2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

HRV analysis: a non-invasive approach to discriminate between newborns with and without seizures*

Lorenzo Frassineti[1], Antonio Lanatà[2] and Claudia Mandredi[2]

Abstract— **Early neonatal seizures detection is one of the most challenging issues in Neonatal Intensive Care Units. Several EEG-based Neonatal Seizure Detectors were proposed to support the clinical staff. However, less invasive and more easily interpretable methods than EEG are still missing. In this work, we investigated if Heart Rate Variability analysis and related measures as input features of supervised classifiers could be a valid support for discriminating between newborns with seizures and seizure-free ones. The proposed methods were validated on 52 subjects (33 with seizures and 19 seizure-free) of a public dataset collected at the Helsinki University Hospital. Encouraging results are achieved using a Linear Support Vector Machine, obtaining about 87% Area Under ROC Curve. This suggests that Heart Rate Variability analysis might be a non-invasive pre-screening tool to identify newborns with seizures.**

Clinical Relevance— **Heart Rate Variability analysis for detecting newborns with seizures in NICUs could speed up the diagnosis process and appropriate treatments for a better neurodevelopmental outcome of the infant.**

I. Introduction

Neonatal seizures are among the most common clinical signs of potential neurological disorders in newborns during their first hours of life. Their detection and timely treatment can reduce possible adverse effects on the infant's neurodevelopment and its neurological conditions even in the long term [1, 2]. However, their detection is still tricky and time-consuming, requiring highly specialised staff available 24/24h [3]. To date, EEG is the gold standard for neonatal seizures diagnosis. In the literature, several machine learning systems were applied to EEG analysis for automatic seizures detection, the so-called Neonatal Seizure Detectors (NSDs) [4, 5, 6]. Different EEG sources were evaluated to achieve more simple, affordable, and less invasive techniques. Noteworthy, Heart Rate Variability (HRV) analysis was already studied in NSDs ECG-based approaches [5, 7, 8]. Unfortunately, the performance of these systems was lower than the EEG-based NSDs, making this approach not yet reliable as a standalone technique, but only as a support to EEG analysis [5, 9]. As stated in [10], changes in the autonomic nervous system could represent a seizure manifestation and thus a possible neonatal seizures detector. Furthermore, new evidence emerged about links between the autonomic nervous system and neonatal seizures [11, 12, 13], indicating that the HRV analysis could reveal hidden relevant information. Indeed, these findings suggest that the HRV analysis might be used to discriminate between newborns with seizures and seizure-free ones. Such systems could be useful in clinical practice as a pre-screening tool in Neonatal Intensive Care Units (NICUs) to quickly identify newborns that need a deeper neurological investigation by continuous EEG (cEEG) or amplitude EEG (aEEG). To this aim, in this work, we investigated if HRV analysis could be effective to detect newborns with seizures. Preliminary results in seizure detection for adults and children were obtained using time, frequency, and nonlinear HRV features. Here we investigated such features as the input dataset of supervised classifiers to recognise newborn with seizures. Proposed methods were trained and validated on a public dataset of neonatal EEG and ECG signals collected in NICU at the Helsinki University Hospital [14]. This paper is organised as follows: in Section II, the proposed HRV analysis tools and supervised classifiers validation workflow are presented. In Section III, statistical results and supervised classifiers' performances are shown. Section IV focuses on HRV analysis as a support to neonatal seizure detection and on comparing our findings with existing literature.

II. Material and Methods

The proposed methods were tested on a public dataset of neonatal data collected in NICUs at the Helsinki University Hospital, one of the more complete public datasets of neonatal seizures [14]. It can be used as a reference set for the validation and reproducibility of algorithms related to NSD. The full dataset consists of 79 at-term newborns evaluated independently by three experts. In this study, we considered only the patients with unanimous consensus between the experts: 39 newborns with seizures and 22 seizure-free. Thus, the remaining 18 patients were discarded. Moreover, in 9 of the 61 patients, the ECG signals were not present or were highly corrupted by noise; therefore, they were discarded. Thus, the methods were defined and tested on 33 patients with seizures and 19 seizure-free subjects. To increase the Signal-to-Noise Ratio (SNR), ECGs were pre-processed and filtered with a band-pass FIR filter in the bandwidth 0,05Hz-45Hz. Then the HRV feature set was extracted with the Kubios software version 2.2 [15]. Statistical analysis, training and validation of classifiers were implemented under the MATLAB 2019b computing environment [16]. According to [11, 17], for further HRV analysis, we defined sliding time windows of 4 minutes of duration without overlap. Moreover, for artefact correction, we implemented a first order de-trending step applying a "medium correction" (for more details, see [15]). For the HRV analysis, we considered the

* This work was funded by POR FSE TOSCANA 2014/2020 and carried on under the PhD Course in Genetics, Oncology and Clinical Medicine, GenOMeC, University of Siena (Italy).

[1]L. Frassineti is with the Department of Information Engineering, Via Santa Marta 3, Universita' degli Studi di Firenze, Firenze, Italy and the Department of Medical Biotechnologies, University of Siena, Siena, Italy (e-mail: lorenzo.frassineti@student.unisi.it).

[2]A. Lanatà and C. Manfredi are with the Department of Information Engineering, Via Santa Marta 3, Universita' degli Studi di Firenze, Firenze, Italy (e-mail: antonio.lanata@unifi.it, claudia.manfredi@unifi.it)

following feature sets (for a complete description, see [15, 17]):

- Statistics Features: mean of RR intervals (mean_RR); standard deviation of RR intervals (std_RR); mean of HRV (mean_HRV); standard deviation of HRV (std_HRV); root mean square of successive RR interval differences (RMSSD); percentage of successive RR intervals that differ more than 50ms (pNN50); HRV triangular index (HRV_tri_ind); baseline width of the RR interval histogram (TINN).
- Frequency Features: peak, absolute and relative powers of Very Low (VLF), Low (LF) and High Frequencies (HF) using AR models of order 16 [15] for the spectrum estimation (e.g., ARLF_peak); the ratio between LF and HF (ARLF_HF_power), the total power (AR_tot_power) and electrocardiogram derived respiration (EDR).
- Nonlinear Features: Poincarè plot standard deviations (Poincare_SD1 and Poincare_SD2); Approximate and Sample Entropy (ApEn, SampEn with embedding dimension two and tolerance 0.2 [15]); Multiscale Entropy (from MSE1 to MSE6, embedding dimension 2 and tolerance 0.2 [15]); Detrending short- and long-term Fluctuation Analysis (DFAα1 and DFAα2); Correlation Dimension (CorDimD2).
- Recurrence Plot Analysis Features: Maximum line length (RPALmax); Mean line length (RPALmean); Divergence (RPADIV); Recurrence rate (RPAREC); Determinism (RPADET) and Shannon entropy (RPAShanEn).

According to [11], we adapted the range of the LF and HF frequency bands to the neonatal case as follows: LF (0.04-0.3) Hz; HF (0.3-1.3) Hz. According to the time window length and proper entropy features estimation, multiscale Entropy features were computed up to level 6 (MSE6) [18]. Statistical significance of each HRV measure in discriminating between patients with seizures and seizure-free patients was performed as follows:

- Test 1 (T1): Mann Whitney Test (significance level α=0.05) and Permutation Test (number of repetitions 1000, significance level α=0.05) between the medians of the seizure-free patients and the medians of the patients with seizures.
- Test 2 (T2): Mann Whitney Test (significance level α=0.05) and Permutation Test (number of repetitions 1000, significance level α=0.05) between the medians of the seizure-free patients and the medians of the patients with seizures but considering only the windows containing one or more seizure events inside them (i.e. discarding the interictal time windows).

Results are reported in Section III. The workflow for the training and validation of the classifiers was set as follows: we considered only the features relevant to the Permutation Test from T2, using the features which may discriminate between a window with seizure events from a seizure-free window. We obtained a matrix 52x13 where 52 were the patients and 13 the medians of the selected features. The classifiers were validated through the Leave-One-Subject-Out Validation (LOSO) to avoid overestimation of the performance for the neonatal seizure detection problem [19]. Before the validation, the features were normalized (zero mean and unit variances) using the training sets' statistics on the validation sets features. The following machine learning models were trained: linear Support Vector Machine (SVM); Linear Discriminant Analysis (LDA); Random Undersampling Boosting (RUSBoost); k-nearest neighbour (kNN) and Random Forest. Hyperparameter optimisation was carried out through the GridSearch method, with the same parameters for each model during the validation procedure. The best model among the set of classifier performance estimates (i.e., accuracy (ACC), F1score, Area Under ROC Curve (AUC), Sensitivity (SEN), and specificity (SPE) see Table II) was selected, based on the AUC score [4, 19]. We remark that the use of the only significant features from Test T2 might not represent the best subset of features for the classifiers and may lead to overfitting, despite the use of LOSO validation. Thus, to increase our models' performance, we retrained and validated the models considering all the features extracted by Kubios and more statistics descriptors besides the median (i.e., mean, standard deviation, maximum, minimum, kurtosis, and skewness), obtaining a matrix of size 52x294. Furthermore, a feature selection minimum-redundancy-maximum-relevance algorithm (MRMR) [20] was implemented to reduce dimensionality, obtaining a final matrix of size 52x29. Afterwards, we repeated the same validation procedure to evenly compare the two approaches. The results and the list of the features considered after the MRMR selection are reported in Section III.

III. Results

Table I shows the Statistical Tests performed on the 33 patients with seizures compared with the 19 seizure-free subjects: only the features with a significant Permutation Test obtained from T2 are shown. Furthermore, we reported the descriptive statistics mean and standard deviation for all the considered patients. For the patients with seizures, we also reported the values considering only the windows with seizure events. Noteworthy, the significant features are almost the same for both T1 and T2. Table II reports the performance of the classifiers, both for the case of features with significant Permutation Test (T2) and that of the MRMR selected features. The Linear SVM with MRMR feature selection showed the highest performance (i.e. 29 predictors). We also evaluated both the case with all the 294 predictors and the case without artefact correction and detrending (Section II) even using, instead of MRMR, Principal Component Analysis (PCA) retaining 95% of the variance. As none of these attempts gave good results, they are not reported in Table II. Finally, a list of features selected by the MRMR algorithm is shown in Table III. The threshold for feature selection was empirically given by the highest AUC score.

IV. Discussion and Conclusion

Our findings suggest that HRV analysis may successfully catch differences between newborns with seizures and seizure-free newborns in the NICUs scenario.

TABLE I. RESULTS OF STATISTICAL TESTS PERFORMED ON THE 33 PATIENTS WITH SEIZURES VS THE 19 SEIZURE-FREE SUBJECTS. ONLY THE FEATURES WITH A SIGNIFICANT PERMUTATION TEST FROM TEST 2 (T2) ARE REPORTED. THE DESCRIPTIVE STATISTICS MEAN (μ) AND STANDARD DEVIATION (σ) ARE SHOWN. MOREOVER, FOR THE PATIENTS WITH SEIZURES THE STATISTICS OF THE SEIZURE WINDOWS ARE SHOWN.

	Mann Whitney (p-value)		**PermTest (p-value)**		**Patients seizure-free (μ±σ)**	**Patients with consensus seizures (μ±σ)**	
Name Feat	**T1**	**T2**	**T1**	**T2**	**All the windows**	**Windows with seizure**	**All the windows**
std_RR (ms)	0.0109	0.0333	0.0110	0.0280	24 ± 15	15 ± 14	13 ± 12
std_HRV (1/min)	0.0092	0.0333	0.0060	0.0490	7.28 ± 4.29	4.60 ± 4.13	4.01 ± 3.68
HRV_tri_ind	0.0010	0.0046	0.0010	0.0090	6.23 ± 2.87	3.94 ± 2.72	3.59 ± 2.75
TINN (ms)	0.0065	0.0255	0.0050	0.0280	130 ± 70	80 ± 70	70 ± 60
ARLF_power_prc (%)	0.0175	0.0046	0.0190	0.0040	31.70 ± 14.48	20.47 ± 10.92	22.80 ± 10.84
Poincare_SD2 (ms)	0.0087	0.0289	0.0050	0.0310	32 ± 20	20 ± 18	17 ± 16
MSE2	N.S.	0.0383	N.S.	0.0360	0.97 ± 0.27	0.79 ± 0.29	N.S.
MSE3	0.0318	0.0016	0.0390	0.0020	1.06 ± 0.28	0.80 ± 0.27	0.89 ± 0.25
MSE4	0.0166	0.0015	0.0160	0.0020	1.15 ± 0.27	0.88 ± 0.29	0.96 ± 0.26
MSE5	0.0226	0.0062	0.0290	0.0050	1.21 ± 0.28	0.96 ± 0.31	1.03 ± 0.26
MSE6	0.0481	0.0062	0.0579	0.0110	1.26 ± 0.28	1.02 ± 0.32	1.10 ± 0.30
CorDimD2	0.0045	0.0205	0.0220	0.0150	0.58 ± 0.60	0.23 ±0.41	0.23 ± 0.55
RPALmean (beats)	N.S.	0.0098	N.S.	0.0320	20.24 ± 9.21	28.78 ± 16.43	N.S.

TABLE II. PERFORMANCE OF LEAVE-ONE-SUBJECT-OUT VALIDATION ON THE 52 PATIENTS: 33 WITH CONSENSUS SEIZURES AND 19 SEIZURE-FREE SUBJECTS. ON THE LEFT: WE USED THE 13 FEATURES WITH SIGNIFICANT PERMUTATION TEST FROM T2. ON THE RIGHT: WE USED 29 FEATURES.

	Performances using the 13 significant features to Permutation Test (T2)					**Performances considering all the Kubios features, with more statistics descriptors and MRMR selection (29 features)**				
MODEL	**ACC**	**$F1_{score}$**	**AUC**	**SEN**	**SPE**	**ACC**	**$F1_{score}$**	**AUC**	**SEN**	**SPE**
Linear SVM	65.38%	68.97%	69.86%	60.61%	73.68%	86.54%	89.23%	87.66%	87.88%	84.21%
LDA	63.46%	70.77%	67.30%	69.70%	52.63%	76.92%	81.82%	74.01%	81.82%	68.42%
RUSBoost	65.38%	73.53%	27.27%	75.76%	47.37%	67.31%	72.13%	70.49%	66.67%	68.42%
Random Forest	69.23%	77.14%	65.07%	81.82%	47.37%	63.46%	70.77%	65.07%	69.70%	52.63%
kNN	75.00%	80.60%	60.85%	81.82%	63.16%	80.77%	84.85%	71.93%	84.85%	73.68%

TABLE III. LIST OF THE FEATURES SELECTED BY THE MRMR ALGORITHM.

Statistics Descriptor	**Feature Name**
Mean	std_HRV, HRV_tri_ind, ARLF_power
Standard Deviation	HRV_tri_ind, CorDimD2
Median	RMSSD, pNN50, ARVLF_peak, ARLF_power
Max	std_RR, ARVLF_peak, ARVLF_power_prc, AR_tot_power, RPALmean
Min	std_HRV, HRV_tri_ind, ARVLF_peak, ARHF_peak, ARVLF_power, ARLF_power, ARLF_power_prc, ARHF_power, MSE6, CorDimD2
Kurtosis	RMSSD, MSE5
Skewness	ARLF_peak, ARLF_power, ARLF_HF_power

In particular, the Linear SVM performance in Table II suggests that this model is suited to this task, reaching the highest score across the tested models. It is worth noting that the AUC value (about 87%) was obtained using all HRV measures with different statistical descriptors. Moreover, feature relationship analysis through MRMR improved the classification performance with respect to the Permutation Test only. Considering the multiscale entropy features (Table I), less complexity was found for seizure windows than for seizure-free windows. This finding confirms the results already highlighted when entropy indexes were applied on EEG signals during seizure events [21]. Noteworthy,

differences between patients with seizures and seizure-free newborns were evident from the second scale, especially between MSE3 and MSE5.

On the contrary, no difference was found without the multiscale analysis (i.e. ApEn, SampEn and MSE1). Furthermore, we found several analogies with [11] in the frequency domain features, although the datasets are slightly different (in [11] also pre-term newborns were considered). Analogies were found for the lower total power for the patients with seizures: mean values of AR_tot_power for the seizure-free subjects were 798 $ms^{2,}$ while for patients with seizures, they were 314 ms^2 (T1 Mann Whitney p-value 0.01). For the HF: the mean values of ARHF_power for the seizure-free subjects were 73 $ms^{2,}$ and for patients with seizures, they were 43 ms^2 (T1 Mann Whitney p-value 0.03). Moreover, we found significant differences also in LF.

Concerning the feature selection: although MRMR already provided consistent improvement on the performance, it did not provide information about the subset relevance. Thus, other methods such as Uncorrelated LDA, Genetic Algorithms and Lasso regression could be used to evaluate different feature subsets and compute their relevance.

Though results are encouraging, more in-depth studies are required to characterize the neuro-vascular mechanisms occurring during neonatal seizures and how they could be related to cardiac parasympathetic outflow [22, 23].
Moreover, future studies could focus on the characterization of differences among seizure windows and interictal windows within each patient that are basic requirements for the definition of an ECG-based and patient-independent NSD [8]. Future perspectives could also aim at integrating the autonomic [24, 25] and central [26] nervous response to achieve a detailed picture of the physiological changes due to seizure events.

In conclusion, the present study shows the feasibility of using HRV analysis as a possible screening tool between patients with and without seizures (see Table II). Taking into account the low cost, low invasiveness, and easier usage of ECG sensors with respect to EEG ones, our findings suggest a possible integration of this approach in NICUs to allow early detection of newborns at risk of seizures.

REFERENCES

[1] S. K. Kang and S. D. Kadam, "Neonatal Seizures: Impact on Neurodevelopmental Outcomes," *Front. Pediatr.*, vol. 3, Nov. 2015.

[2] M.-P. Thibeault-Eybalin, A. Lortie, and L. Carmant, "Neonatal Seizures: Do They Damage the Brain?," *Pediatric Neurology*, vol. 40, no. 3, pp. 175–180, Mar. 2009.

[3] L. Frassineti, A. Parente, and C. Manfredi, "Multiparametric EEG analysis of brain network dynamics during neonatal seizures," *Journal of Neuroscience Methods*, vol. 348, p. 109003, Jan. 2021.

[4] L. Frassineti, D. Ermini, R. Fabbri and C. Manfredi, "Neonatal Seizures Detection using Stationary Wavelet Transform and Deep Neural Networks: Preliminary Results," 2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON), 2020, pp. 344-349.

[5] A. Temko and G. Lightbody, "Detecting Neonatal Seizures With Computer Algorithms," *Journal of Clinical Neurophysiology*, vol. 33, no. 5, pp. 394–402, Oct. 2016.

[6] A. M. Pavel et al., "A machine-learning algorithm for neonatal seizure recognition: a multicentre, randomised, controlled trial," *The Lancet Child & Adolescent Health*, vol. 4, no. 10, pp. 740–749, Oct. 2020.

[7] M. B. Malarvili and M. Mesbah, "Newborn Seizure Detection Based on Heart Rate Variability," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 11, pp. 2594–2603, Nov. 2009.

[8] B. R. Greene, P. de Chazal, G. B. Boylan, S. Connolly, and R. B. Reilly, "Electrocardiogram Based Neonatal Seizure Detection," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 673–682, Apr. 2007.

[9] B. R. Greene, G. B. Boylan, R. B. Reilly, P. de Chazal, and S. Connolly, "Combination of EEG and ECG for improved automatic neonatal seizure detection," *Clinical Neurophysiology*, vol. 118, no. 6, pp. 1348–1359, Jun. 2007.

[10] R. M. Pressler et al., "The ILAE classification of seizures and the epilepsies: Modification for seizures in the neonate. Position paper by the ILAE Task Force on Neonatal Seizures," *Epilepsia*, vol. 62, no. 3, pp. 615–628, Feb. 2021.

[11] R. Statello, L. Carnevali, D. Alinovi, F. Pisani, and A. Sgoifo, "Heart rate variability in neonatal patients with seizures," *Clinical Neurophysiology*, vol. 129, no. 12, pp. 2534–2540, Dec. 2018.

[12] A. Zimmermann, F. Domoki, and F. Bari, "Seizure-Induced Alterations in Cerebrovascular Function in the Neonate," *Dev Neurosci*, vol. 30, no. 5, pp. 293–305, 2008.

[13] M. Lucchini, W. P. Fifer, R. Sahni, and M. G. Signorini, "Novel heart rate parameters for the assessment of autonomic nervous system function in premature infants," *Physiol. Meas.*, vol. 37, no. 9, pp. 1436–1446, Aug. 2016.

[14] N. J. Stevenson, K. Tapani, L. Lauronen, and S. Vanhatalo, "A dataset of neonatal EEG recordings with seizure annotations," *Sci Data*, vol. 6, no. 1, Mar. 2019.

[15] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-aho, and P. A. Karjalainen, "Kubios HRV – Heart rate variability analysis software," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 210–220, Jan. 2014.

[16] MATLAB and Statistics and Machine Learning Toolbox Release 2019b. The MathWorks, Inc., Natick, Massachusetts, United States.

[17] F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Front. Public Health*, vol. 5, Sep. 2017.

[18] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," *Phys. Rev. E*, vol. 71, no. 2, Feb. 2005.

[19] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. B. Boylan, "Performance assessment for EEG-based neonatal seizure detectors," *Clinical Neurophysiology*, vol. 122, no. 3, pp. 474–482, Mar. 2011.

[20] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *J. Bioinform. Comput. Biol.*, vol. 03, no. 02, pp. 185–205, Apr. 2005.

[21] L. Frassineti, C. Barba, F. Melani, F. Piras, R. Guerrini, and C. Manfredi, "Automatic detection and sonification of nonmotor generalized onset epileptic seizures: Preliminary results," *Brain Research*, vol. 1721, p. 146341, Oct. 2019.

[22] M. J. Hilz et al., "Outcome of epilepsy surgery correlates with sympathetic modulation and neuroimaging of the heart," *Journal of the Neurological Sciences*, vol. 216, no. 1, pp. 153–162, Dec. 2003.

[23] G. Valenza, L. Citi, A. Lanatà, E. P. Scilingo, and R. Barbieri, 'A nonlinear heartbeat dynamics model approach for personalized emotion recognition', presented at the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2013.

[24] A. Greco, A. Lanatà, L. Citi, N. Vanello, G. Valenza, and E. Scilingo, "Skin Admittance Measurement for Emotion Recognition: A Study over Frequency Sweep," *Electronics*, vol. 5, no. 4, p. 46, Aug. 2016.

[25] A. Lanatà, A. Armato, G. Valenza, and E. P. Scilingo, "Eye tracking and pupil size variation as response to affective stimuli: a preliminary study," In 5th International Conference on Pervasive Computing Technologies for Healthcare (Pervasive Health) and Workshops (pp. 78-84) IEEE, 2011.

[26] G. Valenza et al., "Combining electroencephalographic activity and instantaneous heart rate for assessing brain–heart dynamics during visual emotional elicitation in healthy subjects," *Phil. Trans. R. Soc. A.*, vol. 374, no. 2067, p. 20150176, May 2016.