Assessment of intra-interobserver reliability of the sonographic optic nerve sheath diameter measurement

Tuba Cimilli Ozturk*, Hasan Demir, Rasim Yorulmaz, Serdar Ozdemir, Gokhan Isat, Ozge Ecme Onur

Department of Emergency Medicine, Fatih Sultan Mehmet Education and Research Hospital, Istanbul, Turkey

Received 4 December 2014; accepted 27 April 2015
Available online 3 July 2015

Abstract  Diagnosis and measuring the level of increase in intracranial pressure (ICP) is critical, especially for the management of trauma patients in the emergency department and intensive care unit. However, measurements are operator-dependent as in all of the sonographic diagnoses. The aim of this study is to assess the operator variations in the measurement of optic nerve sheath diameter (ONSD). There were four emergency medicine specialists involved in the study. Each had at least 1 year of experience of ultrasound scans and performed at least 25 prior ocular scans examining the ONSD. Two measurements were made 1 week apart from both axial and longitudinal planes. Sixty healthy adults were involved in the study and every investigator obtained four measurements from each. Intra-interobserver reliabilities were tested. The investigators performed 60 ocular ultrasounds on individual healthy adults and obtained two measurements in axial and longitudinal planes 1 week apart. Therefore, 960 measurements were analyzed. The levels of compatibilities for most of the measurements were found at acceptable levels statistically. However, it is not possible to say that there was a perfect compatibility among the sonographers according to the previously conducted reliability studies of ultrasound measurements. According to our results, it is hard to say that sonographic measurement of the ONSD is a highly reliable method both in longitudinal and transverse planes.

Copyright © 2015, Kaohsiung Medical University. Published by Elsevier Taiwan LLC. All rights reserved.

Conflicts of interest: All authors declare no conflicts of interest.

* Corresponding author. Fatih Sultan Mehmet Education and Research Hospital, Emergency Medicine Department, E5 Karayolu Uzeri, 34752 Icerenkoy-Atasehir, Istanbul, Turkey.

E-mail address: tcmilliozturk@gmail.com (T. Cimilli Ozturk).

http://dx.doi.org/10.1016/j.kjms.2015.06.004
1607-551X/Copyright © 2015, Kaohsiung Medical University. Published by Elsevier Taiwan LLC. All rights reserved.
Introduction

Intracranial pressure (ICP) elevation is a clinically important phenomenon with high morbidity and mortality, which occurs as a result of any kind of neurologic injury. Diagnosis and measuring the level of increase in the pressure is critical, especially for the management of trauma patients in the emergency department.

Several methods have been used for measuring ICP, such as the radiologic imaging or invasive techniques. The “gold standard” for the diagnosis of elevated ICP is measuring the pressure invasively by intracranial catheters connecting to a pressure transducer [1]. However, the procedure is highly invasive and may result in severe complications such as infection, hemorrhage, and malfunction.

Anatomically the optic nerve is surrounded by cerebrospinal fluid and it is encased in a sheath which is an extension of the dura mater. This uninterrupted connection provides ICP transmission directly into the optic nerve sheath if there is not a cerebrospinal fluid circulation block. Therefore, measuring the changes in the diameter of the optic nerve sheath gives important information about the changes in ICP. The expansion of the sheath was verified to occur immediately in cases of elevated ICP [2].

Measuring the optic nerve sheath diameter (ONSD) in the cranial computed tomography and magnetic resonance imaging sections were found to correlate well with the invasively measured ICP levels in previous studies [3,4]. However, it is obvious that these modalities are not suitable for continuous monitoring of critically ill patients and may not be available in emergency settings. High cost is the other disadvantage of these diagnostic methods.

Sonographic measurement of the ONSD seemed to be a feasible alternative method in the diagnosis of raised ICP, especially in the emergency department and intensive care unit. There are many studies in the literature supporting the use of ultrasound in this field. The results were promising and correlated with the invasively determined ICP levels [5–7]. It has the advantage of being noninvasive, cost-effective, and repeatable. However there has not yet been a consensus on the exact cut-off value for the normal ONSD. This is probably because of the dependence of the technique on the operator’s experiences. Although there are many reliability studies about ultrasound for variable measurements, to our knowledge there is limited number of studies testing the reliability of the sonographic ONSD measurement.

The aim of this study is to assess the operator variations among four ultrasound operators. As we know, this is the first study comparing the four different examiners in the measurement of ONSD.

Methods

This is a prospectively designed study in healthy individuals. Fatih Sultan Mehmet Education and Research Hospital Review Board/Ethics Committee approved the study. There were four emergency medicine specialists involved in the study. The examiners attended the official bedside ultrasound courses organized for all clinicians other than radiologists previously and they were all trained in transorbital sonography before starting the study. Each had at least 1 year experience of ultrasonography scan and performed at least 25 prior ocular scans examining the ONSD. The study group was selected among normal healthy individuals older than 18 years. Two measurements were made 1 week apart from both axial and longitudinal planes. The researchers who were blinded to each other’s results made the measurements in the right eye of the participants in order not to cause any confusion. Sixty healthy individuals were involved in the study and every investigator obtained four measurements from each. We gave codes to the examiners as A, B, C, and D, respectively, in order to ease the interpretation of the results.

The Mindray DC3 gray-scale ultrasonography machine (Shenzhen Mindray Bio-Medical Electronics Co. Ltd., Nanshan, Shenzhen, China) with a 10 MHz linear transducer were used for ultrasonographic measurements. All of the scans were performed in the supine position. The participants were asked to keep their eyes focused on the ceiling as much as possible during the procedure. The transducer was placed perpendicular to the globe both in transverse and longitudinal planes. A transparent tape was applied over the closed eye in order not to discomfort the volunteer, and ultrasound gel was also applied over it (Figure 1). The ONSD was assessed 3 mm behind the papilla as described in the previous studies [8,9] (Figure 2).

Number Cruncher Statistical System (NCSS) 2007 and Power Analysis and Sample Size (PASS) 2008 statistical software (Kaysville, Utah, USA) programs were used for statistical analysis. In addition to the descriptive statistical methods (mean, standard deviation), the paired sample t test was used for the comparison of quantitative data showing normal distribution of changes in the first and second measurements. Cronbach α was used for agreement between three or more observers. The binary compatibilities between the observers, were assessed with the intraclass correlation coefficient [10,11]. The results were...
was a good level of compatibility statistically among the participants. According to the results, an acceptable level of compatibility was determined among the sonographers; Cronbach $\alpha = 0.700$ (95% CI: 0.55–0.81). The mean values were as follows: $5.33 \pm 0.54$ mm, $5.17 \pm 0.69$ mm, $4.79 \pm 0.55$ mm, and $4.24 \pm 0.34$ mm.

Second axial results showed acceptable levels of compatibility with Cronbach $\alpha = 0.683$ (95% CI: 0.53–0.80). The mean measurements of the examiners were $5.37 \pm 0.59$ mm, $5.21 \pm 0.73$ mm, $4.75 \pm 0.52$ mm, and $4.45 \pm 0.34$ mm, respectively.

Finally, the second longitudinal values determined by the participants were also compatible with each other. The mean measurements were $5.34 \pm 0.56$ mm, $5.20 \pm 0.71$ mm, $4.82 \pm 0.57$ mm and $4.27 \pm 0.34$ mm, respectively; Cronbach $\alpha = 0.718$ (95% CI: 0.58–0.82).

As the binary compatibilities of the first axial and longitudinal measurements were interpreted, there were only two incompatibilities observed. Both were found between examiner B and D. Intraclass correlation coefficients of the comparisons were $0.011$ (95% CI: 0.24–0.26) and $0.016$ (95% CI: 0.27–0.24). Similarly, when the second axial and longitudinal measurements were evaluated, there were incompatibilities between examiner B and D. Intraclass correlation coefficients of this comparison were $0.015$ (95% CI: 0.24–0.27) and $0.017$ (95% CI 0.24–0.27). Although the rest of the binary compatibilities seemed to be compatible statistically, they were generally at low to moderate levels. The overall results are summarized in Table 1.

Discussion

Ultrasound is a very useful imaging technique which has many advantages including the bedside availability and the relative ease of performing repeated examinations. Imaging is real-time and free of harmful radiation. There are no documented side effects and discomfort is minimal. However, it is highly operator-dependent. Retrospective review of images provides only limited quality control. There is no scout scan to give a global picture for orientation.

Measurement of ONSD by ultrasound is an easy and convenient method in the diagnosis of raised ICP, especially in the emergency department and intensive care unit. Previously reported results were promising and seemed to correlate well with the invasively determined ICP levels [6,7]. However, there has not been a consensus on the exact cut-off value of ONSD yet. Dubourg and colleagues [12] have been carrying out a project to establish an individual patient-level database from high quality sonographic ONSD measurement studies for the detection of raised ICP. However, the meta-analysis is not reported yet. Normal values reported ranges from 4.0 mm to 5.9 mm in different studies [13–15]. In addition, to our knowledge there are only two studies in the literature that were testing the operator dependence of the technique.

Ballantyne and colleagues [16] searched for the observer variation with three examiners and they concluded with low intra- and interobserver variation. Bäuerle and colleagues [17] conducted another study with two examiners. They found a very high level of intraobserver reliability ranging from 0.92 to 0.97.

Unlike the above mentioned studies, in this study four investigators performed 60 ocular ultrasounds and obtained two measurements in axial and longitudinal planes 1 week
Apart. The results were analyzed to determine the operator dependency of the measurement of ONSD. Comparisons of the measurements obtained by the examiners 1 week apart showed that three of the examiners measured the ONSD nearly the same at two planes. The average increase of 0.05 ± 0.15 mm in the second axial measurements of one examiner was found to be statistically significant. The previous studies in this field showed a high level of intra- and interobserver reliability; unfortunately, our results were obviously inconsistent with these studies.

Cronbach \( z^b \) was used for the agreement among the four observers. The levels of compatibilities for all measurements were found at acceptable levels. Although the results were acceptable statistically, it is not possible to say that there was a perfect compatibility among the sonographers according to the previously conducted reliability studies of ultrasound measurements.

In the case of the analysis of the binary compatibilities which were assessed with intraclass correlation coefficient, we observed statistically significant incompatibilities at all measurements between examiners B and D. Although the rest of the binary comparisons seemed to be compatible statistically, they were generally at low to moderate levels. High correlation was observed only between examiners A and C.

The investigators in this study have the same level of education and experiences on ocular ultrasound. They all used the same technique and ultrasound machine during the examinations. Actually, we are unable to determine the exact reason for these variations. However the most probable explanation is the difficulties in demonstration of the nerve sheath borders both in axial and longitudinal planes. Dealing with very small distances may also contribute to these different results. Even small deflections in the gaze of eye direction may cause some errors in the measurements. Stabilizing the angle between the transducer and the globe in both axes during the whole procedure is probably another technical problem. The pressure applied by the examiner may also contribute to the results. In order to minimize wrong determinations, standardization methods should be further studied in detail. Different axes of measurements may also be helpful.

By contrast, the examiners in our study were trained in orbital ultrasound by studying healthy volunteers, as routinely done in ultrasound teaching and training courses. In fact, Zeiler and colleagues [18] reported a new unique model providing a standardized environment for ultrasound assessment of ONSD. This new model seems to have promising results. Training with a unique reliable model may probably reduce the interobserver variations.

**Limitations**

In this study, unfortunately we were not able to stabilize the gazes of the individuals externally. Also the pressure applied by the examiners could not be measured and standardized.

**Conclusion**

According to our results, it is hard to say that sonographic measurement of the ONSD is a highly reliable method both in longitudinal and transverse planes and training with a unique orbital model may reduce examiner variations.

**References**


