Editorial

Uveitis Treatments: At What Cost Quality?
Russell N. Van Gelder, MD, PhD - Seattle, Washington
Bryan S. Lee, MD, JD - Seattle, Washington

The Agency for Health Care Research and Quality (AHRQ) was created by the US Medicare Prescription Drug, Improvement, and Modernization Act of 2003. This legislation authorized the AHRQ to conduct and support research with a focus on comparing the outcomes and effectiveness of different treatments, as well as communicating its findings to a variety of audiences. This act ushered in the era of comparative effectiveness trials, a cause that the US National Institutes of Health has also championed. One example, the National Institutes of Health-funded Comparison of Age Related Macular Degeneration Treatments Trials, compared ranibizumab and bevacizumab for treatment of exudative macular degeneration and found nearly identical outcomes despite markedly different costs, potentially resulting in tremendous cost savings for payers and patients.

The Multicenter Uveitis Steroid Treatment (MUST) trial is another National Institutes of Health-funded comparative effectiveness trial. It compares systemic medical treatment of noninfectious intermediate and posterior uveitis with a long-acting flucinolone acetonide implant. Initial results from this study demonstrated comparable outcomes for the visual acuity endpoint at 24 months (6 letters gained with implant vs 3.2 gained with systemic treatment, which were not statistically different). Given the absence of a clearly superior treatment, cost becomes a major issue. In this issue of Ophthalmology, the MUST study group reports analysis of the cost and the cost effectiveness of both treatment arms of the study. Using the data collected in the MUST trial, the investigators calculated the total cost of care for the 255 patients in the trial, including medication and device costs, treatment for subsequent surgeries (such as cataract and glaucoma procedures), and hospitalizations. The average treatment cost for patients with bilateral disease was $69,300 for the implant group over the 3 years of the study, compared with $52,500 for the systemically treated group (a difference that was statistically significant). For unilateral disease, average costs were $38,800 for the implant group versus $33,400 for the systemic group; this difference was not significant.

The investigators also analyzed quality of life through utility measures. Utility is a concept derived from decision analysis, which attempts to integrate quality and duration of life by measuring preferences for different health states. A utility of 1 is defined as perfect health, and a utility of 0 is death. Multiplying the utility by life expectancy allows calculation of quality-adjusted life-years (QALY), a common unit for comparing the effectiveness of health interventions across different medical conditions and specialties. The incremental cost-effectiveness ratio (ICER) for comparing 2 treatments is calculated by dividing the incremental cost of the studied intervention by the incremental utility gained. In cost-utility analysis, the ICER is measured in dollars per QALY. The traditional cutoff for proposing that an intervention is “cost effective” is $50,000 to 100,000/QALY. For example, cataract surgery is reported to have a cost utility of approximately $2000/QALY.

Using the EuroQol-EQ5D survey tool, the MUST trial found baseline utility for intermediate and posterior uveitis of approximately 0.82, a figure very close to other studies of intermediate uveitis. (By comparison, using other utility measurement tools, total blindness has a utility of approximately 0.25; in time-tradeoff methodology, this suggests that blind individuals would trade 75% of their remaining lifespan to regain perfect vision.) The MUST group found a small but significant difference in improvement in quality of life favoring the implant, and a marginally significant differential health utility change of +0.02 for the implant and −0.02 for systemic therapy.

Using these numbers, the group calculated the ICER for systemic therapy of noninfectious posterior uveitis versus the flucinolone acetonide implant. The authors found that there was a small, statistically nonsignificant difference in QALY between implant and systemic treatment; the mean difference in QALY was 0.06 for the bilateral cases and 0.13 for unilateral cases, favoring the implant in each instance. Based on these calculations, the authors conclude that the ICER for bilateral disease at 3 years was nearly $300,000 per QALY for those receiving implants, whereas for unilateral disease the ICER was $41,200 per QALY for the group receiving 1 implant. These results suggest that the cost effectiveness of bilateral implantation is limited, and in a resource-limited environment systemic therapy might be favored initially for patients with bilateral disease. Are these conclusions warranted?

Understanding the utility measurement is the key to answering this question. Use of the QALY theoretically allows cross-disciplinary comparison for an entity considering the funding of otherwise disparate interventions; for example, is it more desirable for society to allocate money to second eye cataract surgery or to total knee replacement? For such comparisons, it is critical that the method of estimation of health-related utilities be consistent across all specialties so that a QALY gained by treating orthopedic disease is identical to a QALY gained by treating cataract. This comparability is at the very heart of cost-utility analysis.

The 2 traditional methods for eliciting utilities are the standard gamble and time trade-off. Based on decision theory, the standard gamble asks participants what risk of immediate death they find acceptable in exchange for a...
return from a given health state to perfect health. Although this method has a very strong theoretical basis, it is cognitively challenging, and so the time trade-off was developed as an alternative.13 In this method, the participant is asked the number of years of remaining life willing to be foregone in return for perfect health for that shorter period of time. Time trade-off usually results in a lower utility rating for the same health state.14

The MUST trial used the EQ-5D to assess quality of life (see sample at http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Products/Sample_UK__English_EQ-5D-5L.pdf updated 6/13/14). This brief, 5-question tool asks individuals to rank their level of function in domains of mobility, self-care, activities of daily living, pain, and depression. The EQ-5D is converted to QALY using time trade-off data from a general population sample to value different non—disease-specific health states.15 The EQ-5D’s strengths are its brevity, wide use, and population-based, societal perspective. However, the EQ-5D is known to have a “ceiling effect,” meaning it is not sensitive to mild disease states, and it may not track longitudinal clinical changes well.16,17 A recent comparison of EQ-5D results to another validated, activity-based questionnaire concluded that the EQ-5D is insensitive to quality-of-life changes engendered by low vision.18 An alternative to using generic utility measures is to rely on vision-specific utilities.19,20 By asking participants what risk of instant blindness or instant death they would accept for “perfect vision” instead of “perfect health,” investigators generate quality-adjusted vision-years. Such measures are not true QALYs, meaning that their results are not generalizable or comparable with ICERs generated using the standard death—perfect health scale.21 The literature strongly suggests that utilities measured by different questionnaires are not interchangeable,22 and, even controlling for technique, utilities from dissimilar groups may not be comparable. It would be therefore be desirable in future comparative effectiveness trials in ophthalmology to utilize multiple quality of life measures, including validated vision-specific measures, to assist policymakers with the difficult decisions inherent in interpreting these results. Such an approach has been successfully followed in rheumatology.23

In the MUST trial, although the perceived ICER for bilateral disease seems to be large, favoring systemic therapy, the authors did not design this study to permit subgroup analysis. It is possible that specific subgroups of patients might show greater quality-of-life changes with 1 or the other treatment arm (e.g., specific immunomodulating drugs or doses of these drugs may have differential effects on quality of life) or diagnoses (e.g., patients with comorbid systemic inflammatory diseases may derive incidental benefit from systemic medication). Further, the study did not look at ICER for specific agents; for example, the ICER for implant versus methotrexate therapy may be very different than ICER for implant versus infliximab therapy (because the latter is a much more expensive systemic agent).

Because of these challenges in measuring utility in eye disease, and the resultant complexity of analysis of cost effectiveness, it is critical that analyses such as those presented in the current paper be interpreted cautiously by policymakers. The absolute difference in costs between the 2 arms for bilateral disease was on the order of 30%, but the ICER was magnified by the small change in QALY between groups. As noted, differences in cost or QALY may be much greater for specific subgroups of patients. An in-depth understanding of the particulars of this study and attendant caution in interpretation needs to be applied by policymakers before broad-based policies should be recommended on the basis of this analysis.

References


Footnotes and Financial Disclosures

Financial Disclosure(s): Supported in part by an unrestricted grant from Research to Prevent Blindness.