MYASTHENIA GRAVIS
FOUNDATION OF AMERICA, INC.

MG is an autoimmune neuromuscular disorder.

Symptoms may include:
- double vision
- drooping eyelids
- slurred speech
- difficulty chewing and swallowing
- weakness in arms and/or legs

Community Health Charities Member Agency
MGFA code # 0537

Myasthenia Gravis: Recommendations for Clinical Research Standards

The Task Force of the Medical/Scientific Advisory Board of the Myasthenia Gravis Foundation of America, Inc. is indebted to ***Neurology*** and the ***Annals of Thoracic Surgery*** for publishing their Recommendations simultaneously. As a result, the Recommendations are identical in both journals and readily available to both the neurological and surgical communities.

The **Myasthenia Gravis: Outcomes Analysis** recommendations, which were published by ***Neurology*** on its Web Site, are included in this packet because they are an integral part of the Task Force Recommendations. The Task Force is indebted to the authors of the **Outcomes Analysis** report for this contribution and their assistance in preparing the statistical portions of the Task Force report.

Reprint Index

1. ***Neurology***
   Myasthenia Gravis: Recommendations for Clinical Research
2. ***Neurology***
   Editorial – Treatment of Myasthenia Gravis: A Call to Arms
3. ***Annals of Thoracic Surgery***
   Myasthenia Gravis: Recommendations for Clinical Research
   (1st page of reprint & Invited Commentary) (blue)
4. ***Neurology Web Site***
   Myasthenia Gravis: Outcomes Analysis (beige)

Reprinted from NEUROLOGY
Vol 55 No 1 July 2000

Published by Lippincott Williams & Wilkins Printed in U.S.A.

Views & Reviews

CME

Myasthenia gravis

Recommendations for clinical research standards

Task Force of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America; A. Jaretzki III, MD; R.J. Barohn, MD; R.M. Ernstoff, MD; H.J. Kaminski, MD; J.C. Keesey, MD; A.S. Penn, MD; and D.B. Sanders, MD

The need for universally accepted classifications, grading systems, and methods of analysis for patients undergoing therapy for MG is widely recognized and is particularly needed for therapeutic research trials. The Medical Scientific Advisory Board (MSAB) of the Myasthenia Gravis Foundation of America (MGFA) formed a Task Force in May 1997 to address these issues. Initially, the Task Force planned to develop classifications and outcome measures pertaining only to standardizing thymectomy trials. However, it quickly became apparent that their efforts should apply to all therapeutic trials for MG, and thus the scope of the mission was expanded.

During the development of these recommendations, the Task Force faced numerous dilemmas for which no universally satisfactory solution was available. Dilemmas were defined as "situations that require one to choose between two equally balanced alternatives or predicaments that seemingly defy satisfactory solutions." The Task Force members agreed at the outset, however, that their primary goal was to develop a uniform set of classifications to be used in the comparative analysis of the various therapeutic interventions for MG. With this as the primary goal, a consensus was gradually developed. In developing a consensus, at least two meetings were held each year during a 3-year period. Between meetings there was exchange of all proposals by electronic and surface mail, consultation with national and international experts in the field, critical analysis of all proposals, and many revisions. All conflicts (both minor and major) were resolved by vote. Virtually all issues were eventually approved unanimously; a few received a plurality of six.

This report presents the work of the Task Force and proposes classification systems and definitions of response to therapy designed to achieve more uniformity in recording and reporting clinical trials and outcomes research. Although designed primarily for research purposes, we think physicians may find some of the recommendations useful in the clinical management of patients with MG.

Additional material related to this article can be found on the *Neurology* Web site. Go to www.neurology.org and scroll down the Table of Contents for the July 12 issue to find the title link for this article.

MGFA Clinical Classification. This classification (table 1) is designed to identify subgroups of patients with MG who share distinct clinical features or severity of disease that may indicate different prognoses or responses to therapy. It should not be used to measure outcome. It defers quantitative assessment of muscle weakness to the more precise Quantitative MG Score for Disease Severity, defers response to therapy to the MGFA Postintervention Status and the Quantitative MG Score, and defers the status of medication to the Therapy Status classification.

The fluctuating extent and severity of MG, and the variable predominance of the muscle groups involved, makes it extremely difficult to classify these patients. Most existing classifications are modifications of Osserman's,[1] separating patients with purely ocular involvement from those with generalized weakness, and further separating those with mild, moderate, or severe generalized weakness. Osser-

From the Myasthenia Gravis Foundation of America, Inc., Chicago, IL; the Department of Surgery (Dr. Jaretzki), Columbia Presbyterian Medical Center, New York, NY; the Department of Neurology (Dr. Barohn), University of Texas Southwestern Medical Center, Dallas; the Neurology Service (Dr. Ernstoff), William Beaumont Hospital, Royal Oak, MI; the Departments of Neurology and Neurosciences (Dr. Kaminski), Case Western Reserve University, Department of Veterans Affairs Medical Center, University Hospitals of Cleveland, OH; the Department of Neurology (Dr. Keesey), University of California at Los Angeles; the National Institute for Neurological Disorders and Stroke (Dr. Penn), National Institutes of Health, Bethesda, MD; and the Department of Medicine (Dr. Sanders), Duke University Medical Center, Durham, NC

Supported by the Ringel Foundation.

Received November 16, 1999. Accepted in final form March 17, 2000.

Address correspondence and reprint requests to The Myasthenia Gravis Foundation of America, Inc., 123 West Madison, Suite 800, Chicago, IL 60602; e-mail: myastheniagravis@msn.com

Table 1 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 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 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 other than ocular muscles May also have ocular muscle weakness of any severity
IVa:	Predominantly affecting limb and/or axial muscles 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 by 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

man classifications have included categories based on the course of the disease, such as "acute fulminating" and "late severe," and at times also included categories for muscle atrophy and childhood onset. Experienced clinicians have devised other classifications to monitor response to treatment, some based on the degree of disability or age at onset, and others that include quantitative measurements of specific muscle function, such as arm abduction time and vital capacity.

In general, these classifications use subjective assessments and lack quantification. What one physician may regard as "mild," another might regard as "moderate" or "severe." Furthermore, some experienced clinicians believe that oropharyngeal involvement is more dangerous, and perhaps different than limb weakness, and thus should be identified by the classification system. Some feel that "crisis" (respiratory insufficiency necessitating intubation and assisted ventilation) is the result of coincidental infection or other stress in susceptible patients, whereas others regard crisis as defining the ultimate in disease severity.

The Task Force discussed these issues at length, and the members think that a uniform classification is necessary if meaningful comparison of data is to be achieved. The Task Force sees no alternative but to accept the inherent imprecision of a clinical classification, and it recommends that the MGFA Clinical Classification be used to supplement or to replace the classifications now in use. It also recommends that the most severely affected muscles be employed to define the patient's Class and that the "maximum severity" designation be used to identify the most severe pretreatment clinical classification status. The "maximum severity" designation may be made historically and is employed as a point of reference. The maximum severity remains the point of reference thereafter, with any worsening of the MG being reflected in the postintervention status determination.

The Quantitative MG Score for Disease Severity. A quantitative MG scoring system (QMG Score) is essential in the objective evaluation of therapy for MG. This scoring system is based on quantitative testing of sentinel muscle groups. The QMG Score should be used in conjunction with the Clinical Classification and the Postintervention Status. It does not replace the clinical evaluation of the patient and should not be used to compare severity between patients. To assess the effect of treatment, the QMG Score should be determined before beginning the treatment under study and at appropriate intervals thereafter. As with all patient evaluations, the time of the examination in relation to therapy, and the presence of factors that may alter the clinical course, should be recorded.

The Task Force recommends that a QMG Score be used in all prospective studies of therapy for MG. The specific scoring system recommended (table 2)[2] is a modification of earlier systems developed for this purpose.[3-6] Its interexaminer reliability has been confirmed. A manual and a demonstration video of this system are available from the MGFA.[7]

The Task Force encourages proposals to improve and validate this quantitative scoring system and specifically recommends that studies be performed to determine the value of "weighting" certain subscores of the QMG score (oropharyngeal and respiratory muscle weakness, for example). As it now stands, a patient can have an improved total score but be incapacitated by poor strength in one or two areas. The evaluation and addition of other objective means of testing muscle strength is also recommended. This includes the use of a hand-held dynamometer, the time taken to swallow a fixed volume of liquid,[8] oral muscle function assessment,[9] the measurement of respiratory muscles forces,[10] and objective functional tests in children.

MGFA Therapy Status. The MGFA Therapy Status (table 3) defines the treatment regimen of the

Table 2 Quantitative MG score for disease severity

Test item	None	Mild	Moderate	Severe	Score
Grade	0	1	2	3	
Double vision on lateral gaze right or left (circle one), seconds	61	11–60	1–10	Spontaneous	
Ptosis (upward gaze), seconds	61	11–60	1–10	Spontaneous	
Facial muscles	Normal lid closure	Complete, weak, some resistance	Complete, without resistance	Incomplete	
Swallowing 4 oz. water (½ cup)	Normal	Minimal coughing or throat clearing	Severe coughing/choking or nasal regurgitation	Cannot swallow (test not attempted)	
Speech after counting aloud from 1 to 50 (onset of dysarthria)	None at 50	Dysarthria at 30–49	Dysarthria at 10–29	Dysarthria at 9	
Right arm outstretched (90 deg sitting), seconds	240	90–239	10–89	0–9	
Left arm outstretched (90 deg sitting), seconds	240	90–239	10–89	0–9	
Vital capacity, % predicted	≥80	65–79	50–64	<50	
Rt-hand grip, kgW					
Men	≥45	15–44	5–14	0–4	
Women	≥30	10–29	5–9	0–4	
Lt-hand grip, kgW					
Men	≥35	15–34	5–14	0–4	
Women	≥25	10–24	5–9	0–4	
Head lifted (45 deg supine), seconds	120	30–119	1–29	0	
Right leg outstretched (45 deg supine), seconds	100	31–99	1–30	0	
Left leg outstretched (45 deg supine), seconds	100	31–99	1–30	0	
				Total QMG score (range, 0–39) _____	

patient at a given time and is most useful when used with the MGFA Postintervention Status. The Therapy Status, at any point in time, is defined by using a single designation or a combination of the designations. In addition, the duration of this status, the current doses of all pertinent medications, and the schedule of plasma exchanges and IV immunoglobulin (Ig) should be recorded.

MGFA Postintervention Status. The Postintervention Status (table 4) is designed to assess the clinical state of MG patients at any time after institution of treatment for MG. Use of this classification requires that specific forms of therapy be recorded separately in the Therapy Status section.

The criteria defining the Postintervention Status require that the patient be examined by someone skilled in the evaluation of neuromuscular disease. Criteria for change in the patient's status should be defined in each study protocol based on quantitative assessment of strength in pertinent or sentinel muscles. What constitutes a sustained substantial change in medication should also be defined specifically in each study protocol. Composite analysis ("bundling") of postintervention categories is discouraged because it circumvents the goal of response-specific analysis. However, if composite analysis is employed, the individual categories should also be analyzed. Pending further study, it is recommended that treatment and strength be stable for at least 3 months as a baseline before beginning prospective studies. If the patient has attained CSR,

Table 3 MGFA MG Therapy Status

NT	No therapy
SPT	Status post-thymectomy (record type of resection)
CH	Cholinesterase inhibitors
PR	Prednisone
IM	Immunosuppression therapy other than prednisone (define)
PE(a)	Plasma exchange therapy, acute (for exacerbations or preoperatively)
PE(c)	Plasma exchange therapy, chronic (used on a regular basis)
IG(a)	IVIg therapy, acute (for exacerbations or preoperatively)
IG(c)	IVIg therapy, chronic (used on a regular basis)
OT	Other forms of therapy (define)

Table 4 *MGFA Postintervention 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 eyelid 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 Manifestations (MM)	The patient has no symptoms of 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/day) for at least 1 year.
MM-3	The 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 QMG 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 of **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).

PR, or MM status, the change in status should be indicated as well.

Isolated weakness of eyelid closure was thought not to be a sign of active disease and therefore was not considered an exclusionary criteria from CSR or PR status. In contrast, patients receiving cholinesterase inhibitors are excluded from PR and MM-1 status because these medications mask myasthenic symptoms.

Grouping by age, sex, race, and geography. The sex, race, age at onset, and geographic distribution of patients may be important variables in assessing response to treatment of MG.[11,12] Accordingly, these variables should be identified in all studies and their effect on response to treatment should be evaluated.

At this time, definitions of the age limits of childhood, adolescence, or onset of puberty have not been universally accepted or applied.[11,13] Appropriate age subdivision in children needs to be defined for universal use. Until such time as these age groups can be agreed on, it is recommended that age grouping occur by decade (i.e., up to 10 years, 11 to 20 years, 21 to 30 years, etc.).

Thymic pathology. There are no comprehensive guidelines for the evaluation of the nonthymomatous thymus removed from patients with MG. Incomplete and variable sampling techniques[14] and the lack of comprehensive diagnostic guidelines may account for the conflicting reports regarding the relationship between the histologic changes in the thymus and the response to thymectomy.

To determine the prognostic and therapeutic importance of pathologic changes in the thymus removed during thymectomy, uniform criteria for specimen management, sampling techniques, criteria for the diagnosis of hyperplasia and involution, characterization of immunocytochemical changes, and other determinants need to be defined and standardized.[12,15,16]

The histologic classification and grading of thymic neoplasms, including the identification of noninvasive and invasive thymomas, other neoplasms of the thymus, and their relation to MG results, also need to be standardized.[15,17-24] Patients with thymoma undergoing thymectomy for MG should be analyzed separately from those without thymoma, because combining these patient cohorts has made it difficult

to determine whether the presence of a thymoma, even when noninvasive, alters the prognosis.

The Task Force recommends that the multiple issues involved be addressed with the development of guidelines for the evaluation of the nonthymomatous thymus in MG and a universally acceptable single classification of thymic neoplasms.

MGFA Thymectomy Classification. Multiple techniques are described for removal of the thymus in MG. The debate regarding which technique is preferable is not resolved. Although, classically, "total thymectomy" is considered the goal of surgery, it has not been demonstrated unequivocally that this is necessary, nor is it clear that all the resectional techniques do achieve this goal.

To resolve the issues regarding the choice of thymectomy technique and whether there is a relationship between the resectional technique employed and the rate of remission and improvement, the type of thymic resection used needs to be defined in as objective terms as possible. In addition, multiple resectional techniques should not be reported as a single cohort.

The Thymectomy Classification (table 5) is based on published reports. The techniques are grouped according to the primary approach (transcervical, videoscopic, transsternal, or combinations) and are described briefly. Referenced reports are recommended for details. Because, within each category, there may be variations in the extent of the resection from surgeon to surgeon, the extent of the resection for each patient cohort must be recorded. In all prospective studies it is recommended that detailed descriptions of the operative technique be supplied, accompanied by drawings and photographs of typical specimens. Ideally, a video of the technique should also be available.

At this time, two types of transcervical thymectomy are performed. The "Basic" resection employs an intracapsular extraction of the mediastinal thymus via a cervical incision and is limited to the removal of the central cervical–mediastinal lobes (figure, A and B). No other tissue is removed in either the neck or the mediastinum.[25,26] The original "Extended" resection employs a special manubrial retractor for improved exposure of the mediastinum. The mediastinal dissection is extracapsular and includes resection of the visible mediastinal thymus and fat. Sharp dissection may or may not be performed on the pericardium. The neck exploration and dissection varies in extent and may or may not be limited to exploration and removal of the cervical–mediastinal extensions.[27,28] Variations include the addition of a partial median sternotomy[29] and the associated use of mediastinoscopy.[30]

A number of variations in videoscopic technique are being developed to assist in the performance of a thymectomy. The "Classic" video-assisted thoracic surgery (VATS) technique employs unilateral videoscopic exposure of the mediastinum (right or left) with removal of the grossly identifiable thymus and variable amounts of anterior mediastinal fat. The cervical extensions of the thymus are usually removed from below.[31,32] The Video-assisted Thoracoscopic Extended Thymectomy (VATET) employs bilateral thoracoscopic exposure of the mediastinum for improved visualization of both sides of the mediastinum. Extensive removal of the mediastinal thymus and perithymic fat is described, the thymus and fat being removed separately. A cervical incision is performed with removal of the cervical thymic lobes and pretracheal fat.[33]

There are several variations to the Transsternal Thymectomy approach. The "Standard" technique was originally designed to remove the well-defined central cervical–mediastinal lobes. At this time, although a complete[34] or partial[35,36] sternotomy may be performed, the resection is more extensive than originally described, with removal of all visible mediastinal thymus. Mediastinal fat, varying in extent, may or may not be removed. The cervical extensions of the thymus are removed from below, with or without some adjacent cervical fat. Variations of this technique include a video-assisted technique using a complete median sternotomy via a limited lower sternal transverse skin incision.[37] The "Extended"[38] procedure is also known as Aggressive Transsternal Thymectomy[39] and Transsternal Radical Thymectomy.[40] These resections remove the entire mediastinal thymus and most of the mediastinal perithymic fat. They vary somewhat in extent in the mediastinum and may or may not include all tissue removed by the T-4 techniques. The cervical extensions are removed from below, with or without additional tissue, but without a formal neck dissection.

The combined transcervical and transsternal thymectomy procedures are known as Transcervical–Transsternal Maximum Thymectomy[41] and Extended Cervical–Mediastinal Thymectomy.[42] These resections routinely use wide exposure in the neck and a complete median sternotomy with en bloc removal of all tissue in the neck and mediastinum that may contain gross or microscopic thymus anatomically. The resections include removal of both sheets of mediastinal pleura and sharp dissection of the pericardium. A similar procedure, although a less

Table 5 *Thymectomy Classification*

T-1	Transcervical Thymectomy
	(a)-Basic
	(b)-Extended
T-2	Videoscopic Thymectomy
	(a)-"Classic"
	(b)-"VATET"
T-3	Transsternal Thymectomy
	(a)-Standard
	(b)-Extended
T-4	Transcervical & Transsternal Thymectomy

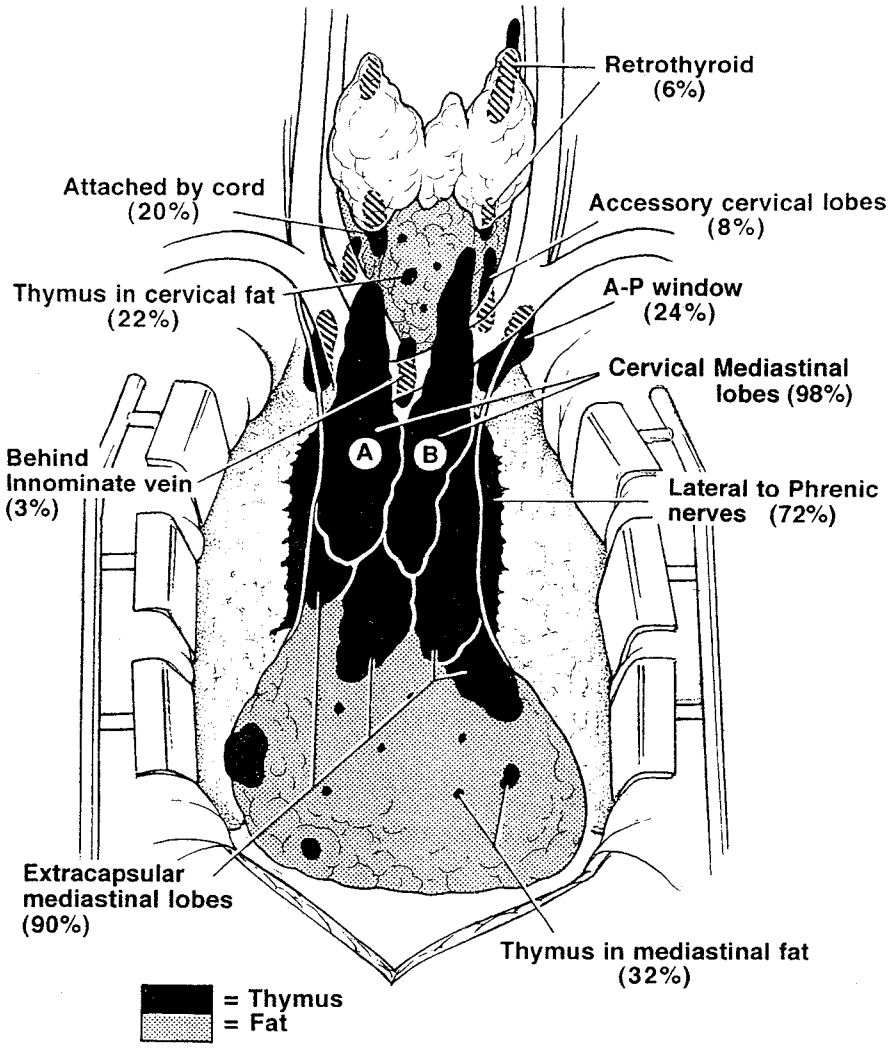


Figure. Anatomy of the thymus. This illustration represents what is now generally accepted as the surgical anatomy of the thymus.[41] The frequencies (percent occurrence) of the variations are noted. Black = thymus; gray = fat that may contain islands of thymus and microscopic thymus. A-P window = aortopulmonary window. Source: Neurology *1997;48(suppl 5):S52–S63.*

extensive resection in the neck and mediastinum, has been described by Lennquist et al.[43]

It is recommended that the thymectomy classification, with modifiers as necessary, be employed when reporting the results of thymectomy for MG.

MGFA Morbidity and Mortality Classification. During the evaluation of therapeutic options, in addition to the determination of the remission and improvement status, quality of life and cost–benefit assessments should be performed.[44] This requires the analysis of, among other things, the number and duration of hospitalizations and intensive care unit stays, and complications related directly to each form of therapy (table 6).

Formal Clinical Trial and Outcomes Research. A prospective, randomized clinical trial remains the preferred method to evaluate therapy (Class I evidence in the American Academy of Neurology [AAN]) nomenclature).[45] When a randomized trial does not appear feasible, a prospective risk-adjusted outcome analysis of nonrandomly assigned treatment (Class II evidence in the AAN nomenclature) is recommended.[46] In prospective studies, in addition to the use of the classification, definitions, grading systems, and methods of analysis recommended herein, the CONSORT guidelines[47,48] are recommended.

"Survival" instruments, which are used in the analysis of remissions, are fundamental in the comparative analysis of therapeutic programs for MG. Although different levels of clinical improvement should be evaluated in the analysis of all forms of therapy, Complete Stable Remission remains the primary focus of the analysis, at least in the assessment of thymectomy.

Qualify-of-life instruments should also be employed because therapy for MG is usually not innocuous and frequently does not produce a completely stable remission. Quality-of-life measures evaluate the impact of intermediate levels of clinical improvement and morbidity of therapy, and complement information provided by remission and clinical improvement analysis. Although a functional status instrument assessing activities of daily living has been developed for MG,[49] there are no disease-specific quality-of-life instruments for MG at this time. The Task Force recommends that these be developed. The steps necessary to accomplish this have been defined.[50]

Table 6 *MGFA Morbidity and Mortality*

Hospitalizations
Number per year (average since onset Rx & no. in the last year)
Days per year (average since onset Rx & no. in the last year)
Intensive care stays
Number per year (average since onset of Rx & no in the last year)
Days per year (average since onset Rx & no. in the last year)
Ventilatory support
Pre-Rx or During Rx
Duration in place (days)
Tracheostomy
Pre-Rx or During Rx
Duration in place (days)
Infections
Pulmonary
"Line"
Other
Therapy-specific Complication (name of drug and complication)
Death (list cause, relation to therapy, and indicate whether therapy related. Include all deaths occurring during a hospitalization.)
Operative/postoperative
Length of surgery (hours/minutes)
Intraoperative complications
Hospital stay (days)
Intensive care stay (days)
Ventilatory support (days)
Infection (location, type, and severity)
Transfusions (number)
Nerve injury (phrenic/recurrent/intercostal—temporary/permanent)
Persistent pain (severity, duration, and therapy)
Chylothorax (severity and duration)
Death (occurring within 30 days of surgery, even if the patient has been discharged, and occurring after 30 days when clearly related to the surgical procedure. Deaths within 90 days of surgery should also be recorded).

An Outcomes Analysis guideline and an accompanying diagram defining their interrelationships were developed to provide background information on the available analytic techniques.[44] Experts in the field of biostatistics and outcomes analysis should be consulted in the design of all studies, and in the collection and evaluation of the data.

Data bank. Multi-institutional studies utilizing the data bank concept and fulfilling the requirements discussed next are recommended. This method of study should be particularly useful and practical for multiple institutions to compare the relative value of the many therapies, including various thymectomy techniques.

For a data bank program to be successful, it must be developed appropriately and monitored rigorously. Computer-based patient records, including the "human language" component,[51] are required. Definitions, classifications, and standardized forms must be agreed on and used. Standardization of numeric grading for all important variables is required. Mechanisms must be in place to review all the clinical records and to monitor the quality of the database, including validation for completeness and accuracy through a rigid auditing process. The monitoring requirement involves a major commitment by the sponsoring institutions, professionals, and staffs. Mechanisms to defray costs would need to be developed.

Amendments. The MSAB of the MGFA has established a Standing Committee for review of the Clinical Research Standards. The goal of any proposed amendment is to improve the guidelines based on demonstrable errors in the existing guidelines, development of new data, and common sense. The Review Committee will review recommendations from national and international neurologic organizations and centers, journal editorial boards, neurologists, immunologists, surgeons, biostatisticians, nurses, respiratory therapists, and others working in the field of MG. The Review Committee will also serve as a clearinghouse for questions, as they arise, concerning the application of the Clinical Research Standards.

Please submit questions and proposals in writing to Chairperson, Clinical Research Standards Review Committee, Medical Scientific Advisory Board, Myasthenia Gravis Foundation of America, Inc., 123 West Madison, Suite 800, Chicago, IL 60602; e-mail: myastheniagravis@msn.com

References

1. Osserman KE. Clinical aspects. In: Osserman KE, ed. Myasthenia gravis. New York, NY: Grune & Stratton, 1958:79–80
2. Barohn RJ, Mcintire D, Herbelin L, Wolfe GI, Nations S, Bryan W. Reliability testing of the quantitative myasthenia gravis score. Ann NY Acad Sci 1998;841:769–772.
3. Besinger UA, Toyka KV, Heininger K, et al. Long-term correlation of clinical course and acetylcholine receptor antibody in patients with myasthenia gravis. Ann NY Acad Sci 1981;377:812–815.
4. Besinger UA, Toyka KV, Homberg M, Heininger K, Hohlfeld R, Fateh-Moghadam A. Myasthenia gravis: long-term correlation of binding and bungarotoxin blocking antibodies against acetylcholine receptors with changes in disease severity. Neurology 1983;33:1316–1321.
5. Hohlfeld R, Toyka KV, Besinger UA, Gerhold B, Heininger K. Myasthenia gravis: reactivation of clinical disease and of autoimmune factors after discontinuation of long-term azathioprine. Ann Neurol 1985;17:238–242.
6. Tindall RSA, Rollins JA, Phillips JT, et al. Preliminary results of a double-blind, randomized, placebo-controlled trial of cyclosporine in myasthenia gravis. N Engl J Med 1987;316:719–724.
7. Barohn RJ. How to administer the quantitative myasthenia test. Video. 1996. Myasthenia Gravis Foundation of America, Inc., 123 West Madison, Suite 800, Chicago, IL 60602; e-mail: myastheniagravis@msn.com
8. Palace J. Azothiaprene vs prednisone—clinical trial. Neurology 1998;50:1778–1783.
9. Weijnen FG, VanDerBilt A, Wokke JHJ, Wassenberg MWM,

Oudenaarde I. Oral functions of patients with myasthenia gravis. Ann NY Acad Sci 1998;841:773–776.
10. Younger DS, Braun NMT, Jaretzki A III, Penn AS, Lovelace AE. Myasthenia gravis: determinants for independent ventilation after transsternal thymectomy. Neurology 1984;34:336–340.
11. Andrews PI, Massey JM, Howard JF, Sanders DB. Race, sex, and puberty influence onset, severity, and outcome in juvenile myasthenia gravis. Neurology 1994;44:1208–1214.
12. Compston DAS, Vincent A, Newsom–Davis J, Batchelor JR. Clinical, pathological, HLA antigen and immunological evidence for disease heterogeneity in myasthenia gravis. Brain 1980;103:579–601.
13. Seybold ME. Thymectomy in childhood myasthenia gravis. Ann NY Acad Sci 1998;841:731–741.
14. Grody WW, Jobst S, Keesey JC, Herrmann C, Naeim F. Pathologic evaluation of thymic hyperplasia in myasthenia gravis and Lambert–Eaton myasthenic syndrome. Arch Pathol Lab Med 1986;110:843–846.
15. Berrih–Aknin S. The in vivo effects of corticosteroids on thymocyte subsets in myasthenia gravis. J Clin Immunol 1984;4: 92–97.
16. Bertho X. Phenotype and immunohistochemical analyses of the human thymus: evidence for an active thymus during adult life. Cell Immunol 1997;179:30–40.
17. Marino M, Muller–Hermelink HK. Thymoma and thymic carcinoma. Relation of thymoma epithelial cells to the cortical and medullary differentiation of the thymus. Virchows Arch A Pathol Anat Histopathol 1985;407:119–149.
18. Lewis JE, Wick MR, Bernd W, Scheithauer MD, Bernatz PE, Taylor WF. Thymoma. A clinicopathologic review. Cancer 1987;60:2727–2743.
19. Muller–Hermelink HK, Marx A, Geuder K, Kirschner T. The pathological basis of thymoma-associatred myasthenia gravis. Ann NY Acad Sci 1993;681:56–65.
20. Close PM, Kirschner T, Uys CJ, Muller–Hermelink HK. Reproducibility of a histiogenic classification of thymic epithelial tumours. Histopathology 1995;26:339–343.
21. Muller–Hermelink HK, Willisch A, Schultz A, Marx A. Characterization of the human thymic microenvironment: lymphoepithelial interaction in normal thymus and thymoma. Arch Histol Cytol 1997;60:9–28.
22. Blumberg D, Burt ME, Bains MS, et al. Thymic carcinoma: current staging does not predict prognosis. J Thorac Cardiovasc Surg 1998;115:303–309.
23. Fukai I, Masaoka A, Fujii Y, et al. Thymic neuroendocrine tumor (thymic carcinoid): a clinicopathologic study. Ann Thorac Surg 1999;67:208–211.
24. Wick MR, Ritter JH. Thymic neuroendocrine tumors: a commentary with current perspectives on neuroendocrine neoplasia. Ann Thorac Surg 1999;67:7–9.
25. Kark AE, Kirschner PA. Total thymectomy by the transcervical approach. Br J Surg 1971;56:321–326.
26. Kark AE, Papatestas AE. Some anatomic features of the transcervical approach for thymectomy. Mt Sinai J Med 1971; 38:580–584.
27. Cooper J, Al-Jilaihawa A, Pearson F, Humphrey J, Humphrey HE. An improved technique to facilitate transcervical thymectomy for myasthenia gravis. Ann Thorac Surg 1988;45:242–247.
28. Bril V, Kojic S, Ilse W, Cooper J. Long-term clinical outcome after transcervical thymectomy for myasthenia gravis. Ann Thorac Surg 1998;65:1520–1522.
29. Durelli L, Maggi G, Casadio C, Ferri R, Rendine S, Bergamini L. Actuarial analysis of the occurrence of remission following thymectomy for myasthenia gravis in 400 patients. J Neurol Neurosurg Psychiatry 1991;54:406–411.
30. Klingen G, Johansson L, Westerholm CJ, Sundstroom C. Transcervical thymectomy with the aid of mediastinoscopy for myasthenia gravis: eight years' experience. Ann Thorac Surg 1977;23:342–347.
31. Yim APC, Kay RLC, Ho JKS. Video-assisted thoracoscopic thymectomy for myasthenia gravis. Chest 1995;108:1440–1443.
32. Mack MJ, Landreneau RD, Yim AP, Hazelrigg SR, Scruggs GR. Results of video-assisted thymectomy in patients with myasthenia gravis. J Thorac Cardiovasc Surg 1996;112:1352–1360.
33. Scelsi R, Ferro MT, Scelsi L, et al. Detection and morphology of thymic remnants after video-assisted thoracoscopic thymectomy (VATET) in patients with myasthenia gravis. Int Surg 1996;81:14–17.
34. Olanow CW, Wechsler AS. The surgical management in myasthenia gravis. In: Sabiston DC Jr, ed. Textbook of Surgery 14th ed. Philadelphia, PA: WB Saunders, 1991:1801–1807.
35. Trastek VF. Thymectomy. In: Kaiser LR, Krin IL, Spray TL, eds. Mastery of cardiothoracic surgery. Philadelphia, PA: Lippincott-Raven, 1998:105–111.
36. LoCicero J III. The combined cervical and partial sternotomy approach for thymectomy. Chest Surg Clin North Am 1996;6: 85–93.
37. Granone P, Margaritora S, Cesario A, Galetta D. Thymectomy (transsternal) in myasthenia gravis via video assisted inframammary cosmetic incision. Eur J Cardiothorac Surg 1999; 15:861–863.
38. Mulder DG. Extended transsternal thymectomy. Chest Surg Clin North Am 1996;6:95–105.
39. Fischer JE, Grinvalski HT, Nussbaum MS, Sayers HK, Cole RE, Samaha FJ. Aggressive surgical approach for drug-free remission from myasthenia gravis. Ann Surg 1987;205:496–503.
40. Hatton P, Diehl J, Daly BDT, et al. Transsternal radical thymectomy for myasthenia gravis: a 15-year review. Ann Thorac Surg 1989;47:838–840.
41. Jaretzki A III, Wolff M. "Maximal" thymectomy for myasthenia gravis: surgical anatomy and operative technique. J Thorac Cardiovasc Surg 1988;96:711–716.
42. Bulkley GB, Bass KN, Stephenson R, et al. Extended cervicomediastinal thymectomy in the integrated management of myasthenia gravis. Ann Surg 1997;226:324–335.
43. Lennquist S, Andaker L, Lindvall B, Smeds S. Combined cervicothoracic approach in thymectomy for myasthenia gravis. Acta Chir Scand 1990;156:53–61.
44. Weinberg A, Gelijns A, Moskowitz A, Jaretzki A. Myasthenia gravis: outcomes analysis. 2000. Available at: http://www.neurology.org.
45. Miller RG, Rosenberg JA, Force APPT. Care of patient with ALS: report of the Quality Standards Subcommittee of the AAN. Neurology 1999;52:1311–1323.
46. Kirklin JW, Blackstone EH. Clinical studies with nonrandomly assigned treatment. In:Kirklin JW, Barrett–Boyes BG, eds. Cardiac surgery. 2nd ed. New York, NY: Churchill-Livingstone, 1993:269–270.
47. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. JAMA 1996;276:637–639.
48. Schulz KF. The quest for unbiased research: randomized clinical trials and the CONSORT reporting guidelines. Ann Neurol 1997;41:569–573.
49. Wolfe GI, Herbelin L, Nations SP, Foster B, Bryan WW, Barohn RJ. Myasthenia gravis activities of daily living profile. Neurology 1999;52:1487–1489.
50. Jaeschke R, Guyatt GH. How to develop and validate a new quality of life instrument. In: Spilker B, ed. Quality of life assessments in clinical trials. New York, NY: Raven Press, 1990:47–57.
51. Kirklin JW, Vicinanza SS. Metadata and computer-based patient records. Ann Thorac Surg 1999;68:S23–S24.

Reprinted from NEUROLOGY
Vol 55 No 1 July 2000

Published by Lippincott Williams & Wilkins Printed in U S A

Editorial

Treatment of myasthenia gravis

A call to arms

John T. Kissel, MD; Gary M. Franklin, MD; and the Quality Standards Subcommittee of the American Academy of Neurology

In this issue of *Neurology*, two articles address important issues in the treatment of myasthenia gravis (MG).[1,2] Gronseth and Barohn review the role of thymectomy[1] and the Task Force of the Medical Scientific Advisory Board of the MG Foundation of America outlines a comprehensive grading system for classifying and following patients.[2] Although these two articles are different in their approach and scope, both are important pieces of work that represent significant contributions to the literature on MG.

MG represents one of the great medical triumphs of the last half century. In 1960, Simpson first proposed that MG was an autoimmune disease and hypothesized that it resulted from an antibody-dependent block in neuromuscular transmission. Investigations in the 1970s demonstrated the deficiency of acetylcholine receptors at the neuromuscular junction in MG, the production of animal models by immunization with acetylcholine receptors, the passive transfer of the disease between species with immunoglobulin G, and the presence of antibodies to acetylcholine receptors in most patients with MG.[3] Subsequently, the immunopathogenic and electrophysiologic mechanisms involved in the disease were elucidated.

Coincident with these discoveries, effective treatments were developed, including acetylcholinesterase inhibitors in the 1950s, and by 1970, prednisone and other immunosuppressive medications were available. In the 1970s, thymectomy—first described as a treatment modality in 1936—became an increasingly accepted form of therapy. In the 1980s, both plasma exchange and intravenous immunoglobulin were used to treat MG, particularly in patients with life-threatening illness.

However, problems and questions remain. The mortality rate for MG has not decreased over the last 50 years[4]; reports suggesting that the documented increased prevalence of MG is due to better survival more likely reflect improved case ascertainment in the elderly.[5] In addition, there are no studies comparing different treatment modalities for ocular versus generalized MG and for differing severity of disease. In fact, only two adequately controlled trials have been performed in MG: one with cyclosporine[6] and the other comparing prednisolone with or without azathioprine.[7] There are, therefore, no evidence-based data on optimum treatment regimens for MG.

Why have there been so few therapeutic trials in MG?[1] 1) MG has a relatively low prevalence (~15/100,000),[8] so recruiting enough patients for a prospective, controlled trial may be difficult. 2) Clinicians may choose not to refer patients with treatable conditions for studies, even if the best treatment is uncertain. Experienced clinicians may feel that they know which treatment is best. 3) The variability of the symptoms and signs of myasthenia make it difficult to determine firm endpoints for a clinical trial. 4) The use of concomitant agents, particularly acetylcholinesterase inhibitors, may be difficult to control. 5) Subsets of MG with different antibody or electrophysiologic profiles make it difficult to select a homogenous patient population for study. 6) Varying thymic pathology (normal, thymic hyperplasia, thyomoma), as well as whether and how the gland is resected, are potential confounding variables.

In this context, both articles[1,2] are a "call to arms" for well-designed therapeutic trials in MG. The Task Force outlines a comprehensive system for succinctly describing and summarizing *most* aspects of the disease pertinent to a clinical trial: a clinical classification scheme; quantitative scoring scale; therapy status descriptors; and a scheme for classifying the various types of thymectomy. It also outlines additional measures crucial for therapeutic trials: standardized hospital morbidity and mortality formats; use of CONSORT guidelines[9]; and quality of life and cost-benefit assessments.

The Task Force recommendations are similar to various grading and staging systems used for cancer treatment protocols. MG patients entering a trial

From the Department of Neurology (Dr. Kissel), The Ohio State University, Columbus; Occupational Epidemiology and Health Outcomes Program (Dr. Franklin), Department of Environmental Health, School of Public Health and Community Medicine, University of Washington, Seattle; and the Quality Standards Subcommittee of the American Academy of Neurology, St. Paul, MN.
Address correspondence and reprint requests to Dr. John Kissel, Department of Neurology, Means Hall, 1654 Upham Drive, Columbus, OH 43210.

would be assigned a grade and stage that would describe their clinical status, treatment, and therapeutic response. Although too complex for routine use, it establishes the standardized criteria needed for the performance of clinical trials in MG. The authors recognize that this is only a first step, and have established a standing committee to make changes that will "improve the guidelines based on demonstrable errors in the existing guidelines, development of new data, and common sense."[2] The complexity of the guidelines highlights the challenge to improve trials for MG.

These new guidelines can be criticized. First, the method used to achieve consensus among expert opinions is not explicit. A formal process, such as a modified Delphi method, would have strengthened the recommendations.[10] Such a process is important when expert opinion, rather than evidence, is the foundation of the analysis. Second, broader international input is needed, particularly if the guidelines are for worldwide use. Third, many of the guidelines are arbitrary; e.g., defining remission as "no symptoms or signs of MG for at least 1 year." Fourth, the reliability of the proposed rating scales needs testing for inter- and intrarater agreement. The recommendations should be validated before they can be applied. Fifth, neither electrophysiologic studies nor acetylcholine receptor antibody status are incorporated in the current version.

The Task Force's stated charge was to develop criteria for thymectomy trials. This initial focus on thymectomy is entirely appropriate because this is usually an early consideration in treating MG.[11] This issue is addressed in the evidence-based review of thymectomy by Gronseth and Barohn.[1] Using Quality Standards Subcommittee Practice Parameter methods,[12] they systematically reviewed 28 articles involving 21 patient cohorts on thymectomy in MG. They categorized the studies according to class of data; reviewed each study in regards to survival, remission (on or off medication), and improvement status after thymectomy; and calculated the relative rates for each outcome in the thymectomy versus nonthymectomy groups. They concluded that thymectomy was associated with disease remission and improvement, with beneficial relative rates between 1.5 and 2.1, and recommend that ". . . thymectomy is an *option* to increase the probability of remission or improvement." Believers in thymectomy will view this analysis as supporting their beliefs.

However, the Gronseth and Barohn article is sobering with regard to the actual evidence favoring thymectomy. In only 7/15 studies describing medication-free remission and in only 4/13 studies describing survival was improvement significant for a thymectomy group. The *majority* of the studies did not show significant benefit. In addition, the methods were not optimum in all of the reports. None was randomized, none used blinded outcome assessments, and none described either patient selection criteria or the status of the thymus gland (normal or hyperplastic). Most studies did not report the number of patients lost to follow-up. In addition, the thymectomy in most studies was performed in younger patients, a bias that skews the data in favor of a positive outcome for thymectomy. The studies also did not standardize medical therapy.

Given these limitations, and the fact that only modest improvement was seen and in only a *minority* of studies, it is surprising that thymectomy is such a widely accepted treatment for MG. Whereas some neurologists feel so strongly about the benefits of thymectomy that they deem a randomized controlled trial unjustified, the data reviewed by Gronseth and Barohn make it clear that there is equipoise over virtually all issues pertaining to thymectomy. Randomized trials are needed.

Both the articles of Gronseth and Barohn[1] and the Myasthenia Task Force[2] demand that neurologists and thoracic surgeons conduct therapeutic trials in MG. Clinicians must not become complacent and, because of the multiple therapeutic options now available, let MG become a victim of its own success, with most of the important issues related to management remaining unanswered.

References

1. Gronseth GS, Barohn RJ. Practice parameter: thymectomy for autoimmune myasthenia gravis (an evidence-base review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2000;55:7–15.
2. Task Force of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America. Myasthenia gravis: recommendations for clinical research standards. Neurology 2000;55:16–23.
3. Engel AG. Acquired autoimmune myasthenia gravis. In: Engel AG, Franzini-Armstrong C, eds. Myology. New York: McGraw-Hill, 1994:1769–1797.
4. Phillips LH II, Torner JC. Epidemiologic evidence for a changing natural history of myasthenia gravis. Neurology 1996;47:1233–1238.
5. Robertson NP, Deans J, Compston DAS. Myasthenia gravis: a population-based epidemiological study in Cambridgeshire, England. J Neurol Neurosurg Psychiatry 1998;15:492–496.
6. Tindall RSA, Rollins JA, Phillips TJ, Greenlee RG, Wells L, Belendiuk G. Preliminary results of a double-blind, randomized, placebo-controlled trial of cyclosporine in myasthenia gravis. N Engl J Med 1987;316:719–724.
7. Pallace J, Newsom-Davis J, Lecky B, and the Myasthenia Study Group. A randomized double-blind trial of prednisolone alone or with azathioprine in myasthenia gravis. Neurology 1998;50:1778–1783.
8. Phillips LH II, Torner JC, Anderson MS, Cox GM. The epidemiology of myasthenia gravis in central and western Virginia. Neurology 1992;42:1888–1893.
9. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. JAMA 1996;276:637–639.
10. Black N, Murphy M, Lamping D, et al. Consensus development methods: a review of best practices creating clinical guidelines. J Health Serv Res Policy 1999;4:236–248.
11. Busch C, Macheus A, Pichlmeier U, Emskotter T, Isbicki JR. Long term outcome and quality of life after thymectomy for myasthenia gravis. Am Surg 1996;224:225–232.
12. Miller RG, Rosenberg JA, Gelinas DF, et al. Practice parameter: the care of the patient with amyotrophic lateral sclerosis (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology: ALS Practice Parameter Task Force. Neurology 1999;52:1311–1323.

Myasthenia Gravis: Recommendations for Clinical Research Standards

Alfred Jaretzki III, MD, Richard J. Barohn, MD, Raina M. Ernstoff, MD, Henry J. Kaminski, MD, John C. Keesey, MD, Audrey S. Penn, MD, and Donald B. Sanders, MD, Task Force of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America

Department of Surgery, Columbia Presbyterian Medical Center, New York, New York; the Department of Neurology, University of Texas Southwestern Medical Center, Dallas, Texas; the Neurology Service, William Beaumont Hospital, Royal Oak, Michigan; the Departments of Neurology and Neurosciences, Case Western Reserve University, Department of Veterans Affairs Medical Center, University Hospitals of Cleveland, Cleveland, Ohio; Department of Neurology, University of California at Los Angeles, Los Angeles, California; the National Institute for Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland; and the Department of Medicine, Duke University Medical Center, Durham, North Carolina

The need for universally accepted classifications, grading systems, and methods of analysis for patients undergoing therapy for MG is widely recognized and is particularly needed for therapeutic research trials. The Medical Scientific Advisory Board (MSAB) of the Myasthenia Gravis Foundation of America (MGFA) formed a Task Force in May 1997 to address these issues. Initially, the Task Force planned to develop classifications and outcome measures pertaining only to standardizing thymectomy trials. However, it quickly became apparent that their efforts should apply to all therapeutic trials for MG, and thus the scope of the mission was expanded.

During the development of these recommendations, the Task Force faced numerous dilemmas for which no universally satisfactory solution was available. Dilemmas were defined as "situations that require one to choose between two equally balanced alternatives or predicaments that seemingly defy satisfactory solutions." The Task Force members agreed at the outset, however, that their primary goal was to develop a uniform set of sponse to therapy designed to achieve more uniformity in recording and reporting clinical trials and outcomes research. Although designed primarily for research purposes, we think physicians may find some of the recommendations useful in the clinical management of patients with MG.

MGFA Clinical Classification

This classification (Table 1) is designed to identify subgroups of patients with MG who share distinct clinical features or severity of disease that may indicate different prognoses or responses to therapy. It should not be used to measure outcome. It defers quantitative assessment of muscle weakness to the more precise Quantitative MG Score for Disease Severity, defers response to therapy to the MGFA Postintervention Status and the Quantitative MG Score, and defers the status of medication to the Therapy Status classification.

The fluctuating extent and severity of MG, and the

INVITED COMMENTARY

Will Rogers once commented that "Everyone talks a lot about the weather, but no one does anything about it!" The same might be said perhaps about those of us involved in the care of patients with myasthenia gravis. We remain somewhat frustrated at our inability to more precisely understand this baffling disease. This is due in part to our lack of uniformity in categorizing the various manifestations and severity of our patients' disease both prior to and following the available medical and surgical interventions.

A major step has been taken to "do something about it" with the formation of the Myasthenia Gravis Task Force. This multi-specialty group of knowledgeable clinicians, after innumerable discussions and consultations with colleagues, and after many months of gestation, have given birth to a comprehensive set of Recommendations for Clinical Research Standards. The delivery of this much needed set of guidelines has been the result of a lengthy and at times painful labor. It is hoped that we will all welcome this new arrival, congratulate its parents, and cooperate in "doing something" about advancing our knowledge and capabilities in caring for our patients with myasthenia.

Donald G. Mulder, MD

UCLA Medical Center
Department of Surgery
620-249 CHS
10833 Le Conte Ave
Los Angeles, CA 90095-1471

Published by Elsevier Science Inc

Ann Thorac Surg 2000;70:327–34 • 0003-4975/00/$20.00
PII S0003-4975(00)01595-2

Myasthenia Gravis: Outcomes Analysis

Alan Weinberg, MS*, Annetine Gelijns, Ph**, Alan Moskowitz**, MD, and Alfred Jaretzki III, MD***. The Division of Biostatistics*, International Center for Health Outcomes and Innovation Research**, and General Thoracic Division***, Department of Surgery, College of Physicians and Surgeons, Columbia University, New York, NY, 10032.

Published on the Neurology Web Site

<http://www.neurology.org>

This Project was made possible by the generous support of the Ringel Foundation

Reprints from: The Myasthenia Gravis Foundation of America, Inc.
123 West Madison. Suite 800
Chicago, IL 60602

<e-mail: myastheniagravis@msn.com>.

I-Statistical Instruments for Analysis of Remissions & Levels of Improvement

A- "Survival" Instruments

Kaplan-Meier Product-Limit Estimator- The Kaplan-Meier method is a nonparametric (actuarial) technique for estimating time-related events (the survivorship function). [1] Ordinarily it is used to analyze death as an outcome. It may be used effectively to analyze time to an endpoint, such as remission.

It is a univariate analysis and is an appropriate starting technique. It estimates the probability of the proportion of individuals in remission at a particular time, starting from the initiation of therapy or the operative date (time zero), is especially applicable when length of follow-up varies from patient to patient, and takes into account those patients lost to follow-up or not yet in remission at end of study (censored patients, assuming the censoring is non-informative). It is therefore the instrument of choice in evaluating remissions following thymectomy. Since the estimated survival distribution for the cohort has some degree of uncertainty, 95% confidence intervals may be calculated for each survival probability on the "estimated" curve.

A variety of tests (log-rank, Wilcoxan and Gehen) may be used to compare two or more Kaplan-Meier "curves" under certain well-defined circumstances. Median remission time (the time when 50% of the cohort has reached remission), as well as quantities such as three, five, and ten year probability of remission, can also be generated from the Kaplan-Meier analysis, provided there has been sufficient follow-up of patients.

The Kaplan-Meier technique is usually only useful as a method of preliminary evaluation, since it is purely a descriptive method for the evaluation of one variable.

Cox Proportional Hazard-Regression Analysis- This is a form of multivariable analysis that can consider many potential predictive variables simultaneously and is required when multiple factors may affect the outcome. [2] The use of this technique usually accompanies the Kaplan-Meier analysis. It is similar to multiple linear regression analysis but differs in that (a) it considers censored data and (b) the outcome is time to a dichotomous endpoint such as remission/no remission. The analysis can be adjusted for certain baseline factors such as severity of illness, age at operation, etc. It determines risk factors and their statistical significance and generates relative risks (hazard ratios) with associated 95% confidence intervals, adjusted for all other covariates.

The model assumes that the ratio of the hazard functions for any two individuals over time should be parallel. It is the opinion of many statisticians that this assumption must be satisfied for verification of the model. There are various methods to test whether these assumptions hold.

Note: The preceding Kaplan-Meier product-limit estimator and the Cox proportional hazard-regression analysis recommendations are predicated on the assumption that relapses are relatively uncommon, as they are after maximal thymectomy [3] and that the "event" is therefore time to first remission as defined. However, if relapses are more common, as may be the case after medical therapy, and perhaps following the more limited thymic resections, a separate analysis would be needed for those patients who subsequently relapse and are therefore not in remission at the end of the study. The endpoint would then be the time of the relapse. Because deaths due to myasthenia gravis are now relatively uncommon, regardless of the form of therapy, they are being considered "censored events". They should, of course, be recorded. If deaths are more common, a "competing risk" analysis would need to be employed. [4]

The UAB Multi-Phase Hazard Regression Model- This method, developed at the University of Alabama at Birmingham by Blackstone, Naftel and Turner, [5] is similar to the Cox Proportional Hazard model. It differs, however, in that it divides the survival experience into at most three distinct hazard phases; early, constant and late. A graphic display of the changing hazard rate (instantaneous rate of remission) across time can be generated. For each phase, a multivariable analysis can be performed. This technique has proved helpful in analyzing the results of cardiac surgery and may prove helpful in analyzing the results of therapy, especially thymectomy, for myasthenia gravis.

The Hazard Rate- This is a univariate estimate of the instantaneous risk (hazard) of an event across time. [6, 7] It is estimated simply as the number of events (remissions) divided by the total duration of follow-up times experienced by all patients in the study group. Therefore, it is expressed as a rate (the number of events per individual at risk per unit time). When comparing two remission rates, the analysis is appropriate only if the hazard rates are constant (*linearized rate*) over time. This may or may not be the case following thymectomy for myasthenia gravis. If this analytic technique is employed, it is recommended that it be used only as a supplement to the Kaplan-Meier analysis and not as a substitute.

Crude Rates- Crude rates should be discontinued in comparing results of therapy, especially thymectomy. It is an incidence rate, which is defined, for instance, as the number of remissions divided by the number of thymic resections, times 100%; or as the number of remissions divided by the number of cases followed, times 100%. These cannot be used interchangeably nor to compare one series to another, since there are no corrections for length of follow-up. They can only be used for comparative analysis if all patients are followed after initiation of therapy or postoperatively for the same length of time relative to disease onset. Under these circumstances, the statistical method of analysis would be the Chi-Square or the Fisher Exact Test; a multivariable approach would be logistic regression. Otherwise crude rates have no place in the comparative analysis of these patients.

B- Instruments to Evaluate Clinical Improvement

In addition to the survival instruments described above, there are a number of methods of capturing and analyzing the intermediate profile (clinical improvement) of patients. These are

especially applicable for those patients who do not go into remission following therapy. The following is a partial list of available options, organized by the type of variable. Many of these analytical techniques are reviewed in standard statistical texts.[8] Individual references are sighted for the more advanced methods of analysis.

Measurements for Continuous Data- Paired and unpaired Student t-tests for means and/or their non-parametric analogues. Analysis of Variance methods (ANOVA), including one-way, two-way, repeated measures and Analysis of Covariance (ANCOVA) for more sophisticated designs.[9] Multivariable methods include multiple linear regression analysis [10] and the Generalized Estimating Equation (GEE) technique for longitudinal data. [11]

Measurement for Discrete Data- Methods include Chi-Square, Fisher's exact test, and McNemar's Chi-Square test for paired data. [12] Multivariable methods include logistic regression [13] and again GEE modeling, including a recent model for clustered ordinal measurements. [14]

II- Functional Status & Quality of Life* Instruments

**The concept of Quality of Life and Quality of Care (as reflected in patient satisfaction and questionnaires) are distinct.*

Functional status measures offer an objective look at patient physical capabilities. In comparison to functional status measures, health status and preference measures capture a broader impact of an illness and its treatment on the well-being of patients. There are two types of health-related quality of life measures: health status indices and patient preference measures.

Functional Status- These measurements characterize the functional capabilities or disabilities of patients. These can be objectively measured by patient examination or subjectively derived from questionnaires which are completed by either the patient or researcher. The Myasthenia Gravis Activities of Daily Living Score is an example of a disease specific functional status measurement score and has been shown to correlate with the myasthenia gravis QMG Score. [15]

Health Status Indices- These are multi-dimensional instruments which characterize quality of life as a series of scores for relevant dimensions or domains (e.g., emotional well being, mobility, self care, pain, anxiety). There are both generic and disease specific measures.

Generic Measures -Three such health profile instruments are the Medical Outcomes Study SF36, [16] the Children's Health Questionnaire, [17] and the Sickness Impact Profile (SIP). [18, 19] These measures are useful for characterizing the impact of disease and treatments on various aspects of health related quality of life.

Disease-Specific Measures –These questionnaires focus on the aspects of quality of life that are affected by one disease. Consequently, they are more responsive to small changes in QoL and therefore more desirable than generic instruments in longitudinal studies of a particular disease. A disease-specific

instrument for myasthenia gravis should be developed since there are none in existence at this time. Steps necessary to develop such a measure have been defined. [20]

Preference (Utility) Measures- These are measurements that capture the overall value or preference that a patient holds for a particular health outcome. [21] Such measures are expressed as numeric values on a uniform scale (0 to 1). They are particularly useful for summarizing overall changes in health related quality of life as they are expressed as a single numeric score. Most importantly, preference scores can serve as quality adjustment factors for calculating quality adjusted survival, measured in quality adjusted life years (QALY). The EuroQoL [22] and Health Utility Index (HUI) [23] are two examples of questionnaires which can be administered to patients to characterize their abilities or disabilities and from which a numeric preference scores for the health states can be calculated.

III- Quality Adjusted Survival & Cost Effectiveness

Quality Adjusted Life Years (QALY)- This measure aggregates both the survival and quality of life experienced by patient or cohort and is expressed as a single numeric value. Patients that survive a year at "full health" have experienced 1 quality adjusted life year (1 QALY). Patients who survive a year but experience only "half of full health" have experienced half of a quality adjusted life year (0.5 QALY). [24] To perform the QALY calculation each of the states of health experienced by the patient during the follow-up period must be valued on a uniform scale from 0 (death) to 1 (full health). States of health perceived to be worse than death have values less than zero. This form of measurement is quite valuable when comparing treatment strategies that have different effects on survival and quality of life, i.e., when one strategy is not dominant in both its effect on survival and quality of life.

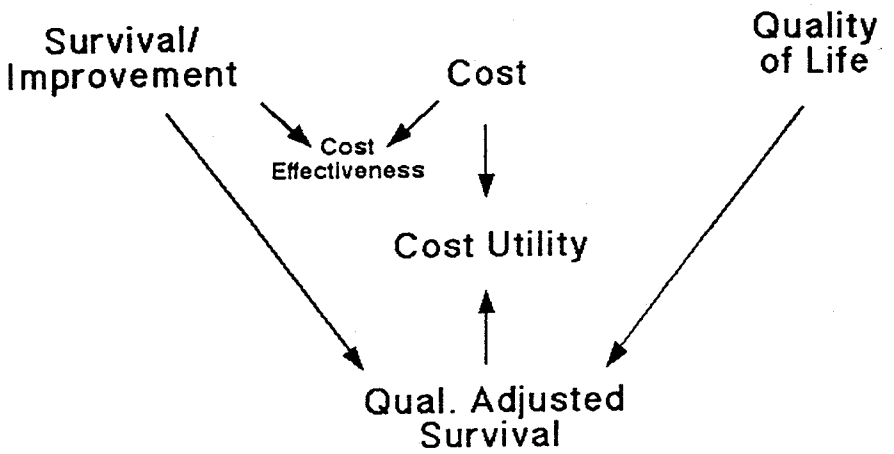
Cost- These are the health care dollars that are used in the care of an illness. They may be measured as charges, payments, or true costs. These are necessary to determine the impact of a disease or proposed treatment on the overall health care budget. There are a variety of techniques for cost calculation. Some studies rely on hospital or payer accounting systems that are in place and some are based on primary collection of health resource use, which are translated into dollars spent or saved.

Cost-effectiveness- This is an analytical technique that looks at the incremental rate paid to obtain an additional unit or measure of health. Such rates are valuable for appropriating health care funds because they allow you to determine the amount of health that your health care dollars can be expected to yield, and, thus, allow you to find optimal way to spend health care dollars to maximize health. This form of analysis is particularly relevant in comparing treatment strategies where one is both more effective and more costly. Under such circumstances, one sees how much more, on average, it cost to treat a patient by the more effective treatment (the incremental cost) and, on average, how much more health each patient obtained (the incremental effectiveness). Typically, costs are expressed as units of currency (e.g., dollars) and effectiveness is expressed in a relevant outcome measure (e.g., years of life saved or years of quality adjusted

life saved). The ratio of the two terms (incremental cost/incremental effectiveness) is the rate paid for each additional measure of effectiveness obtained (e.g. dollars/life-year saved). When quality of life is taken into consideration, the term cost-utility ratio is often used. [25, 26]

Figure

The Interrelationship between Outcomes Instruments--The interrelationship between Survival-Improvement Instruments, the Quality of Life Instruments, and the Quality Adjusted Survival-Cost Effective Instruments is demonstrated.



References

1. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-481.
2. Cox DR. Regression models and life tables (with discussion). J Royal Statis Soc 1972;34:187-220.
3. Jaretzki A, III. Thymectomy for myasthenia gravis: an analysis of the controversies regarding technique and results. Neurology 1997;48(Suppl 5):S52-S63.
4. Allison P. Survival analysis using the SAS system: a practical guide. Cary, NC: SAS Institute Inc. 1995:292.

5. Blackstone EH, Naftel DC, Turner ME. The decomposition of time-varying hazard into phases, each incorpoorating a seeparate stream of concomitant information. AM Statist Assoc 1986;81:615-624.
6. Bodnar E, Frater RWM. Replacement Cardiac Valves. Pergamon Press, Inc 1991:391-434.
7. Edmunds LH, Clark RE, Cohn LH, Miller DC, Weisel RD. Guidelines for reporting morbidity and mortality after cardiac valvular operations. J Thorac Cardiovasc Surg 1988;96:351-353.
8. Daniel WW. A foundation for analysis in the health sciences. 7th. ed. John Wiley & Sons. New York 1998.
9. Fleiss JL. The design and analysis of clinical experiments. John Wiley & Sons, New York 1986.
10. Neter J, Wasserman W, Kutner MH. Applied linear regression models. 2nd. ed. Richard D. Irwin. Homewood, IL 1983.
11. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrica 1986;73:13-22.
12. Fleiss JL. Statistical methods for rates and proportions. 2nd. ed. John Wiley & Sons, New York 1981.
13. Hosmer DW, Lemeshow S. Applied logistic regression. John Wiley & Sons. New York 1989.
14. Haegerty PJ, Zeger SL. Marginal regression models for clustered ordinal measurements. J Am Stat Assoc 1996;91:1024-1036.
15. Wolfe GI, Herbelin L, Nations SP, Foster B, Bryan WW, Barohn RJ. Myasthenia gravis activities of daily living profile. Neurology 1999;52:1487-1489.
16. Brazier JE, Harper R, Jones NM, al e. Validating the SF-36 health survey questionaire: new outcome measures for primary care. Brit Med Jour 1992;305:160-164.
17. Landgraf JM, Abetz L, Ware JE. The Childrfen's Health Questionnaire. 1st. ed. The Health Institute, New England Medical Center. Boston, MA 1996.
18. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: developement and final revision of a health status measure. Medical Care 1981;19:787-805.
19. McGuire D, Garrison L, Armon C, Barohn R, Bryan W, Miller R, Parry G, Petajan J, Ross M, Group SCAS. Relationship of the Tufts Quantitative neuromuscular exam (TQNE) and the sickness impact profile (SIP) in measureing progression of ALS. Neurology 1996;46:1442-1444.
20. Jaeschke R, Guyatt GH. How to develop and validate a new quality of life instrument. Quality of Life Assessments in Clinical Trials, edited by B. Spilker. Raven Press, Ltd. New York 1990;Chapter 5:47-57.
21. Torrance G. Utility approach to measuring health-related quality of life. J Chron Dis 1987;40:593-600.
22. EuroQol G. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
23. Torrance G, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions: Health Utility Index. Pharmaco Economics 1995;7:503-520.
24. Torrance G, Feeny D. Utilities and quality-adjusted life years. Int J of Technology Assessment 1989;5:559-575.
25. Drummond M, Stoddart G, Torrance G. Methods for the economic evaluation of health care programs. Oxford Medical Publications. 1st ed. 1987.
26. Balas E, Kretschmer R, Gnann W, al e. Interpreting cost analysis of clinical interventions. JAMA 1998;279:54-57.