

2014 Neurocritical Care Society Meeting
Poster Presentation

**Interrater Reliability
of Pupillary Assessments
Among
Physicians and Nurses**

Department of Neurology and Neurotherapeutics and Neurological Surgery
University of Texas Southwestern Medical Center

2014 Neurocritical Care Society Meeting, Seattle, Washington

Interrater Reliability of Pupillary Assessments Among Physicians and Nurses

UTSouthwestern
Medical Center

Sonja E. Stutzman¹, DaiWai M. Olson^{1,2,3}, Ciji Saju³,
Margaret Wilson³, Venkatesh Aiyagari^{1,2}



Department of Neurology and Neurotherapeutics¹ and Neurological Surgery², University Hospital³, UT Southwestern Medical Center

Introduction

Observational assessment of pupil size, shape, and reactivity is a fundamental aspect of the neurological examination. In neurological and neurosurgical patients, examination of the pupils is extremely helpful in localizing the lesion, characterizing a differential diagnosis, and following the course of the disease process. However, interrater reliability of these findings has not been established.

Specific Aims

This study seeks to:

1. Determine the interrater reliability of pupillary assessments between two independent observers examining the pupil at the same time.
2. Determine if there is a greater correlation between observers matched by profession (nursing versus medical background)

Background

Serial examination of the pupil is a critical component of the bedside neurological examination. Changes in the pupillary size, shape and/or reactivity can be an early sign of neurological deterioration, increased intracranial pressure, and worsening midline shift. In addition, pupillary findings may also be used to monitor response to therapy. Despite the importance of pupillary assessments, limited previous research has shown that there may be variability in pupillary assessments.

Methods

We examined interrater reliability of bedside pupillary exams as part of a large prospective study. 127 patients with neurological or neurosurgical diagnoses were consented. These patients were observed by staff (RNs and MDs) who were expected to perform a pupillary exam as a normal part of their patient care routine. Two independent observers performed a pupillary exam simultaneously or within 5 minutes of each other. Observers were instructed to perform the exam as they would under normal conditions using the equipment (penlight or flashlight) that they would use if they were typically performing the exam.

Interrater Reliability of Pupillary Assessments Among Physicians and Nurses

Demographics

	All Patients (n=127)	Surgical Procedure (n=57)	Not Surgical (n=70)
Mean Age	55.1 (sd=15.6)	54.8 (sd=15.7)	55.4 (sd=15.6)
% Male	72 (56.7%)	34 (59.6 %)	38 (54.3%)
% Caucasian	103 (81.1%)	47 (82.5%)	56 (80.0%)
<u>Primary diagnosis</u>			
# Ischemic stroke	43 (33.9%)	7 (12.3%)	36 (51.4%)
# Hem. stroke (SAH/ICH)	40 (31.5%)	20 (35.1%)	20 (28.6%)
# Neoplasm	23 (18.1%)	18 (31.6%)	5 (7.1%)
# Other	21 (16.5%)	12 (21.0%)	9 (12.9%)
# Vision or ocular history	9 (7.1%)	5 (8.8%)	4 (5.7%)

Results

- 2332 (1166 paired) observational assessments pupils were completed on 127 consented patients
- There was composite agreement in all three parameters: size (within 1 mm), shape, & reactivity for 463 (39.7%) OS; 470 (40.3%) OD; and 328 (28.1%) OU.
- Interrater reliability Kappa values and % agreement for size (within 1mm) for OS=0.54 (78%), OD=0.54 (79.8%), and OU=0.54 (78.9%).
- Kappa and agreement for shape for OS=0.68 (94.3%), OD=0.06 (97.6%), and OU=0.62 (96%).
- Kappa and agreement for pupil reactivity for OS=0.48 (80%), OD=0.30 (78%), and OU=0.40 (79%).
- Only 49.5% (94/190) pupils assessed as fixed (non-reactive) by one assessor was also scored as fixed by the second assessor; [OS = 73/131 (55.7%), and OD = 21/59 (35.6%)].



Conclusion

These data suggest that there is a surprisingly large degree of disagreement between trained observers when assessing and scoring pupillary findings. Furthermore, for 96 of 190 (51.5%) assessments, there was no agreement on the absence of pupillary reactivity. Interrater reliability was not significantly different within vs between observers. Given the importance of pupillary reactivity in patients with neurological injury, these findings may have significant clinical implications. Future research is needed to standardize pupillary assessments. One mechanism for potential standardization of pupillary assessments is through the use of the Pupillometer.

Interrater Reliability of Pupillary Assessments Among Physicians and Nurses

Results

		Left Eye		Right Eye		Both Eyes	
		Kappa	% agree	Kappa	% agree	Kappa	% agree
Entire Cohort (N=1,166)	Size > 1mm	0.54 (.49-.59)	78.0%	0.54 (.49-.59)	79.8%	0.54 (.50-.58)	78.9%
	Equal in size	-	-	-	-	0.60 (.55-.64)	80.8%
	Shape	0.68 (.61-.76)	94.3%	0.06 (-.07-.19)	97.6%	0.62 (.55-.69)	96.0%
	Reactivity (brisk, slow, fixed)	0.48 (.43-.53)	80.0%	0.30 (.23-.36)	78.0%	0.40 (.36-.44)	79.0%
	Reactive vs not	0.69 (.62-.77)	95.0%	0.51 (.37-.64)	96.7%	0.64 (.58-.71)	95.9%
Only Nurses (n=952)	Size > 1mm	0.54 (.49-.60)	78.6%	0.52 (.46-.58)	79.7%	0.53 (.49-.57)	79.1%
	Equal in size	-	-	-	-	0.60 (.54-.65)	80.8%
	Shape	0.70 (.62-.78)	94.7%	.01 (.005-.01)	97.8%	0.64 (.56-.72)	96.3%
	Reactivity (brisk, slow, fixed)	0.49 (.43-.55)	80.3%	0.33 (.26-.40)	78.7%	0.42 (.38-.47)	79.5%
	Reactive vs not	0.69 (.61-.78)	95.2%	0.59 (.45-.73)	97.1%	0.67 (.59-.74)	96.2%
Only Physician (n=33)	Size > 1mm	0.49 (.20-.78)	74.3%	0.77 (.55-.98)	88.6%	0.63 (.45-.81)	81.4%
	Equal in size	-	-	-	-	0.70 (.42-.97)	87.8%
	Shape	0.64 (.19-1.0)	94.2%	DNC	96.9%	0.55 (.11-.99)	95.6%
	Reactivity (brisk, slow, fixed)	0.73 (.45-1.0)	91.4%	DNC	78.8%	0.54 (.30-.79)	85.3%
	Reactive vs not	0.79 (.38-1.0)	97.1%	DNC	93.9%	0.55 (.09-1.0)	95.6%
One RN and One MD (n=179)	Size > 1mm	0.51 (.39-.64)	76.0%	0.56 (.44-.68)	78.8%	0.54 (.45-.63)	77.4%
	Equal in size	-	-	-	-	0.57 (.44-.69)	79.7%
	Shape	0.61 (.42-.79)	92.1%	0.24 (.14-.63)	96.6%	0.55 (.38-.73)	94.4%
	Reactivity (brisk, slow, fixed)	0.38 (.24-.51)	75.8%	0.07 (-.07-.20)	74.0%	0.25 (.15-.36)	74.9%
	Reactive vs not	0.67 (.49-.85)	93.8%	-0.02 (-0.04-0.002)	94.9%	0.54 (.36-.72)	94.4%

The authors of this study would like to thank Weidan Zhao and the many clinicians that contributed to the pupillary assessments. NeuroOptics, Inc. provide Pupillometers and headrests as part of a larger study from which these data were drawn.