paper did not report the number of animals used. For all excluded studies, we sought further information from the authors, but this was not forthcoming. This analysis is therefore based on 23 publications describing 30 comparisons reporting neurobehavioral scores and 35 comparisons reporting infarct volume.

Study characteristics
Temporary ischemia was the most frequently used animal model of ischemic stroke (20 studies, 889 animals), induced by middle cerebral artery occlusion and reperfusion. The median number of animals receiving edaravone in the treatment and control groups was 10 (range 4–14) and nine (range 4–16), respectively. Eight different neurobehavioral tests were used, most commonly the Longa 5-point scale (31) (five studies, 143 animals) and the Bederson 4-point scale (32) (four studies, 92 animals). Infarct volume was most commonly assessed using 2, 3, 5-triphenyltetrazolium chloride staining (14 studies, 438 animals).
The effects of edaravone on neurobehavioral outcome. Weighted mean difference meta-analysis of the effects of edaravone on neurobehavioral outcome in controlled animal studies (n = 30). Horizontal error bars represent the 95% CI of efficacy in individual comparison; vertical gray bar represents the 95% CI of the pooled estimate of efficacy; symbol size represents the log of the number of animals in that comparison.

**Risk of bias**

The median number of study quality checklist items scored was three of a possible 10 [interquartile range (IQR), 3–4]. Of 23 studies, 13 reported random allocation to group, and only four reported the blinded assessment of outcome. No study reported a sample size calculation or the age of animals used. Only one used spontaneous hypertensive rat models (33); there were no reports using animals with other comorbidities. Two comparisons used ketamine (34), an anesthetic agent with intrinsic neuroprotective properties, and these reported the largest effect sizes (90.4% and 90.9%) for functional outcome.

**Functional outcome**

Overall, edaravone improved neurobehavioral scores by 30.3% (95% CI, 23.4–37.2%, 30 experiments, 503 animals). There was a significant between-study heterogeneity ($\chi^2 = 138.5$, df = 29, $P < 0.0001$) (Fig. 2).

The median number of study quality checklist items scored for neurobehavioral studies was 4 (IQR 3–4). Significant differences between high and low quality studies were observed ($P < 0.0017$), with the highest quality studies reporting the lowest efficacy (13.2%, 95% CI 3.9–22.5%; Fig. 3a). Studies which did not report randomization gave higher estimates of efficacy (37.0%; 95% CI, 13.2%, 95% CI 3.9–22.5%; Fig. 3a). Studies which did not report the blinded assessment of outcome. No study reported a sample size calculation or the age of animals used. Only one used spontaneous hypertensive rat models (33); there were no reports using animals with other comorbidities. Two comparisons used ketamine (34), an anesthetic agent with intrinsic neuroprotective properties, and these reported the largest effect sizes (90.4% and 90.9%) for functional outcome.

**Impact of study design on functional outcome**

In prespecified subgroup analyses, there was no apparent impact of comorbidities, of different types of ischemic injury, or of the use of single, multiple, or continuous dosing. The first-day doses used in the included studies ranged from 0.2 to 60 mg/(kg.24 h), and 3 mg/kg was the most commonly used dose. Similarly, metaregression showed no apparent impact of the dose of edaravone, the duration of ischemia, or the time of assessment.

**Effect of time of administration on functional outcome**

We categorized studies into those which gave edaravone before the induction of ischemia, after ischemia but before reperfusion, and after reperfusion. This stratification accounted for a significant proportion of the observed heterogeneity ($\chi^2 = 17.3$, $P < 0.0017$), with efficacy being highest when edaravone was administered after ischemia and before reperfusion and lowest when edaravone was delivered before ischemia (Fig. 4). However, using metaregression, there was no significant association between time to treatment and efficacy.

**Structural outcome**

Edaravone improved infarct volume by 25.5% (95% CI, 21.1–29.9%, 35 experiments, 503 animals), with substantial heterogeneity between studies ($\chi^2 = 175.1$, df = 34, $P < 0.0015$; Fig. 5).

The median number of quality items scored was 3.5 (IQR 2.3–4.0). Overall, lower quality studies reported higher efficacy ($P < 0.0015$; Fig. 3b). Studies which reported random allocation to group gave lower estimates of efficacy (22.3%; 95% CI 17.93–26.6%) than those that did not (30.6%; 21.7–39.5%). There was no significant difference in efficacy for studies of infarct volume in relation to blinded assessment ($P > 0.20$). Although the highest efficacy was reported in the highest quality grouping, this stratum consisted of a single, unblinded study and gave multiple drug doses. Indeed, taking all studies together, efficacy was highest when edaravone was administered multiple times (33.3%; 95% CI, 4.8–24.0%), compared with studies where it was given as a single dose or in continuous administration ($P < 0.0015$). In stratified meta-analysis and metaregression, there was no significant effect of the type of ischemic model, the use of animals with comorbidities, or timing of administration.

**Publication bias**

There was no evidence of publication bias in either neurobehavioral scores or infarct volume using Funnel plotting (Fig. 6), Egger regression, or ‘trim and fill’ analysis.

**Discussion**

In this systematic review and meta-analysis of edaravone in experimental models of stroke, we found improvement in both functional outcome and structural outcome. This is consistent with the effect observed in human studies and supports the case for further clinical trials.

Our study benefits from having a comprehensive search strategy, prespecified study protocol, and from using a well-established statistical approach. Publications in all languages were included, and two authors independently applied inclusion and exclusion criteria to improve study selection. We assessed both functional and structural outcomes, consistent with recommendations for animal stroke modeling (35), and we explored the impact of study...
design and methodological quality when interpreting the results. Therefore, this minimally biased systematic review and meta-analysis provides important evidence to inform the further development of edaravone, and supplements the currently limited experimental and clinical data.

**Time of administration**

The stage at which edaravone was administered (before, during or after brain artery occlusion) explained a significant proportion of the observed heterogeneity, but the estimates of efficacy for each of these three stages were not substantially different, and there was no effect of the timing of treatment (in relation to the onset of ischemia) in metaregression. Further, these experiments differed in other respects additional to the stage of administration. All estimates of timing of administration have been spurious; it is therefore not possible to draw any conclusion about the optimum timing of treatments, and this is an important area for further research.

**Our previous review of the free radical scavenger tirilazad (36)** suggested maximal efficacy when drug was given before ischemia. Although pretreatment might be expected to improve drug delivery (as the cerebral vasculature is intact), the combination of the short half life of edaravone (less than five-minutes (30)) and the observation that the concentration of superoxide reaches its maximum at the early phase of reperfusion and subsides in a few hours (37) may explain this pattern of efficacy. In mice, the blood concentration of edaravone drops by 95% within 30 mins (38), and treatment might be most beneficial when given shortly before reperfusion.

From included studies, we found insufficient data to draw any meaningful conclusion about the efficacy of the combination of edaravone with thrombolysis/reperfusion therapy (tPA), although Zhang et al. (39) reported that edaravone given every 90 mins between ischemia onset and reperfusion resulted in increased survival, reduced infarct volume, and extended the therapeutic time window for tPA. Indeed, a meta-analysis of...
combination therapy in animals (40) suggests that these might extent the time window for tPA by around three-hours. This should be tested further because it is possible that a treatment known to be safe might be given at the earliest opportunity (for instance in the field) before thrombolysis can be administered; it may be important that the drug is still present at the time of reperfusion.

Limitations of our findings

Our findings suggest that edaravone is effective in animal models of stroke and suggest further work to define the optimal timing of treatment, which might inform future clinical trial design. However, our findings have a number of limitations.

First, our analysis can only be based on published data. Although – in contrast with other drugs tested in animal models of stroke (41) – we found no evidence of publication bias, we only included data from 23 publications, and the methods used to detect publication bias are underpowered when the number of publications is small.

Next, our analysis of the impact of study quality is based on reported study quality. Investigators may in fact have taken some of the measures described by the study quality checklist but not reported them. However, even if this is the case, it is apparent – in this and other meta-analyses – that a failure to report such measures is itself associated with higher reported efficacy and so serves at worse as a surrogate or indirect measure of quality. Reporting of random allocation to group and blinded assessment of outcome has previously been reported to have a substantial impact on reported outcome (42), and these findings are confirmed in the present study. Only one study used animals with comorbidities relevant to stroke patients, and the experiments reporting greatest efficacy used ketamine (43). However, subgroup analysis showed that the use of ketamine only accounted for 0.01% of the total heterogeneity in effect size. Sensitivity analysis by excluding the two comparisons using ketamine showed no significant difference in effect size.

Meta-analysis was developed as a statistical approach to providing summary estimates of efficacy from clinical trials data. As such, it is less well suited to animal studies, where experiments can be orders of magnitude smaller. Further, meta-analysis of clinical trials often seeks to reach a best estimate of the efficacy of a drug rather than (as is the case here) an average efficacy and an understanding of the different factors which impact on that efficacy. To address these issues, we have excluded very small studies (less than three animals), used normalized rather than standardized measures of effect size, and used random rather than fixed effects meta-analysis. Subgroup analyses can lead to spurious findings, but to minimize this risk, we prespecified all subgroup analyses to be conducted and corrected thresholds for significance to account for multiple testing.

Conclusion

Edaravone is effective in improving both functional and structural outcomes in animal models of stroke. However, there is uncertainty about whether timing of administration influences the efficacy of edaravone, and this should be tested in high-quality animal studies before human research. Our study has indicated some methodological weaknesses in current animal studies, to inform further research, therefore, the timely and efficient conduct of these animal studies might best be organized through multicenter animal studies with the standardized study design (44,45).

References


41 Sena ES, van der Worp HB, Bath PMW, Howells DW, Macleod MR. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. PLoS Biol 2010; 8:e1000344.


