



ORIGINAL ARTICLE

Inter-observer reliability for amplitude-integrated electroencephalography in the newborn with perinatal asphyxia^{☆,☆☆}



Carmen Bustamante-Hervás^a, Eva Valverde^{b,c}, Cristina Vega-Del-Val^a,
Susana Schuffelmann^a, Juan Arnaez^{a,c,*}

^a Unidad de Neonatología, Complejo Asistencial Universitario de Burgos, Burgos, Spain

^b Servicio de Neonatología, Hospital Universitario Infantil La Paz, Madrid, Spain

^c Fundación NeNe, Spain

Received 21 December 2020; accepted 25 January 2021

Available online 6 May 2022

KEYWORDS

aEEG;
Amplitude integrated electroencephalography;
Perinatal asphyxia;
Hypoxic-ischemic encephalopathy;
Neonate;
Concordance

Abstract

Introduction: Amplitude integrated electroencephalography (aEEG) is a tool widely used for neuromonitoring in the critical neonate. In the patient with perinatal asphyxia, its interpretation is key to identifying candidates for therapeutic hypothermia, detecting subclinical seizures and providing prognostic information. Our aim was to analyze the concordance in the interpretation of aEEG among neonatologists with different levels of experience.

Material and methods: Unicenter retrospective study of newborns ≥ 35 weeks with perinatal asphyxia included consecutively over a two-year period and monitored with aEEG for at least 6 h. The bedside neonatologist interpreted aEEG regarding background pattern, sleep-wake cycling, and seizures. The aEEG tracings were blindly reviewed by two neonatologists with different experience. The aEEG tracings were divided into periods of 0–3 h and 3–6 h of life, and the concordance (Cohen Kappa coefficient, k), between the two examiners and that of their consensus with the bedside neonatologist, was analyzed.

Results: Seventy-five newborns were included, 5 of them were not aEEG-monitored. 132 tracings were analyzed with a very good concordance between the two examiners in the three characteristics of the aEEG. The k for the bedside neonatologist was very good for background

[☆] Please cite this article as: Bustamante-Hervás C, Valverde E, Vega-Del-Val C, Schuffelmann S, Arnaez J. Concordancia interobservador del electroencefalograma integrado por amplitud en el neonato con asfixia perinatal. *An Pediatr (Barc)*. 2022;96:416–421.

^{☆☆} **Previous presentations:** partial results of this study were presented at the XXIV Congress of Neonatology and Perinatal Medicine and IV Congress of Neonatal Nursing, October 2–4, 2013, Barcelona, Spain.

* Corresponding author.

E-mail address: juan.arnaez@neurologianeonatal.org (J. Arnaez).

pattern ($k=0.93$), moderate ($k=0.52$) for sleep-wake cycling, and weak ($k=0.32$) for seizures. *Conclusions:* This study supports that background pattern is easily interpreted compared to sleep-wake cycling or crisis, improving when targeted training on aEEG is received.

© 2022 Published by Elsevier España, S.L.U. on behalf of Asociación Española de Pediatría. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

PALABRAS CLAVE

aEEG;
Electroencefalografía
integrada por
amplitud;
Asfisia perinatal;
Encefalopatía
hipóxico-isquémica;
Neonato;
Concordancia

Concordancia interobservador del electroencefalograma integrado por amplitud en el neonato con asfisia perinatal

Resumen

Introducción: La electroencefalografía integrada por amplitud (aEEG) es una herramienta utilizada en la neuromonitorización del neonato crítico. En el paciente con asfisia perinatal, su interpretación es clave para identificar a los candidatos a hipotermia terapéutica, detectar crisis subclínicas y aportar información pronóstica. Nuestro objetivo fue analizar la concordancia en la interpretación del aEEG entre neonatólogos con distinto nivel de experiencia.

Material y métodos: Estudio retrospectivo unicéntrico de los recién nacidos ≥ 35 semanas con asfisia perinatal incluidos consecutivamente durante un periodo de dos años y monitorizados con aEEG durante al menos 6 horas. El médico de guardia interpretó el aEEG respecto al trazado de base, los ciclos vigilia-sueño y las crisis. Los aEEG fueron revisados de forma ciega por dos neonatólogas con distinta experiencia. Se analizó la concordancia (coeficiente Kappa de Cohen, k) de los aEEG divididos en periodos de 0–3 horas y 3–6 horas de vida, entre ambas y la de su consenso con el médico de guardia.

Resultados: Se incluyeron 75 neonatos, 5 de ellos no se monitorizaron. Se analizaron 132 trazados con una concordancia muy buena entre las dos examinadoras en las tres características del aEEG. El k respecto al médico de guardia fue muy bueno para el trazado de base ($k=0,93$), moderado ($k=0,52$) para los ciclos vigilia-sueño y débil ($k=0,32$) para las crisis.

Conclusiones: Este estudio apoya una mayor facilidad para interpretar adecuadamente el trazado de base frente a los ciclos vigilia-sueño o las crisis, mejorando cuando se recibe una formación dirigida en el aEEG.

© 2022 Publicado por Elsevier España, S.L.U. en nombre de Asociación Española de Pediatría. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Amplitude-integrated electroencephalography (aEEG) is a technique widely used in neonatal intensive care units for neuromonitoring of critically ill newborns as a supplement to conventional EEG, which, despite the advantages involved in its multichannel setup, has limitations when it comes to prolonged monitoring and requires experienced personnel for its interpretation.^{1–3}

In patients with perinatal asphyxia in particular, aEEG can help identify newborns that may benefit from therapeutic hypothermia, detect subclinical seizures and provide data to predict future neurodevelopmental outcomes.^{1,4,5}

Although it is widely believed that aEEG is easy to interpret, few previous studies have analysed interrater agreement in aEEG interpretation, and there are limits to its interpretation by inexperienced staff, especially when it comes to the detection of seizures.^{6,7}

The aim of our study was to analyse the agreement between the neonatologist on call and 2 trained interpreters in the interpretation of aEEG traces obtained in the first 6 h post birth in a prospective cohort of newborns with perinatal asphyxia, focusing on the background trace (BT), sleep-wake cycling (SWC) and seizures.

Material and methods

The study was conducted in the framework of an integrated care programme for the management of patients with perinatal asphyxia in a tertiary care teaching hospital that houses a referral neurological intensive care unit since 2011. During the period under study, 5 neonatologists provided care to newborn patients in this unit, of who 1 was a specialist in neonatal neurology. We assessed the learning curve for the interpretation of aEEG traces in the rest of the staff of the neonatal unit starting a year before the study through the delivery of a few training sessions on aEEG and the real-world experience in neonatal monitoring in the unit, which included group discussions of clinical cases.

We consecutively included every newborn delivered at or before 35 weeks of gestation in the hospital between June 1, 2011 and May 31, 2013 with perinatal asphyxia, which was defined by the presence of at least one of the following: (1) cord blood pH of 7.00 or less; (2) 5-min Apgar of 5 or less; (3) resuscitation with intubation and/or cardiac massage or delivery of positive pressure after 5 min. We excluded newborns delivered in a different facility and those who were dying at the time of delivery.

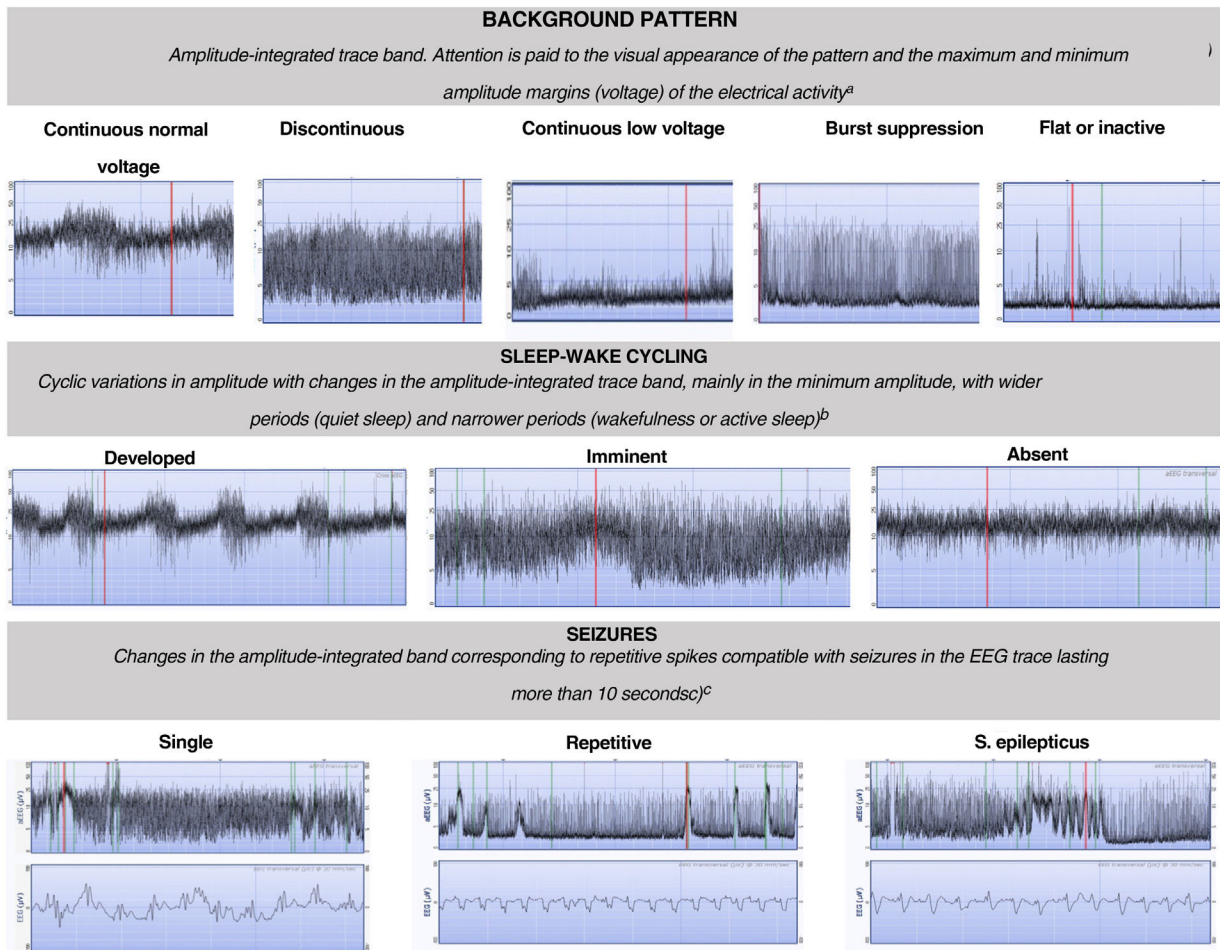


Figure 1 Classification of background, sleep-wake cycling and seizure patterns in amplitude-integrated electroencephalographic traces.^{1,5}

^a *Continuous normal voltage*: continuous activity in a narrow bandwidth with a minimum amplitude of more than 5 μV and a maximum between 10–50 μV . *Discontinuous normal voltage*: discontinuous activity in a broader bandwidth with minimum amplitude variable but of less than 5 μV and a maximum amplitude greater than 10 μV . *Burst suppression*: discontinuous background with amplitude in a narrow bandwidth without variation between 0 and 2 μV and bursts with amplitude greater than 25 μV . *Continuous low voltage*: continuous background pattern of very low voltage with a lower amplitude of less than 5 μV and a maximum of less than 10 μV . *Inactive or flat trace*: isoelectric tracing with background pattern under 5 μV .

^b Developed SWC: clearly identifiable sinusoidal changes between discontinuous and more continuous background in the aEEG with a cycle duration of at least 20 min; imminent/immature SWC: some, but not fully developed, cyclic variation of the lower border; absent SWC: no sinusoidal variations of the background trace.

^c Single seizures: maximum of 1 seizure per hour; repetitive seizures: 2 or more seizures per hour; status epilepticus: continuously ongoing repetitive seizures or a single seizure lasting at least 30 min.

Patients that met the inclusion criteria were admitted to the unit, with initiation of aEEG monitoring at admission and maintained through 6 h post birth or until the background activity normalised. The aEEG was interpreted by the physician on call based on the integrated classification based on voltages and patterns proposed by Hellström-Westas that focuses on 3 features: (1) the background trace (continuous normal voltage [CNV], discontinuous normal voltage [DNV], burst-suppression [BS], continuous low voltage [CLV] and flat trace [FT]); (2) sleep-wake cycling (absent, incipient and developed) and (3) seizures (absent, isolated, recurrent and status epilepticus)^{1,3} (Fig. 1). The presence of seizures was assessed by interpreting not only the integrated trace but

also the raw EEG trace in the monitors employed in the study (Olympic CFM 6000 or Olympic Brainz Monitor). We divided recordings into 2 time intervals for their classification: from 0 to 3 h post birth and from 3 to 6 h post birth.

When the recruitment period closed, the aEEG recordings were reviewed by two neonatologists blinded to previous interpretations. One of them (E.V.) was an expert in neonatal neurologic intensive care from another tertiary care hospital with 6 years' experience in the use of aEEG. The other neonatologist (C.B.) was part of the team of the neonatal unit where the study was conducted and had not had any previous specific training on the use of aEEG until she received a 4-h training given by the neonatologist that

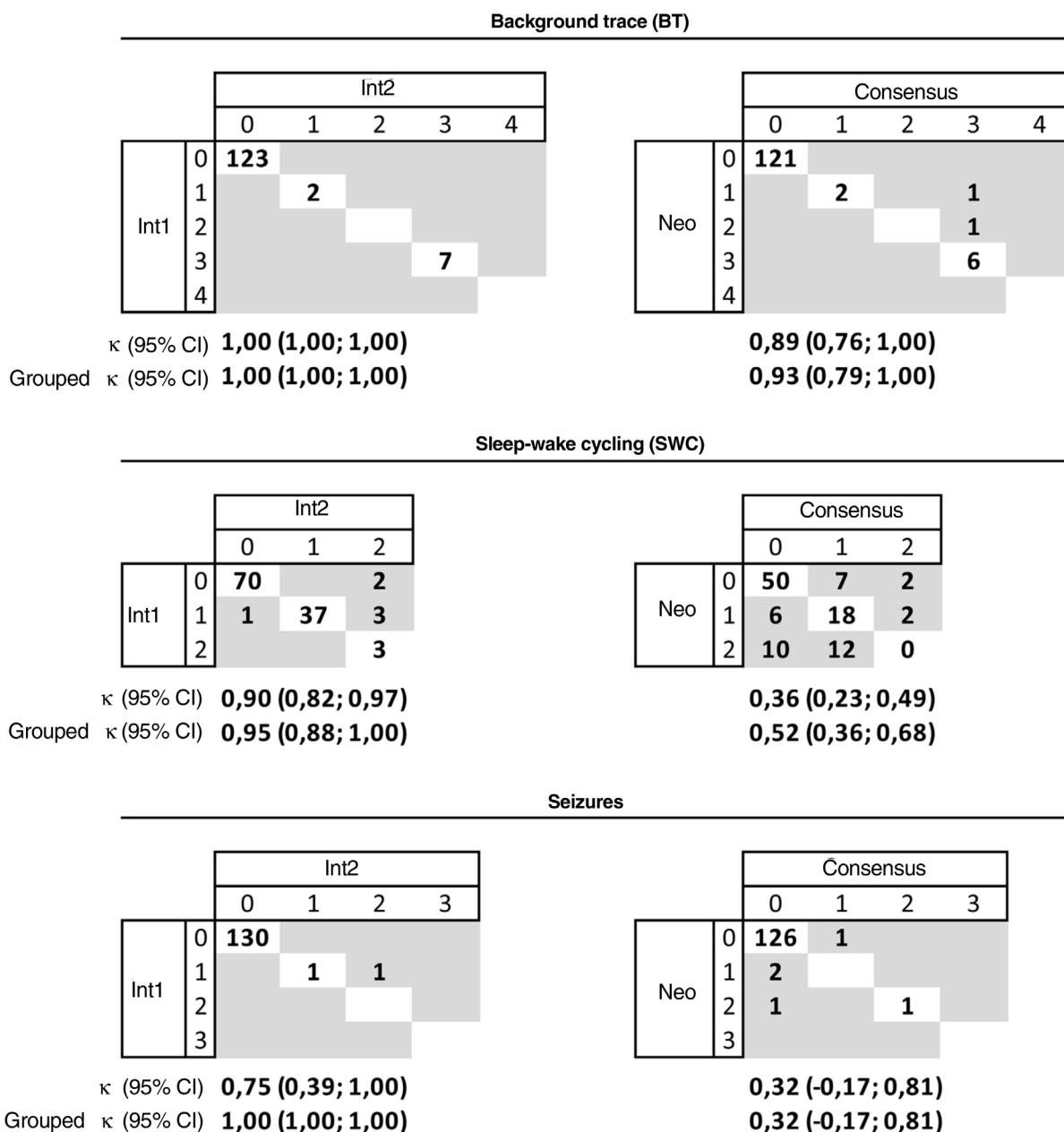


Figure 2 Concordance in the interpretation of the background pattern, sleep-wake cycling and seizures in the aEEG recording between each of the 2 interpreters and between both interpreters and the on-call neonatologist.

Background pattern: 0 (continuous pattern with normal voltage-CNV), 1 (discontinuous normal voltage-DNV), 2 (continuous low voltage-CLV), 3 (burst suppression-BS), 4 (inactive, flat trace-FT). Sleep-wake cycling: 0 (absent), 1 (developed), 2 (imminent/immature). Seizures: 0 (absent), 1 (single), 2 (repetitive), 3 (status epilepticus).

Int1: interpreter 1; Int2: interpreter 2; κ : kappa coefficient; grouped κ : grouped kappa coefficient: kappa coefficient based on the grouped classification: background pattern (CNV/DNV=normal; BS/CLV/FT=pathological), sleep-wake cycling (developed/imminent=normal, absent=pathological), and seizures (absent=normal, single/repetitive/status epilepticus=pathological); Neo: on-call neonatologist.

did have experience with the use of aEEG in our unit. Both classified the traces in the same manner as the physicians on call had before them. If there was disagreement between the two interpreters, a classification was reached by consensus to compare it to the interpretation of the on-call physician.

Statistical analysis

We present descriptive data in the form of percentages for qualitative variables and median and interquartile range or mean and standard deviation for quantitative variables based on whether the data fit a normal distribution.

Table 1 Characteristics of the sample.

Variable	n	Result
Maternal age, median (IQR), years	75	34 (33–35)
Maternal hypertension and/or diabetes, n (%)	75	10 (13%)
Primiparity, n (%)	75	40 (33%)
Gestational age, median (IQR), weeks	75	40 (39–41)
Male sex, n (%)	75	33 (44%)
Birth weight, mean \pm SD, grams	75	3119 \pm 581
Breech presentation, n (%)	75	70 (93%)
Uncomplicated birth, n (%)	75	22 (29%)
Meconium-stained fluid, n (%)	74	10 (14%)
Abnormal CTG trace, n (%)	62	58 (94%)
General anaesthesia, n (%)	75	3 (4%)
Sentinel event, n (%)	74	5 (7%)
Advanced resuscitation, n (%)	75	18 (24%)
1-min Apgar \leq 5, n (%)	75	30 (40%)
5-min Apgar \leq 5, n (%)	75	15 (20%)
Cord blood pH, median (IQR)	75	6.97 (6.91–6.99)
Base deficit $<$ 16 mmol/L, n (%)	75	34 (60%)
Assessment of HIE, n (%)	75	
No HIE		56 (75%)
Mild HIE		12 (16%)
Moderate HIE		4 (5%)
Severe HIE		3 (4%)

CTG, cardiotocography; HIE, perinatal hypoxic-ischaemic encephalopathy; IQR, interquartile range; SD, standard deviation.

We analysed interrater agreement by means of the Cohen kappa (κ) between the two interpreters and the consensus of both interpreters with the on-call neonatologist for each of the parameters considered in the interpretation of the aEEG and for the overall grouped interpretation (BT [CNV/DNV = normal; BS/CLV/FT = pathological], SWC [developed/imminent = normal, absent = pathological], and seizures [absent = normal, single/repetitive/status epilepticus = pathological]). We classified the strength of the agreement as *poor* ($\kappa < 0.20$), *fair* ($\kappa = 0.21–0.40$); *moderate* ($\kappa = 0.41–0.60$); *good* ($\kappa = 0.61–0.80$); and *very good* ($\kappa = 0.81–1.00$). The study was approved by the ethics committee (file no. 1087).

Results

The analysis included 132 traces obtained in 70 newborns out of the 75 in the study, as in the remaining 5 (4 without HIE and 4 with mild HIE) the aEEG recording was unusable. [Table 1](#) summarises the characteristics of the cohort.

The concordance between the 2 interpreters was *very good* for the BT ($\kappa = 1$) and SWC ($\kappa = 0.95$); in 1 case they disagreed about the presence/absence of developed SWC and in 5 cases one interpreter found imminent SWC that the other interpreted as developed or absent. The concordance was *very good* for the seizures: the interpreters only disagreed in 1 case in which one detected a single seizure and the other repetitive seizures. The agreement with the physician on call was *very good* ($\kappa = 0.93$) for the BT, *moderate* ($\kappa = 0.52$) for the SWC and *fair* ($\kappa = 0.32$) for seizures. In 126 traces, the physician on call and the interpreters agreed on not detecting seizure activity, in 3 traces the interpreters

did not detect seizures while the physician on call did, and in 1 the disagreement was reversed, with the interpreters finding seizures and the physician on call not ([Fig. 2](#)).

Discussion

Our findings corroborate the ease of correctly classifying the background pattern, but alert to the difficulty of interpreting SWC and especially seizures in aEEG recordings.

While in patients with HIE the changes in the BT over time are the aEEG feature that is most strongly correlated to the outcome, the interpretation of SWC is not unimportant, as absence of SWC past 72 h post birth is associated with poor neurodevelopmental outcomes.^{4,5}

The introduction of aEEG in clinical practice has marked a turning point in the diagnosis and management of seizures, as it allows their early detection, even in the absence of clinical manifestations, as is usually the case in newborns, and to better adjust antiepileptic treatment.² Our findings are consistent with those of previous studies in finding poor agreement in the interpretation of this feature, identifying this as a potential area of improvement.^{6,7}

The previous literature has already noted the need and demand for training of neonatal care teams in neuromonitoring methods and specifically in aEEG interpretation,⁸ and the fact that the second interpreter exhibited a good level of concordance with the expert following a brief training supports the notion that the implementation of standardised training courses and supervision by experienced professionals are key aspects to enable and improve aEEG interpretation.^{6,9} Furthermore, guidelines and practical atlases may be useful tools to supplement training.^{1,10}

On the other hand, although algorithms for the automated detection of seizures have been improving and their application spreading in recent years, it is important for nurses and physicians to have experience in the interpretation of these events to differentiate them from the artefacts with which they can be confused both in the aEEG and in the raw EEG traces.^{1,11,12} In such cases, the simultaneous use of conventional multichannel EEG could be helpful at the time the aEEG is showing a potential seizure, but unfortunately in most cases the episode has already ended by the time the conventional EEG has been setup and can start.

The main limitations of this study are that it was conducted in a single centre and that there were few pathological aEEG recordings. This resulted from the design of a population-based study focused on newborns with perinatal asphyxia, most of who do not develop HIE. We would have probably found a greater level of disagreement if there had been more abnormal traces, as they are harder to interpret, and artefacts that may be confused with seizures are more likely to be found in abnormal low-voltage traces. Furthermore, our findings can only be extrapolated to units of similar characteristics, as the presence of professionals with specific training in aEEG could have contributed to improving the level of agreement compared to what would be found in other facilities where they are not available.

Conclusion

Our findings establish that following a brief learning curve, the level of interrater agreement between staff with different degrees of experience is very high for the background pattern, with somewhat lower agreement in the identification of sleep-wake cycling and seizures. This corroborates the need to improve standardised training in aEEG to ensure optimal reliability in its interpretation, particularly as concerns the detection of seizures.

Funding

The study was partially funded by the Regional Department of Health of Castilla y Leon (file GRS 827/A/13).

Conflicts of interest

The authors have no conflicts of interest to declare.

References

1. Hellström-Westas L, De Vries L, Rosen I. An atlas of amplitude integrated EEGs in the newborn. 2nd ed. London: CRC Press; 2008.
2. Boylan GB, Kharoshankaya L, Mathieson SR. Diagnosis of seizures and encephalopathy using conventional EEG and amplitude integrated EEG. *Handb Clin Neurol*. 2019;162:363–400.
3. Hellström-Westas L, Rosén I, de Vries LS, Greisen G. Amplitude-integrated EEG Classification and Interpretation in Preterm and Term Infants. *NeoReviews*. 2006;7:e76–87.
4. Del Rio R, Ochoa C, Alarcon A, Arnaez J, Blanco D, Garcia-Alix A. Amplitude integrated electroencephalogram as a prognostic tool in neonates with hypoxic-ischemic encephalopathy: a systematic review. *PLoS One*. 2016;11:e0165744.
5. Thoresen M, Hellstrom-Westas L, Liu X, de Vries LS. Effect of hypothermia on amplitude-integrated electroencephalogram in infants with asphyxia. *Pediatrics*. 2010;126:e131–9.
6. Mehta B, Griffiths N, Spence K, Laing S. Inter-observer reliability in reading amplitude-integrated electroencephalogram in the newborn intensive care unit. *J Paediatr Child Health*. 2017;53:1007–12.
7. Rakshasbhuvankar AA, Wagh D, Athikarisamy SE, Davis J, Nathan EA, Palumbo L, et al. Inter-rater reliability of amplitude-integrated EEG for the detection of neonatal seizures. *Early Hum Dev*. 2020;143:105011.
8. Arnaez J, Herranz-Rubia N, Garcia-Alix A, Grupo de Trabajo E-E, Grupo de Trabajo Esp-Ehi, Unidades de Neonatología. Hospitales terciarios E. (Holistic approach of the care of the infant with hypoxic-ischaemic encephalopathy in Spain). *An Pediatr*. 2020;92:286–96.
9. Hellstrom-Westas L. Amplitude-integrated electroencephalography for seizure detection in newborn infants. *Semin Fetal Neonatal Med*. 2018;23:175–82.
10. Shellhaas RA, Chang T, Tsuchida T, Scher MS, Riviello JJ, Abend NS, et al. The American Clinical Neurophysiology Society's Guideline on continuous electroencephalography monitoring in neonates. *J Clin Neurophysiol*. 2011;28:611–7.
11. Arnaez J, Sanchez-Acosta CG, Fasce J, Garcia-Alix A. Usefulness of two-channel amplitude-integrated EEG recording in a neonatal setting. *J Matern Fetal Neonatal Med*. 2019:1–11.
12. Hagmann CF, Robertson NJ, Azzopardi D. Artifacts on electroencephalograms may influence the amplitude-integrated EEG classification: a qualitative analysis in neonatal encephalopathy. *Pediatrics*. 2006;118:2552–4.