Automatic Detection of Acute Kidney Injury Episodes from Primary Care Data

Santosh Tirunagari*, Simon C Bull*, Norman Poh *
* Department of Computer Science
*University of Surrey, Guildford, Surrey, United Kingdom GU2 7XH.
{santosh.tirunagari, s.c.bull, n.poh}@surrey.ac.uk

Abstract—**Acute kidney injury (AKI) is characterised by a rapid deterioration in kidney function, and can be identified by examining the rate of change in a patient's estimated glomerular filtration rate (eGFR) signal. Due to the potentially irreversible nature of the damage AKI episodes cause to renal function, their detection plays a significant role in predicting a kidney's effectiveness. Although algorithms for this are available for patients under constant monitoring, e.g. inpatients, their applicability to primary care settings is less clear as the eGFR signal often contains large lapses in time between measurements. However, waiting for hospital admittance is undesirable, as detecting AKI early can help to mitigate the degradation of kidney function and the associated increase in morbidity and mortality. Traditionally, a clinician in a primary care setting would manually identify AKI episodes within an eGFR signal. While this approach may work for individual patients, the time consuming nature of it precludes quick large-scale monitoring. We therefore present two alternative automated approaches for detecting AKI: as the outlier points when using Gaussian process regression and using a novel technique called Surrey AKI detection algorithm (SAKIDA). Using SAKIDA, we can identify the exact number of AKI episodes a patient experiences with an accuracy of 70%, when evaluated against the performance of human experts.**

I. INTRODUCTION

Estimated glomerular filtration rate (eGFR) is a derived measurement that characterises the effective functioning of a kidney. It plays a central role in both the management of people with chronic diseases and epidemiology research involving longitudinal data [1], [2]. While an individual's eGFR signal will in general display a gentle downward trend, it may also exhibit brief periods of drastic decrease and recovery. These periods, or acute kidney injuries (AKIs), are accompanied by a substantial decrease in the effectiveness of the kidney, leading to a failure to maintain a normal electrolyte balance, uremia, systemic organ damage and increased mortality [3]–[5]. AKI can occur in any clinical setting, and costs the NHS in England approximately 1 billion per year [6].

For most patients under primary care, their eGFR is measured infrequently and potentially irregularly. However, the kidney function of inpatients is monitored more closely, partly due to the increase in AKI risk factors once admitted to hospital, such as the greater rate of sepsis [7]. Given this closer monitoring and the greater prevalence of AKI amongst inpatients [8], algorithms have been developed that can accurately identify AKI within this population in real-time [9], [10]. While this will improve outcomes for inpatients, approximately two-thirds of AKI episodes are community acquired, i.e. occur outside hospitals, and therefore while the patient is under primary care [11]. Given the better short- and long-term outcomes, including reduced mortality, for patients with community acquired AKI [11], [12], robust detection of AKI within a primary care setting could substantially improve patient outcomes.

Despite the clear benefits derived from identification of AKI in the community, algorithms have not been developed to specifically address this. At present, a GP must manually identify suspect eGFR measurements and evaluate whether they warrant follow up as a potential indicator of AKI. While this may work for individual patients, it is both a task best performed by specialists and too time consuming to allow the quick evaluation of a large number of patients. Although the algorithms developed for hospitalised patients can (and will) be extended to primary care, their reliance on a baseline value estimated from recent measurements makes identification of AKI harder in primary care settings, where eGFR signals often contains large lapses in time between measurements.

In this study, three methods for identifying AKI episodes within a patient's eGFR signal are evaluated against the performance of clinical experts. The first method was developed by NHS England to provide real-time detection and diagnosis of AKI in patients across the National Health Service in England [9]. However, applying it to primary care data can be problematic as it relies on calculating a baseline eGFR value for a patient from their measurements over the last 0-7 or 8-365 days. Given that a patient in primary care may only have a few measurements in the last year, and likely none in the last week, there may be too few measurements available to determine a reliable baseline. Similarly, if the patient experienced an AKI episode in the last year, then the baseline calculation will take into account this period of decreased eGFR, thereby reducing the baseline below its true value and making it more difficult to identify AKI episodes. Additionally, as the algorithm is intended to serve as a real-time warning system, without alteration it is unsuitable for retrospective analysis of an eGFR signal, for example when attempting to identify all AKI episodes a patient has experienced, which may be beneficial for understanding their risk of susceptibility to further degradation in kidney function or kidney failure.

The second and third methods evaluated here, Gaussian process regression (GPR) and the novel Surrey AKI detection algorithm (SAKIDA), are intended not only to be able to provide real-time warnings, but also to be more applicable to a primary care setting and be inherently suitable for retrospective analysis. Previously we have used GPR to classify the overall trend in an eGFR signal [13]. In a similar manner, it can be used to identify AKI episodes by modelling an eGFR signal without them, and then treating any measurements less than the lower confidence limit as AKI episodes. This approach also has the advantage of giving a global trend for a patient's eGFR signal, thereby obviating the need to calculate a time-windowed baseline against which AKI episodes are identified. The final method evaluated was the novel SAKIDA method. Similar to the GPR-based approach, SAKIDA enables retrospective analysis of an entire eGFR signal and can handle the irregularity of primary care eGFR signals. An example of these algorithms applied to a patient's entire eGFR signal can be seen in Figure 1.

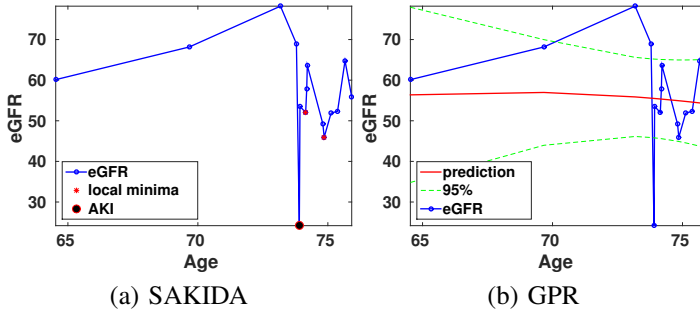Given the need for algorithms that can identify AKI episodes in

Fig. 1. An abrupt change in the eGFR signal can be flagged as an AKI episode. AKI detection using (a) the Surrey AKI detection algorithm and (b) Gaussian process regression. The eGFR measurements outside the confidence bounds indicate the occurrence of AKI.

primary care settings and tools to assist non-specialist GPs in identifying eGFR abnormalities, including AKI episodes, our contributions can be summarised as follows:

1) Proposal and evaluation of novel AKI detection algorithms suitable for primary care data. While outlier detection using GPR is not new, formulating AKI identification as an outlier detection task is. Additionally, in order to generalise the confidence intervals reliably in our application we tune the GPR parameters in a supervised manner, something which to the best of our knowledge has not been reported before.
2) We introduce SAKIDA, a medically inspired algorithm specifically designed to identify AKI episodes in eGFR signals from primary care data.
3) Improved understanding for GPs. Utilising the ability of GPR to provide both predicted values and confidence interval bounds means that GPs can visualise an eGFR signal as stable or unstable [13], while also identifying outlier measurements as potential AKI episodes. Alternatively, a hybrid algorithm consisting of GPR and SAKIDA could be used to visualise eGFR signal stability (GPR) and also localise AKI episodes (SAKIDA).

The organisation of the paper is as follows: In section II, we present and analyse our dataset and illustrate the demographics of the data. In section III, we present the GPR and SAKIDA methodology used in this paper. Experiments and results are discussed in section IV. Finally, in section V, we draw conclusions and discuss their relevance.

II. THE 'HANNAH' DATASET

The dataset used in this work contains the eGFR time signal of 488 patients (275 (56.4%) male and 213 (43.6%) female) treated at East Kent University Hospital, and was collected as part of a study seeking to understand the characteristics of acute kidney injury and its impact on chronic kidney disease. Each patient's eGFR time signal was labelled with the number of AKI episodes experienced: 0, 1, 2, 3 or 4. The distribution of the number of AKI episodes across the 488 patients is given in Table I and Figure 2(h). In total, there were 10,873 eGFR measurements across the 488 patients, with approximately 95% between the values of 25 and 95 mL/min/ $1.73m^2$ and occurring in patients between the ages of 60 and 90. Figure 2 summarises the main characteristics of the dataset.

TABLE I. COUNT OF AKI EPISODES IN THE DATASET

AKI Episodes	0	1	2	3	4	Total
Percentage	55.7377	29.7131	10.2459	3.4836	0.8197	100
Count	272	145	50	17	4	488

III. METHODS

A. Gaussian Process Regression

A Gaussian process (GP) is a collection of random variables, any finite collection of which has a joint Gaussian distribution [14]. GPs characterise the probability distribution over functions by a specified mean function $\bar{d}(\mathbf{x})$ and a covariance function $k(\mathbf{x}, \mathbf{x}')$ [15].

To describe a real process $f(\mathbf{x})$ as a GP, we write: $f(\mathbf{x}) \sim \mathcal{GP}\left(\bar{d}(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right)$. Here $\bar{d}(\mathbf{x}) = \mathrm{E}\{d(\mathbf{x})\}$ and $k(\mathbf{x}, \mathbf{x}') = \mathrm{E}\{(d(\mathbf{x}) - \bar{d}(\mathbf{x}))(d(\mathbf{x}') - \bar{d}(\mathbf{x}'))\}$, where $E\{g(\mathbf{x})\}$ denotes the expectation of a function g over the variable $\mathbf{x}$.

Given a set of measurements $\mathcal{D} = \{(\mathbf{x}_i, d(\mathbf{x}_i)\}_{i=1}^N$, the goal is to estimate the true output $d(x^*)$ at an arbitrary $\mathbf{x}^*$ given the relation:

$$d_{obs}^i = d(\mathbf{x_i}) + \epsilon(\mathbf{x_i}); \quad \epsilon(\mathbf{x_i}) \sim \mathcal{N}(0, \sigma_n^2) \tag{1}$$

The prior distribution of the observed target $d(\mathbf{x})$ is given by:

$$d(\mathbf{x}) \sim \mathcal{N}\left(\bar{d}(\mathbf{x}), k(\mathbf{X}, \mathbf{X}'))\right) \tag{2}$$

where $k(\mathbf{X}, \mathbf{X}')$ is the covariance matrix between all pairs of training points. A squared exponential kernel was used to determine the covariance matrix, where the squared exponential kernel (Gaussian/RBF) is defined as: $\kappa(\mathbf{x}, \mathbf{x}') = \exp(\frac{-(\mathbf{x} - \mathbf{x}')^2}{2\gamma^2})$, with γ the length scale of the kernel.

The distribution of the estimated mean value $\mathbf{d}(\mathbf{x})$ can be computed by conditioning on the training data to get $p(d(\mathbf{x})|\mathbf{x}^*, D)$. The joint distribution over $d(\mathbf{x})$ and the new datapoint $\mathbf{x}^*$ is computed using:

$$\begin{bmatrix} \mathbf{d}_{obs} \\ d(\mathbf{x}^*) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \bar{d}(\mathbf{X}) \\ d(\mathbf{x}^*) \end{bmatrix}, \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I & K(\mathbf{X}, \mathbf{x}^*) \\ K(\mathbf{x}^*, \mathbf{X}) & k(\mathbf{x}^*, \mathbf{x}^*) \end{bmatrix}\right)$$

Here, $\mathbf{d}_{obs} = \left(d_{obs}^1, \ldots, d_{obs}^N\right)^T$; $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, $\bar{\mathbf{d}}(\mathbf{X})_i = \bar{d}(\mathbf{x}_i)$, and $K(\mathbf{X}, \mathbf{X})_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$.

The conditional distribution of Equation 2 allows us to get the distribution of $d(\mathbf{x}^*)$ with the following mean and covariance:

$$d(\mathbf{x}^*) \sim \mathcal{N}\left(E\{d(\mathbf{x}^*)\}, var\{d(\mathbf{x}^*)\}\right) \tag{3}$$

where

$$E\{d(\mathbf{x}^*)\} = \underbrace{\bar{d}(\mathbf{x}^*)}_{prior} + K(\mathbf{x}^*, \mathbf{X})\left[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I\right]^{-1}(\mathbf{d}_{obs} - \bar{\mathbf{d}}(\mathbf{X}))$$

$$var\{d(\mathbf{x}^*)\} = \underbrace{k(\mathbf{x}^*, \mathbf{x}^*)}_{prior} - K(\mathbf{x}^*, \mathbf{X})\left[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I\right]^{-1} K(\mathbf{X}, \mathbf{x}^*)$$

The GPstuff toolbox [1] [16], was used for modelling the eGFR time signals, with the hyperparameters for the squared exponential kernel tuned using maximum a posteriori estimates.

1) Detecting AKIs as GPR outliers: Let $\mathcal{G}$ be the set containing the eGFR signals from all 488 patients and $[\mathcal{G}^0, \mathcal{G}^1, \mathcal{G}^2, \mathcal{G}^3, \mathcal{G}^4]$ denote the sets of eGFR signals containing 0, 1, 2, 3 and 4 AKI episodes as labelled by the human experts. In order to obtain the right confidence intervals to detect the AKI episodes as outliers, we trained the GPR using those patients that have no AKI present in their eGFR signals. Therefore, the dataset was partitioned into training and test sets as follows: $\mathcal{G}_{train} \equiv [\mathcal{G}_{10\%}^0]$ and $\mathcal{G}_{test} \equiv [\mathcal{G}_{90\%}^0, \mathcal{G}^1, \mathcal{G}^2, \mathcal{G}^3, \mathcal{G}^4]$. The GPR was therefore trained using 10% of the 272 patients

---

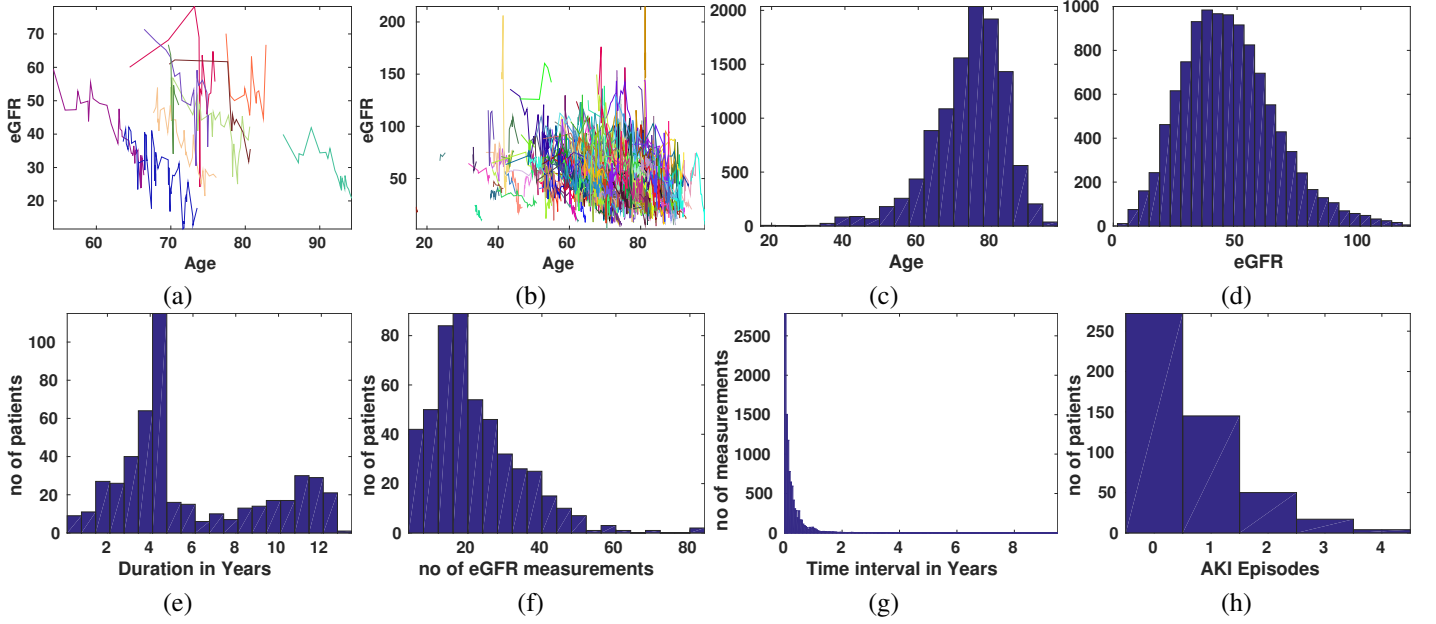[1] http://research.cs.aalto.fi/pml/software/gpstuff/

Fig. 2. Patients' eGFR signals are observed at irregular time intervals and over different age ranges. Each colour represents a single patient (a) for a subset of data and (b) the whole dataset. Dataset characteristics: distributions of the (c) ages over which patients had eGFR measurements recorded, (d) eGFR measurement values, (e) duration of time over which all of a patient's eGFR measurements were recorded, (f) number of eGFR measurements recorded per patient, (g) time intervals between consecutive eGFR measurements, (h) number of AKI episodes per patient.

without an AKI episode. We set the initial parameters $\mathcal{U}$, and trained the regressor to obtain the updated hyperparameters $\hat{\mathcal{U}}$ using maximum a posteriori estimates. Later, using these updated hyper parameters $\hat{\mathcal{U}}$, we predict the number of AKI episodes for each eGFR signal in $\mathcal{G}_{test}$. The mean $E\{\mathbf{x}^*\}$ and variance $var\{\mathbf{x}^*\}$ for each eGFR signal is then calculated, and the 95% confidence intervals calculated as $E\{d(\mathbf{x}^*)\} \pm 2 \times \sqrt{var\{d(\mathbf{x}^*)\}}$. The eGFR datapoints lying outside these confidence intervals are then determined to have been measured during an AKI episode.

B. Surrey AKI Detection Algorithm (SAKIDA)

Let $\mathbf{x} \equiv [x_1, x_3, \ldots, x_m]$ be an eGFR signal of length m. Since, we are interested in finding the local minima (i.e., the points where there is a sudden drop in eGFR signal), we calculate the difference between adjacent elements, of the signal. For this purpose, 1 and -1 are appended to the left and right side of eGFR signal respectively.

$$\hat{\mathbf{x}} \equiv [1, x_1, x_2, \ldots, x_m, -1]. \quad (4)$$

The differences between adjacent datapoints are then calculated as:

$$\widetilde{\mathbf{x}}_{[i]} \leftarrow \hat{x}_{[i+1]} - \hat{x}_{[i]}, \ \forall i \in \{1, \ldots, m+1\}. \quad (5)$$

Any index $\mathbf{i}$ in $\widetilde{\mathbf{x}}$ can then be considered a local minima if $\widetilde{\mathbf{x}}_{[i]}$ is positive and $\widetilde{\mathbf{x}}_{[i-1]}$ negative. Let $\mathcal{J} = [j_1, j_2, \ldots, j_r]$ be the indices of these local minima, i.e.

$$\mathcal{J} \equiv \{i | \widetilde{\mathbf{x}}_{[i]} \leq 0 \wedge \widetilde{\mathbf{x}}_{[i+1]} > 0\}, \ \forall i \in \{1, \ldots, m+1\}. \quad (6)$$

In order to calculate the rate of change for the first local minima, a 1 is prepended to $\mathcal{J}^2$. The rate of change $\mathbf{r}_j$ at a local minima $\mathbf{x}_{[\mathcal{J}_{[j]}]}$ is then calculated as the ratio of the median value of the datapoints in $[\mathbf{x}_{[\mathcal{J}_{[j-1]}]}, \ldots, \mathbf{x}_{[\mathcal{J}_{[j]}-1]}]$ to the value of $\mathbf{x}_{[\mathcal{J}_{[j]}]}$.

$$\mathbf{r}_j = median\big(x_{[\mathcal{J}_{[j-1]}]}, \ldots, x_{[\mathcal{J}_{[j]}-1]}\big)/x_{[\mathcal{J}_{[j]}]}, \ \forall j \in \{2, \ldots, |\mathcal{J}|\}. \quad (7)$$

where $x_{[\mathcal{J}_{[i]}]}$ returns the value of x (the original sequence) at index $\mathcal{J}_{[i]}$.

Finally, the indices $\mathcal{J}_{AKI}$ at which the rate of change is greater than or equal to a threshold Δ are recorded as being an AKI episode.

$$\mathcal{J}_{AKI} \equiv \big\{\mathcal{J}_{[j]} \big| r_j \geq \Delta\big\}, \forall j \in \{2, \ldots, |\mathcal{J}|\} \quad (8)$$

The overall processing steps are shown in Algorithm 1.

C. Groundtruth & Evaluation Criteria

Each eGFR time-signal in the dataset was annotated by a nephrologist to record whether it contained 0, 1, 2, 3 or 4 AKI episodes. Taking y to be the groundtruth and $\hat{y}$ the labels predicted by an algorithm, we evaluated the performance of our proposed algorithms against the experts in the following ways: (i) percentage of correctly labelled eGFR signals ($y = \hat{y}$), (ii) percentage of signals where the algorithm predicted fewer AKI episodes than the experts ($y > \hat{y}$), (iii) percentage of signals where the algorithm predicted a greater number of AKI

---

[2]Note that $\mathcal{J}$ is an ordered list. Since we are describing an algorithm here, we do not explicitly distinguish a list from a set.

Algorithm 1: SAKIDA

Input : An eGFR signal $\mathbf{x}$ & threshold Δ
Output: The indices of AKI episodes $\mathcal{J}_{AKI}$

begin
    /* Append 1 and −1 to the input **x** */
    $\hat{\mathbf{x}} \equiv [1, x_1, x_2, \ldots, x_m, -1]$
    /* Calculate adjacent differences */
    $\widetilde{\mathbf{x}}_{[i]} \leftarrow \hat{x}_{[i+1]} - \hat{x}_{[i]},\ \forall i \in \{1, \ldots, m+1\}$
    /* Obtain indices of local minima */
    $\mathcal{J} \equiv \left\{ i \middle| \widetilde{x}_{[i]} \leq 0 \wedge \widetilde{x}_{[i+1]} \geq 0 \right\},\ \forall i \in \{1, \ldots, m+1\}$
    /* Add index of the first eGFR datapoint */
    $\mathcal{J} \equiv [1, \mathcal{J}]$
    /* Calculate the rates of change */
    for $j \in \{2, \ldots, |\mathcal{J}|\}$ **do**
        $\mathbf{r}_j \leftarrow median\left(x_{[\mathcal{J}_{[j-1]}]}, \ldots, x_{[\mathcal{J}_{[j]}-1]}\right) / x_{[\mathcal{J}_{[j]}]}$
        /* Update rates of change $\geq \Delta$ */
        $\mathcal{J}_{AKI} \leftarrow \mathcal{J}_{AKI} \cup \left\{ \mathcal{J}_{[j]} \middle| r_j \geq \Delta \right\}$
    end
    /* Return rates of change $\geq \Delta$ */
    return $\mathcal{J}_{AKI}$
end

episodes than the experts ($y < \hat{y}$) and (iv) the accuracy with which the algorithm was able to label signals subdivided by the number of AKI episodes ($y_{[j]} = \hat{y}_{[j]},\ \forall j \in \{0, 1, 2, 3, 4\}$).

IV. EXPERIMENTS AND RESULTS

- **Performance of GPR on the training set:** Prior to performing outlier detection for identifying AKI episodes, the performance of the GPR algorithm when detecting no AKI episodes was evaluated. For this purpose, only those patients without an AKI episodes in their eGFR signal were considered. The GPR algorithm was trained on 10% of these patients, and evaluated on the remainder.

- **Effect of the threshold on SAKIDA performance:** According to the NICE guidelines, an AKI episode occurs when the ratio of the reference value calculated from prior measurements to the current measurement is ≥ 1.5. However, this threshold is designed for serum creatinine levels, not the eGFR values derived from them. In order to determine how this threshold influences the algorithms within the eGFR space, thresholds within the set $\{1.1, 1.2, \ldots, 2.0\}$ were evaluated, with the expectation that a threshold around 1.5 will perform best.

- **Comparison of SAKIDA with GPR and NHS England:** In order to determine the algorithm most suitable for detecting AKI episodes within a patient's entire eGFR signal, the performance of SAKIDA was compared to that of GPR and the NHS England algorithm in terms of their ability to mimic the labelling of the expert nephrologists.

A. Performance of GPR on the training set

As the detection of AKI episodes via GPR is reliant on AKI episodes being outliers, and the confidence intervals for determining this are dependent on the model learnt using patients with no AKI episodes, it is worth first assessing the performance of GPR on these patients. Since 10% of the patients with no AKI episodes were to be used for training the final GPR AKI detection model, the same set of patients was used when evaluating the ability of GPR to identify patients without AKI episodes. The models trained via 5-fold cross validation using this set of patients achieved a mean accuracy of $89\% \pm 1.2$.

B. Effect of the threshold on SAKIDA performance

The results of varying the value of the Δ parameter to take on values in the set $\{1.1, 1.2, \ldots, 2.0\}$ can be seen in Table II. The conjecture that a threshold of $\Delta \geq 1.5$ should perform well is borne out by the results, as it results in the algorithm predicting the number of AKI events most accurately (69.67%). However, it can be seen from Figure 3 that whereas SAKIDA's performance when evaluated for $y = \hat{y}$ peaks at a threshold of 1.5, the performance of the NHS England algorithm continues to rise even when the threshold reaches 2.0.

TABLE II.   TUNING OF THRESHOLD ON SAKIDA

$\Delta \geq$	$y = \hat{y}$	$y > \hat{y}$	$y < \hat{y}$
1.1	19.4672	0.6148	79.9180
1.2	42.4180	2.6639	54.9180
1.3	61.4754	8.4016	30.1230
1.4	69.4672	16.3934	14.1393
1.5	**69.6721**	**22.1311**	**8.1967**
1.6	69.0574	26.2295	4.7131
1.7	66.1885	30.5328	3.2787
1.8	65.7787	32.1721	2.0492
1.9	63.5246	35.0410	1.4344
2.0	63.1148	35.8607	1.0246

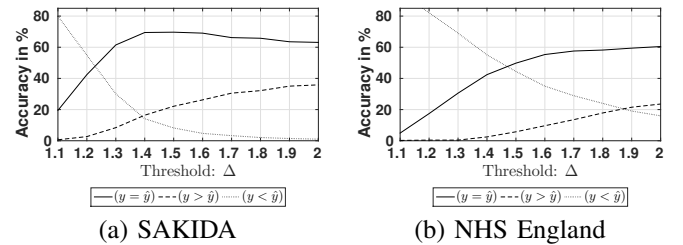(a) SAKIDA   (b) NHS England

Fig. 3.   Threshold Δ parameter tuning for SAKIDA and NHS England.

C. Visual comparison of SAKIDA & GPR

Figure 4 shows example eGFR signals with 1, 2, 3 and 4 AKI episodes, as identified by nephrologists, along with the AKI detection performed using GPR and SAKIDA. In all four cases the abrupt decrease in eGFR that indicates an AKI episode results in measurements that lie outside the confidence interval bounds of the fitted GPR model, indicating the GPR model has accurately identified the time periods over which the AKI episodes occurred. However, in the case of the signal with 3 and 4 AKI episodes, the patients have had

multiple measurements taken during the period in which the AKI episode is influencing the eGFR values, either as they are decreasing or before they have fully recovered, and therefore simply counting the number of outliers does not accurately indicate the number of AKI episodes.
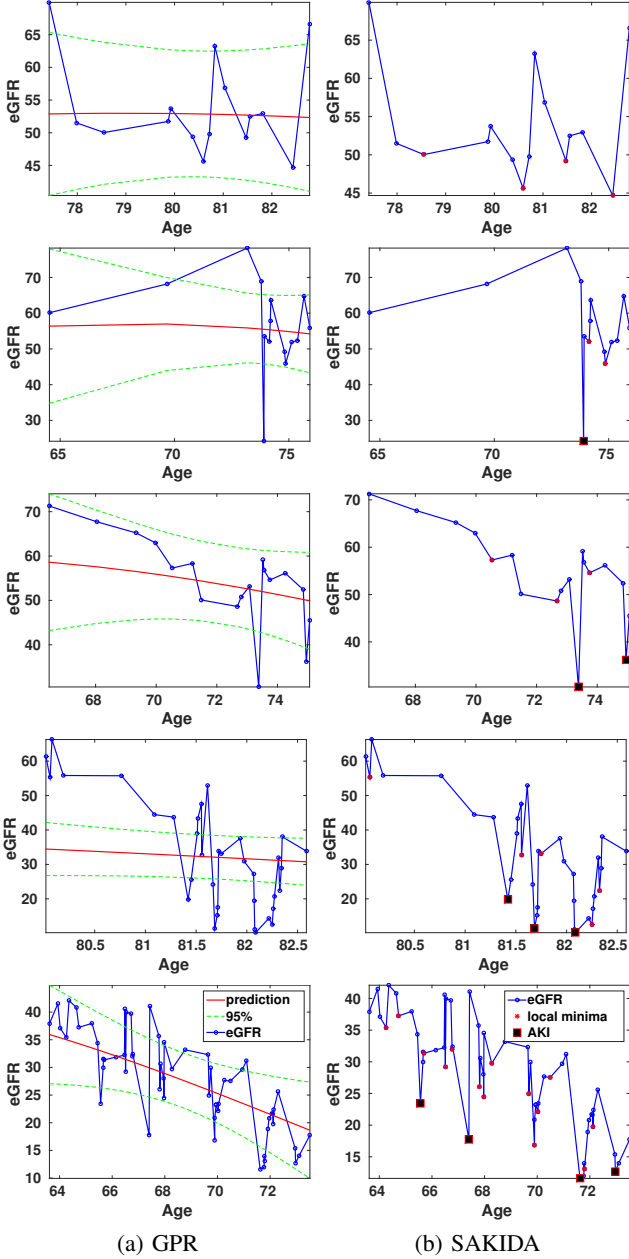

(a) GPR   (b) SAKIDA

Fig. 4. Exemplar eGFR signals with 0, 1, 2, 3 and 4 AKI episodes, top to bottom, along with the results of the AKI detection performed using (a) GPR and (b) SAKIDA

D. Performance comparison of SAKIDA, NHS England & GPR

SAKIDA was able to identify eGFR signal with no AKI episodes with an accuracy of 90.44%, while GPR achieved 83.82% and NHS England 73.53%. SAKIDA also detected no more than 4 AKI episodes per signal, in agreement with the expert's classifications, while GPR and NHS England

detected more in 67 and 99 patients respectively (Figure 5). Interestingly, despite performing worst when detecting 0, 1, 2 and 4 AKI episodes, NHS England detected eGFR signals with 3 AKI episodes with a greater accuracy than that of the other methods tested (17.64%).

TABLE III. PERFORMANCE OF SAKIDA, NHS ENGLAND & GPR IN DETECTING THE NUMBER OF AKI EPISODES.

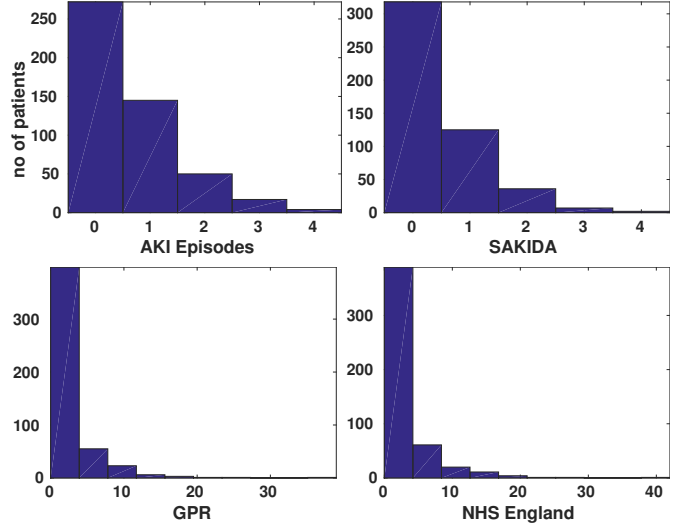| Algorithms	AKI episodes detection accuracy in %				
	0	1	2	3	4
NHS	73.53	22.07	16.00	17.65	0.00
GPR	83.82	28.97	20.00	5.88	0.00
SAKIDA	90.44	53.10	30.00	5.88	25.00

Fig. 5. AKI episodes: Historgrams for the groundtruth and the AKI episode predictions using SAKIDA, GPR and NHS England algorithms.

E. Hybrid visualisation of SAKIDA & GPR

A hybrid system consisting of GPR and SAKIDA could improve clinicians' understanding of the eGFR signal. As GPR is able to provide both predicted values and confidence interval bounds, GPs can help to visualise and classify an eGFR signal as stable or unstable, thereby enabling clinicians to identify patients who are likely to need further monitoring (those with unstable eGFR signals). Additionally, the outliers can be used to identify regions of time where an AKI episode is likely to be occurring or have occurred, while SAKIDA can be used to localise the AKI episodes as shown in Figure 6.

V. CONCLUSIONS & DISCUSSIONS

Considering the severity of the complications arising from it, and the benefits of diagnosing it pre-hospital admission, little attention has been given to improving the detection of AKI within the primary care setting. As most patients will have their health managed in their general practices for the majority of their lives, with hospital admittance possibly not occurring until long after an AKI episode, robust algorithms for real-time warning and retrospective identification of AKI episodes in primary care data are needed. However, primary care data presents a unique challenge, especially when compared with
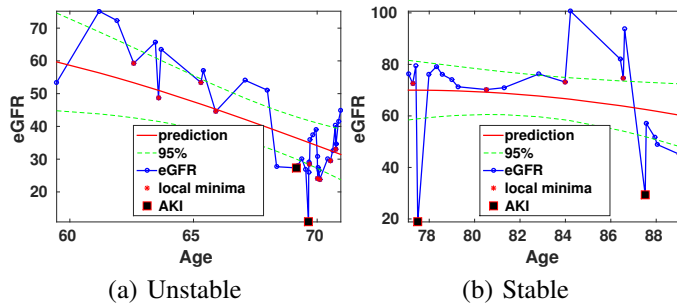
Fig. 6. Hybrid visualisation of SAKIDA and GPR algorithms. (a) unstable and (b) stable eGFR signal trends with GPR model outliers that include AKI episodes. SAKIDA localises the AKI episodes in both signals.

data collected in hospitals. Unlike the constant monitoring of an inpatient, patients in primary care will likely have less frequent and more irregular eGFR measurements along with a greater variability in eGFR values due to weaker controls over pre-test conditions. As demonstrated here, this can prove problematic for algorithms, such as the NHS England algorithm, that rely on the underlying biomedical condition being accurately captured via repeated and frequent measurements. Alternatively, algorithms that treat the AKI episode detection problem as one of signal processing, such as SAKIDA and GPR, fare better.

The main conclusion that can be drawn from the results is that SAKIDA performs better than both GPR and the NHS England algorithms, not only due to the greater accuracy with which it identifies patients with no AKI episodes, but also because it better matches the expert's classifications overall. This indicates that GPR and the NHS England algorithms are likely to be unsuitable for retrospective identification of AKI episodes in eGFR signals from primary care, and are instead more suitable as real-time alert systems. However, while a medical alert system that errs on the side of caution is likely preferable to one that fails to identify risks, our results indicate that both GPR and the NHS England algorithms simultaneously over and under predict the number of AKI episodes. Both methods identify more patients than the experts did as having no AKI episodes, while for 67 patients in the case of GPR and 99 in the case of NHS England too many AKI episodes are detected. Given that SAKIDA closely matches the expert's predictions and can be used for both retrospective analysis and real-time alerts, we believe it should be the method of choice when identifying AKI episodes within primary care data.

REFERENCES

[1] N. Poh and S. de Lusignan, "Calibrating longitudinal egfr in patience records stored in clinical practices using a mixture of linear regressions," in *International Workshop on Pattern Recognition for Healthcare Analytics, 21st International Conference on Pattern Recognition (ICPR)*, 2012.

[2] N. Poh, S. Tirunagari, and D. Windridge, "Challenges in designing an online healthcare platform for personalised patient analytics," in *Computational Intelligence in Big Data (CIBD), 2014 IEEE Symposium on*. IEEE, 2014, pp. 1–6.

[3] W. Druml, "Systemic consequences of acute kidney injury." *Current opinion in critical care*, vol. 20, no. 6, pp. 613–619, Dec 2014.

[4] S. Faubel and P. B. Shah, "Immediate consequences of acute kidney injury: The impact of traditional and nontraditional complications on mortality in acute kidney injury." *Advances in chronic kidney disease*, vol. 23, no. 3, pp. 179–185, May 2016.

[5] C.-C. Shiao, P.-C. Wu, T.-M. Huang, T.-S. Lai, W.-S. Yang, C.-H. Wu, C.-F. Lai, V.-C. Wu, T.-S. Chu, K.-D. Wu, and National Taiwan University Hospital Study Group on Acute Renal Failure (NSARF) and the Taiwan Consortium for Acute Kidney Injury and Renal Diseases (CAKs), "Long-term remote organ consequences following acute kidney injury." *Critical care (London, England)*, vol. 19, p. 438, Dec 2015.

[6] M. Kerr, M. Bedford, B. Matthews, and D. O'Donoghue, "The economic impact of acute kidney injury in england," *Nephrology Dialysis Transplantation*, vol. 29, no. 7, pp. 1362–1368, 2014.

[7] D. McPherson, C. Griffiths, M. Williams, A. Baker, E. Klodawski, B. Jacobson, and L. Donaldson, "Sepsis-associated mortality in england: an analysis of multiple cause of death data from 2001 to 2010." *BMJ open*, vol. 3, no. 8, Aug 2013.

[8] H. E. Wang, P. Muntner, G. M. Chertow, and D. G. Warnock, "Acute kidney injury and mortality in hospitalized patients." *American journal of nephrology*, vol. 35, no. 4, pp. 349–355, Apr 2012.

[9] NHS-England, "Patient safety alert on standardising the early identification of acute kidney injury, (03 august 2016, date last accessed)," 2014. [Online]. Available: http://www.england.nhs.uk/2014/06/09/psa-aki/

[10] P. Subbegowda, R. Royston, D. Waring, R. Challiner, S. Heatley, R. Henney, M. Vincent, C. Butterworth, A. Hutchison, and L. Ebah, "Concurrent use of a local and the national algorithms to detect acute kidney injury," *British Renal Society, Conference on, 2015*, 2015.

[11] N. M. Selby, N. V. Kolhe, C. W. McIntyre, J. Monaghan, N. Lawson, D. Elliott, R. Packington, and R. J. Fluck, "Defining the cause of death in hospitalised patients with acute kidney injury." *PloS one*, vol. 7, no. 11, p. e48580, Nov 2012.

[12] A. Wonnacott, S. Meran, B. Amphlett, B. Talabani, and A. Phillips, "Epidemiology and outcomes in community-acquired versus hospital-acquired aki." *Clinical journal of the American Society of Nephrology : CJASN*, vol. 9, no. 6, pp. 1007–1014, Jun 2014.

[13] S. Tirunagari, S. Bull, and N. Poh, "Automatic classification of irregularly sampled time series with unequal lengths: A case study on estimated glomerular filtration rate," *arXiv preprint arXiv:1605.05142, machine learning for signal processing, IEEE workshop on*, 2016.

[14] C. K. Williams and C. E. Rasmussen, "Gaussian processes for machine learning," *the MIT Press*, vol. 2, no. 3, p. 4, 2006.

[15] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (gpml) toolbox," *The Journal of Machine Learning Research*, vol. 11, pp. 3011–3015, 2010.

[16] J. Vanhatalo, J. Riihimäki, J. Hartikainen, P. Jylänki, V. Tolvanen, and A. Vehtari, "Gpstuff: Bayesian modeling with gaussian processes," *The Journal of Machine Learning Research*, vol. 14, no. 1, pp. 1175–1179, 2013.