Screening Potential Therapies
Lessons Learned From New Paradigms Used in Parkinson Disease

Barbara C. Tilley, PhD; Wendy R. Galpern, MD, PhD

Abstract—In Parkinson Disease (PD) as well as in stroke research there is an urgent need to both optimize the use of resources (number of patients, costs, and time) and select potential effective neuroprotective agents. The processes used to identify and study new therapies for PD may be applicable to the search for new therapies in stroke. The National Institute of Neurological Disorders and Stroke (NINDS) organized the Committee to Identify Neuroprotective Agents for Parkinson’s (CINAPS). CINAPS broadly solicited suggestions for agents and evaluated these agents using rigorous criteria. NINDS also created the NIH Exploratory Trials in PD program (NET-PD), a clinical network where CINAPS recommendations could be tested. Given multiple recommended agents, NET-PD investigators used Phase II futility designs with calibration controls, tested against an historical standard, to screen agents for testing in Phase III trials. Investigators also used the Phase II trial to assess ancillary outcome measures for use in Phase III. The observed value for the calibration controls in the first NET-PD Phase II trial was outside the 95% CI for the historical standard. Using bootstrap methodology it appeared unlikely this outcome happened by chance. The historical standard was updated using more contemporaneous data than were available at the start of the futility trials and a re-evaluation using the new threshold was conducted. Assessment of ancillary outcome measures led to a reduced set of outcome measures for use in Phase III. The CINAPS process, and lessons learned from NET-PD futility studies, particularly with respect to calibration controls and assessment of outcome measures, could enhance the choice and testing of new agents for stroke treatment. (Stroke. 2007;38[part 2]:800-803.)

Key Words: futility studies ■ Phase II clinical trials ■ Parkinson disease ■ stroke

Although the National Institute of Neurological Disorders and Stroke (NINDS) tissue plasminogen activator (t-PA) Stroke Trial showed rt-PA was effective in treating acute ischemic stroke, rt-PA was only tested for use within 180 minutes of stroke onset.1 Although subsequent pooled analyses showed that the window might safely be extended slightly,2 most patients with stroke still arrive at Emergency Departments many hours after stroke onset and outside this treatment window.3 Furthermore, the combination of thrombolysis with neuroprotection may result in a synergistic effect, potentially useful in terms of both safety and the number of patients treatable with thrombolytic agents.4 Thus, the search for new therapies in stroke continues.

In Parkinson disease (PD) there is also an ongoing search for therapies that might modify disease progression.5 The processes used in PD may be applicable to the search for new neuroprotective therapies in stroke.

**Materials and Methods**

Committee to Identify Neuroprotective Agents for Parkinson’s (CINAPS)

To facilitate the search for agents to slow PD progression, the NINDS organized the Committee to Identify Neuroprotective Agents for Parkinson’s (CINAPS) that included experts in PD, clinical trials, and clinical pharmacology. The CINAPS group solicited suggestions from academia (including basic and clinical scientists), industry, clinicians in practice, and from the lay community. The submitted suggestions along with suggestions from the CINAPS group itself were initially evaluated to determine whether the agents could be tested in trials in the near future. Next, agents were systematically evaluated with respect to (1) the scientific rationale, including mechanism and consistency among studies, (2) evidence of blood-brain–barrier penetration, (3) availability of adequate human safety data, (4) efficacy in animal models or preliminary efficacy in humans, including randomized trials or epidemiological studies, (5) feasibility for a long-term trial, and (6) whether or not the agent was already being tested elsewhere.6 After the first CINAPS evaluation, multiple agents were proposed for consideration in further studies.

**NET-PD Phase II Futility Study Design**

In addition to CINAPS, NINDS created the NIH Exploratory Trials in PD (NET-PD) program, a network of clinical sites, along with clinical and statistical coordinating centers, where agents identified by CINAPS could be tested. Faced with multiple agents, the NET-PD Steering Committee evaluated the agents based on practical considerations. Could placebos be obtained? Was there a way to deliver the treatment? Was there a standard clinical dosage? In the interim between the conclusion of CINAPS and the establishment of NET-PD had other clinical trials of the agents begun? This is an
The NET-PD process to select agents for testing.

The NET-PD investigators were concerned that a large Phase III trial of several new untested agents for the many years needed to assess disease progression could be wasteful of resources. Also, the study of multiple agents simultaneously could require additional years to enroll the number of patients necessary to detect small changes, adding to the time until the outcome would be known and potentially discouraging patients and investigators. Thus, a mechanism to rationally select which agents should proceed to Phase III trials was needed. Using a short-term futility trial design, ineffective treatments are screened out more quickly than in a traditional Phase III trial, exposing fewer subjects to ineffective agents. The comparison for futility trials is usually an historical standard. Futility trials are not designed or powered to test efficacy of newly diagnosed PD patients similar to the NET-PD study population, the Deprenyl and Tocopherol Antioxidant Therapy of Parkinsonism (DATATOP) trial, were used to compute an historical standard. Results for the Phase II PD studies were compared against the DATATOP historical standard. As an additional safeguard, given that DATATOP was completed in the early 1990s, NET-PD investigators included a calibration control group for each Phase II study. Equal numbers of subjects were randomized to each arm including placebo. If the observed calibration control value fell outside the 95% confidence limits on the historical standard, the historical standard would be updated using a Bayesian approach.

Because there is no accepted single measure of neuroprotection, NET-PD investigators planned to use multiple measures of disease progression when studying selected agents in Phase III trials. Not all of these potential measures had been used extensively in PD clinical trials. Thus, futility studies were used to assess each instrument with respect to ease of administration and ability to detect change.

Results

In the NET-PD Phase II trials, the primary outcome for the placebo group, change in total UPDRS score, was observed to be outside 95% CIs on the historical standard, and less than (ie, better than) the Bayesian update of the historical mean from DATATOP. To determine whether values as small or smaller than the observed calibration control value could have occurred by chance we used the data from the DATATOP trial. We drew 4000 random samples of size 67 (sample size for the calibration control group) from the DATATOP data set, replacing those samples after they were drawn. We tabulated the percent of samples that had a mean value less than or equal to the calibration placebo mean. Based on this bootstrap approach, only 4% of the samples had a value as small or smaller than that observed for the calibration control in the first futility study. Thus, it is unlikely that a mean of the magnitude observed or smaller would occur by chance alone. NET-PD investigators obtained additional placebo data from another recently completed PD trial involving a similar patient population. These data were not available at the time the NET-PD futility studies were initiated. Consistent with the NET-PD futility studies, placebo mean change in UPDRS at one year was also less than that observed in DATATOP. Combining calibration control data with these contemporaneous data, and using the updated boundary, 3 of the 4 agents tested (minocycline, Coenzyme Q10, and GPI-1485) were found to be futile. Based on the updated boundary only creatine could not be considered futile.

Although no formal registry of placebo data existed for PD at the time the Phase II trials were being designed, investigators across the world studying PD were very generous in sharing data from their trials. The availability of the data allowed the assessment of the validity of the calibration control group and the setting of a new standard for comparison. Since the development of the futility trials, The Parkinson’s Disease Data and Organizing Center (PDDOC) has been developing a collaboration among researchers. PDDOC will host a shared database for observational studies and clinical trials that will include PD subjects and controls and will catalog tissue and biological information, including DNA findings, to facilitate clinical-pathological and genotype-phenotype correlations. This sharing of data strengthens the work of all investigators in PD.

The assessment of outcome measures in the Phase II trials led to the elimination of a few outcome measures that were
too cumbersome for use in long-term studies in PD and led to an increased interest in using the modified Rankin Scale (mRS). Further investigation of the properties of the mRS when used in PD are underway. Although unplanned at the start of the trials, the Phase II process allowed assessment of site-specific training needs before proceeding to larger and longer-term Phase III trials.

Discussion

Lessons Learned That May be Applicable to Stroke

In addition to CINAPS, similar processes are being used as a first step in assessing novel treatments for amyotrophic lateral sclerosis and Huntington Disease. These evaluative approaches provide a model that could be used in finding agents to study in stroke.

There are several lessons learned from the PD futility trial experience that may be applicable to stroke. Although the Phase II futility trial approach is not new to stroke, the use of calibration placebo groups is not standard. For example, the Interventional Management of Stroke Study (IMS) investigators did not include a calibration control group because they believed the placebo data from the NINDS t-PA Stroke Trial would be applicable. As progress is made in the development of stroke therapeutics, and as poststroke outcomes improve, there may be an increasing need for more contemporaneous placebo data.

Currently there is an evolving registry, the Virtual International Stroke Trials Archive, collecting contemporaneous data on placebos that could be useful for planning Phase II studies for stroke research. To date, this registry has included more pharmaceutical than NIH trials, although the NINDS t-PA Stroke Trial was recently added. Greater participation in this registry could facilitate the development of future futility trials in stroke.

The PD experience provides methods for evaluating and possibly updating historical standards should there be an unexpected positive outcome in the calibration control relative to the historical control. Note that if the calibration control group has an outcome much worse than hypothesized based on the historical standard, the Phase II trial should be repeated to avoid the risk of falsely concluding an effective treatment is futile or ineffective.

There is also an opportunity to apply the PD example to the study of stroke outcomes. Just as multiple measures are needed to assess disease progression in PD, multiple measures of outcome are typically used in stroke. It has been said that the ideal poststroke outcome is a measure of whether “my mother is still my mother” after a stroke, referring to the many subtle and unmeasured changes that occur poststroke (Brott T, personal communication, 1993). Whereas some participants at the recent Princeton Conference believe the mRS is a measure of the latter qualities, others believe the mRS is too broad a measure and that such a measure or set of measures is still to be developed.

A recent review of functional outcome measures for stroke concluded that the mRS had excellent reliability, adequate validity, adequate responsiveness to change and required no special tools or training. The weakness cited was lack of clear criteria to assign grades.

The mRS could be considered as a binary or ordinal outcome for futility studies in stroke, particularly if the mRS is planned as the primary outcome for Phase III. The recently completed SAINT trials used the mRS as an ordinal scale for the primary outcome measure. Whether a formal futility study was conducted before the first randomized trial was not reported. The choice of a binary or ordinal mRS would be based on the question to be addressed, the magnitude of effect expected, and the expected distribution of outcomes. For example, if the outcomes for a Phase III trial were expected to fall in a J-shaped distribution such as observed in the NINDS t-PA Stroke Trial (most patients at the extremes of the scale) a smaller sample size would be required for a binary outcome (n = 286) than for an ordinal outcome (n = 394). Calculations were based on achieving 90% power to detect an effect similar to the NINDS trial, α = 0.05, two-sided.

In PD Phase III trials of treatments to slow progression require 5 to 7 years of observation, much longer than the 1 year generally required for futility studies. The outcomes tested in PD futility studies cannot always be the same as the outcomes used for the long-term Phase III trial. In stroke research, futility outcomes and Phase III outcomes are both generally measured at 3 months poststroke. Thus, Phase II trials present a rich opportunity for continued development of new measures and testing of existing measures, particularly the measures to be used in Phase III.

Conclusions

Although there have been many trials of potential neuroprotective agents in stroke, the majority of agents tested have not been effective and a need remains to identify new therapies to test in Phase III trials. The CINAPS process and the use of Phase II futility trials, building on the expanded methodology used in the NET-PD trial, may facilitate this search.

Sources of Funding

Funding for this study was provided by NINDS U01NS043127 & U01NS043128, and U10NS44415 through 44555.

Disclosures

None.

References


Screening Potential Therapies: Lessons Learned From New Paradigms Used in Parkinson Disease
Barbara C. Tilley and Wendy R. Galpern

Stroke. 2007;38:800-803
doi: 10.1161/01.STR.0000255227.96365.37
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2007 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/38/2/800

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

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

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