PROTOCOL TITLE:

A Multi-Center, Single-Blind, Randomized Study Comparing Thymectomy to No Thymectomy in Non-Thymomatous Myasthenia Gravis (MG) Patients Receiving Prednisone

SHORT TITLE:

Thymectomy in non-Thymomatous MG patients receiving prednisone

EXECUTIVE COMMITTEE:

Study Chair: John Newsom-Davis, MD
Vice Chair: Gil Wolfe, MD
Surgical Chair: Alfred Jaretzki MD
Ancillary Studies: Henry Kaminski, MD
Director, Data Coordinating Center: Gary Cutter, PhD

NINDS COLLEAGUES: Robin Conwit, MD
Joanne Odenkirchen

Project Manager: Greg Minisman, MA

Data and Safety Monitoring Committee: To be selected by NINDS

Investigator Sites and Center PIs: See separate Excel file

TABLE OF CONTENTS

1.0 Introduction
1.1 Summary
1.2 Background
1.2.1 Immunopathogenesis: the thymus in MG
1.2.2 Thymectomy as therapy for myasthenia gravis
1.2.3 Shortcomings of thymectomy studies
1.2.4 An evidence-based approach
1.2.5 Nonsurgical treatment of myasthenia gravis
1.2.6 Complications of corticosteroid treatment
1.2.7 Clinical importance of the study
1.3 Rationale
2.0 Objectives
2.1 Primary Objectives
2.2 Secondary Objectives

3.0 Study Design
3.1 Duration
3.2 Subject Management During the Study
3.3 Subject Withdrawal from the Study
3.4 Discontinuation of the Study
3.5 Ancillary Studies

4.0 Patient Selection
4.1 Number of Subjects
4.2 Inclusion Criteria
4.3 Exclusion Criteria
4.4 Screening Log

5.0 Therapy
5.1 Treatment Schedule
5.1.1 Treatment initiation
5.2 Extended Transsternal Thymectomy (ETTX)
5.3 Prednisone Dosing
5.4 Tapering schedule for pyridostigmine and prednisone
5.5 Escalation of prednisone if Minimal Manifestation (MM) status lost
5.6 Additional Therapies
5.7 Discontinuation of study drug

6.0 Study Medication, Description, and Allocation
6.1 Drug Formulation
6.2 Storage
6.3 Dispensing of Medication
6.4 Drug Accountability

7.0 Schedule of Events
7.1 Study Personnel
7.1.1 Medical Coordinator (MC)
7.1.2 Blinded Evaluator (BE)
7.2 Tests and Evaluations
7.2.1 Subject Eligibility
7.2.2. Patient Allocation
7.2.3 Schedule of Events
7.2.4 Blinding
7.2.5 Selection Bias
7.3 Patient-related issues:
7.3.1 Recruitment rate
7.3.2 Compliance
7.3.3 Maintaining Patient Participation
7.4 Patient Monitoring
7.4.1 Chest CT
7.4.2 Laboratory Monitoring
7.5 Data Coordinating Center Activities

8.0 Safety Assessments
8.1 Clinical Safety Assessments
8.1.1 Appointment of a Data and Safety Monitoring Board
8.2 Adverse Events: Definition and Management

9.0 Statistical Statement and Analytical Plan
9.1 Sample Size
9.1.1 Two-sided Hypothesis
9.1.2 Is 30% decrease a sufficient
9.1.3 Penalty for use of drugs in addition to steroids
9.2 Interim Analyses
9.3 Statistical Methods to be used in Analyses
9.3.1 Analysis Populations
9.3.2 Statistical Analysis
9.3.3 Additional Safety Analyses
9.3.4 Lab Abnormalities

10.0 Ethical Requirements
10.1 Declaration of Helsinki
10.2 Subject Information and Consent
10.3 Maintenance of Study Documentation and Supplies

11.0 Further Requirements and General Information
11.1 Changes to the Protocol
11.2 Record Retention
11.3 Protocol Completion

12.0 Investigator’s and Sponsors Agreement

References

APPENDIX I Quantitative MG Score
APPENDIX II Trial Specific Adverse Events
APPENDIX III Trial Specific Adverse Symptoms
APPENDIX IV MGFA Post-Intervention Status
APPENDIX V MGFA Clinical Class
APPENDIX VI Data Coordinating Center
APPENDIX VII MG-ADL score
APPENDIX VIII Declaration of Helsinki
ACRONYMS AND ABBREVIATIONS

AchR-Ab  Acetylcholine receptor antibody
AUDTC  Area under the prednisone dose-time curve,
AUQMG  Area under the QMG weakness score-time curve
BE  Blinded Evaluator
BRN  Blinded Research Nurse
°C  Degree centigrade
CRF  Case report form
CT  Computerised Tomographic (scan)
DCC  Data Coordinating Center
DMC  Data Monitoring Committee
DSMC  Data and Safety Monitoring Committee
EC  Ethics Committee
ECG  Electrocardiogram
ETTX  Extended Transsternal Thymectomy
e.g.  For example
°F  Degrees Fahrenheit
IRB  Institutional Review Board
IVIG  Intravenous immunoglobulin infusion
kg  Kilogram
MC  Medical Coordinator
mg  Milligram
MG  Myasthenia Gravis
MGFA  Myasthenia Gravis Federation of America
MM  Minimal Manifestation (status)
mm³  Cubic millimeter
MC  Medical Coordinator (usually the center PI)
PI  Principal Investigator
QMG  Quantitative Myasthenia Gravis (Weakness Score)
QOL  Quality of Life
REB  Research Ethics Board
RCT  Randomized Clinical Trial
TSAEs  Trial Specific Adverse Events
TSASs  Trial Specific Adverse Symptoms
UAB  University of Alabama at Birmingham
ULN  Upper limit of normal
USA  United States of America
USP  United States Pharmacopoeia
1.0 Introduction

1.1 Summary

Protocol Title: A Multi-Center, Single-Blind, Randomized Study Comparing Thymectomy to No Thymectomy in Non-Thymomatous Myasthenia Gravis (MG) Patients Receiving Prednisone

Study Design: Multi-Centered, Single-Blind, Randomized, Two Arm Trial

Study Objectives: Myasthenia gravis (MG) is an autoimmune disease in which 85% of patients have antibodies to muscle acetylcholine receptors (AchR-Ab) that interfere with neuromuscular transmission. MG frequently causes severe disability that can be life-threatening. Thymectomy has been established therapy in non-thymomatous MG since 1940, based on retrospective, non-randomized studies. Corticosteroids are now being used increasingly either as the sole treatment or in conjunction with thymectomy. Both therapies have associated adverse effects, and indications for their use based on randomized trial data are lacking.

There is therefore an important clinical need to establish the place of thymectomy in the management of MG patients receiving prednisone, specifically extended transsternal thymectomy (ETTX) because it provides the greatest thymic resection with low morbidity and limited risk of nerve injury. To address this, we aim to answer three questions in a rigorously executed 5 year study in which each patient will be followed for at least 3 years:

Question 1: Does ETTX combined with prednisone result in a greater improvement in myasthenic weakness, compared to prednisone alone?

Question 2: Does ETTX combined with prednisone result in a lower total dose of prednisone, thus decreasing the likelihood of concurrent and long-term toxic effects, compared to prednisone alone?

Question 3: Does ETTX combined with prednisone enhance quality of life by reducing adverse events and symptoms associated with the therapies, compared to prednisone alone?

No. of Subjects: Target of 200

Study Population: Male and female subjects aged 18 to 60 years inclusive, with non-thymomatous MG (MGFA class II-IV at entry, AchR-Ab positive, receiving optimal oral anticholinesterase treatment with or without oral prednisone)

Treatment Groups: Subjects will be randomized to:

- ETTX to occur within 30 days of randomization plus Prednisone Therapy
- Prednisone Therapy alone

Visit Schedule: Screening visit(s), Randomization (Month 0), Months 1, 2, 3, 4, 6, and then every 3 months to at least Month 36 (End of Study Visit).
Outcomes:

Primary:
- Stage 1: Comparison of the prednisone treatment alone to thymectomy (ETTX) plus prednisone treatment, based on the clinical response to therapy measured by the Area Under the Quantitative Myasthenia Gravis (QMG) Weakness Score (AUQMG); see Appendix I
- Stage 2: Testing the difference of total prednisone used over the 3 year trial period measured by pill count from blister packs (Area Under the prednisone Dose Time Curve, AUDTC) conditional on the results of comparing AUQMG.

Secondary:
- Frequency of Trial Specific Adverse Events (TSAEs) (see Appendix II)
- Trial Specific Adverse Symptoms questionnaire score (see Appendix III).
- Change in QMG and MG-ADL (Appendix VII) over time and at months 12, 24 and 36
- Time from day 0 to reach initial Minimal Manifestation (MM) status (see Appendix IV)
- MM status at months 12, 24 and 36
- Actual prednisone dose at month 36
- Quality of life assessment (SF 36) at months 12, 24 and 36 (Permission to use SF36 obtained from Dr John E. Ware, Medical Outcomes Trust)
- Cumulative days in hospital for treatment of, or complications related to, MG by months 24 and 36.
- Number of plasma exchanges and IVIG infusions, and total dose of IVIG from day 0 to month 36.

Brief outline of the trial:
This will be a multi-centered, multi-racial, single-blind, randomized, two arm trial in which ~70 centres worldwide will participate. Non-thymomatous acetylcholine receptor antibody positive MG patients (disease duration less than 3 years) aged 18-60 will be recruited, in whom thymectomy and/or prednisone would normally be considered as conventional treatment options. At entry, all will be receiving optimal oral anticholinesterase (pyridostigmine) treatment with or without oral prednisone. After eligibility has been established (all exclusion criteria negative), all required baseline data documented, and myasthenic weakness measured by a Blinded clinical Evaluator (BE) using the Quantitative MG (QMG) score (see Appendix I), the patient will be randomized (day 0) via the Data Coordinating Center (University of Alabama at Birmingham, USA) to extended transsternal thymectomy (ETTX) or no thymectomy. ETTX has been selected because it provides the greatest thymic resection with low morbidity and limited risk of nerve injury, and is already in use at all participating centres. We expect to recruit a minimum of 3-4 patients per centre, each patient to be followed for at least 3 years.

Both groups will receive incrementing alternate day (AD) oral prednisone according to the same protocol (dose increased in 10 mg steps to an initial maximum of 100 mg or 1.5 mg/kg, whichever is the lower). Prednisone will be dispensed in study-provided blister packs to enhance compliance and pill counting; trial treatments will be initiated as soon as
possible after randomization (and before day 30). The Medical Coordinator (MC) will evaluate patients, obtain pill counts from the blister packs, and adjust medication according to the protocol at months 1-3 when blinding to ETX is insecure. They will also report Serious Adverse Effects (SAEs) throughout the trial and document Trial Specific Adverse Events (TSAEs), Trial Specific Adverse Symptoms (TSASs), MG-ADL and SF-36, at the scheduled visits as indicated in Table 1.

At day 0 and from month 4 onwards, each patient, wearing a polo-necked garment and suitably instructed to avoid unmasking of the treatment group, will be assessed by the BE at their centre at fixed time intervals (month 4, 6 and then every 3 months to month 36; see Table 1). The BE will (a) obtain the QMG score at each visit, (b) apply the criterion score (QMG < day 0 value and <14) in conjunction with clinical assessment to determine whether Minimal Manifestation (MM) status has been reached (defined as having 'no symptoms or functional limitations from MG but may have some weakness on examination of some muscles'; see Appendix IV), (c) record the prednisone dose taken since the previous visit by a pill count via the returned blister packs which all subjects will be instructed to bring to the clinic, and (d) adjust the trial medications according to the protocol rules. When MM status is achieved, irrespective of initial randomization group, pyridostigmine (anticholinesterase) medication will first be reduced to a maximum of 240 mg/day provided MM status is maintained. Thereafter prednisone dose reduction (by 10 mg at 2-weekly intervals to 40 mg AD and by 5 mg at monthly intervals thereafter) will be initiated according to the protocol that is identical for both groups, and maintained unless MM status is lost. In that circumstance, the prednisone dose will be increased (10 mg every 2 weeks) until MM status is again achieved, prednisone reduction then beginning after an interval of 1 month. Patients intolerant of prednisone or failing to achieve MM status at 1 year will be prescribed azathioprine 2.5 mg/kg or (if intolerant) cyclosporine. Plasmapheresis or IVIg infusion can be used if clinically indicated and deviation protocols are met.

The primary outcome determinant will be the total prednisone exposure of the two treatment groups over the three years (measured by the Area under the prednisone Dose Time Curve, AUDTC). However, this indicator of concurrent and long term potential adverse effects of prednisone will take into account the Stage 1 analysis of the clinical response to therapy, measured as the Area under the QMG weakness score (AUQMG). In this analysis, we will compute a 99.5% confidence interval on the difference in AUQMG between the groups to estimate the differences, if any, observed in the QMG weakness measure over the three year duration of treatment. Because of the effectiveness of prednisone therapy and the fact that the treatment protocol - aimed at achieving MM status - is the same in both groups, we do not expect the outcome (AUQMG) in the two groups to differ significantly in this Stage 1 comparison (but if differences should occur, these will be taken into account in interpreting our primary comparison of the AUDTC in Stage 2; see Table 2). In Stage 2, we will test the primary hypothesis, looking for differences in the total prednisone dose (AUDTC) between the two treatment arms. Finally, the assessment of the two treatments will be augmented by information on the frequency of the TSAEs associated with the therapies over the 3 year period. A sample size of 200 patients randomized to one of two groups provides adequate power (over 90%) to demonstrate a difference of a 30% reduction in AUDTC between these two groups using a two-tailed Type I error of 0.05%.

Secondary outcome measures include a questionnaire-based comparison by gender of TSASs (see Protocol Appendix III), quality of life measures, time to reach MM status, duration (percent of time) in MM status, days in hospital due to MG, number of plasmaphereses and IVIg infusions.
1.2 Background
MG is characterized by weakness and fatigability of ocular, bulbar, respiratory, trunk and limb muscles.\textsuperscript{18,29} It is the most common disorder of the neuromuscular junction, affecting all races. A United Kingdom investigation estimated a prevalence of 15 per 100,000 and an incidence of 1 per 100,000.\textsuperscript{39} A central and western Virginia population, which closely approximated the Caucasian and African-American distribution of the general US population, identified a prevalence rate of 14.2 per 100,000.\textsuperscript{36} Despite its relative rarity, the financial burden of the disorder to society and the compromise to individual productivity and quality of life is high, although formal analyses of such issues are lacking. Life-threatening weakness may necessitate hospitalization, intensive care monitoring and assisted ventilation when respiratory and/or bulbar muscles are severely affected, so-called 'myasthenic crisis'. One of every five MG patients will suffer myasthenic crisis at some point of their illness.\textsuperscript{46} With treatment, the patient’s strength usually improves, but the illness still often restricts activity, including employment.

1.2.1 Immunopathogenesis: the thymus in MG
Several lines of evidence support the central role of the thymus in the pathogenesis of MG.\textsuperscript{10,11,35} Most patients have thymic abnormalities; 70% have lymphoid follicular hyperplasia, and 10% have a thymoma (in patients with thymoma, removal of the tumor is mandatory and these patients are excluded from this investigation). The usual thymic abnormalities in MG suggest an active immune response. Hyperplastic MG thymuses have more numerous mature T cells (CD4\textsuperscript{+} CD8\textsuperscript{-} or CD4\textsuperscript{-} CD8\textsuperscript{+}, usually present in peripheral lymphoid organs) and less numerous immature T cells (CD4\textsuperscript{+} CD8\textsuperscript{+}) than thymuses of healthy subjects. The morphological changes of thymuses with lymphoid follicular hyperplasia suggest an ongoing immune response: the perivascular spaces are extended, disrupted and filled with lymphoid tissue that resembles peripheral immune organs, and active germinal centers are present, similar to the secondary lymph follicles of peripheral lymphoid organs. The thymuses of MG patients contain all the functional components for the development of an immune response to AchR, and thymic cells in culture spontaneously generate AchR-Ab.\textsuperscript{30,42} Transplantation of thymus fragments from MG patients into SCID mice induces appearance of human AchR-Ab, which binds the mouse neuromuscular junction and causes AchR loss.\textsuperscript{43}

1.2.2 Thymectomy as therapy for myasthenia gravis
In 1939, Blalock et al. reported improvement of generalized MG in a 21 year-old woman following removal of a cystic thymic tumor.\textsuperscript{4} Blalock et al expressed caution: "\textit{We wish to emphasize again the absence of conclusive proof that the improvement noted in our patient is due to the removal of the tumor from the thymic region.}" In a subsequent study by Blalock, of six MG patients who underwent a thymectomy, one became symptom-free, two significantly improved, and two had mild benefit.\textsuperscript{3}

The worldwide subsequent use of thymectomy in MG patients without thymoma has been based on non-randomized case series that have produced a consensus among MG experts that the procedure is sometimes beneficial, with one-third of patients or fewer developing complete remission. Of 821 patients from a large series who underwent thymectomy, 28 percent went into remission.\textsuperscript{25} Although similar remission rates following thymectomy have been published in recent literature,\textsuperscript{24} there have been reported rates as high as 50-80%.\textsuperscript{16} A higher percentage of patients are reported to demonstrate some benefit following surgery but criteria for improvement
are mostly subjective and have differed widely among studies. Randomized controlled studies to
define the role of thymectomy in MG patients have never been performed. For this reason,
doubts exist about the role of thymectomy in MG management for patients without thymoma. In a
survey of 56 MG experts, only three had no reservations in using thymectomy in patients with
generalized MG. Over 20% of the experts surveyed had significant reservations about
recommending the procedure. Only one-third of the experts referred more than two-thirds of their
patients for thymectomy. Reflecting the consensus that the procedure is beneficial, the majority of
experts stated they referred between one-third and two-thirds of the patients for the procedure.

1.2.3 Shortcomings of thymectomy studies
The many shortcomings of existing thymectomy studies are well described. Confounding
differences between surgical patients and non-surgical controls exist in most prior work. Some
studies have used historical controls from a time when outcomes from medical therapy were likely
to have been less favorable. Women and young patients who tend to do better clinically are more
likely to undergo thymectomy, a factor that would favor surgical outcomes. However, surgical
patients more often are generalized with severe disease, a factor that would negatively influence
surgical outcomes. Many of the thymectomies previously analyzed were, by present day
standards, incomplete and therefore would also negatively influence surgical outcome since there
is recent evidence that the more extensive the resection, the better the results. Outcome measures are often
crudely defined. While remission has been a gold-standard, the definition of remission has varied between
studies. Fixed residual weakness or the continuation of medical therapy for MG has been allowed in some
studies. The length of an asymptomatic period has varied from one day to one year before a remission is
declared. Because of this confusion, it has been recommended that a complete stable remission be
defined as at least 6 months to 12 months of no signs or symptoms off all medications. The multiple clinical
grading systems and "improved" categories employed have been overly subjective and often
conflicting, especially in the setting of unblinded observers. The duration of preoperative
symptoms and the loss of patients to follow-up are often not reported. Finally, techniques of
data analysis have frequently been flawed.

1.2.4 An evidence-based approach
A recent evidence-based practice parameter study from the American Academy of Neurology
analyzed large, retrospective, controlled, non-randomized studies of thymectomy in MG. A total of
28 studies published between 1953 and 1998 were identified, describing 21 MG cohorts. The effect of surgery was broadly favorable in most series. However, the benefit of surgery was
generally small. For example, the median relative rate favoring surgery over non-surgical
treatment was 2.1 (mid-halftile 1.4-2.7) for achieving remission, but the gain was modest because
the median remission rate in the non-thymectomized groups was 10%. Other median relative rates
were 1.6 for asymptomatic status (0.9-2.2), 1.7 for improvement (1.2-2.3), and only 1.1 (1.0-1.2)
for survival. The patient subgroup analysis indicated that only those with moderate weakness or
greater (Osserman 2b) showed a significant improvement following thymectomy compared with
controls. Importantly, the modest benefits ascribed to thymectomy were confounded by baseline
differences between the surgical and non-surgical groups described above, as well as the limited

<table>
<thead>
<tr>
<th>Table 1: Shortcomings of Previous Thymectomy Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Confounding differences between surgical versus non-surgical controls</td>
</tr>
<tr>
<td>• Use of historical controls for comparisons</td>
</tr>
<tr>
<td>• Focus on women and younger patients</td>
</tr>
<tr>
<td>• Outcome measures poorly defined</td>
</tr>
<tr>
<td>• Incomplete thymectomies by present day standards</td>
</tr>
<tr>
<td>• Flaws in data analysis</td>
</tr>
</tbody>
</table>

thymic resection performed in a significant percent of the patients. There were no blinded assessments, and outcome measures included the unreliable asymptomatic and improvement categories. In addition, patients undergoing thymectomy tended to be female and of younger age, both factors that would favor the surgical arm. Whereas the thymectomy patients tended to be enrolled later, many also had thymomas and a number were in Class I, all unfavorable to the thymectomy group. The studies also failed to control for the nature of medical therapy. In addition, the few studies that employed a matched design with an attempt to control for multiple confounding variables failed to demonstrate a consistent benefit from thymectomy. As a result, Gronseth and Barohn expressed uncertainty as to whether claims of improved MG outcomes were a result of thymectomy or related to differences in baseline characteristics between surgical and nonsurgical groups. The authors concluded that the benefit of thymectomy in non-thymomatous MG remains unproven. As a practice recommendation, they proposed that thymectomy be considered a treatment option in autoimmune MG patients without thymoma and called for a prospective, randomized study that standardizes medical therapy for all patients and employs rigidly defined and executed outcome measures.

1.2.5 Nonsurgical treatment of myasthenia gravis
Acetylcholinesterase inhibitors are the initial therapy for nearly all patients. While effective in mild cases of MG, these only control symptoms and do not influence its pathogenesis. The introduction of corticosteroids in the treatment of MG was a major clinical advance. Prednisone is the most commonly used and the most clearly reliable immunosuppressive agent; it is associated with improvement or remission in many patients, the proportions varying between reports. Azathioprine, cyclophosphamide, cyclosporine, and mycophenolate mofetil have also demonstrated effectiveness in MG. Essentially all these medications were evaluated in patients with severe MG resistant to corticosteroid treatment, and a reduction of corticosteroid dose was used as a measure of clinical utility of the drug as proposed in the present investigation. Specifically, in the double blind randomized trial in MG of azathioprine plus prednisolone versus prednisolone alone, the primary outcome measure was the prednisolone dose at 3 years needed to maintain remission. The result showed a highly significant reduction in the prednisolone dose in the group receiving the combined treatment.

1.2.6 Complications of corticosteroid treatment
Despite their effectiveness in treating MG, the corticosteroid regimens of most patients are prolonged and often entail high doses, leading to frequent complications and difficulty in tapering without disease recurrence. Corticosteroids have adverse effects such as generalized immunosuppression, hyperglycemia, hypertension, renal insufficiency, hepatotoxicity, mood and sleep disturbance and mutagenicity. A study of 116 MG patients treated with a typical regimen (60-80 mg of prednisone daily then switched to every other day equivalent dose and tapered 5 mg/month) identified serious complications in most patients: Cushingoid appearance (33%), cataracts (18%), weight gain (18%), diabetes (12%), hypertension (12%), and compression fractures (5%). Two more recent series of MG patients identified significant adverse events from corticosteroids in over 40% of patients. The most common side effects in these series were weight gain (19-22%); diabetes (33% in the Japanese series), and hyperlipidemia (25% in the Japanese series). High adverse event rates to corticosteroids are observed across medical scenarios. Of 191 Icelandic patients who received prednisolone for at least 3 months a year or for repeated periods (for a total of 3 months annually), only one-third had no corticosteroid complications. Serious complications included osteoporosis (26%), Cushingoid appearance (26%), fractures (20%), edema (19%), infections (13%), gastrointestinal complaints (10%),
cataracts/glaucoma (9%), hypertension (8%), and hyperglycemia/diabetes (8%).\textsuperscript{14} Twenty-five of 43 consecutive patients (58\%) with temporal arteritis developed major corticosteroid complications after a mean follow-up of 3 years.\textsuperscript{28} Fractures and severe infections were the most common events. Six of nine patients with infections died as a result. Iatrogenic loss of bone density and the increased risk of osteoporosis are the most significant long-term sequelae of corticosteroid therapy,\textsuperscript{37} and current evidence indicates that the deleterious effects of corticosteroids on bone are cumulative and to an extent, dose-dependent, a key fact strongly supportive of the correlation between SAEs, cumulative dose and presumed long term complications.\textsuperscript{38} In the Icelandic population, 46\% had either an osteoporosis-related fracture or corticosteroid-induced osteoporosis.\textsuperscript{14} Other adverse effects include sleep disturbance and mood change that can include suicidal depression and, in rare cases, mania.

1.2.7 Clinical importance of the study
Thymectomy is used worldwide with prednisone therapy to manage non-thymomatous MG. Nonetheless, previous studies in this field have been less than rigorous, both in patient selection and scientific design. Thus, the clinically important information to underpin evidence-based clinical care has been lacking for both patient and physician. Neurologists and thoracic surgeons have not had the data needed to advise patients on the optimal indications for these therapies. We believe that the eagerness of a large number of MG specialists and surgeons to participate in this study provides direct evidence of that deficiency. The outcome of the proposed randomized and rigorously executed study will thus fulfill an important clinical and public health need, and will unquestionably impact on future patient management.

1.3 Rationale for the study
Myasthenia gravis (MG) is an autoimmune disease in which 85\% of patients have antibodies to muscle acetylcholine receptors (AchR-Ab) that interfere with neuromuscular transmission. MG frequently causes severe disability that can be life-threatening. Thymectomy has been established therapy in non-thymomatous MG since 1940, based on retrospective, non-randomized studies. Corticosteroids are now being used increasingly either as the sole treatment or in conjunction with thymectomy. Both therapies have associated adverse effects, and indications for their use based on randomized trial data are lacking.

There is therefore an important clinical need to establish the place of thymectomy, (specifically extended transternal thymectomy [ETTX] because it provides the greatest thymic resection with low morbidity and limited risk of nerve injury), in the management of MG patients receiving prednisone. To address this, we aim to answer three specific questions in a rigorously executed 3 year study:

**Question 1:** Does ETTX combined with prednisone result in a greater improvement in myasthenic weakness, compared to prednisone alone?

**Question 2:** Does ETTX combined with prednisone result in a lower total dose of prednisone, thus decreasing the likelihood of concurrent and long-term toxic effects, compared to prednisone alone?

**Question 3:** Does ETTX combined with prednisone enhance quality of life by reducing adverse events and symptoms associated with the therapies, compared to prednisone alone?
To answer these questions, we propose a multicenter, multiracial, international, single-blinded, clinical trial in which patients with generalized AchR-Ab positive non-thymomatous MG are randomized to ETTX or no thymectomy, with both groups receiving prednisone according to the same set protocol designed to achieve and maintain Minimal Manifestation (MM) status (see Appendix IV) over a 3 year period.

We will address these questions in a stepwise fashion. We first estimate the difference in the average myasthenic weakness, measured as the Area Under the QMG weakness score (AUQMG), of the thymectomy plus prednisone group and the prednisone alone group. Next we compare the risk associated with prednisone therapy as represented by total exposure to prednisone, measured by the Area Under the prednisone Dose Time Curve, (AUDTC). The AUDTC provides a measure not only of the potential for concurrent adverse effects of prednisone therapy, but also for long-term effects not necessarily manifest within the time period of the study (e.g. osteoporosis, cataracts, etc), but likely a consequence of the total prednisone dose received. These two outcomes should establish the optimal therapy, but will be evaluated within the context of the frequency of trial specific adverse events (TSAEs) and adverse symptoms associated with treatment over the 3 year period.

We considered but rejected other possible outcome measures. A clinical outcome measure alone, such as the number of patients in MM status, is unlikely to be sufficiently informative because we expect the majority of patients in both groups to achieve MM status or substantial improvement in myasthenic weakness by the end of the 3 year study, due principally to the effectiveness of the immunosuppressive actions of prednisone. In the prospective study by Palace et al. in which prednisone was prescribed according to a similar protocol (one group being randomized to azathioprine rather than to ETTX), only one patient failed to achieve the equivalent of MM status at the end of the 3 year study, and no difference could be detected in objective or subjective strength measurement scores between groups. Thus a clinical outcome measure alone will not be sufficiently sensitive to detect reliably differences between the groups and this is the rationale for our using a stepwise approach, conditioning our comparison of the AUDTC on the outcome of the differences, if any, observed in the average clinical weakness.

Furthermore, no instrument exists, and none can be devised, that allows an absolute measurement of muscle strength to be used as a reliable index of "remission." Although the Quantitative Myasthenia Gravis (QMG) score quantifies weakness in a variety of muscles and is a valuable and reliable way of measuring change with time, it cannot be used in isolation to guide prednisone dosing because scores for patients in MM status vary considerably (range 1-13) and overlap with scores for patients not in MM status.

The results will, however, include a questionnaire-based comparison by gender of patients' adverse symptom experiences that can occur with the therapies used and quality of life measures. These patient evaluations together with the stepwise primary endpoint will help determine the role of thymectomy as adjuvant treatment for MG patients previously or concurrently receiving prednisone medication. A study such as this, that may draw conclusions from not showing a difference between groups, requires high power to ensure that minimal but important differences are not overlooked; equally, it must be capable of detecting minimal but important differences if they do exist. A greater improvement in strength, fewer adverse effects or symptoms of the therapies, or a meaningful reduction in the total prednisone dose at 3 years in the operated group would determine the immunotherapeutic benefits of thymectomy in this patient population. Subgroup analysis may show whether benefits are confined to those who are prednisone naïve at entry or who are in a particular age group. A positive outcome would also provide indirect
evidence of the possible benefits of ETTX in those not receiving prednisone medication. Conversely, failure to demonstrate a significant difference between groups in the variables measured would suggest that thymectomy is an unnecessary procedure in the population studied. Future patients in this category would therefore be spared the pain and potential adverse effects associated with the operation, and there would be considerable cost-savings to health providers. Thus the results will have a substantial impact on recommended clinical practice.

2.0 Objectives

2.1 Primary objective
The primary objective of this study is to determine in the non-thymomatous MG patient population studied whether ETTX combined with prednisone therapy should be preferred to prednisone therapy alone

2.2 Secondary objectives
The secondary objectives are (i) to determine the efficacy of the therapies by documenting their effects on myasthenic weakness and (ii) to determine their safety by documenting total exposure to prednisone and by recording the Trial Specific Adverse Events (TSAEs) and Adverse Symptoms (TSASs). Importantly, we shall also investigate whether outcomes vary between different subgroups by the following planned sub-group analyses:

- Use of corticosteroids vs. none prior to entering the study
- Male versus female
- Age <40 years and age >40 years at disease onset, (because of documented differences in HLA class and thymic pathology in these groups30,42)

3.0 Study design

3.1 Duration
The study is a multi-centered, Single Blind, Randomized, two-armed treatment trial. The duration of a subject’s participation in this study will be a minimum of 36 months and a maximum of 51 months.

3.2 Subject management during the study
Subjects will be managed according to the study treatment protocol, other care being provided as required by the clinical centers’ usual medical care regimens. Up to 70 recruitment sites are expected to participate in this study.

3.3 Subject withdrawal from the study
Subjects must be withdrawn from the study if either of the following conditions apply:

- The subject desires to discontinue participation in the study.
- The subject is unable to comply with the protocol.

3.4 Discontinuation of the study
This study may be terminated by the Executive Committee and/or Data and Safety Monitoring Committee at any time with the agreement of NINDS.
3.5 Ancillary studies
Ancillary studies will be encouraged so long as they do not interfere with the primary objectives of the trial. ‘Interference’ can be maintaining the subjects in the trial or performing studies on patients that interfere with the goals or objectives of this study. All ancillary studies proposed will be reviewed by the Executive Committee, the NINDS and the DSMB.

4.0 Study population

4.1 Number of subjects
A minimum of 200 subjects will be recruited for the trial.

4.2 Inclusion criteria
To be eligible for entry into this study, candidates must meet the following eligibility criteria at the time of enrollment:

- Male and female MG patients aged 18 to 60 years inclusive (thymectomy is not regularly performed in non-thymomatomous MG patients > ~60 years of age)
- Onset of generalized MG within the last 3 years
- Positive serum anti-acetylcholine receptor binding antibodies (AchR Ab =/> 1.0 nmol/L)
- MGFA class II-IV at entry, using the MG Foundation of America (MGFA) classification (see Appendix V), while receiving optimal anti-cholinesterase treatment with or without oral prednisone

Justification for including patients already receiving prednisone at entry:
In current clinical practice, many patients are already receiving prednisone at the time of thymectomy. In a 2 year retrospective survey undertaken by 35 center PIs, 54% of patients were receiving prednisone at the time of thymectomy, and a further 6% received it within 2 years. Because prednisone treatment exerts actions on the cellular composition of the thymus, it is essential to include this group of pre-treated patients to investigate whether prior treatment influences the outcome of thymectomy and to establish whether or not thymectomy is a redundant procedure in the context of well-established corticosteroid therapy.

- Subject has given written informed consent prior to any testing under this protocol, including screening tests and evaluations that are not considered part of the subject's routine care.

4.3 Exclusion criteria
Candidates will be excluded from study entry if any of the following exclusion criteria exist at the time of enrollment:

- Ocular MG without generalized weakness (MGFA Class I; see Appendix V) or minimal weakness that would not require the use of corticosteroids
- Myasthenic weakness requiring intubation (MGFA Class IV) in the prior month
- Immunosuppressive therapy other than corticosteroids in the preceding year
- Medically unfit for thymectomy
- Chest CT evidence of thymoma. CT examinations will conform with the following guidelines: (a) 5-7.5 mm slice thickness through a region encompassing at a
minimum the suprasternal notch through the domes of the diaphragm; (b) contrast enhancement is not required, but contrast material may be given if deemed necessary by the investigator; (c) a spiral CT scanner will be preferred when available, but non-spiral examinations are acceptable

- Pregnancy or lactation; contraindications to the use of corticosteroids, unless postmenopausal or surgically sterile, unwillingness to practice effective contraception, as defined by the investigator, during the study. The rhythm method is not to be used as the sole method of contraception. Women considering becoming pregnant during the period of the study are to be excluded.
- A serious concurrent medical, neurological or psychiatric condition that would interfere with thymectomy or subsequent clinical assessments
- Unwillingness or incapacity to participate, to agree to necessary follow-up visits, or to give written and informed consent
- Current alternate day dose of prednisone > than 1.5 mg/kg or 100 mg or the equivalent daily doses (> 0.75 mg/kg or 50 mg).
- Participation in another experimental clinical trial
- History of alcohol or drug abuse within the 2 years prior to randomization.
- Unwillingness or inability to comply with the requirements of this protocol including the presence of any condition (physical, mental, or social) that is likely to affect the subject's returning for follow-up visits on schedule.

4.4 Screening log
Participating centers are required to document (i) all screened candidates, and to forward this to the data coordinating center and (ii) all subjects initially considered for inclusion in this study but then excluded, specifically stating the reason(s) for their exclusion. There are 3 basic groups of exclusions:

- Group 1: Those patients who don’t meet the Inclusion Criteria
- Group 2: Those who do, but don’t satisfy all the Exclusion Criteria
- Group 3: Those who should be screened but are not for some reason (PI decision, unavailable, etc.)

5.0 Therapy

5.1 Treatment schedule

The MC and BE will be assisted by a detailed Manual of Operations that contains flowcharts to indicate the actions to be followed at each point in the treatment protocol.

For each potentially eligible patient, review of inclusion and exclusions will be conducted and entered into the data management system in order to enable a close monitoring of recruitment efforts. Once eligibility is established by the MC and confirmed by their signing the completion checklist, patient data will be entered into the data management system and automatically randomized to one of the two treatment groups.

The two arms will be:

- Extended transsternal thymectomy (ETTX) plus prednisone treatment
- Prednisone treatment alone
Patients will be receiving either optimal anticholinesterase medication (pyridostigmine) at randomization, or no anticholinesterase medication if the patient's status without it is considered optimal by the MC. The treatment protocol for prednisone treatment following randomization will be the same for both arms.

5.1.1 Treatment initiation
Day 0 is the randomization date. Protocol treatments (prednisone dosing alone or ETTX plus prednisone dosing) will start as soon as possible after day 0 and before day 30.

5.2 Extended transsternal thymectomy (ETTX)
Several surgical techniques exist for removal of the thymus in non-thymomatous MG, and debate continues on which is preferable. Although removal of all thymic tissue is the accepted goal, it has not been proven that complete resection is necessary or that all operative techniques accomplish it. The debate is compounded by the relatively recent demonstration that the thymus is not a well-encapsulated bi-lobed gland but is usually multi-lobulated and associated with islands of thymus in surrounding fat.

There are three primary surgical approaches: transcervical, transsternal, and videoscopic. The "maximal" approach combines the transcervical and transsternal operations and is considered the benchmark against which other resections are measured. The transsternal approaches, especially the extended transsternal operation, are the most commonly employed resections at leading MG centers. The transcervical and videoscopic techniques have been developed as minimally invasive to reduce the morbidity of the surgery and eliminate an unpleasant scar. The proponents of these procedures claim equal resections and equal results, although this has not been confirmed. There is some evidence to support the hypothesis that the more extensive the resections, the better the results. Complication rates for thymectomy are much lower now than they were prior to 1970, when perioperative mortality ranged between 5 and 15%. Mortality rates are now below 1%. Common morbidities include acute respiratory failure from MG crisis in 6%, infection in 11%, and recurrent laryngeal or phrenic nerve injury in 0-2%. Further, uncomplicated thymectomy is an expensive procedure requiring hospitalization for 3 to 7 days, usually with 1 day of intensive care monitoring. The extended transsternal approach (ETTX) is selected for this study because it provides the greatest thymic resection with low morbidity and limited risk of nerve injury. This approach is already in use at all centers participating in this study, and the surgeons participating are experienced in the method.

Each thoracic surgeon will have agreed to perform an ETTX as prescribed and have been previously certified to participate (see Appendix IX). A complete median-sternotomy is to be employed with the goal of an en-bloc resection of all tissue in the mediastinum that anatomically may contain gross and/or microscopic thymus. These resections will include removal of both sheets of mediastinal pleura and sharp dissection on the pericardium. Surgeons have been advised that extreme care be taken to avoid injury to the phrenic and left vagus nerves. As emphasized by the proponents of ETTX resections, it is preferable to leave behind some tissue containing thymus or suspected thymus than to injure these structures. Photographs of the gross specimen will be taken using a study-provided digital camera. These will be transmitted into the data management system as JPG files; copies will be sent to the Surgical Coordinator by the data coordinating center, who will be blinded to the clinical outcome, source of photos, etc. The Surgical Coordinator will analyze the extent of thymic resections based on the photographs, mandated detailed operative reports, and
gross determinations. The unblinded MC will provide required neurologic consultation for patients in the perioperative period through discharge from the hospital. Should a thymoma be discovered unexpectedly, the patient will remain in the study, according to an intention-to-treat model. Thymomas detected subsequently in the non-thymectomy group will be allowed standard of care treatment via thymectomy. We do not expect many, if any, of these cross-overs, but the slightly increased sample size has been added to assist in overcoming these and other theoretically potential occurrences.

5.3 Prednisone dosing
Either oral prednisone or prednisolone can be used, referred to here as prednisone. Potency is equivalent (Physicians' Desk Reference, 2004). The study drug will be dispensed in study-provided blister packs for patient convenience and pill-counting/monitoring efficiency. For patients not already receiving prednisone at entry, a starting dose of 10 mg of prednisone will be prescribed by the MC. On an alternate day (AD) regimen, each successive dose of prednisone will be increased by 10 mg to 1.5 mg/kg body weight AD or to 100 mg AD, whichever is the lower. For patients already receiving an AD prednisone dose prior to Day 0, each successive dose will be similarly increased by 10mg to a peak dose of 1.5 mg/kg AD or to 100 mg AD, whichever is the lower. Patients who are on a daily prednisone dose at entry will be switched to the equivalent AD dose, so that the average dose over two days will match their prior daily dose; subsequent doses will be increased as above. The maximum dose (100 mg AD) can be increased to 120 mg AD in patients failing to achieve MM status by month 4. Patients will not be enrolled if they are already receiving prednisone at doses that exceed the study’s maximum dosing protocol.

5.4 Tapering schedule for pyridostigmine and prednisone
The MC rather than the Blinded Evaluator (BE) will assess the patient at the scheduled visits at months 1, 2, and 3 (Table 1) and adjust the medications according to the protocol because it is not feasible to maintain blinding in the first three months following thymectomy owing to clinical clues such as chest discomfort, surgical dressings, etc.

The BE will take over the assessment and implementation of the protocol at the month 4 visit and all subsequent trial visits (Table 1) to the end of the study (month 36). He/she will be responsible for (a) obtaining QMG scores at each trial visit, (b) applying the criterion score (QMG < baseline (day 0) value and < 14) in conjunction with clinical assessment to determine whether MM status has been reached, and recording this for use in the primary outcome measure (c) recording the prednisone dose taken since the previous visit by a pill count from the used and unused portion of the blister packed dose cards, and (d) adjusting trial medications according to the protocol as follows. Assistance can be provided by the optional Blinded Research Nurse (BRN) for (a) and (c) above. When a patient has reached MM status, the daily dose of pyridostigmine (anticholinesterase medication) will be reduced by 60 mg at weekly intervals. If MM status is lost during the tapering of pyridostigmine, the dose will be adjusted to regain control, and the tapering schedule resumed after an interval of one month. Once pyridostigmine has been fully withdrawn or is at a maximum daily dose of 240 mg without return of MG symptoms (i.e. MM status is maintained), the dose of prednisone will be reduced. The prednisone dose will be tapered by 10 mg at 2-weekly intervals to 40 mg AD and by 5 mg at monthly intervals thereafter, provided that MM status is maintained. To illustrate, a patient in MM status on a maximal prednisone dose of 100 mg AD who begins the taper at the month 3 visit will be on 10 mg AD at the end of year 1 and off prednisone completely by the end of month 13, assuming MM status is maintained.
5.5 Escalation of prednisone if MM status lost
If symptoms of myasthenic weakness recur while prednisone is tapered, the patient will be re-examined, if necessary at a clinic visit outside those specified in the protocol. If significant symptoms or signs of myasthenic weakness are confirmed by the BE, the prednisone dose will be increased by 10 mg every 2 weeks on an AD basis as required to regain control and achieve MM status (after the unscheduled visit the patient will return to the planned visit schedule as soon as possible). Escalating doses of prednisone will be limited to the maximum level previously defined. Pyridostigmine will not be increased above a maximum daily dose of 240 mg a day. When MM status has been restored for a period of one month, prednisone dose reduction will resume at the same rate defined in the prior section. Patient attendance at the study site is mandatory at all times specified in Table 1 unless an extra visit occurs within 2 weeks before the specified visit.

5.6 Additional therapies:
Immunosuppressive medication
For a BE to prescribe this, one of the following deviation protocols must be observed:

- **Patient has not achieved MM status during year 1.** Maintain maximum prednisone dosage according to the trial protocol. Add azathioprine (dose 2.5 mg/kg body weight), or cyclosporine (dose 3.5 mg/kg) if azathioprine intolerant. Continue to follow prednisone dose adjustment protocol.
- **Patient experiencing intolerable prednisone adverse effects.** Reduce prednisone to maximum tolerated dose. Add azathioprine (dose 2.5 mg/kg body weight), or cyclosporine (dose 3.5 mg/kg) if azathioprine intolerant. Inform the Data Coordinating Center. How this is counted in the analysis is discussed below.

Plasma exchange (PE) or intravenous immunoglobulin (IVIg).
The MC is permitted to prescribe a course of PE or IVIg between day 0 and day 120 (month 4) if one of the following deviation protocols is met: (a) Patient's clinical state requires this treatment in preparation for thymectomy or postoperatively; (b) Patient's clinical state judged to require this treatment prophylactically during prednisone initiation; (c) Patient's clinical state and QMG score have both deteriorated from baseline (day 0) and patient is judged to be at risk from life-threatening bulbar or respiratory muscle weakness.

From month 4 through to month 36, the BE is permitted to recommend to the MC a course of PE or IVIg if the following deviation protocol is met: Patient's clinical state and QMG score have both deteriorated from that on day 0, and patient is judged to be at risk from life-threatening bulbar or respiratory muscle weakness. Patients requiring these treatments will be managed by the MC and not by the BE (to avoid the risk of unblinding). Existing protocol will continue unchanged. These treatments may not be used to maintain MM status and the “penalty” for using these treatments on the outcome measure is discussed below. Quantification of these treatments is a secondary outcome measure.

Late thymectomy
An option for patients in the non-surgical group to undergo thymectomy during the 3 year period of the study is not offered as part of the trial protocol. Patients would be instructed at entry not to raise this issue with the BE, but could do so with the MC. The MC would be required to point out that the benefits of thymectomy in this circumstance are unproven, and that thymectomy would still be an option for the patient at the end of the study if judged to be clinically indicated. If thymectomy was undertaken during the 3 year trial period, the patient would continue to be followed. If the patient demands thymectomy, a protocol deviation will be documented in the data management system. If the MC feels that for best medical care a thymectomy is essential, a
request to the Executive Committee via the Data Management Committee will be made, partly in order to discourage any such use and to be sure we understand any criteria being used. Investigators will be advised of the risk to the entire study for unnecessary cross-over. It is likely that most clinicians participating in this study would wait until the end of year 3 before recommending thymectomy. Nevertheless, we will conduct an intention to treat analysis, but of course monitor for any crossovers as they clearly would impact the trial.

5.7 Discontinuation of study medication
Subjects who modify or prematurely discontinue prescribed study medications should remain in the study and continue the protocol-specified follow-up evaluations.

Modification or discontinuation of medication may be necessary in the following circumstances:

- The subject desires to modify/discontinue medication treatment under this protocol.
- The subject experiences a medical emergency that necessitates permanent discontinuation of study medication
- At the discretion of the investigator, the Executive Committee and Data Coordinating Center, for medical reasons.

6.0 Study medication, description, and allocation

6.1 Drug formulation
Prednisone 10mg Tablets in Blister packs of 100 tablets per pack.

6.2 Storage
Drugs may be stored at room temperature until expiration date provided with drug supplies.

6.3 Dispensing of medication
Medications may be dispensed by a licensed pharmacy, or by the clinic with appropriate controls and documentation of drugs prescribed and accountability.

6.4 Drug accountability
The study site must maintain accurate records demonstrating dates and amount of study drug received, to whom dispensed (subject-by-subject accounting), and accounts of any study drug accidentally or deliberately destroyed. Unless otherwise notified, all drug packs, both used and unused, must be saved for drug accountability.

7.0 Schedule of events

7.1 Study personnel
The following personnel will be involved in the conduct of this study. Where specified, only the personnel indicated must perform evaluations described in this section.

7.1.1 Medical coordinator (MC)
The MC will be a neurologist, usually also the Center PI, who will have overall clinical responsibility for subject screening and safety and for the conduct of the trial at their center. They will oversee subject management, including the assessment, documentation and treatment of adverse events and symptoms. Also, for the first 3 months of the trial (when blinding of the BE would be insecure), the MC will be responsible for pill counting.
and for adjusting drug dosages following the protocol based increases and decreases in prednisone.

7.1.2 Blinded evaluator (BE)
The BE will be a clinical neurologist trained in the use of the Quantitative MG Score (QMG), who will perform all Neurological examinations in a blinded manner. The BE will not engage in conversations that might unmask the patient’s treatment assignment and will remind the patient that they are not to know the status of their treatment.

7.2 Tests and evaluations

7.2.1 Subject eligibility - screening and randomization
The required Case Report Forms (CRFs) including the eligibility form must be completed on every patient. Patients deemed ineligible will be documented as to reasons for ineligibility, as mentioned above.

Data to be collected via the Case Report Forms (CRFs) will include:
- Medical history.
- A complete physical examination.
- Historical and concomitant medications.

7.2.2 Patient allocation (randomization)
Random allocation using a distributed web-like data management system will be stratified by center and balanced in time using short block sizes (2 or 4), but will not be revealed to the sites. Once eligibility has been verified at the time of entry of the baseline study data, an automated randomization will be made. Each clinical individual certified to enter data (certification will occur at the training meetings) will be the only ones allowed to conduct the randomization. The blocking will be used to achieve a reasonably balanced allocation within centers to help ensure adequate numbers for an investigation of variation in intra-center treatment group comparisons. No other stratifications are considered practical or necessary. Stratification by prednisone dose was considered, but because practices cluster heavily within site, it was deemed redundant and potentially problematic. Information will be collected on relevant demographic and clinical factors to permit stratified analyses (post-stratification) to reduce variation due to their effects, and to check the assumption of a uniform difference in treatment effects across centers and other subgroups.

7.2.3 Schedule of visits is shown in the table below:
Table 1: Schedule of visits and major outcome variables:

<table>
<thead>
<tr>
<th>Year 01</th>
<th>Day 0</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M6</th>
<th>M9</th>
<th>M12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical assessment by BE or MC</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Pill count by BE or MC</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>QMG by BE or MC</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>MG-ADL by MC</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>
Key: Day 0 is Randomization; M=month; treatment protocol initiated by day 30 (M1); BE, Blinded Evaluator; MC, Medical Coordinator. To maintain BE blinding, the clinical assessment, QMG and pill count at Months 1, 2 & 3 will be done by the MC, indicated by (+) and (MC).

7.2.4 Blinding
Because patients cannot be blinded in this thymectomy + prednisone versus prednisone alone trial, all evaluations that contribute to the primary outcome measure will be carried out by a Blinded Evaluator. Blinding the evaluator and having firm criteria for dose alterations is therefore crucial and the following strengthening measures have been put in place to maintain this.

- The BE will have no duties on the relevant medical and surgical wards.
- The month 1, month 2, and month 3 evaluations cannot be blinded because of the possibility of recent chest surgery; these will be undertaken by the MC and will not contribute to the outcome measures (see Table 1).
- The day 0 and month 4 through month 36 evaluations (see Table 1) that contribute directly to the primary outcome measure will be performed by the BE.
- All evaluations by the BE will take place in an office that is remote from the medical and surgical wards.
- The BE will not have access to the hospital case notes.
- The BE and other medical and nursing staff with whom the BE might come in contact will be informed of the crucial importance of rigorously maintaining blinding. The BE will be instructed to report this to their MC if he/she believes blinding has been broken with respect to a particular patient. Should that occur, a substitute BE would be appointed and trained to complete the evaluation of the patient in question.
- All centers will be provided with trial polo-necked garments for use by all patients during evaluation visits.
- Patients will be informed verbally and in writing of the necessity to avoid revealing to the BE either directly or indirectly whether or not they have undergone
In addition, all patients will receive a further reminder before each evaluation visit.

- Because completion of the MG-ADL and SF36 forms might provide clues regarding previous surgery, these will be completed by the patient, and then checked for completeness by the MC who will ensure their entry into the data management system and who will store the hard copies locally; this information would not be seen by the BE.

7.2.5 Selection bias
To avoid selection bias and to ensure that all eligible patients attending the MC's clinic are offered entry, the MC will log all such patients and the logs will be entered into the data management system. If a patient is excluded from the study because of one or more of the exclusion criteria, an explanation will be recorded on the study form and the Data Coordinating Center will monitor the frequency of various reasons for exclusion and to ensure that consecutive patients are being rigorously considered for the trial.

7.3 Patient-related issues

7.3.1 Recruitment rate
The number of participating centers has been increased to 69 to enhance recruitment and ensure accrual within 18 months. With a mean figure of 4.6 eligible patients per centre per year (based on our previous survey), and assuming a 50% participation rate (lower than the 67% figure obtained in a Preliminary survey of 86 patients which suggest a slightly higher acceptance rate) we estimate recruiting approximately 150 patients per annum and completing recruitment within 1.5 years.

7.3.2 Compliance
We do not anticipate that lack of compliance with study procedures will be a substantial problem. Both arms will be receiving active treatment for MG, and non-compliance, which should be low, is likely to be similar between them. Further, MG patients have a built in biofeedback loop to compliance because, when they fail to take medication, symptoms increase. Thus, MG patients tend to be more compliant with medication than patients suffering, for example, from hypertension. Compliance is an issue of course, because patients may stop taking or reducing medications because of side effects. However, this form of non-compliance is part of the reason we are using drugs taken rather than drugs prescribed for deriving the prednisone AUDTC.

7.3.3 Maintaining patient participation
Strenuous efforts will be made to avoid trial dropouts by maintaining strong links (i) with individual PIs via e-mail, telephone calls and faxes as necessary, (ii) with other study group staff through the prompting by the data management system and (iii) with patients through newsletters, birthday cards etc.

A contribution towards the patient's travel expenses for trial visits will encourage attendance, as will the special care they will receive. During training, ideas for the clinics to use, such as making these patients feel 'special' with reduced waiting time or other benefits will be suggested to enhance the patient's experience with the clinic. Patients who change residence can transfer to another participating center if suitably located. In the Palace et al.31 randomized, controlled trial of adjunctive azathioprine therapy in MG, the drop-out rate was relatively high, principally due to death or to significant intercurrent illness that were confined to subjects 63 years or older. The upper age limit for the proposed trial is 60 years which will reduce the higher dropout within this
cohort. Given the treatment provisions available in the trial and the tendency of MG patients to maintain close follow-up with their providers, we anticipate that a drop-out rate no more than 10-15% at 3 years is possible. To be conservative, however, sample size calculations are based on a drop-out rate of 20% at 3 years.

7.4 Patient monitoring

7.4.1 Chest CT
At the end of the 3 year study period, as part of standard care, patients randomized to medical therapy will have a repeat chest CT to assess for thymic changes over the course of the study and to confirm that a thymoma has not developed. Guidelines for the follow-up scan are identical to those listed in the exclusion criteria. At any time during the study, if clinically indicated, a repeat chest CT may be requested for clinical indications and reviewed by the MC.

7.4.2 Laboratory monitoring
Complete blood count, glucose, glycosylated hemoglobin, and potassium will be measured on Day 0 or earlier (baseline), at least monthly from day 0 to month 3, and every 3 months thereafter. Metabolic derangements related to the medications will be managed according to routine practices. TSAEs identified by laboratory monitoring will be reported as required by the protocol (Table 1). The investigators are very familiar with potential adverse events from prednisone, given its frequent use in MG management.

For patients prescribed azathioprine according to the deviation protocol, liver function tests (LFTs: bilirubin, alkaline phosphatase, aspartate-aminotransferase) and complete blood count will be checked every 2 weeks for the first 2 months, at month 3, and then on an every 3-month schedule. If azathioprine intolerance develops, the drug should be withdrawn, and a TSAE form completed and entered into the data management system. Mild stable elevations of LFTs, defined as less than twice the upper limit of normal for the laboratory used by the MC, are not an indication for drug withdrawal unless associated with systemic symptoms. Moderate disturbances of LFTs (defined as 2-5 times the upper limit of normal) will result in temporary discontinuation of azathioprine. Azathioprine will be restarted when LFTs have returned to normal, at a dose that is 50 mg lower than before withdrawal. Azathioprine will be permanently withdrawn if adverse effects re-develop at this lower dose, or for any LFT disturbance defined as severe (laboratory testing > 5 times the upper limit of normal). If the WBC count falls below 3.0x10^9/L or the absolute neutrophil count falls below 1.5x10^9/L, azathioprine will be stopped until values normalize. Once white cell counts return to normal, azathioprine will be restarted at a dose 50 mg below the prior dose. If values again fall below the guideline values given above, azathioprine will be permanently withdrawn.

If cyclosporine is used in place of azathioprine, blood pressure and BUN/Creatinine levels will be added to the laboratory monitoring described above for azathioprine. Cyclosporine doses will be reduced appropriately should the patient develop toxicity (Creatinine >1.5mg/dl or uncontrolled hypertension) and the primary analysis will take their inclusion into account. TSAEs will continue to be documented.

7.5 Data coordinating center activities
University of Alabama at Birmingham (UAB) will serve as the Data Coordinating Center (DCC) for this study. Details of the role and functions are provided in Appendix VI. These activities have included planning and organization of the submission and design, sample size estimation, etc. The orchestration of the training, preparation of manuals, certification of the investigative teams
etc. will all be carried out as part of the DCC responsibilities. A web-based data entry and randomization system will be utilized, which is HIPAA compliant and affords use worldwide with minimal hardware requirements. All of the proposed participants have sufficient internet connections as has been evident in communications in preparation for this submission. The system to be used in MGTX is currently being used in 4 large multicenter trials (SPS3, TEAR, CREST and CombiRx) and is operational and tested. The system facilitates monitoring, feedback, quality control and assurance, drug distribution, safety monitoring and analyses. The DCC will also be responsible for reporting to the NINDS, DSMB and the study as a whole. Publications and presentations will be supported by the DCC and ancillary studies, if approved, will be managed through this same system.

8.0 Safety assessments

8.1 Clinical safety assessments

The following clinical safety assessments will be performed during the study.

8.1.1 Appointment of a data and safety monitoring board:

A Data and Safety Monitoring Board (DSMB) will be appointed by the NINDS. This team should include a statistician, several clinicians, an ethicist and/or patient advocate. At the request of NINDS, we have nominated individuals for the DSMB, recognizing that conflict of interest must be avoided. The role of the DSMB is to advise the NINDS and, through their feedback, to advise study leadership on difficult clinical and ethical issues that may arise during the study. The board will also be provided with interim safety analyses. We note that because of the important long term nature of the question, we do not expect an interim analysis for effectiveness to lead to stopping the study. We recognize that in many studies of long duration, DSMBs often require an interim look. Since this is predominantly a safety trial with an irreversible treatment in one arm, any potential decision to stop will by necessity take this into account. Nevertheless if there was overwhelming clinical benefit for the ETTX group, the DSMB and the study executive committee would act to stop the trial.

A plan for reporting serious adverse events (SAEs) has been formalized and is automatically built into the data management system, and also described in the manual of operations for the study included with this grant submission (see also section 8.2). [Serious adverse events (SAEs) in this context represent only some of those included in the Trial Specific Adverse Events (TSAEs)]. The SAEs will be reported promptly to the IRBs, chief investigator, and the chairman of the DSMB and others as appropriate or directed by the NINDS. The data management system automatically emails notification of any such events and will continue to email reminders to the specific person until the email is opened. This will facilitate timely reporting in this international trial and will be conducted over secure on-line reporting systems to ensure confidentiality. To prevent problems of confidentiality or, in the US, Health Insurance Portability and Accountability Act (HIPAA) violations – the data management system is designed to not have access to local databases – only the local center’s system initiates data transmissions, scheduled software updates, etc. A formal report of SAEs will be prepared by the study statistician for the DSMB to review every three months or on any schedule deemed more appropriate or desirable by the DSMB.

At regular intervals (every 6 months), reports will be prepared for the DSMB to ensure the safety of participants in the trial. Individual reporting to the DSMB and IRBs in the US will also be done as required. The groups will be compared with respect to the incidence of SAEs. This will involve comparing their distributions among the SAE categories (including none) and various incidences using $\chi^2$ tests, investigating sub-groups where it seems appropriate. Because the SAEs are a
safety issue, these tests will use a 0.1% significance level to avoid repeated testing generating false-positive results, but the DSMB will be given all the results so they can use their discretion independent of the statistician’s categorization of SAEs. A report of SAEs will be provided to all members of the DSMB every six months. A teleconference or face-to-face meeting to discuss these is planned at least every 6 months. More frequent communication through email, phone conferencing, and faxed documents will be used as needed.

8.2 Adverse events: definition and management

Serious adverse event (SAE)
Serious adverse events are defined as the following (in accordance with 21 CFR §600.80 and the recommendations of the International Conference on Harmonization [Federal Register, October 7, 1997, Vol. 62, No. 194, pp. 52239-45]):

- Any death.
- Any life-threatening event, i.e., an event that places the subject, in the view of the investigator, at immediate risk of death from the event as it occurred (does not include an event that, had it occurred in a more severe form, might have caused death).
- Any event that requires or prolongs in-patient hospitalization.
- Any event that results in persistent or significant disability/incapacity.
- Any congenital anomaly/birth defect diagnosed in a child of a subject who participated in this study.
- Any event that necessitates medical or surgical intervention to prevent one of the outcomes listed above.

Reporting of serious adverse events (SAEs)
In order to adhere to all applicable laws and regulations for adverse event reporting, the Data Coordinating Center must immediately be made aware of adverse events that are SERIOUS as defined above. Therefore, it is the investigator’s responsibility to ensure the following:

Any SERIOUS adverse event including DEATH that occurs from the time of randomization up to the final visit inclusive, must be reported to the Data Coordinating Center within 24 hours following report of the event, regardless of the severity or relationship to the therapies.

In reporting a serious adverse event via the web-based system to the Data Coordinating Center, the Center PI must provide specific information regarding the event as per the Serious Adverse Event form discussed in the Manual of Operations. For deaths, a Record of Death CRF as well as a copy of the autopsy report (when available) must also be submitted to the Data Coordinating Center. The Center PI must keep a copy of all documentation related to the adverse event in the site’s study files.

The Center PI must also notify the local review committee, i.e., the Institutional Review Board (IRB), Research Ethics Board (REB), or Ethical Committee (EC) as per local requirements. Documentation of these reports must be kept in the site’s study files.
9.0 Statistical statement and analytical plan

Myasthenia gravis (MG) is an autoimmune disease in which 85% of patients have antibodies to muscle acetylcholine receptors (AchR-Ab) that interfere with neuromuscular transmission. MG frequently causes severe disability that can be life-threatening. Thymectomy has been established therapy in non-thymomatous MG since 1940, based on retrospective, non-randomized studies. Corticosteroids are now being used increasingly either as the sole treatment or in conjunction with thymectomy. Both therapies have associated adverse effects, and indications for their use based on randomized trial data are lacking.

There is therefore an important clinical need to establish the place of thymectomy, (specifically extended transsternal thymectomy [ETTX] because it provides the greatest thymic resection with low morbidity and limited risk of nerve injury), in the management of MG patients receiving prednisone. To address this, we aim to answer three specific questions in a rigorously executed 3 year study:

**Question 1.** Does ETTX combined with prednisone result in a greater improvement in myasthenic weakness, compared to prednisone alone?

**Question 2.** Does ETTX combined with prednisone result in a lower total dose of prednisone, thus decreasing the likelihood of concurrent and long-term toxic effects, compared to prednisone alone?

**Question 3** Does ETTX combined with prednisone enhance quality of life by reducing adverse events and symptoms associated with the therapies, compared to prednisone alone?

The approach is a stepwise assessment. The first stage is to estimate the difference in the average myasthenic weakness of the thymectomy plus prednisone group and the prednisone alone group. Based on the outcome of this analysis, the second stage will be a conditional test of prednisone use over the first 3 years determined by pill counts. This approach is used because we anticipate that the stage 1 analysis will show that the two groups will have quite similar average myasthenic weakness scores over time, due principally to the effectiveness of the immunosuppressive actions of prednisone and to the fact that the drug treatment protocol for both groups is identical. Thus, our interpretation of the hypothesis test posed in stage 2 will be interpreted differently depending on whether the outcome of the first stage shows reasonable equivalence in average weakness or more weakness in one arm or another. Thus, we have used an estimation approach for stage 1 and then, given that result, we test the hypothesis of prednisone exposure for our primary outcome consideration (see 9.1.1).

That pill counts could be a potential source of error is acknowledged, but we believe that this does not invalidate their use as a study outcome. A clinical measure alone cannot be used as the primary endpoint in this study because it takes no account of the prednisone dose required to achieve whatever measure is used (for example, the number of patients in Minimal Manifestation status at year 3 or QMG score integrated over time). Thus a means of quantifying the prednisone dose is an essential component of the primary consideration.

**Justification for using pill counts**

(a) We would like to have a more precise measure of patient exposure to prednisone than pill counts, but we do not believe it exists. Measuring prednisone blood levels at trial visits was
considered as an alternative measure of adherence and even dosage, but because of the short half-life of oral prednisone (6 hours)\textsuperscript{44}, the alternate day dosing, the varying time interval between dosing and the serum sample being taken, and the practical difficulties in making reliable measurements, this method was not thought to be adequately informative. In addition, the costs for serial blood measurements would also be substantial. Even if such a test were possible, we would need to use random testing, and with the extremely tight logistics to get the patients to the clinic to accommodate the half life exposure and treatment schedule, the serum levels would be likely to suffer greater problems than pill counts do. Furthermore, to gauge actual long term exposure these assays would need to be conducted quite frequently, potentially diminishing attendance. (b) We argue that pill counts are preferable to recording the dose prescribed because it more sensibly assesses what the patients took. (c) Pill counts should not be differentially biased; while they may introduce some errors due to inaccuracies, these should be broadly similar for the two groups since treatment and treatment rules are the same. (d) The pill counts that we have proposed will be based on blister packs, a device that helps the patient to take the required dose and helps investigators to calculate the patient’s compliance. (e) Many studies have reported that pill counts correlate well with treatment prescriptions, for example in hypertension where the disease is silent and often without symptoms.\textsuperscript{1,7,45} Furthermore, unlike hypertensive patients, those with myasthenia gravis have another medication reminder, namely the occurrence of symptoms. This provides a built-in feedback loop for failure to take medication, since symptoms will recur and/or worsen.

9.1 Sample size

A composite primary endpoint was not selected because it could lead to inconsistent results. For example, ETTX could have an improved clinical QMG score, but if this was attributable to using more prednisone then no clear decision of benefit could be made. Further, the time to first MM status would place undue emphasis on initial control and not long term control of the disease. However, while we have argued above that total exposure to prednisone measured by pill counts is essential, a conditional approach in analyzing this outcome is feasible and addresses the questions we are asking.

We therefore propose a two stage approach that first asks about differences in the clinical outcome based on the Area Under the QMG weakness score time curve (AUQMG) (a clinical measure) and then uses the Area Under the prednisone Dose Time Curve (AUDTC) to answer the first two of our primary questions. We start by estimating the difference in the mean of the clinical outcomes, based on the AUQMG, to validate our supposition that each of the two treatments (thymectomy plus prednisone or prednisone alone) results in a comparable clinical status. In the second stage we perform a test to determine whether the amount of drug in each treatment group necessary to achieve this comparable clinical status is different. Thus, stage 1 estimates the differences found between the prednisone alone versus the ETTX plus prednisone group. These estimates will provide strong statistical evidence as to whether the two treatments are comparable, inferior or superior based on the QMG scores. We will assess these differences by constructing a 99.5% confidence interval for the difference (ETTX plus prednisone group minus prednisone alone group) in the mean AUQMG scores. If this confidence interval contains zero, we conclude that the clinical score for the two treatment groups are comparable. Otherwise, we conclude that one treatment is superior to the other in the clinical score, that is, ETTX plus prednisone is superior (inferior) to prednisone alone if the interval is negative (positive). We then move to stage 2 to determine the better treatment based on the exposure to prednisone using a two sided t-test of the AUDTC with a Type I error of 0.05 conditional on the results of stage 1. The null hypothesis for this conditional test is that the mean AUDTC for the thymectomy plus prednisone Group is equal to the prednisone alone Group. For instance, if the confidence interval for the difference in the
The AUQMG score contains zero so that the average clinical scores for the two groups can be considered comparable, then the two-tailed test of AUDTC will give us the following conclusions. If we fail to reject the null hypothesis of stage 2 then prednisone alone is better than ETTX plus prednisone because one could argue that surgery is likely unnecessary as prednisone alone only was able to achieve the same results. If we reject the null hypothesis because the average AUDTC is lower in the medicine only group than the ETTX plus prednisone group, then prednisone alone is again the better treatment. Finally, if we reject the null hypothesis because the average AUDTC is lower for the surgery group, then ETTX plus prednisone is a better treatment than prednisone alone. The outcomes of the two stage testing are shown in Table 2 below.

Note: While this strategy is equivalent to a formal test of a null hypothesis of equality, we have not used a hypothesis testing approach because we would proceed to stage 2 irrespective of the outcome of such a test, and the information necessary to interpret the results of stage 2 is merely the size of the difference established by the stage 1 analysis, if such a difference exists. Thus, we state that the results of stage 2 are conditional on the results derived from stage 1. Further, we rejected the null hypothesis by using a formal test of equivalence, which requires specifying a tolerable difference between the treatment groups and then having two one sided tests of inferiority rejected to support a formal claim of equivalence. Investigations into this approach showed that we could fail to establish equivalence when the AUQMG between the two groups differed with a p-value of less than 0.005 (with our proposed sample size) and with a magnitude of difference that would be clinically relevant. Since we have no need to formally establish equivalence on the AUQMG, the variable used to drive drug treatment, we chose to estimate the differences and interpret the results of the formal decision making test (Stage 2: AUDTC) conditional on the findings from Stage 1.

Table 2: Study Outcomes
where P = prednisone alone, TX+P = ETTX plus prednisone

<table>
<thead>
<tr>
<th>2 sided Hypothesis Test of the AUDTC</th>
<th>Stage 1: Confidence Interval Results (TX+P minus P)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 2</strong> H₀: P=TX+P</td>
<td>99.5% CI above 0 P better QMG than TX+P</td>
</tr>
<tr>
<td>Reject H₀</td>
<td>99.5% CI contains 0 P and TX+P QMG comparable</td>
</tr>
<tr>
<td>P&gt;TX+P</td>
<td>99.5% CI below 0 TX+P better QMG than P</td>
</tr>
<tr>
<td>Inconclusive study look at TSAEs</td>
<td>TX+P recommended</td>
</tr>
<tr>
<td>P recommended</td>
<td>P recommended</td>
</tr>
<tr>
<td>Inconclusive study look at TSAEs</td>
<td>TX+P recommended</td>
</tr>
<tr>
<td>Reject H₀ TX+P &gt; P</td>
<td>P recommended</td>
</tr>
<tr>
<td>Inconclusive study look at TSAEs</td>
<td>P recommended</td>
</tr>
</tbody>
</table>

Finally, we confirm the consequences of our effectiveness analyses by examining the secondary endpoints, especially the Trial Specific Adverse Events and Serious Adverse Events.

9.1.1 Two-tailed hypothesis
The sample size for this trial was estimated based on the AUDTC for prednisone over 3 years. A reduction of the mean AUDTC of 30% or greater in favor of one treatment compared to the other would be clinically valuable. The sample size calculations are based on the analysis being a two-group comparison of the mean AUDTC at 3 years. The sample size calculation assumes a two-
group comparison of the treatment means, with the distribution of the AUDTC values assumed to be approximately normal. This assumption was satisfactorily tested in the Palace Trial \cite{Palace2005} and the standard deviations of the AUDTC values at 24 months were obtained from the non-azathioprine group.

Data from that study on the AUDTC (the prednisolone + placebo group in the earlier study) showed a mean that is nearly 3 times its standard deviation. To be conservative, the sample size was calculated assuming a mean/SD ratio of 2.0 which is close to the value obtained using the SD from the previous trial pooled from both treatment groups. \cite{Palace2005} On this basis for 90% power to obtain a significant result at the 5% two-tailed level, if the true difference between the treatment effects on the AUDTC at 3 years is 30% of the baseline mean, the trial requires 60 subjects in each arm or 120 total. Inflating this for up to 20% dropouts over 3 years yields a sample size of 75 per group. In order to provide a reasonable test of the effect for those pre-medicated with prednisone versus those not on prednisone, we have increased the sample size to 100 per group. This allows for approximately 50 patients in the subgroup pretreated with prednisone and not pretreated within each treatment arm. For the same difference, 50 patients would yield a power of 84% within each subgroup.

We use a 99.5% confidence level for the first stage, and apply a Type I error level of 5% in the second stage of drug exposure testing. No penalty is taken for the Type I error as the result is a conditional result – that is, conditional in that the ability to control the patients clinically is indeed equivalent or better for one group, and that the relevant clinical question is then the amount and risk of drug that is necessary to achieve that.

We have considered a simultaneous approach to testing both outcomes, but feel that the sequential conditional approach most directly reflects the clinical issues in comparing these treatments.

In addition to the primary analyses above, we will perform secondary analyses that assess the results at 36 months, and will use longitudinal methods to assess the measures at each point in time by treatment condition to assess the overall pattern of response and adherence. Importantly, secondary outcomes will now include a formal comparison by gender of the occurrence of any of 27 adverse symptoms (TSASs) that can occur with prednisone, azathioprine and cyclosporin therapy, the three agents that study subjects may receive in the trial. These adverse events will be evaluated by the patients for symptom frequency and symptom distress questionnaire, \cite{TSAS2007} together with quality of life measures and clinical indicators (weight gain, fasting blood sugar).

**9.1.2 Is a 30% decrease in steroid use satisfactory?**

Our figure of 30% was a consensus reached by participating PIs representing MG specialists worldwide. At the individual level, we fully understand that a patient with a maximum starting alternate day (AD) dose of 100 mg who reduces only 30% is still at a high dose. However, we are looking at the average amount of prednisone over time in the two treatment groups. For example, in the comparable study by Palace et al., \cite{Palace2005} the average dose at baseline was around 90 mg declining to 45 mg at 1 year and 40 mg at 3 years in the Prednisone plus Placebo (PP) group. In the active arm (Prednisone plus azathioprine), the baselines were comparable declining to 37.5 mg at year 1 and to 0 at year 3. The estimated average AUDTC for the placebo group was 152.5 compared to 101.25, a difference of 33.6% for the study groups, but representing a 56% percent reduction in the PP group and a 100% reduction in the active arm. Thus, we feel that a 30%
decrease as a minimal decrease between the groups (not within a group or for an individual patient) is a reasonable minimal effect to detect.

9.1.3 Penalty for use of drugs in addition to steroids
We will provide for penalty doses of drugs as described above for those requiring drugs other than prednisone to control their clinical condition. While the rule is a bit complex, it doesn’t appear biased between the two groups unless there is a differential use of these drugs across treatment arms. We will do a sensitivity analysis and can ensure that we are not biasing the results. Using the highest dose of prednisone, as was suggested, is one option and we will adopt this for consistency sake and accept the advice of the NINDS review committee. Using the algorithm we discussed will lower the dose applied for the area under the curve and we will also assess the treatments in this manner to be sure that we are not endorsing a treatment in our now, clinically conditional approach (that is, we have already demonstrated comparability of the two treatment or superiority of one treatment group only on clinical grounds when testing dosage results) because of the assumptions we are making about ancillary drug use. To declare superiority at this stage, there must not be more auxiliary drugs used by the preferred treatment.

9.2 Interim analyses
Stopping early, even for an extremely small p-value, would not be a likely or desired decision for the investigative team. Suppose, as was theorized by the review committee, that a large difference existed during the first 12 months, would we as investigators wish to stop the study, even if the one arm was strikingly better? The answer is no, for at least two reasons: symptom reduction may not last and, while there could be a short term benefit, the long term consequences are important to the decision. Nevertheless, we understand the unease that a DSMB would have without rules. Thus, we propose to use Haybittle-Peto boundaries for the DSMB’s efficacy inspection. We propose a p=0.0005, such that the DSMB can take up to 5 interim inspections of the data and the final p-value will remain virtually unaffected by the interim monitoring.

9.3 Statistical methods to be used in analysis
All demographic and basic clinical data will be summarized in order to characterize the study population. The techniques used to summarize this data will utilize basic summary statistics such as percentages, means, standard deviations, medians, and ranges.

9.3.1 Analysis populations
Every effort will be made to keep subjects in the study for the entire follow-up period of evaluations. Even if they choose not to follow the prescribed medication, they will be asked to return for QMG measurements and characterization by side effects etc. Based on the understanding that this is not always possible, the following populations will be formed for the purpose of data analysis:
An intent-to-treat cohort, defined as all subjects who were recruited and randomized into the study, will be used for the primary analyses. Patients withdrawn for adverse reactions and/or dropout will be considered as failures and evaluated as such in this assessment. Safety will be monitored at both the group level and the individual level.

9.3.2 Statistical analysis
Analyses will commence in stages during the trial. Initially, we will be concerned with recruitment rates, the characteristics of patients entered versus those not enrolled and the similarity of enrollment over the sites. Baseline comparisons between the two groups, the quality control and assurance of the data collection, management and responsiveness to edits and corrections, will be conducted. The characterization of the cohort through univariate and multivariate (subgroups)
analyses will be conducted and displayed. Safety monitoring will commence from day 0 and appropriate tables and summarizations of this data will be routinely presented to the DSMB.

The definitive analyses will be performed using the “intention-to-treat” principle. The primary outcome measure for this study is a stepwise assessment of response for each patient based on the clinical response (AUQMG) and the risk of exposure to prednisone (AUDTC) from day 0 to month 36. The AUQMG will be computed from the standardized blinded QMG scorings. The AUDTC will be computed from the pill counts with a second AUDTC computed from prescribed dose. The prescribed dose AUDTC is a secondary analysis, but avoids the concerns expressed regarding pill counts and will be used to affirm the results and conclusion, as a sensitivity analysis against the pill counts. In terms of the AUDTC, previous results have shown this sum to be reasonably normally distributed, but the assumptions of normality will be assessed. Similarly, the AUQMG should be approximately normal, again resulting from the averaging of scores declining over time and linked by protocol to the AUDTC. In addition, the frequency of Serious Adverse events (see Section 8.2) and the number of TSAEs will be summed over the 36 month interval.

As noted above, we will perform a sensitivity analysis replacing the AUDTC based on pill counts by the same quantity based on prescribed drug. We will examine differences and compare the average difference between these two measures (pill count based versus prescribed) by treatment group based on actual deviations and the variability within groups of these two measures. In the event that this affects the conclusions, these differences will be discussed and considered as to which more accurately represents the trial results. We do not expect this to happen, but are sensitive to the concerns and weaknesses of pill counts.

Owing to missed follow-up visits or other reasons, there may be gaps in some of the individual sequences needed to calculate the AUDTC and AUQMG. In this case, the AUDTC will be calculated as an average for the specified time period and entered into the analysis with weights appropriate to the number of points used in its calculation. We will also use longitudinal data methods to impute this information over time, but because of the complexity of such analyses and the clinical need to discuss the results in a summarized manner, we propose to use these methods as confirmatory.

We will compare the frequency of other drug use between the two arms. If this is not statistically significant, we analyze the data on the AUDTC without a penalty dose (discussed below). We will pick up the TSAEs caused by the prednisone which led to the lowering and alteration of planned therapy enabling us to include some measure of the impact of dosage adjustment and additional drugs via this secondary outcome measure. In addition, the other medications have potential adverse effects and these will be included in the TSAEs (Appendix II) and counted under our intent to treat principle.

Two confirmatory analyses will be conducted. First, covariates will be added to our test of the hypothesis to account for the amount of other drugs utilized to examine the sensitivity of the results to these additions. Second, we will perform a confirmatory analysis using a penalty given for the use of other active agents, lest a lower dose of prednisone (for adverse effects or other reasons) could provide inappropriate information. Thus, in computing the AUDTC when an alternative drug is used (azathioprine, cyclosporine, etc.), the dosage of prednisone to be used in calculation of the AUDTC is the prescribed dose of prednisone averaged over the last two visits prior to the prescription of the alternative drug. While considered by the previous NINDS review as complicated, we used the last two doses to ensure representation of more than what may have been a lower dose that was insufficiently effective; we were concerned that the dose should be
representative. In this revision, we have also counted the TSAEs that resulted in these patients having to receive these additional agents. Further, we will use prescribed dose rather than actual dose via the pill count, because the patient may lower compliance because of adverse effects; this means that the dosage penalty is independent of compliance.

This penalized measure is independent of the choice of alternative drugs and treatment group, since all patients after the initial surgery are following the same protocol. As noted above, we will examine the frequency of occurrence of other drug use, which may involve preferences, training, formulary preferences, etc. that also may cluster by site, but it should not be biased towards either treatment arm. This scoring allows reasonable prednisone values to be inserted into the calculation irrespective of alternative drug choice and without arbitrary assignment of prednisone equivalents. Further, it is the best representation of where a person was when alternative drugs were needed.

The overall analysis will use a general linear model (SAS PROC MIXED) nesting patients within clinical centers and assessing the treatment effects. This analysis will enable testing of the subgroup hypotheses and if violations of model assumptions are present (i.e. excess skewness), suitable data transformations will be done prior to the analyses. To ensure no potential bias can enter due to the MC examinations over the first 3 months and potentially less prednisone stemming from a waiting period for surgery, all analyses on the primary endpoint will be repeated using month 4 to 36 AUDTC. These results must be the same or stronger to declare a positive result.

Secondary analyses include: (1) Individual AUQMG, TSAEs and AUDTC from day 0 to month 12 and to month 24, and (2) actual prednisone dose at month 36, total TSAEs and 36 month QMG or MM status. These can be tested using the same models as subsets. In addition, these methods enable the longitudinal assessments of both the prescribed dosages and pill counts, where each point in time is used within a patient. This longitudinal analysis may provide increased understanding of patterns over time that may enhance our understanding of the primary hypothesis. For example, while unlikely, the AUDTC could show no difference between the two treatment groups, because of the presence of a time by treatment interaction. These secondary analyses will ensure that any conclusions consider the potential for such differing trends over time.

Cox proportional hazards models will be used to evaluate the time from day 0 to reach initial MM status. Two specific hypotheses are of interest: first, does ETTX decrease the time until initial MM status and, a key second one, does the use of prednisone prior to surgery in combination with thymectomy reduce the time until MM status is achieved and does it impact on the duration of MM status over the trial. For this analysis we will calculate the proportion of time spent in MM status for each treatment group. This analysis coupled with the primary dosage hypothesis is also key because the duration of time spent in MM status should also be as good or better if one treatment is superior to another. While this is a similar measure, for each group the probability of cycling in and out of MM status will be computed and compared among clinics and overall. Cross sectional analyses will examine MM status at months 12, 24 and 36 using logistic regression, and to enable the use of data which minimizes the impact of dropouts or censored observations, generalized estimating equations (PROC GENMOD in SAS) will be used to assess the achievement of MM status over time adjusting for the clustering of these results within patient. The frequency of the TSAEs and the time until TSAEs occur will be examined from randomization (which will include the surgical impacts) and with a guarantee time to examine the pattern of TSAEs that occur in course of therapy after the impact of the surgery.
Lower dosages of prednisone and increased time spent in MM status should be meaningful to the patients. These should be reflected in their self reported quality of life (SF-36, MG-ADL) and symptom checklists (TSAS score). These will be tested using our quality of life assessment tool (SF-36) and the symptom checklists at months 12, 24 and 36 using generalized linear models. We expect improvements in the subscales of the SF-36, especially the physical component. All of these should also be clinically demonstrated with greater improvements in measured function and performance. These will be compared using change in the QMG and improvements in the MG-ADL over time and at months 12, 24 and 36 using similar techniques as above.

If we can demonstrate that patient related factors are improved for ETTX over drugs alone, we want to look at the costs to the system to achieve this benefit. This would encompass indicators of cost as we are not attempting a cost benefit analysis due to the complexity of collecting such data in a multinational trial. Nevertheless, we will compare the cumulative number of days in hospital for treatment of, or complications related to, MG by months 24 and 36; the number of plasma exchanges and IVIG infusions, and total dose of IVIG from day 0 to month 36.

Additional analyses will be conducted comparing the extent of the thymectomy and other parameters assessed from the surgical notes and scoring of the surgeries. These include within the ETTX group analyses of predictors of overall AUDTC from the surgical information, and predictors of response within group from other baseline characteristics. This cohort will also represent one of the largest series of prospectively treated MG patients ever assembled and, as such, will provide a valuable database resource that, consistent with NIH Policy, will be made public with adequate documentation for its use in future research.

9.3.3 Additional safety analyses
All clinical adverse events and laboratory abnormalities will be evaluated for safety during the 6 month combination therapy period.

9.3.4 Laboratory abnormalities
Laboratory evaluations will be assessed to determine incidence of clinically notable abnormalities that emerge during the course of the study.

10.0 Ethical requirements

10.1 Declaration of Helsinki
See Appendix VIII

10.2 Subject information and consent
Prior to any testing under this protocol, including screening tests and evaluations, written informed consent must be obtained from the subject in accordance with local practice and regulations.

Whenever possible, the treating neurologist (usually the Center PI) will also be involved in this procedure. The background of the proposed study and the benefits and risks of the procedures and study will be explained to the subject. A copy of the informed consent document signed and dated by the subject must be given to the subject. Confirmation of a subject’s informed consent must also be documented in the subject’s medical records prior to any testing under this protocol, including screening tests and evaluations.

10.3 Maintenance of study documentation and supplies
The study will be conducted according to the Good Clinical Practices as outlined by the Food and Drug Administration. It is the responsibility of the investigator to maintain a comprehensive and centralized filing system of all study relevant documentation. Such documentation includes:

1. Case Report Forms (CRF)

CRFs for individual subjects will be entered via a web-based system and CRFs will be generated off this system as provided by the Data Coordinating Center. CRF’s are used to record study data and are an integral part of the study and subsequent reports. Therefore, the CRFs must be legible and complete.

All forms should be filled out using black ballpoint pen. Errors should be lined out but not obliterated, and the correction inserted, initialed and dated by designated study personnel. Further data corrections will be performed on special “Data Clarification Forms” (DCF’s). Such DCF’s will be dispatched by the Data Coordinating Center. After the Center PI makes the correction on the CRF, a copy of it will be kept in the CRF book and the original entered into the system.

A CRF must be completed and signed by the Center PI for each subject enrolled, including those removed from the study for any reason. The reason for removal must be noted on the study conclusion CRF by the investigator for each subject.

CRFs must be kept current to reflect the subject’s status at each phase during the course of the study. Subjects are not to be identified on CRFs by name; appropriately coded identification must be used. The Center PI must keep a separate log of the subjects’ names and addresses.

Because of the potential for errors, inaccuracies, and illegibility in transcribing data onto CRFs, originals of laboratory and other test results must be kept on file with the subject’s CRF. CRFs and copies of test results must be available at all times for inspection by the Data Coordinating Center or study monitor.

2. Subject Hospital Files - which substantiate the data entered in the CRFs with regard to lab data, subject history, treatment regimens, subject follow-up, etc.

3. Subject Exclusion Record - which should reflect the reason any subject screened for the study was found to be ineligible.

4. Subject Failure-to-Screen Record – which should document why screening for the study was not undertaken

5. Subject Identification Record - which should allow linking of subject study number and subject name and date of birth.

6. Drug Dispensing Log - which will reflect the total amount of study medication received, and the amount administered to each subject. This information must agree with the information entered in the Drug Accountability CRF.
7. Informed Consent Forms - which must be available for each subject and be verified for proper documentation.

11.0 Further requirements and general information

11.1 Changes to final study protocol
All protocol amendments must be submitted to the IRB/REB/EC. Protocol modifications that impact on subject safety, the scope of the investigation, or affect the scientific quality of the study must be approved by the IRB/REB/EC and submitted to the appropriate regulatory authorities before implementation of such modifications to the conduct of the study. However, the Data Coordinating Center, at any time, may amend this protocol to eliminate an apparent immediate hazard to a subject. In this case, the appropriate regulatory authorities will be subsequently notified. In the event of a protocol modification, the subject consent form may require similar modifications.

11.2 Record retention
The Institution shall preserve all records relating to the Trial and the subjects participating therein as required by the applicable Protocol and otherwise in accordance with good clinical practices, e.g., a minimum of fifteen (15) years after the completion or termination of the Trial, and thereafter shall offer such records to Data Coordinating Center and/or NINDS before destroying or disposing thereof.

11.3 Protocol completion
The IRB/REB/EC must be notified of completion or termination of the protocol. Within 3 months of protocol completion or termination, the Center PI must provide a final clinical summary report to the IRB/REB/EC. The Principal Investigator at each center must maintain an accurate and complete record of all submissions made to the IRB/REB/EC, including a list of all reports and documents submitted.

12.0 Principal Investigator’s and sponsor’s agreement
I have carefully read and thus have understood the provisions of this protocol, and am prepared to follow it in every detail in the conduct of this study.

____________________________________  ____________
Principal Investigator                                                                    Date

_______________________________________  ____________
Study Chairman                                                                            Date
REFERENCES


### Quantitative MG Weakness Score

<table>
<thead>
<tr>
<th>TEST ITEMS WEAKNESS</th>
<th>NONE</th>
<th>MILD</th>
<th>MODERATE</th>
<th>SEVERE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double vision on lateral gaze right or left (circle one), seconds</td>
<td>61</td>
<td>11-60</td>
<td>1-10</td>
<td>Spontaneous</td>
</tr>
<tr>
<td>Ptosis (upward gaze), seconds</td>
<td>61</td>
<td>11-60</td>
<td>1-10</td>
<td>Spontaneous</td>
</tr>
<tr>
<td>Facial Muscles</td>
<td>Normal lid closure</td>
<td>Complete, weak, some resistance</td>
<td>Complete, without resistance</td>
<td>Incomplete</td>
</tr>
<tr>
<td>Swallowing 4 oz. water (1/2 cup)</td>
<td>Normal</td>
<td>Minimal coughing or throat clearing</td>
<td>Severe coughing/choking or nasal regurgitation</td>
<td>Cannot swallow (test not attempted)</td>
</tr>
<tr>
<td>Speech following counting aloud from 1-50 (onset of dysarthria)</td>
<td>None at #50</td>
<td>Dysarthria at #30-49</td>
<td>Dysarthria at #10-29</td>
<td>Dysarthria at #9</td>
</tr>
<tr>
<td>Right arm outstretched (90° sitting), seconds</td>
<td>240</td>
<td>90-239</td>
<td>10-89</td>
<td>0-9</td>
</tr>
<tr>
<td>Left arm outstretched (90° sitting), seconds</td>
<td>240</td>
<td>90-239</td>
<td>10-89</td>
<td>0-9</td>
</tr>
<tr>
<td>Vital Capacity (% predicted)</td>
<td>≥80%</td>
<td>65-79%</td>
<td>50-64%</td>
<td>&lt; 50%</td>
</tr>
<tr>
<td>Rt hand grip: male (KgW) female</td>
<td>≥45</td>
<td>15-44</td>
<td>5-14</td>
<td>0-4</td>
</tr>
<tr>
<td></td>
<td>≥30</td>
<td>10-29</td>
<td>5-9</td>
<td>0-4</td>
</tr>
<tr>
<td>Left hand grip: male (KgW) female</td>
<td>≥35</td>
<td>15-34</td>
<td>5-14</td>
<td>0-4</td>
</tr>
<tr>
<td></td>
<td>≥25</td>
<td>10-24</td>
<td>5-9</td>
<td>0-4</td>
</tr>
<tr>
<td>Head, lifted (45° supine), seconds</td>
<td>120</td>
<td>30-119</td>
<td>1-29</td>
<td>0</td>
</tr>
<tr>
<td>Right leg outstretched (45° supine), seconds</td>
<td>100</td>
<td>31-99</td>
<td>1-30</td>
<td>0</td>
</tr>
<tr>
<td>Left leg outstretched (45° supine), seconds</td>
<td>100</td>
<td>31-99</td>
<td>1-30</td>
<td>0</td>
</tr>
</tbody>
</table>

TOTAL QMG SCORE (range 0-39) _________
APPENDIX II

Trial Specific Adverse Events

Aseptic necrosis
Assisted ventilation
Bone marrow suppression requiring withdrawal of medication
Cataract (patient reports impaired vision; ophthalmological opinion to confirm)
Cyclosporin associated encephalopathy
Death due to MG
Diabetes mellitus requiring medication (oral hypoglycaemic agents or insulin)
Empyema
Fractures (number of bones)
Glaucoma (ophthalmological opinion to confirm)
Hemothorax
Herpes zoster
Hospitalization other than for ETTX and/or initiation of prednisone therapy
Hypertension (>150/90 or requiring hypotensive therapy)
Infection requiring intravenous antibiotics
Intestinal perforation
Liver function test abnormalities requiring withdrawal of medication
Lymphoma
Pancreatitis
Persistent thoracic pain (more than 4 weeks)
Pneumothorax
Prominent (keloid) scar
Rash
Recurrent nerve injury
Renal failure
Re-operation, any cause
Serious mental symptoms requiring psychiatric referral
Sleep disturbance requiring referral or treatment
Sternal dehiscence
Sternal wound infection
Tendon rupture
Thoracic duct injury
Tracheotomy
Upper gastrointestinal hemorrhage
Weight gain >7% above baseline weight at study entry (scores positive if present on two consecutive visits)
APPENDIX III

Adverse Symptoms of the therapies to be included in a questionnaire to be completed by patients. Modified from (Moons et al).

Each item is graded for (a) frequency:  
- never
- sometimes
- regularly
- almost always
- always

(b) symptom distress from 0-4 (not distressing at all to very much distressing)

- acne
- back pain
- bruises
- changed appearance
- changed taste
- decreased interest in sex
- depression
- diarrhea
- fatigue
- fragile skin
- gingival hyperplasia
- gingival swelling
- headache
- impotence/painful menstruation
- increased appetite
- increased hair growth
- inflammation
- mood swings
- moon face
- muscle weakness
- painful/inflamed/prominent scar
- palpitations
- persistent chest pain
- poor appetite
- poor concentration
- poor vision
- sleeplessness
- stomach complaints
- swollen ankles
- tremor
APPENDIX IV

MGFA Post-Intervention Status

Complete Stable Remission (CSR)  The patient has had no symptoms or signs of MG for at least 1 year and has received no therapy for MG during that time. There is no weakness of any muscle on careful examination by someone skilled in the evaluation of neuromuscular disease. Isolated weakness of eye closure is accepted.

Pharmacologic Remission (PR)  The same criteria as for CSR except that the patient continues to take some form of therapy for MG. Patients taking cholinesterase inhibitors are excluded from this category because their use suggests the presence of weakness.

Minimal Manifestation (MM)  The patient has no symptoms or functional limitations from MG but has some weakness on examination of some muscles. This class recognizes that some patients who otherwise meet the definition of CSR or PR do have weakness that is only detectable by careful examination.

MM-0: The patient has received no MG treatment for at least 1 year.
MM-1: The patient continues to receive some form of immunosuppression but no cholinesterase inhibitors or other symptomatic therapy.
MM-2: The patient has received only low dose cholinesterase inhibitors (< 120 mg pyridostigmine per day), for at least 1 year.
MM-3: Patient has received cholinesterase inhibitors or other symptomatic therapy and some form of immunosuppression during the past year.

Change in Status

Improved (I)  A substantial decrease in pretreatment clinical manifestations or a sustained substantial reduction in MG medications as defined in the protocol. In prospective studies, this should be defined as a specific decrease in QMG score.

Unchanged (U)  No substantial change in pretreatment clinical manifestations or reduction in MG medications as defined in the protocol. In prospective studies, this should be defined in terms of a maximum change in score.

Worse (W)  A substantial increase in pretreatment clinical manifestations or a substantial increase in MG medications as defined in the protocol. In prospective studies, this should be defined as a specific increase in QMG score.

Exacerbation (E)  Patients who have fulfilled criteria for CSR, PR, or MM but subsequently developed clinical findings greater than permitted by these criteria.
| Died of MG (D of MG) | Patients who died of MG, of complications of MG therapy, or within 30 days after thymectomy. List the cause (see Morbidity and Mortality table) |
APPENDIX V

MGFA Clinical Classification

Class I: Any ocular muscle weakness; may have weakness of eye closure. All other muscle strength is normal.

Class II: Mild weakness affecting muscles other than ocular muscles; may also have ocular muscle weakness of any severity.
  IIa. Predominantly affecting limb, axial muscles, or both. May also have lesser involvement of oropharyngeal muscles.
  IIb. Predominantly affecting oropharyngeal, respiratory muscles, or both. May also have lesser or equal involvement of limb, axial muscles, or both.

Class III: Moderate weakness affecting muscles other than ocular muscles; may also have ocular muscle weakness of any severity.
  IIIa. Predominantly affecting limb, axial muscles, or both. May also have lesser involvement of oropharyngeal muscles.
  IIIb. Predominantly affecting oropharyngeal, respiratory muscles, or both. May also have lesser or equal involvement of limb, axial muscles, or both.

Class IV: Severe weakness affecting muscles other than ocular muscles; may also have ocular muscle weakness of any severity.
  IVa. Predominantly affecting limb, axial muscles, or both. May also have lesser involvement of oropharyngeal muscles.
  IVb. Predominantly affecting oropharyngeal, respiratory muscles, or both. May also have lesser or equal involvement of limb, axial muscles, or both.

Class V: Defined as intubation, with or without mechanical ventilation, except when employed during routine postoperative management. The use of a feeding tube without intubation places the patient in class IVb.
APPENDIX VI

Data Management and Quality Control:
The focus of the biostatistical and data management activities in the DCC include quantitative analytic areas comprised of database design and management, experimental design, sampling, quality control, statistical analysis for the planning of the project, specific analyses, cost effectiveness analyses, and manuscript preparation. More specifically, we will:

a. Manage data obtained from the Clinical Sites, and the photographs from the surgeries and/or auxiliary data collected for specific individual ancillary projects that may be developed as part of this Clinical Trial. When data are transmitted from the Clinical Sites electronically (or via transfer of paper forms as a backup), they will be managed in a timely fashion by the DCC. The electronic transfer of data is already functioning as part of several ongoing clinical trials including SPF3, funded by NINDS, the TEAR trial, funded by Amgen, and CombiRX, funded by NINDS. The DCC will enhance and apply quality control algorithms for verifying receipt, entry, and transmission specific to this trial. The DCC will document any errors and ensure that error messages are generated; when corrections are not made in a timely fashion, emails will be sent and, if necessary, calls made to adjudicate correctable problems. Errors will be identified, even if uncorrectable, to educate the Clinical Sites on proper forms completion. This will be in addition to the summary reports periodically generated and reviewed.

b. Maintain integrity, confidentiality, and security of all data for the overall project. Our web-based system is HIPAA compliant and no access is initiated by the DCC. We have no direct access to the systems on site and all communication is initiated by the center. No personal identifiers such as name, initials or birth date appear on any forms or in any transmissions; only a study ID with appropriate check digit are used in communication. Age and birth date are collected for eligibility, but only age is transmitted to the DCC in order to maintain HIPAA compliance.

c. Assume responsibility for implementation of quality control procedures and periodic review of all data transmitted to the DCC by the Clinical Sites.

d. Assist investigators and the Steering Committee in ensuring appropriate methods of experimental design, data collection, processing, monitoring, statistical analysis, and presentation of results during the course of the overall project.

e. Work with the research team in the effectiveness and cost effectiveness assessments and with the Executive Committee on larger issues as appropriate.

f. Collaborate with the investigative teams, including those from NINDS, in preparing periodic project summary reports and research papers for publication.

g. Develop and implement long-term storage protocols for securing critical documents and all machine readable databases.

h. Develop a public distribution mechanism for the completion of the study to adhere to the NINDS mandate that all research data be placed in the public domain; we have actively discussed and planned for satisfaction of this obligation.
Functions performed by the DCC personnel includes: management of a web-based (using secure socket technology) randomization system that only allows randomization after online documentation of eligibility and baseline completion of required information, management of an online electronic and email notification system of serious adverse events, tracking of drugs, compliance and pill counts, inventories, and transfer of routine visit data in real time from all locations. From this, the DCC will be monitoring recruitment and adherence to study protocols, as outlined above, enabling this multinational study without the need for large personnel costs and extensive personnel. The DCC is both the study police and cheerleader and will seek to balance these two roles to maintain the quality and enthusiasm necessary for a trial that will run longer than most MG clinical trials to date.

Systems for data transmission will be employed to facilitate timely key information transfer and provide the least burden on the clinics, while maintaining patient confidentiality. We will have direct control over the randomization process, and will utilize a web-based system, centralizing the randomization and data collection, while distributing the labor. This data entry burden at the site level is quite low given the number of patients at any one site. This system, which is being used currently in several trials, will enable the DCC to have up-to-date information on recruitment, randomization, drug accountability, and patient tracking and follow-up.

The forms will be received on site before entry, and the staff will be aware that this is a key important step before the patients leave the clinic. The online system performs quality control checks (range, logical consistency, and cross form validation) at the time of entry and is managed by the DCC. Ideally, the forms will be entered within 24 hours of a visit and monitoring reports will be generated among clinical sites to provide performance feedback. To ensure the highest quality forms, a structured process will be used prior to implementation. The forms developed will be circulated between the sites (the Investigators, their staff, the Executive Committee, NINDS and the DCC) for final comments (and additional pilot testing, where appropriate) and ultimately approval.

Centralized training is extremely beneficial in standardizing exam techniques and procedures across centers, as well as fostering a unified, cohesive effort and attitude; this will be conducted prior to starting the trial. Although forms have been designed and pre-tested during the planning phase, an important part of the process of achieving appropriate forms and finalizing exam components is the process of allowing Centers to feel comfortable with the instruments. Centers will be required to go through a practice use of the forms on actual patients after they have been trained; the resulting feedback for this practice run is an important part of the implementation process. We will assess the performance of the instruments and help identify any final changes that may improve their performance.

Many of the forms have been used in other studies conducted by the team; others have been adapted or taken from existing studies. Nevertheless, it is important that all Centers be familiar with the complete package of forms, the processes of randomization, arranging for surgery, drug dispensing and dose titration, reporting, correcting and obtaining feedback. The final protocol and forms will also be presented to the DSMB and NINDS for approval before starting actual recruitment.

In an effort to create an awareness of internal and external quality control measures, a special session designed to educate all exam personnel in the importance of quality control will be incorporated into the centralized training. There are three main purposes for quality control: to prevent errors from occurring, to detect and eliminate errors that have occurred or are occurring,
and to document the quality of data already collected. Each of the actions we plan for quality control will address at least one of these three purposes. Preventing errors will be the primary focus of the centralized training. Effective protocols, forms that minimize the potential for errors, efficient data collection, and dedicated training sessions will provide the primary tools of prevention. The forms will serve as the primary source documents and the data entry system is designed to ensure appropriate form completion.

Clinics will understand that data quality is their responsibility and they will be trained to maximize the quality of recording and to check and recheck data before they are entered and transmitted. To insure adequate entry, we will periodically review a sample of their work by having forms faxed to Birmingham, double entered here, and compared to the transmitted data files; error rates must be below 5%. Quality and timeliness will be reinforced with monitoring visits. After the record undergoes the Clinical Site’s quality control procedures, inculcated during initial training and certification meetings, the records entered, and the site and systems prompts satisfied, the data will be released for transfer. Records will be sent immediately (the system polls every few minutes, when online, back to the local transmissions) to the DCC. Once the DCC receives the information, centralized data editing will also be conducted on participant records, to ensure that all data collected are consistent via dates, missing components, auxiliary forms, etc. All range checks and skip logic that need exceptions require a password obtained from the DCC. After a form is “released,” changes and exceptions to the database will only be allowed with a password (which changes for each exception, etc.). These will be obtained via email to the DCC, allowing for all exceptions and overrides to be completely documented. This system has been implemented and carefully documented to ensure that such exceptions are confirmed by the Clinical Center. Periodically, project reports tracking accrual, quality control, follow-up, and baseline analyses of participant characteristics will be distributed. Analysis of the information will be presented and discussed in cooperation with, and on the timetable set by, the Executive Committee, NINDS and/or the DSMB. The statisticians, clinicians and database programmers will design and prepare reports in collaboration with the Executive Committee, NINDS and/or the DSMB.

The System has been developed to include multiple redundancies in data backup and archiving procedures, as well as handle problems in computer communications. The DCC will be responsible for ensuring adherence to standard protocols for confidentiality and off-site storage, and the maintenance of data for disaster recovery. Complete documentation of the database management system is, and will be, carried out.

APPENDIX VII

MG Activities of Daily Living Score
### Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talking</td>
<td>Normal</td>
<td>Intermittent slurring or nasal speech</td>
<td>Constant slurring or nasal, but can be understood</td>
<td>Difficult to understand speech</td>
<td></td>
</tr>
<tr>
<td>Chewing</td>
<td>Normal</td>
<td>Fatigue with solid food</td>
<td>Fatigue with soft food</td>
<td>Gastric tube</td>
<td></td>
</tr>
<tr>
<td>Swallowing</td>
<td>Normal</td>
<td>Rare episode of choking</td>
<td>Frequent choking necessitating changes in diet</td>
<td>Gastric tube</td>
<td></td>
</tr>
<tr>
<td>Breathing</td>
<td>Normal</td>
<td>Shortness of breath with exertion</td>
<td>Shortness of breath at rest</td>
<td>Ventilator dependence</td>
<td></td>
</tr>
<tr>
<td>Impairment of ability to brush teeth or comb hair</td>
<td>None</td>
<td>Extra effort, but no rest periods needed</td>
<td>Rest periods needed</td>
<td>Cannot do one of these functions</td>
<td></td>
</tr>
<tr>
<td>Impairment of ability to arise from a chair</td>
<td>None</td>
<td>Mild, sometimes uses arms</td>
<td>Moderate, always uses arms</td>
<td>Severe, requires assistance</td>
<td></td>
</tr>
<tr>
<td>Double vision</td>
<td>None</td>
<td>Occurs, but not daily</td>
<td>Daily, but not constant</td>
<td>Constant</td>
<td></td>
</tr>
<tr>
<td>Eyelid droop</td>
<td>None</td>
<td>Occurs, but not daily</td>
<td>Daily, but not constant</td>
<td>Constant</td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL MG-ADL SCORE** (range 0-24) ________

### APPENDIX VIII

DECLARATION OF HELSINKI
INTRODUCTION

It is the mission of the physician to safeguard the health of the people. His or her knowledge and conscience are dedicated to the fulfillment of this mission.

The Declaration of Geneva of the World Medical Association binds the physician with the words: "The health of my patient will be my first consideration", and the International Code of Medical Ethics declares that "A physician shall act only in the patient's interest when providing medical care which might have the effect of weakening the physical and mental condition of the patient".

The purpose of biomedical research involving human subjects must be to improve diagnostic, therapeutic and prophylactic procedures and the understanding of the etiology and pathogenesis of disease.

In current medical practice most diagnostic, therapeutic or prophylactic procedures involve hazards. This applies especially to biomedical research.

Medical progress is based on research which ultimately must rest in part on experimentation involving human subjects.

In the field of biomedical research a fundamental distinction must be recognized between medical research in which the aim is essentially diagnostic or therapeutic for a patient, and medical research, the essential object of which is purely scientific and without implying direct diagnostic or therapeutic value to the person subjected to the research.

Special caution must be exercised in the conduct of research, which may affect the environment, and the welfare of animals used for research must be respected. Because it is essential that the results of laboratory experiments be applied to human beings to further scientific knowledge and to help suffering humanity, the World Medical Association has prepared the following recommendations as a guide to every physician in biomedical research involving human subjects. They should be kept under review in the future. It must be stressed that the standards as drafted are only a guide to physicians all over the world. Physicians are not relieved from criminal, civil and ethical responsibilities under the laws of their own countries.

Basic Principles

Biomedical research involving human subjects must conform to generally accepted scientific principles and should be based on adequately performed laboratory and animal experimentation and on a thorough knowledge of the scientific literature.

The design and performance of each experimental procedure involving human subjects
should be clearly formulated in an experimental protocol which should be transmitted for consideration, comment and guidance to a specially appointed committee independent of the investigator and the sponsor provided that this independent committee is in conformity with the laws and regulations of the country in which the research experiment is performed.

Biomedical research involving human subjects should be conducted only by scientifically qualified persons and under the supervision of a clinically competent medical person. The responsibility for the human subject must always rest with a medically qualified person and never rest on the subject of the research, even though the subject has given his or her consent.

Biomedical research involving human subjects cannot legitimately be carried out unless the importance of the objective is in proportion to the inherent risk to the subject.

Every biomedical research project involving human subjects should be preceded by careful assessment of predictable risks in comparison with foreseeable benefits to the subject or to others. Concern for the interests of the subject must always prevail over the interests of science and society.

The right of the research subject to safeguard his or her integrity must always be respected. Every precaution should be taken to respect the privacy of the subject and to minimize the impact of the study on the subject's physical and mental integrity and on the personality of the subject.

Physicians should abstain from engaging in research projects involving human subjects unless they are satisfied that the hazards involved are believed to be predictable. Physicians should cease any investigation if the hazards are found to outweigh the potential benefits.

In publication of the results of his or her research, the physician is obliged to preserve the accuracy of the results. Reports of experimentation not in accordance with the principles laid down in this Declaration should not be accepted for publications.

In any research on human beings, each potential subject must be adequately informed of the aims, methods, anticipated benefits and potential hazards of the study and the discomfort it may entail. He or she should be informed that he or she is at liberty to abstain from participation in the study and that he or she is free to withdraw his or her consent to participation at any time. The physician should then obtain the subject's freely given informed consent, preferably in writing.

When obtaining informed consent for the research project the physician should be particularly cautious if the subject is in a dependent relationship to him or her or may consent under duress. In that case, the informed consent should be obtained by a physician who is not engaged in the investigation and who is completely independent of this official relationship.

In the case of legal incompetence, informed consent should be obtained from the legal guardian in accordance with national legislation. Where physical or mental incapacity makes it impossible to obtain informed consent, or when the subject is a minor, permission from the responsible relative replaces that of the subject in accordance with national
Whenever the minor child is in fact able to give a consent, the minor's consent must be obtained in addition to the consent of the minor's legal guardian.

The research protocol should always contain a statement of the ethical considerations involved and should indicate that the principles enunciated in the present Declaration are complied with.

Medical Research Combined With Professional Care (Clinical Research)

In the treatment of the sick person, the physician must be free to use a new diagnostic and therapeutic measure, if in his or her judgment it offers hope of saving life, reestablishing health or alleviating suffering.

The potential benefits, hazards and discomfort of a new method should be weighed against the advantages of the best current diagnostic and therapeutic methods.

In any medical study, every patient - including those of a control group, if any - should be assured of the best proven diagnostic and therapeutic methods.

The refusal of the patient to participate in a study must never interfere with the physician patient relationship.

If the physician considers it essential not to obtain informed consent, the specific reasons for this proposal should be stated in the experimental protocol for transmission to the independent committee (I, 2).

The physician can combine medical research with professional care, the objective being the acquisition of new medical knowledge, only to the extent that medical research is justified by its potential diagnostic or therapeutic value for the patient.

Non-Therapeutic Biomedical Research Involving Human Subjects (Non-Clinical Biomedical Research) In the purely scientific application of medical research carried out on a human being, it is the duty of the physician to remain the protector of the life and health of that person on whom biomedical research is being carried out.

The subjects should be volunteers - either healthy persons or subjects for whom the experimental design is not related to the patient's illness.

The investigator or the investigating team should discontinue the research if in his/her or their judgment it may, if continued, be harmful to the individual.

IN RESEARCH ON MAN, THE INTEREST OF SCIENCE AND SOCIETY SHOULD NEVER TAKE PRECEDENCE OVER CONSIDERATIONS RELATED TO THE WELL-BEING OF THE SUBJECT.

Additional Human Subjects Issues
Inclusion of women. There will be no restriction to the enrollment of women in this study. In accord with other autoimmune disorders, MG predominantly affects women, especially in younger age groups. Since the study intends to enroll MG patients between the ages of 18 and 60, it is anticipated that a large number of study subjects will be women.

Inclusion of minorities. Likewise, there will be no restriction to the enrollment of minorities in this study. As a multinational study with centers in several European countries, in South Africa, and in all major regions of North America, in Australia and in the Far East (Japan and Taiwan) a wide variety of ethnic groups will be encountered. It is the investigators’ policy that all eligible patients be considered for enrollment in the trial, and that any exclusions be specifically stated. The investigators are committed to enrolling as varied an ethnic population as possible. Proposed dates of enrollment are February 2005 through July 2006. The MGFA and international components of the Muscular Dystrophy Association will be used to publicize the trial. An article describing the planning of the study has already been published in the MGFA national newsletter.19

Participation of children. Individuals above age 18 will be enrolled in the study. Children younger than 18 years are being excluded because of the nature of the procedure being tested (a major surgical procedure) and concerns regarding the risks of thymectomy in individuals who have not reached full maturity and are still refining their immunological repertoire. Thymectomy remains a controversial procedure in children, especially those who are pre-pubertal. Many MG experts will not perform thymectomy on children unless there is uncontrolled severe disability or a thymoma is evident. Therefore, the investigators considered this study was not appropriate for individuals less than 18 years of age.

Protection of Human Subjects.

1. Although ETTX and prednisone are widely used in the management of generalized MG, the mode of their use in this trial differs in some respects from that in normal practice, notably the requirement initially to increase the prednisone dose by 10 mg per dose to a protocol maximum, achieved within 20 days or less (depending on the starting dose). In the previous application this was set at 120 mg ad. However, this could mean that some patients are required to take a higher dose than would have been the case if they had not been in the trial. Although this means that they may achieve MM status more rapidly, it also means that they could experience unnecessary prednisone adverse effects. To take account of this, the maximum dose has now been reduced to 100 mg AD, with the option to increase to 120 mg AD if MM status is not reached within 4 months (BE will be required to follow a Deviation Protocol).

Only patients with generalized MG will be recruited for the study. Patients will have the option not to participate in the study. In such a case, their management is likely to include some or all of the procedures described in the study plan.

The only specimens to be collected will be blood specimens and histology from the thymic resection. While these specimens will be collected for research studies (if additional funding is obtained), the thymic tissue would be available simply from the standpoint of the patient having undergone thymectomy.

To minimize potential risks, patients at all centers will be followed closely during the 3-year study with frequent laboratory monitoring to guard against toxicity from prednisone or other immunosuppressive medications should these be prescribed (via Deviation Protocols). These
studies will be performed on Day 0 or earlier (baseline), at least monthly for 3 months after Day 0, and every three months thereafter while the patient remains in the study. Provisions have been made in the study plan should patients worsen during the tapering of prednisone. In addition, plasma exchange, and intravenous immunoglobulin are treatment options (but require following Deviation Protocols).

2. Informed consent will be required before any eligible subject is enrolled in the study. This informed consent will be obtained by the MC or his/her clinical research staff. Full disclosure of the procedures involved, alternative care, and potential benefits and risks of participating will be provided on consent forms constructed to satisfy the requirements of the local Institutional Review Boards (IRBs).

Appropriate steps to maintain patient confidentiality will be taken. Patients will be identified by study number when information is transmitted to the data management center.

Plans for a Data and Safety Monitoring Board are included in section (D30, page 23). A report of serious adverse events will be provided to all members of the DSMB every three months. A teleconference or face-to-face meeting is planned at least every 12 months for the DSMB. More frequent communication will be arranged as needed.

3. As described in Section 8B, the benefit of thymectomy in non-thymomatous MG remains unclear. A recent practice parameter from the American Academy of Neurology has called for a randomized, controlled study to examine this issue. Thymectomy is a major surgical procedure with potentially life-threatening adverse effects. Prednisone too has adverse effects and patients often receive both therapies. While the trial protocol uses standardized treatments that are likely to differ in detail from those the patients might have received had they been treated outside the trial, the treatments themselves (ETTX and alternate day prednisone) are in regular use in the management of MG as observed above. Set against the potential risks of this standardized therapy is the prospect that the results of the study will make a substantial difference to the future management of many MG patients. Thus a 30% or greater reduction in the total prednisone dose at 3 years in the operated group would establish the immunotherapeutic benefits of thymectomy in this patient population, and the risk to these patients of prednisone adverse effects would be substantially reduced or removed. In addition, such an outcome would provide indirect evidence of its possible benefits in those not receiving prednisone medication. Conversely, failure to demonstrate a meaningful reduction in prednisone dosage would suggest that thymectomy is an unnecessary procedure in the population studied, and would challenge the use of thymectomy in this group of patients, while subgroup analysis may show whether benefits are confined to those who are prednisone naïve at entry or who are in a particular age group. Thus the results should impact on recommended clinical practice to the benefit of future MG patients.

4. If this study determines that thymectomy does not confer additional benefits for patients beyond pharmacologic immunosuppressive treatment alone, the risk of surgery to patients and substantial cost to society will be reduced. On the other hand, if the study demonstrates benefit related to the surgical procedure, it will reinforce the use of this intervention in the management of generalized MG and clarify communication between clinicians and patients when thymectomy is raised as a treatment option.