Pupillometer Inter-Device Reliability Assessment

Weidan Zhao BA, Sonja Stutzman PhD, Daiwai Olson RN PhD, Ciji Saju RN, Margaret Wilson RN, Venkatesh Aiyagari MBBS DM FAHA
The University of Texas Southwestern Medical Center, Dallas, Texas

Abstract

Objective: To explore the inter-device reliability of NPi™-100 pupillometers (NeuroOptics, Inc.).

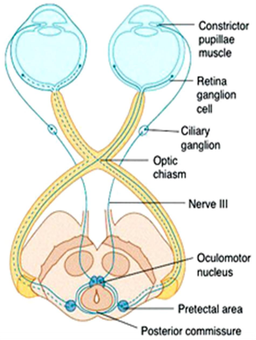
Background: Current evidence suggests that the traditional bedside examination of the pupil with a flashlight has limited inter-device reliability. Automated pupillometers were developed to provide an objective scoring of pupil reactivity. However, there are no data examining inter-device reliability of automated pupil assessments.

Design/Methods: 210 paired pupillometer measurements were obtained by 33 practitioners from 20 patients at risk for cerebral edema.

Results: There was no statistically significant difference between the mean maximum pupil size at rest, the minimum pupil size during light stimulation, and the mean pupil reactivity, for both eyes, when assessed by two raters using two different pupillometers. In addition, Cohen's Kappa assessments of pupil size and reactivity revealed an almost perfect agreement between the two pupillometers for the maximum pupil size, the minimum pupil size, and for pupil reactivity for both eyes.

Conclusions: There is a high inter-device reliability of automated pupillary assessments by two practitioners examining the same patient using different NPi™-100 pupillometers.

Pupillary Reflex



The pupillary reflex involves CN II and III. Under normal conditions, the afferent signal is carried by CN II to the pretectal area and then activates the Edinger-Westphal nucleus of CN III. The efferent signal is carried along CN III and synapses with the ciliary postganglionic nerves to innervate the ciliary body and iris, leading to pupil constriction in response to light stimulation.

When both the left and right afferent and efferent pathways are functionally intact, both pupils should be equal in size and constriction when stimulated with the same light source. Differences in either measurements are often linked to neurological conditions, such as ischemic and hemorrhagic stroke, traumatic brain injury, neoplasm, and CNS infection. These conditions can cause cerebral edema and increased intracranial pressure (ICP), which in turn can lead to compression of CN III due to uncal herniation, or distortion of the midbrain from shift of the midline structures. [1-3]

It is widely recognized that early identification of any change in the neurologic status is central to initiating therapies aimed to ensure that ICP remains within normal limits. The serial neurologic exam, performed across the globe, is designed to detect even minute changes in neurological function. A hallmark of this exam is the periodic assessment of the pupillary light reflex.

Traditional Pupillary Exam

	Right Eye	Left Eye
Pupil Size (mm)		
Pupil Shape	Round Irregular	Round Irregular
Pupil Reactivity	Brisk Sluggish Non-reactive	Brisk Sluggish Non-reactive

Subjective assessment of the size and shape of the pupil prior to the manual application of a light source and the speed of pupil reactivity when exposed to light.

Studies suggest that the manual examination may have limited inter-device reliability, as the examiners are of different levels of skill and training and are allowed to use a variety of non-standardized flashlights, penlights, and handheld light sources.[4-5]

We have earlier shown that among trained observers, when the pupil is considered to be non-ractive by one examiner, another examiner agrees only 49.7% of the time. Therefore, in an attempt to standardize the pupillary exam, automated pupillary assessment technology has been developed that is able to provide an objective measurement of the initial and final size of the pupil and to grade the speed of pupil contraction in response to a light stimulus.

NPi™-100 Pupillometers

Portable
Battery-operated
LCD screen
Digital Video Camera
Automated light source

Measures, records, and analyzes both the pupil size and reactivity.

The pupil light response is calculated and a Neurological Pupil index (NPi™) is reported immediately.

The NPi™ is based on a multi-dimensional normative model of pupil reaction to light, and measurements are rated on a scale: 0-5.

- A score ≥ 3 indicates that the pupil measurement is within the boundaries of normal pupil behavior.
- A score < 3 suggests that the reflex is abnormal.
- A score of 0 suggests an absence of constriction.

Additionally, a difference between the left and right NPi™ scores (even if both are above 3) is a sign of pupil abnormalities.

Participants & Method

210 paired pupillometer measurements obtained between July and August of 2014 by 33 practitioners: 28 RNs, 2 MDs, 1 NP, 1 medical student, and 1 PhD research coordinator.

20 patients were consented to the study from the Neurocritical Care Unit at Zale Lipshy Hospital, Acute Stroke Unit at Zale Lipshy Hospital, and the Surgical Intensive Care Unit of Parkland Memorial Hospital located in Dallas, TX.

Practitioners were provided with a five-minute training session of how to use the pupillometers. The pupillometer device was set to a research mode, blinding the users to the results. Paired pupillometer assessments were completed within a five-minute period and consisted of two separate assessments by two different practitioners. To assess inter-device reliability, each staff member used a different pupillometer device on the same subject. Four different pupillometers were used during the study.

Results

	Left Eye		P-Value	Right Eye		P-Value
	Device 1	Device 2		Device 1	Device 2	
Max Size (mm)	3.8 (1.1)	4.0 (1.6)	0.27	3.6 (1.1)	3.8 (1.1)	0.74
Min Size (mm)	2.8 (1.2)	3.0 (1.5)	0.64	2.6 (1.0)	2.6 (0.6)	0.44
NPi™	3.9 (1.2)	3.9 (1.4)	0.36	4.2 (1.0)	4.3 (0.8)	0.82

	Left Eye			Right Eye		
	Max Size	Min Size	NPi™	Max Size	Min Size	NPi™
Simple Kappa	0.97	0.96	1.00	0.91	0.98	0.90
Weighted Kappa	0.97	0.93		0.90	0.96	

Cohen's Kappa (k) assessments of size and reactivity were examined for agreement using interpretation described by Viera and Garrett[6] based on work of Landis and Koch[7].

Discussion

This study is one of its first in assessing the inter-device reliability of pupillometers, and its results show that there is an extremely high inter-device reliability.

The high correlation between different pupillometer measurements of the NPI™ and pupil sizes for both the maximum and minimum and for both the right and left eyes attests the devices' reliability in measuring pupil and reactivity under both resting conditions and during stimulations. This suggests the possibility that the pupillometers are reliable for usage in various situations: outpatient clinics, routine checkups, emergency assessments, or chronic monitoring, etc. In addition, different observers, even using different devices, are likely to receive very similar findings. This is a significant improvement compared to the traditional method of pupillary examination.

The pupillometer wasn't able to produce a pupillary assessment of both size and reactivity in about 4% of patients. It's possible that there is a learning curve to using the device so perhaps over time, there will be fewer dropped values as examiners become more experienced with the device. In addition, as with current manual examinations, pupillary assessment with the pupillometer on agitated or confused patients was challenging. And using the device to assess patients with slit pupils, irregularly shaped pupils, and patients with significant cataract was especially troublesome.

Conclusion and Future Work

There is a high inter-device reliability between two pupillometers when used on the same subject. The data provide sufficient support to conclude that when an automated pupillary reading is obtained, the results can be interpreted with confidence that any other provider would get the same or similar results using another automated device by the same manufacturer. These findings may support the routine use of the pupillometer in patients instead of relying on the traditional method. The reasons for an inability to obtain complete pupillometer readings for 4.3% of paired observations is unclear and requires additional study.

Acknowledgements

We thank UT Southwestern Medical Center and Parkland Memorial Hospital clinicians for their assistance. Funding for this project was provided by the UT Southwestern Summer Medical Student Research Program. NPi™-100 Pupillometers were provided by NeurOptics Inc.

References

1. Loewenfeld IE. (1993). The pupil. Anatomy, Physiology, and Clinical Applications. Wayne State University Press, Detroit.
2. Geobert HW Jr. (1970). Head injury associated with a dilated pupil. *Surgery Clinics of North America*. 50(2), 427-432.
3. Manley GT, Larson MD. (2002). Infrared pupillometry during uncal herniation. *Journal of Neurosurgical Anesthesiology*, 14(3), 223-228.
4. A Clark, T. N. S. C., B Gregson, P N A Hooker, I R Chambers. (2006). Variability in pupil size estimation. *Emergency Medicine Journal, 23*(6), 440-441. doi: 10.1136/emj.2005.030247
5. Wilson SF, Amling JK, Floyd SD, McNair ND. (1988). Determining interrater reliability of nurses' assessments of pupillary size and reaction. *Journal of Neuroscience Nursing*, 20(3). 189-192
6. Viera AJ, Garrett JM. (2005). Understanding interobserver agreement: The kappa statistic. *Family medicine*. 37, 360-363
7. Landis JR, Koch GG. (1977). The measurement of observer agreement for categorical data. *Biometrics*.33, 159-174