Accepted: 30 October 2017
DOI: 10.1111/epi.14049

Check for updates

Epilepsia®

SUPPLEMENT ARTICLE

Standards for testing and clinical validation of seizure detection devices

Sándor Beniczky[1] | **Philippe Ryvlin**[2,3]

[1]Department of Clinical Neurophysiology, Danish Epilepsy Center, Dianalund and Aarhus University Hospital, Aarhus, Denmark

[2]Department of Clinical Neurosciences, Vaud University Hospital Center, Lausanne, Switzerland

[3]European Epilepsy Monitoring Association, Lyon, France

Correspondence
Sándor Beniczky, Danish Epilepsy Center and Aarhus University, Dianalund, Denmark.
Email: sbz@filadelfia.dk

Summary

To increase the quality of studies on seizure detection devices, we propose standards for testing and clinical validation of such devices. We identified 4 key features that are important for studies on seizure detection devices: subjects, recordings, data analysis and alarms, and reference standard. For each of these features, we list the specific aspects that need to be addressed in the studies, and depending on these, studies are classified into 5 phases (0-4). We propose a set of outcome measures that need to be reported, and we propose standards for reporting the results. These standards will help in designing and reporting studies on seizure detection devices, they will give readers clear information on the level of evidence provided by the studies, and they will help regulatory bodies in assessing the quality of the validation studies. These standards are flexible, allowing classification of the studies into one of the 5 phases. We propose actions that can facilitate development of novel methods and devices.

KEYWORDS
clinical trials, mobile health, seizure detection, standards, validation studies, wearable devices

1 | INTRODUCTION

Recent advances in technology and in signal analysis have made seizure detection a realistic goal. Several large-scale surveys of patients and caregivers have clearly demonstrated the need for reliable automated seizure detection using wearable devices.[1-4] Recently, there has been a considerable increase in the number of publications on seizure detection.[5,6] However, the way studies are designed and reported is very heterogeneous and often confusing. It can be difficult to understand the level of evidence these studies provide.

To help in designing and reporting studies on seizure detection devices, we propose a set of standards, specific for this field. The goal is to improve the quality of these studies, and to provide readers a clear-cut picture of the position of the studies in the clinical validation process. This could be useful information for regulatory bodies too. Developing standards for testing and validating seizure detection devices requires in-depth knowledge of the field, because standards from trials on other types of medical devices can be difficult to extrapolate to seizure detection.

We provide a list of key features that are essential for studies on seizure detection, and depending on how these aspects are addressed, classify studies into 5 categories (phases) reminiscent of those used in therapeutic intervention clinical trials. The framework entails some flexibility for the first 2 phases, encouraging researchers and technical experts to engage in developing novel methods. The phase 3 pivotal trials confirming safety and accuracy are more strictly regulated. Phase 4 denotes in-field studies of seizure detection devices in the home environment of the patients, focusing on aspects related to usability (corresponding to the open-label phase of therapeutic trials). Obviously, the level of evidence needed depends largely on the study objectives. For seeking approval from regulatory bodies, or for devices where seizure detection is directly connected to a therapeutic decision or intervention, phase 3 studies should be required.

© 2018 International League Against Epilepsy

2 | KEY FEATURES OF SEIZURE DETECTION STUDIES AND THE 5 PHASES

The essential features of studies on seizure detection devices are grouped as follows: (1) subjects, (2) recordings, (3) analysis and alarms, and (4) reference standard. Table 1 lists the aspects related to these features, and how the robustness of the design can be increased. Similar to the therapeutic trials, we suggest classifying seizure detection trials into 5 phases. Although some aspects (like evaluation of safety) are common for therapeutic and for diagnostic trials, many aspects differ (pharmacokinetics and pharmacodynamics, vs accuracy measures and device deficiency). Therefore, there is a need to adapt these principles specifically for trials on seizure detection devices. Table 2 lists the 5 phases (0-4) of seizure detection studies, depending on the key features of seizure detection studies. Not all 5 phases need to be reported for all devices. A phase can be skipped or merged with another phase. Nevertheless, it is important to know the highest level of evidence available for a particular device.

It is also important to define the primary goal of the device, and the setting in which it is intended to be used (eg, alarm caregivers to intervene, provide objective seizure counts for treating physicians, aid in a clinical situation on a ward or in the epilepsy monitoring unit [EMU], nocturnal use vs whole day).

Phase 0 corresponds to initial studies on starting up or developing a novel method. Simulated data or recordings from normal subjects, using conventional or already existing devices, can be included. Testing and training the algorithm can be done on the same database (eg, using leave-one-group-out method). There is no need for a gold standard.

Phase 1 corresponds to proof-of-principle studies. These are still initial studies, but they include real seizure data from patient(s) with epilepsy; thus, a low number (1-10) of patients with seizures is accepted at this phase and recording/analysis of continuous data is not yet required. The same dataset can be used for testing and training, and the analysis is typically retrospective. There is, however, need for a reference standard for identifying seizures (including onset and

Key Points

- Standards are needed for testing and validating seizure detection devices
- Key features are: subjects, recordings, data analysis, alarms, and reference standard
- Five phases of seizure detection studies are proposed according to key features and outcome measures

TABLE 1 Key features of studies on testing and clinical validation of seizure detection devices

Key features	Design
Subjects	Simulated data or healthy subjects → Low number of patients with seizures → High number of patients with seizures
Recordings	Conventional method → Dedicated device
	Discontinuous → Continuous
	Single center → Multicenter
	Retrospective → Prospective
Analysis & alarms	Training & testing using the dataset → Predefined algorithm and cutoff values
	Offline (after the recording) → Real time (simultaneous with the recording)
	Not blinded → Blinded
Reference standard	None → Information from patient and caregiver → Video or Video-EEG recordings

EEG, electroencephalographic.
Arrows indicate increase in robustness of the studies.

TABLE 2 Five phases of development and clinical validation of seizure detection devices, depending on the key features

	Phase				
Study design	**0**	**1**	**2**	**3**	**4**
Subjects					
Simulation/ healthy subjects	+	±	–	–	–
Number of patients with seizures	–	≥1	≥10	≥20	≥50
Number of seizures	–	≥1	≥15	≥30	≥75
Recordings					
Conventional methods	±	±	–	–	–
Dedicated device	±	±	+	+	+
Continuous	±	±	±	+	+
Multicenter	±	±	±	+	na
Offline/retrospective	+	+	+	–	–
Analysis & alarms					
Training & testing using the dataset	+	+	+	–	–
Predefined algorithm and cutoff values	–	–	–	+	+
Real time	–	–	–	+	+
Blinded	–	–	–	+	+
Reference standard					
Video or video-EEG recordings	±	+	+	+	±
Information from patient and caregivers	±	–	–	–	±

EEG, electroencephalographic; na, not applicable.
+, compulsory; ±, optional; –, excluded;

offset time point) using video-electroencephalographic (EEG) recordings. For motor (convulsive) seizures, with unequivocal semiology, video recordings are sufficient.

Phase 2 needs to use a dedicated seizure detection device, on a somewhat larger number of patients with seizures (≥10), including at least 15 seizures into the analysis. Safety of the device has to be addressed. Retrospective analysis and using the same dataset for training and validation are still allowed at this phase. There is a strict need for a reference standard.

Phase 3 is the final confirmation of the safety and accuracy. It needs to address all outcome measures (see below), and it needs a strict study design. Recordings need to be continuous, using the dedicated seizure detection device. Analysis and seizure detection need to be prospective and real time (ie, logging the seizure detection time points during the recording) and use predefined algorithms, with predefined cutoff values. Training of the algorithm on the same dataset is not allowed at this phase. For studies on algorithms using individualized cutoff values or subject-specific training of the algorithm, data from the baseline phase in which the algorithm or the cutoff values are fine-tuned are not included in calculation of the primary endpoint measures. These studies need to involve patient recruitment and inclusion at several centers. Logging of the seizure detection time points is done blinded to all other data, and experts providing the reference standard (video or video-EEG recordings) need to be blinded to all data from the device. The minimum numbers of seizures and of patients with seizures vary depending on the intra- and interindividual variability of the parameter used, the sensitivity expected, and the error accepted. Therefore, a sample size calculation needs to be specified in the study. For an expected sensitivity of 95% and with a confidence interval of 10%, at least 20 patients with seizures have to be included.[7] To reliably determine the false alarm rate, at least 3000 hours of continuous recording from at least 30 patients is required. Physical exercises and activities resembling the patients′ daily routines need to be included during the long-term video-EEG monitoring. Home video or video-EEG monitoring is a good alternative to the EMU, provided the whole dataset is reviewed for inferring the reference standard. Some patients need to have multiple seizures, to determine the intraindividual consistency of seizure detections. Therefore, we propose that at least 30 seizures are recorded in phase 3 studies.

Phase 4 consists of in-field studies of seizure detection devices in the home environment of the patients, addressing aspects related to usability. This is the equivalent of the open-label phase of therapeutic trials. Patients who completed phase 3 studies can continue using the device in their homes; thus, studies can combine phases 3 and 4 in the same trial. The advantage of phase 4 is the potentially large number of patients, the longer recording time, and the real setting in which the devices work. This can be at the cost of the reference standard, because home video/video-EEG surveillance will be available only for a small group of patients, or at least not continuously (eg, monitoring only during the night). Reports of family members or caregivers alerted by the alarm can be a surrogate for the absolute reference standard. Standardized questionnaires can be useful for systematically assessing the usability of the devices. Changes in the patients' quality of life, as well as impact of the device on stress, anxiety, depression, seizure activity, and medical management, can be addressed in phase 4 studies.

3 | EXAMPLES OF THE 5 PHASES OF SEIZURE DETECTION STUDIES

We will use here examples from the development of commercially available, wearable seizure detection devices to illustrate the 5 phases.

Using surface electromyographic (EMG) data from seizures simulated by healthy volunteers, Conradsen et al[8] showed promising results for detecting tonic–clonic seizures (TCS). However, because EMG signals were not recorded from real, epileptic TCS, but from simulated seizures, this study qualifies for phase 0. Subsequently, EMG-based seizure detection algorithms were tested in small numbers (1-6) of patients, demonstrating that algorithms can correctly identify real TCS, detected from patients with epilepsy.[9,10] Although the number of participants was low, using data from real epileptic seizures qualifies these studies for phase 1.[9,10] Two studies[11,12] included a larger number of patients (n = 11) with recorded TCS (22 and 21 TCS, respectively). EMG data were recorded using conventional amplifiers, data analysis was retrospective and offline, and the threshold values for triggering seizure alarms were determined from the same dataset; sensitivity was 95%-100%. These studies[11,12] qualify for phase 2. Larger multicenter studies (69-199 patients) used dedicated wearable devices (sEMG and accelerometer combined with electrodermal activity, respectively).[13,14] However, data analysis was offline (not real time) from archived data, and the algorithm was tested retrospectively, across a wide range (95-255) of threshold-values,[13] or optimized using the leave-one-out method.[14] Therefore, these studies also qualify for phase 2.[13] Another multicenter study included 71 patients, using a wearable device, and real-time seizure detection, based on predefined threshold values,[15] thus qualifying for phase 3. All of the above studies were conducted in EMUs, to enable comparison with an unequivocal reference standard (video-EEG). Currently, there are no studies using EMG-based wearable devices in the home

environment of the patients. In this issue of *Epilepsia*, Meritam et al[16] present such a study for an accelerometry-based wearable device that has previously been validated in EMUs. This study qualifies for phase 4.

4 | OUTCOME MEASURES

It is of an utmost importance that studies account for the outcome measures that are necessary to estimate the accuracy and performance of seizure detection devices:

1. Total recording time or total time when the device was switched on. Specify also: range (minimum – maximum) and mean or median recording time per patient.
2. Device deficiency time: the proportion of the time period when the device was not functioning. Specify the reasons for the deficiency, when possible.
3. Safety: specify all adverse effects with information on their severity, seriousness, relation to the device, outcome, and impact on study withdrawal.
4. Sensitivity: number (and percentage) of all detected seizures/number of all seizures recorded during the study, belonging to the seizure type the device is targeting. Specify 95% confidence intervals. Calculate sensitivity for each patient and specify range, mean, or median of sensitivity per patient.
5. False alarm rate: the number of false alarms per 24 hours of recording. Specify also the number (and percentage) of patients experiencing false alarms, and the range, mean, or median of false alarms/patient, and describe the source of the false alarms. For patients who had seizures, specify the ratio between the true alarms and false alarms. If seizure types other than the one targeted by the device occurred during the recording, specify whether they were detected.
6. Detection latency: time (in seconds) from the seizure onset of the targeted seizure type (determined by the reference standard) to the detection time. Specify range, mean, or median. Large surveys of patients and their caregivers showed that the required detection latency was less than 10 seconds or 30 seconds.[2,3]
7. In-field studies should also address the experience of the users with handling the device and the impact of the alarms on the daily life of patients and caregivers. Standardized scales (for example 7-point Likert scale) can be used to assess these aspects. In addition, information on the time the device has been used and the number (and percentage) of users who stopped the device should be added. Kaplan-Meyer plots showing the drop in time of patients still using the device can be useful. The reasons why users opted to stop using the device should be listed.

5 | REPORTING STUDIES ON SEIZURE DETECTION DEVICES

Although seizure detection devices are typically used in patients in whom epilepsy has already been diagnosed, they distinguish seizure periods from nonseizure periods and thus belong to the realm of diagnostic devices for which clinical trials should be reported according to Standards for the Reporting of Diagnostic Accuracy Studies (STARD) criteria.[17]

A 30-item checklist is provided to account for essential aspects that need to be reported in diagnostic studies. In addition, a standardized flow diagram gives a full overview on the study.

Reporting studies according to the STARD criteria increases the quality of the reports and helps readers identify the level of evidence the study provides. Table 2 helps to adapt these general guidelines to the specific case of seizure detection. When the device specifically targets one seizure type, this needs to be stated in the title and the abstract. In the introduction, one should describe what the pathophysiological basis of the seizure detection method is, when this is possible. The Web address of the registry and study protocol has to be specified, in accordance with the STARD criteria.

6 | FACILITATING FUTURE STUDIES ON SEIZURE DETECTION DEVICES

An open, large database containing annotated signals recorded with conventional methods and possibly also with dedicated devices, from patients with epilepsy, including both annotated seizure periods, and continuous, nonseizure data, could facilitate development of new methods. This could considerably decrease the time and costs of developing new seizure detection devices, because phase 0-2 studies could be completed using such databases. Therefore, we would like to encourage the scientific community to create such open databases. An example is EPILEPSIAE, an extensive EEG database of epilepsy patients.[18]

Large research consortia, containing multidisciplinary expertise, could facilitate rapid and efficient development of seizure detection devices, and we would like to appeal to the international research community to create such consortia. These should also include lay associations, because the point of view of the patients is essential for designing and developing useful seizure detection devices.

ACKNOWLEDGMENTS

We would like to thank the participants of the First International Congress on Mobile Health Devices and Seizure

Detection in Epilepsy (Copenhagen, 2017) who contributed with excellent suggestions and comments on this topic during the panel discussion.

DISCLOSURE

None of the authors has any conflict of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

REFERENCES

1. Hoppe C, Feldmann M, Blachut B, et al. Novel techniques for automated seizure registration: patients' wants and needs. Epilepsy Behav. 2015;52:1–7.
2. Van de Vel A, Smets K, Wouters K, et al. Automated non-EEG based seizure detection: do users have a say? Epilepsy Behav. 2016;62:121–8.
3. Tovar Quiroga DF, Britton JW, Wirrell EC. Patient and caregiver view on seizure detection devices: a survey study. Seizure. 2016;41:179–81.
4. Patel AD, Moss R, Rust SW, et al. Patient-centered design criteria for wearable seizure detection devices. Epilepsy Behav. 2016;64:116–21.
5. Jory C, Shankar R, Coker D, et al. Safe and sound? A systematic literature review of seizure detection methods for personal use. Seizure. 2016;36:4–15.
6. Van de Vel A, Cuppens K, Bonroy B, et al. Non-EEG seizure detection systems and potential SUDEP prevention: state of the art: review and update. Seizure. 2016;41:141–53.
7. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J. 2003;20:453–8.
8. Conradsen I, Beniczky S, Wolf P, et al. Seizure onset detection based on a uni- or multi-modal intelligent seizure acquisition (UISA/MISA) system. Conf Proc IEEE Eng Med Biol Soc. 2010;2010:3269–72.
9. Conradsen I, Beniczky S, Hoppe K, et al. Seizure onset detection based on one sEMG channel. Conf Proc IEEE Eng Med Biol Soc. 2011;2011:7715–8.
10. Conradsen I, Beniczky S, Wolf P, et al. Automatic multi-modal intelligent seizure acquisition (MISA) system for detection of motor seizures from electromyographic data and motion data. Comput Methods Programs Biomed. 2012;107:97–110.
11. Conradsen I, Beniczky S, Hoppe K, et al. Automated algorithm for generalized tonic-clonic epileptic seizure onset detection based on sEMG zero-crossing rate. IEEE Trans Biomed Eng. 2012;59:579–85.
12. Szabó CÁ, Morgan LC, Karkar KM, et al. Electromyography-based seizure detector: preliminary results comparing a generalized tonic-clonic seizure detection algorithm to video-EEG recordings. Epilepsia. 2015;56:1432–7.
13. Halford JJ, Sperling MR, Nair DR, et al. Detection of generalized tonic-clonic seizures using surface electromyographic monitoring. Epilepsia. 2017;58:1861–9.
14. Beniczky S, Conradsen I, Henning O, Fabricius M, Wolf P. Automated real-time detection of tonic-clonic seizures using a wearable EMG device. Neurology. 2017;90:e428–34.
15. Onorati F, Regalia G, Caborni C, et al. Multicenter clinical assessment of improved wearable multimodal convulsive seizure detectors. Epilepsia. 2017;58:1870–9.
16. Meritam P, Ryvlin P, Beniczky S. User-based evaluation of applicability and usability of a wearable accelerometer device in detecting bilateral tonic-clonic seizures: a field study. Epilepsia. 2018;59(Suppl. 1):48–52.
17. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.
18. Klatt J, Feldwisch-Drentrup H, Ihle M, et al. The EPILEPSIAE database: an extensive electroencephalography database of epilepsy patients. Epilepsia. 2012;53:1669–76.

How to cite this article: Beniczky S, Ryvlin P. Standards for testing and clinical validation of seizure detection devices. *Epilepsia*. 2018;59(S1): 9–13. https://doi.org/10.1111/epi.14049