Developing an international consensus guidance for myasthenia gravis using RAND/UCLA appropriateness method

Wei-Bin Liu1,4, Hao Ran2, Chuang-Yi Ou1, Li Qiu1, Zhi-Dong Huang1, Zhong-Qiang Lin1, Yin-Kai Li1,4, Xiao-Xi Liu1,4, Hao Huang1,4, Wei Fang3

1Department of Neurology, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, Guangdong, China.
2School of Pharmacy, Sun Yat-Sen University, Guangzhou 510275, Guangdong, China.
3School of Earth and Environmental Sciences, Queens College, City University of New York, NY 11367, USA.
4Guangdong Provincial Key Laboratory for Diagnosis and Treatment of Major Neurological Diseases, Guangzhou 510080, Guangdong, China.

Correspondence to: Dr. Wei Fang, School of Earth and Environmental Sciences, Queens College, City University of New York, 65-30 Kissena Blvd, Flushing, NY 11367, USA. E-mail: wei.fang@qc.cuny.edu


Aim: Myasthenia gravis (MG) is a rare and heterogeneous disease for which there is no generally accepted standard of care. Thus, it is critical that MG experts develop consensus guidelines based on their practice and disease management to assist clinicians and provide advice for insurance companies, health organizations and institutional review boards.

Methods: An international treatment guidance was developed based on national guidelines established in the US, Denmark, Norway, Germany, Japan, Netherlands, United Kingdom and Europe. The RAND/UCLA appropriateness method (RAM) was applied to reach consensus among 15 worldly renowned experts and experienced clinicians.

Results: This paper introduced the RAM procedure with its principles and applications and conducted a brief review of the resulting 2016 international consensus guidance for MG in comparison to clinical experience and management of Chinese MG patients.

Conclusion: The 2016 international consensus guidance is a major contribution to the treatment and management of MG, providing an up-to-date expert consensus to assist clinicians around the world, especially those with limited experience and/or practice in countries/regions that have limited resources to develop local treatment guidelines. It is also an important contribution showing how RAM can help to develop consensus guidance for treatment of rare diseases based on scientific findings and expert experience.

Key words: International treatment guidance, myasthenia gravis, RAND/UCLA appropriateness method, developing consensus

ABSTRACT

Aim: Myasthenia gravis (MG) is a rare and heterogeneous disease for which there is no generally accepted standard of care. Thus, it is critical that MG experts develop consensus guidelines based on their practice and disease management to assist clinicians and provide advice for insurance companies, health organizations and institutional review boards.

Methods: An international treatment guidance was developed based on national guidelines established in the US, Denmark, Norway, Germany, Japan, Netherlands, United Kingdom and Europe. The RAND/UCLA appropriateness method (RAM) was applied to reach consensus among 15 worldly renowned experts and experienced clinicians.

Results: This paper introduced the RAM procedure with its principles and applications and conducted a brief review of the resulting 2016 international consensus guidance for MG in comparison to clinical experience and management of Chinese MG patients.

Conclusion: The 2016 international consensus guidance is a major contribution to the treatment and management of MG, providing an up-to-date expert consensus to assist clinicians around the world, especially those with limited experience and/or practice in countries/regions that have limited resources to develop local treatment guidelines. It is also an important contribution showing how RAM can help to develop consensus guidance for treatment of rare diseases based on scientific findings and expert experience.

For reprints contact: service@oaepublish.com
INTRODUCTION

Myasthenia gravis (MG) is a neuromuscular transmission disorder. The incidence of MG ranges between 0.3-2.8 per 100,000, affecting more than 700,000 people worldwide. The prognosis for patients with MG has been improved tremendously in recent years due to the increasing use of immunomodulating treatments. However, there is no optimal treatment approach for all patients due to disease heterogeneity, thus an internationally recognized standard of care for MG is still missing.

The MG symptoms and the “morning light evening heavy” characters were first described as early as 1672 by British clinician Thomas Willis, whereas the cause of the disease remained a mystery until the 1960s. At that time, MG was described as the result of antibodies binding to the neuromuscular junction, most commonly against the acetylcholine receptor. However, up to date management of MG still remains a great challenge.

There are many reasons behind the need to develop an international consensus guideline for MG management. Firstly, early expert treatments can significantly improve the prognosis of a MG patient, as some physicians have not seen enough MG patients to be familiar with all its cardinal features. This is due to MG low incidence and heterogeneity. Secondly, uncontrolled clinical trials may be potentially biased, while the few successful randomized controlled trials (RCTs) cannot be generalized to assess the effectiveness and safety of multi-regimens, and to select the best treatment for each patient. Thus, it is critically important that experts share their knowledge and competence to improve the management of MG. Such expert-developed guidelines not only will help the clinician, but will also represent a unique resource for third-party payers such as insurance companies, governmental health organizations and institutional review boards.

METHODS

In October 2013, a Task Force of the Myasthenia Gravis Foundation of America (MGFA) assembled a panel of 15 internationally recognized MG experts, chaired by Donald Sanders of the Duke University and Gil Wolfe of State University of New York at Buffalo, and moderated by Pushpa Narayanaswami of Harvard University. The main goal of this panel was to develop treatment guidance statements based on formalized consensus. The guideline development employed the RAND/UCLA Appropriateness Method (RAM), which was established and refined by RAND and the University of California, Los Angeles (UCLA) in the 1980s.

The first meeting was held in February 2013 in Durham, North Carolina, to make decisions on cardinal definitions that were going to be instrumental for subsequent guidance treatment statements: goals of treatment, minimal manifestations, remission, ocular MG, impending and manifest myasthenic crises and refractory MG. Definitions without consensus were modified upon the panelists’ suggestions and shared with the panel for subsequent voting rounds.

The first draft of the MG guidance treatment statements was prepared by the two executive chairmen and the Harvard University service providers, based on recent publications and guidelines from the US, Denmark, Norway, Germany, Japan, Netherlands, United Kingdom and Europe. The following three assumptions were agreed a priori: (1) treatment costs and availability would not be considered; (2) clinical examination is performed by experienced physicians for the evaluation of neuromuscular disease; (3) the MGFA Clinical Classification refers to the state of the patient at the time of evaluation. Guidance statements were developed for the following seven topics: MG symptomatic and immunosuppressive (IS) treatment, IV immunoglobulin (IVIg) and plasma exchange (PLEX), impending and manifest myasthenic crisis, thymectomy, juvenile MG (JMG), muscle-specific tyrosine kinases (MuSK) antibody-positive MG and MG in pregnancy.

The consensus guidance statements were refined using a quantitative evaluation system of the RAM program. Initial and revised statements were voted and commented on anonymously at least three times during the process. The facilitator (Dr. Pushpa Narayanaswami) was the only person who made announcements about the statements and collected votes and feedbacks. Dr. Narayanaswami was not allowed to vote or participate in discussions and feedbacks to ensure the maximum objectivity of the process. The chairmen gathered the votes anonymously and revised the statements, seeking for ultimate consensus. A second face-to-face meeting and panel discussion was held in March 2014 after the first round of vote. The second and third round of votes were solicited after statements revision to reflect all panelists’ comments and experiences. After all three rounds, consensus was reached on all definitions and guidance statements. So far, this represents the first official international MG treatment guidance and as such it was published in its final form on June 29, 2016.
in the journal of Neurology.[1]

RAM was developed to combine the best available scientific evidence, even when the randomized controlled trials (the golden standard in evidence-based medicine) are not available or cannot provide enough detailed guidance for everyday clinical practice. The RAM is based on the collective judgment of experts with the common objective to release statements regarding the appropriateness of following a procedure, using the multiple-rounds Delphi polling sessions to assess the treatment rationality.[5]

**The procedure of RAM**

There were 2 major concepts during the RAM: appropriateness and agreement. The median rating at each round is the appropriateness score and the summary of the appropriateness scores for each recommendation is its agreement score. Appropriateness ratings are collected from each panelist to quantitatively assess the relative harm or benefit of a particular intervention. Each recommendation is rated on a 9 point scale: 1-3 are extremely inappropriate to inappropriate (i.e. risks > benefit); 4-6 are uncertain (i.e. risks ≈ benefit); 7-9 are appropriate to extremely appropriate (i.e. benefit > risks) [Table 1].

The detailed procedure of RAM is listed in the flow chart of Figure 1.

Disagreements and uncertainty ratings assist in determining “grey” areas for future research. Panel consensus is NOT forced. Rather, the degree of agreement is used to define the strength of the recommendations. Agreement: ≤ 3/13 panelists or ≤ 4/14 panelists rate the recommendation outside the 3 point region containing the median score of appropriateness. Disagreement: ≥ 4/13 panelists or ≥ 5/14 panelists rate the recommendation in the 1-3 region and the same number in the 7-9 region of appropriateness.

Therefore, if the appropriate scores on a particular recommendation have a median of 7-9, the recommendation is considered to be appropriate without disagreement, instead, if the appropriate scores have a median of 4-6, it could be either uncertain (if the actual scores are within 4-6) or disagreement (if some scores are in 7-9 region and others are in 1-3 region), and the recommendation must be revised for future consensus. Similarly, if the appropriate scores have a median of 1-3, the recommendation is considered to be inappropriate without disagreement.

It is also possible to reach an agreement about a recommendation being appropriate, inappropriate or even uncertain. Facilitator manages voting but does not vote. Cost and availability considerations are NOT used at this stage and all options are assumed to be affordable and freely available.

**RESULTS**

**The application of RAM**

RAM is useful for rare cases with low morbidity, lack of RCT clinical research, pooling available research evidence and expert experience to draw up a guideline,

---

**Table 1: RAND/UCLA appropriateness method rationality and its meaning of the 9-point score (adopted from[5])**

<table>
<thead>
<tr>
<th>Inappropriate</th>
<th>1</th>
<th>Extremely inappropriate (risk greatly exceed benefits)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>Moderately inappropriate</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Slightly inappropriate</td>
</tr>
<tr>
<td>Uncertain</td>
<td>4</td>
<td>May be inappropriate</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Uncertain/equivocal (benefit and risk about equal)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>May be appropriate (expected health benefits to an average patient exceed the expected health risks by a sufficiently wide margin to make intervention worthwhile and the intervention is superior to alternatives, including no intervention)</td>
</tr>
<tr>
<td>Appropriate</td>
<td>7</td>
<td>Slightly appropriate</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Moderately appropriate</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Extremely appropriate (benefit greatly exceed risks)</td>
</tr>
</tbody>
</table>

---

**Figure 1:** Flow chart of the RAND/UCLA rationality approach to develop a consensus (adopted from Dr. Sanders’ presentation with permission)
solicit the expert panel’s opinion and quantify the level of approval. However, if this does not meet the stated objectives, the recommendations of the consolidated expert group will be revised and the above steps will be repeated so that the views will be recognized to the maximum benefit.

The advantages of group decisions are obvious. A group is less likely than an individual to draw a wrong conclusion. If panelists are properly chosen, they can represent a wide range of knowledge and experience. Their interaction stimulates debate and consideration of many opinions that may challenge previously well-accepted ideas and stimulate new ones.

However, formal consensus has its own pitfalls: (1) only one person can speak at a time, limiting the number of ideas expressed and discussed; (2) a social pressure might induce to agree with the majority or a “powerful” voice in public; (3) the desire to reach agreement may override concerns about the accuracy of the result and may result in premature closure without consideration of all possible alternatives.

When uncertainty and differences of opinions exist, the RAM process of summarizing judgements helps to identify areas of agreement and establish areas of disagreement. The combination of face-to-face discussion at early stage and the solicitation of anonymous votes and comments handled by a non-voting facilitator are effective in maximizing the input of all experts’ knowledge and experience.

The case of developing MG international consensus guidance using RAM

More than two years passed between the initial appointment of the MGFA Task Force to develop treatment guidance for MG in October 2013 and the final acceptance of the publication of the international consensus guidance in July 2016. At the beginning, all definitions obtained consensus easily and all guidance statements were eventually agreed upon as being appropriate by the panel by the time of publication. However, not all topics took the same effort to reach consensus. Here are some examples of extreme cases during the RAM process.

Example 1 – easy consensus

The panelists easily reached consensus on statements about immunotherapy.

Round 1 statement: "If high steroid doses are needed chronically to achieve or maintain an adequate response, a steroid-sparing agent should be added, typically along with the steroid, to permit subsequent reduction of the steroid dose to the lowest necessary to maintain an adequate response."


Consensus had been achieved; however, based on panel input and discussion, the statements were modified and re-voted.

Round 2 statement: "A non-steroid IS agent should be used alone when steroids are contraindicated or refused. A non-steroid IS should be used initially in conjunction with steroids when the risk of steroid side effects is high based on medical co-morbidities. A non-steroid IS should be added to steroids when: (a) steroid side effects, deemed significant by the patient or the treating physician, develop; (b) response to an adequate trial of steroids is inadequate; (c) symptoms relapse upon steroid taper."


Final statement on the publication: "A non-steroid IS agent should be used alone when corticosteroids are contraindicated or refused. A nonsteroidal IS agent should be used initially in conjunction with corticosteroids when the risk of steroid side effects is high based on medical comorbidities. A nonsteroidal IS agent should be added to corticosteroids when: (a) steroid side effects, deemed significant by the patient or the treating physician, develop; (b) response to an adequate trial of corticosteroids is inadequate; or (c) the corticosteroid dose cannot be reduced due to symptom relapse."

Example 2 – difficult consensus

Considerable effort was needed to reach consensus on statements about thymectomy in childhood MG.

Round 1 statement: "In children and adolescents aged 5-10 years, thymectomy should be considered only after failure of symptomatic therapy and immunotherapy."

Round 1 votes: median 6, range 1-9, uncertain/equivocal.

Round 2 statements (modified based on discussion): "(A) in patients under 15 years of age, thymectomy should be considered in generalized MG after unsatisfactory response to AChEs and immunotherapy; (B) there is wide consensus that thymectomy is indicated in peri-pubertal and post-pubertal children with moderate to severe AChR-ab+ MG; (C) published reports also suggest that early thymectomy (within the first 12 months of onset of symptoms) is more effective
than delayed thymectomy; (D) for seronegative children, there is always a risk that some will have a CMS and not immune-mediated JMG; (E) evaluation at a centre specializing in childhood neuromuscular diseases should be considered before recommending thymectomy in young patients with seronegative MG.”

Round 2 votes: median 8, range 2-9, 4 panelists rated 2-4, and the rest in the 7-9 range. There was still a disagreement, no consensus.

Round 3 statements (modified based on discussion): “(A) the value of thymectomy in the treatment of pre-pubertal MG patients is unclear, but thymectomy should be considered in children with generalized AChRantibody* MG either if: the response to AChE inhibitor and immunosuppressive is unsatisfactory, or If there is a need/desire to avoid potential complications of immunosuppressive therapy; (B) for children diagnosed as seronegative GMG, the possibility of a congenital myasthenic syndrome or other neuromuscular condition should be entertained, and evaluation at a center specializing in neuromuscular diseases is of value prior to thymectomy.”

Round 3 votes: median 8, range 7-9, appropriate with consensus.

Final statement on the publication: “(A) the value of thymectomy in the treatment of pre-pubertal patients with MG is unclear, but thymectomy should be considered in children with generalized AChR antibody-positive MG. (a) if the response to pyridostigmine and IS therapy is unsatisfactory; or (b) in order to avoid potential complications of IS therapy. (B) For children diagnosed as seronegative GMG, the possibility of a congenital myasthenic syndrome or other neuromuscular condition should be entertained, and evaluation at a center specializing in neuromuscular diseases is of value prior to thymectomy.”

DISCUSSION

Preliminary definitions

Among the preliminary definitions compiled for the 2016 International Consensus Guidance for MG, for the first time two concepts are given clear definitions and provide highly valuable guidance to the clinical practice of treating MG patients.

The first concept is impending myasthenic crisis. It is defined as “Rapid clinical worsening of MG that, in the opinion of the treating physician, could lead to crisis in the short term (days to weeks).” In the past, crisis in MG is only referred to as manifest myasthenic crisis, defined as “MGFA Class V Worsening of myasthenic weakness requiring intubation or non-invasive ventilation to avoid intubation”. The concept of “impending myasthenic crisis” will raise the awareness of the physicians who can take proactive approach to intervene before crisis actually takes place.[12]

Another concept is refractory MG. It is defined as “PIS is unchanged or worse after corticosteroids and at least 2 other IS agents, used in adequate doses for an adequate duration, with persistent symptoms or side effects that limit functional, as defined patient and physician.” Refractory MG has been the focus of several discussions,[13,14] although without a specific definition until the 2016 Guidance. The definition of refractory MG could be furthered developed and improved, however the one currently approved provides a common denominator for MG specialists.

Guideline topics

The consensus guidance treatment statements were developed around the following seven major topics: symptomatic and IS treatment of MG, IVIg and PLEX, impending and manifest crisis, thymectomy in MG, juvenile MG, MG with MuSK antibodies and MG in pregnancy. The following four topics require further discussions.

Symptomatic and IS treatment of MG

The statement on the use pyridostigmine is straightforward and relatively easy to reach consensus. It is almost always the first choice in treating MG patients. However, when pyridostigmine is not readily available due to various social-economical reasons (for example in recent months in mainland China), physician may directly prescribe nonsteroidal IS agents.

We totally agree with the statements on the use of IS treatment, especially statement 5 on IS agent dosage and duration of treatment. It is highly desirable to prescribe a low dose of corticosteroids and dosage adjustments should not be made too frequently and abruptly (“no more frequently than every 3-6 months”) based on our decades of clinical experience in China, although there are very different views and approaches regarding dosage and duration of treatment among Asian physicians.[8,15,16]

IVIg and PLEX

Although the guidance was developed with a priori agreement of not considering treatment costs and availability, it is worth noting that IVIg and PLEX are not covered by the Chinese insurance system. Since they are both expensive procedures (about $4,400-$7,300/IVIg and $1,500/PLEX), their applications have been limited.
Statement 1 mentioned that PLEX and IVIg are used as short-term treatment in MG patients with life-threatening signs. Our clinical experiences suggest that PLEX and IVIg alone are not sufficient in manifest MG crisis cases.[17,18] There are many complications during MG crisis (such as pneumonia, pneumothorax, heart failure and renal failure etc.) when PLEX and IVIg are not appropriate. Instead, a comprehensive treatment plan should be designed for each individual MG patient undergoing crisis.

Statement 2 mentioned that the choice between PLEX and IVIg depends on conditions of individual patient such as sepsis and renal failure. In addition, our clinical experiences suggest that IVIg is safer than PLEX to patients with cardiovascular disorders.[17]

Thymectomy
We totally agree with the statements about prepubertal patients with generalized AchR antibody-positive MG and all patients with MG with thymoma. Our research has shown that thymectomy on juvenal generalized MG patients did not affect their growth.[19,20]

Statement 5 mentioned that less invasive thymectomy approaches such as endoscopic and robotic approaches appear to yield similar results to more aggressive approaches and show a good track record for safety in experienced center. However, based on our experiences, endoscopic and robotic approaches to thymectomy are much less desirable for thymoma than extended thymectomy.[21-23]

MG with MuSK antibodies
In Chinese or Asian population in general, MG patients with MuSK antibodies are rarely seen. However, due to the lack of universal testing kits, the actual percentage is yet to be determined. As far as we know up to date, the testing center for MG patients in the US does not provide testing kits for other countries. Over a dozen of Chinese MG patients were found to be double seropositive using testing kits made in China (a surprisingly much higher positive rate than using testing kits made in UK) while only two of them had thymoma. This was in marked contrast to experts’ experiences (personal communication with Dr. Donald Sanders). Testing kits made in Germany would not be available to Chinese patients until 2017. An internationally recognized standard testing kits for MuSK antibodies is yet to be developed.

In summary, the 2016 international consensus guidance is a major contribution to the treatment and management of MG. It provides an up-to-date expert consensus to guide clinicians throughout the world, especially to those who have limited experiences or practice in countries or regions that have limited resources to develop local treatment guidelines. More importantly, it is an extraordinary example of how the RAND/UCLA appropriate methods can help to develop a consensus guidance for treatment of rare diseases, bringing together the best of existing scientific results with the experience of specialists around the world.

Authors’ contributions
Conceived the manuscript: W.B. Liu, W. Fang
Wrote the first draft in Chinese: W.B. Liu
Translated the first draft into English: H. Ran
Developed the conceptual structure and revised the manuscript: W. Fang
Participated in the clinical work and approved the final draft: C.Y. Ou, L. Qiu, Z.D. Huang, Z.Q. Lin, Y.K. Li, X.X. Liu, H. Huang

Financial support and sponsorship
This paper is supported by grants from the China National Natural Sciences Foundation (30870850, 81071002, 81371386, 81620108010) and the Clinical study of 5010 plan of Sun Yat-sen University (2010003).

Conflicts of interest
There are no conflicts of interest.

Patient consent
There is no patient data involved.

Ethics approval
There is no ethics issue in this paper.

REFERENCES


