Self-Reporting Technologies for Supporting Epilepsy Treatment

Jonathan Bidwell, Eliana Kovich, Cam Escoffery, Cherise Frazier and Elizabeth D. Mynatt

Abstract

Epilepsy diagnosis and treatment relies heavily on patient self-reporting for informing clinical decision-making. These self-reports are traditionally collected from handwritten patient journals and tend to be either incomplete or inaccurate. Recent mobile and wearable health tracking developments stand to dramatically impact clinical practice through supporting patient and caregiver data collection activities. However, the specific types and characteristics of the data that clinicians need for patient care are not well known. In this study, we conducted interviews, a literature review, an expert panel, and online surveys to assess the availability and quality of patient-reported data that is useful but reported as being unavailable, difficult for patients to collect, or unreliable during epilepsy diagnosis and treatment, respectively. The results highlight important yet underexplored data collection and design opportunities for supporting the diagnosis, treatment, and self-management of epilepsy and expose notable gaps between clinical data needs and current patient practices.

Keywords: health tracking, patient self-reporting, clinical data indicators, neurocognitive, neurophysiological, implications for design

1. Introduction

Health tracking technologies such as wrist-worn seizure detection devices stand to play an increasingly important role in epilepsy diagnosis and treatment as tools that can assist patients with reporting physical [1] and psychiatric symptoms [2]. Neurologists rely on patient self-reporting to document a wide range of symptoms and triggers [3, 4] such as medication
intake, sleep, exercise, missed meals and stress levels. In practice, most patients struggle with reporting these types of information [5], and these reports are known to be highly inaccurate [6–10].

These shortcomings present two sets of challenges for technology developers and healthcare providers alike. First, the type, priority, and characteristics of clinically relevant patient data are not well documented. This presents problems for technology developers, who must understand these needs for establishing appropriate technology design requirements. Second, the performance of health tracking devices is not well explored with respect to clinical seizure reporting needs. This presents problems for healthcare providers, who require an understanding of current reporting capabilities for recommending appropriate patient self-reporting tools.

In this chapter, we present our research with clinicians to first establish clinical patient self-reporting needs with respect to clinical decision-making during epilepsy diagnosis and treatment, and second to investigate the extent that the performance of current seizure detection and classification devices may be suitable for addressing these needs.

Our study included a literature review and interviews with clinicians to identify relevant epilepsy-related symptoms and triggers; a card sorting session to prioritize these symptoms and triggers; a technology review of seizure detection devices; and finally, a pair of online surveys was administered for establishing further characteristics of clinical patient self-reporting needs during epilepsy diagnosis and treatment, respectively.

Our results include a consensus on the types, priorities, and characteristics of clinical patient self-reporting needs during anti-epileptic drug (AED) selection and treatment along with a comparison of seizure detection devices and patient self-reporting performance. These findings are intended to provide a useful reference for both developing future patient self-reporting tools and to highlight the extent that current technologies may be suitable for addressing the clinical needs that we identified within the practice. Finally, we conclude with a discussion of technology recommendations that could stand to benefit mainstream epilepsy treatment.

The main contribution of this work is a “roadmap” for developing technologies that support epilepsy treatment. The results include:

1. Identification of the type, priority, and characteristics of self-reported data that clinicians need from patients as a reference for technology developers.

2. Evaluation of current device and patient seizure reporting performance as a reference for providers.


Moreover, our study demonstrated the importance of stakeholder engagement. The role of clinical patient self-reporting is important yet often undocumented in literature. In our findings, we help to address this gap by presenting specific data collection strategies to help
providers to collect high priority patient self-reports that are not well supported by current health reporting mechanisms and technologies. This research is unique in that we investigate both the information needs of neurologists during clinical treatment and the performance statistics that technologists need for guiding development efforts.

2. Related work

2.1. Patient self-reporting challenges

Health information can also be challenging for patients to collect and therefore be unreliable even when available, or it may not be collected frequently enough to be informative [9, 11]. Handwritten and electronic patient “seizure diaries” tends to be either incomplete or inaccurate [5, 12]. In practice, many patients (30–50%) fail to report seizures during the day [6–9] while most patients fail to report seizures at night (85.8%) [9]. Eyewitness accounts often disagree on important details of how a seizure presents [9, 13], and observation is often difficult at night [5, 14]. Reminding patients to fill in reports may be ineffective as consciousness can be impaired both during and following a seizure [9]. These collective challenges present problems for clinicians as important patient information is often unavailable, unreliable, or not collected frequently enough to be informative during diagnosis and treatment [9, 11].

The types and priorities of patient self-reporting information are therefore important to understand for informing care along with the following self-reporting characteristics:

1. **Reporting availability**: Self-reports may not be available due to patient non-compliance, inability to document or observe the requested data [9], or limited awareness regarding the types of information that would be most relevant to collect in preparation for appointments.

2. **Reporting usefulness**: Self-reports may not be useful for clinical management. For example, patients may be unsure of what data to collect and fail to report important indicators [15] while devices may report data in a manner that requires considerable interpretation for answering clinical questions.

3. **Reporting reliability**: Self-reports may not be reliable due to issues such as recall bias [16]. Moreover, self-reporting performance can be difficult to assess given the absence of readily quantified measurements or validated study designs [5].

4. **Reporting difficulty**: Self-reporting may be too difficult or burdensome for patients to collect between appointments [17]. Neurologists often ask patients to document and report data such as the time, date, and a description of symptoms before, during, and after clinical presentations of symptoms [5, 18–20].

5. **Reporting frequency**: Finally, patient health data must be sampled or collected at an appropriate frequency for clinical interpretation. For example, clinicians require frequent and detailed seizure and developmental reports when treating patients with infantile spasms [24].
2.2. Health tracking design challenges

Health tracking technologies could help answer these questions but many practical challenges remain [17]. Health tracking priorities among clinicians and the patient’s role in self-reporting are each often under-specified in the literature. There is considerable interest in behavioral surveillance [21] as input for both assessing chronic conditions and evaluating self-management during treatment [22].

These advancements stand to greatly inform traditional diagnosis and treatment, but current health tracking measurements only touch on a small subset of health indicators that are relevant for patient care. The Chronic Care Model [23] is helpful in describing the role of clinical systems, healthcare communications, and self-management in patient care, but it is not instructive in terms of describing what clinical, self-management, and electronic health record (EHR) information is most important to keep track of for achieving positive long-term outcomes.

Current developments can be leveraged for greatly enhancing the capabilities of the existing systems. For example, inertial-based seizure detection wristbands are increasingly capable of detecting convulsive seizures [24]. Most patients have access to smartphones with increasingly powerful sensing capabilities [9, 10]. Well-designed health tracking [25] and health reporting tools [26] have the potential to greatly reduce the burden placed on patients to collect clinically significant health information [27]. It is, therefore, important for researchers to establish an understanding of clinical information needs and health tracking performance for developing appropriate and effective health tracking tools.

3. Methods

In this section, we present a multiphase, sequential mixed method study design. The study included a total of 16 clinicians who specialized in pediatric and adult epilepsies.

Our study included two main parts: The first part of the study investigated the types, priorities, and characteristics of useful clinical indicators during epilepsy diagnosis and treatment, respectively; while the second part investigated the performance of current seizure detection technologies as compared with current patient self-reporting.

3.1. Part 1: establishing self-reporting types, priorities, and characteristics

The first step was establishing the types, priorities, and characteristics of useful patient indicators that clinicians need during diagnosis and treatment. This included

• Interviews and a literature review to identify symptoms and triggers.

• Interviews with subject matter experts to identify five characteristics of self-reporting challenges.

The complete list of symptoms and triggers is available upon request. The resulting findings are intended to provide technology developers with insights for anticipating clinical patient self-reporting needs.
3.1.1. Investigating self-reporting needs

We conduct interviews over a 2-month period. The interviews included one-on-one meetings with one nurse practitioner specializing in pediatric epilepsy at the Children’s Healthcare of Atlanta (CHOA), Georgia and two attending specializing in adult epilepsy at Emory University Hospital, Georgia. These meetings highlighted important patient self-reporting characteristics that we would later include in our online survey.

Next, we conducted a literature review to generate a list of patient symptoms and triggers that clinicians might find as useful feedback during diagnosis and treatment. Inclusion criteria for the symptoms included any health indicators that described a specific aspect of the condition such as duration and quality; while triggers included any factors that were known to impact the likelihood of symptoms such as physical activity, sleep quality [28], and self-management behaviors [29, 30]. The literature review resulted in a list of 48 symptoms and 11 triggers that may be of interest during either diagnosis or treatment.

3.1.2. Investigating self-reporting priorities

The next step was to establish the clinical priority of these symptoms and triggers during diagnosis and treatment. A one-hour card sorting session was conducted with six pediatric epilepsy care specialists at CHOA. Informed consent was obtained from all participants. The participants included four nurse practitioners and two epileptology attendings.

The card sorting was conducted as follows. First, we printed the list of symptoms and triggers from the literature review on two separate stacks of notecards. The same card sorting exercise was conducted twice, with one stack of cards being sorted in terms of usefulness to prioritize data needs during diagnosis, and the second stack of cards being sorted to prioritize the same data during treatment. Each card contained a single symptom or trigger, and each set of notecards was shuffled beforehand.

The clinicians were asked to order the notecards in terms of “most-to-least” useful patient-reported symptoms and triggers during diagnosis and treatment, respectively. Notecards of equal importance were stacked on top of one another. New note cards were added to both piles if the clinicians believed we had overlooked any important symptoms and triggers from our literature review. Likewise, irrelevant or difficult to understand notecards were discarded from both piles, as shown in Figure 1.

The priority ranking for each card was then computed transcribing the notecards into a three-columned Excel spreadsheet that contained: (1) symptom/triggers names, (2) notecard positions during diagnosis, and (3) notecard positions during treatment, respectively and then summing the two sorted card indices for diagnosis and treatment as shown in Eq. (1).

\[
\text{Priority ranking} = \text{diagnosis card index} + \text{treatment card index}
\]

The spreadsheet was then sorted by our resulting priority ranking column from least to greatest for establishing a list of clinical data indicators that were considered important during both diagnosis and treatment. It should be noted that the exercise could have been accomplished...
3.1.3. Establishing self-reporting consensus

Next, we conducted an online survey with the aim of further understanding several practical characteristics of these clinical data collection needs. The survey was administered to 6 clinicians over a 5-week period and included the following 5 questions for each of the “top 20” highest ranked symptoms and triggers:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Availability</td>
<td>Is this information available?</td>
</tr>
<tr>
<td>2. Reliability</td>
<td>Is this information useful?</td>
</tr>
<tr>
<td>3. Usefulness</td>
<td>Is this information reliable in your opinion?</td>
</tr>
<tr>
<td>4. Difficulty</td>
<td>Is it easy or hard for patients to report?</td>
</tr>
<tr>
<td>5. Frequency</td>
<td>How frequently would you ideally like this information to be collected?</td>
</tr>
</tbody>
</table>

The survey had two pages and was designed to take less than 15 min. The first page included demographics questions. The second page contained a 20 row by 5 column table of multiple-choice questions with symptoms and triggers as rows and questions as columns.

3.2. Part 2: investigating seizure reporting performance and capabilities

The second part of our study specifically investigated clinical patient self-reporting needs surrounding patient seizure reporting. This included:
• Interviews and a literature review to identify aspects of seizure reporting.
• Technology review and a meta-analysis to present common performance statistics.

Moreover, we discussed our findings with clinicians. This feedback highlighted several important yet underexplored data collection opportunities for supporting diagnosis, treatment, and self-management.

The findings are intended to help providers to assess the extent that current seizure detection devices may be suitable for complementing patient self-reporting capabilities.

3.2.1. Investigating seizure reporting needs

Interviews, a literature review, and an online survey were conducted as background for establishing clinical seizure reporting needs during diagnosis and treatment.

The interviews included two fellows and one attending at the Emory School of Medicine and provided us with an opportunity to discuss seizure reporting practices among current patients and caregivers. The literature review included 27 papers and focused on identifying seizure reporting needs for informing clinical decision-making. The most common clinical information needs were seizure frequency, duration, type, and ability to observe seizure progression over time.

Next, we administered an online survey among an additional group of clinicians to further assess the perceived importance and accuracy of these seizure reporting measures. The survey was administered to 10 epileptologists at Emory (5 residents, 1 fellow, and 4 attendings) and included 23 Likert scale ratings. The Likert ratings were presented on a scale from 1 to 5 with 1 being “not important” and 5 being “most important” while ratings of self-reporting accuracy ranged between <20 and >80% with 5 even intervals. The respondents were also asked which type of patient reporting error would be the most detrimental during treatment and then given three choices: (a) patient over reporting, (b) patient underreporting, or (c) both errors are equally detrimental.

The online survey results highlighted the need for a follow-on technology review. Most notably, while survey respondents indicated a strong need for more accurate patient and caregiver seizure movements and seizure counts, limited research was available regarding the applicability of current technologies for addressing these needs. In addition, the literature did not specifically compare patient self-reporting to system performance [37].

3.2.2. Evaluating seizure reporting technologies

The technology review addressed these shortcomings by evaluating the performance of current systems for detecting and counting seizures, characterizing patient seizure motion, and comparing performance against that of current patient self-reporting capabilities.

Inclusion criteria included all systems that had been evaluated within a home or clinical setting. Exclusion criteria included vagus nerve [31] and brain stimulation [32] systems that required permanent surgeries and electroencephalogram (EEG) systems that can be burdensome for patients during long-term use [33, 34].
The first step was to choose performance measures that would address two sets of findings from our earlier research. First, our survey respondents showed no consensus regarding the relative impact of over or under reporting seizures. Second, our interviews with clinicians indicated that most patients and caregivers report seizures themselves without the help of seizure detection devices [3, 5]. It was, therefore, important for us to choose performance metrics that would both quantify over and under reporting as well as support comparison between seizure reporting systems and patient self-reporting rates from the literature [9].

To address these requirements, we evaluated each system in terms of three statistics: precision, recall, and F-score. Recall or sensitivity is the fraction of all seizures that were detected. High recall values reflect a low chance of under reporting or missing a seizure. Missed seizure events are problematic as untreated seizures can have serious long-term health consequences.

\[
\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \tag{2}
\]

Precision is the fraction of all relevant seizures that are detected. High precision values reflect a low chance of over reporting seizures or triggering false alarms. Low false alarm rates are important to avoid changing already effective medication.

\[
\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \tag{3}
\]

The F-score balances over and under reporting and is expressed as:

\[
F = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{4}
\]

In practice, notable inconsistencies between studies required making several assumptions. Many systems did not report precision and recall directly. In some cases, these rates had to be calculated based on information in the papers. Next, several studies presented statistics in terms of only those patients with seizures (PWS) [38–40] while other studies reported statistics for all patients in a study [41–43]. Including all patients meant that some patients without seizures might also contribute false positives. To address this discrepancy, we recomputed precision to include only those false positives from patients with seizures. For example, Poh et al. [41] reported performance for all patient and precision subsequently increased 24.54% when calculated among only those patients with seizures.

Next, we calculated patient self-reporting performance based on previous studies [18]. In this case, we assumed perfect self-reporting precision. Blum et al. [7, 9] evaluated seizure awareness among 31 patients with partial and generalized type epilepsies and observed that patients never falsely reported seizures. Next, we calculated based on observations from a similar study from Hoppe et al. [9] in which 91 patients with focal type epilepsies failed to report 32.0% of seizures during the day and 85.8% of seizures while asleep at night. This resulted in a precision of 100% for both day and night time reporting, recall values of 68.0 and 14.5% and F-scores of 0.25 and 0.81 for day and night time reporting, respectively.
4. Results

4.1. Part 1: self-reporting types, priorities, and characteristics

This section summarizes our key research findings. Figure 2 presents the type, priority, and characteristics of important information that clinicians need patients to report along with notable perceived patient self-reporting challenges and agreement between participants.

4.1.1. Self-reporting types

The first step for our analysis was establishing the types of patient self-reported data that clinicians need from patients. The bottom row of Figure 2 shows a sorted list with highest to lowest priority clinical information needs.

![Figure 2](image-url)
4.1.2. Self-reporting priorities

Next, we investigated the priority of the patient self-reported data that clinicians need from patients. The “top 20” highest priority symptoms and triggers are shown in Figure 2.

4.1.3. Self-reporting characteristics

The online survey established a consensus regarding several important self-reporting characteristics. The top row of Figure 2 shows clinician perceptions regarding the “top 20” highest ranked symptoms and triggers in terms of availability, reliability, difficulty and desired frequency; while the bottom row shows the same characteristics but categorized in terms of “unavailable”, “difficult” for patients to collect” or “unreliable”, respectively.

4.1.4. Self-reporting challenges

Next, we identified the pair of symptoms and triggers with the highest number of critical clinical responses. The most frequent clinician survey responses are shown in Table 1.

The first row in Table 1 highlights patient reporting challenges associated with information access. This includes the symptom or trigger with the greatest number of “unavailable” and “difficult” responses. The second row, further accounts for problems associated with data collection performance. This includes the symptom or trigger with the greatest number of “unavailable”, “difficult”, and “unreliable” responses, respectively. The results highlight “suicide attempts” and “seizure onset time at night” as two important unmet clinical needs.

4.1.5. Self-reporting themes

Mental health and sleep-related symptoms and triggers each appeared among the “top 20” highest ranked symptoms and triggers. Icons above the bar graphs in Figure 2 denote mental health-related symptoms and triggers such as “depression symptoms” with red circles and sleep-related symptoms and triggers such as “impaired sleep and daytime alertness” and “impaired sleep quality” with blue diamonds.

4.2. Part 2: seizure reporting technology review capabilities

This section summarizes our research findings and highlights how inaccurate patient and caregiver seizure reporting impacts clinical decision-making for prescribing and adjusting

<table>
<thead>
<tr>
<th>Epilepsy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unavailable + difficult</td>
<td>Suicide attempts</td>
</tr>
<tr>
<td>Unavailable + difficult + unreliable</td>
<td>Seizure onset time at night</td>
</tr>
</tbody>
</table>

Table 1. Most frequent clinician reported survey responses.
AEDs. Here our key findings were that limited technologies exist for supporting the process of characterizing patient seizure type, and while most seizure detection devices are more accurate than patients for nighttime reporting, these devices must be made more accurate to be beneficial for daytime use.

The results in Figure 3 provide a comparison of seizure detection device and patient self-reporting capabilities on an F-score axis between 0 and 1. The results also account discrepancies in study population size by computing performance for only those patients with seizures (PWS) as opposed to all patients that participated in each study. The following subsections describe inertial systems, video systems, and multimodal systems.

4.2.1. Inertial systems

Inertial systems utilize one or more wrist and/or chest-worn motion sensor [36, 44] and detect seizure-like convulsions as intense, repetitive limb, and torso movements with F-scores ranging from 0.133 to 0.990. These systems offer the benefit of being able to measure motion under blankets for nighttime use [36] and typically measure limb motion using an accelerometer [42], and/or gyroscope [45]. The two highest performing research systems in our review were from Schulc et al. [45] and Dalton et al. [46]. Schulc et al. [45] instrumented patients with a

![Image](Image.png)

**Figure 3.** Seizure reporting performance comparison: multiple types of non-EEG seizure detection systems are compared against patient self-reporting on a continuous F-score scale from 0.0 to 1.0, read left to right, where 0.0 is worst and 1.0 is the best performance. Each seizure detection system is represented as a circle for given class technology. The circle texture indicates the time of day that the system was evaluated and diameter reflects the relative number of patients that had at least one seizure during each study. Self-reporting performance is shown using vertical lines. Daytime performance is shown as a vertical white line with a black border while nighttime performance is shown as a solid black line, respectively.
single sensor on the forearm (98.00% precision, 100.00% recall) while Dalton et al. [46] instrumented patients with a pair of wrist-worn sensors (84.0% precision, 91% recall). The highest performing commercial product is Epi-care Free. Epi-care Free is a single wrist sensor with similar performance (81.95% precision, 89.74% recall) [43]. High false positive rates remain a challenge as rhythmic activities such as brushing teeth [42, 43] and exercise [41] are often responsible for triggering false alarms.

4.2.2. Video systems

Marker and markerless video systems been developed for detecting and classifying a range of seizure types [39] with F-scores between 0.201 and 0.964. These systems had lower overall performance than other alternatives such as inertial systems, but modern computer vision techniques are making these systems increasingly flexible and attractive for long-term use.

Markerless video systems can be trained to reliably detect patient seizure movement without the need to wear sensors on the body. For example, while prior systems were restricted to specific settings such as specific Neonatal Intensive Care Units [47, 48], more recent systems such as the one from Cuppens et al. [77] use image features that are more robust to lighting and viewpoint changes and thus applicable to different bedrooms.

Marker-based video systems, by contrast, require patients to wear active or passive markers for measuring patient motion but provide among the few examples of systems that also classify types of detected seizures [35, 38, 51]. Rémi et al. [35] used an infrared camera and retroreflective markers to track and classify different types of patient limb movements during seizures. The video recordings were analyzed to track the position of each marker over time. The relative movement of these markers between video frames was then used to discriminate between motor characteristics during different types of convulsive seizures.

4.2.3. Multimodal systems

Multimodal systems utilize inputs from multiple types of sensors thereby improving seizure detection performance with F-scores ranging from 0.083 to 0.560. Poh et al. [41] showed that electrodermal activity (EDA), in conjunction with an accelerometer, could detect seizures better than using accelerometry alone [41]. EDA measures autonomic arousal and could play a role in detecting seizures with subtle motor movement. In addition, future research may highlight differences between EDA responses on both wrists and legs for differentiating generalized and partial seizures, and for characterizing seizure laterality [52].

The MP5 system [54, 55] consisted of an under mattress microphone and accelerometer according to Ref. [56], although performance was comparatively poor (average F-score = 0.234). More recently, Pavlova et al. showed that respiration can complement video EEG during seizure diagnosis [57]. Heart rate variability [58], EDA, and respiration may enable systems to recognize life-threatening postictal depression following seizures [59, 60].
4.2.4. Audio, ECG, EMG, pressure systems

van Elmpt et al. [62] used ECG measurements for detecting the onset of heart rate changes associated with seizures and achieved competitive performance with inertial sensors (F-score = 0.391). Heart rate was observed to increase (tachycardia) at the onset of seizures and decrease following seizures (postictal bradycardia). Muscle activated sensors have been used to detect seizures [63], however, no further efforts have been made, perhaps due to adhesive EMG sensors being cumbersome to wear for long periods of time.

Mattress pressure pads have achieved mid-level performance for generalized tonic clonic (GTC) seizures [64, 65] with F-scores ranging from 0.580 to 0.78. These sensors present the added benefit of not requiring patients to wear sensors and increased privacy over having a camera installed in bedrooms. Most mattress systems, however, report false positive rates that are notably higher than inertial and video-based systems [66], due to pillows dampening pressure readings or the patient sitting up in bed [64].

4.2.5. Seizure reporting comparison between devices and patient self-report

Table 2 presents a set of statistics for comparing each system to patient [9] seizure reporting performance. Each row contains an F-score along with precision, recall, and number of patients with seizures and modality or type of system and is sorted by descending F-score for reference. Next, Table 3 presents statistics for comparing performance between each type of system. Each row includes the mean, standard deviation, minimum and maximum values together with two sets of p-values from a one-sided t-test. The p-values report the likelihood that each type of system would achieve a higher average F-score performance than that of patient self-reporting [9]. It should be noted that the t-test could not be computed for EMG and ECG as we only evaluated a single system from each of these categories.

<table>
<thead>
<tr>
<th>Systems</th>
<th>F-score</th>
<th>Precision</th>
<th>Recall</th>
<th>PWS</th>
<th>Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schult [45]</td>
<td>0.990</td>
<td>0.980</td>
<td>1.000</td>
<td>3</td>
<td>Inertial</td>
</tr>
<tr>
<td>Cuppens [66]</td>
<td>0.964</td>
<td>0.931</td>
<td>1.000</td>
<td>5</td>
<td>Video</td>
</tr>
<tr>
<td>Lu [67]</td>
<td>0.933</td>
<td>0.933</td>
<td>0.933</td>
<td>5</td>
<td>Video</td>
</tr>
<tr>
<td>Cattani [76]</td>
<td>0.921</td>
<td>0.932</td>
<td>0.910</td>
<td>1</td>
<td>Video</td>
</tr>
<tr>
<td>Karayiannis [48]</td>
<td>0.900</td>
<td>0.900</td>
<td>0.900</td>
<td>54</td>
<td>Video</td>
</tr>
<tr>
<td>Dalton [46]</td>
<td>0.874</td>
<td>0.840</td>
<td>0.910</td>
<td>5</td>
<td>Inertial</td>
</tr>
<tr>
<td>Beniczky [43]</td>
<td>0.854</td>
<td>0.814</td>
<td>0.897</td>
<td>20</td>
<td>Inertial</td>
</tr>
<tr>
<td>Kramer [72]</td>
<td>0.811</td>
<td>0.714</td>
<td>0.938</td>
<td>15</td>
<td>Inertial</td>
</tr>
<tr>
<td>Cuppens [77]</td>
<td>0.797</td>
<td>0.850</td>
<td>0.750</td>
<td>3</td>
<td>Video</td>
</tr>
<tr>
<td>Nijsen [85]</td>
<td>0.788</td>
<td>0.650</td>
<td>1.000</td>
<td>7</td>
<td>Inertial</td>
</tr>
<tr>
<td>Van de Vel, Emfit [65]</td>
<td>0.780</td>
<td>0.780</td>
<td>0.780</td>
<td>1</td>
<td>Pressure</td>
</tr>
<tr>
<td>Cuppens [73]</td>
<td>0.737</td>
<td>0.600</td>
<td>0.952</td>
<td>7</td>
<td>Inertial</td>
</tr>
<tr>
<td>Systems</td>
<td>F-score</td>
<td>Precision</td>
<td>Recall</td>
<td>PWS</td>
<td>Modality</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>Van de Vel [82]</td>
<td>0.721</td>
<td>0.578</td>
<td>0.957</td>
<td>7</td>
<td>Inertial</td>
</tr>
<tr>
<td>Conradsen [63]</td>
<td>0.682</td>
<td>0.750</td>
<td>0.625</td>
<td>2</td>
<td>EMG</td>
</tr>
<tr>
<td>Jallon [78]</td>
<td>0.639</td>
<td>0.717</td>
<td>0.577</td>
<td>2</td>
<td>Inertial</td>
</tr>
<tr>
<td>Narechania [64]</td>
<td>0.580</td>
<td>0.430</td>
<td>0.890</td>
<td>13</td>
<td>Pressure</td>
</tr>
<tr>
<td>Van de Vel, VARIA [65]</td>
<td>0.560</td>
<td>0.560</td>
<td>0.560</td>
<td>1</td>
<td>Multimodal</td>
</tr>
<tr>
<td>Poh [41]</td>
<td>0.508</td>
<td>0.349</td>
<td>0.938</td>
<td>7</td>
<td>Multimodal</td>
</tr>
<tr>
<td>Nijsen [84]</td>
<td>0.492</td>
<td>0.350</td>
<td>0.830</td>
<td>18</td>
<td>Inertial</td>
</tr>
<tr>
<td>Nijsen [83]</td>
<td>0.487</td>
<td>0.350</td>
<td>0.800</td>
<td>36</td>
<td>Inertial</td>
</tr>
<tr>
<td>Bruijne Screams [61]</td>
<td>0.459</td>
<td>0.300</td>
<td>0.980</td>
<td>17</td>
<td>Audio</td>
</tr>
<tr>
<td>Van de Vel, Epi-care Free [65]</td>
<td>0.410</td>
<td>0.410</td>
<td>0.410</td>
<td>1</td>
<td>Inertial</td>
</tr>
<tr>
<td>Van de Vel, Epi-care [65]</td>
<td>0.400</td>
<td>0.400</td>
<td>0.400</td>
<td>1</td>
<td>Inertial</td>
</tr>
<tr>
<td>van Elmp [62]</td>
<td>0.391</td>
<td>0.900</td>
<td>0.250</td>
<td>3</td>
<td>ECG</td>
</tr>
<tr>
<td>Carlson MP5 [55]</td>
<td>0.385</td>
<td>0.278</td>
<td>0.625</td>
<td>4</td>
<td>Multimodal</td>
</tr>
<tr>
<td>Pisani [47]</td>
<td>0.201</td>
<td>0.117</td>
<td>0.714</td>
<td>12</td>
<td>Video</td>
</tr>
<tr>
<td>Lockman [42]</td>
<td>0.133</td>
<td>0.072</td>
<td>0.875</td>
<td>6</td>
<td>Inertial</td>
</tr>
<tr>
<td>Fulton MP5 [54]</td>
<td>0.083</td>
<td>1.000</td>
<td>0.043</td>
<td>15</td>
<td>Multimodal</td>
</tr>
<tr>
<td>Fulton ST-2 [54]</td>
<td>0.043</td>
<td>1.000</td>
<td>0.022</td>
<td>15</td>
<td>Inertial</td>
</tr>
<tr>
<td>Bruijne Lip smacking [61]</td>
<td>0.039</td>
<td>0.020</td>
<td>0.980</td>
<td>17</td>
<td>Audio</td>
</tr>
</tbody>
</table>

**Self-reporting**

Daytime reporting [9] 0.810 100.00 68.00 91 Patient

Nighttime reporting [9] 0.253 100.00 14.50 91 Patient

**Table 2.** System and patient self-reporting performance comparison.

<table>
<thead>
<tr>
<th>Modality</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>Right-tail hypothesis test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p-Value day</td>
</tr>
<tr>
<td>Inertial</td>
<td>0.598</td>
<td>0.282</td>
<td>0.043</td>
<td>0.99</td>
<td>0.008</td>
</tr>
<tr>
<td>Video</td>
<td>0.786</td>
<td>0.292</td>
<td>0.201</td>
<td>0.964</td>
<td>0.426</td>
</tr>
<tr>
<td>Pressure</td>
<td>0.68</td>
<td>0.142</td>
<td>0.58</td>
<td>0.78</td>
<td>0.209</td>
</tr>
<tr>
<td>EMG</td>
<td>0.682</td>
<td>0</td>
<td>0.682</td>
<td>0.682</td>
<td>–</td>
</tr>
<tr>
<td>Multimodal</td>
<td>0.384</td>
<td>0.214</td>
<td>0.083</td>
<td>0.56</td>
<td>0.014</td>
</tr>
<tr>
<td>Audio</td>
<td>0.249</td>
<td>0.297</td>
<td>0.039</td>
<td>0.459</td>
<td>0.114</td>
</tr>
<tr>
<td>ECG</td>
<td>0.391</td>
<td>0</td>
<td>0.391</td>
<td>0.391</td>
<td>–</td>
</tr>
<tr>
<td>All systems</td>
<td>0.585</td>
<td>0.288</td>
<td>0.039</td>
<td>0.99</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 3.** System F-score and p-value statistics by modality.
The resulting tables can then be used for more closely examining system performance with respect to under and over reporting. High recall systems with low precision [41, 61] seldom miss seizures for addressing the concern of underreporting yet tend to overcompensate and over report seizures due to false alarms. High-precision systems with low recall [54, 62] have the opposite problem and address the concern of over reporting seizures at the risk of missing seizures. High F-score systems [45, 66, 67] have high-precision and recall values and therefore perform well without over or under reporting.

5. Discussion

5.1. Part 1: self-reporting types, priorities, and characteristics

The multiphase structure of our study was instrumental in translating our interviews, literature review, and expert panel findings into effective online survey questions. The key findings included the types, priorities, and characteristics of self-reported data that clinicians need from patients as shown in Figure 2.

The remainder of this section highlights notable patient self-reporting challenges as well as subsequent feedback after sharing these findings with clinicians.

5.1.1. Self-reporting availability

Many symptoms and triggers were reported as “useful” but “unavailable” as shown in orange in Figure 2. “Academic decline” was said to be unavailable (five out of six respondents). These findings highlight the need for patient data that may already be collected but unavailable to clinicians. Improved interoperability between electronic health records (EHRs) and electronic grading systems could alert clinicians to changes in patient grades during appointments.

5.1.2. Self-reporting difficulty

There were several symptoms and triggers that were reported as “difficult” for patients to report as shown in yellow in Figure 2. Notably, “Seizure onset at night” and “excessive sleep movements” were said to be “difficult” to report among most epilepsy specialists (five out of six respondents). These findings highlight the inherent difficulty of patient data collection while sleeping or unconscious. Introducing automated wrist-worn devices such as the Empatica E4 [68] and ActiGraph Link [69] could stand to increase patient self-reporting performance by detecting events such as seizure and unusual sleep movements, while also reducing patient and caregiver data collection burden.

5.1.3. Self-reporting reliability

Next, many symptoms and triggers were reported as being useful but unreliable when self-reported as shown in gray in Figure 2. All epilepsy specialists (six out of six respondents) agreed that patient and caregiver reports of patient “memory impairment” were “unreliable”.
These findings suggest a need for more reliable and simple patient data collection tools. For example, introducing automated data collection tools could help to increase reliability for clinicians by making data collection more accurate and consistent. The forward auditory Digit Span task, WISC-R subtest [70] could be administered as a smartphone unlock screen for periodically assessing short-term memory.

5.1.4. Self-reporting desired frequency

Finally, the majority response for desired patient self-reporting frequency is shown in each column of Figure 2. Epilepsy clinicians desired daily reports for over half of all “top 20” indicators (11 out of 20 items). New onset “viral infections” and “status seizures” require immediate medical attention for managing seizure control medications.

Finally, “episodic” patient self-reporting was most frequently desired by psychologists. Many patient behavior changes are highly context driven such as “loss of interest in activities”. These findings have implications for displaying patient health dashboards for clinicians within these respective specialties as episodic changes may be more difficult to anticipate than daily, weekly, and monthly data collection on a pre-defined schedule.

5.1.5. Self-reporting challenges

Mental health indicators and patient seizure reports may not always be available to clinicians during pediatric epilepsy treatment.

Most notably, suicide attempts were reported as useful, unavailable, and difficult for patients to collect at monthly intervals by all but one clinician. Medication side effects can often trigger depression and there is a high prevalence of depression among patients with epilepsy. If symptoms are known, then clinicians can consider prescribing a seizure control medication that may be less effective but help to stabilize mood. The main challenge for clinicians is that this type of data is often not available or difficult to collect depending on the patient’s age and caregiver situation. For example, a primary caregiver may be knowledgeable of the patient’s mental health but a patient may be accompanied by an uninformed family member.

The onset of nighttime seizure reports was also reported as useful, unavailable, and difficult for patients to collect at daily intervals by most clinicians. All respondents indicated that patient reports were unreliable due to age and cognition. This reliability is important for treating and thereby helping to reduce the risk of a condition called sudden unexpected death in epilepsy (SUDEP).

These results suggest the need for more reliable and easy to use mental health and seizure reporting tools. For example, automated weekly or monthly validated mental health surveys such as the Personal Health Questionnaire PHQ-9 [71] could be emailed or assessed in clinic on a tablet to increase reliability when it comes to screening for suicide attempts and depression. Finally, clinicians could suggest that patients wear seizure detection wristbands at night for detecting and reporting convulsive type seizures [41].
5.2. Part 2: seizure reporting performance and capabilities

Interviews and our literature review and card sorting exercise with clinicians helped us characterize the types of information that neurologists deem to be the most important during typical stages of epilepsy treatment, how likely they are to have access to this information, and the perceived accuracy of patient reports.

Most neurologists reported having access to EEG reports and verbal descriptions of seizures during treatment. However despite this information, neurologist also expressed a need for more and more detailed characterization of patient movement during seizures and more accurate seizure counts over time.

These needs were then further reflected in our literature review as we explored current methods for characterizing motion during seizures and compared existing patient seizure counting performance to current seizure detection systems. Moreover, our online survey results highlighted two important self-reporting challenges. First, there are limited recording and annotation tools available for characterizing patient motion during seizures. Second, seizure detection systems tend to have false positives and therefore over report seizures. Introducing video capture systems that are triggered by wearable seizure detection sensors may prove beneficial in both cases. More accurate seizure data could, therefore, present new opportunities for informing clinical decisions.

5.2.1. Self-reporting needs

Neurologists reported mixed reliance on patient and caregiver reports when making decisions during treatment. In our questionnaire, 70% of neurologists rated patient and caregiver self-reporting as playing a significant role when determining the best course of AED treatment (4 or greater on a scale of 5), however there was considerable in terms of how frequently these initial self-reports included patient movement characteristics during seizures (SD = 1.10) and/or described the evolution of the seizure over time (SD = 0.78). This finding suggests that, while neurologists perceive self-reporting as important, they also emphasize the need for evaluating the validity of patient reports.

5.2.1.1. Informing initial AED selection

Neurologists from our survey indicated that support for characterizing patient seizure type could be beneficial for selecting the most suitable initial AED based on the patient’s seizure symptoms. The survey respondents ranked seizure type and movement characterization as the most important information during the initial diagnosis and AED selection phase. Most respondents had access to EEG reports (7/10) and MRI reports (5/10) and verbal accounts of seizures (80%). Less than one-third of neurologists had access to hospital records, imaging records, blood work, seizure diaries, and video of patient seizure events. Most notably, while all neurologists (10/10) expressed a desire for supplemental video only 3/10 respondents had regular access to such video for informing diagnosis and treatment.
These findings stress a need for capturing additional patient information prior to diagnosis and have implications for patient and caregiver data collection efforts. MRI and EEG may not be available for first-time general practitioner referrals. Initial outpatient EEG sessions tend to be short, ~20 min. Moreover, even with routine activation procedures such as patient hyperventilation, photic stimulation, and sleep withdrawal, many patients may not show symptoms during a single visit and require further observation. It may, therefore, be helpful for patients to collect additional seizure observations such as video recordings prior to initial appointments.

5.2.1.2. Informing AED adjustment

Neurologists ranked seizure frequency as the most important patient self-reported information available to them (100% rated 5 out of a scale of 5) for making AED adjustments. Most neurologists (8/10) estimated that patients failed to report between 40 and 60% of seizures overall (given 5 uniform ranges between 0 and 100%). This estimate agreed with Hoppe et al.’s findings that 55% of patients failed to document 55% of seizures overall [9]. Most neurologists also agreed that an ictal description of a seizure is the most difficult for a patient to report, and 66% of the surveyed neurologists said that patients or caregivers report less than 60% of their seizures. It may, therefore, be helpful to introduce seizure detection devices that address specific patient challenges such as nighttime reporting.

5.2.2. Seizure reporting shortcomings

Major shortcomings of current seizure classification and detection technologies include (1) limited capture and playback solutions for characterizing seizure type and (2) inaccurate seizure detection for counting seizures and limited support for identifying seizure types that do not exhibit limb movement.

5.2.2.1. Limited tools for AED selection

The prospect of developing motion characterization tools for informing initial AED selection remains largely unexplored. Efforts have been limited to active and passive motion tracking as additional feedback for EEG technicians [38, 51]. To date, existing research and commercial systems have not focused on the problem of motor characterization for initial partial versus generalized seizure characterization. There is, therefore, a need to utilize additional video and motion tracking technologies for informing AED selection.

5.2.2.2. Inaccurate seizure counts for AED adjustment

Neurologists from our survey agreed that accurate seizure counts are the most important feedback. In our review, inertial seizure detection systems [43, 45, 46, 72] achieved higher performance than embedded mattress devices [55, 65] and multimodal devices [41]. Inertial devices also tend to support daytime use, while mattress and video systems are often limited nighttime use within bedrooms [47, 65].

High false positive rates remain a problem. More accurate seizure counts could better inform AED treatment. We contend that false positives remain a problem despite studies with higher
precision, but fewer numbers of patients [45, 46, 73]. In turn, more work is needed for reducing false positives among all classes that we surveyed. Table 3 shows that many of the best performing systems utilize video with an average F-score of 0.79 (SD = 0.29) while audio-based systems performed the worse with an average F-score performance of 0.25 (SD = 0.30) with precision as low as 2% for detecting audible lip smacking [61]. High p-values above 0.05 in Table 3 suggest that mean F-scores for each type of system share a greater than chance probability of performing better than self-reporting at night while low p-values suggest that systems will perform worse than patients during the day on average.

Most systems performed better than patient reporting during the night but notably worse than patients during the day. In our review, all but two systems achieved higher F-score performance at night while only four inertial systems performed better during the day. The average F-score for all systems was 0.59 (SD = 0.29); this reflects a notable improvement over patient self-reporting at night (F-score 0.25) yet remains significantly worse than self-reporting during the day (F-score 0.81). Inertial systems are shown to perform well across both day and nighttime studies [41, 43], however, as noted more work must be done for reducing false alarms during daily activities [40].

High-performance variability was observed between the same types of systems. These discrepancies can largely be explained by the following four contributing factors:

1. **Day versus night**: Many systems were only evaluated at night [55, 65, 73], or strictly during the day, [36] while others were evaluated during the night and day [41–43]. Nighttime studies tended to perform better than daytime studies with an average F-score of 0.62 as compared to 0.56 during studies that included daytime monitoring. This makes direct comparison difficult because daytime seizure detectors must also distinguish non-seizure events such as exercise and teeth brushing which were less prevalent at night. For example, Van De Vel et al.’s [65] evaluation of the Emfit pressure mat highlighted false positives when sitting up in bed, Pisani et al. [47]’s video analysis confused random infant movements, Lockman et al.’s inertial wristband [42] reported false positives during rhythmic activities such as brushing teeth and pen tapping while Poh et al. [41] reported similar false positives during dice rolling and video game activities.

2. **Alerting versus reporting**: Many systems are primarily designed for alerting caregivers rather than accurately reporting seizure counts. Existing commercial systems are designed for alerting caregivers to ongoing seizures [43, 53, 67]. The caregiver is often burdened with adjusting system-specific threshold settings for minimizing false positives [51, 74–75]. This, in turn, may result in missing facial tics and other less apparent symptoms.

3. **Patient age**: We observed considerable variation between the age and number of patients enrolled in studies. For example, Cuppens et al. [73] and Lockman et al. [42] each developed similar inertial-based systems, however, Cuppens et al. [73] studied patients aged 5–16 while Lockman et al. [35] studied ages 3–85. It may be reasonable to expect that differences in muscle development and limb length between these age groups could have resulted in slightly different movement characteristics during seizures.

4. **Patient count**: The number of patients with seizures also varied between studies with a single patient having seizures at night. For example, seven studies had less than four
participants [45, 62, 63, 65, 76–78]. Van De Vel et al. [65] and Narechanie et al. [64] each evaluated pressure sensing mattress inserts, however, Van De Vel et al. [65] included only 1 patient with an F-score of 0.78 while Narechanie et al. [64] included 51 patients with a perfect F-score of 1.0 for reporting seizure counts at night.

5.2.2.3. Limited diversity of seizure types

Most patients have focal types seizures (70%>50%) [80, 81]; however, only some but not all focal seizures involve limb movement. This presents a challenge as most systems to date are limited to measuring seizures based on limb movements. More reliable metrics or a combination of metrics should be studied for capturing non-motor seizure symptoms.

To date, there has been limited work on detecting seizures using non-inertial and video sensors. Bruijnne et al. analyzed [61] audio for detecting “lip smacking” and “screams” however; the performance was among the poorest of all the systems that we evaluated (F-score = 0.04). To the best of our knowledge, there are no non-EEG devices for detecting symptoms (e.g. subtle face or hand movement during partial seizures or behavioral arrest).

Inertial seizure detection wristbands [43] and nighttime video recording could provide a promising short-term solution for increasing the accuracy of patient reporting. Most patients are seen by a general practitioner and are later referred to see a neurologist [79]. This gap presents an opportunity to equip patients with data collection systems for detecting and recording patient seizures in the home prior to an initial neurology visit. For example, the open source OpenSeizureDetector [52] inertial wristband could be used in conjunction with an already available and bedroom instrumented camera such as the SAMi [49] or OpenSeizureDetector [50] detect seizures and trigger video recording. In turn, neurologists could review video for characterizing the seizure prior to treatment. Movements during seizures could be captured and reviewed.

Finally, seizure reporting video annotation tools could enable patients, caregivers, and neurologists to label the start and stop of seizure events could improve seizure detection performance over time and address the problem of having to manually adjust thresholds as in commercial products [42, 43] as the system will be trained for a particular individual.

6. Conclusions

The role of clinical patient self-reporting is important yet often undocumented in literature. Health tracking technologies such as wrist-worn seizure detection devices stand to play an increasingly important role in epilepsy treatment and diagnosis as data collection tools that can help patients and caregivers to collect self-report seizure counts and other high priority health indicators for informing clinical decision-making.

In this paper, we conducted a multiphase study that included interviews with clinicians, two literature reviews, a card sorting exercise and online surveys for investigating clinical patient self-reporting needs within the context of epilepsy diagnosis and treatment. In our work with clinicians, we identified a need for more reliable mental health reporting
and sleep indicators during epilepsy treatment. In our technology review, we surveyed seven types of seizure detection sensing modalities and identified a strong need for more accurate and reliable seizure reporting and motion characterization during diagnosis and treatment.

The key challenges faced by technology developers and providers are:

1. Identifying the specific types, priorities, and characteristics of data that clinicians need from patients.
2. Establishing the extent that current health tracking devices are suitable for addressing these needs is similarly unknown.

The findings from our research highlighted important patient self-reporting needs among a diverse set of clinicians for epilepsy diagnosis and treatment and in turn, may provide clinicians and technology developers with a useful reference for aligning development efforts with clinical information needs within epilepsy treatment. High false positives remain a problem for seizure detection devices; however low-cost hardware may be able to mitigate these issues. For example, inertial sensors [42] and the bedroom instrumented camera [49] could be sent home with patients prior to treatment; such a combination of tools could be vital information aiding neurologist in how best to treat the patient.

Acknowledgements

This chapter is taken in part from our paper on, “Seizure reporting technologies for epilepsy treatment: A review of clinical information needs and supporting technologies” [24].

Author details

Jonathan Bidwell*, Eliana Kovich², Cam Escoffery³, Cherise Frazier² and Elizabeth D. Mynatt¹

*Address all correspondence to: bidwej@gatech.edu
1 Georgia Institute of Technology, School of Interactive Computing, Atlanta, GA, USA
2 Children’s Healthcare of Atlanta, GA, USA
3 Emory University, Rollins School of Public Health, Atlanta, GA, USA

References


Panayiotopoulos CP. A Clinical Guide to Epileptic Syndromes and their Treatment. London: Springer; 2010


Poh M-Z. Continuous Assessment of Epileptic Seizures with Wrist-Worn Biosensors. Massachusetts Institute of Technology; 2011


Karayiannis NB, Tao G, Frost JD Jr, Wise MS, Hrachovy RA, Mizrahi EM. Automated detection of videotaped neonatal seizures based on motion segmentation methods.


[57] Pavlova M. et al. Advantages of respiratory monitoring during video-EEG evalua
tion to differentiate epileptic seizures from other events. Epilepsy and Behavior. Mar 2014;32142-32144


