Elmer Press

Appropriateness of Treatments for Patients With Degenerative Ataxias: Recommendations by a Panel of Experts

Maria M. Trujillo-Martin[a, b, e], Pedro Serrano-Aguilar[b, c], Fernando Monton-Alvarez[d],
Juan Manuel Ramos-Goni[a, b]

Abstract

Background: The available scientific evidence regarding treatment of degenerative ataxias is scarce. Appropriateness of therapy options for degenerative ataxias were evaluated by using the RAND/UCLA Method.

Methods: After a systematic literature review, a list of clinical scenarios were developed to simulate situations most likely to arise in clinical practice. An 8-members expert panel rated, in a two rounds process, the appropriateness of each clinical scenario-treatment combination or indication. Analysis used the ratings to categorize each indication as appropriate, of uncertain appropriateness or inappropriate.

Results: Final rankings for the indications were as follow: 26/154 (18.3%) appropriate, 36/154 (25.4%) uncertain and 92/154 (56.3%) inappropriate. The agreement rate was 66.2%. For patients with Friedreich ataxia, physostigmine, 5-hydroxytryptophan and amantadine were rated inappropriate while *L*-carnitine was rated appropriate only for asymptomatic patients or for patients with gait ataxia. Panelists recommended idebenone therapy for Friedreich ataxia complicated by cardiomyopathy. Therapy with 4-aminopyridine was rated inappropriate for episodic ataxia type 1 but it was rated appropriate for type 2. In the treatment of other ataxias, such as autosomal dominant ataxias and the autosomal recessive ataxias not Friedreich, physostigmine, acetazolamide and *L*-carnitine were rated inappropriate while amantadine was inappropriate only in pa-tients without gait ataxia. All other combinations were considered uncertain.

Conclusions: Within the limits of expert opinion, these guidelines provide direction for some common clinical uncertainties in the treatment of degenerative ataxias.

Keywords: Rare diseases; Degenerative ataxia; Treatment effectiveness; Consensus; RAND method; Appropriateness

Introduction

Degenerative ataxias are a heterogeneous group of hereditary and non-hereditary rare diseases that cause progressive imbalance as a result of the degeneration of the cerebellar cortex and its various connecting pathways. In addition to the difficulties with balance, other clinical signs are the impairments of speech, gait, swallowing and coordination of the limbs [1, 2]. In some instances, vision loss and cognitive decline may occur as well. Overall, the prevalence of those disorders in Spain is 20.2 cases/100,000 population [3].

Besides the uncertainty about the origin of these rare diseases and limited availability of diagnostic and therapeutic technologies, scientific evidence about the benefits of treatments in use is lacking [1, 3]. Research on treatment for degenerative ataxias have been characterized by its shortage and overall low scientific quality due to small sample sizes, heterogeneity of recruited patients, lack of adequate measurement instruments and short follow up [4]. Though some new trials could offer relevant information on new therapeutic agents in the next future, clinical decisions need to be made currently and on a daily basis [5]. In Spain, these limitations seem to be contributing to the presence of variations in treatment prescription and access to different services for degenerative ataxia patients across geographical areas [6].

The aim of the present study was to combine the best available scientific evidence with a structured expert-opinion process to develop appropriateness therapeutic criteria to manage scientific uncertainty and standardize clinical decision-making for degenerative ataxia patients.

Table 1. An Example of a Page From the Matrix

Chapter 1

Ataxia Friedreich	Appropriateness of treatment				
	Idebenone	Physostigmine	5-HTP	Amantadine	L-Carnitine
A. Gait ataxia absent and understandable speech	123456789	123456789	123456789	123456789	123456789
B. Gait ataxia present					
1. Understandable speech	123456789	123456789	123456789	123456789	123456789
2. No understandable speech	123456789	123456789	123456789	123456789	123456789
C. Impossible gait					
1. Understandable speech	123456789	123456789	123456789	123456789	123456789
2. No understandable speech	123456789	123456789	123456789	123456789	123456789
D. Gait ataxia present					
1. Understandable speech					
a. Asymptomatic cardiomyopathy present					
Respiratory failure absent	123456789	123456789	123456789	123456789	123456789
Respiratory failure present	123456789	123456789	123456789	123456789	123456789
b. Symptomatic cardiomyopathy present					
Respiratory failure absent	123456789	123456789	123456789	123456789	123456789
Respiratory failure present	123456789	123456789	123456789	123456789	123456789

5-HTP: 5-hydroxytryptophan.

Table 2. Appropriateness Classification

Agreement level	Panel median		
	1 - 3	4 - 6	7 - 9
Agreement	Inappropriate	Uncertain	Appropriate
Indeterminate	Inappropriate	Uncertain	Appropriate
Disagreement	Uncertain	Uncertain	Uncertain

Materials and Methods

To reach a formal collective judgment of experts, a well-established method, the RAND/UCLA Appropriateness Method (RAM) was used [7]. This method, developed in the mid-1980s by the RAND Corporation and the School of Medicine of the University of California at Los Angeles (UCLA), combines the best available scientific evidence with the practical experience from experts in the field to yield a statement regarding the appropriateness or inappropriateness of medical and surgical procedures. It has been verified that appropriateness guidelines developed by means of RAM are reproducible [8], clinically consistent [9], and correlated with clinical outcomes [10]. The RAM has different stages: 1) systematic review of the available scientific literature on the procedure to be rated; 2) development of a list of clinical scenarios which categorise patients likely to be encountered in clinical practice for the procedure in question in terms of their specific symptoms and signs, medical history and test results; 3) selection and setting of an expert panel; and 4) rating by the experts the benefit-to-harm ratio of the procedure for each clinical scenario, following a two-round modified Delphi technique [7].

To evaluate and synthesize the scientific evidence on the effectiveness and safety of treatments for degenerative ataxia patients, we performed an extensive systematic review that focused on the different options available, i.e., pharmacological, rehabilitation and psychological support services [4]. Based on this literature review and consultation with experts in the field, researchers developed a list of clinical scenarios based on combinations of clinical factors (specific types of ataxia, severity of ataxia, systemic comorbidities and symptoms) which might be taken into account in deciding the appropriate treatment. In a matrix table each clinical scenario was paired with every treatment alternative selected from the systematic review (Table 1).

As we required an expert panel of 8 - 12 experts, we approached 15 clinicians from different specialties (neurologists, cardiologists, pediatric neurologists and rehabilitation physicians) and diverse geographic sites within Spain. The eventual panel was composed by 5 clinic neurologists, 2 research neurologists and 1 rehabilitation physician.

Three documents were produced to be mailed to each panelist for the first round of the consultation process: 1) the systematic review of the evidence for effectiveness and safety of treatment alternatives for degenerative ataxias, 2) the matrix containing the clinical scenario-treatment combinations and 3) a list of concise and explicit definitions of the terms used in the scenarios. Each participating panelist individually rated each indication on a nine points appropriateness scale, ranging from 1 to 9 on the basis of increasing appropriateness, according to the scientific evidence provided by the systematic review and his/her own experience. A treatment is considered appropriate for a specific clinical scenario when its expected benefits exceed the expected harms (risks) by a sufficiently wide margin and, conversely, is considered inappropriate when its expected harms exceed any likely benefit. A score of 5 means that benefits and harms were thought to be about equal or that the panelist was unable to make a judgement for the situation described. Financial cost implications must not be considered in making judgements. No interaction among the experts occurred before rendering opinions. The score sheets were then returned by mail to the project coordinator for data entry and analysis.

The eight panelists then met for a second round of rating and group discussion, led by a moderator experienced in the method. Each panelist was provided with his or her first round own ratings together with the frequency distribution of all the experts' responses. For each clinical scenario-treatment combination, the median score was given, interpreted as appropriate [7-9], uncertain [4-6] or inappropriate [1-3]. In addition, there was an assessment of agreement/dispersion of views using the following definitions: (1) Agreement: no more than two panelists rate the indication outside the 3-point region [1-3, 4-6, 7-9] containing the median, (2) Disagreement: at least two panelists rated the indication in the 1 - 3 region, and at least two rated it in the 7 - 9 region, and (3) Indeterminate: covers all other eventualities. The convened panel discussed indications where there was disagreement, made several revisions to the clinical scenarios and then individually re-rated the revised indications.

The collected assessments were aggregated to obtain a composite judgment. The final appropriateness classification was based on the median panel rating and level of agree-

Table 3. Final Variables Involved in the Scenarios Rated for Degenerative Ataxia Patients

Chapter 1: Patients with FA*	Chapter 2: Patient with EA†	Chapter 3: Other Ataxias‡
	Type of EA	
Level of gait ataxia	EA1	
Ataxia absent	EA2	
Ataxia present	Stage of EA	Level of gait ataxia
Impossible gait (wheelchair)	Critic	Ataxia absent
Level of speech ataxia	Intercritic	Ataxia present
Understandable speech	Comorbidities	Impossible gait (wheelchair)
No understandable speech Systemic	Dyskinesias	Level of speech ataxia
comorbidities	Neuromuscular hyperexcitability	Understandable speech
Asymptomatic cardiomyopathy	Epilepsy	No understandable speech
Symptomatic cardiomyopathy	Dizziness	
Respiratory failure	Generalized weakness	
	Migraine	

FA: Friedreich ataxia; EA: Episodic ataxia. * No. of scenarios: 21, No. of indications: 105; † No. of scenarios: 24, No. of indications: 24; ‡ No. of scenarios: 5, No. of indications: 25.

ment for each indication in the second round of rating. An appropriate indication had a median rating of 7 - 9 without disagreement among panel members, an inappropriate indication had a median rating of 1 - 3 without disagreement and, finally, uncertain indications were those for which there was either a median rating of 4 - 6 or disagreement among panel members (Table 2).

Statistical analysis

Analysis of the rounds was carried out in a non parametric way. The median and interquartile range were calculated for each panel score. The type of analysis was selected due to the nature of response variable. Pearson's chi-squared test was used to detect changes in participants' opinions between the first and the second rounds. Statistical analyses were performed using software package STATISTICA 8.0 (StatSoft Inc, Tulsa, Okla, USA).

Results

The first cycle round was performed on a total of 308 clinical scenarios grouped into three main chapters: 1) Friedreich Ataxia (FA), that addressed 5 therapy options: idebenone, physostigmine, 5-hydroxytryptophan (5-HTP), amantadine and L-carnitine; 2) Episodic Ataxia (EA), subdivided into EA type 1 (EA1) and EA type 2 (EA2), that addressed treatment with 4-aminopyridine; and 3) other ataxias (mainly including the autosomal dominant ataxias (SCAs) and the autosomal recessive ataxias not Friedreich), that addressed 5 therapy options: physostigmine, acetazolamide, 5-HTP, amantadine and L-carnitine, giving a total of 1,396 scenario-treatment combinations or indications. Overall agreement, indeterminacy and disagreement between experts about appropriateness concerning treatments were 31.1%, 68.1% and 0.8%, respectively. Based on the above definitions, a total of 74 (5.3%) of the 1396 indications were rated appropriate, 368 (26.4%) inappropriate and 954 (68.3%) uncertain.

For the second round, at a subsequent 1-day meeting, the group of panelists reviewed the first-round ratings, revised the structure of the clinical scenarios, modified some definitions of key terms and discussed reasons for the degree of agreement or disagreement in the ratings from the first round. After a detailed deliberation of clinical scenarios, there was wide agreement among panelist about the convenience of not taking some variables into consideration because they were not relevant for the selection of specific treatment among available alternatives and, therefore, there would be no significant differences in their ratings between scenarios. Such variables were: neuropathy, oscillopsia, spasticity, parkinsonism, depression and neuropsychiatric disorders. Besides, a set of scenarios was removed because they represented extremely rare scenarios in clinical practice and therefore the scientific evidence and clinical experience of experts on treatments were too scarce. Consequently, the list was shortened to 50 final scenarios, resulting in 154 indications to be evaluated. Finally, experts confidentially re-

Table 4. Changes in Experts' Responses Between Two Rounds (P = 0.0026)

		2nd Round			Total 1st Round
		Inappropriate	**Uncertain**	**Appropriate**	
1st Round	Inappropriate	62 (100%)	0	0	62 (38.0%)
	Uncertain	30 (45.5%)	36 (54.5%)	0	66 (43.7%)
	Appropriate	0	0	26 (100%)	26 (18.3%)
	Total 2nd Round	92 (56.3%)	36 (25.4%)	26 (18.3%)	154

Percentages in parentheses are calculated in relation of each row, except the total that is calculated in relation to the total number of scenario-treatment associations rated.

rated indications. Final considered clinical factors are shown in Table 3.

Of the 154 indications, 26 (18.3 %) were rated appropriate, 36 (25.4 %) were rated uncertain and 92 (56.3 %) were rated inappropriate. Agreement was reached on 66.2 % of the indications, disagreement on 0% and inconclusive agreement or disagreement on 33.8%. Comparing only the 154 indications rated in both rounds, statistically significant changes in panelists' responses between the first and the second rounds were obtained (Table 4). Clinical scenario-treatment pair associations evaluated as appropriate or inappropriate in the first round remained equally classified after the second round. However, among indications that were considered uncertain in the first round, 41.9% were classified as inappropriate after the second round, while the rest (58.1%) remained uncertain. Specific results for each chapter are described below.

Friedreich's ataxia

Treatment with idebenone was rated as appropriate for patients with cardiomyopathy complicated or not with respiratory failure. However, experts could not determine whether it is appropriate or inappropriate even when there have been no such symptoms. Physostigmine, 5-HTP and amantadine were rated as inappropriate for all FA scenarios. Finally, treatment with *L*-carnitine was scored as inappropriate only for asymptomatic patients or patients with gait ataxia with or without speech alteration, while the remaining scenarios constituted an area of uncertainty for the panelists.

Episodic ataxia

Other episodic ataxias different from EA1 or EA2 were not considered because of their rarity.

For all EA1 clinical scenarios, treatment with 4-aminopyridine was rated as inappropriate.

For EA2, experts assessed 4-aminopyridine as appropriate treatment in the critical phase as well as in the intercritical phase with gait ataxia. However, the presence of gait ataxia influenced the appropriateness rating for 4-aminopyridine treatment because scenarios with intercritical EA2 without gait ataxia were scored in the intermediate region and, therefore, these scenarios stayed uncertain.

SCAs and autosomal recessive ataxias not friedreich

The experts considered inappropriate the use of physostigmine, acetazolamide and *L*-carnitine in any of the clinical stages of these types of degenerative ataxias. Treatment with 5-HTP was rated as uncertain for all scenarios. Finally, amantadine therapy was assessed as inappropriate for asymptomatic patients and uncertain for all other scenarios.

Discussion

The concept of appropriate care is fundamental to the creation of an efficient and equitable health-care delivery system and, in the RAM, it refers to the relative weight of the benefits and harms of a medical or surgical intervention. The underlying principle supporting the RAM is that although strong scientific evidence about the clinical value of many medical procedures is lacking, health care professionals must nonetheless make decisions every day about when, how and to who apply them. A rigorous combination of the available scientific evidence and the collective opinion of experts is useful, and might be the only option, in guiding individual decisions in those cases where high-quality scientific information is incomplete, conflicting or non-existent [11]. However, the purpose of RAM is not to force consensus, but to find scenarios where there is agreement on the appropriateness or inappropriateness of a treatment and scenarios where conversely there is disagreement. Besides the role of RAM

for clinical guidelines development, there is also a growing recognition of this method to assess quality of care linked to reimbursement [12].

Degenerative ataxias, like most rare diseases, lack of sufficient and valid scientific evidence to inform clinical decision-making, contributing to variations in clinical practice and access to different services [4-6]. Despite this, the appropriateness criteria developed in RAM has not been used to guide clinical decision-making in the field of rare diseases, as we have verified after revision of two consecutive systematic review on the application of RAM [13-14]. This work represents an initial exploratory study and the results presented here might be used as a starting point to assist clinicians in their interactions with degenerative ataxia patients in order to select the most appropriate treatment, contributing to harmonize clinical recommendations.

This consensus development on clinical management of degenerative ataxia patients classified 18.3% of clinical scenario-treatment indications as appropriate and 56.3% of indications as inappropriate, with a 25.4% of uncertain indications. The rate of inappropriate recommendations in this study is high, confirming that whereas high quality scientific evidence are most often helpful in illuminating what clinicians should do in practice, expert opinion help to determine what they should not do and can delineate uncertain indications for which new trials are needed. Surprisingly the percentage of uncertain pairs, which could be interpreted as an expression of the lack of information and/or clinical experience, is not higher than the observed results published for other consensus panels on chronic and frequent diseases or common procedures [12, 15, 16]. In our case, since there was no disagreement among the panelists, these uncertain indications are due to median scores in the 4 - 6 range.

Panelists held that physostigmine, 5-HTP and amantadine were inappropriate in FA patients and *L*-carnitine was considered a treatment option only in asymptomatic patients or in patients with gait ataxia. Idebenone was rated appropriate in FA patients with gait ataxia and cardiomyopathy whether they had respiratory failure or not. The panel agreed that 4-aminopyridine has no role in the management of EA1 but it was rated appropriate for patients with EA2. In the treatment of other ataxias, physostigmine, acetazolamide and *L*-carnitine were rated inappropriate while amantadine was inappropriate only in patients without gait ataxia and without speech disturbance. All other combinations were considered uncertain.

The low frequency of these rare diseases together with their heterogeneous etiology, variable natural history, progression and comorbilities, as well as their different prognosis make difficult the rating of indications as appropriate or inappropriate. Overall, our panel recognized that the specific therapeutic options available for degenerative ataxia are still imperfect. Except for specific deficient ataxias (for example, vitamine E and Co Q10 deficiency), no treatment has sci-

entifically demonstrated to change the natural course of the disease. These ataxias were not included in the study. Therefore, the panel suggested that, currently, improvement can only be achieved by symptomatic treatment.

Our study has several limitations. First, no scenarios rated by the experts represent the complexities of individual patients. In order to evaluate a manageable number of clinical scenarios a wide variety of clinical entities were grouped under the chapter of "other ataxias". Besides, several variables, like time from diagnosis, quality of life and tolerability, were not explicitly addressed by this panel and some others that were included in the first round were removed from the final matrix because of the very low frequency of the resulting clinical scenarios and therefore the lack of clinical experience of experts regarding these specific scenarios. These variables could influence the choice of therapy in particular circumstances. Second, the development of scenarios is based on combinations of clinical factors that could affect patients' net benefits from the procedure of interest. However, this is a subjective process. A different research team might have developed a different list of scenarios.

Despite these limitations, we found that RAM is a useful, systematic and reproducible methodology for developing recommendations to assist clinicians in reducing variations in decision-making and optimize the utilization of health care resources, particularly in clinical conditions such as degenerative ataxia in which there is still little or low quality scientific evidence. The present study also provides important insights for investigators dealing with degenerative ataxias, highlighting the gaps between currently available data and the evidence needed for clinical decision-making. Future trials should be designed to address these needs. Since some of the patients´ self-perceived degenerative ataxia-related health problems have never been investigated so far [17], they should be considered to improve future research projects and adapt them to the patients' needs.

Acknowledgement

Neurobiology Ramon y Cajal, Madrid); Juan Jesus Vilchez, MD (Neurology Service, University Hospital La Fe, Valencia).

Competing Interests

Grant Support

References

1. Harding AE. Classification of the hereditary ataxias and paraplegias. Lancet. 1983;1(8334):1151-1155.
2. Klockgether T. Recent advances in degenerative ataxias. Curr Opin Neurol. 2000;13(4):451-455.
3. Polo JM, Calleja J, Combarros O, Berciano J. Hereditary ataxias and paraplegias in Cantabria, Spain. An epidemiological and clinical study. Brain. 1991;114 (Pt 2):855-866.
4. Trujillo-Martin MM, Serrano-Aguilar P, Monton-Alvarez F, Carrillo-Fumero R. Effectiveness and safety of treatments for degenerative ataxias: a systematic review. Mov Disord. 2009;24(8):1111-1124.
5. Boddaert N, Le Quan Sang KH, Rotig A, Leroy-Willig A, Gallet S, Brunelle F, Sidi D, et al. Selective iron chelation in Friedreich ataxia: biologic and clinical implications. Blood. 2007;110(1):401-408.
6. Lopez-Bastida J, Perestelo-Perez L, Monton-Alvarez F, Serrano-Aguilar P. Social economic costs and health-related quality of life in patients with degenerative cerebellar ataxia in Spain. Mov Disord. 2008;23(2):212-217.
7. Fitch K, Bernstein SJ, Aguilar MS, Burnand B, La Calle JR, Lazaro P, Van het Loo M, et al. The RAND/UCLA Appropriateness Method User's Manual. http://rand.org/publications/MR/MR1269/index.html (accessed 7 May 2012).
8. Shekelle PG, Kahan JP, Bernstein SJ, Leape LL, Kamberg CJ, Park RE. The reproducibility of a method to identify the overuse and underuse of medical procedures. N Engl J Med. 1998;338(26):1888-1895.
9. Hodgson DC, Brierley JD, Cernat G, Bondy S, Slaughter PM, Pinfold SP, Paszat LF. The consistency of panelists' appropriateness ratings: do experts produce clinically logical scores for rectal cancer treatment? Health Policy. 2005;71(1):57-65.
10. Hemingway H, Crook AM, Feder G, Banerjee S, Dawson JR, Magee P, Philpott S, et al. Underuse of coronary revascularization procedures in patients considered appropriate candidates for revascularization. N Engl J Med. 2001;344(9):645-654.
11. Naylor CD. What is appropriate care? N Engl J Med. 1998;338(26):1918-1920.
12. Furst DE, Halbert RJ, Bingham CO, 3rd, Fukudome S, Anderson L, Bonafede P, Bray V, et al. Evaluating the adequacy of disease control in patients with rheumatoid arthritis: a RAND appropriateness panel. Rheumatology (Oxford). 2008;47(2):194-199.
13. Van het Loo M, Kahan JP. The RAND appropriateness method: an annotated bibliography through 1999. RAND Europe Santa Monica: RAND Distribution Services; 1999. Available at: http://www.rand.org/pubs/rand_europe/RE99-010 (accessed 17 mayo 2012).
14. Gonzalez N, Quintana JM, Lacalle JR, Chic S, Maroto D. [Review of the utilization of the RAND appropriateness method in the biomedical literature (1999-2004)]. Gac Sanit. 2009;23(3):232-237.
15. Halbert RJ, Figlin RA, Atkins MB, Bernal M, Hutson TE, Uzzo RG, Bukowski RM, et al. Treatment of patients with metastatic renal cell cancer: a RAND Appropriateness Panel. Cancer. 2006;107(10):2375-2383.
16. Caprilli R, Angelucci E, Cocco A, Viscido A, Annese V, Ardizzone S, Biancone L, et al. Appropriateness of immunosuppressive drugs in inflammatory bowel diseases assessed by RAND method: Italian Group for IBD (IG-IBD) position statement. Dig Liver Dis. 2005;37(6):407-417.
17. Serrano-Aguilar P, Trujillo-Martin MM, Ramos-Goni JM, Mahtani-Chugani V, Perestelo-Perez L, Posada-de la Paz M. Patient involvement in health research: a contribution to a systematic review on the effectiveness of treatments for degenerative ataxias. Soc Sci Med. 2009;69(6):920-925.