Supplementary Online Content


**eTable 1.** Adjudication rules used in the grading of e-ROP study images.

**eTable 2.** Pre-certification and certification scores of Trained Readers (TRs) in the e-ROP study.

**eFigure.** Study flow from enrollment through grading.

**eAppendix 1.** Grading process.

**eAppendix 2.** Grading image quality and morphological features of ROP.

This supplementary material has been provided by the authors to give readers additional information about their work.
**eTable 1. Adjudication rules used in the grading of e-ROP study images.**

<table>
<thead>
<tr>
<th>Question</th>
<th>Grader 1 Record</th>
<th>Other Grader Record</th>
<th>Final Record</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image Quality</strong></td>
<td>Values Agree</td>
<td>Poor</td>
<td>Adjudicate</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing Image</td>
<td>Good, Fair, or Poor</td>
<td></td>
<td>Adjudicate</td>
</tr>
<tr>
<td></td>
<td>Fair</td>
<td></td>
<td>Fair</td>
</tr>
<tr>
<td></td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field Definition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupil diameter</td>
<td>If difference ≤ 3 mm</td>
<td>Average of 2 Values</td>
<td></td>
</tr>
<tr>
<td>Pupil diameter</td>
<td>If difference &gt; 3 mm</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td>In Retinal images</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>disc placement in</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>clock hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from preferred site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In Retinal images</td>
<td>If difference ≤ 3 DD</td>
<td>Average of 2 Values</td>
<td></td>
</tr>
<tr>
<td>disc distance from</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preferred site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ROP</strong></td>
<td>Values Agree</td>
<td>Retain Value</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Values Disagree</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td><strong>Vessels</strong></td>
<td>Values Agree</td>
<td>Retain Value</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>Preplus</td>
<td>Preplus</td>
<td>Preplus</td>
</tr>
<tr>
<td>Normal</td>
<td>Plus</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>CD</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td>Preplus</td>
<td>Plus</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td>Preplus</td>
<td>CD</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td>Plus</td>
<td>CD</td>
<td>Adjudicate</td>
<td></td>
</tr>
<tr>
<td><strong>Zones Vascularized</strong></td>
<td>Values Agree</td>
<td>Retain Value</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Values Disagree</td>
<td>Adjudicate</td>
<td></td>
</tr>
</tbody>
</table>
**eTable 2.** Pre-certification and certification scores of Trained Readers (TRs) in the e-ROP study

<table>
<thead>
<tr>
<th>Trained Reader</th>
<th>Plus Disease Zone 1 ROP</th>
<th>Any Stage 3/4/5 Referral Warranted ROP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre Cert</td>
<td>Cert</td>
</tr>
<tr>
<td>1</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>3</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Cert = Certification. Gradings were compared with the “Standard Criterion” which was based on the consensus score among 3 ROP specialist expert graders, the Reading Center Director and the StudyChair. Note: Certification values are for the 3 TRs who completed the study.
**eFigure.** Study flow from enrollment through grading.

<table>
<thead>
<tr>
<th>Enrollment</th>
<th>1285 infants enrolled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic Exam</td>
<td>1257 infants had 4263 diagnostic exams</td>
</tr>
<tr>
<td></td>
<td>244 infants had RW-ROP</td>
</tr>
<tr>
<td></td>
<td>1013 infants did not have RW-ROP</td>
</tr>
<tr>
<td>Imaging</td>
<td>242 infants with RW-ROP imaged</td>
</tr>
<tr>
<td></td>
<td>999 infants without RW-ROP imaged</td>
</tr>
<tr>
<td></td>
<td>All 242 infants with RW-ROP image sets selected</td>
</tr>
<tr>
<td></td>
<td>613 infants without RW-ROP had image sets selected</td>
</tr>
<tr>
<td>Grading</td>
<td>1759 image sets From 454 eyes with RW-ROP selected</td>
</tr>
<tr>
<td></td>
<td>150 image sets From 30 eyes without RW-ROP selected</td>
</tr>
<tr>
<td></td>
<td>3611 image sets From 1226 eyes without RW-ROP selected</td>
</tr>
<tr>
<td></td>
<td>5250 image sets From 855 eyes of 1710 infants graded</td>
</tr>
</tbody>
</table>
**eAppendix 1. Grading process.**

Each image set was assigned to a TR reading queue for initial grading. Queues were prioritized by visit date, with earlier visit dates given priority for grading. The TRs logged in using a unique user ID and password each day to check for and access image sets available for reading. Image sets were randomly assigned to specific TR pairs to distribute the workload based on reader capacity. Image sets for each eye were graded independently from its fellow eye. TRs graded independently and were masked from all clinical data and each other’s results. Once an image set was assigned to a TR, it was unavailable for grading to other TRs until the initial image set grading was finalized. Once finalized the image set was released and assigned to another TR’s reading queue. Quality control image sets were selected using weighted sampling and assigned by the Data Manager. These quality control image sets were randomly mixed in with the initial grading assignments. At any given time the TR reading queues could include image sets for initial grading and quality assurance image sets.

After image set grading was finalized by two TRs, the grading responses were automatically compared. If the responses agreed, a final reading record was generated. If there were discrepancies, the software analyzed the types of discrepancies and created a final, adjudicated reading record (eTable1). If discrepancies were above a pre-determined threshold, the software automatically assigned that image set to the RC Director Adjudication Queue for adjudication of the discrepant fields. The software then used the RC Director’s responses to the discrepant fields to generate a final adjudicated reading record. If the RC Director needed additional review of discrepancies by the Study Chair, these image sets were added to a Study Chair Adjudication Queue for adjudication of discrepancies. The software used these responses to generate a final adjudicated reading record. Results finalized from adjudicated queues were considered to be the final grading for that image set and were used in the statistical analysis.
eAppendix 2. Grading image quality and morphological features of ROP.

Image Quality:
The quality of each of the six images in an image set was evaluated as submitted, without using image enhancement to alter contrast, brightness, red-free filtering or magnification. Missing images from each set were noted. Image quality was determined based on focus, clarity and field definition. Pupil image quality was determined based on the ability to determine pupil diameter. The quality of each retinal image’s focus and clarity were subjectively categorized as good, fair or poor while quality of field definition was based on the presence and placement of the optic disc in each image, in relation to its expected location.

Morphologic Features of ROP:
The presence of three key morphologic features was evaluated by TRs, including: 1) the degree of vascular dilation and tortuosity in the center of the posterior pole (normal, pre-plus disease, plus disease); the lowest zone where ROP was located (zone I, II or III or none); and 3) the highest stage of retinopathy seen (none, stages 1-5).

Vascular dilation and tortuosity: TR’s followed a specific grading sequence for evaluating plus disease by first assessing all five retinal images and then evaluating the tortuosity and dilation of vessels within the circular area of the disc center image. A circular area of the retina with a radius of 3.5 disc diameters from the center of the disc was used to identify dilation and tortuosity of arterioles and venules with grading assigned for each quadrant. Plus disease was defined as severe dilation and tortuosity of arterioles and venules that was equal to or greater than that of the referenced ICROP (International Classification of Retinopathy of Prematurity) images in in two or more quadrants. Pre-plus disease was defined as vascular dilation or tortuosity that was insufficient for the diagnosis of plus disease. Posterior pole vessels that did not demonstrate dilation or tortuosity were graded as normal.
The morphological features related to the severity of ROP were evaluated using all of the five retinal images of an image set. A demarcation line (Stage 1 ROP) was defined as a thin sharp white line between the vascularized and non-vascular retina. A ridge (Stage 2 ROP) was wider than a line and white or yellowish in color. A shadow-like appearance in a ridge suggested elevation. Specific morphological lesions such as a “popcorn” lesion or a “sinusoidal” lesion were considered variants of a ridge unless there was vasculature associated with the lesion. Popcorn lesions were white or yellow areas less than 1 disc area that were posterior to but not connected to the ridge and obscured retinal details beneath them. Sinusoidal lesions were oblong with rounded edges located posterior to the ridge and were larger than one disc area.

Extraretinal fibrovascular proliferation (EFP) (Stage 3 ROP) described proliferative tissue that extended from the ridge and gave the posterior aspect of the ridge a thicker and more irregular appearance. EFP typically had a pink hue due to vascularization, although a whitish appearance was seen in areas acquiring a fibrotic component. Flat neovascularization (Stage 3 ROP) was a lacy reddish network of blood vessels seen at the junction between vascularized and non-vascular retina that appeared as flat, wide areas that extended circumferentially and almost always associated with plus disease. Retinal detachment (Stage 4 and 5 ROP) was defined as elevation and separation of the retina from the underlying retinal pigment epithelium.

Zones of vascularization were also determined using all five retinal images: Zone I was defined as the area of retina demarcated by a circle with a radius that was twice the distance from the center of the optic disc to the foveal center. Since the fovea is poorly developed in most premature infants, locating the foveal center was a critical step for defining the extent of zone I. When present, five cues, used in combination, helped TRs determine the center of the fovea: 1) using the curvature of the macular temporal
arcades, extrapolating a circle and locating the center of the circle as the fovea; 2) using the light reflex from the clivus of the macula (the thickened slope of retina in the parafovea, circumferentially located about 500 microns from the center of the fovea); 3) identifying the area of increased pigmentation (located in the retinal pigment epithelial layer) circumferentially located around the center of the fovea; 4) luteal or yellowish pigment in the center of the macula; 5) identifying the avascular foveal zone in the center of the macula. Retinal areas beyond zone I but within a circle with a diameter from disc center to nasal ora serratta were considered to be in zone II. The remaining temporal crescent of retina not included in zone II was defined as zone III. This area was unlikely to be consistently visualized in the images due to technical limitations of the camera’s optical system but was included in the grading in case findings in zone III were present.