CLINICAL ARTICLE

Flash visual evoked potentials are unreliable as markers of ICP due to high variability in normal subjects

Linnea Andersson · Johanna Sjölund · Josefin Nilsson

Abstract

Background Previous publications have suggested a high correlation between flash visual evoked potential (F-VEP) N2 peak latency and intracranial pressure. This would enable F-VEP to be used as a non-invasive and inexpensive method to estimate ICP in a number of settings. However, basic knowledge about variability across subjects and test-retest properties of the F-VEP is lacking.

Methods Fifteen healthy adult subjects were tested on three different occasions. F-VEP responses were recorded using international standards.

Findings For the tested population, mean N2 latency was 65.7 ms (SD 10.7 ms) and the range was 48–110 ms. Intra-individual variability was high, in four of the 15 subjects more than 15 ms between testing sessions. The same was found for P2 latency and for N2 and P2 amplitudes. The response waveform was very variable and unambiguous marking of peaks was often diffucult. One out of the 15 subjects had a very poorly developed F-VEP response, but a normal pattern-reversal VEP response.

Conclusions F-VEP has a wide range of latency, amplitude and waveform across normal subjects. A large proportion of subjects also had a high intra-individual variability over time. This variability makes F-VEPs unreliable as a marker for intracranial pressure, and caution in interpreting F-VEP changes in clinical work is advised.

Keywords Flash visual evoked potentials · Intracranial pressure · N2 latency · Intracranial hypertension

L. Andersson · J. Sjölund · J. Nilsson
Department of Clinical Neuroscience,
Institute of Neuroscience and Physiology,
The Sahlgrenska Academy at the University of Gothenburg,
Göteborg, Sweden

J. Nilsson
Department of Clinical Neurophysiology,
Sahlgrenska University Hospital,
Göteborg, Sweden

J. Nilsson (✉)
Department of Clinical Neurophysiology,
Blå stråket 7-3 tr,
413 45 Göteborg, Sweden
e-mail: josefin.nilsson@neuro.gu.se

Introduction

Visual evoked potentials (VEPs) represent the responses of the visual cortex to stimuli presented in the visual field. In awake subjects, many different types of visual stimuli can be used to elicit a VEP, with the most common being the checkerboard pattern-reversal VEP; however, in comatose or sedated patients a robust stimulus such as a bright flash is needed to elicit a response. The typical VEP response to a flash stimulus is characterized by a waveform seen in Fig. 1 [11].

Somatosensory evoked potentials (SEP) are well documented as prognostic markers of cerebral damage and widely used in intensive care settings as a marker for outcome and in both children and adults [1, 13].
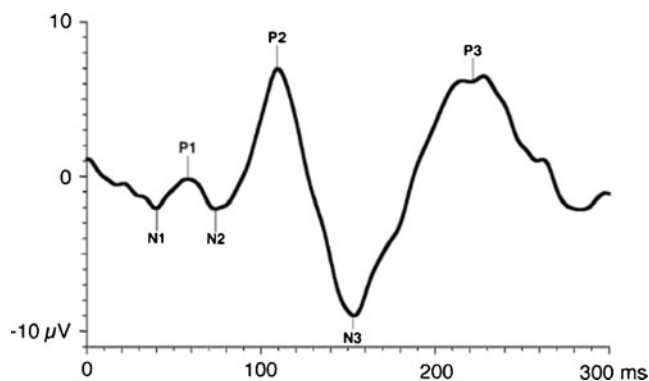
During the last 30 years there has been scattered publications of correlations between intracranial pressure (ICP) and flash VEP (F-VEP), with elevated ICP leading to delayed F-VEP latencies [4, 12, 15, 16, 18, 20]. Possible mechanisms behind this reported correlation are insufficiently studied. In children with hydrocephalus, stretching of visual pathways around dilated ventricles has been suggested [3]; other authors have suggested diffuse neural trauma or cerebral oedema [2, 19]. Recently a study in a large population reported a remarkable correlation between

Fig. 1 The typical waveform of a flash VEP [1]

invasively measured ICP and F-VEP latency estimated ICP, with $r=0.97$ [22]. The authors suggest F-VEP to be an effective method for monitoring ICP changes in seriously injured patients [22].

A correlation as strong as this would enable measurement of F-VEP latency as a non-invasive method to establish ICP. Moreover, F-VEP is inexpensive, it has a minimal risk of complications and can be performed in situations and countries where modern intensive care is uncommon. Another field of application is to monitor individuals who have shunted hydrocephalus or children with craniosynostosis in order to detect increased ICP before the appearance of obvious clinical symptoms and risk of brain damage [6, 8].

However, basic knowledge of the F-VEP is lacking and needs to be investigated before these conclusions can be drawn. It is well known that pattern reversal VEP has a low variability of latency and wave morphology, both within and across subjects [11, 14]. F-VEP, on the other hand, is known to be highly variable in waveform across the population [11]. There is one previous publication on F-VEP variability in children and adolescents with hydrocephalus; however, this group cannot be considered as normal controls and the presence of a ventricular shunt may be a bias. To the best of our knowledge, reports on intra-individual variability of F-VEP latency over time in healthy subjects is lacking. In order to suggest latency shifts in F-VEP as a marker for change of ICP, it is necessary to show if latency is stable over time in normal subjects. In this study we aim to investigate test-retest properties for F-VEP responses in healthy subjects.

Materials and methods

Subjects

Fifteen healthy students without any known ocular or neurological diseases were included. The subjects were 23–27 years old, of which eight were women and seven were men. Subjects with refractive errors were tested wearing habitual optical correction. The study was approved by the Regional Ethical Rewiev Board at the University of Gothenburg, and written informed consent was obtained from all participating indviduals.

Methods

F-VEPs were recorded on the Espion E2 workstation (Diagnosys LLC, Cambridge, UK) following ISCEV standards [11]. The flash stimulus was elicited by a Grass PS40 photic stimulator (Grass Technologies, Rodgau, Germany) with a stimulus rate of 1 Hz. Ambient lighting was used in the testing room. Three silver/silver-chloride electroencephalograph (EEG) electrodes were placed at O1, Oz and O2 according to the international 10–20 EEG system. The reference electrode was placed at Fz, and a ground electrode on the back of the hand. The amplifier bandpass filter was set at 1–100 Hz. Electrode impedance was kept below 5 kΩ. The artefact rejection filter was set at ±100 mV amplitude. The recording window was 300 ms. Two trials to each binocular, right eye and left eye stimulation were recorded (in order to prove reproducibility). Binocular recordings were always performed first, for the second run right- or left-eye stimulation was chosen arbitrarily. For each trial 64 reversals was collected. The same tester conducted all three tests for any subject to minimise investigator dependent variability.

All analyses were performed after testing was completed. An average of the two reproducible trails for each condition was obtained for marking. Marking of peaks was first done independently by two authors together (L.A. and J.S.) and then checked by an experienced neurophysiologist (J.N.). All results for the first testing session were analysed for all subjects before the next session in order to not bias (remember) the results for an individual subject. Markings were not changed when results for different testing sessions were superimposed.

Statistics

The study was dimensioned based on results from a previous pilot study [10], where the standard deviation (SD) was estimated to be 7.8 ms. Testing 15 subjects on three different occasions, a two-sided 95% confidence interval for the estimated SD will have an expected width of 4.2 ms. Microsoft Excel was used for production of graphs and for statistical analyses.

Results

The VEP measurements were conducted at three different occasions for all subjects with at least 2 days and at most

**Table 1** The study population's mean ± SD, median and range of N2 and P2 latency (*lat*) and amplitude (*amp*)

n=15	N2 lat	N2 amp	P2 lat	P2 amp
Mean ± SD	65.73±10.68	5.25±3.18	88.86±10.29	16.01±5.21
Median	61.12	4.36	85.2	16.35
Range (min-max)	48.13-109.7	0.44-20.93	74.11-125.7	4.58-31.39

10 days apart. The mean time between tests 1 and 2 was 3.5 days and between tests 2 and 3 was 3.7 days (NS, $P=0.85$). Results for the tested population are shown in Table 1.

Variance of N2 and P2 latencies for all 15 subjects are shown in Fig. 2.

For an individual subject, the greatest range of latency for N2 was between 62.8–84.04 ms and for P2 79.9–103.1 ms. The greatest amplitude range seen for N2 was between 0.4 to 9.6 μV and for P2 11.3-31.4 μV.

Figure 3 shows the waveforms from the three different recording sessions for six of our subjects.
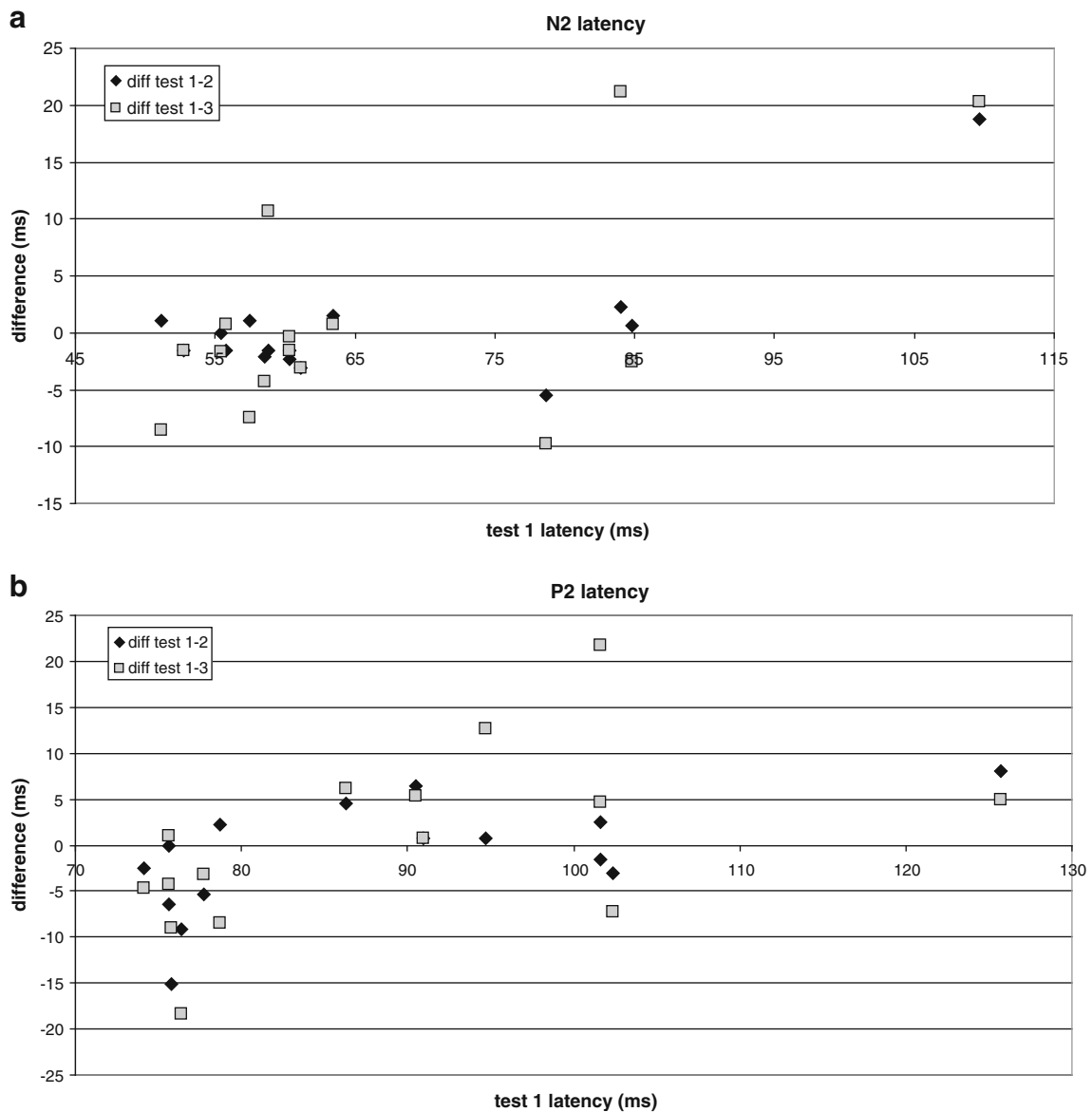
Fig. 2a, b Bland-Altman plots of variance. Averages of first, second and third recording sessions are plotted against differences between first, second and third recording sessions for each subject. The baseline is equivalent to test 1 and the markers are equivalent to the differences between tests
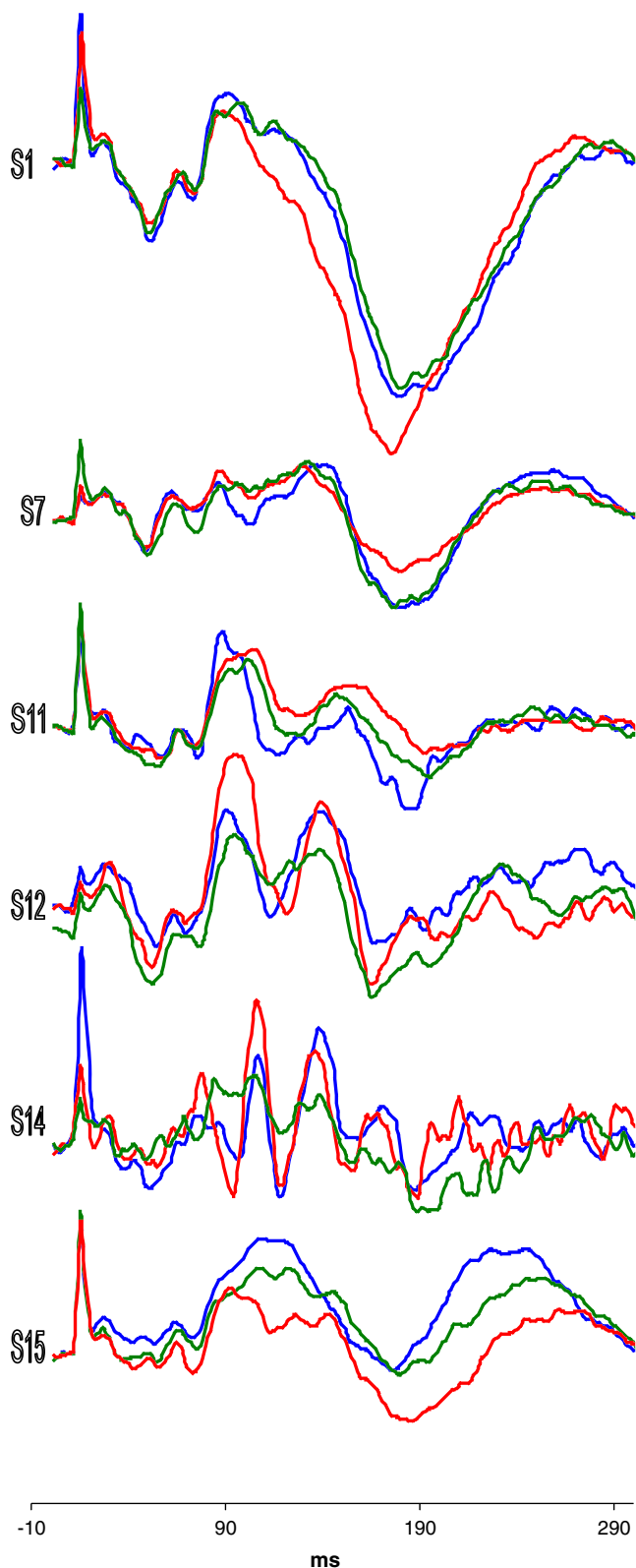
Fig. 3 Superimposed waveforms from the three different recording sessions of both eyes in six subjects (electrode Oz). *Blue trace* test 1, *red trace* test 2, *green trace* test 3

In subject 1, the N2 and P2 latencies and amplitudes were relatively stable over time.

In subject 7, the N2 latency varied significantly over time (89–110 ms), whereas the P2 latency, N2 and P2 amplitudes were more stable. Possible reasons for the large N2 variability are addressed in the discussion.

Subject 11 had a stable N2 and a relatively stable P2; however the waveform is different compared with subject 1.

In subject 12 there is a clear variability of N2 latency.

Subject 14 had a poorly developed waveform and marking the response was difficult. Interestingly, this subject had a normal pattern-reversal VEP to 30' checks (Fig. 4).

In subject 15, the P2 latency was very variable (80–103 ms), which could be explained by the double-peaked P2 and the difficulty to identify the exact position of P2.

Discussion

Most previous papers on correlations between ICP and F-VEP latency have suggested the N2 peak of the F-VEP to best correlate to ICP [7, 18, 20, 22]. In our study population, the mean latency for N2 was 65.7±10.7 ms and the range was 48.1-109.7 ms. Our results thus demonstrate a wide inter-individual range of latency, which confirms the finding from a previous pilot study [11]. Analysing intra-individual variability (test-retest), we found both subjects with very large variability of N2 latency (62.8–84.0 ms) and others with very stable results across testing sessions (60.4-61.9 ms). The same is seen for P2 latency, where the mean latency in the study population was 88.9±10.3 ms and the range was 74.1-125.7 ms. The subject with the widest range for P2 latency varied from 79.9 to 103.1 ms and the subject with the smallest range varied from 90.2-90.9 ms. A subject with a stable N2 latency over time could have a variable P2 latency over time and vice versa. According to previous published results [2, 13] and personal communication, a N2 latency shift of approximately 10 ms would indicate a clinically significant change in ICP (10–15 mmHg). However, this suggestion would be quite problematical in a clinical situation since four of our control subjects had an intra-individual variability of N2 latency of more than 15 ms. One previous publication has suggested a subpotential of the VEP, the "P-prime" (which is not an accepted waveform component by the ISCEV), to correlate with ICP [16]. We could not systematically identify a P-prime peak in our results.

Some of the variability may be a result of difficulties in marking an isolated testing session. This can be illustrated by subject 15, where the P2 peak was once presented as a triple peak, once as a double peak and once as a normal
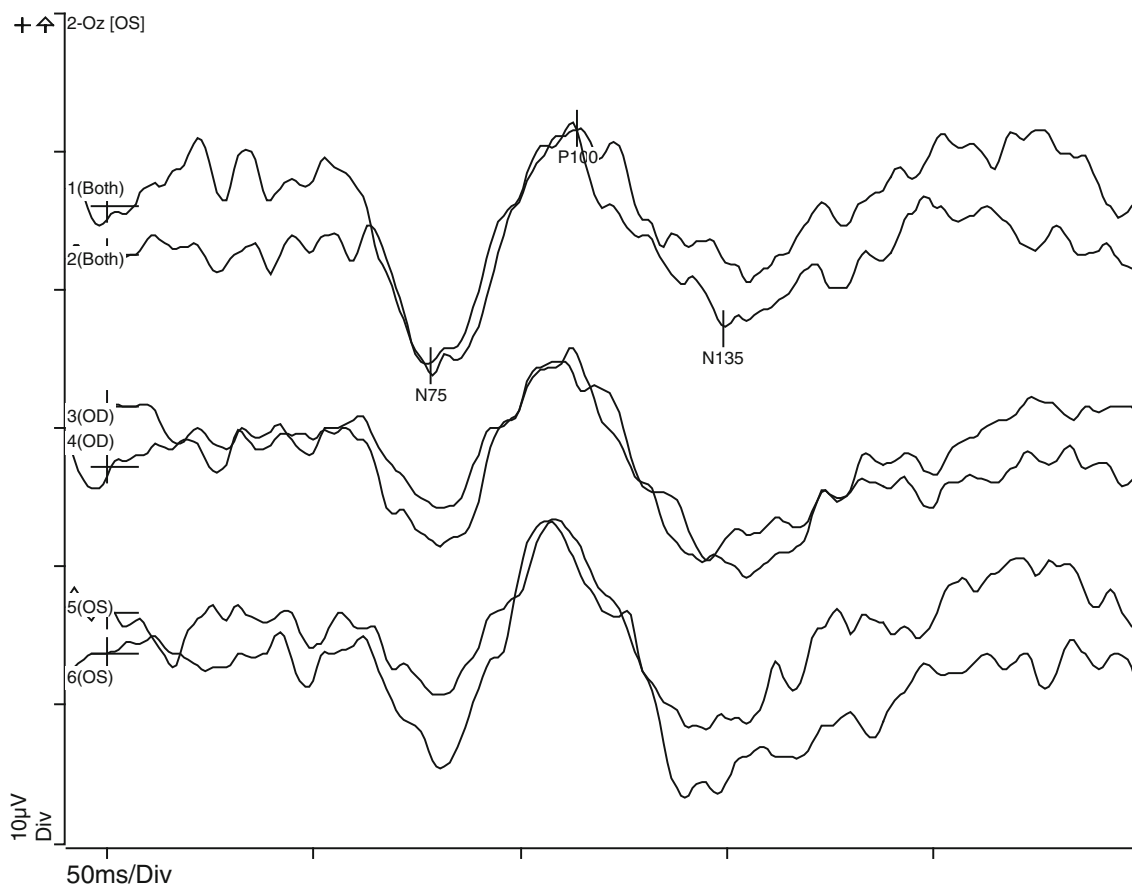
Fig. 4 Pattern-reversal VEP results for subject

single peak (Fig. 3). When we superimpose all three testing occasions, it is quite clear where to mark P2, but when the responses are presented separately, especially in a follow-up patient, it is impossible to know if the response reflects a normal variation or a real change in the response, perhaps accompanying a deterioration of the patient's clinical state. The same was true for subject 7, where it was difficult to differentiate between P1 and a double-peaked P2. We did not re-mark the waves after superimposing the three sessions, but kept our initial assessments since this better reflects the clinical situation with consecutive testing sessions. N2 was, in general, less ambiguous to mark than P2, and our data thus correlate with the opinion of N2 being overall a better choice for marking than P2. On the other hand, our results indicated a less stable N2 than has been suggested in previous studies. Desch et al. [6] have investigated F-VEP stability over time in children and adolescents with hydrocephalus. The mean difference of latency of N2 was in this study 6.5 ms and the range was between 5.3 to 16.2 ms. In our study the mean difference of N2 was 3.7 ms, but the range was 0–18.9 ms. As many as four out of our 15 subjects had an N2 variability of more

than 15 ms between testing sessions! To know if the patient of interest has a high or low intra-individual variability beforehand is of course not possible. Therefore, in a previously untested critically ill patient, one cannot tell whether the difference of latency from one time to another is a marker of a changing ICP, or a normal variation in a specific patient.

One other difficulty is illustrated in subject 14 who proved to have very poorly developed VEP responses to flash stimulation (Fig. 3). When tested with pattern-reversal VEP , the response to 30' checks was perfectly normal. Thus, one in our study population (total of 15) had a F-VEP which was very difficult to assess. This suggests that in approximately 7% of potential patients this method would be unreliable.

When it comes to correlation between ICP and amplitudes of N2 and P2 previous results are contradicting. One study of children with craniosynostosis suggested that amplitude reduction of N2 and P2 implied an elevation of ICP [8]. On the contrary, another study reported no patterns in the wave amplitude changes correlating with ICP values [22]. As with latencies, our results show both large inter-individual variability and large test-retest variability in healthy subjects,

which would make it very problematical to rely on the response as a diagnostic tool in suspected ICP elevation.

If there is a correlation between VEP and ICP, it is important that ICP is stable at the time of testing. In this study we took care to minimise ICP variations caused by external factors. Posture may of course influence ICP, during testing subjects were always seated supine and relaxed in the same chair. Due to ethical and practical reasons, circadian influence on ICP has not been studied in normal humans. In rats, ICP was found to be stable over a 24-h period [9]. Regarding circadian influence on VEP in humans, a previous study [17] has shown that there might be a variation in latency of VEP caused by circadian factors with delays during early morning hours (2–5 am). Therefore, our recording sessions were performed during the afternoon and within the same three hours for the same subject.

In a previous study, Davenport et al. [5] monitored ICP in patients with renal failure receiving haemofiltration and F-VEP sessions were conducted before, during and after treatment. Their results show that an acute ICP elevation did not directly induce VEP changes (amplitude and latency), but first after several hours. According to this study there seem to be a delay between ICP changes and VEP-changes. The same relationship was seen after treatment when ICP decreased and VEP stayed abnormal for several hours. These findings may suggest that VEP is unsuitable for intensive care patients, where clinical changes are are very rapid and intervention is needed quickly.

A limitation of this study is the fact that we were unaware of the subject's ICP during measurements. We assume that healthy individuals without optical or neurological disease have a stabile ICP. Another limitation is the length of the testing period, which did not exceed 2 weeks. A longitudinal study might show different results in VEP variability. Also, subjects requiring ICP monitoring are usually sedated and treated with other pharmacological agents, while the subjects we tested were not. There is, however, no available evidence suggesting that variability is smaller in sedated patients compared with non-sedated. Finally, since the method could be used in monitoring children with potential elevated ICP, our exemption of younger subjects can be seen as a limitation. However, test-retest studies of young children is challenging since there is considerable development of latency and wave morphology during the first 2 years of life, which is the age of most interest especially for children with craniofacial malformations.

There are several stimulation methods to produce an F-VEP, including a hand-held flash-lamp (as used in this study), LED goggles or Ganzfeld bowl flash. Previous studies on potential correlations of F-VEP and ICP have mainly used LED goggles or a hand-held flash-lamp. Whether variability of VEP latencies differs with different stimulating methods has, to the best of our knowledge, not

been studied. However, provided that the stimulating flash strength follows ISCEV standards[1], the method of flash generation is of minor importance for the resulting VEP waveform. (Limits of normal must, however, be established in each lab in order to take into account local differences in recording parameters.)

There are previous publications dealing with advanced mathematical comparisons of VEP waveforms [21]. We have not used any of these techniques since it may not be relevant for the application of this study. Previous work have shown the N2 component of the flash VEP to correlate with ICP, while the P2 component has not been to shown to have the same correlation. Mathematical comparisons of waveforms analyses the whole VEP response (0–300 ms), including both earlier and later components than the N2 peak. Results from this type of mathematical variability evaluation method may thus be misleading for the purpose of ICP estimation.

Conclusions

Our results suggest the F-VEP response to have a wide range of both latency, amplitude and waveform across normal adult subjects. Moreover, in a significant amount of subjects there is a high intra-individual variability of F-VEP over time. This variability makes F-VEPs unreliable as a marker for ICP, and caution in interpreting F-VEP changes in clinical work is advised.

Conflicts of interest None.

References

1. Carrai R, Grippo A, Lori S, Pinto F, Amantini A (2010) Prognostic value of somatosensory evoked potentials in comatose children: a systematic literature review. Intensive Care Med 36 (7):1112–1126
2. Connolly MB, Jan JE, Cochrane DD (1991) Rapid recovery from cortical visual impairment following correction of prolonged shunt malfunction in congenital hydrocephalus. Arch Neurol 48 (9):956–957
3. Constantini S, Umansky F, Nesher R, Shalit M (1987) Transient blindness following intracranial pressure changes in a hydrocephalic child with a V-P shunt. Childs Nerv Syst 3(6):379–381
4. Coupland SG, Cochrane DD (1987) Visual evoked potentials, intracranial pressure and ventricular size in hydrocephalus. Doc Ophtalmol 66:321–330
5. Davenport A, Bramley PN (1993) Cerebral function analysing monitor and visual evoked potentials as a noninvasive method of

detecting cerebral dysfunction in patients with acute hepatic and renal failure treated with intermittent machine hemofiltration. Ren Fail 15(4):515–522

6. Desch LW (2001) Longitudinal stability of visual evoked potentials in children and adolescents with hydrocephalus. Dev Med Child Neurol 43:113–117

7. Gumerlock MK, York D, Durkis D (1994) Visual evoked responses as a monitor of itracranial pressure during hyperosmolar blood-brain barrier disruption. Acta Neurochir Suppl (Wein) 60:132–135

8. Liasis A, Thompson DA, Hayward R, Nischal KK (2003) Sustained raised intracranial pressure implicated only by pattern reversal visual evoked potentials after cranial vault expansion surgery. Pediatr Neurosurg 39:75–80

9. Lin JS, Liu JHK (2010) Circadian variations in intracranial pressure and translaminar pressure difference in Sprague-Dawley Rats. Invest Ophthalmol Vis Sci 51(11):5739–5743

10. Nilsson J (2006) Intraindividual variability of flash-VEP latency over time in normal subjects. Clin Neurophys 117(Suppl 1):166

11. Odom JV, Bach M, Brigell M, Holder GE, McCulloch DL, Tormene AP, Vaegan (2010) ISCEV standard for clinical visual evoked potentials (2009 update). Doc Ophthalmol 120:111–119

12. Rizzo PA, Pierelli F, Pozzessere G, Sancesario G, Boatta M, Santarcangelo G (1984) Pattern visual evoked potentials in pseudotumor cerebri. Acta Neurol 84:57–63

13. Rothstein TL (2000) The role of evoked potentials in anoxic-ischemic coma and severe brain trauma. J Clin Neurophysiol 17 (5):486–497

14. Sarnthein J, Andersson M, Zimmermann MB, Zumsteg D (2009) High test-retest reliability of checkerboard reversal visual evoked potentials (VEP) over 8 months. Clin Neurophys 120:1835–1840

15. Soelberg Sörensen P, Trojaborg W, Gjerris F, Krogsaa B (1985) Visual evoked potentials in pseudotumor cerebri. Arch Neurol 42:150–153

16. Sjöström A, Uvebrant P, Roos A (1995) The light-flash-evoked response as a possible indicator of increased intracranial pressure in hydrocephalus. Childs Nerv Syst 11(7):381–387, discussion 387

17. Stolz G, Aschoff JC, Born J, Aschoff J (1988) VEP, physiological and psychological circadian variations in humans. J Neurol 235 (5):308–313

18. York D, Pulliam MW, Rosenfeld JG, Watts C (1981) Relationship between visual evoked potentials and intracranial pressure. J Neurosurg 55:909–916

19. Wohns RN, Colpitts M, Colpitts Y, Clement T, Blackett WB (1987) Effect of high altitude on the visual evoked responses in humans on Mt. Everest Neurosurgery 21(3):352–356

20. York D, Legan M, Benner S, Watts C (1984) Further studies with a noninvasive method of intracranial pressure estimation. Neurosurgery 14(4):456–461

21. Young RS, Kimura E (2010) Statistical test of VEP waveform equality. Doc Ophthalmol 120(2):121–135

22. Zhao YL, Zhou JY, Zhu GH (2005) Clinical experience with the noninvasive ICP monitoring system. Acta Neurochir (Wein) 95:351–355