Prediction of patients' admission to ICU for sepsis cases

1. Use case description

The underlying use case is from a Dutch hospital with about 50,000 patients per year and 700 beds. The process focuses on sepsis cases. Sepsis can be life-threatening and occurs when the body reacts to an infection and injures its organs or tissue.

To track all performed events, the hospital uses an Enterprise Resource Planning system. The process contains logistical activities, that is, the patient's pathway through the hospital, and medical, that is, which blood values were measured and which treatments were performed.

First, patients are registered at the Emergency Room (ER). Next, they go through ER triage and ER sepsis triage. Their blood values in regards to Lactic Acid, Leucocytes, and CRP are measured recurringly throughout the process. Further, patients can receive IV Liquid, IV Antibiotics, or both during the treatment process. Based on their blood values, patients are admitted to the Normal Care Unit (NCU) or Intensive Care Unit (ICU). Patients are released from the hospital with one of five release types. Finally, it is tracked if patients return to the ER after the release. See Figure 1 for an exemplary scenario.

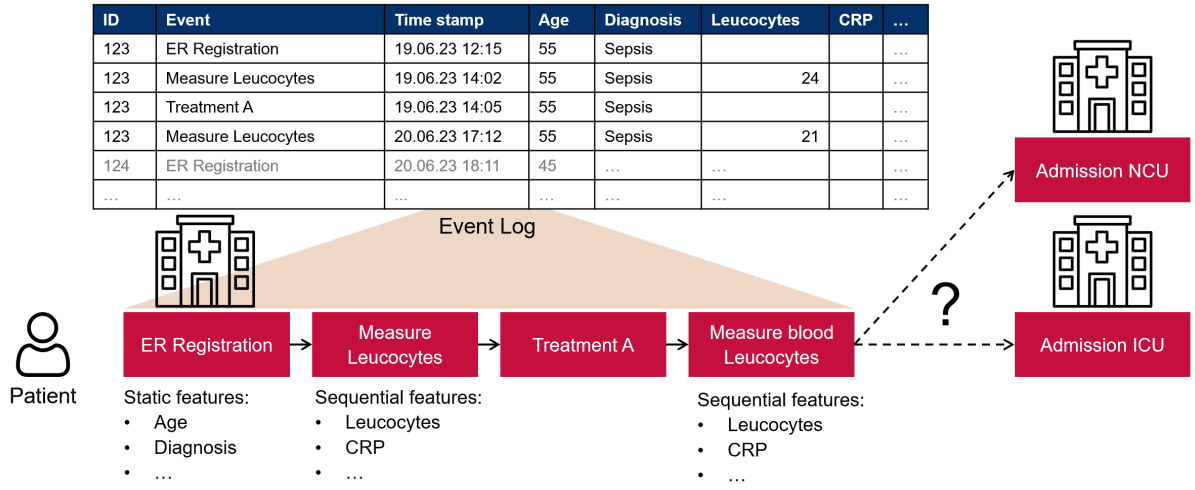

Figure 1: Exemplary scenario.

The patient events of the sepsis event log can be differentiated into 16 activities with different purposes, e.g., release type, measurement of *CRP*, *Leucocytes*, and *Lactic Acid*, or stating whether the patient was admitted to normal care, which represent sequential features:

- 5 activities regarding the release type
- 3 activities regarding the measurement of CRP, Leucocytes, and Lactic Acid
- 3 activities regarding the registration and triage
- 2 activities stating if the patient was admitted to ICU or NCU
- 2 activities regarding the treatment with IV Liquid or Antibitotics
- One activity stating whether the patient had to return to ER

In addition to the control-flow information, the patient event log contains 27 features. These can be **static** or **sequential**. Static features do not change over time (i.e., the patient pathway), e.g. a patient's age, whereas sequential ones can change over time, e.g. measurements of blood values. In the underlying case, three features are sequential and numerical and represent the measured values of *CRP*, *Leucocytes*, and *Lactic Acid*.

The remaining 24 features are static and categorical and only the patient's *age* is a numerical feature. The characteristics of the numerical features (three sequential features and one static) are presented in Table 1.

Table 1: Summary statistics of the numerical features.

Feature	Number of Observation	Mean	Standard Dev.	Percentile				
				5%	25%	50%	75%	95%
Age	724	72.12	15.48	40.0	65.0	75.0	85.0	90.0
CRP	2,388	111.66	83.53	12.0	44.0	94.0	156.0	276
LacticAcid	992	1.98	1.49	0.7	1.1	1.6	2.3	4.7
Leucocytes	2,525	13.24	16.87	2.8	7.6	11.0	15.1	24.9

Note: Some data was anonymized due to data security issues. For instance, Release Type A means a patient was released from the hospital as healthy. Further release types can be death, transferral to other hospitals, etc.

2. Process model

The process model looks as follows (with rectangles representing activities in the process); one walk through the process model refers to one patient pathway.

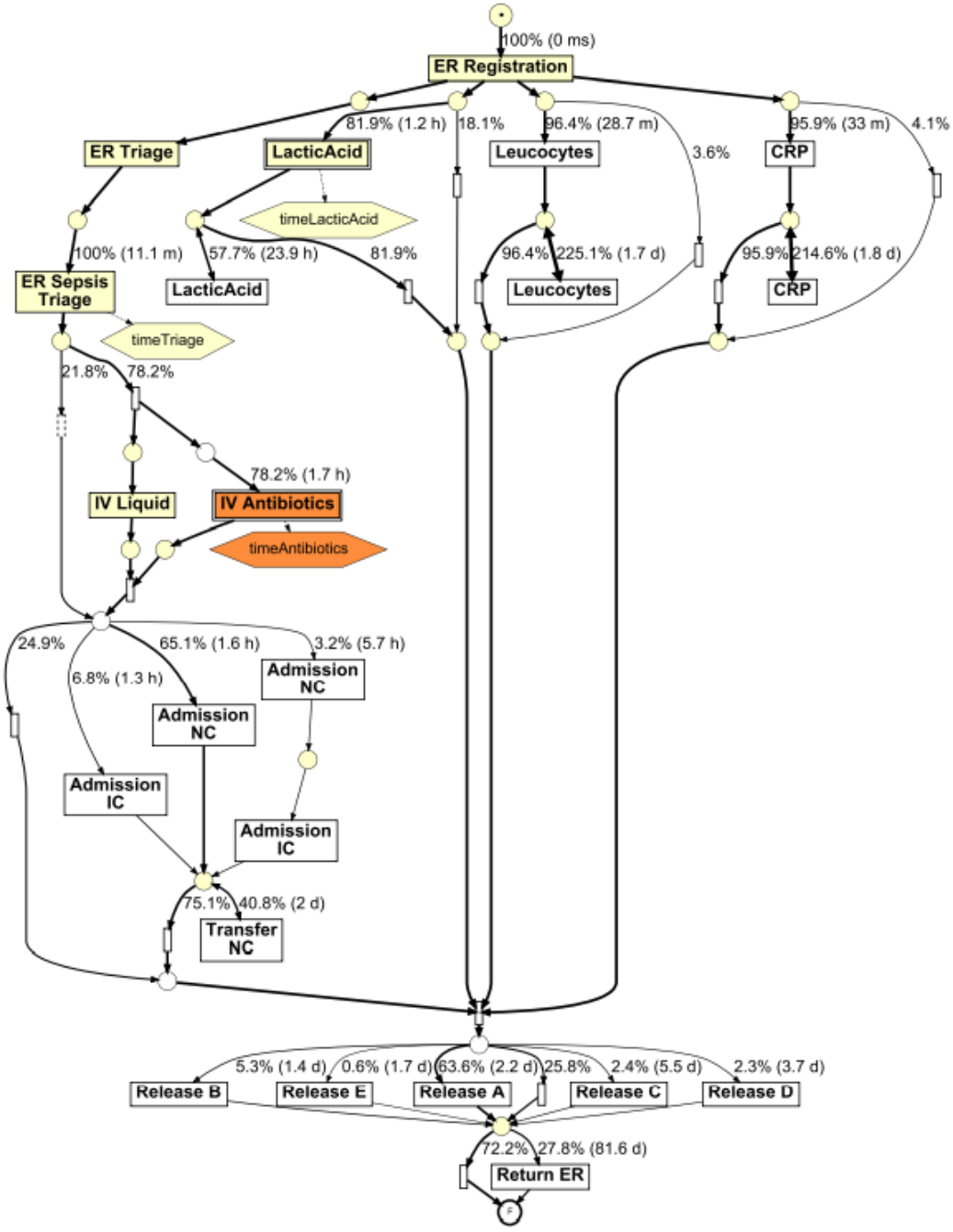

Figure 2: Process model.

Source: Mannhardt, Felix & Blinde, Daan. (2017). Analyzing the Trajectories of Patients with Sepsis using Process Mining.

3. Prediction

In our paper, we aim to predict the patient pathway. For this purpose, we have created a method based on Artificial Intelligence. This method uses the data gathered in the process as input and creates a prediction. These data include not only the conducted activities but also further information like age, blood values, etc. In the underlying use case, we aim to predict if the patients will be admitted to the ICU during their stay at the hospital.

> ***Q1:*** *Is the prediction of admission to the intensive care unit relevant in general? If so, in what way would it support medical staff in their decisions/ tasks?*
> ***Q2:*** *Would you trust the prediction without further explanation on how it was created?*
> ***Q3:*** *Can you think of other prediction tasks which may be of use to practitioners?*

4. Prediction with explanation

Research has shown that practitioners do not trust such predictions because the reasons remain unclear to the user. Therefore, we have created a method that not only predicts the patients' pathways but also provides an explanation of the prediction to the user.

The explanation is provided to the user in the form of a dashboard as depicted in Figure 3.

Figure 3: Developed dashboard.

The dashboard consists of four main parts.

Part a) provides general (static) information on a patient, such as their age, height or weight as well as their history during their current stay, for example when they have been admitted or when certain measurements have been taken.

Part b) shows the prediction of whether the patient will be admitted to ICU shortly and the probability.

Parts c) and d) provide the explanation plots. Specifically, c) shows the overall importance of certain medical indicators on the prediction and d) demonstrates further explanation details on specific indicators, such as a patient`s age or blood measurements. Both will be described in more detail in what follows.

4.1 Importance of medical indicators (Part c) in the dashboard)

Figure 4 presents the importance for medical indicators on the prediction. This plot shows the importance on the x-axis and the indicator name on the y-axis for static and sequential features. The indicators on the y-axis are sorted by their importance with the color indicating the importance as well. The features *Oligurie*, *Hypotensie*, and *Leucocytes* emerge as the three top contributors with the most substantial impact on the prediction (admission to ICU).

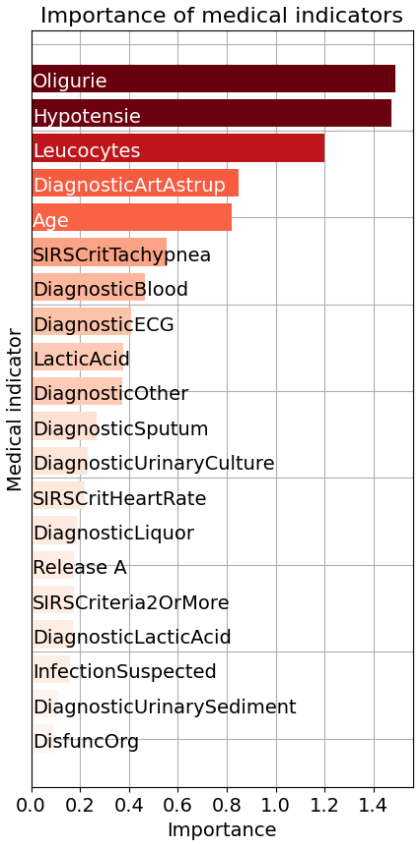

Figure 4: Importance for static and sequential medical indicators.

> ***Q4:*** *In general, are explanations in the form provided above helpful to you and provide additional insights?*
> ***Q5:*** *Would explanations in the form provided above influence your trust/ acceptance of predictions?*

4.2 Explanation per medical indicator (Part d) in the dashboard)

We distinguish between sequential and static medical indicators in the type of explanation plots.

4.2.1 Sequential indicators

The three explanation plots for sequential indicators will be described for the medical indicator *Leucocytes* exemplarily (see Figure 5). These plots are specifically created for a patient, which means that they are based on a patient's history.

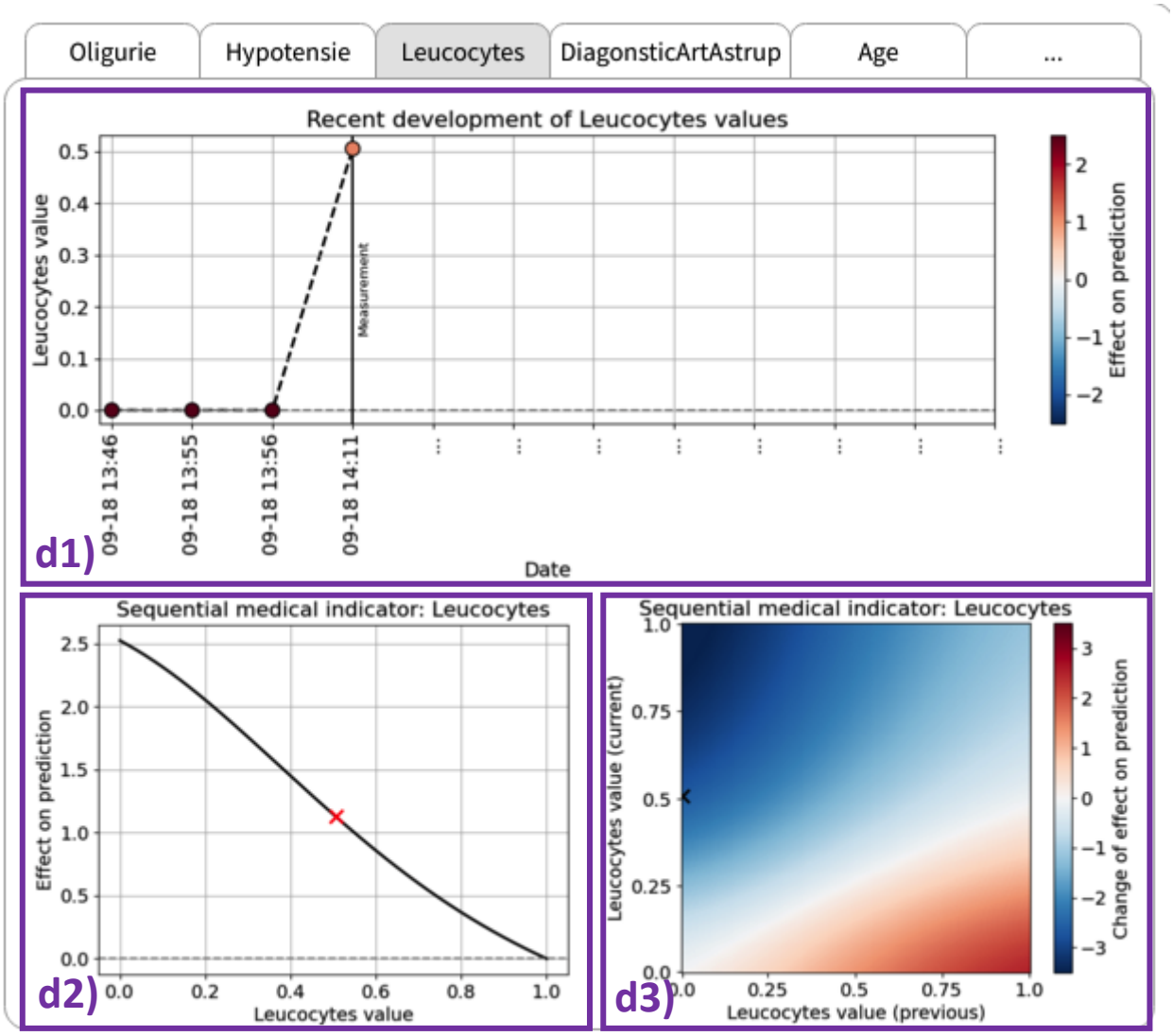

Figure 5: Explanation plots for sequential indicator Leucocytes in the dashboard.

In Figure 5, the explanation plot d1) presents the development of the values for the sequential medical indicator *Leucocytes* over time. This plot shows on the date of measurement on the x-axis and on the y-axis the Leucocytes value. The color thereby indicates a single measurement's effect on the prediction. For example, the Leucocytes value measured on 09-18 14:11 has a lower effect on the prediction than the three previous values, indicating initial values before the first Leucocytes value could be measured by medical staff.

> ***Q6:*** *In general, are explanations in the form provided above helpful to you and provide additional insights?*
> ***Q7:*** *Would explanations in the form provided above influence your trust/ acceptance of predictions?*

In Figure 5, the explanation plot d2) shows the shape function for the sequential medical indicator *Leucocytes*. This plot shows the Leucocytes value on the x-axis and the effect on the prediction on the y-axis. The red cross indicates a specific patient's Leucocytes value and the respective effect on the prediction.

> ***Q8:*** *In general, are explanations in the form provided above helpful to you and provide additional insights?*
> ***Q9:*** *Would explanations in the form provided above influence your trust/ acceptance of predictions?*

In Figure 5, the explanation plot d3) presents the change of effect on the prediction for the medical indicator *Leucocytes*, from the previous measurement value to the current measurement value. This plot shows the previous Leucocytes value, the current Leucocytes values, and the change of effect on the prediction on the x-axis, y-axis, and z-axis, respectively. The plot illustrates that an elevation in the Leucocytes value (from a previous measurement value of 0.0 to a current measurement value of 1.0) corresponds to a decreased probability of the patient requiring ICU admission. The black cross indicates a specific patient's previous and current Leucocytes value.

> ***Q10:*** *In general, are explanations in the form provided above helpful to you and provide additional insights?*
> ***Q11:*** *Would explanations in the form provided above influence your trust/ acceptance of predictions?*

4.2.2 Static indicators

The explanation plot for static indicators will be described for the medical indicator *Age* exemplarily.

In Figure 6, the explanation plot shows the shape of the static indicator *Age*. This plot shows the value on the x-axis and the effect on the prediction on the y-axis. The red cross indicates a specific patient's age and the respective effect on the prediction.

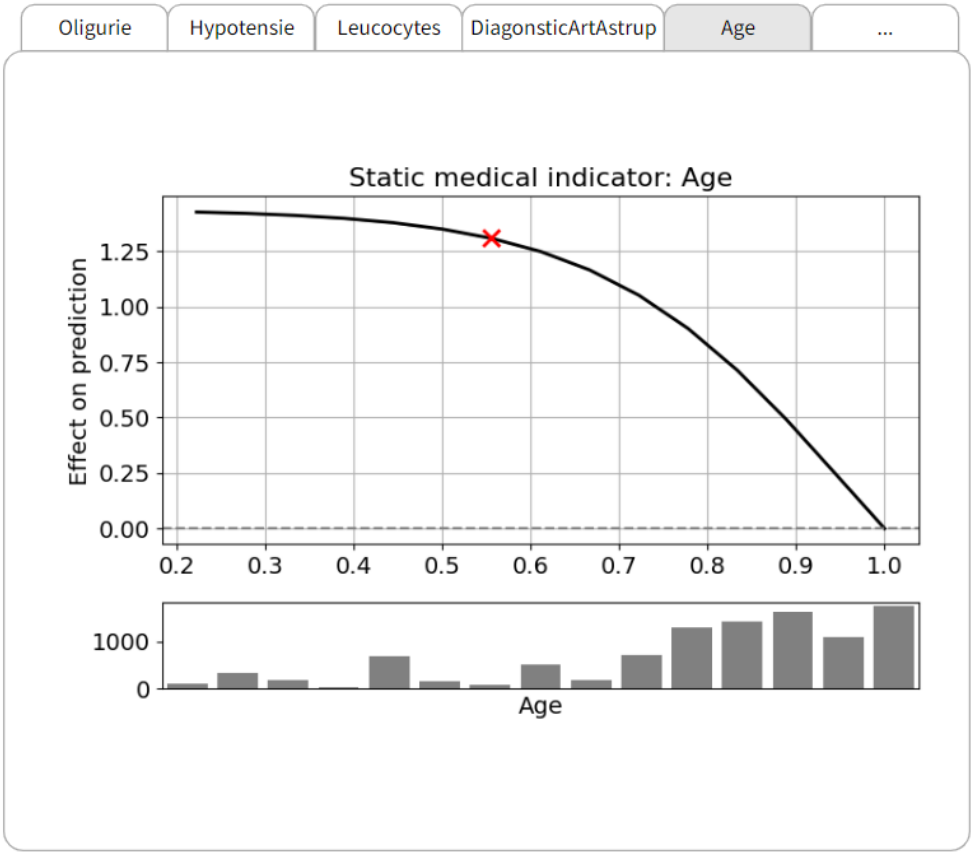

Figure 6: Explanation plot for static indicator Age in the dashboard.

> ***Q12:*** *In general, are explanations in the form provided above helpful to you and provide additional insights?*
> ***Q13:*** *Would explanations in the form provided above influence your trust/ acceptance of predictions?*

Further, our explanation plot for the medical indicator *Age* is compared with an explanation plot, created with one of the most common methods in the field of Explainable Artificial Intelligence. In Figure 7, the explanation plot created with our method is presented on the left side and the explanation plot created with the other method is presented on right side.

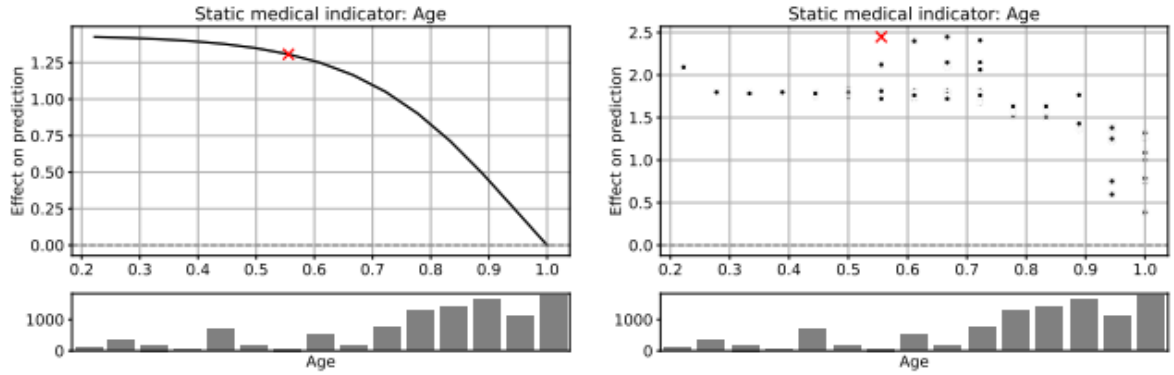

Figure 7: Two different explanation plots for the medial indicator Age (left our plot, right other plot).

***Q14:** Which plot (left or right) would you prefer and why?*

4.3 General questions

To conclude, we are interested in some general questions regarding the dashboard.

Q15: *How useful would you say is the dashboard in general?*
Q16: *Do you think that the dashboard is easy to read and comprehend?*
Q17: *Would you need more visual guidelines (e.g. more icons or different colors)?If so, please specify.*
Q18: *Is there anything missing in the dashboard?*
Q19: *Is there anything else that you would like to add regarding this interview?*