Cryotherapy proved to be far less successful in zone 1 disease than in zone 2 disease, where fortunately greater than 75% of the disease occurs. It is important to note that unfavorable outcomes as stage 5 (total retinal detachment) outnumbered stage 4B (a retinal fold extending through the macula) by a ratio of 13 to 1. The treated unfavorable results were far more common proportionately in zone 1 disease than in zone 2 disease.

The results of this study proved for the first time the availability of a treatment that improved the outlook for vision significantly over the natural history of the disease. The results of the study were quickly adopted both in the United States and worldwide as the standard of care, a tribute to the meticulous planning and execution by the investigators involved in carrying it out. Treatment took a further step forward when the indirect diode laser became available and was shown in a number of observational studies to be as effective as cryotherapy and easier in application for the infant and therapist. The CRYO-ROP trial proved to be the template for studies to follow. It also furnished a cadre of ophthalmologists trained in the design and implementation of randomized trial methods and the application of the methods to other problems in pediatric ophthalmology as they arose. It stands as one of the most successful trials undertaken with support of the National Eye Institute.

Submitted for Publication: November 22, 2006; accepted November 22, 2006.
Correspondence: John T. Flynn, MD, Department of Ophthalmology, Harkness Eye Institute, 635 W 165th St, New York, NY 10032 (jtf38@columbia.edu).
Financial Disclosure: None reported.
Funding/Support: This work was supported in part by the Communities Foundation of Texas.

REFERENCES


Evaluating the Cryotherapy for Retinopathy of Prematurity Study (CRYO-ROP)

Monte D. Mills, MD

The Multicenter Trial of Cryotherapy for Retinopathy of Prematurity Study (CRYO-ROP) remains one of the largest investigations ever organized for a pediatric ophthalmic disease. Spanning more than 20 years so far, and involving hundreds of ophthalmologists, neonatologists, photographers, visual acuity testers, and other investigators in 23 clinical centers across the United States, this gargantuan effort has produced high-quality data about the benefits of treatment of retinopathy of prematurity (ROP) with peripheral retinal ablation as well as the natural history of ROP and the development and measurement of visual function in young children with developmental and visual impairment. In addition, this trial led to innovations in the design and organization of randomized interventional trials that have gained wide acceptance and application in other fields.

What is the best measure of the success of a clinical trial? Formally, trials must be judged first on the ability to achieve the prospective goals of the investigation, generally a question of benefit and safety of the intervention being tested. The quality of the information measured, including its statistical measurement and clinical relevance; the number and quality of publications produced from the data and their subsequent citation; the development of new tools for future clinical research and patient care; and secondary trials founded on the initial results are also important measures of the influence of a clinical investigation. By all these measures, CRYO-ROP has been highly successful, both in achieving its formal goals as well as developing information to support a new understanding of ROP and the development of visual impairment in young children.

Ultimately, however, the most important measure of the impact of a clinical investigation is its influence on the clinical practice of medicine. CRYO-ROP changed tremendously the way we treat ROP from the previously highly variable strategies, including everything from non-intervention to treatment of even mild cases, to the current nearly universal application of peripheral retinal ablation at specific levels of severity based on evidence from this and subsequent randomized trials. Despite the evo-
ulation of better methods of peripheral ablation, with cryotherapy having been replaced nearly completely by laser photocoagulation, this study marks the point at which the treatment of ROP became data-driven and when the application of systematic screening and treatment for ROP became the standard of care in clinical practice of neonatology. By this measure, CRYO-ROP has been extremely successful.

**STUDY GOALS**

The CRYO-ROP study was developed in the 1980s to resolve a debate about the role of interventions for ROP. Case series, mostly from Japan in the 1970s and early 1980s, supported the concept of peripheral retinal ablation for ROP during the acute phase, prior to the development of retinal detachment and retrolental fibroplasia. However, even in treated patients, outcomes were frequently poor. A lack of standardized terminology and grading, and especially a lack of comparative clinical trial data, led to the reluctance of many clinicians to embrace this treatment.

At the time of the development of the CRYO-ROP study in the mid 1980s, accumulating anecdotal experience and small case series suggested the procedure was safe and showed possible benefit, but at the time of the initiation of this trial, there was no widely endorsed treatment strategy and therefore no universal screening criteria.

The primary goal of the CRYO-ROP study, as stated in the initial publication, was “to resolve uncertainty about the value of this (peripheral retinal ablation) treatment.” This clearly stated, clinically relevant primary goal was achieved in the initial publication from the study, which demonstrated a statistically and clinically relevant 49.3% reduction in the rate of unfavorable outcome in the treated eyes when compared with the control eyes. This primary goal has been a consistent part of the subsequent major publications from the study. Although the exact rate of reduction in unfavorable outcomes has varied depending on the length of follow-up and type of outcome measured, the CRYO-ROP study has been very successful in demonstrating the clinically relevant benefit of treatment in both anatomical and functional outcomes.

A secondary goal of CRYO-ROP has been to determine the long-term outcome of eyes with severe ROP both with and without ablative treatment. The very high completion rate of 5- and 10-year follow-up examinations for the study, the development of the subjects to the point at which more sophisticated visual function tests can be done, and the time elapsed from the initial treatment have allowed the study to achieve this long-term descriptive goal. These data remain the best long-term information available about the consequences of ROP with and without treatment against which other ROP interventions are measured.

**STUDY DESIGN**

Clinical investigations in ROP confront major barriers to study related to the age and relative immaturity of the subjects, the acute nature and rapid course of the disease, and the associated medical and developmental consequences of prematurity. The outcome of greatest interest, long-term visual function, cannot be measured until years after treatment. These technical and logistical limitations caused compromises in the design of the CRYO-ROP study related to recruitment and randomization, treatment, choice of outcomes to be measured, ethical factors, and statistical analysis.

**Patient Selection and Randomization**

Ideally, one would select perfectly matched cases to randomize for an interventional trial. In the case of ROP, patients could present with symmetrical involvement of both eyes, allowing randomization of one eye for treatment and the other for control. Randomizing one of symmetrically involved eyes also allows analysis based on discordant outcome between the two eyes, which would otherwise be assumed to have similar prognosis. However, some patients presented with asymmetric involvement, requiring a different randomization scheme, randomizing to treatment or no treatment in the single randomized eye.

Patients were recruited from admissions to the participating centers, potentially introducing a bias due to their tertiary nature and referral patterns. Post hoc analysis of enrollment suggests possible overrepresentation of black infants, multiple births, and boys.

The study design required staging by strict clinical criteria, but the nature of the disease and its rapid pace did not allow masked, photographic interpretation prior to randomization as has been done in other large randomized trials of retinal disease such as the Diabetic Retinopathy Treatment Study. A second examination by a different clinician was used as a measure to control for bias in staging. The lack of reproducible, objective assessments of pretreatment staging is a limitation of the study design and may have introduced bias and variability in the data despite the efforts in the study to train and standardize classification among the investigators and centers.

**Treatment Methods**

The stage of severity of ROP at which randomization occurred, now widely known as “threshold ROP,” was defined by the study as ROP in zone 1 or 2 with stage 3 in at least 5 contiguous or 8 discontinuous clock-hour segments (30° segments of the retinal circumference) with plus characteristics (dilation and tortuosity of posterior vessels equal to or greater than those in a standard photograph). This point was selected based on unpublished data and expert opinion that approximately 50% of eyes reaching this stage would be expected to progress to stages 4 and 5 (retinal detachment) or macular fold, which was defined in the study as an unfavorable outcome. However, the study designers were limited by the lack of large-scale natural history information. Retrospective analysis of data from CRYO-ROP, however, demonstrates that this was correct because 53% of untreated control eyes progressed to unfavorable outcome at the 3-month assessment.

The interpretation of the results of this study is limited to the benefit of treatment at the defined threshold. It is possible that treatment at an earlier stage may also have been effective in some cases, and this is the subject of a subsequent, second-generation study of laser pan-
retinal photocoagulation for ROP, the Early Treatment for Retinopathy of Prematurity trial (ETROP).9

The method of treatment used in the study, peripheral retinal transscleral cryotherapy, was the form of retinal ablation most widely used at that time for treatment for ROP, and virtually all of the earlier clinical reports had used cryotherapy. However, at about the same time as the trial, transpupillary laser photocoagulation was introduced as an alternative treatment.10 Subsequently, the wide availability of portable lasers with indirect ophthalmoscopic delivery, which have fewer associated local and systemic adverse effects, including pain and swelling, and do not require conjunctival incision, allowed laser photocoagulation to become the most widely used method of ROP treatment.11 Laser photocoagulation remains the predominant method of treatment to the present time, and cryotherapy is generally reserved for patients with opaque media or other technical limitations to transpupillary photocoagulation. Transscleral laser photocoagulation has also been reported.12

Data comparing cryotherapy with laser photocoagulation for ROP are limited. It is possible that laser may be associated with less peripheral visual field constriction, less pathological myopia, and lower risk of macular retinopathy,13,14 but it is unlikely that a large-scale comparative trial would be feasible because of the almost universal adoption of laser photocoagulation. The existing information suggests that laser treatment is equivalent in effectiveness to cryotherapy.15,16 Had the trial taken place just a few years later, the treatment selected for study would have almost certainly been photocoagulation rather than cryotherapy.

**Outcome Measures**

Masked interpretation of fundus photographs taken at 3 and 12 months was chosen as the primary end point of the study. This allowed standardization of interpretation and masking but was limited by the technical difficulty of photographing these young infants. At the 12-month outcome measurement, of the 246 infants (492 eyes) tested, only 385 eyes (78%) had photographs that were gradable.16 More significantly, this choice of anatomic rather than functional outcome seems to have overestimated the treatment benefit. In the initial 3-month analysis, the treatment benefit (reduction in rate of unfavorable outcome) was 39.5%,8 but at 5½ years, the visual acuity measurement indicated treatment benefit of 23.7%.17 Unfavorable anatomic outcome correlates best with severe vision loss, but “favorable” anatomic outcome does not mean normal vision. The choice of an anatomic outcome was appropriately expedient, allowing earlier analysis, and the benefit of treatment has remained clinically and statistically relevant in subsequent reports of visual function. However, the magnitude of the benefit of treatment was exaggerated by this design when compared with the vision outcomes.

Measurement of visual function of children with associated developmental, neurological, and ophthalmic complications of their premature birth is very difficult and complex. Recognition acuity was measurable at the 3½-year interval in an surprisingly large number of subjects: using a single-letter HOTV recognition test, almost half of the eyes that were not blind could be tested, and about 83% could be tested with preferential gratings (Teller Acuity Card procedure).18 Many anatomically intact eyes showed less than normal acuity, suggesting that a major factor in long-term visual function may be nonocular and may interfere with the estimate of treatment benefit in this study. The difficulty in precisely measuring visual acuity and the effect of nonocular visual pathway abnormalities in these children limit the interpretation of vision as a secondary outcome. Despite these limitations, a statistically significant and clinically relevant beneficial effect on visual function is seen when comparing the rate of severe vision loss (unfavorable outcome) between treated and control eyes at the 1-year, 3½-year, 5-year, and 10-year intervals, demonstrating a strong correlation between visual function at those intervals with the primary anatomic outcome at 3 months. Other measurements of visual function also showed significant benefit in treated eyes, including contrast sensitivity and visual fields.19,20

**Ethical Considerations**

Ethical concerns also complicated the development of this study design. If, as the trial goals intended, a benefit was seen with cryotherapy, randomized patients were eligible to have this benefit only in one eye and were not eligible to be treated in the opposite control eye. Asymmetrical patients assigned to the untreated control group also might suffer if treatment was beneficial. Because of this concern and others related to subject safety and ethical oversight, a special oversight committee was established within the study. This data and safety monitoring committee, composed of individuals with no other connection to the study, was empowered to monitor data during the recruitment phase and, if necessary, stop recruitment and complete data analysis prematurely if statistical analysis suggested that the goal of the study (determining benefit of treatment) could be reached with fewer than the planned number of eyes randomized or if other significant safety or ethical concerns existed.

The use of a separate data safety monitoring committee with authority to prematurely halt the trial if treatment success could be determined with less than complete enrollment was a relatively new innovation at the time of this trial, although it has become routine now. When, as anticipated by this study design, it became apparent that the benefit of treatment was significantly greater than initially estimated, the data safety monitoring committee exercised its prerogative, and subject recruitment and randomization was stopped.21 This event was a sensation widely discussed throughout the pediatric ophthalmology community and interpreted as evidence that cryotherapy was beneficial. Thus, even before the publication of the initial manuscript, many clinicians had adopted screening and treatment guidelines similar to those used in the study in anticipation of favorable results.

**Statistical Measures**

The study design, which included both infants with bilateral, symmetrical threshold ROP and infants with only one eye with threshold ROP, allowed analysis based on a paired sample (comparison of the outcomes of the treated
and untreated eyes of symmetrical cases) and independent case/control analysis of patients with asymmetrical ROP. The paired analysis provides a strong control for environmental covariables in these complex cases with many factors that might influence visual outcomes.

**STUDY RESULTS**

**Primary Outcome Measures**

The primary outcome measure in the CRYO-ROP trial was the masked grading of photographs taken at the 3-month and 12-month follow-up examinations. These results were reported in the initial manuscripts published from the study. No other large randomized trial of cryotherapy for ROP exists to compare the results, but the findings of 49.3% reduction in the unfavorable outcome rate at 3 months in the treated eyes vs the untreated eyes (21.8% compared with 43.0%) was dramatic and was sufficiently greater than expected by the statistical plan to allow early termination of the study. This result is also very clinically significant, especially considering that peripheral retinal ablation was the only effective treatment for acute retinopathy of prematurity at that time.

The early results of the trial also showed what had become apparent clinically to the ophthalmologists with experience cryotherapy: many eyes (21.8%) did not have sufficient response to treatment to prevent the macular distortion or retinal detachment that is the major cause of long-term vision loss from ROP.

What the primary outcomes did not demonstrate, but what has become clearer as the trial subjects have been followed up to an age when functional vision testing is possible, is that some patients with anatomically favorable outcomes still have poor vision. At the 10-year outcome measurement, 44.4% of eyes had unfavorable visual function outcomes (distance acuity ≤20/200), but only 27.2% showed unfavorable anatomical outcomes.

The primary anatomic outcome measure did not define the significant proportion of eyes, both in the treated and control groups, with favorable visual acuity and anatomical outcomes but moderately impaired vision (visual acuity of 20/40 to 20/200). The limitations of using an anatomical primary outcome, and the practicality of waiting until visual function could be measured in the infant subject, meant that this finding of a substantial group of eyes with reduced vision but better than “unfavorable” visual outcome (vision less than 20/40 but better than 20/200) is a frequent visual outcome at the 10-year measurement (30.4% of treated eyes, 14.2% of control eyes). Thus, the choice of an anatomical primary outcome tends to overestimate the functional benefit of treatment. Review of the 10-year outcome data suggests that most of the difference between the treated and control eyes is “rescue” of eyes from functional blindness (20/200 or worse) to moderate visual impairment with similar rates of normal vision in treated and control eyes.

**Secondary Outcome Measures**

Many other secondary outcomes have been reported in the publications of study data. The most important early visual function outcome reported was visual measurement using forced preferential looking tests of grating resolution (Teller Acuity Card procedure). Psychophysical measures of visual function, including grating resolution, have been used experimentally in prior studies to estimate visual acuity in nonverbal infants. However, this was the first major multicenter clinical trial to use a standardized procedure, including masked examiners and standardized presentation of gratings. The very high rate of testability and the precision of the test results when compared with optotype recognition provided a major validation of the Teller Acuity Card procedure, which has subsequently become widely used as a practical clinical tool and clinical research tool in pediatric ophthalmology.

Long-term assessment of ocular fundus findings reported over the length of the study have also been very useful in understanding the lifelong anatomical issues related to ROP. The large treatment effect has persisted throughout the period reported with significant reduction in retinal detachment and macular folds associated with treatment. However, long-term follow-up has identified a low but persistent incidence of late retinal detachment in both treated and untreated eyes.

Abnormal ocular growth and refractive error related to treated ROP has been another area of secondary outcome that has been illuminated by the CRYO-ROP study. Long-term measurement of refractive error suggests that both treated and untreated eyes with severe ROP have a high risk for high myopia.

The differences in treatment outcome among the eyes with different stages of ROP, especially zone 1 vs zone 2, have also been an important finding in this study. The study identified the zone 1 eyes to have the worst prognosis both with and without treatment. This important finding led to the development of the ETROP and revised treatment recommendations for earlier treatment of patients with involvement of zone 1.

**Natural History Control Outcomes**

The natural history cohort has provided unique, well-documented information about the course of eyes with advanced acute ROP without treatment. This cohort will likely never be reproduced because as a rule severe ROP is universally treated at the present time. The expectation of the planners of the trial, that approximately 50% of eyes would develop unfavorable outcome without treatment, was vindicated because 53% of untreated eyes had unfavorable photographic fundus appearance at the first measurement at 3 months. However, the primary outcome measure underestimates the functional outcome. At the 10-year interval, 47.9% of untreated eyes had retinal detachment or macular fold, and 62.1% had visual acuity of 20/200 or less.

Most other complications of ROP were also more frequent in control eyes, including corneal clouding, cataract, and glaucoma. Nevertheless, there remain a substantial group of eyes, 23.7% at 10 years, with good visual outcomes (visual acuity 20/40 or better) in untreated eyes.

The epidemiological description of patients reaching threshold ROP in the study prior to randomization was...
an important collection of data and has been used to develop screening recommendations. Screening recommendations from the CRYO-ROP study were the first evidence-based screening criteria and were extremely influential in developing the examination guidelines used currently throughout the United States and abroad. Recently, new information from the ETROP study has modified screening recommendations, but the CRYO-ROP data have remained the most influential evidence in the development of guidelines for ophthalmic consultation and screening in neonatal units.

STUDY CONCLUSIONS AND RECOMMENDATIONS

The initial conclusion of the expedited early publications and 3- and 12-month primary outcomes is that cryotherapy for threshold ROP is safe and effective when used in one eye. The authors of the reports are extremely cautious in recommending bilateral treatment, stating, “There are insufficient long-term data at present for us to reach a conclusion on the wisdom of routinely applying cryotherapy to both eyes of a patient with bilateral threshold ROP.” The authors refer to the rate of spontaneous resolution with good outcome among untreated eyes as well as the treated patients with progression and poor outcome and the unknown effects of cryotherapy in justifying this conservative recommendation. Some of this caution was removed in the conclusions of the 3-month interval publication. In this publication, the authors recommended bilateral treatment if there is threshold ROP in zone 1 but were still suggesting treatment of at least one eye for threshold ROP zone 2 and also recommended the uniform implementation of screening and access to cryotherapy for all neonatal units. The screening recommendations were implemented, and most children in neonatal units had access to treatment within a short period of time following the initial publication of the study.

INFLUENCE OF THE STUDY ON SUBSEQUENT CLINICAL INVESTIGATIONS

In addition to its huge impact on the clinical practice ofROP and pediatric ophthalmology, the CRYO-ROP study has influenced clinical trial design. The action of the data safety monitoring committee to terminate enrollment early due to statistically relevant early analysis, in one of the first instances of its kind in an ophthalmologic clinical trial, demonstrated the statistical and ethical value of this approach to clinical trials. CRYO-ROP has influenced the direction of clinical research in ROP. It led to the publication of 50 or more manuscripts directly and has influenced many more publications and studies indirectly. At least 3 additional large, multicenter trials sponsored by the National Institutes of Health (ETROP, Light-R, and STOP-ROP) have been direct offspring of the CRYO-ROP study, measuring various other parameters influencing ROP incidence and outcome. Vision testing protocols from CRYO-ROP have been widely used in other trials measuring visual acuity in young children, including the Teller Acuity Card procedure.

INFLUENCE OF THE STUDY ON CLINICAL PRACTICE IN MANAGEMENT OF ROP

The influence of a clinical trial on the broader management of patients is the ultimate measure of the study's impact. The CRYO-ROP study produced the essential data that led to the implementation of neonatal screening and peripheral retinal ablation for acute ROP throughout the United States and beyond. It is difficult to overestimate the impact of this well-designed trial in taking the concept of ablative treatment from controversial to near universal acceptance within a short period of time. By judicious choice of a threshold for treatment that allowed a statistically and clinically relevant improvement in outcomes, and by choosing outcome measures that were both expeditious (fundus appearance) and relevant (visual function), the study planners created a framework for answering the important question of treatment benefits from ablative therapy. The quality of the data obtained, the study design, and the decision to terminate enrollment early when the benefit of treatment was apparent set standards against which current and future clinical treatment trials will be measured.

Despite the cautious initial recommendations of the authors that routine treatment of both eyes with threshold ROP could not be supported by early study data, ablative treatment in fact became the standard of clinical care for all eyes with threshold ROP within the years after the initial results of the study were known and remains so today. Few studies have so influenced the clinical care of ophthalmic disease and been so broadly accepted within the clinical community in such a short period. Ablative treatment modalities have evolved as well, and currently laser retinal ablation is far more frequently used than cryotherapy, although the findings of the CRYO-ROP study are still the primary evidence used to support ablative treatment.

In addition to treatment of threshold ROP, the CRYO-ROP study collected the natural history and epidemiologic data that permitted evidence-based screening for ROP. These guidelines have evolved based on subsequent studies, but the basic CRYO-ROP epidemiologic data remain the foundation of the existing clinical guidelines for examination of premature infants. Long-term results of CRYO-ROP have also influenced the management of affected patients throughout the remainder of childhood, including long-term surveillance for late retinal changes, refractive error, and amblyopia management. We anticipate that the study subjects will continue to be measured and data continue to be reported as the cohort reaches visual maturity and adulthood.

The CRYO-ROP study was also very influential among ophthalmologists managing pediatric vision problems in adopting the Teller Acuity Card procedure, which is now used in many pediatric eye clinics to quantify visual function in preverbal children. The Teller Acuity Card procedure and similar forced preferential looking grating tests, while not perfect proxies for optotype visual acuity, have contributed to the clinical treatment of amblyopia, cataracts, and other vision problems in children in addition to ROP.

©2007 American Medical Association. All rights reserved.
Submitted for Publication: October 17, 2006; final revision received November 9, 2006; accepted November 18, 2006.

Correspondence: Monte D. Mills, MD, Division of Ophthalmology, Children’s Hospital of Philadelphia, 3400 Civic Center Blvd, Philadelphia, PA 19104 (mills@email.chop.edu).

Financial Disclosure: None reported.

REFERENCES