

An Integrated Medical CPS for Early Detection of Paroxysmal Sympathetic Hyperactivity

Author

Abstract—Paroxysmal sympathetic hyperactivity (PSH) is an important clinical problem of severe traumatic brain injury (TBI) which incurs approximately 90% of all TBI-related costs. However, current detection approach is hampered by no consensus clinical diagnostic criteria, paroxysmal episode feature with complex manifestations, and already overloaded clinical activities. Thus, these limitations cause delayed recognitions which result in poor clinical outcomes. In this paper, we design an integrated Medical Cyber Physical System (Medical CPS) for early detection of paroxysmal sympathetic hyperactivity patients. First, a formal model is proposed to describe clinical diagnostic criteria. With the formalized models employed, we implement an early detector and integrate it with revised medical device adapters. Our system will monitor patient condition automatically and continuously to relieve medical staff from the heavy burden of clinical activities and provide timely decision supports. Evaluations on 107 clinical cases extracted from medical publications demonstrate the effectiveness and the efficiency of our integrated system.

Keywords—Paroxysmal Sympathetic Hyperactivity; Medical Cyber Physical System; Early Detection

I. INTRODUCTION

Paroxysmal sympathetic hyperactivity (PSH) is a syndrome manifested with: simultaneously increased heart rate, respiratory rate, blood pressure, body temperature and other clinical features, such as severe sweating, posturing, etc. It is primarily caused by traumatic brain injury (TBI). Every year, TBI affects millions of Americans, in which up to 33% severe patients have been reported with PSH [9], [15]. Delayed recognition of PSH may increase morbidity, resulting in long-term disability, even death¹. This situation causes approximately 90% of all TBI-related costs which is worth millions of dollars [12].

To detect PSH, physicians have proposed a number of diverse criteria sets according to clinical experience. In [25] a group of six PSH experts present an approach to calculate the probability of a patient as PSH in order to overcome the inconsistencies between different criteria sets. However, through the discussions with physicians, underdiagnoses and misdiagnoses are particularly common in current clinical diagnosis and treatment. First, although a number of criteria sets have been proposed, no universally accepted clinical diagnostic criteria exist because of the limited evidence of pathophysiology and the evolution of PSH definition. Thus, many physicians learn a little background on it. Furthermore, PSH is a syndrome with complex manifestations to which a lot of conditions have similar appearances. Additionally, the paroxysmal clinical feature requires symptoms to be recurrent and episodic to make a diagnosis. Unfortunately, medical staff are already overloaded at hospitals, and it is impractical to

¹Cases reported in [24] showed that only 7% of PSH patients achieved a moderate or good recovery, however, 45% with severe disability, 30% with persistent vegetative state and 18% with death.

perform frequent clinical monitoring activities. Therefore, all these limitations hamper the awareness of a diagnosis and it is not easy for medical staff to make an early detection of PSH.

In this paper, we design an integrated Medical Cyber Physical System (Medical CPS) for early detection of PSH based on the existing medical knowledge. Efforts to improve medical aspects are out of the scope of this paper. First, because of the lack of pathophysiology, we propose a formal diagnostic criteria model to describe diverse clinical diagnostic criteria sets uniformly. Based on the multiple formalized models with a weight vector predefined by physicians, we provide an algorithm to compute the patient conditions on a set of vital signs. Because some clinical signs and symptoms like sweating cannot be monitored automatically, we allow the physicians to customize the alert constraint on performing manual checks with a configuration file. Then, we implement a monitoring detector with the formalized models generated and employed automatically, and integrate it into a medical CPS. In our system, medical device adapters which are revised on Integrated Clinical Environment [1] will sample patient vital signs. With the patient data parsed and computed by monitoring detector, our system display real-time patient data on the monitor screen and send the results to physicians to relieve medical staff from the heavy burden of clinical monitoring activities and provide timely decision supports. In summary, our work contributes the following:

- 1) **Formal diagnostic criteria models.** We propose a formal model to describe diverse clinical criteria sets uniformly. Physicians can utilize multiple well-known clinical diagnostic criteria sets with a weight vector to monitor a patient. It can reduce physicians' memory load and augment the detection capacity.
- 2) **An integrated Medical CPS.** We implement an early detector in terms of the formal models and integrate it into a medical CPS to observe patient vital signs automatically and continuously, which reduces medical staff burden and provides timely decision supports. In our system, we accumulate patient data and usages of clinical criteria in a format which can be used by data science research. To the best of our knowledge, this is the first study on applying Medical CPS to perform early detection of PSH.
- 3) **Experimental evaluation.** We evaluate our approach on 97 real-world clinical cases extracted from medical publications on PSH and 10 cases from overlap syndromes which have similar clinical features. Compared to the current approach, our work is able to early detect 17.5% more PSH patients with almost the same false positive

rate (3.12%). The result shows the effectiveness and the efficiency of our integrated system.

This paper is organized as follows. Related work is introduced in Section II. In Section III, we describe the background of PSH. Section IV presents a formal diagnostic criteria model to describe diverse clinical criteria and Section V presents our integrated Medical Cyber Physical System for early detection of PSH based on the models. Evaluation on real world clinical cases is given in Section VI and we conclude the paper in Section VII.

II. RELATED WORK

Criteria Evolution: PSH is an important clinical problem which has been studied for more than sixty years. Previously, it has been identified as general dysautonomia, but now it is considered as a specific syndrome and also referred as dysautonomia [4], PAID [8], etc. Since the first diagnostic criteria were presented in 1993 [13], many criteria sets have been proposed such as [8], [26] or modified [14] to help physicians make diagnoses. However, no universally accepted diagnostic criteria exist and the definition of PSH is in evolution. In this paper, we formalize the existing diagnostic criteria in a formal model for early detection of PSH and efforts to improve medical aspects are out of the scope of this paper.

PSH Detection: Currently, detection of PSH is by manual checks according to diagnostic criteria presented in literature or modified. In [5], [11], [22], cases of PSH patients were successfully detected and differentiated from control groups. To our best knowledge, the only “tool” applied to detection of PSH is PSH-AM, which is composed of two components: Diagnosis Likelihood Tool addressing the probability and Clinical Feature Scale assessing the severity to estimate the diagnostic likelihood of PSH [25]. It is not a runnable tool and the authors envisage that PSH-AM would be completed daily by medical staff at a standardized time. However, it is impractical for medical staff to monitor a huge number of vital signs frequently because of tremendous stress and overloaded clinical activities. Moreover, statistics show that records stored in a handwritten format usually cause significant under-reporting of adverse clinical events [6]. The incompleteness and inaccuracy hamper the awareness of PSH and result in underdiagnoses and misdiagnoses. In this work, we design an integrated Medical CPS to automatically monitoring patient vital signs. For some features beyond automatic observation, physicians can customize the constraint on performing manual checks and learn background knowledge from our system. In this way, real-time patient data will be sampled and computed, thus reducing the burden of clinical monitoring activities.

Medical CPS: With the proliferation of measuring devices, a lot of researchers have been paying attention to Medical CPS to provide continuous high-quality care for patients [20]. One of the main applications focuses on anomaly detection. In [17], Ivanov et al. proposed a predictive monitor to detect sharp decreases in $C_a O_2$ caused by a pulmonary shunt in infants. Based on the model characterized by unknown patient specific parameters, they first applied parameter-invariant technique in Medical CPS. Jiang et al. [18] applied runtime verification technique for medical decision support systems

to ensure complex temporal properties in medical guidelines. With medical practice scenarios described in a domain specific language DRTV, event sequences and runtime property verifier automata can be generated to rigorously verify the properties automatically. Data-driven machine learning approaches are also utilized to perform anomaly detection [7], [10]. In contrast to the approaches mentioned above, our work focuses on a more complicated syndrome where some features need manual checks. Moreover, it requires symptoms to be recurrent and episodic to make a diagnosis. However, there are no existing pathophysiologic models or medical guidelines to apply runtime verification technique and not enough data with accurate event annotations to train data-driven models. Therefore, we propose a different strategy, involving physicians to customize the formal models and the constraints to perform manual checks for early detection of PSH.

III. BACKGROUND

Paroxysmal sympathetic hyperactivity is a syndrome where transient nervous system activity occurs manifested with simultaneously increased heart rate, respiratory rate, blood pressure, body temperature and other clinical features, such as severe sweating, posturing and so on. Through the systematic literature review, the duration of each episode is on average 30 minutes and its frequency is on average 3-8 times/day.

To detect PSH, physicians have proposed many clinical diagnostic criteria sets. In Table I, we present three of widely used ones. Same as other criteria, all of the three require paroxysmal episodes manifested with a group of simultaneous patient signs and symptoms to make a diagnosis. While there is a strong agreement on simultaneous and paroxysmal feature, some inconsistencies are illustrated as below. Duration and frequency of episodes are different. In [5], episodes need to persist for more than two weeks, but more than 1 daily episode for at least 3 days in [22]. Another primary inconsistency is on clinical features such that composition and severity of symptoms must be present during an episode. Some features like heart rate, sweating are commonly used, but some vary between criteria sets, like pupillary dilation and muscle tone.

TABLE I: Three widely used diagnostic criteria sets

Reference	Diagnostic Criteria
Baguley [5]	simultaneous, paroxysmal increases in at least five out of the seven reported features (1) heart rate > 120 beats/min, (2) respiratory rate > 30 breaths/min, (3) temperature > 39 °C, (4) systolic blood pressure > 160 mmHg, (5) posturing, (6) dystonia, and (7) sweating, with episodes persisting for at least 2 weeks
Fernandez [14]	short episodes of at least 5 of the 8 following signs and symptoms: (1) tachycardia, (2) arterial hypertension, (3) tachypnea, (4) reduced level of consciousness, (5) muscle rigidity, (6) fever, (7) hyperhidrosis and (8) pupillary dilation; and parameters have to be abrupt and simultaneous, and not otherwise explainable.
Lv [22]	simultaneous occurrence of 5 or more of the following features: (1) tachycardia (heart rate > 120 beats/min), (2) tachypnea (respiratory rate > 30 breaths/min), (3) hyperthermia (temperature > 38.5 °C), (4) hypertension (blood pressure > 160 mmHg), (5) increased muscle tone, (6) decerebrate or decorticate posturing, and (7) excessive sweating at least 1 daily paroxysm that occurs for at least 3 days.

However, severity for the same feature may be distinct, like 38.5°C of body temperature in [22], 39°C in [5] and undefined in [14], respectively.

This situation confuses physicians and hampers the awareness of PSH patients. Unfortunately, it is difficult to create a consensus criteria set because of the lack of evidence in pathophysiology. Therefore, a formal model is desired to describe criteria uniformly and to be easily extended when the definition of PSH evolves.

IV. FORMAL DIAGNOSTIC CRITERIA MODEL

In this section, we introduce the formal diagnostic criteria model to describe the current widely used clinical criteria sets. First, we present the formalization and construction of the model. Then, the semantics and model combination are illustrated.

a) Model Formalization: As discussed in Section III, clinical criteria consist of a group of clinical features with thresholds to confirm an episode, and duration and frequency of the episodes to confirm a diagnosis. However, each of them may differ between criteria sets. Therefore, we propose a formal diagnostic criteria model. Our formal model is specified as an extended automata and the definition is a tuple $\mathcal{M} = \langle S, s_0, s_f, E, C, G, T \rangle$, where:

- S is a set of states: $S = \{s_{1,1}, s_{1,2}, \dots, s_{i,j}\} \cup \{s_0, s_f\}$, where $s_{i,j}$ is a state corresponding to the episode condition of a patient, indicating the occurrence of the j -th episode in the i -th day.
- s_0 is the initial state where a model starts, also labeled as $s_{0,0}$.
- s_f is the final state where a model reaches indicating PSH confirmed, also labeled as $s_{max,0}$.
- E is a set of events: $E = \{e_1, e_2, \dots, e_n\}$, where e_i is an event to trigger a transition. In each event, there is a set of vital signs of a patient labeled as $e.vs$ with a timestamp of these vital signs labeled as $e.tt$.
- C is a tuple of criteria variables: $C = \langle d, F, \xi \rangle$, where d is the least episode duration to make a diagnosis; F is an array of frequencies, where f_i is the least episode frequency of the i -th day; ξ is a predicate to describe the occurrence of an episode under a given event.
- G is a set of guards: $G = \{g_{(s_0, s_{1,1})}, \dots, g_{(s_{d,f}, s_f)}\}$, where $g_{(s_{i,j}, s_{m,n})}$ is a Boolean expression defined in Equation 1 on an event e to guard the transition between two states labeled as $s_{i,j}$ and $s_{m,n}$. A guard is composed of three predicates referring to episode occurrence constraint as $C.\xi$, episode frequency constraint as fre and episode duration constraint dur , respectively. We set $g_{(s_0, s_{1,1})}$ as *True* to initialize our model.
- T is a set of transitions: $T = \{t_1, t_2, \dots, t_n\}$, where t_i is the transition between two states triggered by an event e and guarded by a guard g , as $t_i \in S \times E \times G \times S$.

$$g_{(s_{i,j}, s_{m,n})}(e) = C.\xi(e.vs) \wedge fre_{(s_{i,j}, s_{m,n})}(e) \wedge dur_{(s_{i,j}, s_{m,n})}(e) \quad (1)$$

$$fre_{(s_{i,j}, s_{m,n})}(e) = \begin{cases} \text{True , if } P_1 \vee P_2 \vee P_3 & ^2 \\ \text{False, otherwise.} & ^3 \end{cases} \quad (2)$$

$$dur_{(s_{i,j}, s_{m,n})}(e) = \begin{cases} \text{True , if } P_1 \vee P_4 \vee P_5 & ^3 \\ \text{False, otherwise.} & \end{cases} \quad (3)$$

$$\left\{ \begin{array}{l} P_1 = (i == C.d) \vee (j == C.f_d) \\ \quad \vee (m == max) \vee (n == 0), & ^4 \\ P_2 = (i == m) \vee (j == n - 1), \\ P_3 = (i == m - 1) \vee (j == C.f_i) \vee (n == 1), \\ P_4 = (i == e.tt) \vee (j == n - 1) \vee (m == e.tt), \\ P_5 = (i == e.tt) \vee (j == C.f_i) \\ \quad \vee (m == i + 1) \vee (n == 1). \end{array} \right. \quad (4)$$

In our formal model, criteria variables tuple C is configured by a physician to describe the medical constraints to make diagnoses. Thus, when the pathophysiology develops, our model can easily be extended by modifying the predicate ξ . And other elements can be generated and calculated from C automatically. We label the initial state s_0 as $s_{0,0}$ and the final state s_f as $s_{max,0}$ for a unified equation to calculate the guards. A guard is composed of three parts as defined in Equation 1, which are the constraint on a single occurrence in $C.\xi$, the constraint on episode frequency in Equation 2 and the constraint on episode duration in Equation 3, respectively. In Equation 4, P_1 constrains the transition to the final state s_f . P_2 and P_4 cooperate to constrain the transitions between states in the same day, and P_3 and P_5 for the transitions between two continuous days. We will describe the semantics for these predicates in the *Model Semantics* part.

b) Model Construction: We take the clinical criteria from [22] in Table I as an example to demonstrate the formalization and construction of our formal model. Through the natural language description, we can extract the criteria variables C as

$$C = \langle 3, [1, 1, 1], \xi \rangle$$

indicating that at least 1 episode for each day lasting three days and ξ is a predicate that 5 or more symptoms presented in the table are satisfied. By duration and frequency array defined, we can automatically generate the state set S as

$$S = \{s_{1,1}, s_{2,1}, s_{3,1}\} \cup \{s_0, s_f\}$$

where s_0, s_f are the start and final state and $\{s_{1,1}, s_{2,1}, s_{3,1}\}$ are used as labels corresponding to episode conditions, such as $s_{1,1}$ for the first episode in the first day. We can calculate the guards on transitions between states by Equation 1. We use the guard between $s_{1,1}$ and $s_{2,1}$ as an example. With an event e as a trigger, it depends on the $e.vs$ to judge the episode occurrence constraints. In Equation 4, P_4 is *True* indicating that episode frequency constraint is *True*. Episode duration constraint depends on the $i == e.tt$, because P_6 is *False* and P_5 depends on the $i == e.tt$. Therefore, the guard between $s_{1,1}$ and $s_{2,1}$ is:

$$g_{(s_{1,1}, s_{2,1})}(e) = \xi(e.vs) \wedge True \wedge (e.tt == 1)$$

² P are predicates on i, j, m, n defined in Equation 4

³ P are predicates on i, j, m, n defined in Equation 4

⁴ max is used to label the s_f state. In practice, we assign $max = C.d + 1$

To initialize the model, we set the guard between s_0 and $s_{1,1}$ as *True*. The visualization of this formal model is illustrated in Figure 1.

c) Model Semantics: The execution semantics of our formal diagnostic criteria model can be considered as a labelled transition system [27]. The model starts from the initial state to monitor a patient under a series of events recording the patient vital signs lasting criteria-duration days. When a state transits to another, it implies that an episode occurs and parts of episode duration and frequency constraints are satisfied to make a diagnosis. Finally, if the model reaches the final state, a diagnosis is confirmed because all the constraints are satisfied.

According to the semantics, except for the transition to the final state, valid transitions are specified between j -th and $(j+1)$ -th episode in the same day, and between the last episode in the i th day and the first in the $(i+1)$ -th day. We generate the predicates on episode frequency and duration constraints defined in Equation 4.

As visualized in Figure 1, the formal model of criteria in [22] starts from state s_0 . After the initialization, the model reaches state $s_{1,1}$ to wait for the first episode of the first day. Given a set of events in which each event e contains patient signs vs with a timestamp tt , our model transits between states guarded by the guard in square brackets generated from equations defined above. If the guard between $s_{1,1}$ and $s_{2,1}$ is satisfied on an event e , the model transits to $s_{2,1}$ implying that one episode happens in the first day and our model waits for the first episode of the second day. Finally, if the model arrives at s_f , the patient is confirmed as PSH.

d) Model Combination: For the lack of pathophysiology knowledge, it is hard to create a consensus clinical criteria. One alternative approach is utilizing multiple clinical criteria sets with a weight vector. We present our diagnostic computation of combined models in Algorithm 1. The inputs are a series of events \mathbb{E} in which each event e records patient signs (vs) with a timestamp (tt), a set of clinical criteria \mathbb{C} with the

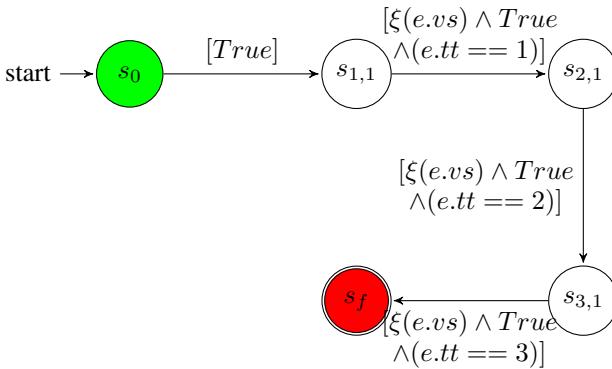


Fig. 1: A simplified visualization of the formal diagnostic criteria model in [22]. Guards are presented in square brackets and transitions with contradiction guard are eliminated. Event e contains the patient vital signs as vs and the timestamp as tt to trigger the transitions. ξ is the predicate that whether the vital signs cross the thresholds defined in the clinical criteria indicating an episode occurrence.

Algorithm 1 Diagnostic computation of combination model

Input: Event set \mathbb{E} , clinical criteria sets \mathbb{C} , weight vector \mathbb{V} and result threshold τ

Output: Diagnosis \mathfrak{D}

```

1: function computation()
2:    $M \leftarrow construct(\mathbb{C})$ 
3:    $score \leftarrow initScore()$ 
4:   for  $m \in M$  do
5:      $r\_tem \leftarrow check(m, \mathbb{E})$ 
6:      $score \leftarrow updateScore(\mathbb{V}, r\_tem)$ 
7:   end for
8:    $\mathfrak{D} \leftarrow judge(score, \tau)$ 
9:   return  $\mathfrak{D}$ 
10: end function
11: function check( $m, \mathbb{E}$ )
12:   for  $E \leftarrow Extract(\mathbb{E}, m \rightarrow d)$  do
13:      $current \leftarrow initCurrent()$ 
14:     for  $e \in E$  do
15:        $current \leftarrow transit(e)$ 
16:       if  $current$  is  $m \rightarrow s_f$  then
17:         return True
18:       end if
19:     end for
20:   end for
21:   return False
22: end function
  
```

weight vector \mathbb{V} and a confirmed threshold τ . The result is computed by function *computation* in statements 1-10. First, we generate formal diagnostic criteria models from clinical criteria by function *construct* in statement 2 as discussed in *Model Construction* part. A variable *score* is initialized to combine the each criteria diagnostic result in statement 3. For each formal model, we compute the diagnostic result under the given events by function *check* in statement 5 and update the *score* according to the weight vector by function *updateScore* in statement 6. As different clinical criteria may propose its specific episode duration, a subset of events will be extracted to meet the time constraint by *Extract* in statement 12. Then, the model and events are used in function *check* illustrated in statements 11-22 to make a diagnosis based on the model semantics. After all the clinical criteria sets have been computed, we make the final diagnosis result according to the relationship between *score* and τ by function *judge* in statement 8.

V. INTEGRATED MEDICAL CPS

In this section, we will present an integrated Medical CPS for early detection of PSH. First, we present an overview of our system and interactions between components: Revised ICE Device Adapter, Monitoring Detector, Model Generator. The following part describes the work-flow of the detector and the implementation of our system.

a) System Structure and Interactions: As illustrated in Figure 2, our system consists of three main components: revised ICE device adapter, model generator and monitoring detector. Modern hospitals have been equipped with a

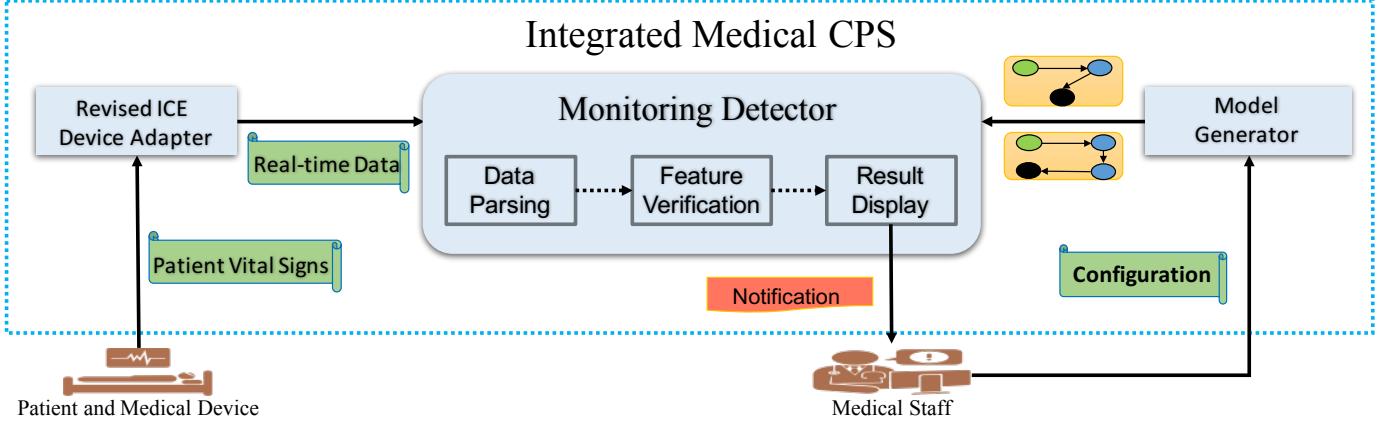


Fig. 2: Components and interactions of our integrated Medical CPS for early detection of PSH

number of advanced medical devices to automatically observe patient vital signs by sensors and display the data on device monitors. However, few of these devices provide specific diagnostic analysis functionality. In order to comprehensively utilize these devices, a medical devise adapter is revised on Integrated Clinical Environment(ICE) to extract patient vital signs from different medical devices. Because some clinical features like sweating and posturing cannot be monitored automatically, we allow physicians to customize the manual check alert constraints with other necessary system properties like data sampling frequencies in configuration files. Then, model generator will automatically parse the configurations into the formal diagnostic criteria models defined in Section IV to minimize the impact from PSH definition evolution. With the formalized models employed, we implement an early monitoring detector. With all the components integrated, our Medical CPS will sample the real-time patient data to relieve medical staff from the heavy burden of monitoring activities and provide timely decision supports.

b) Monitoring Detector Work-Flow: Monitoring Detector is the kernel component in our system, employing the formalized models and monitoring the patient condition to provide decision supports for physicians. We present the workflow details below:

- 1) **Data Parsing:** This step processes patient vital signs. When the predefined sampling period arrives, Monitoring Detector will read patient data from Revised ICE Device Adapter and preprocess all the data, such as checking data rationality, filling in empty values, interpreting data and so on. All the processed data are sent to the next step and stored for further use.
- 2) **Feature Verification:** This step uses the formal diagnostic criteria models generated from clinical criteria to make an early detection of PSH. Even though medical devices can monitor some quantitative clinical signs like blood pressure and heart rate successfully, some qualitative ones are not available like sweating or posturing. Therefore, with the data from the first step, we automatically verify the patient condition to decide whether it needs a further manual verification. If

the patient is safe, the system jumps to Result Display step because of the simultaneity feature where all the signs and symptoms occur simultaneously to confirm an episode which is described in Section III. If the patient needs further clinical tests, medical staff will be alerted. In the manual verification part, medical staff check the patient condition with a checking list automatically generated from diagnostic criteria by our system. With all the vital signs ready, we encapsulate them with history data as events to compute the diagnostic results. The results are sent to the next step.

- 3) **Result Display:** This step displays all the processed data and analysis results. In our system, we update the patient vital signs in the main window timely. It is critical that Medical CPS should not mislead the medical staff. Especially, the reasons for no anomaly are the patient in good condition or system failures. Therefore, we display runtime system state to overcome this limitation. For the computation result, Monitoring Detector will send it to physicians along with analysis procedure in an easy-understanding format to provide timely decision supports. All the results will be saved for review and further use.

c) Implementation: As described above, we implement our system in Java platform which can be deployed on any java-capable computers. We separate our system into several components to decrease the dependency. Currently, we design a medical device adapter based on the revised Integrated Clinical Environment for Phillips IntelliVue MP70 to continually get patient vital signs. We implement a GUI tool for physicians to generate clinical criteria and all the diagnostic criteria generated by our system are stored in a readable file format. The raw data and computation results are stored for further use.

VI. EVALUATION

In this section, we evaluate the performance of our approach by patient data extracted from medical publications. First, we describe the experimental setup, including the clinical diagnostic criteria, the patient data composition and the system



Fig. 3: Deployment for simulation and evaluation.

deployment in lab environment for simulation. Finally, the results and discussions are presented.

A. Experimental Setup

A number of diverse criteria sets are provided by physicians. In our evaluation, we choose two of the most widely used ones [14], [22] described in Table I. In [14], we replace qualitative values “tachycardia, hypertension, tachypnea, fever” with most widely used quantitative values (heart rate > 120 beats/min, blood pressure > 160 mmHg, respiratory rate > 30 breaths/min, temperature > 38.5 °C) to reduce subjective impartiality.

We use patient data extracted from medical publications illustrated in Table II to evaluate our approach. References are recommended by physicians and identified through searches of online databases PubMed [2] by use of the keywords paroxysmal sympathetic hyperactivity, case studies, etc. For references without concrete patient signs and symptoms, but with statistics distributions and episode descriptions, we automatically generate data satisfied the constraints. We use most widely used values counted from references as default values for quantitative features described in natural languages. Precisely, that is 120 beats/min, 160 mmHg, 30 breaths/min, 38.5 °C for increased heart rate, blood pressure, respiratory rate, body temperature or other terms with the similar meaning, respectively. For the qualitative clinical features which are not mentioned in the case description, we treat the patients without those symptoms. We ensure that there are more than 3 episodes in each day for 3 days to meet the episode duration thresholds in the criteria sets. The composition of patient data is presented in Table II which consists of 97 PSH cases and 10 non-PSH cases with clinical features overlapped for false alarm testing.

As presented in Figure 3, our system is deployed in a lab environment to simulate real world monitoring scene. We utilize the two criteria sets, and set weight vector \mathbb{V} as [1, 1] and result threshold τ as 1 to perform the detection as described in Algorithm 1, indicating that a patient is conformed if any of the criteria sets is satisfied. Our system samples patient data every 30 minutes, because the duration of each episode

is on average 30 minutes as describe in Section III. To our best knowledge, there is no runnable tool to detect PSH up to now. The current clinical approach is based on checking patient signs and symptoms manually. Therefore, we invite volunteers with medical experience to manually check patient data every hour⁵ using the same criteria sets to verify the effectiveness and efficiency of our system.

B. Results

All the experiment results are presented in Table II. The first two columns are patient data composition. In the third and fourth columns, we illustrate the results of our work and the average results of the current approach performed by volunteers, respectively. We discuss the the results as follows,

a) Effectiveness: In order to compare our work and current approach, we use *Precision* and *Recall* defined in Equations 5, 6 to compare the effectiveness.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

Precision is a ratio between correctly detected PSH patients and the number of patients assigned as PSH. As illustrated in Table II, for this evaluation cases set, our precision is 93/96 (96.88%) compared to 76/78 (97.44%) for manually checking. Therefore, in terms of the correctness, our approach performs almost the same as manually checking, namely, we will not burden medical staff with much more false alarms than the current approach. *Recall* is a ratio between correctly detected PSH patients and the number of PSH patients in our case set. With 97 cases as PSH patients, we successfully detect 93 of them with a recall as 95.88%, which is better than manually checking with 76/97 (78.35%). In terms of the detection ability, our approach performs better than the current approach. In the medical domain, human safety is the

⁵After the discussions with physicians, we learn that the most frequent manual checking routine in their department is 1 hour. Therefore, we manually check patient data every hour in our evaluation.

TABLE II: Detection results on cases from medical publications

Publication	Cases			PSHMonitor				Manually			
	PSH	N-PSH	Total	TP	TN	FP	FN	TP	TN	FP	FN
Lee [21]	2	0	2	2	0	0	0	1	0	0	1
Hughes [16]	44	0	44	44	0	0	0	37	0	0	7
Baguley [4]	15	0	15	14	0	0	1	13	0	0	2
Blackman [8]	20	0	20	20	0	0	0	16	0	0	4
Deepika [11]	4	0	4	3	0	0	1	1	0	0	3
Lv [22]	6	0	6	6	0	0	0	4	0	0	2
Baguley [3]	6	0	6	4	0	0	2	4	0	0	2
Umbriaco [28]	0	5	5	0	4	1	0	0	4	1	0
Martin [23]	0	5	5	0	3	2	0	0	4	1	0
Summary	97	10	107	93	7	3	4	76	8	2	21

PSH is for patient confirmed as PSH and N-PSH is used for other cases like *Sepsis* and *Malignant hyperthermia*. TP is true positive for PSH cases detected as PSH. TN is true negative for N-PSH cases detected as N-PSH. FP is N-PSH cases misdiagnosed as PSH and FN is PSH cases underdiagnosed.

first factor. With a reasonable false alarm rate, a higher recall indicates a better solution.

b) *Efficiency*: With the same diagnostic criteria, manually checking underdiagnosed 17 cases (17.53%) in all PSH cases than us, which indicates that our approach can perform better in terms of early detection of PSH. In an effort to understand the primary reason, we illustrate the patient data for the first 3 hours in a case from [21] in Figure 4. We note that heart rate (HR), respiration rate (RR) and blood pressure (BP) crossed the threshold at 1.5 hours and 1 hour, respectively. In Section III, the average duration of each episode is 30 minutes. We successfully detected this episode because our system combined the data in the near 30 minutes. However, it is unreasonable for medical staff to observe every patient every 30 minutes. Thus, the volunteers missed the syndromes, precisely high HR and RR, resulting in an underdiagnosis. From another perspective, we calculate the average detection time of true positive groups respectively. The result shows that

manual checks have an average 4.5 hours delay than our system. Along with the paroxysm and complex clinical features, we believe the real world situation is worse. Additionally, during the manual checks process, a volunteer mixed up two criteria sets resulting in under-recognition of a case. Therefore, our approach provides a benefit of steady performance and detects anomalies earlier.

c) *False Alarms*: In all the cases, we have 3 false positives and 4 false negatives. After reviewing the cases manually, we note that for all the false positives, they have crossed the PSH criteria. For example, the patient in [28] was confirmed as Sepsis with BT 39.2 °C, HR 190, RR 35, posturing and poor response to interactions. However, these symptoms are satisfied to be confirmed as PSH. Therefore, the reason for false positives is there are many overlapped clinical features between PSH and Sepsis, resulting in misdiagnoses. It is the future work of us to provide a relevant analysis component to show the possibilities faced with the same clinical features. For the 4 false negatives, we note that all of them are lack of enough vital signs to meet the criteria, which are important to do the computation. These cases are also missed by manual checks. To detect these cases, we can decrease the thresholds of each vital signs. However, it will disturb physicians with many *false positives*. In the future, we will carry out more experiments to balance this issue.

According to the results in this part, it is reasonable to draw the conclusion that our work performs better in early detection of PSH compared with current approach.

C. Discussion

a) *Limitation*: In our evaluation, we used the real-world cases extracted from medical publications. However, some of them lacked concrete data values and were described in human natural language. With the default values generated based on statistics, we filled these cases and sampled in every 30 minutes to make a diagnosis. All these processes may affect evaluation results. Currently, we are working with physicians to deploy our system in hospitals for a further evaluation.

b) *Lessons*: During the discussions with physicians and evaluation procedure, we learn many lessons. (1) Medical CPS is desired. Medical devices have been equipped in hospitals to provide the information to improve the health care. However, during the discussions with physicians, they need more

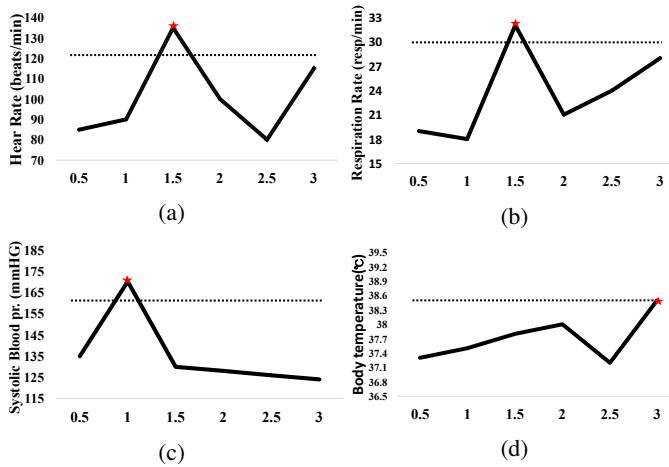


Fig. 4: Patient vital signs for the first three hours in a case from [21], which was missed by manual checking but detected by our system. (a) is heart rate, (b) is respiratory rate, (c) is systolic blood pressure, (d) is body temperature, where red stars indicate outlier points above criteria. Gradations on X axis indicate 30 min intervals. The patient got sweating and posturing during the first and the second hour.

“intelligent tools” to help them make better decisions with reasonable disturbances. (2) False alarms are hard. According to our experience, the most efficient solution is providing an easy access to the thresholds and inviting physicians to define the thresholds. (3) Simple is better. Computer technologies are totally unknown domain to medical staff. Therefore, we need to develop applications with user-friendly interfaces. Moreover, an extensible structure is demanded to cope with the volatile clinical requirements such as the evolution in definitions of diseases.

VII. CONCLUSION

In this paper, we presented an integrated Medical Cyber Physical System for early detection of paroxysmal sympathetic hyperactivity. We proposed a formal diagnostic criteria model to describe diverse widely used clinical criteria. With the models employed, we implemented an early detector and integrated it into a Medical CPS to monitor patient vital signs and help physicians make diagnoses. The evaluation on cases from medical publications shows the effectiveness and the efficiency of our approach. In the future, we will evaluate on real-world patient data in hospitals to strengthen our work.

ACKNOWLEDGMENT

The authors thank

REFERENCES

- [1] ICE: Integrated Clinical Environment, <http://www.mdppnp.org/mdice.html>
- [2] PubMed: a free public interface to MEDLINE, <http://www.ncbi.nlm.nih.gov/pubmed>
- [3] Baguley I J, Heriseanu R E, Gurka J A, et al. Gabapentin in the management of dysautonomia following severe traumatic brain injury: a case series[J]. *Journal of Neurology, Neurosurgery & Psychiatry*, 2007, 78(5): 539-541.
- [4] Baguley I J, Nicholls J L, Felmingham K L, et al. Dysautonomia after traumatic brain injury: a forgotten syndrome?[J]. *Journal of Neurology, Neurosurgery & Psychiatry*, 1999, 67(1): 39-43.
- [5] Baguley I J, Sleva-Younan S, Heriseanu R E, et al. The incidence of dysautonomia and its relationship with autonomic arousal following traumatic brain injury[J]. *Brain injury*, 2007, 21(11): 1175-1181.
- [6] Benson M, Junger A, Fuchs C, et al. Using an anesthesia information management system to prove a deficit in voluntary reporting of adverse events in a quality assurance program[J]. *Journal of clinical monitoring and computing*, 2000, 16(3): 211-217.
- [7] Burgos A, Goi A, Illarramendi A, et al. Real-time detection of apneas on a PDA[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2010, 14(4): 995-1002.
- [8] Blackman J A, Patrick P D, Buck M L, et al. Paroxysmal autonomic instability with dystonia after brain injury[J]. *Archives of Neurology*, 2004, 61(3): 321-328.
- [9] Centers for Disease Control and Prevention. Rates of TBI-related emergency department visits, hospitalizations, and deaths by sexUnited States, 2001-2010. 2014[J]. 2014.
- [10] Convertino V A, Moulton S L, Grudic G Z, et al. Use of advanced machine-learning techniques for noninvasive monitoring of hemorrhage[J]. *Journal of Trauma and Acute Care Surgery*, 2011, 71(1): S25-S32.
- [11] Deepika A, Mathew M J, Kumar S A, et al. Paroxysmal sympathetic hyperactivity in pediatric traumatic brain injury: A case series of four patients[J]. *Autonomic Neuroscience*, 2015, 193: 149-151.
- [12] Faul M, Wald M M, Rutland-Brown W, et al. Using a cost-benefit analysis to estimate outcomes of a clinical treatment guideline: testing the Brain Trauma Foundation guidelines for the treatment of severe traumatic brain injury[J]. *Journal of Trauma and Acute Care Surgery*, 2007, 63(6): 1271-1278.
- [13] Fearnside M R, Cook R J, McDougall P, et al. The Westmead Head Injury Project outcome in severe head injury. A comparative analysis of pre-hospital, clinical and CT variables[J]. *British journal of neurosurgery*, 1993, 7(3): 267-279.
- [14] Fernandez-Ortega J F, Prieto-Palomino M A, Garcia-Caballero M, et al. Paroxysmal sympathetic hyperactivity after traumatic brain injury: clinical and prognostic implications[J]. *Journal of neurotrauma*, 2012, 29(7): 1364-1370.
- [15] Hinson H E, Sheth K N. Manifestations of the hyperadrenergic state after acute brain injury[J]. *Current opinion in critical care*, 2012, 18(2): 139-145.
- [16] Hughes J D, Rabinstein A A. Early diagnosis of paroxysmal sympathetic hyperactivity in the ICU[J]. *Neurocritical care*, 2014, 20(3): 454-459.
- [17] Ivanov R, Weimer J, Simpao A, et al. Early detection of critical pulmonary shunts in infants[C]//Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems. ACM, 2015: 110-119.
- [18] Jiang Y, Liu H, Kong H, et al. Use runtime verification to improve the quality of medical care practice[C]//Proceedings of the 38th International Conference on Software Engineering Companion. ACM, 2016: 112-121.
- [19] Konstantas D, Herzog R. Continuous monitoring of vital constants for mobile users: the MobiHealth approach[C]//Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE. IEEE, 2003, 4: 3728-3731.
- [20] Lee I, Sokolsky O. Medical cyber physical systems[C]//Proceedings of the 47th Design Automation Conference. ACM, 2010: 743-748.
- [21] Lee S, Jun G W, Jeon S B, et al. Paroxysmal sympathetic hyperactivity in brainstem-compressing huge benign tumors: clinical experiences and literature review[J]. SpringerPlus, 2016, 5(1): 1.
- [22] Lv L Q, Hou L J, Yu M K, et al. Hyperbaric oxygen therapy in the management of paroxysmal sympathetic hyperactivity after severe traumatic brain injury: a report of 6 cases[J]. *Archives of physical medicine and rehabilitation*, 2011, 92(9): 1515-1518.
- [23] Martin S N, Vane E A. Malignant hyperthermia: a case study[C] Seminars in perioperative nursing, 2000, 9(1): 27-36.
- [24] Perkes I, Baguley I J, Nott M T, et al. A review of paroxysmal sympathetic hyperactivity after acquired brain injury[J]. *Annals of neurology*, 2010, 68(2): 126-135.
- [25] Perkes I E, Menon D K, Nott M T, et al. Paroxysmal sympathetic hyperactivity after acquired brain injury: a review of diagnostic criteria[J]. *Brain Injury*, 2011, 25(10): 925-932.
- [26] Rabinstein A A. Paroxysmal sympathetic hyperactivity in the neurological intensive care unit[J]. *Neurological research*, 2013.
- [27] Tretmans J. Conformance testing with labelled transition systems: Implementation relations and test generation[J]. *Computer networks and ISDN systems*, 1996, 29(1): 49-79.
- [28] Umbrico F, Andreoni C. Pediatric sepsis: A case study[J]. *Advanced emergency nursing journal*, 2013, 35(4): 303-313.