*Volumen 13, número 1 de 2026*
https://doi.org/10.53853/encr.13.1.994

REVISTA COLOMBIANA de Endocrinología Diabetes & Metabolismo

Consensos

# Consensus-based recommendations for the medical management of moderate-to-severe thyroid eye disease using the RAND/UCLA Appropriateness Method

María Del S. Cabarcas-Solano[1,2], Alejandro Román-González[3], Marta L. Muñoz-Cardona[3,4], María G. Mejía-López[5,6], Natalia Aristizábal-Henao[4,7], Katherine Restrepo-Erazo[8], Jennifer Camargo González[9], Hernando Vargas-Uricoechea[10], Henry M. Arenas-Quintero[11], Alejandro A. Castellanos-Pinedo[12], Carlos E. Builes-Montaño[3,13]

[1]Universidad del Norte, Barranquilla, Colombia.
[2]Universidad San Martín, Barranquilla, Colombia.
[3]Universidad de Antioquia, Medellín, Colombia.
[4]Universidad Pontificia Bolivariana, Medellín, Colombia.
[5]Fundación Universitaria de Ciencias de la Salud, Bogotá, Colombia.
[6]Hospital San José, Bogotá, Colombia.
[7]Clínica Las Américas-AUNA, Medellín, Colombia.
[8]Universidad Santiago de Cali, Cali, Colombia.
[9]Hospital Universitario San Ignacio, Bogotá, Colombia.
[10]Universidad del Cauca, Popayán, Colombia.
[11]Universidad Tecnológica de Pereira, Pereira, Colombia.
[12]Hospital San Jerónimo, Montería, Colombia.
[13]Hospital Pablo Tobón Uribe, Medellín, Colombia.

**How to cite this article:** Cabarcas-Solano MDS, Román-González A, Muñoz-Cardona ML, Mejía-López MG, Aristizábal-Henao N, Restrepo-Erazo K, et al. Consensus-based recommendations for the medical management of moderate-to-severe thyroid eye disease using the RAND/UCLA Appropriateness Method. Rev Colomb Endocrinol Diabet Metab. 2026;13(1):e994. https://doi.org/10.53853/encr.13.1.994

**Submitted:** 14/September/2025 **Accepted:** 12/February/2026 **Published:** 1/April/2026

## Abstract

**Background:** Thyroid eye disease (TED) is the most common extrathyroidal manifestation of Graves' disease and a major cause of morbidity. Management of active, moderate-to-severe thyroid eye disease remains challenging due to clinical heterogeneity, a variable response to glucocorticoids, and an evolving evidence base for alternative therapies. Although international guidelines provide general recommendations, uncertainty persists regarding the role of biologics, immunosuppressants, and local interventions.

## Highlights

- Intravenous glucocorticoids were rated as necessary as a first-line therapy for active, moderate-to-severe thyroid eye disease.
- Teprotumumab and tocilizumab were consistently rated as

**Corresponding author:** Carlos E. Builes-Montaño, CL 78B #69 - 240, Hospital Pablo Tobón Uribe, Cordoba, Medellín, Colombia. E-mail: esteban.builes@udea.edu.co

E-ISSN: 2805-5853 | L-ISSN: 2389-9786

**Objective:** To assess the appropriateness and necessity of medical therapies for active, moderate-to-severe thyroid eye disease using the RAND/UCLA Appropriateness Method, integrating systematic evidence with expert consensus.

**Methods:** A systematic review of randomized trials, meta-analyses, systematic reviews, and observational studies was conducted following PRISMA standards. A multidisciplinary panel (n=10) rated 14 clinical scenarios across 238 intervention–scenario pairs in three rounds using a 9-point scale. Appropriateness was defined by median scores and the interpercentile range; necessity was evaluated for appropriate interventions according to RAND/UCLA criteria.

**Results:** Intravenous glucocorticoids were consistently rated as necessary as a first-line therapy. Teprotumumab and tocilizumab were considered appropriate in multiple scenarios, with teprotumumab frequently meeting necessity criteria. Mycophenolate was appropriate but not necessary, reflecting concerns about the robustness of the evidence and its modest efficacy. Rituximab was rated conservatively due to heterogeneous evidence and safety concerns, including dysthyroid optic neuropathy. Local therapies, such as periocular triamcinolone and botulinum toxin, were appropriate in several scenarios and necessary in selected cases.

**Conclusions:** Intravenous glucocorticoids remain the cornerstone of TED management. This consensus clarifies the role of emerging therapies and provides context-specific recommendations that complement existing guidelines.

**Keywords:** Thyroid eye disease, Graves' orbitopathy, Glucocorticoids, Teprotumumab, Tocilizumab, RAND/UCLA Appropriateness Method.

- appropriate, with teprotumumab often deemed necessary in refractory or severe scenarios.
- Mycophenolate was rated appropriate but not necessary, reflecting concerns over the quality of the evidence.
- Rituximab was appraised conservatively due to inconsistent efficacy and safety concerns, particularly the risk of dysthyroid optic neuropathy.
- Local therapies, including periocular triamcinolone and botulinum toxin, were judged appropriate in multiple scenarios and necessary in selected cases, supporting their role as adjunctive treatments.

# Recomendaciones basadas en consenso para el manejo médico de la oftalmopatía tiroidea moderada a severa mediante el método de adecuación RAND/UCLA

## Resumen

**Antecedentes:** la enfermedad ocular tiroidea (EOT) es la manifestación extratiroidea más frecuente de la enfermedad de Graves y una causa importante de morbilidad. Su manejo en fases activas moderadas a graves es complejo por la heterogeneidad clínica, la respuesta variable a los glucocorticoides y la evidencia aún limitada para terapias alternativas. Persisten incertidumbres sobre el papel de los biológicos, inmunosupresores e intervenciones locales.

**Objetivo:** evaluar la adecuación y necesidad de terapias médicas para la enfermedad ocular tiroidea activa moderada a grave mediante el método RAND/UCLA, integrando evidencia sistemática y consenso experto.

**Métodos:** se realizó una revisión sistemática de ensayos clínicos, metaanálisis, revisiones sistemáticas y estudios observacionales según PRISMA. Un panel multidisciplinario (n = 10) evaluó 14 escenarios clínicos y 238 pares intervención–escenario en tres rondas con escala de 9 puntos. La adecuación se definió por la mediana y el rango interpercentil; la necesidad se evaluó según criterios RAND/UCLA.

**Resultados:** los glucocorticoides intravenosos se clasificaron como una terapia de primera línea necesaria. Teprotumumab y tocilizumab fueron considerados adecuados en múltiples escenarios, con teprotumumab alcanzando frecuentemente criterios de necesidad. El micofenolato fue adecuado pero no necesario, reflejando limitaciones en la evidencia y una eficacia moderada. Rituximab recibió valoraciones conservadoras por evidencia heterogénea y preocupaciones de seguridad. Las terapias locales (triamcinolona periocular, toxina botulínica) fueron adecuadas y, en algunos casos, necesarias.

## Destacados

- Los glucocorticoides intravenosos fueron calificados como una terapia de primera línea necesaria para la enfermedad ocular tiroidea activa, moderada a grave.
- Teprotumumab y tocilizumab fueron consistentemente considerados apropiados, y teprotumumab alcanzó el nivel de necesidad en escenarios refractarios o graves.
- El micofenolato fue calificado como apropiado pero no necesario, debido a preocupaciones sobre la calidad de la evidencia.
- Rituximab fue evaluado de manera conservadora por la inconsistencia en su eficacia y preocupaciones de seguridad, especialmente el riesgo de neuropatía óptica distiroidea.

**Conclusiones:** los glucocorticoides intravenosos siguen siendo la base del tratamiento. Este consenso delimita el rol de terapias emergentes y aporta recomendaciones contextuales que complementan las guías.

**Palabras clave:** oftalmopatía tiroidea, orbitopatía de Graves, glucocorticoides, teprotumumab, tocilizumab, método RAND/UCLA de adecuación.

- Las terapias locales, incluyendo triamcinolona periocular y toxina botulínica, fueron consideradas apropiadas en múltiples escenarios y necesarias en casos seleccionados, apoyando su rol como tratamientos adyuvantes.

## Introduction

Thyroid eye disease (TED), also referred to as Graves' orbitopathy or thyroid-associated ophthalmopathy, is the most frequent extrathyroidal manifestation of Graves' disease (GD), affecting up to 40% of patients with GD over the course of their illness (1). TED is an autoimmune condition driven by shared antigenic targets in the thyroid and orbit, particularly the thyrotropin receptor (TSHR) and the insulin-like growth factor 1 receptor (IGF-1R), leading to orbital fibroblast activation, adipogenesis, and glycosaminoglycan deposition (2). This process results in the expansion of orbital tissues, causing a spectrum of ocular manifestations such as proptosis, eyelid retraction, diplopia, ocular surface exposure, and, in severe cases, optic neuropathy (3). Risk factors associated with TED include smoking, older age, high thyroid-stimulating immunoglobulin levels, and radioiodine therapy (1).

The clinical expression of TED varies in activity and severity, with approximately 5–10% of cases progressing to a moderate-to-severe form that significantly compromises quality of life and may threaten vision (1). Moderate-to-severe TED is typically defined by the presence of active inflammation (Clinical Activity Score ≥3) and sight-threatening features, including significant diplopia, restrictive myopathy, corneal exposure, or optic nerve compression. Management of these patients requires a tailored approach, with intravenous glucocorticoids recommended as the first-line therapy for active, moderate-to-severe disease (3). Nonetheless, a substantial proportion of patients exhibit a suboptimal response or intolerance to steroids, prompting the investigation and use of alternative treatments such as biologic therapies and targeted immunomodulators (2, 3).

In this context, decision-making in moderate-to-severe TED is complicated by variability in clinical presentation, heterogeneity in patient response, and limited comparative evidence across therapeutic options (4). In this setting, formal consensus methodologies offer a valuable approach to guide clinical practice when evidence alone is insufficient. The RAND/UCLA Appropriateness Method (RAM) is particularly suited for this purpose, as it combines a systematic literature review with expert clinical judgment to assess the appropriateness of interventions across a range of patient scenarios. Its structured, iterative process facilitates transparent recommendations while capturing the diversity of expert perspectives. Given the growing interest in optimizing individualized therapy in TED and the lack of universally accepted algorithms beyond guideline statements, RAM provides a rigorous yet flexible framework to support consensus-based clinical decision-making in this complex field (5, 6). The objective of this study was to apply the RAND/UCLA Appropriateness Method to evaluate and generate consensus-based recommendations on the medical management of patients with active, moderate-to-severe TED. While the study was conceived in the context of the Colombian health care system, with its unique challenges related to access, availability of biologics, and referral pathways, the methodological framework and findings may also be applicable and transferable to other health systems with comparable characteristics in Latin America and beyond.

## Methods

### Evidence synthesis process

A systematic review was conducted to synthesize the available evidence on medical treatments for active, moderate-to-severe TED, following PRISMA guidelines. Eligible studies included randomized controlled trials, meta-

analyses, systematic reviews, and observational studies. Risk of bias was assessed using the RoB 2, ROBINS-I, and AMSTAR 2 tools. The full review was made available to all panelists before rating and is published separately for in-depth appraisal.

## Expert panel

An expert panel of ten specialists (three ophthalmologists and eight endocrinologists) was convened. Members were identified through stakeholder analysis (7), prioritizing expertise, influence, and relevance to TED. All participants declared their conflicts of interest. A three-member coordinating committee oversaw the review, structured the scenarios, and managed the process.

## RAND/UCLA Appropriateness Method

A modified RAND/UCLA Appropriateness Method (RAM) was employed, encompassing the following phases:

**Phase 1. Preparation of the information package:** The project coordinating committee compiled an information package that included the results of the systematic review with references, a standardized set of clinical definitions, a list of 14 clinical scenarios, and detailed instructions for appropriateness rating. Definitions covered disease activity and severity, exophthalmos thresholds, response to corticosteroids, contraindications, and other relevant clinical parameters. Panelists independently reviewed this material and were allowed to provide feedback and suggest modifications to the scenarios, which were incorporated before voting.

**Phase 2. First-round ratings:** To assess the appropriateness of pharmacological interventions for moderate-to-severe TED, panelists used a 9-point Likert scale adapted from the RAND/UCLA Appropriateness Method, where 1 indicated that risks clearly outweigh benefits, 5 indicated that benefits and risks are approximately equivalent or that evidence is insufficient to draw conclusions, and 9 signified that benefits clearly outweigh risks. Each panelist independently reviewed the clinical scenarios and supporting evidence materials and rated each proposed intervention using an online evaluation form. They were instructed to apply their clinical judgment in an average setting, without considering cost implications.

**Phase 3. Panel discussion:** A face-to-face meeting was held, moderated by a member of the coordinating committee. All panelists provided informed consent for the recording of the discussion, which was transcribed using the AI tool, Read AI.

**Phase 4. Second-round ratings:** Following the discussion, panelists re-rated each scenario. The average rating from the first round was provided for reference.

**Phase 5. Necessity ratings:** In a third round, panelists evaluated the necessity of treatments deemed appropriate in the second round. The necessity of treatment was assessed using a 9-point Likert scale, where 1 indicated that the intervention was clearly not necessary and 9 stated that it was clearly necessary. A treatment was considered essential if omitting it would constitute improper care, there was a reasonable likelihood of patient benefit, and the expected benefit was clinically significant (5).

## Analysis

For each clinical scenario–intervention pair, individual panel ratings were aggregated to calculate the median and the interpercentile range (IPR), defined as the difference between the 70th and 30th percentiles of the distribution. To assess disagreement, the asymmetry index (AI) was calculated as the absolute difference between the midpoint of the IPR and the central point of the scale (5.0). The interpercentile range adjusted for symmetry (IPRAS) was then computed as:

$$IPRAS = 2.35 + (1.5 \times AI)$$

Disagreement was deemed present if the IPR > IPRAS.

Appropriateness classifications followed RAND/UCLA rules: appropriate: median 7–9, without disagreement; uncertain: median 4–6, or any median with disagreement; inappropriate: median 1–3, without disagreement.

In the third round, necessity ratings were analyzed for interventions classified as *appropriate*

in the second round. The same analytical procedure (median, IPR, AI, IPRAS) was applied. A treatment was classified as *necessary* if it met all of the following criteria: median rating 7–9; no disagreement (IPR ≤ IPRAS); and fulfilled RAM necessity criteria, defined as follows: omitting the intervention would constitute improper care, a reasonable likelihood of benefit exists, and the expected benefit outweighs potential harm by a clinically significant margin.

All analyses were performed using R (version 4.3.1), with custom scripts implementing the RAND/UCLA statistical approach (5).

## Results

### Evidence synthesis

A comprehensive systematic review of 58 studies informed the consensus process, including randomized controlled trials, meta-analyses, systematic reviews, and observational studies evaluating pharmacologic treatments for active, moderate-to-severe TED. Interventions were assessed using a semi-quantitative framework across four domains: proptosis, Clinical Activity Score (CAS), diplopia, and safety/tolerability.

Intravenous glucocorticoids (IVGCs) remain the most consistently supported first-line therapy. Compared with oral regimens, IVGCs demonstrated superior control of inflammatory activity, with better tolerability and fewer systemic adverse effects. Their strongest benefit was observed in reducing CAS, with moderate impact on proptosis and diplopia. No clear superiority was identified between weekly versus monthly regimens or across higher cumulative doses.

Biologic therapies have expanded options in glucocorticoid-refractory or intolerant patients. Teprotumumab showed the most robust and consistent efficacy across all domains, particularly in proptosis reduction, though access, cost, and adverse effects (e.g., hearing-related events) remain important considerations. Tocilizumab and rituximab demonstrated moderate efficacy, primarily in early active or steroid-resistant disease.

Among non-biologic immunosuppressants, mycophenolate mofetil (MMF) showed consistent benefits across inflammatory outcomes, while methotrexate and cyclosporine produced more variable results. Adjunctive local therapies (e.g., periocular triamcinolone, botulinum toxin) may provide targeted benefits in selected clinical contexts but are supported by limited evidence from smaller studies.

To facilitate clinical comparison, a semi-quantitative synthesis of efficacy and safety across the four domains is presented in table 1. This structured framework supports therapeutic prioritization, particularly in the absence of direct head-to-head comparative trials.

**Table 1.** Qualitative summary of the efficacy and safety of therapeutic options for moderate-to-severe thyroid eye disease

| Intervention | Proptosis response | CAS change | Diplopia response | Safety / tolerability | Supporting evidence |
|---|---|---|---|---|---|
| IV glucocorticoids | +++ | ++++ | ++ | +++ | RCTs and meta-analyses |
| Oral glucocorticoids | + | ++ | + | + | RCTs and meta-analyses |
| Rituximab | ++ | +++ | ++ | ++ | RCTs, meta-analyses, observational |
| Tocilizumab | ++ | +++ | ++ | ++ | RCT, systematic review, observational |

| | | | | | |
|---|---|---|---|---|---|
| Teprotumumab | ++++ | ++++ | +++ | ++ | RCTs and MAIC |
| Adalimumab | - | + | - | ++ | Case series |
| Mycophenolate Mofetil | ++ | +++ | ++ | +++ | Meta-analysis |
| Methotrexate | + | ++ | + | ++ | RCT + observational |
| Sirolimus | ++ | +++ | + | ++ | observational |
| Cyclophosphamide | + | ++ | + | + | Meta-analysis |
| Cyclosporine | + | ++ | ++ | + | RCTs |
| Intravenous immunoglobulin | + | ++ | + | +++ | Observational |
| Somatostatin Analogues | - | ++ | - | ++ | RCTs + observational |
| Botulinum toxin | + | + | + | +++ | Systematic review |
| Colchicine | - | ++ | - | +++ | RCT |
| Triamcinolone | + | +++ | ++ | ++ | Observational |
| Plasma filtration | - | ++ | - | ++ | RCT |

RCT: Randomized controlled trial; MAIC: Matching-adjusted indirect comparison.

**Note.** Efficacy ratings are based on reported outcomes in proptosis response, CAS reduction, and improvements in diplopia. Safety/tolerability ratings reflect adverse event profiles as reported in clinical trials or observational studies. Ratings are semi-quantitative (- to ++++) and derived from the overall body of supporting evidence.

**Source:** Authors.

## Consensus based on the RAND/UCLA Appropriateness Method

A total of 238 scenario–intervention combinations were evaluated across 14 clinical scenarios. After the second rating round, 49 combinations (20.6%) were classified as appropriate, 145 (60.9%) as inappropriate, and 44 (18.5%) as uncertain, of which 40 (16.8%) corresponded to intermediate scores and 4 (1.7%) were due to panel disagreement. The distribution of appropriateness ratings across categories is presented in table 2. The 14 clinical scenarios were grouped into four categories reflecting disease severity, treatment refractoriness, comorbidities, and risk of vision loss. The scenario–intervention combinations rated as appropriate and necessary within each category and scenario are summarized in supplementary table 1. Detailed RAND/UCLA second-round ratings for each intervention in all scenarios are provided in supplementary table 2.

**Table 2.** Interventions rated as "appropriate" by clinical category and scenario

| Category | Scenario | Interventions rated appropriate |
|---|---|---|
| **1. Moderate TED without functional urgency or refractoriness** | 1 | Intravenous glucocorticoids; Teprotumumab. |
| | 2 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Mycophenolate mofetil; Periocular triamcinolone. |
| | 8 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Periocular triamcinolone; Botulinum toxin. |
| | 11 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Periocular triamcinolone. |
| **2. Moderate TED with refractoriness, comorbidities, or risk factors** | 4 | Teprotumumab; Tocilizumab; Mycophenolate mofetil. |
| | 5 | Teprotumumab; Tocilizumab; Mycophenolate mofetil. |
| | 7 | Teprotumumab; Tocilizumab; Mycophenolate mofetil; Periocular triamcinolone. |
| | 9 | Teprotumumab; Tocilizumab; Mycophenolate mofetil. |
| | 12 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab;Mycophenolate mofetil; Periocular triamcinolone; Botulinum toxin. |
| **3. Severe TED with acute visual compromise or optic neuropathy** | 3 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Periocular triamcinolone. |
| | 6 | Teprotumumab; Tocilizumab; Mycophenolate mofetil. |
| | 10 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Mycophenolate mofetil. |
| | 14 | Intravenous glucocorticoids; Teprotumumab. |
| **4. Severe TED with major corneal complications** | 13 | None. |

**Source:** Authors.

Of the 49 combinations rated appropriate, 33 (67.3%) were also classified as necessary. The distribution of essential ratings across categories is shown in table 3.

**Table 3.** Interventions rated as "necessary" by clinical category and scenario

| Category | Scenario | Interventions rated necessary |
|---|---|---|
| **1. Moderate TED without functional urgency or refractoriness** | 1 | Intravenous glucocorticoids; Teprotumumab. |
| | 2 | Intravenous glucocorticoids; Teprotumumab; Periocular triamcinolone. |
| | 8 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Periocular triamcinolone. |
| | 11 | Intravenous glucocorticoids; Teprotumumab; Periocular triamcinolone. |
| **2. Moderate TED with refractoriness, comorbidities, or risk factors** | 4 | Teprotumumab; Tocilizumab. |
| | 5 | Teprotumumab; Tocilizumab. |
| | 7 | Teprotumumab; Tocilizumab; Periocular triamcinolone. |
| | 9 | Teprotumumab; Tocilizumab. |
| | 12 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab; Botulinum toxin. |
| **3. Severe TED with acute visual compromise or optic neuropathy** | 3 | Intravenous glucocorticoids; Teprotumumab; Tocilizumab. |
| | 6 | Teprotumumab; Tocilizumab. |
| | 10 | Intravenous glucocorticoids; Teprotumumab. |
| | 14 | Intravenous glucocorticoids. |
| **4. Severe TED with major corneal complications** | 13 | None. |

**Source:** Authors.

## Patterns of uncertainty and disagreement

In the second rating round, a total of 43 scenario–intervention combinations were classified as uncertain. The majority (39/43; 90.7%) corresponded to intermediate median scores (4–6) without disagreement, indicating primarily insufficient or inconclusive evidence rather than divergent clinical positions among panelists.

In contrast, four interventions (9.3%) were classified as uncertain due to disagreement, as their interpercentile ranges exceeded the IPRAS threshold. These cases occurred across different clinical categories:

- In Category 1 (moderate TED without functional urgency or refractoriness), mycophenolate mofetil in scenario 11 (patients with moderate TED, with cosmetic impact but no functional urgency) showed disagreement.
- In Category 2 (moderate TED with refractoriness or comorbidities), periocular triamcinolone in scenario 5 (patients with partial response to IV glucocorticoids and uncontrolled diabetes mellitus) also showed disagreement.
- In Category 3 (severe TED with acute optic neuropathy or visual risk), both mycophenolate mofetil and teprotumumab in scenario 14 (patients with globe luxation and acute optic neuropathy) were classified as uncertain due to disagreement.

## Analysis by categories

### *Category 1 – Moderate TED without functional urgency or refractoriness*

Across scenarios, IVGCs and teprotumumab were consistently rated necessary, reflecting strong alignment between trial-level evidence and panel consensus. Teprotumumab demonstrated superior efficacy across proptosis, CAS, and diplopia, whereas IVGCs retained a central role in inflammatory control. Tocilizumab, MMF, and periocular triamcinolone were considered appropriate in selected contexts, particularly when the inflammatory burden predominated without visual compromise. Purely cosmetic presentations lacked strong evidence for systemic escalation, though panel ratings suggested variability in expert perception of disease impact.

### *Category 2 – Moderate TED with refractoriness, comorbidities, or risk factors*

In steroid-refractory or contraindicated patients, biologics predominated. Teprotumumab and tocilizumab were consistently rated necessary across scenarios, supported by randomized data in resistant disease. MMF and periocular triamcinolone were appropriate in selected contexts, particularly where systemic steroids posed safety concerns (e.g., uncontrolled diabetes). These findings underscore the importance of individualized therapy in moderate but complex presentations.

### *Category 3 – Severe TED with acute visual involvement or optic neuropathy*

In vision-threatening disease, IVGCs remained foundational when not contraindicated. Teprotumumab and tocilizumab were rated necessary in multiple scenarios, particularly in refractory cases. MMF was considered appropriate but not necessary, reflecting its slower onset of action and more limited evidence in acute settings. Scenarios involving extreme emergencies (e.g., globe luxation with neuropathy) showed the highest disagreement, highlighting limited high-quality evidence in these rare but critical contexts.

### *Category 4 – Severe TED with major corneal complications*

No systemic pharmacologic therapy was rated appropriate or necessary in corneal perforation. Management relies on urgent surgical and local interventions. Systemic immunotherapy may be considered after structural stabilization.

## Consensus-based treatment algorithm

This algorithm integrates three sources: (i) formal case definitions and clinical scenarios used to elicit ratings, (ii) RAND/UCLA consensus classifications (appropriate vs. necessary), and (iii) the systematic review summarizing efficacy and safety across drug classes. Active thyroid eye disease (TED) was defined as CAS ≥3/7; moderate disease encompassed typical inflammatory signs and functional impact, whereas severe disease required the presence of optic neuropathy or major corneal complications (e.g., perforation). Refractoriness was operationalized as no or partial response after a standardized 12-week course of intravenous glucocorticoids (IVGCs), and absolute contraindications to systemic glucocorticoids were specified a priori. These criteria, together with scenario-level modifiers (e.g., smoking, uncontrolled diabetes, cosmetic burden, corneal exposure), were used to structure the decision nodes.

Among moderate cases without refractoriness or critical comorbidities, IVGCs and teprotumumab were consistently rated necessary, aligning with

the evidence base supporting IVGCs as first-line therapy and the superior efficacy of teprotumumab across proptosis, diplopia, and CAS outcomes. Tocilizumab, mycophenolate mofetil (MMF), and periocular triamcinolone were considered appropriate in selected contexts, reflecting their safety-efficacy balance in non-urgent presentations (figure 1).

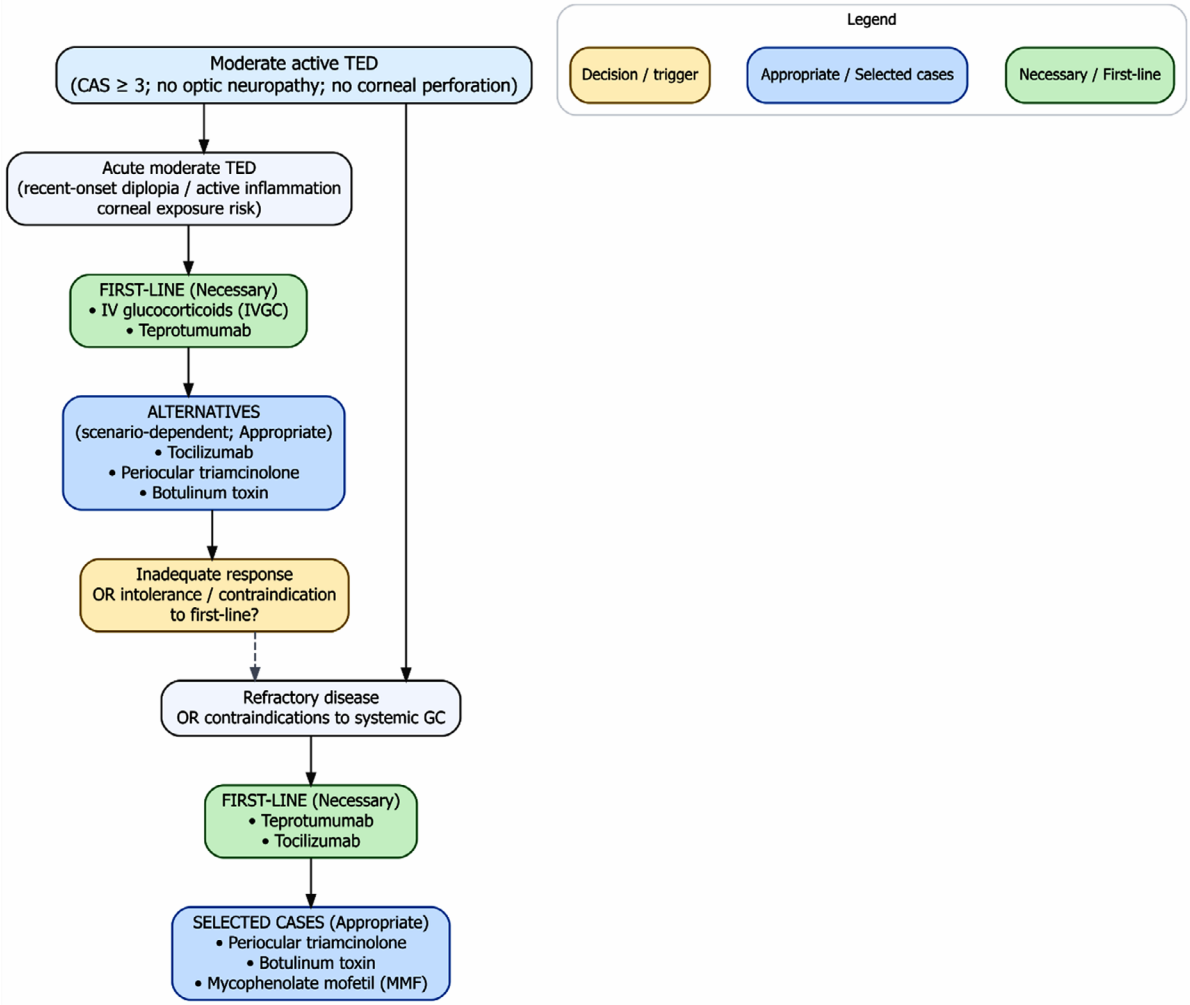


**Figure 1.** Consensus-based treatment algorithm for moderate active thyroid eye disease

**Note.** This figure presents the detailed consensus-derived treatment algorithm for moderate active thyroid eye disease without optic neuropathy or corneal perforation. Therapeutic prioritization is based on RAND/UCLA appropriateness and necessity ratings across clinical scenarios, incorporating refractoriness to IVGC, contraindications to systemic glucocorticoids, inflammatory activity, and risk modifiers. Green boxes denote treatments classified as necessary (first-line). Blue boxes indicate appropriate interventions in selected cases. Dashed connectors represent escalation pathways following inadequate response or intolerance.

**Source:** Authors.

In moderate disease with refractoriness, contraindications, or risk-modifying comorbidities, biologics predominated. Teprotumumab and tocilizumab were rated necessary across scenarios, supported by randomized evidence in steroid-refractory contexts, while MMF and periocular triamcinolone were appropriate in selected profiles (e.g., uncontrolled diabetes favoring local steroid delivery and corneal exposure allowing broader therapeutic flexibility, including the use of botulinum toxin when deemed necessary).

For severe disease with acute optic neuropathy, IVGCs (when not contraindicated), teprotumumab, and tocilizumab were consistently rated necessary, in line with the standard-of-care role of IVGCs and trial-level data supporting biologics in refractory cases. MMF was deemed appropriate but not necessary, reflecting weaker evidence and a slower onset. Extreme emergency scenarios, such as globe luxation with acute neuropathy, showed the most extraordinary polarization, with uncertain-by-disagreement ratings recorded for MMF and teprotumumab in the global uncertainty set, reflecting the paucity of robust data in extreme emergencies (figure 2).

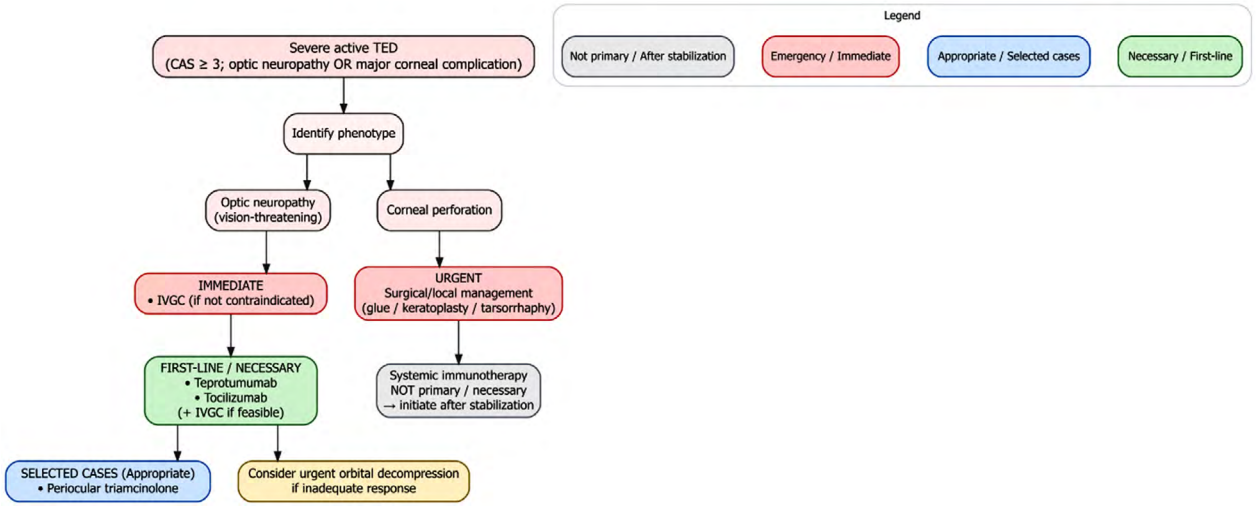


**Figure 2.** Consensus-based treatment algorithm for severe active thyroid eye disease

**Note.** This figure outlines the consensus-based management algorithm for severe active thyroid eye disease, defined by the presence of optic neuropathy or major corneal complications. Immediate interventions for vision-threatening disease are highlighted, followed by escalation strategies for refractory or contraindicated cases. Red boxes denote emergency or immediate management steps. Green boxes indicate therapies rated as necessary. Blue boxes represent appropriate treatments in selected scenarios. Gray boxes identify interventions not considered primary therapy in the acute setting.

**Source:** Authors.

Finally, in cases of severe corneal perforation, no systemic pharmacologic intervention was rated appropriate or necessary, indicating surgical or local management as the preferred approach. After the perforation is controlled, prompt systemic therapy should be initiated.

To facilitate bedside clinical application of these consensus recommendations, a simplified clinical decision pathway is presented in figure 3. This practical algorithm integrates disease severity, refractoriness, and urgency into a streamlined framework intended for rapid clinical use, while detailed scenario-based consensus algorithms are provided in figures 1 and 2.

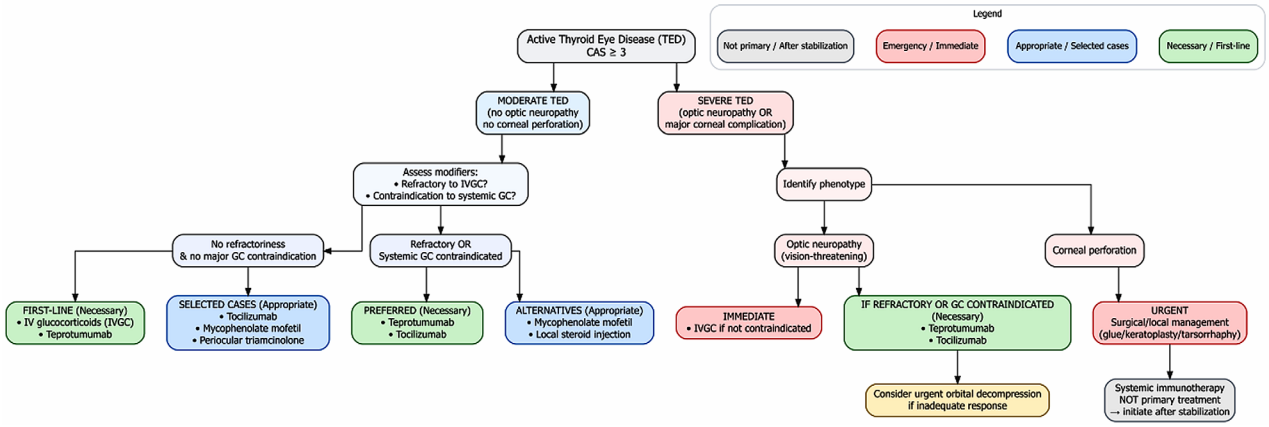


**Figure 3.** Practical clinical decision pathway for active moderate-to-severe thyroid eye disease

**Note.** This simplified clinical algorithm summarizes consensus-based therapeutic priorities for active thyroid eye disease (TED) (Clinical Activity Score ≥3). The pathway integrates disease severity (moderate vs. severe), refractoriness to intravenous glucocorticoids (IVGC), and urgency of visual compromise into a streamlined framework intended for rapid bedside application. Green boxes indicate therapies rated as necessary or first-line according to the RAND/UCLA consensus process. Blue boxes represent appropriate treatments in selected clinical contexts. Red boxes denote emergency or immediate interventions.

Gray boxes indicate therapies that are not considered first-line treatments in the specified scenario.

For detailed scenario-level consensus algorithms and full therapeutic stratification, see figures 1 and 2.

**Source:** Authors.

## Discussion

Our consensus study using the RAND/UCLA Appropriateness Method provides a structured appraisal of medical therapies for active, moderate-to-severe TED in the context of the Colombian health care system. The systematic review confirmed intravenous glucocorticoids (IVGCs) as the cornerstone of treatment while also highlighting the selective role of biologics such as teprotumumab and tocilizumab, and the consistent benefit of mycophenolate as an adjunct or alternative immunosuppressant. Panel deliberations further refined these findings, distinguishing between scenarios where therapies were deemed appropriate, uncertain, or inappropriate, thereby offering practical guidance for clinical decision-making across heterogeneous clinical presentations. Notably, ratings were based exclusively on clinical efficacy and safety, as panelists were explicitly instructed not to consider costs or availability.

When compared with both the EUGOGO 2021 guidelines (4) and the ATA/ETA 2022 consensus statement (8), our findings align in reaffirming intravenous glucocorticoids as the cornerstone of therapy for active moderate-to-severe TED and in supporting the selective use of biologics such as teprotumumab and tocilizumab. A key point of divergence relates to mycophenolate mofetil (MMF). EUGOGO recommends MMF, in combination with IVGCs, as a preferred first-line regimen, while ATA/ETA considers it an acceptable therapy, particularly in combination with glucocorticoids. In contrast, our panel did not classify MMF as necessary. This more cautious stance reflected two main concerns: the retraction of pivotal publications, which undermined confidence in the robustness of the evidence (8), and the perception among several panelists that the therapeutic effect of MMF appeared less consistent when compared with the more consistent benefits demonstrated with teprotumumab or IVGC monotherapy. This divergence illustrates how evolving evidence and expert interpretation can shift the weighting of the same therapeutic options, even within consensus frameworks.

Another relevant divergence concerns the use of rituximab. Both the EUGOGO 2021 guidelines and the ATA/ETA 2022 consensus statement acknowledge rituximab as an acceptable option in selected cases of active, moderate-to-severe TED, particularly in patients resistant to glucocorticoids. Our panel, however, expressed greater reservations. While recognizing that some randomized trials and observational studies suggest benefit, the group discussed at length the potential risk of dysthyroid optic neuropathy (DON) associated with rituximab therapy (9, 10), a complication reported in clinical practice and highlighted by panelists as a serious safety concern. This perception, coupled with the inconsistency of trial results, led to a more conservative classification of rituximab in our consensus, underscoring the importance of weighing not only efficacy signals but also the potential for rare but sight-threatening adverse events in therapeutic decision-making.

The emergence of biologic therapies is reshaping the therapeutic landscape of TED, with particular attention to tocilizumab and teprotumumab. Tocilizumab has shown consistent benefits in patients with glucocorticoid-resistant disease, including a randomized trial demonstrating significant reductions in CAS (11). Yet, it remains without formal regulatory approval for TED in Colombia. Teprotumumab, on the other hand, has been approved in several countries, including Colombia, as the first targeted therapy specifically indicated for TED and has demonstrated robust efficacy across all clinical domains (12–16). Although not widely available, its broader incorporation into clinical practice would represent a paradigm shift, especially for patients with rapidly progressive disease or an inadequate response to glucocorticoids. Our consensus classified both agents as appropriate in defined scenarios, underscoring their potential role in advancing individualized care and highlighting the need for regulatory pathways and access strategies to ensure their eventual integration into clinical practice.

Beyond systemic interventions, our consensus also highlighted the value of local therapies, which remain underrepresented in international guidelines. Both periocular triamcinolone and botulinum toxin were classified as appropriate in several scenarios and even necessary in some contexts, underscoring their perceived utility as adjuncts. Evidence from the systematic review further supports these findings. A recent meta-analysis of 30 studies (n=299 patients treated for eyelid retraction; n=205 for TED-associated strabismus) reported an 84% pooled success rate of BTX (Botulinum toxin type A) for eyelid retraction, with a mean MRD (margin reflex distance) reduction of 2.4 mm and mainly transient adverse events such as ptosis (13%) and diplopia (2%) (17). In contrast, outcomes in strabismus were less consistent, with only 24% achieving resolution of diplopia and the majority eventually requiring surgery.

For periocular triamcinolone, the prospective study by Bordaberry *et al.* (18) evaluated repeated peribulbar injections in patients with active, moderate-to-severe TED, including those with optic neuropathy. Significant reductions in CAS and improvement in visual function were observed, with two-thirds of patients experiencing complete resolution of optic neuropathy and minimal adverse events. Nevertheless, the non-randomized design, lack of controls, and potential for selection and detection bias limit confidence in these results. Despite these limitations, both the systematic review and the consensus panel converged on the view that periocular triamcinolone represents a valuable local option, particularly when systemic steroids are contraindicated or insufficient.

Collectively, these findings position local therapies as pragmatic, well-tolerated adjuncts for selected patients with TED. Their inclusion among interventions rated appropriate or necessary by our panel emphasizes the importance of broadening treatment algorithms beyond systemic agents and highlights the need for controlled studies to define their optimal role.

A key strength of this study is the use of the RAND/UCLA Appropriateness Method, which integrates the best available evidence with structured expert judgment (6, 19). This approach goes beyond narrative expert consensus by applying iterative rating rounds, moderated discussion, and quantitative measures of agreement and disagreement (IPR, IPRAS, AI). As such, it provides a transparent and reproducible framework that captures both convergence and divergence of expert opinion, allowing for nuanced recommendations that reflect the complexity of TED management.

Nonetheless, certain limitations should be acknowledged. The panel size was relatively small, and although composed of experts in endocrinology and ophthalmology, broader representation might have yielded additional perspectives. The study was conceived in the Colombian health care setting, which may affect generalizability; however, the methodology ensures relevance to similar contexts. Notably, costs and availability were intentionally excluded from consideration, consistent with RAND/UCLA principles, which may limit direct applicability to policy-making. Finally, for several therapies, especially local interventions and some immunomodulators, the evidence base remains limited and heterogeneous, thereby constraining the strength of the recommendations (20).

These findings have important implications for both clinical practice and research. They provide structured, scenario-based recommendations that can inform local and regional guidelines in Latin America and other comparable health systems. They also emphasize the role of emerging biologics, the more cautious appraisal of older immunosuppressants, and the potential utility of local therapies such as periocular triamcinolone and botulinum toxin.

Future research should prioritize randomized controlled trials of these interventions, as well as long-term comparative effectiveness studies and consensus updates that incorporate evolving therapies and regulatory approvals, ensuring that therapeutic algorithms for TED remain both evidence-based and adaptable.

## Conclusions

This RAND/UCLA consensus study provides a structured evaluation of medical therapies for active, moderate-to-severe thyroid eye disease. The panel confirmed intravenous glucocorticoids as the necessary first-line therapy while highlighting the emerging role of biologics such as teprotumumab and tocilizumab and the selective use of local therapies including periocular triamcinolone and botulinum toxin. Mycophenolate was judged appropriate but not necessary, reflecting concerns about the reliability of its evidence base. Rituximab was appraised conservatively due to inconsistent trial results and safety concerns, particularly the risk of dysthyroid optic neuropathy.

By integrating systematic evidence with structured expert judgment, this study offers scenario-based recommendations that complement and refine international guidelines. These findings may serve as a reference for clinicians in Colombia and other comparable health systems, while also emphasizing the need for ongoing research to clarify the role of immunomodulators, biologics, and local therapies in individualized treatment algorithms.

## Authors' contributions

Carlos E. Builes-Montaño: Conceptualization, methodology, project administration, and supervision; María G. Mejía-López: Methodology, project administration; Alejandro Román-González: Methodology, project administration, final approval; María Del S. Cabarcas-Solano: Writing, review, and editing, final approval; Marta L. Muñoz-Cardona: Writing, review, and editing, final approval; Natalia Aristizábal-Henao: Writing, review, and editing, final approval; Katherine Restrepo-Erazo: Writing, review, and editing, final approval; Jennifer Camargo González: Writing, review, and editing, final approval; Hernando Vargas-Uricoechea: Writing, review, and editing, final approval; Henry M. Arenas-Quintero: Writing, review, and editing, final approval; Alejandro A. Castellanos-Pinedo: Writing, review, and editing, final approval.

## Ethical implications

This work was a systematic review and consensus statement; therefore, no specific ethical considerations were identified. The study did not directly involve patients, human participants, or animals and was based exclusively on a review of the scientific literature and contributions from multidisciplinary experts.

## Funding

Festina Lente support was funded by the Asociación Colombiana de Endocrinología, Diabetes y Metabolismo. This work was funded by the Asociación Colombiana de Endocrinología, Diabetes

y Metabolismo, through an unrestricted educational grant from Amgen.

## Conflicts of interest

The authors declare the following conflicts of interest: Natalia Aristizábal-Henao has received academic consulting fees from Amgen. Katherine Restrepo-Erazo has participated in conferences and advisory boards for AstraZeneca, Boehringer Ingelheim, Novo Nordisk, Merck, Sanofi, Recordati, Eli Lilly, Bayer, and Abbott. Hernando Vargas-Uricoechea has received speaking honoraria from Sanofi, Abbott, and Boehringer Ingelheim. Carlos E. Builes-Montaño has received consulting and speaking fees from Sanofi, Novo Nordisk, Novartis, Recordati Rare Diseases, Valentech, Janssen, Abbott, and Boehringer Ingelheim, and is a shareholder of Festina Lente. Marta Lucía Muñoz-Cardona has received travel support from Amgen. María del Socorro Cabarcas-Solano, Maria G. Mejía-López, Alejandro Román-González, Jennifer Camargo, and Alejandro Castellanos-Pinedo report no conflicts of interest.

## Data statement

No data was collected in the development of this manuscript.

## Use of artificial intelligence (AI)

The authors acknowledge the use of Grammarly, an artificial intelligence–based tool, exclusively for language editing purposes. This tool was not used for generating ideas, analyzing data, or creating scientific content. All intellectual contributions, including study design, data interpretation, conclusions, and final approval of the manuscript, remain the sole responsibility of the authors.

## Acknowledgements

Festina Lente provided methodological support and was responsible for medical writing and editorial assistance. The authors assume full responsibility for the content and conclusions expressed in this manuscript.

## References

[1] Chin YH, Ng CH, Lee MH, Koh JWH, Kiew J, Yang SP, *et al.* Prevalence of thyroid eye disease in Graves' disease: A meta-analysis and systematic review. Clin Endocrin. 2020;93(4):363-374. https://doi.org/10.1111/cen.14296

[2] Bartalena L, Piantanida E, Gallo D, Lai A, Tanda ML. Epidemiology, natural history, risk factors, and prevention of Graves' orbitopathy. Front Endocrinol (Lausanne). 2020;11(2020):615993. https://doi.org/10.3389/fendo.2020.615993

[3] Wiersinga WM, Eckstein AK, Žarković M. Thyroid eye disease (Graves' orbitopathy): Clinical presentation, epidemiology, pathogenesis, and management. Lancet Diabetes Endocrinol. 2025;13(7):600-614. https://doi.org/10.1016/S2213-8587(25)00066-X

[4] Bartalena L, Kahaly GJ, Baldeschi L, Dayan CM, Eckstein A, Marcocci C, *et al.* The 2021 European Group on Graves' orbitopathy (EUGOGO) clinical practice guidelines for the medical management of Graves' orbitopathy. Eur J Endocrinol. 2021;185(4):G43-G67. https://doi.org/10.1530/eje-21-0479

[5] Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, *et al.* RAND/UCLA appropriateness method user's manual. RAND corporation Santa Monica, CA; 2000.

[6] Sparks JB, Klamerus ML, Caverly TJ, Skurla SE, Hofer TP, Kerr EA, *et al.* Planning and reporting effective web-based RAND/UCLA Appropriateness Method panels: Literature review and preliminary recommendations. J Med Internet Res. 2022;24(8):e33898. https://doi.org/10.2196/33898

[7] Bourne L. Making projects work: Effective stakeholder and communication management. CRC press; 2015. https://doi.org/10.1201/b18100

[8] Burch HB, Perros P, Bednarczuk T, Cooper DS, Dolman PJ, Leung AM, *et al.* Management of thyroid eye disease: A consensus

statement by the American Thyroid Association and the European Thyroid Association. Thyroid. 2022;32(12):1439-1470. https://doi.org/10.1089/thy.2022.0251

[9] Zhang B, Li Y, Xu W, Peng B, Yuan G. Use of rituximab after orbital decompression surgery in two Grave's ophthalmopathy patients progressing to optic neuropathy. Front Endocrinol (Lausanne). 2020;11(2020):583565. https://doi.org/10.3389/fendo.2020.583565

[10] Pelewicz-Sowa M, Miśkiewicz P. Dysthyroid optic neuropathy: Emerging treatment strategies. J Endocrinol Invest. 2023;46(7):1305-1316. https://doi.org/10.1007/s40618-023-02036-0

[11] Perez-Moreiras JV, Gomez-Reino JJ, Maneiro JR, Perez-Pampin E, Romo Lopez A, Rodríguez Alvarez FM, *et al.* Efficacy of tocilizumab in patients with moderate-to-severe corticosteroid-resistant Graves orbitopathy: A randomized clinical trial. Am J Ophthalmol. 2018;195:181-190. https://doi.org/10.1016/j.ajo.2018.07.038

[12] Smith TJ, Kahaly GJ, Ezra DG, Fleming JC, Dailey RA, Tang RA, *et al.* Teprotumumab for thyroid-associated ophthalmopathy. New Engl J Med. 2017;376(18):1748-1761. https://doi.org/10.1056/NEJMoa1614949

[13] Douglas RS, Kahaly GJ, Patel A, Sile S, Thompson EHZ, Perdok R, *et al.* Teprotumumab for the treatment of active thyroid eye disease. New Engl J Med. 2020;382(4):341-352. https://doi.org/10.1056/NEJMoa1910434

[15] Kahaly GJ, Douglas RS, Holt RJ, Sile S, Smith TJ. Teprotumumab for patients with active thyroid eye disease: A pooled data analysis, subgroup analyses, and off-treatment follow-up results from two randomised, double-masked, placebo-controlled, multicentre trials. Lancet Diabetes Endocrinol. 2021;9(6):360-372. https://doi.org/10.1016/S2213-8587(21)00056-5

[15] Douglas RS, Kahaly GJ, Ugradar S, Elflein H, Ponto KA, Fowler BT, *et al.* Teprotumumab efficacy, safety, and durability in longer-duration thyroid eye disease and re-treatment: OPTIC-X Study. Ophthalmology. 2022;129(4):438-449. https://doi.org/10.1016/j.ophtha.2021.10.017

[16] Douglas RS, Batten R, Qadeer RA, Cameron C. Meta-Analysis of proptosis response in thyroid eye disease: Efficacy of EUGOGO recommended treatment regimen with IV methylprednisolone. J Endocr Soc. 2021;5(s1):A842. https://doi.org/10.1210/jendso/bvab048.1718

[17] Zong AM, Giannakakos VP, Delbourgo Patton C, Barmettler A. Botulinum toxin treatment in thyroid eye disease: A systematic review and meta-analysis. Ophthalmic Plast Reconstr Surg. 2025;41(3):250-257. https://doi.org/10.1097/IOP.0000000000002852

[18] Bordaberry M, Marques DL, Pereira-Lima JC, Marcon IM, Schmid H. Repeated peribulbar injections of triamcinolone acetonide: A successful and safe treatment for moderate to severe Graves' ophthalmopathy. Acta Ophthalmol. 2009;87(1):58-64. https://doi.org/10.1111/j.1755-3768.2008.01171.x

[19] Jandhyala R. Delphi, non-RAND modified Delphi, RAND/UCLA appropriateness method and a novel group awareness and consensus methodology for consensus measurement: A systematic literature review. Curr Med Res Opin. 2020;36(11):1873-1887. https://doi.org/10.1080/03007995.2020.1816946

[20] Tamhankar MA, Raza S, Brutsaert E, Urdániz E, Vainilovich Y, Heyes A, *et al.* The burden of illness in thyroid eye disease: Current state of the evidence. Front Ophthalmol (Lausanne). 2025;5(2025):1565762. https://doi.org/10.3389/fopht.2025.1565762

**Supplementary Table 1.** Grouping of clinical scenarios into categories

| Category | Description | Included Scenarios |
|---|---|---|
| **1. Moderate TED without functional urgency or refractoriness** | Patients with active moderate disease, without imminent visual threat, with good response or no previous treatment, and without critical comorbidities. This category represents typical cases of moderate TED in early or subacute phases, with controllable symptoms and without treatment failure or visual emergencies. | **Scenario 1** – Active moderate TED, <6 months of evolution, intermittent diplopia, without optic neuropathy, without risk factors, stabilized thyroid status.<br>**Scenario 2** – Active moderate TED, >6 months of evolution, progressive exophthalmos (>24 mm), without diplopia or neuropathy, active smoker.<br>**Scenario 8** – Active moderate TED, without relevant exophthalmos, but with persistent diplopia and CAS=3, without other comorbidities.<br>Scenario 11 – Active moderate TED, without significant functional or visual symptoms, but with exophthalmos or eyelid edema that impacts quality of life for cosmetic reasons. CAS=3. Non-smoker, euthyroid. |
| **2. Moderate TED with refractoriness, comorbidities, or risk factors** | Moderate cases with complicating factors for treatment choice: Therapeutic failure, limiting comorbidities, active smoking, or high functional impact. This category includes patients in whom treatment indication is complex due to insufficient response to glucocorticoids, conditions limiting their use, or risks modifying the benefit–risk balance. | **Scenario 4** – Active moderate TED, significant diplopia and high CAS, no response after 12 weeks of IV glucocorticoids.<br>**Scenario 5** – Active moderate TED, partial response to IV glucocorticoids, but with uncontrolled diabetes mellitus.<br>**Scenario 7** – Active moderate TED, exophthalmos and orbital pain, active smoker, no response to initial treatment with glucocorticoids.<br>**Scenario 9** – Patient with active moderate-to-severe TED who presents an absolute contraindication or documented intolerance to systemic glucocorticoids.<br>**Scenario 12** – Active moderate TED with unilateral or bilateral corneal thinning due to exposure. |

| | | |
|---|---|---|
| **3. Severe TED with acute visual involvement or optic neuropathy** | Severe disease scenarios with clear visual risk (optic neuropathy, acute visual loss), with or without contraindication to standard treatment. This category gathers ophthalmologic emergencies in which timely intervention is critical to preserve vision. | **Scenario 3** – Active severe TED, with incipient optic neuropathy, <6 months of evolution, without contraindications to glucocorticoids.<br>**Scenario 6** – Active severe TED with established optic neuropathy, and absolute contraindication to IV glucocorticoids.<br>**Scenario 10** – Patient with active severe TED, bilateral optic neuropathy with rapid visual loss (days) and imaging findings suggestive of orbital compression. No contraindication to IV glucocorticoids.<br>**Scenario 14** – Severe TED with globe luxation and acute optic neuropathy due to stretching. |
| **4. Severe TED with severe corneal complications** | Cases of TED with severe ocular surface involvement: corneal thinning or perforation, with structural and visual risk. This category distinguishes a less frequent but highly severe subgroup, in which corneal involvement predominates and requires a different management approach compared to optic neuropathy. | **Scenario 13** – Active severe TED with unilateral or bilateral corneal perforation. |

**Source:** Authors.

Supplementary Table 2. RAND/UCLA Round 2 ratings by category and scenario

| Category | Scenario | Intervention | Median | IPRAS | Disagreement | Classification |
|---|---|---|---|---|---|---|
| 1 | 1 | Intravenous glucocorticoids | 9 | 8.35 | FALSE | Appropriate |
| | 1 | Teprotumumab | 9 | 7.60 | FALSE | Appropriate |
| | 1 | Adalimumab | 3 | 5.95 | FALSE | Inappropriate |
| | 1 | Somatostatin analogues | 1 | 7.60 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 1 | Cyclophosphamide | 2 | 6.85 | FALSE | Inappropriate |
| | 1 | Cyclosporine | 2 | 7.45 | FALSE | Inappropriate |
| | 1 | Oral glucocorticoids | 3 | 5.50 | FALSE | Inappropriate |
| | 1 | Intravenous immunoglobulin | 2 | 6.85 | FALSE | Inappropriate |
| | 1 | Plasmapheresis | 1 | 7.30 | FALSE | Inappropriate |
| | 1 | Rituximab | 3 | 5.35 | FALSE | Inappropriate |
| | 1 | Sirolimus | 3 | 5.95 | FALSE | Inappropriate |
| | 1 | Colchicine | 4 | 4.60 | FALSE | Uncertain |
| | 1 | Methotrexate | 4 | 4.75 | FALSE | Uncertain |
| | 1 | Mycophenolate mofetil | 6 | 3.25 | FALSE | Uncertain |
| | 1 | Tocilizumab | 6 | 4.75 | FALSE | Uncertain |
| | 1 | Botulinum toxin | 5 | 4.60 | FALSE | Uncertain |
| | 1 | Periocular triamcinolone | 6 | 4.00 | FALSE | Uncertain |
| | 2 | Intravenous glucocorticoids | 9 | 8.20 | FALSE | Appropriate |
| | 2 | Mycophenolate mofetil | 7 | 6.10 | FALSE | Appropriate |
| | 2 | Teprotumumab | 9 | 7.60 | FALSE | Appropriate |
| | 2 | Tocilizumab | 8 | 6.25 | FALSE | Appropriate |
| | 2 | Periocular triamcinolone | 7 | 5.20 | FALSE | Appropriate |
| | 2 | Adalimumab | 2 | 6.85 | FALSE | Inappropriate |
| | 2 | Somatostatin analogues | 2 | 7.45 | FALSE | Inappropriate |
| | 2 | Cyclophosphamide | 2 | 7.45 | FALSE | Inappropriate |
| | 2 | Cyclosporine | 2 | 7.45 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 2 | Colchicine | 2 | 6.70 | FALSE | Inappropriate |
| | 2 | Intravenous immunoglobulin | 2 | 7.60 | FALSE | Inappropriate |
| | 2 | Methotrexate | 3 | 4.90 | FALSE | Inappropriate |
| | 2 | Plasmapheresis | 2 | 6.55 | FALSE | Inappropriate |
| | 2 | Sirolimus | 2 | 6.10 | FALSE | Inappropriate |
| | 2 | Oral glucocorticoids | 4 | 4.45 | FALSE | Uncertain |
| | 2 | Rituximab | 4 | 4.00 | FALSE | Uncertain |
| | 2 | Botulinum toxin | 6 | 3.85 | FALSE | Uncertain |
| | 8 | Intravenous glucocorticoids | 8 | 7.60 | FALSE | Appropriate |
| | 8 | Teprotumumab | 9 | 8.20 | FALSE | Appropriate |
| | 8 | Tocilizumab | 8 | 6.85 | FALSE | Appropriate |
| | 8 | Botulinum toxin | 7 | 3.70 | FALSE | Appropriate |
| | 8 | Periocular triamcinolone | 8 | 6.70 | FALSE | Appropriate |
| | 8 | Adalimumab | 3 | 6.10 | FALSE | Inappropriate |
| | 8 | Somatostatin analogues | 2 | 6.85 | FALSE | Inappropriate |
| | 8 | Cyclophosphamide | 3 | 6.70 | FALSE | Inappropriate |
| | 8 | Cyclosporine | 3 | 6.85 | FALSE | Inappropriate |
| | 8 | Colchicine | 3 | 6.70 | FALSE | Inappropriate |
| | 8 | Oral glucocorticoids | 3 | 5.95 | FALSE | Inappropriate |
| | 8 | Intravenous immunoglobulin | 2 | 6.55 | FALSE | Inappropriate |
| | 8 | Plasmapheresis | 2 | 5.35 | FALSE | Inappropriate |
| | 8 | Methotrexate | 4 | 5.35 | FALSE | Uncertain |
| | 8 | Mycophenolate mofetil | 6 | 3.10 | FALSE | Uncertain |

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 8 | Rituximab | 4 | 4.00 | FALSE | Uncertain |
| | 8 | Sirolimus | 4 | 5.35 | FALSE | Uncertain |
| | 11 | Intravenous glucocorticoids | 9 | 7.45 | FALSE | Appropriate |
| | 11 | Teprotumumab | 9 | 8.20 | FALSE | Appropriate |
| | 11 | Tocilizumab | 8 | 6.70 | FALSE | Appropriate |
| | 11 | Periocular triamcinolone | 8 | 6.70 | FALSE | Appropriate |
| | 11 | Adalimumab | 3 | 5.35 | FALSE | Inappropriate |
| | 11 | Somatostatin analogues | 2 | 7.45 | FALSE | Inappropriate |
| | 11 | Cyclophosphamide | 3 | 6.70 | FALSE | Inappropriate |
| | 11 | Cyclosporine | 3 | 6.85 | FALSE | Inappropriate |
| | 11 | Colchicine | 2 | 6.10 | FALSE | Inappropriate |
| | 11 | Oral glucocorticoids | 2 | 5.80 | FALSE | Inappropriate |
| | 11 | Intravenous immunoglobulin | 2 | 6.55 | FALSE | Inappropriate |
| | 11 | Methotrexate | 2 | 5.35 | FALSE | Inappropriate |
| | 11 | Plasmapheresis | 2 | 5.35 | FALSE | Inappropriate |
| | 11 | Rituximab | 3 | 4.75 | FALSE | Inappropriate |
| | 11 | Sirolimus | 3 | 5.35 | FALSE | Inappropriate |
| | 11 | Botulinum toxin | 3 | 3.85 | FALSE | Inappropriate |
| | 11 | Mycophenolate mofetil | 4 | 2.95 | TRUE | Uncertain (disagreement) |
| 2 | 4 | Mycophenolate mofetil | 7 | 5.65 | FALSE | Appropriate |
| | 4 | Teprotumumab | 9 | 8.35 | FALSE | Appropriate |
| | 4 | Tocilizumab | 9 | 7.60 | FALSE | Appropriate |
| | 4 | Somatostatin analogues | 2 | 7.60 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 4 | Cyclophosphamide | 2 | 6.70 | FALSE | Inappropriate |
| | 4 | Cyclosporine | 2 | 7.30 | FALSE | Inappropriate |
| | 4 | Colchicine | 2 | 7.30 | FALSE | Inappropriate |
| | 4 | Intravenous glucocorticoids | 3 | 5.80 | FALSE | Inappropriate |
| | 4 | Oral glucocorticoids | 2 | 7.00 | FALSE | Inappropriate |
| | 4 | Intravenous immunoglobulin | 2 | 5.95 | FALSE | Inappropriate |
| | 4 | Plasmapheresis | 2 | 5.35 | FALSE | Inappropriate |
| | 4 | Sirolimus | 2 | 6.70 | FALSE | Inappropriate |
| | 4 | Adalimumab | 5 | 4.15 | FALSE | Uncertain |
| | 4 | Methotrexate | 4 | 5.35 | FALSE | Uncertain |
| | 4 | Rituximab | 6 | 2.95 | FALSE | Uncertain |
| | 4 | Botulinum toxin | 5 | 3.55 | FALSE | Uncertain |
| | 4 | Periocular triamcinolone | 5 | 2.65 | FALSE | Uncertain |
| | 5 | Mycophenolate mofetil | 7 | 6.10 | FALSE | Appropriate |
| | 5 | Teprotumumab | 9 | 8.35 | FALSE | Appropriate |
| | 5 | Tocilizumab | 8 | 7.60 | FALSE | Appropriate |
| | 5 | Adalimumab | 3 | 4.75 | FALSE | Inappropriate |
| | 5 | Somatostatin analogues | 2 | 6.85 | FALSE | Inappropriate |
| | 5 | Cyclophosphamide | 2 | 5.95 | FALSE | Inappropriate |
| | 5 | Cyclosporine | 3 | 6.10 | FALSE | Inappropriate |
| | 5 | Colchicine | 2 | 5.35 | FALSE | Inappropriate |
| | 5 | Intravenous glucocorticoids | 3 | 5.50 | FALSE | Inappropriate |
| | 5 | Oral glucocorticoids | 2 | 7.60 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 5 | Intravenous immunoglobulin | 2 | 6.55 | FALSE | Inappropriate |
| | 5 | Plasmapheresis | 2 | 5.35 | FALSE | Inappropriate |
| | 5 | Methotrexate | 5 | 4.45 | FALSE | Uncertain |
| | 5 | Rituximab | 6 | 2.95 | FALSE | Uncertain |
| | 5 | Sirolimus | 4 | 4.75 | FALSE | Uncertain |
| | 5 | Botulinum toxin | 5 | 4.60 | FALSE | Uncertain |
| | 5 | Periocular triamcinolone | 5 | 2.35 | TRUE | Uncertain (disagreement) |
| | 7 | Mycophenolate mofetil | 7 | 5.20 | FALSE | Appropriate |
| | 7 | Teprotumumab | 9 | 8.35 | FALSE | Appropriate |
| | 7 | Tocilizumab | 8 | 6.85 | FALSE | Appropriate |
| | 7 | Periocular triamcinolone | 7 | 4.15 | FALSE | Appropriate |
| | 7 | Adalimumab | 3 | 4.75 | FALSE | Inappropriate |
| | 7 | Somatostatin analogues | 2 | 6.85 | FALSE | Inappropriate |
| | 7 | Cyclophosphamide | 2 | 6.70 | FALSE | Inappropriate |
| | 7 | Colchicine | 3 | 5.35 | FALSE | Inappropriate |
| | 7 | Intravenous glucocorticoids | 3 | 6.85 | FALSE | Inappropriate |
| | 7 | Oral glucocorticoids | 1 | 8.35 | FALSE | Inappropriate |
| | 7 | Intravenous immunoglobulin | 3 | 5.95 | FALSE | Inappropriate |
| | 7 | Plasmapheresis | 2 | 5.20 | FALSE | Inappropriate |
| | 7 | Rituximab | 3 | 5.20 | FALSE | Inappropriate |
| | 7 | Sirolimus | 3 | 5.50 | FALSE | Inappropriate |
| | 7 | Cyclosporine | 4 | 6.10 | FALSE | Uncertain |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 7 | Methotrexate | 4 | 4.75 | FALSE | Uncertain |
| | 7 | Botulinum toxin | 5 | 2.50 | FALSE | Uncertain |
| | 9 | Mycophenolate mofetil | 7 | 6.10 | FALSE | Appropriate |
| | 9 | Teprotumumab | 9 | 8.35 | FALSE | Appropriate |
| | 9 | Tocilizumab | 9 | 8.20 | FALSE | Appropriate |
| | 9 | Adalimumab | 3 | 4.15 | FALSE | Inappropriate |
| | 9 | Somatostatin analogues | 2 | 6.10 | FALSE | Inappropriate |
| | 9 | Cyclophosphamide | 2 | 6.70 | FALSE | Inappropriate |
| | 9 | Cyclosporine | 2 | 6.70 | FALSE | Inappropriate |
| | 9 | Colchicine | 3 | 5.50 | FALSE | Inappropriate |
| | 9 | Intravenous glucocorticoids | 1 | 8.35 | FALSE | Inappropriate |
| | 9 | Oral glucocorticoids | 1 | 8.35 | FALSE | Inappropriate |
| | 9 | Intravenous immunoglobulin | 2 | 6.40 | FALSE | Inappropriate |
| | 9 | Periocular triamcinolone | 2 | 6.25 | FALSE | Inappropriate |
| | 9 | Methotrexate | 4 | 4.00 | FALSE | Uncertain |
| | 9 | Plasmapheresis | 5 | 4.60 | FALSE | Uncertain |
| | 9 | Rituximab | 6 | 3.85 | FALSE | Uncertain |
| | 9 | Sirolimus | 4 | 5.35 | FALSE | Uncertain |
| | 9 | Botulinum toxin | 4 | 3.85 | FALSE | Uncertain |
| | 12 | Intravenous glucocorticoids | 8 | 6.85 | FALSE | Appropriate |
| | 12 | Mycophenolate mofetil | 7 | 3.85 | FALSE | Appropriate |
| | 12 | Teprotumumab | 9 | 7.60 | FALSE | Appropriate |
| | 12 | Tocilizumab | 7 | 5.80 | FALSE | Appropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 12 | Botulinum toxin | 7 | 5.05 | FALSE | Appropriate |
| | 12 | Periocular triamcinolone | 7 | 6.10 | FALSE | Appropriate |
| | 12 | Adalimumab | 3 | 4.60 | FALSE | Inappropriate |
| | 12 | Somatostatin analogues | 2 | 6.25 | FALSE | Inappropriate |
| | 12 | Cyclophosphamide | 2 | 6.70 | FALSE | Inappropriate |
| | 12 | Cyclosporine | 2 | 6.70 | FALSE | Inappropriate |
| | 12 | Colchicine | 3 | 5.95 | FALSE | Inappropriate |
| | 12 | Oral glucocorticoids | 2 | 6.85 | FALSE | Inappropriate |
| | 12 | Intravenous immunoglobulin | 3 | 6.55 | FALSE | Inappropriate |
| | 12 | Methotrexate | 3 | 5.35 | FALSE | Inappropriate |
| | 12 | Plasmapheresis | 3 | 5.35 | FALSE | Inappropriate |
| | 12 | Rituximab | 4 | 5.20 | FALSE | Uncertain |
| | 12 | Sirolimus | 4 | 4.90 | FALSE | Uncertain |
| | 12 | Botulinum toxin | 7 | 6.60 | FALSE | Appropriate (7–9) |
| 3 | 3 | Intravenous glucocorticoids | 9 | 8.35 | FALSE | Appropriate |
| | 3 | Mycophenolate mofetil | 7 | 5.05 | FALSE | Appropriate |
| | 3 | Teprotumumab | 9 | 8.35 | FALSE | Appropriate |
| | 3 | Tocilizumab | 8 | 6.70 | FALSE | Appropriate |
| | 3 | Periocular triamcinolone | 7 | 3.85 | FALSE | Appropriate |
| | 3 | Adalimumab | 2 | 6.25 | FALSE | Inappropriate |
| | 3 | Somatostatin analogues | 2 | 7.45 | FALSE | Inappropriate |
| | 3 | Cyclophosphamide | 3 | 6.70 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | 3 | Cyclosporine | 3 | 6.85 | FALSE | Inappropriate |
| | 3 | Colchicine | 3 | 6.25 | FALSE | Inappropriate |
| | 3 | Oral glucocorticoids | 3 | 5.35 | FALSE | Inappropriate |
| | 3 | Intravenous immunoglobulin | 3 | 6.70 | FALSE | Inappropriate |
| | 3 | Methotrexate | 3 | 4.90 | FALSE | Inappropriate |
| | 3 | Plasmapheresis | 3 | 6.55 | FALSE | Inappropriate |
| | 3 | Rituximab | 3 | 5.95 | FALSE | Inappropriate |
| | 3 | Sirolimus | 3 | 6.25 | FALSE | Inappropriate |
| | 3 | Botulinum toxin | 4 | 2.95 | FALSE | Uncertain |
| | 6 | Mycophenolate mofetil | 7 | 5.50 | FALSE | Appropriate |
| | 6 | Teprotumumab | 9 | 8.35 | FALSE | Appropriate |
| | 6 | Tocilizumab | 8 | 7.00 | FALSE | Appropriate |
| | 6 | Adalimumab | 3 | 5.95 | FALSE | Inappropriate |
| | 6 | Somatostatin analogues | 2 | 6.85 | FALSE | Inappropriate |
| | 6 | Cyclophosphamide | 2 | 5.95 | FALSE | Inappropriate |
| | 6 | Cyclosporine | 2 | 6.70 | FALSE | Inappropriate |
| | 6 | Colchicine | 2 | 5.35 | FALSE | Inappropriate |
| | 6 | Intravenous glucocorticoids | 1 | 8.35 | FALSE | Inappropriate |
| | 6 | Oral glucocorticoids | 1 | 8.35 | FALSE | Inappropriate |
| | 6 | Intravenous immunoglobulin | 3 | 5.95 | FALSE | Inappropriate |
| | 6 | Sirolimus | 3 | 5.35 | FALSE | Inappropriate |
| | 6 | Periocular triamcinolone | 2 | 7.60 | FALSE | Inappropriate |
| | 6 | Methotrexate | 4 | 5.35 | FALSE | Uncertain |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | 6 | Plasmapheresis | 5 | 4.00 | FALSE | Uncertain |
| | 6 | Rituximab | 5 | 3.85 | FALSE | Uncertain |
| | 6 | Botulinum toxin | 5 | 4.60 | FALSE | Uncertain |
| | 10 | Intravenous glucocorticoids | 9 | 8.35 | FALSE | Appropriate |
| | 10 | Mycophenolate mofetil | 7 | 4.75 | FALSE | Appropriate |
| | 10 | Teprotumumab | 9 | 8.20 | FALSE | Appropriate |
| | 10 | Tocilizumab | 8 | 6.25 | FALSE | Appropriate |
| | 10 | Somatostatin analogues | 2 | 6.25 | FALSE | Inappropriate |
| | 10 | Cyclophosphamide | 2 | 6.70 | FALSE | Inappropriate |
| | 10 | Cyclosporine | 2 | 6.70 | FALSE | Inappropriate |
| | 10 | Colchicine | 2 | 6.70 | FALSE | Inappropriate |
| | 10 | Oral glucocorticoids | 2 | 6.70 | FALSE | Inappropriate |
| | 10 | Intravenous immunoglobulin | 2 | 6.55 | FALSE | Inappropriate |
| | 10 | Plasmapheresis | 2 | 5.35 | FALSE | Inappropriate |
| | 10 | Rituximab | 2 | 6.70 | FALSE | Inappropriate |
| | 10 | Botulinum toxin | 3 | 4.75 | FALSE | Inappropriate |
| | 10 | Periocular triamcinolone | 2 | 5.80 | FALSE | Inappropriate |
| | 10 | Adalimumab | 4 | 4.15 | FALSE | Uncertain |
| | 10 | Methotrexate | 4 | 5.35 | FALSE | Uncertain |
| | 10 | Sirolimus | 4 | 5.35 | FALSE | Uncertain |
| | 14 | Intravenous glucocorticoids | 9 | 8.35 | FALSE | Appropriate |
| | 14 | Tocilizumab | 8 | 4.75 | FALSE | Appropriate |
| | 14 | Adalimumab | 2 | 7.45 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | 14 | Somatostatin analogues | 2 | 7.45 | FALSE | Inappropriate |
| | 14 | Cyclophosphamide | 1 | 7.30 | FALSE | Inappropriate |
| | 14 | Cyclosporine | 1 | 6.70 | FALSE | Inappropriate |
| | 14 | Colchicine | 2 | 7.45 | FALSE | Inappropriate |
| | 14 | Oral glucocorticoids | 2 | 6.85 | FALSE | Inappropriate |
| | 14 | Intravenous immunoglobulin | 2 | 6.10 | FALSE | Inappropriate |
| | 14 | Methotrexate | 2 | 6.10 | FALSE | Inappropriate |
| | 14 | Plasmapheresis | 2 | 7.15 | FALSE | Inappropriate |
| | 14 | Rituximab | 2 | 6.85 | FALSE | Inappropriate |
| | 14 | Sirolimus | 2 | 5.50 | FALSE | Inappropriate |
| | 14 | Botulinum toxin | 2 | 6.70 | FALSE | Inappropriate |
| | 14 | Periocular triamcinolone | 2 | 6.85 | FALSE | Inappropriate |
| | 14 | Mycophenolate mofetil | 3 | 3.85 | TRUE | Uncertain (disagreement) |
| | 14 | Teprotumumab | 8 | 3.10 | TRUE | Uncertain (disagreement) |
| 4 | 13 | Adalimumab | 1 | 7.45 | FALSE | Inappropriate |
| | 13 | Somatostatin analogues | 2 | 7.60 | FALSE | Inappropriate |
| | 13 | Cyclophosphamide | 1 | 7.60 | FALSE | Inappropriate |
| | 13 | Cyclosporine | 1 | 7.60 | FALSE | Inappropriate |
| | 13 | Colchicine | 1 | 7.60 | FALSE | Inappropriate |
| | 13 | Intravenous glucocorticoids | 2 | 7.60 | FALSE | Inappropriate |
| | 13 | Oral glucocorticoids | 1 | 7.60 | FALSE | Inappropriate |
| | 13 | Intravenous immunoglobulin | 2 | 7.60 | FALSE | Inappropriate |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4 | 13 | Methotrexate | 1 | 7.60 | FALSE | Inappropriate |
| | 13 | Mycophenolate mofetil | 2 | 7.45 | FALSE | Inappropriate |
| | 13 | Plasmapheresis | 2 | 7.60 | FALSE | Inappropriate |
| | 13 | Rituximab | 2 | 6.55 | FALSE | Inappropriate |
| | 13 | Sirolimus | 1 | 7.60 | FALSE | Inappropriate |
| | 13 | Teprotumumab | 2 | 6.70 | FALSE | Inappropriate |
| | 13 | Tocilizumab | 2 | 7.45 | FALSE | Inappropriate |
| | 13 | Botulinum toxin | 2 | 7.30 | FALSE | Inappropriate |
| | 13 | Periocular triamcinolone | 1 | 7.60 | FALSE | Inappropriate |

**Note.** IPRAS: interpercentile range adjusted for symmetry; Disagreement; Refers to panel disagreement per RAND/UCLA, where TRUE, items are labeled "Uncertain (disagreement)".

**Source:** Authors.