



# HHS Public Access

Author manuscript

*Pharmacoepidemiol Drug Saf.* Author manuscript; available in PMC 2017 July 31.

Published in final edited form as:

*Pharmacoepidemiol Drug Saf.* 2016 April ; 25(4): 363–371. doi:10.1002/pds.3959.

## Using a patient-centered approach to benefit-harm assessment in treatment decision-making: a case study in uveitis

Tsung Yu<sup>1,2</sup>, Janet T. Holbrook<sup>1</sup>, Jennifer E. Thorne<sup>1,3</sup>, and Milo A. Puhan<sup>2</sup>

<sup>1</sup>Department of Epidemiology, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland <sup>2</sup>Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland <sup>3</sup>Department of Ophthalmology/Wilmer Eye Institute, The Johns Hopkins University School of Medicine, Baltimore, Maryland

### Abstract

**Background**—Synthesizing evidence from comparative effectiveness trials can be difficult since multiple outcomes of different importance are to be considered. The goal of this study was to demonstrate an approach to conducting quantitative benefit-harm assessment that considers patient preferences.

**Methods**—We conducted a benefit-harm assessment using data from the Multicenter Uveitis Steroid Treatment Trial that compared corticosteroid implant versus systemic corticosteroids and immunosuppression in non-infectious intermediate, posterior, and panuveitis. We focused on clinical outcomes considered important to patients, including visual acuity, development of cataracts/glaucoma, need for eye surgery, prescription-requiring hypertension, hyperlipidemia and infections. Patient preferences elicited in a recent survey were then incorporated into our assessment of the benefit-harm balance.

**Results**—Benefit-harm metrics were calculated for each time point that summarized the numbers of outcomes, caused or prevented by implant therapy versus systemic therapy if 1000 patients were treated. The benefit-harm metric was -129 (95% CI: -242 to -14), -317 (-436 to -196), -390 (-514 to -264) and -526 (-687 to -368) at 6, 12, 18, and 24 months follow-up, respectively, suggesting that systemic therapy may have a better benefit-harm balance. However, measures of quality of life for patients treated with implant therapy were found to be better than patients treated with systemic therapy over the same time period.

**Conclusions**—Results of benefit-harm assessment were different from the prospectively collected quality of life data during trial follow-up. Future studies should explore the reasons for such discrepancies and the strength and weakness of each method to assess treatment benefits and harms.

---

**Corresponding Author:** Janet T. Holbrook, PhD, MPH, Phone: 443-287-5791, Fax: 775-269-0836, jholbro1@jhu.edu, Address: 415 N. Washington St, 2nd Floor, Baltimore, MD 21231.

**Ethics Statement:**

The MUST trial has been registered on clinicaltrials.gov (identifier NCT00132691) and this study was approved by the institutional review board at the Johns Hopkins Bloomberg School of Public Health.

## Keywords

Benefit-harm assessment; patient-centered; decision-making

---

## INTRODUCTION

Randomized comparative effectiveness trials are used to make head-to-head comparisons between competing treatments in terms of effectiveness and safety.[1] To synthesize different benefit and harm outcome data and to assess the overall benefit-harm balance for patients in a comparative effectiveness trial, a number of quantitative approaches for benefit-harm assessment have been developed.[2–4] Some of these approaches use a single metric such as an index or a probability of one treatment providing more benefits than harms as compared against another treatment, to aid the interpretation of comparative treatment effectiveness.[2, 3] Since patients may perceive each outcome differently, such summary metric of benefits and harms also allows for taking into account the patient preferences by weighting the outcomes differently depending on how important the outcomes are to patients.[3]

Joint occurrence of benefit and harm outcomes is another important, though usually neglected, issue to consider when doing quantitative benefit-harm assessment.[3–7] Occurrence of outcomes can be correlated rather than independent if, for instance, they typically co-occur with each other within patients. Nonetheless, treatment effects on each outcome are conventionally reported separately in the literature.[3, 8] Some approaches that account for such joint occurrence of outcomes have been developed, yet are not commonly applied.[9–13] For example, Chuang-Stein et al. illustrated their method to simultaneously consider patient's response (treatment benefits) and side effects (harms) in a clinical trial of antihypertensive[10] and generated summary measures that facilitate decision-making.

The objective of the present study was to demonstrate an approach to benefit-harm assessment that specifically incorporates patient preferences. We used available outcome data from a clinical trial, the Multicenter Uveitis Steroid Treatment (MUST) Trial, which compared two treatment strategies in patients with non-infectious intermediate, posterior, and panuveitis.[14] The MUST Trial randomized patients to either fluocinolone acetonide implant or systemic corticosteroids plus immunosuppression when indicated and followed patients for 24 months. No significant difference in visual acuity change from baseline was detected. Ocular adverse effects (e.g., cataracts and glaucoma) were more likely with implant therapy, and surprisingly, except for an increased use of antibiotics, there was no increased risk of systemic adverse events with systemic therapy.[15] Regarding patient-reported outcomes, the study found that patients in the implant group, on average, had better vision-related and generic health-related quality of life than patients in the systemic group. [15] Given the different types of outcomes with varying treatment effects and the uncertainty about the overall benefit-harm balance between the two treatment strategies, the MUST Trial is a good case study for testing methods for a structured and quantitative benefit-harm assessment.

## METHODS

### Data Source: The MUST Trial

Details on the study design and primary results of the MUST Trial were reported previously. [14, 15] In brief, patients aged 13 or older who had non-infectious intermediate, posterior, or panuveitis in at least one eye and who were indicated for systemic corticosteroids were randomized. Patients randomized to the implant therapy group received in the eligible eye a surgical fluocinolone acetonide implant (0.59 mg) that delivers corticosteroids intravitreally. Patients randomized to the systemic therapy group were treated with oral corticosteroids (prednisone) supplemented with immunosuppressive agents.[14, 16] The primary outcome was change in best-corrected visual acuity from baseline to 24 months.

### Approaches Used and Outcomes Included for Benefit-Harm Assessment

We used an analytical approach that was similar to the one proposed by Gail and colleagues, [2, 17] which, according to our categorization of benefit-harm methodologies,[3] compares multiple benefit and harm outcomes on a common metric (referred to as the “trade-off index” by others[4]). Besides, we tested the approach proposed by Chuang-Stein et al. that can account for joint occurrence of outcomes.[10, 12] We focused on available clinical outcome data measured in the MUST Trial that we deemed important to patients, as opposed to biomarkers or surrogate outcomes that are not necessarily linked to clinical outcomes. We did not include quality of life measures in our assessment because they reflect the consequences of many benefits and harms that are difficult to disentangle, but these measures served as a good comparison for our results of benefit-harm assessment.

The goal of treating these patients is to preserve their vision from getting worse and, at the same time, to minimize adverse effects caused by the treatments. For the vision outcome at each time point, we categorized the patients into “remaining at better than 20/40 or improving to better than 20/40” and “remaining at worse than 20/40 or decreasing to worse than 20/40”, for their best-corrected visual acuity in the better-seeing eye. The 20/40 cut-off was chosen since it is commonly used to define low vision in previous research[18] and also a cut-off that we can associate with the vision standard for driving when communicating with patients. For ocular and systemic adverse effects of treatments, we examined at each time point the proportion of patients who had the following outcomes: incident cataracts, incident glaucoma, requiring cataract surgery, requiring intraocular pressure-lowering surgery (glaucoma surgery), prescription-requiring hypertension, prescription-requiring hyperlipidemia, and prescription-requiring infections. The time frames of our assessment of these outcomes were 6, 12, 18, and 24 months after randomization. The unit of the analysis was the individual patient.

### Relative Importance for Outcomes

In a benefit-harm assessment where multiple outcomes are put on a common metric, these outcomes should be weighted properly based on their relative importance. Therefore, in our analysis we assigned the weights using data from a survey that elicited patient ratings of adverse treatment outcomes in non-infectious uveitis.[19] Briefly, we surveyed 182 patients with non-infectious uveitis and used best-worst scaling method to elicit their ratings of the

outcomes included in our assessment. The survey results suggested that most patients considered impaired vision, development of glaucoma, and needing eye surgery more worrying outcomes as compared against development of cataracts, needing medicine for hypertension/hyperlipidemia and systemic infections (e.g., sinusitis). We used these estimates of relative importance to derive weights for our benefit-harm assessment.

### Benefit-Harm Metric

We summarized the treatment effects on different outcomes (weighted by each outcome's relative importance) in a "benefit-harm metric" that reflects the benefit-harm balance. First, based on the MUST Trial data, we calculated the number of outcomes (outcome  $x$ ) if 1000 patients were treated with implant therapy ( $N_{X,IMP}$ ) or systemic therapy ( $N_{X,SYS}$ ). Second, we calculated the outcomes prevented or caused if 1000 patients were treated with implant therapy versus systemic therapy ( $N_x = N_{X,SYS} - N_{X,IMP}$ ). A positive number represents the number of outcomes prevented and a negative number represents the number of outcomes caused by implant therapy. Third, we assigned weights ( $W_x$ , relative importance) to these outcomes according to the estimates obtained in the patient preferences survey. We then computed a benefit-harm metric that summarizes the overall numbers caused or prevented by implant therapy and that incorporates the relative importance of outcomes. If the benefit-harm metric is positive, it suggests that implant therapy is superior to systemic therapy since the implant therapy prevented more outcomes overall.

We computed the benefit-harm metrics at 6, 12, 18, and 24 months after randomization. In addition, we varied the weights assigned to outcomes as sensitivity analyses to evaluate whether our study conclusions would change with regards to different assigned weights. We used bootstrapping approach to incorporate the statistical uncertainty where we obtained 10000 replicates of the metric to compute its 95% confidence interval (CI) and the probability that the metric is positive. Analyses were performed using Stata 11.2 (StataCorp LP, College Station, TX) and R statistical software 3.0.1.

### Joint Occurrence of Benefits and Harms

To examine the joint occurrence of benefit and harm outcomes in patients in the MUST Trial, we defined two benefit categories (based on patients' vision outcome) and three harm categories (based on their experience with adverse effects). Definition of each category can be found in Table 1. We then created "benefit-harm categories" that consider benefits and harms jointly: two benefit categories  $\times$  three harm categories, e.g., "With benefits/No harms", "With benefits/Minor harms", "With benefits/Moderate harms", and so on. One more category ("Missing data") was created that included patients with missing data of their visual acuity. Based on their experience with benefit and harm outcomes during follow-up, patients in the MUST Trial were then assigned to each of the seven benefit-harm categories. We compared the distributions between the two treatment groups at each time point.

## RESULTS

The clinical outcome data (proportion of patients who ever had each outcome at 6, 12, 18 and 24 months follow-up) stratified by the randomized treatment group are given in Table 2.

In summary, there was little difference between the two groups in outcomes of visual acuity, prescription-requiring hypertension and hyperlipidemia. But the implant therapy group was associated with higher risks for requiring eye surgery and for developing glaucoma and cataracts. The systemic group was associated with a higher risk for prescription-requiring infections. Table 2 also shows the weights assigned in the main analysis derived from the estimates we obtained in the preference-elicitation survey. We varied the weights in the first sensitivity analysis by assigning 1.0 to more worrying outcomes and 0.5 to less worrying outcomes. In sensitivity analysis two, we assigned 1.0 to the visual acuity outcome (as this is the primary outcome in the trial) and 0.5 to other more worrying outcomes and 0.25 to less worrying outcomes.

An example of calculation of the benefit-harm metric (main analysis, 24 months follow-up) is provided in Table 3. We calculated the numbers of outcomes if 1000 patients were treated with implant or systemic therapy and the numbers of outcomes caused or prevented by implant therapy. We assumed at baseline 82% of patients had already had cataracts and 3% of patients had had glaucoma (based on data from the MUST Trial). For example, our calculations show that there would be 226 and 60 patients with incident glaucoma if 1000 patients were treated with implant and systemic therapy, respectively. Thus, 166 incident glaucoma would be in excess if 1000 patients were treated with implant therapy versus systemic therapy. The numbers caused or prevented by implant therapy were then multiplied by the weights, and were summed to compute the benefit-harm metric.

Results of the main and sensitivity analyses of the benefit-harm metric at each time point are shown in Table 4. In the main analysis, the benefit-harm metric is -129, -317, -390 and -526 at 6, 12, 18, and 24 months follow-up, respectively, suggesting that implant therapy may have a worse benefit-harm balance than systemic therapy. The 95% CIs and the probability that the metric is positive, meaning that implant therapy would be superior to systemic therapy, is 1%, 0%, 0%, and 0%, respectively. Results of the sensitivity analyses are similar. The benefit-harm metrics are more and more distant from 0 (negative) across the 6, 12, 18, and 24 months follow-up, and the probabilities of the index being positive are all small or 0%.

We plotted the distribution of patients to assigned benefit-harm categories by treatment group over time in the Figure, in which the joint occurrence of benefits and harms can be examined. For example at 24 months follow-up, proportions of patients (66% in implant group versus 71% in systemic group) assigned to the “With benefits” category were similar. But after combined with their experience with harms, most patients (49%) in the implant group were in the “With benefits/Moderate harms” category, while only 23% in systemic group were in this category.

## DISCUSSION

Our approach to benefit-harm assessment used clinical events from a multicenter trial as well as patient ratings of the importance of these events to assign weights. The benefit-harm assessment suggested that systemic therapy may have a better benefit-harm balance than implant therapy for patients with non-infectious intermediate, posterior, and panuveitis. This

is contrary to findings from the main report of the MUST Trial, which showed better health-related quality of life for patients treated with implant therapy compared to patients treated with systemic therapy (see Table 5).[15]

A key factor that may bias our results and lead to the discrepancy with the quality of life data is the selection of the actual outcomes for inclusion in our quantitative benefit-harm assessment. Outcome selection should be done with great caution by investigators to define the possible outcomes that influence the benefit-harm balance while avoiding double counting.[4] It is dependent upon the decision-making context and the perspective of decision-makers (patient's, clinician's or policy-maker's perspective), and it requires a comprehensive literature review, consultation with many stakeholders, or conducting additional qualitative studies to identify the appropriate outcomes.[4] Our selection of outcomes was limited to the available trial data since we did a *post hoc* analysis as this is currently still common for benefit-harm assessments. Had we had the chance to plan the study before the trial started, patients could have been involved to define the important outcomes that should be considered. We think future benefit-harm assessment studies should include *a priori* planning to construct the metric that includes a rigorous procedure to define important outcomes and to complete outcome ascertainment.

For our study we chose, though without any intent to influence the results of the study, more ocular events (5) than systemic events (3). Since ocular events had higher incidence rates in both therapy groups and were significantly increased in the implant group, the benefit-harm balance is more likely to be found against implant therapy. Furthermore, the systemic events included in the preference weighting were rather minor events whereas ocular events included a more diverse set of potentially severe harms including glaucoma and diminished vision. Janz and colleagues[20] showed in their study that at least moderate worry about blindness was reported by 34% of the patients newly diagnosed with glaucoma, and even after five years with proper treatment, 48% of patients reported still to be at least a little worried. Our preference based approach to conducting benefit-harm assessment may be biased by such overwhelming fear of ocular events and blindness. This can explain to some extent the discrepant results of our analysis compared to the quality-of-life data that express how patients actually felt.

Another factor that may bias our results is the often hypothetical nature of preference elicitation surveys. It is uncommon that patients will have experienced all outcomes addressed in a survey. Indeed, almost no patients in our survey had had experienced all outcomes. Although our survey[19] did not suggest a difference between the preferences of patients who had longstanding disease and thus experienced more outcomes and the preferences of less experienced patients, more evidence is needed to elucidate the impact of being more or less familiar with specific benefit and harm outcomes. On the other hand, although quality of life measures do reflect patient's actual experience, intrinsically integrating over the clinical events and functional effects, most of these instruments were designed to be used across different diseases and do not reflect all outcomes directly relevant to specific treatments and patient groups. Therefore, these instruments are likely to measure different constructs rather than specific benefit and harm outcomes. An advantage of quality

of life measures though is that they usually have been well-validated and used in a consistent manner across trials and cost-effective analyses.[21]

The treatment effects are commonly reported in clinical trials as relative risks. But because the same relative risk can translate into considerably different effects as the respective absolute risks are different, it is necessary to use absolute risks to put multiple outcomes on the same metric.[2] In our analysis, we created a common metric that synthesized and weighted different outcomes properly to estimate the benefit-harm balance. This approach is transparent in that the specific outcomes considered, weights assigned, and the uncertainty of the data are clearly laid out. In contrast, for the more qualitative interpretation of a benefit-harm balance, i.e., without a quantitative synthesis of the outcome data and outcome importance, the assumptions and judgment made at every step of the analysis are often not explicit or made transparent.

Our study also demonstrated how the joint occurrence of benefits and harms can be examined when doing benefit-harm assessment. The evaluation of treatment effect is usually done separately for each outcome.[8] However, it may sometimes be of interest to patients and clinicians if data of the joint impact of benefit and harm outcomes in the *same* individual are available.[8] For example, we computed and plotted the distribution of benefit-harm categories (in which the joint probability of the occurrence of benefits and harms was examined) by treatment group. Such information is not commonly reported in the current literature of trials, but may be helpful when patients and clinicians desire to make personalized treatment decisions.

Assigning weights to outcomes is another controversial part of a benefit-harm assessment and is inevitably subjective. Nevertheless, it remains essential, as in clinical practice, because it is not sensible if all outcomes are considered to be of equal importance. We decided to conduct a preference-elicitation survey of patients with non-infectious uveitis to help us choose appropriate weights for each outcome. We were comfortable with using the findings from the survey to determine the weights since we found the relative importance of outcomes was comparable across different patient groups.[19] Another advantage of using a quantitative approach to benefit-harm assessment is its transparency and reproducibility because the analysis can easily be repeated and modified if anyone disagrees with the outcomes included or the weights assigned to outcomes. In addition, a reproducible approach greatly facilitates sensitivity analyses, which are essential to assess if the benefit-harm balance changes as different assumptions are made or different data are chosen.

This study has some limitations. The metric used for our benefit-harm assessment was “event-based” and focused on clinical outcomes deemed important to patients. We were unable to capture in our analysis some subtle issues and treatment burdens (e.g., sleep disorder and mood change), which may be associated with higher-dose oral corticosteroid therapy. The study time frame is 24 months after randomization, the same time points at which the primary outcome was measured. Therefore, our model does not capture the potential systemic complications that may occur years later after treatment. For example, there is a debate about whether an increased risk of cancer is associated with immunosuppressive agents in patients with uveitis.[22] Such concerns about severe adverse

events in the future can be one important factor that affects patients' decision-making, which we did not consider in this analysis, and may alter the patient's assessment of benefits versus harms. Finally, when generating the benefit-harm categories and computing the benefit-harm metrics, we categorized visual acuity outcome and combined different harm outcomes. This can lead to much information loss but it is almost inevitable in any benefit-harm assessment when trying to reduce some of the multi-dimensionality and to facilitate treatment decision-making.

In summary, this case study demonstrates an approach to select outcomes, consider the joint occurrence of benefits and harms and incorporate patient preferences for a quantitative benefit-harm assessment. In line with the recent interest in patient-centered outcomes research,[23, 24] our approach may be useful and deserves future replications. The finding that the results of the benefit-harm assessment were different from assessments of quality of life in the same participants over the same period highlights the need for more research. Specifically, future studies should explore the methods for properly identifying outcomes that should be included in a benefit-harm assessment and explore the underlying constructs of common quality of life measures across different diseases.

## Acknowledgments

### Financial Support:

Supported by cooperative agreements from the National Eye Institute to Mount Sinai School of Medicine (U10 EY 014655), The Johns Hopkins University Bloomberg School of Public Health (U10 EY 014660), and the University of Wisconsin, Madison, School of Medicine (U10 EY 014656).

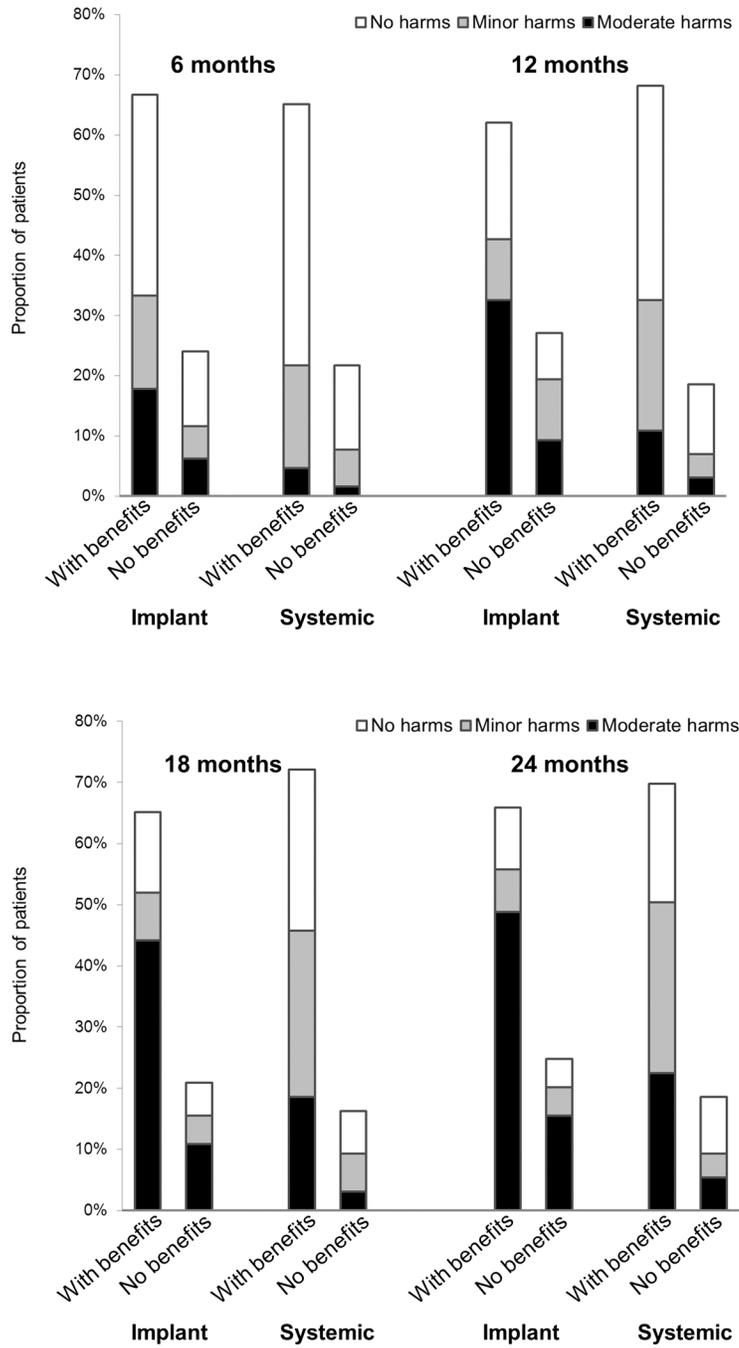
## References

1. IOM (Institute of Medicine). Initial National Priorities for Comparative Effectiveness Research. Washington DC: The National Academies Press; 2009.
2. Gail MH. Using absolute risks to assess the risks and benefits of treatment. *Thorax*. 2014 Jul; 69(7): 604–5. [PubMed: 24550059]
3. Puhan MA, Singh S, Weiss CO, Varadhan R, Boyd CM. A framework for organizing and selecting quantitative approaches for benefit-harm assessment. *BMC Med Res Methodol*. 2012; 12:173. 2288-12-173. [PubMed: 23163976]
4. Mt-Isa S, Hallgreen CE, Wang N, Callréus T, Genov G, Hirsch I, Hobbiger SF, Hockley KS, Luciani D, Phillips LD, Quartey G, Sarac SB, Stoeckert I, Tzoulaki I, Micallef A, Ashby D. IMI-PROTECT benefit-risk participants. Balancing benefit and risk of medicines: a systematic review and classification of available methodologies. *Pharmacoepidemiol Drug Saf*. 2014 Jul; 23(7):667–78. [PubMed: 24821575]
5. Chuang-Stein, C. Proceedings of the Fourth Seattle Symposium in Biostatistics: Clinical Trials. New York: Springer; 2013. Quantitative risk/benefit assessment: where are we?; p. 119-135.
6. Guo JJ, Pandey S, Doyle J, Bian B, Lis Y, Raisch DW. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy-report of the ISPOR risk-benefit management working group. *Value Health*. 2010; 13(5):657–66. [PubMed: 20412543]
7. Ouellet D. Benefit-risk assessment: the use of clinical utility index. *Expert Opin Drug Saf*. 2010; 9(2):289–300. [PubMed: 20175698]
8. Kraemer HC, Frank E. Evaluation of comparative treatment trials: assessing clinical benefits and risks for patients, rather than statistical effects on measures. *JAMA*. 2010; 304(6):683–4. [PubMed: 20699462]

9. Boers M, Brooks P, Fries JF, Simon LS, Strand V, Tugwell P. A first step to assess harm and benefit in clinical trials in one scale. *J Clin Epidemiol*. 2010 Jun; 63(6):627–32. [PubMed: 19800197]
10. Chuang-Stein C, Mohberg NR, Sinkula MS. Three measures for simultaneously evaluating benefits and risks using categorical data from clinical trials. *Stat Med*. 1991; 10(9):1349–59. [PubMed: 1925166]
11. Frank E, Kupfer DJ, Rucci P, Lotz-Wallace M, Levenson J, Fournier J, Kraemer HC. Simultaneous evaluation of the harms and benefits of treatments in randomized clinical trials: demonstration of a new approach. *Psychol Med*. 2012; 42(4):865–73. [PubMed: 21861951]
12. Wisniewski SR, Chen CC, Kim E, Kan HJ, Guo Z, Carlson BX, Tran QV, Pikalov A. Global benefit-risk analysis of adjunctive aripiprazole in the treatment of patients with major depressive disorder. *Pharmacoepidemiol Drug Saf*. 2009 Oct; 18(10):965–72. [PubMed: 19662630]
13. Yu T, Fain K, Boyd CM, Singh S, Weiss CO, Li T, Varadhan R, Puhan MA. Benefits and harms of roflumilast in moderate to severe COPD. *Thorax*. 2014 Jul; 69(7):616–22. [PubMed: 24347460]
14. Multicenter Uveitis Steroid Treatment Trial Research Group. The multicenter uveitis steroid treatment trial: rationale, design, and baseline characteristics. *Am J Ophthalmol*. 2010; 149(4):550–561. e10. [PubMed: 20097325]
15. Multicenter Uveitis Steroid Treatment (MUST) Trial Research Group. Randomized comparison of systemic anti-inflammatory therapy versus fluocinolone acetonide implant for intermediate, posterior, and panuveitis: the multicenter uveitis steroid treatment trial. *Ophthalmology*. 2011; 118(10):1916–26. [PubMed: 21840602]
16. Jabs DA, Rosenbaum JT, Foster CS, Holland GN, Jaffe GJ, Louie JS, Nussenblatt RB, Stiehm ER, Tessler H, Van Gelder RN, Whitcup SM, Yocum D. Guidelines for the use of immunosuppressive drugs in patients with ocular inflammatory disorders: recommendations of an expert panel. *Am J Ophthalmol*. 2000; 130(4):492–513. [PubMed: 11024423]
17. Gail MH, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, Vogel V. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst*. 1999 Nov 3; 91(21):1829–46. [PubMed: 10547390]
18. Klein R, Klein BE. The prevalence of age-related eye diseases and visual impairment in aging: current estimates. *Invest Ophthalmol Vis Sci*. 2013 Dec 13; 54(14):ORSF5-ORSF13.
19. Yu T, Holbrook JT, Thorne JE, Flynn TN, Van Natta ML, Puhan MA. Outcome Preferences in Patients With Noninfectious Uveitis: Results of a Best-Worst Scaling Study. *Invest Ophthalmol Vis Sci*. 2015 Oct 1; 56(11):6864–6872. [PubMed: 26501236]
20. Janz NK, Wren PA, Guire KE, Musch DC, Gillespie BW, Lichter PR. Collaborative Initial Glaucoma Treatment Study. Fear of blindness in the Collaborative Initial Glaucoma Treatment Study: patterns and correlates over time. *Ophthalmology*. 2007 Dec; 114(12):2213–20. [PubMed: 17490746]
21. Sugar EA, Holbrook JT, Kempen JH, Burke AE, Drye LT, Thorne JE, Louis TA, Jabs DA, Altaweel MM, Frick KD. Multicenter Uveitis Steroid Treatment (MUST) Trial Research Group. Cost-effectiveness of fluocinolone acetonide implant versus systemic therapy for noninfectious intermediate, posterior, and panuveitis. *Ophthalmology*. 2014 Oct; 121(10):1855–62. [PubMed: 24908205]
22. Khachatryan N, Kempen JH. Immunosuppressive therapy and cancer risk in ocular inflammation patients: fresh evidence and more questions. *Ophthalmology*. 2015 Feb; 122(2):219–21. [PubMed: 25618424]
23. Methodology Committee of the Patient-Centered Outcomes Research Institute (PCORI). Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA*. 2012 Apr 18; 307(15):1636–40. [PubMed: 22511692]
24. Gagnon MP, Desmartis M, Lepage-Savary D, Gagnon J, St-Pierre M, Rhainds M, Lemieux R, Gauvin FP, Pollender H, Légaré F. Introducing patients' and the public's perspectives to health technology assessment: A systematic review of international experiences. *Int J Technol Assess Health Care*. 2011 Jan; 27(1):31–42. [PubMed: 21262085]

### Key Points

- Assessing treatment benefits and harms can be challenging since multiple outcomes with different incidences and importance are at play.
- We conducted a benefit-harm assessment that considered patient preferences.
- Benefit-harm assessment provides a transparent way to show the benefit-harm balance of treatments and may be helpful to decision-makers.



**Figure.** For implant group (n=129) and systemic group (n=126), we categorized the patients first into “With benefits” or “No benefits” and further sub-categorized them into “No harms” “Minor harms” or “Moderate harms”. The definition of benefit and harm categories can be found in the “Methods” section. We plotted the distributions over time.

**Table 1**

Benefit-harm categories

|                                                                                                                                         | No harms                                                                                         | Minor harms                                                                                                                                                                                                | Moderate harms                                                                                                                                       |
|-----------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                                         | Patients had no adverse events of interest                                                       | Patients only had the adverse events that are less worrying including incident cataracts, prescription-requiring hypertension, prescription-requiring hyperlipidemia, or prescription-requiring infections | Patients had any of the adverse events that are more worrying including incident glaucoma, requiring cataract surgery and requiring glaucoma surgery |
| <b>With benefits</b><br>Patients' visual acuity of the better-seeing eye remained at better than 20/40 or improved to better than 20/40 | <b>(1) With benefits/No harms</b>                                                                | <b>(2) With benefits/Minor harms</b>                                                                                                                                                                       | <b>(3) With benefits/Moderate harms</b>                                                                                                              |
| <b>No benefits</b><br>Patients' visual acuity of the better-seeing eye remained at worse than 20/40 or decreased to worse than 20/40    | <b>(4) No benefits/No harms</b>                                                                  | <b>(5) No benefits/Minor harms</b>                                                                                                                                                                         | <b>(6) No benefits/Moderate harms</b>                                                                                                                |
|                                                                                                                                         | <b>(7) Missing data</b>                                                                          |                                                                                                                                                                                                            |                                                                                                                                                      |
|                                                                                                                                         | Patients died or were loss to follow-up, or their visual acuity was not or could not be measured |                                                                                                                                                                                                            |                                                                                                                                                      |

**Table 2**

Data input for benefit-harm assessment

| <b>Patient experiences with treatment outcomes over time for the two randomized treatment groups</b>                           |                                                              |                                                               |
|--------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|---------------------------------------------------------------|
|                                                                                                                                | <b>Implant therapy group<br/>Number of patients, n/N (%)</b> | <b>Systemic therapy group<br/>Number of patients, n/N (%)</b> |
| <b>Visual acuity of the better-seeing eye remained at worse than 20/40 or decreased to worse than 20/40 at each time point</b> |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 33/121 (27%)                                                 | 28/115 (24%)                                                  |
| <b>12 months</b>                                                                                                               | 37/119 (31%)                                                 | 24/115 (21%)                                                  |
| <b>18 months</b>                                                                                                               | 29/114 (25%)                                                 | 21/117 (18%)                                                  |
| <b>24 months</b>                                                                                                               | 33/118 (28%)                                                 | 24/114 (21%)                                                  |
| <b>Number of patients who ever had the event(s) at each time point</b>                                                         |                                                              |                                                               |
| <b>Incident glaucoma</b>                                                                                                       |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 0/116 (0%)                                                   | 0/115 (0%)                                                    |
| <b>12 months</b>                                                                                                               | 1/116 (1%)                                                   | 0/115 (0%)                                                    |
| <b>18 months</b>                                                                                                               | 1/116 (1%)                                                   | 0/114 (0%)                                                    |
| <b>24 months</b>                                                                                                               | 27/116 (23%)                                                 | 7/114 (6%)                                                    |
| <b>Requiring cataract surgery</b>                                                                                              |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 23/124 (19%)                                                 | 6/121 (5%)                                                    |
| <b>12 months</b>                                                                                                               | 43/124 (35%)                                                 | 16/120 (13%)                                                  |
| <b>18 months</b>                                                                                                               | 65/123 (53%)                                                 | 25/120 (21%)                                                  |
| <b>24 months</b>                                                                                                               | 74/121 (61%)                                                 | 30/119 (25%)                                                  |
| <b>Requiring glaucoma surgery</b>                                                                                              |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 8/124 (6%)                                                   | 3/121 (2%)                                                    |
| <b>12 months</b>                                                                                                               | 21/124 (17%)                                                 | 3/120 (3%)                                                    |
| <b>18 months</b>                                                                                                               | 33/123 (27%)                                                 | 4/120 (3%)                                                    |
| <b>24 months</b>                                                                                                               | 40/121 (33%)                                                 | 8/119 (7%)                                                    |
| <b>Incident cataracts</b>                                                                                                      |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 11/24 (46%)                                                  | 5/21 (24%)                                                    |
| <b>12 months</b>                                                                                                               | 22/24 (92%)                                                  | 8/21 (38%)                                                    |
| <b>18 months</b>                                                                                                               | 22/24 (92%)                                                  | 8/21 (38%)                                                    |
| <b>24 months</b>                                                                                                               | 23/24 (96%)                                                  | 10/21 (48%)                                                   |
| <b>Prescription-requiring hypertension</b>                                                                                     |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 2/124 (2%)                                                   | 3/121 (2%)                                                    |
| <b>12 months</b>                                                                                                               | 4/124 (3%)                                                   | 4/120 (3%)                                                    |
| <b>18 months</b>                                                                                                               | 4/123 (3%)                                                   | 5/120 (4%)                                                    |
| <b>24 months</b>                                                                                                               | 5/121 (4%)                                                   | 9/119 (8%)                                                    |
| <b>Prescription-requiring hyperlipidemia</b>                                                                                   |                                                              |                                                               |
| <b>6 months</b>                                                                                                                | 1/124 (1%)                                                   | 3/121 (2%)                                                    |
| <b>12 months</b>                                                                                                               | 2/124 (2%)                                                   | 3/120 (3%)                                                    |
| <b>18 months</b>                                                                                                               | 3/123 (2%)                                                   | 7/120 (6%)                                                    |
| <b>24 months</b>                                                                                                               | 3/121 (2%)                                                   | 8/119 (7%)                                                    |
| <b>Prescription-requiring infections</b>                                                                                       |                                                              |                                                               |

**Patient experiences with treatment outcomes over time for the two randomized treatment groups**

|                  | <b>Implant therapy group<br/>Number of patients, n/N (%)</b> | <b>Systemic therapy group<br/>Number of patients, n/N (%)</b> |
|------------------|--------------------------------------------------------------|---------------------------------------------------------------|
| <b>6 months</b>  | 25/124 (20%)                                                 | 27/121 (22%)                                                  |
| <b>12 months</b> | 32/124 (26%)                                                 | 38/120 (32%)                                                  |
| <b>18 months</b> | 38/123 (31%)                                                 | 52/120 (43%)                                                  |
| <b>24 months</b> | 45/122 (37%)                                                 | 57/119 (48%)                                                  |

**Weights assigned in the main and sensitivity analyses**

| <b>Outcome</b>                                                                                               | <b>Main analysis<br/>(weights obtained from<br/>the patient preferences<br/>survey)*</b> | <b>Sensitivity<br/>analysis one</b> | <b>Sensitivity<br/>analysis two</b> |
|--------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------|
| <b>Visual acuity of the better-seeing eyes remained at worse than 20/40 or decreased to worse than 20/40</b> | 0.8                                                                                      | 1.0                                 | 1.0                                 |
| <b>Incident glaucoma</b>                                                                                     | 0.7                                                                                      | 1.0                                 | 0.5                                 |
| <b>Requiring cataract surgery</b>                                                                            | 0.6                                                                                      | 1.0                                 | 0.5                                 |
| <b>Requiring glaucoma surgery</b>                                                                            | 0.6                                                                                      | 1.0                                 | 0.5                                 |
| <b>Incident cataracts</b>                                                                                    | 0.3                                                                                      | 0.5                                 | 0.25                                |
| <b>Prescription-requiring hypertension</b>                                                                   | 0.3                                                                                      | 0.5                                 | 0.25                                |
| <b>Prescription-requiring hyperlipidemia</b>                                                                 | 0.3                                                                                      | 0.5                                 | 0.25                                |
| <b>Prescription-requiring infections</b>                                                                     | 0.2                                                                                      | 0.5                                 | 0.25                                |

\* The preference estimates are on a -5 to 5 scale. We rescaled the median of the preference estimates to a 0 to 1 scale. For example, the median of the preference estimate of incident glaucoma is 2, and the weights we assigned in analysis is  $2 - (-5)/10$ , or 0.7.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Calculation of the benefit-harm metric at 24 months follow-up

Table 3

| Outcome                                                                                               | Number of outcomes if 1000 patients were treated   |                                                     | Number of outcomes caused or prevented by implant therapy (N <sub>X</sub> = N <sub>X,SYS</sub> - N <sub>X,IMP</sub> )* | Weight assigned to the outcome (W <sub>X</sub> ) | N <sub>X</sub> * W <sub>X</sub> |
|-------------------------------------------------------------------------------------------------------|----------------------------------------------------|-----------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|---------------------------------|
|                                                                                                       | Treated with implant therapy (N <sub>X,IMP</sub> ) | Treated with systemic therapy (N <sub>X,SYS</sub> ) |                                                                                                                        |                                                  |                                 |
| Visual acuity of the better-seeing eyes remained at worse than 20/40 or decreased to worse than 20/40 | 280                                                | 211                                                 | -69                                                                                                                    | 0.8                                              | -55                             |
| Incident glaucoma <sup>‡</sup>                                                                        | 226                                                | 60                                                  | -166                                                                                                                   | 0.7                                              | -116                            |
| Requiring cataract surgery                                                                            | 612                                                | 252                                                 | -359                                                                                                                   | 0.6                                              | -216                            |
| Requiring glaucoma surgery                                                                            | 331                                                | 67                                                  | -263                                                                                                                   | 0.6                                              | -158                            |
| Incident cataracts <sup>‡</sup>                                                                       | 173                                                | 86                                                  | -87                                                                                                                    | 0.3                                              | -26                             |
| Prescription-requiring hypertension                                                                   | 41                                                 | 76                                                  | 34                                                                                                                     | 0.3                                              | 10                              |
| Prescription-requiring hyperlipidemia                                                                 | 25                                                 | 67                                                  | 42                                                                                                                     | 0.3                                              | 13                              |
| Prescription-requiring infections                                                                     | 369                                                | 479                                                 | 110                                                                                                                    | 0.2                                              | 22                              |
| <b>Benefit-harm metric (I<sub>X</sub>)</b>                                                            |                                                    |                                                     |                                                                                                                        |                                                  | <b>-526</b>                     |

\* Numbers of outcomes caused are negative and numbers of outcomes prevented are positive.

<sup>‡</sup> Assuming 97% of patients at risk at baseline.

<sup>‡</sup> Assuming 18% of patients at risk at baseline.

**Table 4**

Main and sensitivity analyses of benefit-harm metrics at different time points after randomization

|                                 | <b>Benefit-harm metric</b> | <b>95% confidence interval*</b> | <b>Probability of the metric being positive*</b> |
|---------------------------------|----------------------------|---------------------------------|--------------------------------------------------|
| <b>Main analysis</b>            |                            |                                 |                                                  |
| <b>6 months</b>                 | -129                       | -242 to -14                     | 1%                                               |
| <b>12 months</b>                | -317                       | -436 to -196                    | 0%                                               |
| <b>18 months</b>                | -390                       | -514 to -264                    | 0%                                               |
| <b>24 months</b>                | -526                       | -687 to -368                    | 0%                                               |
| <b>Sensitivity analysis one</b> |                            |                                 |                                                  |
| <b>6 months</b>                 | -201                       | -362 to -39                     | 1%                                               |
| <b>12 months</b>                | -482                       | -665 to -298                    | 0%                                               |
| <b>18 months</b>                | -603                       | -800 to -412                    | 0%                                               |
| <b>24 months</b>                | -808                       | -1049 to -570                   | 0%                                               |
| <b>Sensitivity analysis two</b> |                            |                                 |                                                  |
| <b>6 months</b>                 | -115                       | -245 to 13                      | 4%                                               |
| <b>12 months</b>                | -292                       | -421 to -161                    | 0%                                               |
| <b>18 months</b>                | -339                       | -467 to -209                    | 0%                                               |
| <b>24 months</b>                | -439                       | -588 to -294                    | 0%                                               |

\* 95% confidence interval and the probability of the metric being positive were calculated based on the 10000 bootstrapping replicates.

Change in quality of life and weight over time for the implant and systemic treatment arms during the initial 2 years of follow-up (as randomized analysis)

**Table 5**

|                                                 |         | Estimated mean (SE) |             | Estimated mean change from enrollment (SE) |                   | Estimated treatment effect <sup>†</sup> (95% CI) |                    | P Value |
|-------------------------------------------------|---------|---------------------|-------------|--------------------------------------------|-------------------|--------------------------------------------------|--------------------|---------|
| Sample size*                                    | Implant | Systemic            | Implant     | Systemic                                   | Implant: Systemic | Systemic                                         | Implant: Systemic  | P Value |
| <b>Vision-related QoL</b>                       |         |                     |             |                                            |                   |                                                  |                    |         |
| <b>VFQ-25 (Overall Composite Score)</b>         |         |                     |             |                                            |                   |                                                  |                    |         |
| Enrollment                                      | 255     | 60.6 (2.4)          | 64.9 (2.5)  | n/a                                        | n/a               | n/a                                              | n/a                | n/a     |
| 6 months                                        | 233     | 71.4 (2.4)          | 66.3 (2.5)  | 10.77 (1.31)                               | 1.41 (1.38)       | 9.37 (5.64, 13.09)                               | 9.37 (5.64, 13.09) | <0.001  |
| 12 months                                       | 235     | 72.7 (2.5)          | 69.7 (2.5)  | 12.13 (1.60)                               | 4.86 (1.38)       | 7.27 (3.11, 11.42)                               | 7.27 (3.11, 11.42) | 0.001   |
| 24 months                                       | 232     | 72.1 (2.5)          | 71.7 (2.6)  | 11.44 (1.67)                               | 6.80 (1.58)       | 4.64 (0.14, 9.15)                                | 4.64 (0.14, 9.15)  | 0.043   |
| <b>Generic health-related QoL</b>               |         |                     |             |                                            |                   |                                                  |                    |         |
| <b>SF-36 (Mental Component Summary Score)</b>   |         |                     |             |                                            |                   |                                                  |                    |         |
| Enrollment                                      | 255     | 47.5 (1.3)          | 48.3 (1.2)  | n/a                                        | n/a               | n/a                                              | n/a                | n/a     |
| 6 months                                        | 236     | 51.1 (1.2)          | 45.5 (1.4)  | 3.58 (1.12)                                | -2.8 (1.14)       | 6.33 (3.19, 9.47)                                | 6.33 (3.19, 9.47)  | <0.001  |
| 12 months                                       | 234     | 51.0 (1.1)          | 46.5 (1.3)  | 3.52 (1.05)                                | -1.8 (1.11)       | 5.33 (2.33, 8.33)                                | 5.33 (2.33, 8.33)  | 0.001   |
| 24 months                                       | 232     | 50.0 (1.2)          | 47.2 (1.4)  | 2.55 (1.11)                                | -1.1 (1.15)       | 3.62 (0.49, 6.76)                                | 3.62 (0.49, 6.76)  | 0.023   |
| <b>SF-36 (Physical Component Summary Score)</b> |         |                     |             |                                            |                   |                                                  |                    |         |
| Enrollment                                      | 255     | 46.5 (1.2)          | 48.4 (1.1)  | n/a                                        | n/a               | n/a                                              | n/a                | n/a     |
| 6 months                                        | 236     | 46.4 (1.2)          | 46.6 (1.1)  | -0.07 (0.73)                               | -1.8 (0.67)       | 1.73 (-0.21, 3.67)                               | 1.73 (-0.21, 3.67) | 0.079   |
| 12 months                                       | 234     | 47.5 (1.2)          | 46.6 (1.1)  | 1.02 (0.78)                                | -1.9 (0.87)       | 2.89 (0.60, 5.18)                                | 2.89 (0.60, 5.18)  | 0.013   |
| 24 months                                       | 232     | 47.6 (1.2)          | 46.6 (1.2)  | 1.15 (0.83)                                | -1.8 (0.90)       | 2.95 (0.54, 5.36)                                | 2.95 (0.54, 5.36)  | 0.016   |
| <b>Health utility</b>                           |         |                     |             |                                            |                   |                                                  |                    |         |
| <b>EuroQol (Visual Analog Scale)</b>            |         |                     |             |                                            |                   |                                                  |                    |         |
| Enrollment                                      | 253     | 72.7 (1.9)          | 74.3 (2.0)  | n/a                                        | n/a               | n/a                                              | n/a                | n/a     |
| 6 months                                        | 233     | 75.0 (2.1)          | 72.5 (2.0)  | 2.34 (1.61)                                | -1.85 (1.90)      | 4.19 (-0.69, 9.07)                               | 4.19 (-0.69, 9.07) | 0.092   |
| 12 months                                       | 234     | 77.3 (1.9)          | 71.3 (2.2)  | 4.66 (1.36)                                | -3.00 (1.78)      | 7.65 (3.26, 12.05)                               | 7.65 (3.26, 12.05) | 0.001   |
| 24 months                                       | 232     | 78.0 (1.8)          | 73.4 (1.9)  | 5.29 (1.28)                                | -0.88 (1.78)      | 6.17 (1.87, 10.47)                               | 6.17 (1.87, 10.47) | 0.005   |
| <b>EuroQol (EQ-5D Health Utility Index)</b>     |         |                     |             |                                            |                   |                                                  |                    |         |
| Enrollment                                      | 254     | 0.81 (0.02)         | 0.83 (0.02) | n/a                                        | n/a               | n/a                                              | n/a                | n/a     |

|                    | Estimated mean (SE) |             |             | Estimated mean change from enrollment (SE) |             |                     | Estimated treatment effect <sup>†</sup> (95% CI) |                   |       | P Value |
|--------------------|---------------------|-------------|-------------|--------------------------------------------|-------------|---------------------|--------------------------------------------------|-------------------|-------|---------|
|                    | Sample size*        | Implant     | Systemic    | Implant                                    | Systemic    | Implant             | Systemic                                         | Implant: Systemic |       |         |
| <b>6 months</b>    | 236                 | 0.83 (0.02) | 0.82 (0.02) | 0.01 (0.01)                                | -0 (0.02)   | 0.03 (-0.02, 0.07)  |                                                  |                   | 0.25  |         |
| <b>12 months</b>   | 235                 | 0.83 (0.02) | 0.80 (0.02) | 0.02 (0.02)                                | -0 (0.02)   | 0.05 (0.00, 0.10)   |                                                  |                   | 0.041 |         |
| <b>24 months</b>   | 232                 | 0.84 (0.02) | 0.81 (0.02) | 0.02 (0.02)                                | -0 (0.02)   | 0.04 (-0.01, 0.09)  |                                                  |                   | 0.060 |         |
| <b>Weight (kg)</b> |                     |             |             |                                            |             |                     |                                                  |                   |       |         |
| <b>Enrollment</b>  | 255                 | 87.8 (2.6)  | 86.0 (2.7)  | n/a                                        | n/a         | n/a                 |                                                  |                   | n/a   |         |
| <b>6 months</b>    | 231                 | 88.2 (2.7)  | 87.0 (2.7)  | 0.40 (0.54)                                | 0.96 (0.75) | -0.56 (-2.38, 1.26) |                                                  |                   | 0.54  |         |
| <b>12 months</b>   | 229                 | 87.4 (2.7)  | 86.4 (2.7)  | -0.38 (0.70)                               | 0.38 (0.81) | -0.76 (-2.87, 1.34) |                                                  |                   | 0.48  |         |
| <b>24 months</b>   | 229                 | 87.3 (2.7)  | 86.2 (2.8)  | -0.41 (0.80)                               | 0.20 (0.86) | -0.61 (-2.91, 1.7)  |                                                  |                   | 0.61  |         |

CI = confidence interval; QoL = quality of life; SE = standard error.

\* Sample size: the number of individuals with data available at each visit.

<sup>†</sup> At each of the follow-up time points, the treatment effect is the model-based comparison of within treatment group change from enrollment (the difference of differences). A positive number favors implant.

Reprinted from Ophthalmology, 118/10, Multicenter Uveitis Steroid Treatment (MUST) Trial Research Group. Randomized comparison of systemic anti-inflammatory therapy versus fluocinolone acetonide implant for intermediate, posterior, and panuveitis: the multicenter uveitis steroid treatment trial, 1916-26, Copyright (2011), with permission from Elsevier.