

Supplementary Online Content

Kim SJ, Campbell JP, Kalpathy-Cramer J, et al; Imaging and Informatics in Retinopathy of Prematurity (i-ROP) Research Consortium. Accuracy and reliability of eye-based vs quadrant-based diagnosis of plus disease in retinopathy of prematurity. *JAMA Ophthalmol*. Published online April 26, 2018. doi:10.1001/jamaophthalmol.2018.1195

eTable 1. Intragrader reliability between eye-based and quadrant-based diagnosis

eTable 2. Intragrader agreement between individual quadrant assessment and eye-based approach for plus disease diagnosis

eTable 3. Agreement between individual quadrant assessment and reference standard diagnosis

eTable 4. Intragrader agreement between individual quadrant assessment and eye-based approach for plus disease diagnosis in 31 images with reference standard diagnosis of plus disease

eTable 5. Agreement between individual quadrant assessment and reference standard diagnosis in 31 images with reference standard diagnosis of plus disease

eTable 6. Diagnostic accuracy of eye-based vs quadrant-based plus disease diagnosis compared with a consensus reference standard diagnosis

This supplementary material has been provided by the authors to give readers additional information about their work.

eTable 1. Intragrader reliability between eye-based and quadrant-based diagnosis

	Eye-based diagnosis	Quadrant-based diagnosis, Number of eyes (%)		Absolute agreement (95% CI)	Cohen's kappa (95% CI)	P value by McNemar test	Probability of under-diagnosis by quadrant-based diagnosis (95% CI by binomial test) ^a
		No plus	Plus				
Grader 1	No plus	166 (84.3)	1 (0.5)	0.89 (0.85 - 0.94)	0.44 (0.22 - 0.67)	<0.001	0.95 (0.76 – 1.00)
	Plus	20 (10.2)	10 (5.1)				
Grader 2	No plus	169 (85.8)	1 (0.5)	0.94 (0.91 - 0.97)	0.69 (0.53 - 0.86)	0.006	0.92 (0.62 – 1.00)
	Plus	11 (5.6)	16 (8.1)				
Grader 3	No plus	154 (78.2)	0 (0)	0.83 (0.78 - 0.88)	0.32 (0.11 - 0.53)	<0.001	1.00 (0.90 – 1.00)
	Plus	33 (16.8)	10 (5.1)				
Grader 4	No plus	152 (77.2)	5 (2.5)	0.89 (0.84 - 0.93)	0.61 (0.46 - 0.76)	0.017	0.77 (0.55 – 0.92)
	Plus	17 (8.6)	23 (11.7)				
Grader 5	No plus	149 (75.6)	13 (6.6)	0.92 (0.88 - 0.96)	0.75 (0.63 - 0.87)	0.021	0.19 (0.04 - 0.46)
	Plus	3 (1.5)	32 (16.2)				
Grader 6	No plus	157 (79.7)	12 (6.1)	0.92 (0.88 - 0.96)	0.70 (0.56 - 0.84)	0.077	0.25 (0.07 – 0.52)
	Plus	4 (2.0)	24 (12.2)				

CI=confidence interval.

^a Probability of “plus” disease by eye-based diagnosis but “no plus” by quadrant-based diagnosis. Probability of > 0.5 means under-diagnosis of plus disease by quadrant-based diagnosis.

Six expert graders evaluated 197 posterior retinal images (31 with reference standard diagnosis of plus disease, 166 with reference standard of pre-plus or normal).

eTable 2. Intragrader agreement between individual quadrant assessment and eye-based approach for plus disease diagnosis

	Superonasal quadrant		Inferonasal quadrant		Superotemporal quadrant		Inferotemporal quadrant	
	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)
Grader 1	0.88 (0.83 - 0.92)	0.33 (0.07 - 0.58)	0.87 (0.83 - 0.92)	0.31 (0.06 - 0.56)	0.89 (0.84 - 0.93)	0.38 (0.14 - 0.63)	0.89 (0.84 - 0.93)	0.43 (0.20 - 0.65)
Grader 2	0.93 (0.89 - 0.96)	0.62 (0.42 - 0.81)	0.90 (0.86 - 0.94)	0.40 (0.16 - 0.65)	0.92 (0.88 - 0.96)	0.56 (0.35 - 0.77)	0.92 (0.88 - 0.96)	0.58 (0.38 - 0.78)
Grader 3	0.82 (0.76 - 0.87)	0.23 (0.01 - 0.46)	0.81 (0.75 - 0.86)	0.17 (-0.07 - 0.41)	0.84 (0.79 - 0.89)	0.38 (0.18 - 0.58)	0.83 (0.77 - 0.88)	0.31 (0.10 - 0.52)
Grader 4	0.87 (0.83 - 0.92)	0.55 (0.39 - 0.72)	0.86 (0.81 - 0.91)	0.45 (0.26 - 0.64)	0.85 (0.80 - 0.90)	0.44 (0.25 - 0.62)	0.90 (0.86 - 0.94)	0.67 (0.53 - 0.81)
Grader 5	0.90 (0.86 - 0.94)	0.68 (0.55 - 0.81)	0.88 (0.84 - 0.93)	0.60 (0.45 - 0.76)	0.91 (0.87 - 0.95)	0.73 (0.60 - 0.85)	0.90 (0.86 - 0.94)	0.69 (0.56 - 0.82)
Grader 6	0.90 (0.86 - 0.94)	0.58 (0.41 - 0.76)	0.92 (0.88 - 0.96)	0.62 (0.44 - 0.80)	0.90 (0.86 - 0.94)	0.63 (0.47 - 0.78)	0.89 (0.85 - 0.94)	0.63 (0.49 - 0.78)
Mean ± SD	0.88 ± 0.04	0.50 ± 0.18	0.87 ± 0.04	0.43 ± 0.17	0.89 ± 0.03	0.52 ± 0.14	0.89 ± 0.03	0.55 ± 0.15
Range	0.82 – 0.93	0.23 – 0.68	0.81 – 0.92	0.17 – 0.62	0.84 – 0.92	0.38 – 0.73	0.83 – 0.92	0.31 – 0.69

CI=confidence interval.

Six expert graders evaluated 197 posterior retinal images (31 with reference standard diagnosis of plus disease, 166 with reference standard diagnosis of pre-plus or normal).

eTable 3. Agreement between individual quadrant assessment and reference standard diagnosis

	Superonasal quadrant		Inferonasal quadrant		Superotemporal quadrant		Inferotemporal quadrant	
	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)
Grader 1	0.88 (0.84 - 0.93)	0.37 (0.13 - 0.61)	0.89 (0.84 - 0.93)	0.41 (0.18 - 0.64)	0.88 (0.84 - 0.93)	0.37 (0.13 - 0.61)	0.88 (0.84 - 0.93)	0.41 (0.19 - 0.64)
Grader 2	0.91 (0.87 - 0.95)	0.55 (0.35 - 0.75)	0.89 (0.84 - 0.93)	0.41 (0.18 - 0.64)	0.91 (0.87 - 0.95)	0.55 (0.35 - 0.75)	0.91 (0.87 - 0.95)	0.56 (0.37 - 0.76)
Grader 3	0.88 (0.83 - 0.92)	0.33 (0.08 - 0.58)	0.87 (0.82 - 0.92)	0.25 (-0.03 - 0.52)	0.90 (0.86 - 0.94)	0.52 (0.31 - 0.72)	0.90 (0.86 - 0.94)	0.48 (0.27 - 0.70)
Grader 4	0.92 (0.88 - 0.96)	0.68 (0.53 - 0.83)	0.87 (0.83 - 0.92)	0.42 (0.21 - 0.63)	0.92 (0.88 - 0.96)	0.64 (0.47 - 0.81)	0.93 (0.89 - 0.96)	0.72 (0.59 - 0.86)
Grader 5	0.90 (0.86 - 0.94)	0.67 (0.53 - 0.81)	0.88 (0.84 - 0.93)	0.59 (0.43 - 0.75)	0.91 (0.87 - 0.95)	0.72 (0.59 - 0.84)	0.91 (0.87 - 0.95)	0.72 (0.59 - 0.84)
Grader 6	0.89 (0.85 - 0.94)	0.58 (0.41 - 0.75)	0.92 (0.89 - 0.96)	0.66 (0.50 - 0.83)	0.92 (0.89 - 0.96)	0.73 (0.60 - 0.86)	0.91 (0.87 - 0.95)	0.70 (0.56 - 0.83)
Mean ± SD	0.90 ± 0.02	0.53 ± 0.15	0.89 ± 0.02	0.46 ± 0.15	0.91 ± 0.01	0.59 ± 0.14	0.91 ± 0.02	0.60 ± 0.13
Range	0.88 - 0.92	0.33 - 0.68	0.87 - 0.92	0.25 - 0.66	0.88 - 0.92	0.37 - 0.73	0.88 - 0.93	0.41 - 0.72

CI=confidence interval.

Six expert graders evaluated 197 posterior retinal images (31 with reference standard diagnosis of plus disease, 166 with reference standard diagnosis of pre-plus or normal).

eTable 4. Intragrader agreement between individual quadrant assessment and eye-based approach for plus disease diagnosis in 31 images with reference standard diagnosis of plus disease

	Absolute agreement (95% CI)			
	Superonasal quadrant	Inferonasal quadrant	Superotemporal quadrant	Inferotemporal quadrant
Grader 1	0.39 (0.32-0.46)	0.35 (0.29-0.42)	0.45 (0.38-0.52)	0.45 (0.38-0.52)
Grader 2	0.55 (0.48-0.62)	0.35 (0.29-0.42)	0.48 (0.41-0.55)	0.52 (0.45-0.59)
Grader 3	0.35 (0.29-0.42)	0.29 (0.23-0.35)	0.52 (0.45-0.59)	0.42 (0.35-0.49)
Grader 4	0.68 (0.61-0.74)	0.45 (0.38-0.52)	0.48 (0.41-0.55)	0.68 (0.61-0.74)
Grader 5	0.90 (0.86-0.94)	0.74 (0.68-0.80)	0.94 (0.90-0.97)	0.87 (0.82-0.92)
Grader 6	0.68 (0.61-0.74)	0.58 (0.51-0.65)	0.71 (0.65-0.77)	0.81 (0.75-0.86)
Mean ± SD	0.59 ± 0.21	0.46 ± 0.29	0.60 ± 0.19	0.62 ± 0.19
Range	0.35 - 0.90	0.29 - 0.74	0.45 - 0.94	0.42 - 0.87

CI=confidence interval.

eTable 5. Agreement between individual quadrant assessment and reference standard diagnosis in 31 images with reference standard diagnosis of plus disease

	Absolute agreement (95% CI)			
	Superonasal quadrant	Inferonasal quadrant	Superotemporal quadrant	Inferotemporal quadrant
Grader 1	0.26 (0.20-0.32)	0.29 (0.23-0.35)	0.26 (0.20-0.32)	0.32 (0.26-0.39)
Grader 2	0.42 (0.35-0.49)	0.29 (0.23-0.35)	0.42 (0.35-0.49)	0.45 (0.38-0.52)
Grader 3	0.23 (0.17-0.28)	0.16 (0.11-0.21)	0.39 (0.32-0.46)	0.35 (0.29-0.42)
Grader 4	0.68 (0.61-0.74)	0.39 (0.32-0.46)	0.55 (0.48-0.62)	0.74 (0.68-0.80)
Grader 5	0.87 (0.82-0.92)	0.71 (0.65-0.77)	0.90 (0.86-0.94)	0.90 (0.86-0.94)
Grader 6	0.61 (0.54-0.68)	0.58 (0.51-0.65)	0.84 (0.79-0.89)	0.87 (0.82-0.92)
Mean ± SD	0.51 ± 0.25	0.40 ± 0.20	0.56 ± 0.26	0.61 ± 0.26
Range	0.23 - 0.87	0.16 - 0.71	0.26 - 0.90	0.32 - 0.90

CI=confidence interval.

eTable 6. Diagnostic accuracy of eye-based vs quadrant-based plus disease diagnosis compared with a consensus reference standard diagnosis

	Eye-based diagnosis		Quadrant-based diagnosis	
	Absolute agreement (95% CI)	Cohen's kappa (95% CI)	Absolute agreement (95% CI)	Cohen's kappa (95% CI)
Grader 1	0.94 (0.91 - 0.98)	0.79 (0.66 - 0.91)	0.90 (0.86 - 0.94)	0.48 (0.27 - 0.70)
Grader 2	0.98 (0.96 - 1.00) ^a	0.92 (0.84 - 1.00) ^a	0.93 (0.89 - 0.96) ^a	0.67 (0.51 - 0.84) ^a
Grader 3	0.90 (0.86 - 0.94)	0.67 (0.53 - 0.81)	0.89 (0.85 - 0.94)	0.45 (0.22 - 0.67)
Grader 4	0.93 (0.90 - 0.97)	0.78 (0.66 - 0.89)	0.93 (0.90 - 0.97)	0.74 (0.60 - 0.88)
Grader 5	0.97 (0.95 - 0.99) ^a	0.89 (0.80 - 0.98)	0.91 (0.87 - 0.95) ^a	0.71 (0.58 - 0.84)
Grader 6	0.97 (0.95 - 1.00)	0.90 (0.81 - 0.99)	0.93 (0.90 - 0.97)	0.77 (0.64 - 0.89)
Mean ± SD	0.95 ± 0.03	0.83 ± 0.10	0.92 ± 0.02	0.64 ± 0.14
Range	0.90 - 0.97	0.67 - 0.92	0.89 - 0.93	0.45 - 0.77

CI=confidence interval.

^a No overlaps in 95% CIs between the two diagnostic approaches.

Six expert graders evaluated 197 posterior retinal images (31 with RSD of plus disease, 166 with RSD of pre-plus or normal).