

## RECOMMENDATIONS FOR MYASTHENIA GRAVIS CLINICAL TRIALS

MICHAEL BENATAR, MBChB,<sup>1</sup> DONALD B. SANDERS, MD,<sup>2</sup> TED M. BURNS, MD,<sup>3</sup> GARY R. CUTTER, PhD,<sup>4</sup> JEFFREY T. GUPTILL, MD,<sup>2</sup> FULVIO BAGGI, PhD,<sup>5</sup> HENRY J. KAMINSKI, MD,<sup>6</sup> RENATO MANTEGAZZA, MD,<sup>5</sup> MATTHEW N. MERIGGIOLI, MD,<sup>7</sup> JOANNE QUAN, MD,<sup>8</sup> GIL I. WOLFE, MD<sup>9</sup> and THE TASK FORCE ON MG STUDY DESIGN OF THE MEDICAL SCIENTIFIC ADVISORY BOARD OF THE MYASTHENIA GRAVIS FOUNDATION OF AMERICA

<sup>1</sup> Department of Neurology, University of Miami School of Medicine, Miami, Florida, USA

<sup>2</sup> DUMC 3403, Division of Neurology, Duke University Medical Center, Durham, North Carolina, USA

<sup>3</sup> Department of Neurology, University of Virginia Medical School, Charlottesville, Virginia, USA

<sup>4</sup> University of Alabama School of Public Health, Birmingham, Alabama, USA

<sup>5</sup> Neurology IV, Fondazione Istituto Neurologico "Carlo Besta," Milan, Italy

<sup>6</sup> Department of Neurology, George Washington University School of Medicine, Washington, DC, USA

<sup>7</sup> Department of Neurology and Rehabilitation, University of Illinois College of Medicine, Chicago, Illinois, USA

<sup>8</sup> Bayhill Therapeutics, Inc., Palo Alto, California

<sup>9</sup> Dept. of Neurology, University at Buffalo/SUNY School of Medicine and Biomedical Sciences, Buffalo, NY

Accepted 24 January 2012

**ABSTRACT:** The recommendations for clinical research standards published in 2000 by a task force of the Medical Scientific Advisory Board (MSAB) of the Myasthenia Gravis Foundation of America (MGFA) were largely successful in introducing greater uniformity in the recording and reporting of MG clinical trials. Recognizing that changes in clinical trial design and implementation may increase the likelihood that new therapies are developed for MG, the MGFA MSAB Task Force here presents updated recommendations for the design and implementation of clinical trials in MG, including (a) the use of a quantitative measure, such as the MG-Composite, that is weighted for clinical significance and incorporates patient reported outcomes; (b) consideration of nontrial strategies; and (c) development of biomarkers that support mechanistic studies of pharmacotherapies. The hope is that these updated task force recommendations will expedite the development and acceptance of more effective and less noxious therapies for MG.

*Muscle Nerve* 45: 909–917, 2012

In 2000, a task force of the Medical Scientific Advisory Board (MSAB) of the Myasthenia Gravis Foundation of America (MGFA) proposed a series of classification systems and definitions of response to therapy for myasthenia gravis (MG).<sup>1</sup> These recommendations for clinical research standards included the MGFA Clinical Classification (a tool that identifies subgroups of MG patients who share distinct clinical features), the Quantitative MG Scale (QMG—a quantitative score of disease severity), and

the MGFA Post-Intervention Status (a system to classify clinical status after therapy). These research recommendations, designed to achieve greater uniformity in the recording and reporting of MG clinical trials and outcomes research, have been adopted widely. However, the success of therapeutic development for MG has been relatively limited, and there is broad recognition that changes in the design and implementation of clinical trials may increase the likelihood that new therapies are developed for MG. To better define these issues and to update standards for MG clinical research, the Task Force of the MSAB of the MGFA held a Consensus Conference in Durham, North Carolina on November 19–20, 2010 and a follow-up meeting in Stresa, Italy on June 14–15, 2011, which was jointly supported by the MGFA and Fondazione Istituto Neurologico "Carlo Besta" (INCB)/Associazione Italiana Miastenia Grave (AIM). Attendees included neurologists from the United States and Europe, statisticians, industry representatives, patients with MG, and members of the lay board of the MGFA and the European Myasthenia Gravis Association (EuMGA). This report summarizes the content of these conferences and the consensus recommendations of the Task Force (See Supplemental Materials).

**Abbreviations:** AChR, acetylcholine receptor antibody; AIM, Associazione Italiana Miastenia Grave; AUDTC, area under dose time curve; CDE, common data element; EAMG, experimental autoimmune myasthenia gravis; EuMGA, European Myasthenia Gravis Association; GMG, generalized MG; INCB, Fondazione Istituto Neurologico "Carlo Besta"; IVIg, intravenous immunoglobulin; MG, myasthenia gravis; MGC, MG composite; MGFA, Myasthenia Gravis Foundation of America; MSAB, Medical/Scientific Advisory Board; MuSK, muscle specific tyrosine kinase; PIS, Post-intervention status; PLEX, plasma exchange; QMG, quantitative MG score; RCT, randomized controlled trial; SFEMG, single fiber electromyography; TACT, TREAT-NMD Advisory Committee for Therapeutics; TPMT, thiopurine S-methyltransferase

**Key words:** all clinical trials, clinical trials methodology/study design, myasthenia, outcome measures, quality of life

**Correspondence to:** D.B. Sanders; e-mail: donald.sanders@duke.edu

© 2012 Wiley Periodicals, Inc.

Published online 30 January 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/mus.23330

### MYASTHENIA GRAVIS TRIALS: PAST AND PRESENT (ONLINE TABLE 1)

There are 2 published randomized controlled trials (RCTs) for each of azathioprine,<sup>2,3</sup> mycophenolate mofetil,<sup>4,5</sup> and cyclosporine.<sup>6,7</sup> RCTs of methotrexate,<sup>8</sup> eculizumab,<sup>9</sup> and thymectomy<sup>10</sup> are ongoing, and a trial of steroids in ocular myasthenia has begun.<sup>11</sup>

In stable or worsening MG, there have been 3 RCTs of high-dose intravenous immunoglobulin (IVIg)<sup>12–14</sup> and several studies that compared IVIg with plasma exchange (PLEX).<sup>15–17</sup> The overall impression gleaned from review of these trials is

that: (a) in most trials recruitment was much slower and more difficult than anticipated; (b) the therapeutic effect of prednisone alone in improving myasthenic symptoms may be greater than anticipated,<sup>4,5,18</sup> making it more difficult to discern the incremental therapeutic effect of an additional immune suppressive agent; (c) trials tend to be too short for a steroid-sparing effect to become apparent;<sup>4,5,18</sup> (d) trial selection criteria produce a restricted population and create difficulty in generalizing trial results to a broader population; and (e) outcome measures may have been either insufficiently sensitive or too narrowly defined to detect an important therapeutic effect.<sup>4</sup>

## OUTCOME MEASURES (ONLINE TABLE 2)

**Clinical Improvement.** The current trend in clinical assessment of improvement in MG has moved beyond simple ordinal measures that quantify muscle strength and endurance/fatigue such as the QMG score<sup>1,19,20</sup> and MG-manual muscle testing.<sup>21</sup> Interest has shifted toward instruments that assess quality of life<sup>22–25</sup> or scales weighted for clinical significance that incorporate patient-reported outcomes, such as the MG-Composite<sup>26,27</sup> and the INCB-MG score.<sup>28</sup> Because MG symptoms fluctuate and are usually more evident to the patient than to the physician, it is important to incorporate subject responses into trial outcome measures.<sup>29</sup> Moreover, impairment in different domains such as eyelid elevation, swallowing, or breathing are not equally significant, limiting, bothersome, or dangerous; this is the basis for the argument that some functional items should carry more weight than others.

The 2000 MGFA Task Force report recommended using the Post-Intervention Status (PIS) to "...assess the clinical state of MG patients at any time after institution of treatment for MG."<sup>1</sup> The PIS captures information about current clinical status (Complete Stable Remission, Pharmacological Remission and Minimal Manifestation status) and change in clinical status (Improved, Unchanged, Worse, Exacerbation, Death), incorporating therapies used/needed to produce or maintain a particular clinical state.<sup>1</sup> The PIS did not define criteria for "Improvement" and "Worsening," thus it required use of a predefined increase/decrease in a quantitative measure, such as the QMG score, to detect the minimal difference that is clinically important and biologically plausible in a clinical trial. Based on deliberations at the recent consensus conferences, *the Task Force recommends using the MG-Composite as the quantitative measure for determining improvement and worsening* for patients with generalized disease. Available data support the use of a  $\geq 3$ -point change as the criterion for a clinically sig-

nificant treatment effect.<sup>27</sup> The MG-Composite is favored above (a) the QMG, because it is weighted for clinical significance and incorporates patient reported outcomes; and (b) the INCB-MG, given the greater body of literature supporting its development, implementation and validation. In making this recommendation, however, *the Task Force encourages further study of existing measures, for example by incorporating these measures for comparison in future trials, and development of new measures that might also be used to define "Improvement" and "Worsening" on the PIS.*

Recent trials have used a global impression of change as a secondary outcome measure.<sup>4,5</sup> Global impressions of change are integrated measures, whether provided by the patient or the clinician. They are subject to the placebo effect and bias but may provide an overall perspective often missed by tailored instruments.

**Steroid-Sparing Effects.** Although steroid-sparing effects of immune suppressive agents are often studied, little attention has been paid to how this steroid-sparing effect should be measured or quantified. Trials to date have simply measured the steroid dose at a particular point in time,<sup>2,7</sup> the area under the dose-time curve (AUDTC)<sup>8,10</sup> or the proportion of subjects achieving a desired response while taking an acceptably-low dose of prednisone (e.g.,  $\leq 7.5$  mg/day).<sup>4</sup> Depending on how dropouts are handled, the AUDTC for cumulative steroid dosage can under- or over-estimate steroid exposure. Repeated measures analyses are better able to summarize the exposures, but they are often difficult for clinicians to understand, especially when the steroid dose decreases over time. *The Task Force recommends the use of an AUDTC or repeated measures analysis to quantitate steroid exposure,* recognizing that such an approach will not distinguish between daily and alternate day steroid dosing paradigms. A "delayed start" approach may be appropriate, in which measurement of the AUDTC commences only after the steroid-sparing effect of the tested drug is expected to begin.<sup>8</sup>

Because the adverse biological effects of steroids represent the major limitation to the use of these agents, *the Task Force recommends active monitoring for steroid side effects in all clinical trials in which they are used.* These may be quantitated by monitoring frequencies of common steroid side-effects such as weight gain, impaired glucose tolerance, and diabetes, although this approach will likely not identify those side effects that develop mainly after prolonged use, such as cataracts. Also, because some steroid-related adverse events are more significant to the patient (e.g., development of diabetes) than others (e.g., bruising), consideration

**Table 1.** Biomarkers.

Biomarker	Rationale	Potential applications	Limitations
Clinical: Weakness	Clinically relevant marker of the underlying aberrant immune response	Prior successful application in randomized controlled trials	Requires moderate number of subjects, often followed over an extended period of time
Electrophysiological: SFEMG	Sensitive marker of impaired neuromuscular transmission	Pharmacodynamic marker of response to treatment	Limited availability of expertise
Genetic: Pharmacogenomic <sup>48</sup>	Genetic susceptibility may influence efficacy and/or toxicity of therapeutic agent	TPMT metabolizes azathioprine; accelerated activity might predispose to drug toxicity or no response	Direct relevance of pharmacogenomic markers to toxic and/ or therapeutic effects of drugs has not been established
Immunological AChR antibodies MuSK antibodies	Pathogenic antibody in majority of GMG patients	Established utility for diagnosis; potential utility in monitoring response to therapy is not yet established	Single lab to perform all assays at the same time using full dilution to ascertain titers; relationship to disease severity remains uncertain
T-regulatory cells <sup>40</sup>	Evidence of impaired T-regulatory cell suppressive function	Reflection of disease severity/ response to treatment	Studies of T-regulatory cell function in MG are in their infancy
Thymic pathology <sup>50</sup>	Thymic lymphoreticular hyperplasia is common in GMG	Define treatment responsive subgroups	Limited to those undergoing thymectomy; relationship of pathology to therapeutic response is unknown

AChR, acetylcholine receptor antibodies; GMG, generalized myasthenia gravis; MuSK, muscle specific tyrosine kinase; SFEMG, single fiber electromyography; TPMT, thiopurine S-methyltransferase.

should be given to weighting their clinical importance. Steroid toxicity may also be ascertained by incorporating both efficacy (or lack thereof) and drug toxicity in a definition of “treatment failure,” to be used as an outcome measure.<sup>11</sup> The optimal outcome measure for studies using steroid-sparing endpoints in MG will likely depend on the specific question being asked.

Finally, the Task Force recognizes that it is not possible to recommend a single steroid tapering schedule suitable across all clinical trials, and that the tapering strategy should be tailored to the study design.

## BIOMARKERS

Biomarkers are increasingly recognized as a tool to facilitate drug development.<sup>30,31</sup> A biological marker, or biomarker, is defined as a characteristic that is objectively measured and evaluated as an indicator of a normal or pathogenic biological process, or of a biological response to a therapeutic intervention.<sup>32</sup> A valid biomarker might be useful as a measure of disease severity or therapeutic response and would be particularly instructive during early or middle development to help select compounds worthy of further investigation. Biomarkers might also be useful in later development, for example by identifying MG subpopulations with a good response to a therapeutic agent. The principal difficulty with biomarkers is their validation as an indicator of therapeutic effect, which could require extensive therapy-specific study.

Potential biomarkers in MG can be of several different types—clinical, electrophysiological, genetic or immunological (Table 1). Acetylcholine receptor (AChR) antibody titers have been used as a marker of the therapeutic response, but the utility of this measure has not been confirmed. Single fiber EMG (SFEMG) represents the most robust biomarker of neuromuscular transmission, but the paucity of technical expertise limits its widespread use. Biological markers of the underlying aberrant immune response (e.g., T-regulatory cell function) are of great interest but are in their infancy.

Experimental autoimmune MG (EAMG) can be induced in rodents by immunization with AChRs derived from mammals, the electric organ of eels or rays, or peptide fragments of AChR subunits. EAMG recapitulates many of the clinical and immunologic features of human MG.<sup>33,34</sup> Relevant immunological biomarkers in the EAMG animal model effectively reflect the biology of human MG and can be identified and measured serially in experimental animals before and during the development of disease.<sup>34</sup> The limitations that distinguish EAMG from MG, however, must be appreciated: (1) exogenous antigenic stimulation is required to induce disease; (2) stimulation with antigen must be repeated to perpetuate the autoimmune response; (3) the fluctuations of weakness and autoimmune activity that characterize the human disease are not seen in animals; and (4) the rodent immune system is different from humans. Therefore, caution must be used in extrapolating observations in EAMG to human MG.

*The Task Force recommends that future clinical trials incorporate studies of potential biomarkers and develop repositories of biological samples for future use.*

### **TRIAL MECHANICS**

The Task Force recognized the need to improve several aspects of the design and conduct of clinical trials, irrespective of the stage of development.

**Study Population.** A tendency in prior trials was to define inclusion and exclusion criteria so narrowly as to exclude all but a small proportion of patients. This hampered recruitment and created uncertainty about the appropriateness of a therapy for excluded subpopulations. The Task Force recognizes the need to balance selection criteria to yield a relatively homogeneous study population, especially 1 likely to demonstrate efficacy, while avoiding overly restrictive criteria that inhibit recruitment and raise doubt as to broad application of trial data to the larger MG population. The Task Force believes efforts to broaden entry criteria are important.

**Recruitment.** Recruitment of adequate subjects at a sufficiently swift pace is pivotal to a trial's success. After considering several recruitment strategies, the Task Force *recommends: (a) establishment of a network of sites that have a demonstrated track record of successful recruitment and that are "trial ready"; (b) increasing the number of sites beyond that which is anticipated when a study is planned, rather than waiting until recruitment lags before doing so; and (c) defining a priori for each trial a point at which additional sites will be added or the inclusion criteria will be reconsidered, if recruitment lags.*

**Trial Duration.** It is concerning that trials of steroid-sparing agents are not of sufficient duration to demonstrate a steroid-sparing effect. The Task Force recognizes that the duration of follow-up may be constrained by costs but recognizes that it may become more difficult to launch a follow-up trial of a given agent if a prior study ended with inconclusive results. *The Task Force recommends that the duration of trials of steroid-sparing agents be tailored to the anticipated onset of action of the investigative agent, recognizing that if this is impractically long, then the trial may not be feasible.* Investigations may then focus on agents expected to have a speedier onset of action.

**Steroids.** Although their efficacy has not been demonstrated in an RCT, there is general agreement that steroids are effective in the treatment of MG. Their utility, however, is limited by both their short and long-term toxicity. Nevertheless, the availability of such an effective therapy complicates studies of novel therapeutic agents. It would be dif-

ficult, for example, to randomize patients in need of treatment to receive a new drug or placebo without including prednisone in both treatment groups. This would especially be true for experimental agents with a long latency to onset of action. One approach might be to incorporate the need for rescue treatment with steroids into the trial design. For example, the time until the need for steroid therapy (a form of treatment failure) might be used as an outcome measure. If a trial design requires that all participants receive steroids, the treatment effect of steroids alone should not be underestimated.<sup>4,5,18</sup>

**Concomitant and Recent Use of Immune Modulating Therapy.** Most previous clinical trials have required either that patients be naive to immune suppressive/modulating therapy or that such therapy be withdrawn for some minimum period of time (typically at least 2–3 months) before randomization. The rationale is the need to remove the potential confounding effect of the nonexperimental therapy, as well as to minimize complications posed by multiple immune suppressive therapies. Apart from the restrictions that this approach places on recruitment, it also stands in contrast to routine clinical practice—if 1 immune suppressive/modulating agent yields insufficient benefit after an adequate trial, a different agent is substituted or added. An “overlapping” approach is used if a relatively long onset latency is anticipated for the new drug. If a novel immune suppressive/modulating agent is being evaluated in a trial, then it would be appropriate to withdraw any potentially confounding medication well before the trial outcome is measured. In trials of long duration, permitting early use of concomitant immune suppressive/modulating agents followed by their withdrawal during the trial can improve recruitment, more closely mirrors clinical practice and improves generalizability of trial results. However, depending on the biological characteristics of the agent, this approach could bias results; some potent immunosuppressive agents have demonstrated biologic activity for up to 2 years after drug withdrawal.<sup>35</sup>

**Cholinesterase Inhibitor Therapy.** The Task Force recognized several areas of controversy regarding the use of cholinesterase inhibitor therapy during a clinical trial.

1. *The Task Force recommends that cholinesterase inhibitor dosage should be optimized before patients enter a clinical trial.* The rationale is that patients may be overdosed at study entry, and symptoms may improve as a result of tapering these medications. Optimizing cholinesterase inhibitor

dosing before trial enrollment may also reduce dosage adjustments during the study.

2. *The Task Force believes that timing the last dose of cholinesterase inhibitor may be less critical than previously believed, at least in trials that aim to demonstrate a disease modifying effect rather than simply a symptomatic effect.* This conclusion is supported in part by shifting focus from outcome measures of strength/fatigue (e.g., QMG) to more global measures (e.g., MG-Composite) that integrate function and general well being over a more extended period of time. If there remains concern about the timing of cholinesterase inhibitor therapy in the trial design *the Task Force recommends that the latency from cholinesterase inhibitor administration to outcome assessment be included as a covariate in analysis.*
3. *The Task Force also supports allowing dosing adjustments for cholinesterase inhibitor therapy based on symptom severity and disease control during a trial.* This approach mirrors clinical practice. The confounding effect of varying cholinesterase inhibitor doses in a clinical trial may be mitigated by measuring outcomes “off” cholinesterase inhibitor therapy or by including both timing and dosage of cholinesterase inhibitor as covariates in the analysis.

**Regression to the mean.** Is a statistical phenomenon that describes the tendency of extreme measurements to trend toward their average when they are repeated. For example, imagine that eligibility criteria require a QMG score  $\geq 12$ . A patient might be evaluated on several occasions, initially with a QMG score  $< 12$ , but then on 1 occasion is found to have a QMG score  $> 12$  due, for example, to day to day disease variation or measurement error, and is enrolled in the trial. Subsequent QMG scores are likely to be less severe (i.e.,  $< 12$ ) even without effective therapy. A change in QMG score, or any outcome measure, may then be erroneously interpreted as an improvement due to the intervention when it might simply represent regression toward the mean. Several approaches might be adopted to mitigate the impact of regression to the mean. These include: (a) use of an outcome measure that was not used to select participants for inclusion and which correlates poorly with the selection criteria; (b) inclusion of a lead-in phase to document variability in the outcome measure and to produce a baseline that is a truer measure of disease severity. A placebo control validates the difference between 2 groups even in the presence of regression toward the mean, but it fails to accurately assess the amount of change produced by the regression.

*The Task Force recommends that future clinical trials recognize the potential impact of regression to the mean and/or take active steps to mitigate this effect.*

#### **DRUG SELECTION**

In view of the limited size of the MG population and the typically slow rate of recruitment for MG trials, it is advisable to limit competition between studies vying for the same patient subgroups. Recognizing that no individual or organization should *control* which drugs come to trial, the Task Force nevertheless believes it would be valuable to establish a scientific review process, something akin either to what the ALS community has developed to prioritize drugs for clinical trials<sup>36</sup> or the broader neuromuscular community in Europe through the TREAT-NMD Advisory Committee for Therapeutics (TACT).<sup>37</sup> The goal would be to provide transparent and consistent guidance to the neuromuscular community on the readiness of drugs and/or therapeutic targets for trials in MG.

#### **DRUG DEVELOPMENT: EARLY AND LATE DEVELOPMENTAL TRIALS**

Animal models may be valuable in the early stage of drug development by providing evidence that novel immunotherapeutic agents may be useful and appropriate for early phase clinical trials. EAMG offers a potentially useful approach to screening compounds for efficacy and has been used to a limited extent in drug development in MG.<sup>38–41</sup>

The optimal strategy for drug development depends on the breadth, number, and nature of compounds in the experimental pipeline. A sensible general approach is to first establish optimal dosing based on pharmacokinetics and pharmacodynamics and then proceed to preliminary efficacy and safety studies with the goal of eliminating ineffective drugs at an early stage. Useful designs for early developmental trials include futility, pick-the-winner, and some adaptive designs (see below). Two challenges for drug development in MG arise from: (a) the difficulty of distinguishing symptomatic from disease-modifying effects; and (b) latency from the time therapy is initiated to the onset of benefit. A delay of 15 months in 1 of the azathioprine trials is a good example of the latter.<sup>2</sup> Both problems could be addressed by the development of a robust, responsive, valid biomarker that predicts later clinical improvement. For example, an immunologic biomarker could be used to move drugs with a disease-modifying effect into early phase development. At a predetermined point of a clinical trial such a biomarker could help to determine whether a trial will be futile in reaching delayed clinical endpoints and should be stopped early, saving time and expense.

As discussed earlier, experience from MG trials raises questions about methods for selecting drugs for comparative trials, optimal study duration, definition of the study population and selection of outcome measures. Many of these questions might be addressed using appropriate adaptive designs. The central tenet of an adaptive design is that a prespecified change is made to a trial in progress based on observed outcomes. For example, a trial might begin with multiple drugs or a single drug at several different dosages, and the adaptation is to select a particular drug or dosage for continued study. A selection paradigm analogous to that used in the I-SPY2 breast cancer trial<sup>42</sup> would require collaboration by biotech/pharmaceutical companies to compare their drugs in a single trial, but the adaptive design can also help select the correct dosage. Such an approach also offers the advantage of economy by using a common placebo group. We have previously raised those issues surrounding selection criteria and the need to broaden inclusion and exclusion criteria to recruit a more heterogeneous population (see above). Adaptive designs could identify a subset of treatment responders from this broader population. A potential benefit of an adaptive design could be a reduction in sample size, either by using seamless phases (maximizing information from study subjects participating in multiple phases of a trial) or by early stopping rules. Stopping futile treatment arms early also permits allocation of remaining trial participants to treatment arms that hold more promise. The slow recruitment rate of MG trials might actually prove useful if it provides time to observe a treatment response in early enrollees and allows modification of the design before enrollment of the full population. Despite these potential advantages, adaptive design trials are logistically more complex and require greater planning and expense. Planning for middle and late development might mitigate some of the extra costs and shorten drug development time, which would then justify the added complexity. In MG, given the long-term nature of many of the outcomes and the lack of biomarkers or surrogate outcomes, it may not be clear when an adaptive design should be used. Furthermore, the costs of such adaptations must be planned in advance, possibly making it more difficult to secure funding. Any protocol design under consideration must be acceptable to relevant regulatory agencies; a recent FDA Guidance on adaptive trial designs should help in this regard.<sup>43</sup>

#### **STRATEGIES OTHER THAN RANDOMIZED CONTROLLED TRIAL DESIGN**

Randomized controlled trials (RCTs) are not the only method for advancing therapies. This is not

to say that RCTs should not remain the gold standard for evaluating drug efficacy, but rather to emphasize the value of other approaches, particularly for a rare disease such as MG in which large and adequately powered RCTs may not always be feasible. The Task Force discussed the following strategies and approaches that might complement RCTs.

**Patient Registry.** The Task Force recognizes the value of patient registries to facilitate recruitment for trials. A patient registry is "...an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, exposure, and that serves 1 or more predetermined scientific, clinical, or policy purposes".<sup>44</sup> Through a registry, MG patients could indicate interest in trial opportunities, facilitating communication between trial investigators and potential study subjects. The registry should permit transmission of information about trials from investigators to potential participants and also offer registrants broad access to information about all ongoing trials. Ideally, such a registry should be hosted by an independent agency to minimize any proprietary claims or limitations that might result from private ownership. Cross-sectional or one-time surveys are often plagued by many biases, and longitudinal registries that enable assessments of patients over time are preferred.

The lack of a common "technical" language among existing MG registries represents a major drawback in unifying existing local registries into broader national or multi-national databases. Broad agreement on common data elements (CDEs) (see below) would help to overcome this problem. These may be constructed using a layered approach in which the more essential data elements are considered "core," and less crucial data elements are regarded as "supplementary." Such an approach was used in developing the European MG database, a sub-project of the European Myasthenia Gravis Network (EU Grant 2005105 - DG SANCO - Health Information). MG patient advocacy groups should play a pivotal role in disseminating information about registries and encouraging participation. The MGFA is developing a patient driven registry, to be used in support of research, advocacy and public awareness.<sup>45</sup>

**Common Data Elements.** The National Institute of Neurological Disorders and Stroke has launched an ambitious project to establish CDEs for a range of neurological disorders, including neuromuscular disease. The goal is to standardize data collection to more effectively aggregate data and facilitate comparison of results across studies. The Task

Force recognizes that extending this effort to include CDEs for MG databases and trials will be of value in optimizing the use of clinical and trial data for several of the approaches outlined below (see *Retrospective Analyses*). CDEs already collected in existing registries represent the inner core of a wider registry. CDEs can be further enhanced by adding Supplemental Data Elements that are recorded but not required in individual registries, and Exploratory Data Elements, which are under development or not yet validated and are proposed for consideration to the wider clinical community.

**Existing Data Sources.** The Task Force recognizes the potential value of mining existing data sources to develop and validate outcome measures, for preliminary data to justify future RCTs, and to inform sample size calculations in the planning stages of clinical trials. Construction of the MG-QoL15 and the MG-Composite, for example, relied heavily on data from the 2 mycophenolate trials, which were made available for this purpose by the sponsors. Existing data sources might prove valuable for development of other instruments as well. Data from placebo-treated patients in prior trials might also serve as historical controls or provide information for early phase drug development trials including adaptive design and futility studies. Moreover, datasets from prior trials could be of high value, because their accuracy and completeness likely surpass that from clinic databases or registries. *The Task Force recommends establishing a mechanism to pool and share data from completed trials and strongly encourages pharmaceutical and biotech companies to commit to a policy of data sharing once the trial is complete and primary results are published.*

**Retrospective Analyses.** The Task Force considered methods such as blinded prospective review and propensity score matching as potential tools for comparative effectiveness research in the absence of RCT data. The blinded prospective review is designed to mimic an RCT; after initial identification of potentially eligible subjects, a blinded review of potential cases is performed to ensure that participants selected for inclusion in the final study cohort would have been eligible to receive either of the 2 treatments being evaluated. Such an approach has been used to compare endovascular coil embolization and surgical clipping of unruptured cerebral aneurysms<sup>46</sup> but has not been applied to patients with MG. Propensity score matching is a statistical method for analyzing observational data that attempts to remove bias in estimates of treatment effects. It is similar to covariate adjustment in regression analyses, but it matches on predictors of exposure to the drug to create a model that predicts receipt or no receipt

of the therapy under study. In theory, balancing the groups for predictors of exposure to drug eliminates some biases that confound nonrandomized studies. Once the model predicting the choice of therapy is created, subjects are matched on their propensity to receive drug “A” or not. If 2 patients are identical in presentation, they have the same chance of receiving the drug, similar to randomization. The limitation of this approach is that it cannot adjust for what has not been measured or for unknown determinants of offering a particular treatment.<sup>47</sup>

#### **TRIAL ETHICS**

All protocols should discuss the risks of withholding therapy that is considered standard of care, e.g., cholinesterase inhibitors, steroids, thymectomy, and immunosuppressive medications, especially if there is evidence that delaying any of these could adversely affect the ultimate outcome. Adequate rescue therapy must be a part of any protocol. Investigators and study subjects should be fully informed about posttrial access to agents that are demonstrated to be effective before regulatory approval, and if subjects will have access to open-label study drug if a trial is terminated early for considerations other than safety/tolerability. Timely publication of results is also an ethical responsibility of both investigators and sponsors.

#### **RECOMMENDATIONS**

The Task Force’s recommendations are summarized here, with more details in the relevant sections of this manuscript.

##### **1. Outcome Measures**

- a. The MG-Composite should be used as the quantitative measure for determining the PIS definitions of Improvement and Worsening. A  $\geq 3$ -point change is considered clinically meaningful.
- b. The Task Force encourages further study of existing measures (e.g., the INCB-MG) and development of new ones to better define Improvement and Worsening on the PIS.
- c. An AUDTC or repeated measures analysis should be used to quantitate steroid exposure in trials of steroid-sparing agents.
- d. Active monitoring for steroid side effects should be included in all clinical trials in which they are used.

##### **2. Biomarkers**

- a. Priorities for research in MG include developing biomarkers capable of supporting mechanistic studies of pharmacotherapies and monitoring disease severity. Steps to facilitate biomarker development include incorporation of biomarker studies into clinical trials

- and establishment of a repository of biological samples collected from trial participants.
- b. SFEMG is currently the best biomarker for monitoring disease activity in MG. Because the requisite expertise is not widely available, which limits its use, SFEMG would be most useful in smaller, early phase studies.
3. Trial Mechanics.
    - a. The Task Force recommends (i) establishment of a network of sites that have a demonstrated track record of successful recruitment and are “trial ready,” (ii) increasing the number of sites beyond that which is anticipated when a study is planned, rather than waiting until recruitment lags before doing so, and (iii) defining *a priori* for each trial a point at which additional sites will be added or the inclusion criteria will be reconsidered, if recruitment lags.
    - b. Investigators should attempt to broaden trial inclusion and exclusion criteria, with the goal of enrolling a study population that resembles the population of MG patients in need of new therapies.
    - c. Trials of steroid-sparing agents should be of sufficient duration to have a reasonable expectation that the steroid-sparing effect will be demonstrable.
    - d. Cholinesterase inhibitor dosage should be optimized before patients enter a clinical trial, and the latency from cholinesterase inhibitor to outcome assessment should be included as a covariate in analysis.
    - e. Protocols should consider allowing concurrent use of immunosuppressive therapies early in MG trials, provided that they are withdrawn within time to minimize the risk of confounding outcome measurements.
    - f. Investigators and study sponsors should commit to timely publication of results from all trials.
    - g. Study sponsors should commit to a policy of data sharing after a trial is completed and primary results have been published.
    - h. A mechanism should be established to pool and share study data from completed trials.
  4. Drug Selection
    - a. An unbiased and transparent scientific review process would benefit the research community in prioritizing drugs for clinical trials.
  5. Trial Design
    - a. The merits of an adaptive design approach should be considered in development of future MG RCTs.
    - b. Common data elements (CDEs) should be developed and used in future MG trials (as well as in registries).

- c. Future clinical trials should recognize the potential impact of regression to the mean and/or take active steps to mitigate this effect.
6. Non-Trial Strategies
    - a. An MG registry, with prospective ascertainment of patients, should be developed in which they register their interest in being contacted about research opportunities.

\*MGFA Task Force on MG Study Design. Chair: Donald B. Sanders; Co-Chair – Michael Benatar; Members: Richard J. Barohn, Ted M. Burns, Gary R. Cutter, Jeffrey T. Guptill, James F. Howard, Jr., Henry J. Kaminski, Renato Mantegazza, Matthew N. Merrigioli, Steven Piantadosi, Alexis Smirnow, and Gil I. Wolfe. Dr. Benatar has received consultation fees from Bayhill Therapeutics and Cytokinetics Inc. and receives grant support from the Food & Drug Administration, the Muscular Dystrophy Association and the ALS Association. Dr. Sanders is on the Speakers Panel for Athena Diagnostics and is a consultant for Accordant Health Services, Alexion Pharmaceuticals Inc., Cytokinetics Inc., GlaxoSmithKline, Jacobus Pharmaceuticals and Talecris Biotherapeutics. Dr. Burns is a consultant for Talecris Biotherapeutics. Dr. Cutter is a member of Data and Safety Monitoring Committees for Biogen-Idex, Sanofi-Aventis, Cleveland Clinic, Daichi-Sankyo, GlaxoSmithKline Pharmaceuticals, Genmab Biopharmaceuticals, EliLilly, Medivation, Modigenetech, Ono Pharmaceuticals, PTC Therapeutics, Teva, Vivus, University of Pennsylvania, NHLBI, NINDS and NMSS, and has received consulting, speaking or advisory board fees from Alexion Pharmaceuticals Inc., Bayhill Therapeutics, Bayer, Novartis, Nuron Biotech, Consortium of MS Centers (grant), Genzyme, Klein-Buendel Incorporated, Peptimmune, Somnus Pharmaceuticals, Sandoz, Spiniflex Pharmaceuticals, Teva Pharmaceuticals, University of Texas Southwestern and Visioneering Technologies, Inc. Dr. Guptill receives funding from the American Academy of Neurology Foundation and the Myasthenia Gravis Foundation of America. Dr. Kaminski is a consultant for Bayhill Therapeutics, Cytokinetics, GlaxoSmithKline and Varleigh Limited. His research has been supported by NIH EY14837. Drs. Baggi and Mantegazza have nothing to disclose. The European Database on Myasthenia Gravis is funded by EU-FP6 DG-SanCo Contract No. 2005105 (Euromyasthenia Gravis Network) and EU-FP7 HEALTH-F2-2010-242210 (FIGHT-MG). Dr. Merigioli is a consultant for Athena Diagnostics, Talecris Biotherapeutics and Gliknik Inc. His research is funded by NIH K08NS058800 and the Muscular Dystrophy Association. Dr. Quan is an employee of Bayhill Therapeutics. Dr. Wolfe is a consultant for Bayhill Therapeutics and Talecris Pharmaceuticals Inc. This work was supported by a grant from the Myasthenia Gravis Foundation of America.

## REFERENCES

1. Jaretzki A, Barohn RB, Ernstoff RM, Kaminski HJ, Keeseey JC, Penn AS, et al. Myasthenia gravis. Recommendations for clinical research standards. Task Force of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America. *Neurology* 2000;55:16–23.
2. Palace J, Newsom-Davis J, Lecky B, The Myasthenia Gravis Study Group. A randomized double-blind trial of prednisolone alone or with a zathioprine in myasthenia gravis. *Neurology* 1998;50:1778–1783.
3. Myasthenia Gravis Clinical Study Group. A randomised clinical trial comparing prednisone and azathioprine in myasthenia gravis. Results of the second interim analysis. *J Neurol Neurosurg Psychiatry* 1993;56:1157–1163.
4. Sanders DB, Hart IK, Mantegazza R, Shukla SS, Siddiqi ZA, De Baets MHV, et al. An international, phase III, randomized trial of mycophenolate mofetil in myasthenia gravis. *Neurology* 2008;71:400–406.



5. Muscle Study Group. A trial of mycophenolate mofetil with prednisone as initial immunotherapy in myasthenia gravis. *Neurology* 2008; 71:394–399.
6. Tindall RSA, Rollins JA, Phillips JT, 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. Tindall RSA, Phillips JT, Rollins JA, Wells L, Hall KA. A clinical therapeutic trial of cyclosporine in myasthenia gravis. *Ann N Y Acad Sci* 1993;681:539–551.
8. Barohn RJ. Efficacy of methotrexate in myasthenia gravis. In: *ClinicalTrials.gov* [Internet]. Bethesda (MD): National Library of Medicine (US). [Accessed 2008 Dec 23].
9. Alexion Pharmaceuticals. Safety and efficacy study of eculizumab in patients with refractory generalized myasthenia gravis. In: *ClinicalTrials.gov* [Internet]. Bethesda (MD): National Library of Medicine (US). [Accessed 2008 July 30].
10. Aban IB, Wolfe GI, Cutter GR, Kaminski HJ, Jaretzki A, Minisman G, et al. The MGTX experience: challenges in planning and executing an international, multicenter clinical trial. *J Neuroimmunol* 2008; 201:80–84.
11. Benatar M. Efficacy of prednisone in the treatment of ocular myasthenia (EPITOME). In: *ClinicalTrials.gov* [Internet]. Bethesda (MD): National Library of Medicine (US). [Accessed 2009 Oct 14].
12. Wolfe GI, Barohn RJ, Foster BM, Jackson CE, Kissel JT, Day JW, et al. Randomized, controlled trial of intravenous immunoglobulin in myasthenia gravis. *Muscle Nerve* 2002;26:549–552.
13. Gajdos P, Tranchant C, Clair B, Bolgert F, Eymard B, Stojkovic T, et al. Treatment of myasthenia gravis exacerbation with intravenous immunoglobulin. A randomized double-blind clinical trial. *Arch Neurol* 2005;62:1689–1693.
14. Zinman L, Ng E, Bril V. IV immunoglobulin in patients with myasthenia gravis: a randomized controlled trial. *Neurology* 2007;68:837–841.
15. Ronager J, Ravnborg M, Hermansen I, Vorstrup S. Immunoglobulin treatment versus plasma exchange in patients with chronic moderate to severe myasthenia gravis. *Artif Organs* 2001;25:967–973.
16. Gajdos P, Chevret S, Clair B, Tranchant C, Chastang C. Clinical trial of plasma exchange and high-dose intravenous immunoglobulin in myasthenia gravis. Myasthenia Gravis Clinical Study Group. *Ann Neurol* 1997;41:789–796.
17. Barth D, Nabavi Nouri M, Ng E, Nwe P, Bril V. Comparison of IVIg and PLEX in patients with myasthenia gravis. *Neurology* 2011;76:2017–2023.
18. Sanders DB, Siddiqi ZA. Lessons from two trials of mycophenolate mofetil in myasthenia gravis. *Ann N Y Acad Sci* 2008;1132:249–253.
19. Bedlack RS, Simel DL, Bosworth H, Samsa G, Tucker-Lipscomb B, Sanders DB. Quantitative myasthenia gravis score: assessment of responsiveness and longitudinal validity. *Neurology* 2005;64:1968–1970.
20. Barohn RJ, McIntire D, Herbelin L, Wolfe GI, Nations S, Bryan WW. Reliability testing of the quantitative myasthenia gravis score. *Ann N Y Acad Sci* 1998;841:769–772.
21. Sanders DB, Tucker-Lipscomb B, Massey JM. A simple manual muscle test for myasthenia gravis: validation and comparison with the QMG score. *Ann N Y Acad Sci* 2003;998:440–444.
22. Paul RH, Nash JM, Cohen RA, Gilchrist JM, Goldstein JM. Quality of life and well-being of patients with myasthenia gravis. *Muscle Nerve* 2001;24:512–516.
23. Padua L, Evoli A, Aprile I, Caliandro P, Mazza S, Padua R, et al. Health-related quality of life in patients with myasthenia gravis and the relationship between patient-oriented assessment and conventional measurements. *Neurol Sci* 2001;22:363–369.
24. Mullins LL, Carpentier MY, Paul RH, Sanders DB; Muscle Study Group. A disease-specific measure of quality of life for myasthenia gravis. *Muscle Nerve* 2008;38:947–956.
25. Burns TM, Conaway MR, Cutter GR, Sanders DB; Muscle Study Group. Less is more, or almost as much: a 15-item quality-of-life instrument for myasthenia gravis. *Muscle Nerve* 2008;38:957–963.
26. Burns TM, Conaway M, Cutter GR, Sanders DB; Muscle Study Group. Construction of an efficient evaluative instrument for myasthenia gravis: the MG composite. *Muscle Nerve* 2008;38:1553–1562.
27. Burns TM, Conaway M, Sanders DB; MG Composite and MG-QOL, 15 Study Group. The MG composite: a valid and reliable outcome measure for myasthenia gravis patient care. *Neurology* 2010; 74:1434–1440.
28. Mantegazza R, Antozzi C, Peluchetti D, Sghirlanzoni A, Cornelio F. Azathioprine as a single drug or in combination with steroids in the treatment of myasthenia gravis. *J Neurol* 1988;235:449–453.
29. Burns TM. History of outcome measures for myasthenia gravis. *Muscle Nerve* 2010;42:5–13.
30. US Food and Drug Administration. Innovation or stagnation: Challenge and opportunity on the critical path to new medicinal products. In: [FDA.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm](http://FDA.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm). [Internet] [Accessed 2011 Apr 20].
31. US Food and Drug Administration. Innovation or stagnation: critical path opportunities list. In: [FDA.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/default.htm](http://FDA.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/default.htm). [Internet] [Accessed 2011 Apr 20].
32. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69:89–95.
33. Lindstrom JM. Acetylcholine receptors and myasthenia. *Muscle Nerve* 2000;23:453–477.
34. Christadoss P, Poussin M, Deng C. Animal models of myasthenia gravis. *Clin Immunol* 2000;94:75–87.
35. Keymeulen B, Vandemeulebrouck C, Kaufman L, Hale G, Gorus F, Goldman M, et al. Insulin needs after CD3-antibody therapy in new-onset Type 1 diabetes. *N Engl J Med* 2005.
36. Traynor BJ, Bruijn L, Conwit R, Beal F, O'Neill G, Fagan SC, et al. Neuroprotective agents for clinical trials in ALS: a systematic assessment. *Neurology* 2006;67:20–27.
37. TACT: the TREAT-NMD Advisory Committee for Therapeutics. In: [TREAT-NMD.eu/resources/TACT/introduction/](http://TREAT-NMD.eu/resources/TACT/introduction/) [Internet]. [Accessed 2011 Aug 31].
38. Rowin J. Etanercept treatment in myasthenia gravis. *Ann N Y Acad Sci* 2008;1132:300–304.
39. Janssen SPM, Phernambucq M, Martinez P, De Baets MH, Losen M. Immunosuppression of experimental autoimmune myasthenia gravis by mycophenolate mofetil. *J Neuroimmunol* 2008;201:111–120.
40. Zhou Y, Gong B, Lin F, Rother RP, Medof ME, Kaminski HJ. Anti-c5 antibody treatment ameliorates weakness in experimentally acquired myasthenia gravis. *J Immunol* 2007;179:8562–8567.
41. Hepburn NJ, Chamberlain-Banoub JL, Williams AS, Morgan BP, Harris CL. Prevention of experimental autoimmune myasthenia gravis by rat Cry-Ig: a model agent for long-term complement inhibition in vivo. *Mol Immunol* 2008;45:395–405.
42. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin Pharmacol Ther* 2009;86:97–100.
43. US Food and Drug Administration. Guidance for industry adaptive design clinical trials for drugs and biologics. In: [FDA.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm](http://FDA.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm). [Internet] [Accessed 2011 Apr 20].
44. Gliklich RE, Dreyer NA. Registries for evaluating patient outcomes: a user's guide. 2nd ed. AHRQ Publication No.10-EHC049. Rockville, MD: Agency for Healthcare Research and Quality. 2010.
45. Myasthenia Gravis Foundation of America. In: [myasthenia.org](http://myasthenia.org) [Internet] [Accessed 2011 Aug 31].
46. Johnston SC, Wilson CB, Halback VV, et al. Endovascular and surgical treatment of unruptured cerebral aneurysms: comparison of risks. *Ann Neurol* 2000;48:11–19.
47. D'Agostini RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998;17:2265–2281.
48. Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. *N Engl J Med* 2011;364:1144–1153.
49. Ballandina A, Sandrine L, Darteville P, Saoudi A, Berrih-Aknin S. Functional defect of regulatory CD4+CD25+ T cells in the thymus of patients with autoimmune myasthenia gravis. *Blood* 2005;105:735–741.
50. Strobel P, Moritz R, Leite MI, Willcox N, Chuang WY, Gold R, et al. The ageing and myasthenic thymus: a morphometric study validating a standard procedure in the histological workup of thymic specimens. *J Neuroimmunol* 2008;201–202:64–73.