

Uncovering Complex Relations in Patient Pathways based on Statistics: the Impact of Clinical Actions

Jelmer J. Koorn
Utrecht University
j.j.koorn@uu.nl

Xixi Lu
Utrecht University
x.lu@uu.nl

Felix Mannhardt
Technical University
Eindhoven
f.mannhardt@tue.nl

Henrik Leopold
Kühne Logistics
University
henrik.Leopold@the-klu.org

Hajo A. Reijers
Utrecht University
h.a.reijers@uu.nl

Abstract

Process mining is a family of techniques that can aid healthcare organizations in improving their processes. Most existing process mining techniques do not provide insights into the impact that activities can have on the process. Some novel techniques try to address this issue, but these techniques are either not generic in their approach or cannot provide insights into complex relations in organizational processes. We propose a novel and generic approach with the goal of producing insights into statistical relations within healthcare processes. We apply the approach on a public data set on sepsis in an emergency room. We find that the hospital might optimize its process in two respects: (1) their cost-benefit balance for patient care by considering their activities in terms of continuous monitoring and substance administration, and (2) their policies on discharging patients as to ensure patients are not discharged too early and return to the emergency room.

1. Introduction

Process mining is a popular family of techniques that uses event log data to visualize and analyze organizational processes [1]. Also, in the field of healthcare, process mining techniques have become increasingly popular in recent years [2, 3]. The healthcare field offers many opportunities for process mining to support data-driven analyses to optimize complex healthcare processes [4]. Many healthcare case studies (e.g. [5]) have been performed and even new techniques and methodologies have been proposed (e.g. [6]). As such, the healthcare domain and process mining techniques have proven to be a fruitful combination.

However, most process mining techniques focus on the control-flow perspective, i.e. they focus on the order in which activities are performed. To investigate processes from this perspective, a number of discovery techniques have been proposed over the

years. Prominent examples include the heuristic miner [7], fuzzy miner [8], and inductive miner [9]. In a healthcare setting, these techniques can help healthcare professionals, among others, to understand possible patient pathways [10]. So-called conformance checking techniques can also show whether medical guidelines and procedures are appropriately followed [11].

Once such insights are obtained, natural follow-up insights are often required to deepen the understanding of a process. In that light, practitioners are often interested in finding out how particular actions impact the patient pathways and their well-being. The control-flow perspective does not reveal the impact that actions may have on the process. Recognizing this, existing process mining techniques have proposed several alternative ways to uncover these insights. From a causal mechanisms perspective, a number of techniques have been proposed that try to find statistically significant patterns that show the impact of treatments in processes [12, 13]. Composite state machine techniques also aim to tackle the problem by identifying state transitions [14].

Nonetheless, both strands of techniques face some fundamental limitations that prohibit their applicability in uncovering the impact of actions on the process. Firstly, the techniques related to causal mechanisms are often not generic in their applicability. To exemplify, the authors of [12] propose a technique that can only analyze repetitive processes and single actions (rather than multiple actions). Secondly, the composite state machine techniques can capture the status of a patient state and the transition between patient states, but they cannot show how process activities influence state transitions. The limitations of the existing process mining techniques motivate the proposal of a novel approach to gain an understanding of the complex healthcare processes. Therefore, we propose a novel and generic state abstraction approach to generate new insights in terms of the impact of actions on the transition between states in organizational (healthcare) processes. In particular, we propose a statistical

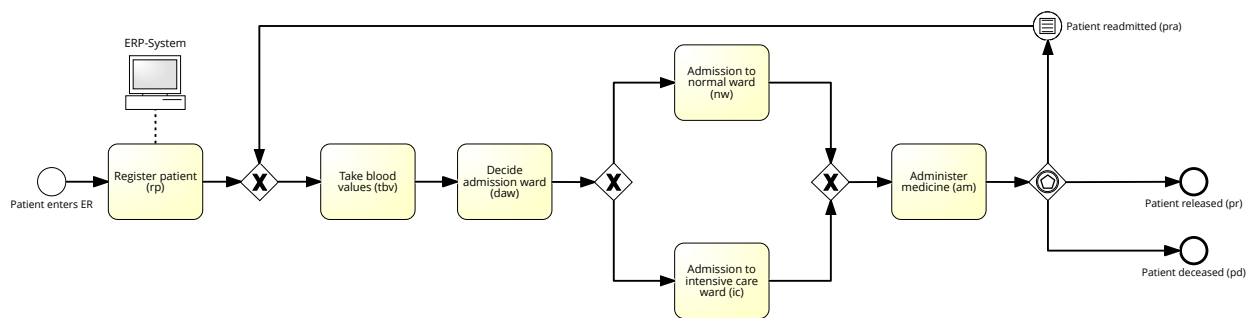


Figure 1. Simplified sepsis process as a BPMN model

approach to provide insights into the complex relations that influence the state transitions. We apply this technique on a public data set on sepsis and uncover previously hidden relations in the healthcare process. These new insights can form the basis for the hospital to improve their sepsis process two ways: (1) to optimize their cost-benefit balance in patient care, and (2) to review the way they discharge patients to avoid early returns to the emergency room.

The rest of the paper is organized as follows. In Section 2, we illustrate the problem we address in more detail. In Section 3, we introduce our approach from a conceptual point of view. In Section 4, we report on the results of our case study on data from an emergency room where we show that the proposed approach can be applied to real-world data and provides us with new insights on the impact of actions on state transitions. In Section 5 we discuss related work before in Section 6 we discuss the limitations and conclude the paper.

2. Problem Statement

To illustrate the problem we address in this paper, consider the scenario of a patient being admitted to the emergency room (ER) of a hospital. In Figure 1 we visualize such a healthcare process. What we see from the process model in Figure 1 is that the process starts when a patient is registered (*rp*). After registration, the medical team is required to perform initial blood tests to gain an understanding of the situation of the patient (*tbv*). Then, the medical team faces a decision to admit the patient to a normal care ward or an intensive care ward (*nw* and *ic*). Next, the medical team performs a number of activities that have to do with providing a treatment to a patient in the form of administering medicine (*am*). For example, they can administer antibiotics through an intravenous line (IV). The end of the process, for most patients, is the discharge from the ER (*pr*). However, some patients return to the ER at a later stage (*pra*) or pass away (*pd*).

Process mining techniques can help healthcare practitioners and managers to obtain better insights into this process. Based on data logged by IT systems so-called process discovery techniques can detect and visualize which specific activities were performed and in which order these activities were executed. These activities are often a mix of medical activities (e.g. treatments) and logistic activities (e.g. transferring patients). Figure 1 represents the outcome of such a process discovery technique. We can see that the resulting process model gives an overview of the main process flow. This focus on activities and their order is commonly referred to as the *control-flow perspective*.

Visualizing the order in which activities are performed already aids in gaining an overall understanding of the process. However, it does not reveal the impact of certain decisions on the process flow. Given the potential severity of decisions in an ER, it is very important to understand which decisions lead to desirable and undesirable outcomes. For example, doctors may want to understand if admitting a patient to an intensive care ward (*ic*) increases or decreases the chances of that patient needs to return to the ER at a later stage (*pra*). To understand such aspects, several process mining techniques complementing the control-flow perspective have been developed. We can identify two relevant research directions in this regard: (1) causal mechanism, and (2) artifact-centric process discovery.

Previous research has looked into this phenomenon from a *causal mechanism* perspective [12, 13]. One technique is proposed by the authors of [12]. In their approach, the authors study repetitive aggressive behavior of clients by trying to identify statistically significant patterns in actions of clients, responses of caregivers, and future aggression of clients. The technique provides insights into the process of aggressive behavior, but is limited in its applicability to other scenarios for two reasons. First, the technique focuses on *repetitive* processes. As such, the outcome

of the process (the effect) must equal the input (action) of the next iteration of the same process. Second, the technique only considers the impact of singular actions, but ignores the possibility of sets of actions. Given these limitations, we cannot study our sketched problem setting. The sketched healthcare process is not a repetitive process as its input (the activities conducted in the context of the ER) and outcome (the discharge type) are not equal. Secondly, the activities performed in the process can also be sets of activities rather than a single activity, e.g. taking a number of blood tests rather than a single blood test.

Another research direction that relates to problems similar to ours are *artifact-centric* discovery techniques [14]. These techniques use composite state machines to capture and visualize different perspectives on a process and how the relations in a process change. The advantage of such techniques is that, by using states, a generic approach is provided. The notion of states is interesting in this setting as we cannot capture the status of a patient in terms of activities, only in terms of states (e.g. *pr* or *pd*). In the sketched problem we are interested in how the process activities (e.g. *ic* or *mw*) influence the state of a patient. However, the composite state machine techniques do not provide insights into the activities that impact the state transitions. In the above problem scenario, we would be able to understand state transitions, but not the underlying causes for those transitions.

In sum, to obtain the insights required in our scenario, existing process mining techniques are either not sufficiently generic or do not provide insights into the complex relations. The combination of these two factors is crucial to aid practitioners and managers in gaining an in-depth understanding of the process and to generate actionable insights. In this research, we introduce a novel approach that explicitly differentiates between *states* and *actions* and automatically discovers whether (sets of) actions lead to different states.

3. Approach

In Figure 2 we show which steps we take in the approach. Below, we go into detail on every step.

3.1. Step 1: State-action log

The input for our approach is an event log L that captures how the considered process was executed. Formally, we can define L based on the universe of all events \mathcal{E} . The events recorded for a single execution (i.e., an instance) of the process is called a *trace*, which is modeled as a sequence of events. Therefore, we denote a trace with n events as $\sigma = \langle e_1, \dots, e_n \rangle$,

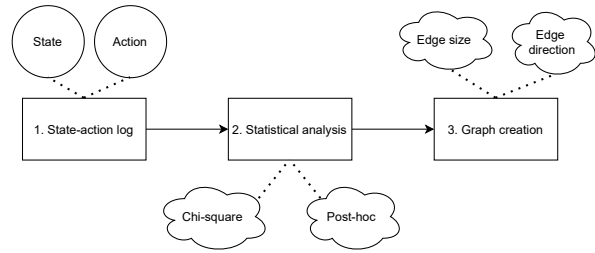


Figure 2. Proposed approach

Table 1. Running example sepsis chi-square table.

Observed	Normal discharge	Deceased	Readmission	Total
Normal ward	350	50	90	490
Intensive care	350	150	60	560
Total	700	200	150	1,050

Expected	Normal discharge	Deceased	Readmission	Total
Normal ward	326.7	93.3	70.0	490
Intensive care	373.3	106.7	80.0	560
Total	700	200	150	1,050

where each event $e \in \sigma$ is part of \mathcal{E} . However, for the purpose of our analysis, we require an adapted notion of such a traditional event log. Therefore, we build the state-action log. Specifically, we need to distinguish which events from a trace σ represent *states* and which events from σ represent *actions* that lead to state changes. To this end, we introduce a function α , which returns for each $e \in \mathcal{E}$ whether e represents a *state* or an *action*. To illustrate this, consider the trace $\sigma_1 = \langle rp, tbv, daw, nw, am, pr \rangle$, which represents a possible trace according to the process shown in Figure 1. As pointed out above, we are particularly interested in the impact of the activities conducted in the ER on the discharge type (*pre*, *pra*, or *pd*). Therefore, we define α in such a way that it maps the three discharge types to *states* and all other events to *actions*. As a result, we can analyze whether and which activities lead to one or the other discharge type.

Due to the complexity of real-world processes, the definition of α requires input from domain experts. This means that we ask the user to define which events represent *actions* according to our definition. While this is certainly associated with some effort, it makes sure that our technique can detect practically relevant relationships.

3.2. Step 2: Statistical analysis

Once the state-action log is created, the approach moves to the analysis step in which we perform statistical tests to discover significant relations. Here, we perform chi-square tests, which are subject to a number of assumptions. This step is followed by

post-hoc tests. The *Chi-square* test is a well-established statistical test proposed halfway in the last century [15]. The Chi-square test has many advantages [16], the main reasons for its use in this approach are: (1) it is very robust due to its non-parametric nature (i.e. it does not make assumptions about data distribution), (2) the computation of the statistic is relatively easy and fully transparent, (3) the test provides rich information, and (4) the test is very suitable for in-depth post-hoc tests.

The test takes two inputs: (1) a set of observed behavior for two categorical variables, and (2) the chi(χ) distribution. The categorical variables in the context of our approach are the state and the action. The data that is used in the chi-square test represents how often each combination of state and event are observed in reality. This is referred to as the *observed frequency*. The observed frequencies are converted into an $n \times m$ table where n represents the number of action sets and m represents the number of states. This is called the observed frequency table, see for example the top of Table 1. Based on this table the chi-square distribution is introduced. In our example scenario we have two actions; *admission to normal ward* and *admission to intensive care unit*, and we have three states; *normal discharge*, *deceased*, and *readmission*. Table 1 shows artificially generated data for illustrative purposes.

The chi-square distribution is used to calculate the *expected frequency* for each cell (i.e. combination of action and response). The expected frequency calculator takes the distribution and the observed frequency grand total and total of each column and row. Using these number the following formula is applied: $\frac{N_r \times N_c}{E[s][a]}$ where

N_r equals the expected row total, N_c equals the column total and $E[s][a]$ equals the grand total.

Take for example the action *normal ward* and state *normal discharge*. The observed frequency is 350. The expected frequency is calculated based on the column total (700), row total (490) and grand total (1,050). Applying the formula on the example gives: $\frac{490 \times 700}{1050} = 326.7$ The result of this exercise is represented in a separate table called the expected frequency table, see the bottom of Table 1.

Now, the chi-square test compares the difference between the observed and expected frequency for each cell. These differences are all summed and result in the chi-square score. The formula used for this is:

$$\chi_c^2 = \frac{(O_{s1,a1} - E_{s1,a1})^2}{E_{s1,a1}} + \dots + \frac{(O_{sn,an} - E_{sn,an})^2}{E_{sn,an}} \quad (1)$$

In the formula, $O_{sn,an}$ describes the observed value and $E_{sn,an}$ describes the expected value for a combination of state and action where 1 to n denote the individual

action or state. If we apply this formula on our example case for all the combination of states and actions, i.e. from state *normal discharge (ND)* to state *readmission (R)* and for action *normal ward (NW)* and action *intensive care (IC)*. This results in the following equation:

$$\chi_3^2 = \frac{(O_{ND,NW} - E_{ND,NW})^2}{E_{ND,NW}} + \dots + \frac{(O_{R,IC} - E_{R,IC})^2}{E_{R,IC}} \quad (2)$$

We know from Table 1 that there are three states and two actions, so the degrees of freedom: $c = (3 - 1) \times (2 - 1) = 3$.

$$\chi_3^2 = \frac{(350 - 326.7)^2}{326.7} + \dots + \frac{(60 - 80.0)^2}{80.0} = 51.56 \quad (3)$$

If the observed and expected frequency values are sufficiently different, a large test score is returned. The next step is to check which pairs of observed and expected frequency are sufficiently different. This step is quite complex and describes how the chi-test score is compared to the chi-distribution, see [17] for more details. The larger the chi-square test score, the more likely it is a significant score. In this study we test on the alpha = 0.05 level. If the chi-square test score is significant, this indicates that there is at least one set of actions for which we observe a significantly different frequency of states than expected. In the example above, the p-value is $6.4 * 10^{12}$ which is well below the alpha level of 0.05. Thus, we can conclude that there is a significant relation between at least one set of actions and states.

The chi-square test is subject to a number of data assumptions, which need to be checked [16]. Most assumptions have to do with the way the data is collected and stored. They can only be manually checked by the analyst. However, one assumption relates to the sample size and can be automatically checked. Specifically, the assumption states that in the expected frequency table in 80% of the cells a minimum value of 5 must be present. In this work, we apply a heuristic selection criterion to test this. If the assumption is not met, an NA value is returned and no further test results are returned.

The next step in the analysis is to perform *post-hoc* tests. The goal is to determine which specific combinations of sets of actions and states are significantly related. To that end, a detailed statistic called the adjusted standardized residual (ASR) is calculated [18]. The ASR standardizes the difference between the observed and expected frequency using the

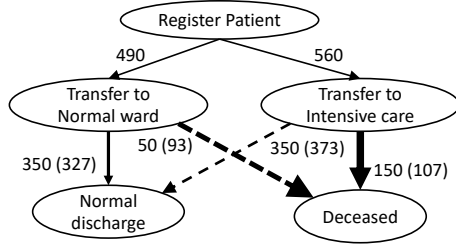


Figure 3. The resulting graph of the running example.

following formula:

$$ASR = \frac{O_{s1,a1} - E_{s1,a1}}{\sqrt{E_{s1,a1} * (1 - \frac{E_{s1,a1}}{N_c}) * (1 - \frac{E_{s1,a1}}{N_c})}} \quad (4)$$

To apply it to our example case scenario, let us consider the state action combination of *normal discharge* and *normal ward*:

$$ASR = \frac{350 - 326.7}{\sqrt{326.7 * (1 - \frac{326.7}{700}) * (1 - \frac{326.7}{700})}} = 5.70 \quad (5)$$

With the ASR score, three metrics are produced: (1) its significance, (2) its value, and (3) its size. In the analysis part of the algorithm, only the significance of the ASR is used. The ASR score is then normalized and compared to a standardized p-value score (critical score) [18]. The significance of the ASR score determines if it has a significant impact on the overall chi-square test score. In the example scenario above, the 5.70 score is significant. If the test result is significant, this means that the tested combination of state and set of actions is significantly related. Below, we will go into detail about how the graphical representations are created for all logs in which at least one ASR score is significant.

As we perform multiple statistical tests on the same data set we have to correct for the multiple comparisons problem [19]. The multiple comparisons problem states that with each test we increase the chance of finding something by chance. To counteract this, we adjust the p-value we set as significance barrier with the Bonferroni correction [20]. The test corrects the p-value according to the following formula: $p = \frac{\alpha}{n}$ where α is the significance level, n equals the number of tests, and p equals the new p-value used for testing. Applying this on the running example, we perform tests on each cell ($n=6$) and test on a 0.05 alpha level. Thus: $p\text{-value} = \frac{0.05}{6} = 0.008$.

3.3. Step 3: Graph creation

The goal of the *graph creation* is to visualize the uncovered statistical relations from the analysis step. Existing work of [12] proposes a technique to visualize statistical relations in process mining. In this work, we build on that particular visualization technique. The aim of graph creation is to increase the understandability of the graph to the end-user. To that end, the significance, value, and size of the ASR and the frequencies of the arcs are used to create graphical representations.

The *significance* determines which arcs need to be drawn for each node. If the ASR value is significant, an arc is drawn. If the ASR value is insignificant, no arc is drawn. The *value* of the ASR can be either positive or negative and determines its *direction*. The direction of the ASR cannot be negative as that would indicate a non-significant value and thus no arcs would be drawn. If the ASR value is positive, a solid arc is drawn to indicate a positive relational direction. If the ASR value is negative, a dotted arc is displayed to indicate a negative relational direction. Returning to the example scenario described in the previous section, a 5.70 significant score would result in a solid arc to indicate a positive relational direction between normal ward and normal discharge. In practical terms, after a normal ward, we see an increased likelihood that a normal discharge follows.

The *size* of the ASR value corresponds with the thickness of the arcs. In total there are six thickness classes, three for positive and three for negative values. These classes are based on the maximum and minimum ASR values and the critical value. A simple algorithm calculates the difference between the maximum and the critical score (usually $|2.57|$) and determines three equally large ranges for three classes. The classes reflect the effect size of the relation. A thicker arc means a stronger relation. In our case, 5.70 is the largest ASR value. Thus, it would be represented as a thick solid arc in the graphical representation.

Note that a thicker dotted arc means a stronger negative relation. To exemplify, a thick dotted arc between a set of actions and a state reflects that there is (highly) decreased likelihood that this state occurs after that set of actions. Finally, the observed and expected *frequencies* that are calculated in the second step are displayed on each drawn arc. This helps to gain an understanding as to the prevalence of each combination of state and action.

If we apply this approach on the running example, as depicted in Table 1, we get the representation that can be found in Figure 3. Here, we see for example that admission to a normal ward ($n = 490$) (slightly)

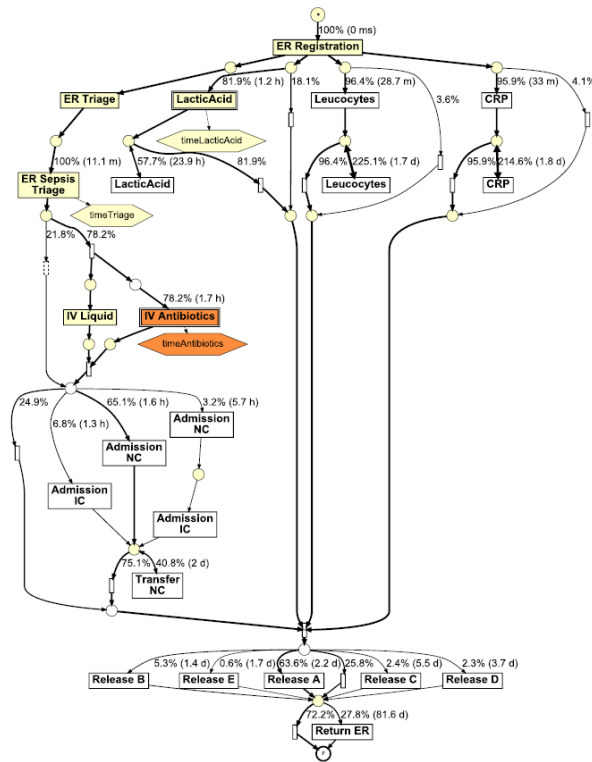


Figure 4. Multi-perspective process model used for decision mining on the sepsis process data in [21].

increases the chances of a normal discharge (observed = 350, expected = 350) and (largely) decreases the chances that the patient deceases (observed = 50, expected = 93). In addition, the same analysis can be applied for the cases that are admitted to the intensive care unit. Finally, we see that readmission is not included in the graph, which means that it has no effect on either normal discharge or death of a patient.

4. Case study

The goal of this section is to twofold: (1) to demonstrate the applicability of the proposed approach in a real-world scenario, and (2) to generate new insights into the complex relation between state and action. To this end, the sepsis data from [21] is used in this research. Below, we will first go into detail on the data set and its characteristics. As will be explained below, the data can be analyzed in multiple ways. Therefore, we present two different ways to create a state-action log based on the sepsis event log in the pre-processing section. First, we present the results where we look into the effect of *continuous monitoring and substance administration* of a patient on the way a patient is discharged. Then, we present the results where we

look into the impact of the *decisions of the medical team* on whether or not the patient returns to the ER room. Here, we study two main decision moments: (1) to which type of ward the patient is admitted, and (2) in what way the patient is discharged. The goal of looking into these challenges is to show that our approach can produce new work hypotheses that can be used to enhance the understanding of healthcare processes. Ideally, these insights form a starting point of a collaboration effort with healthcare practitioners and/or managers to uncover and understand the complex relations in their healthcare processes.

4.1. Case study data & context

The data for this project was collected in another study in which the researchers collaborated with a medium-sized regional hospital in the Netherlands [21]. The project was aimed at studying the trajectories of emergency room patients, more specifically, those patients with symptoms for sepsis. Sepsis is a life-threatening condition where the body reacts to an infection by damaging its own tissue. Sepsis requires continuous monitoring and a treatment with antibiotics. Data was taken from three systems: emergency room, laboratory, and other ward. This data was anonymized and combined into one data warehouse after which an event log in XES format was created. In total, the log contains traces for 1,050 cases and 15,214 events that are recorded in 1.5 years of patient records (from November 2013 to June 2015). The log contains 16 unique activities and 846 distinct variants. More details on the data and its collection can be found in [21].

The event log contains sixteen activities: three activities for the registration and triaging, three activities for taking certain blood measures for patient monitoring (leucocytes, CRP, and lactic acid), two activities for administration of substances to the patient (IV liquid and IV antibiotics), two activities for admission or transfer to normal ward or intensive care unit (admission NC or IC), five types of discharge (activities release A-E) from the hospital, and one activity capturing if patients returned to the ER at a later stage (within 28 days).

In their project, the authors of [21] mainly focus on three challenges: (1) conformance checking in terms of adherence to medical guidelines, (2) uncovering the various trajectories that patients can flow through, and (3) to discover decision rules to detect returning patients [21]. They create a process model depicted in Figure 4 to study these challenges. For our work, we aim to expand on this work by focusing on the latter two challenges. These two challenges are of interest as the impact of certain actions can influence both the

Table 2. The mapping from activities to states and actions of analyses

	Activities mapped to states	Activities mapped to actions
Analysis 1	'ER Registration', 'Admission NC', 'Admission IC', 'Release A', 'Release B', 'Release C', 'Release D', 'Release E'	'Lactic Acid', 'IV Liquid', 'IV Antibiotics', 'CRP', 'Leucocytes'
Analysis 2	'ER Registration', 'return ER'	'Admission NC', 'Admission IC', 'Release A', 'Release B', 'Release C', 'Release D', 'Release E'

trajectory of a patient as well as the patient return rate of patients to the ER. As for the *patient trajectories*, the authors *describe* the various types and frequencies of the paths that patients can take [21]. They also indicate that some paths are more desirable than others. In this light, it is interesting to determine if there are certain actions that lead to a more desirable patient path. However, from the work of [21] no insights can be obtained from their findings as to which specific actions impact the path of the patient.

The first challenge addresses the patient trajectories, turning to the second challenge where the authors focus on *patients returning to the ER*. The authors of the original study try to apply decision mining techniques [22, 23] to discover if there were certain rules that could give insights into whether or not a patient would return to the ER within a certain time frame. Ideally, rules would be discovered that helped doctors gain insights into which actions to perform or decisions to take in order to reduce the number of returning patients. However, the authors could *not* find any such rules based on, amongst others, triage documentation and values from metrics taken. In this work we will expand on the challenge of returning patients by studying if monitoring the patient continuously and/or administering substances through IVs is correlated with a lower likelihood of the patient returning to the ER.

4.2. Pre-processing

As explained in our approach, we try and tackle two challenges: (1) patient trajectories, and (2) patients returning to the ER. For each challenge we create a corresponding state-action log by mapping the activities to the set of states and the set of actions. For this cases study, we created two state-action logs, one for each analysis. For the *first analysis*, we mapped the ER registration, admission, and discharge activities as states and considered the tests and the medicines as the actions. This mapping allows us to find whether the lab tests and the medicines have any influence on the admissions (to IC or NC) and the type of discharge. For the *second analysis*, we mapped the admission and

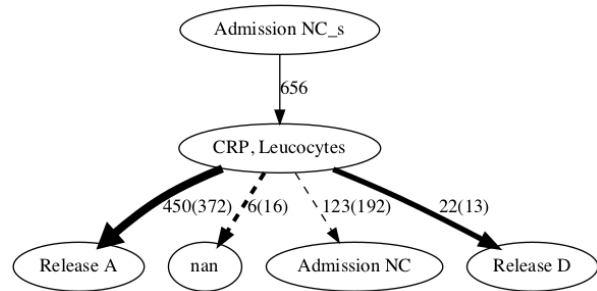


Figure 5. Sepsis analysis when we use metrics as actions and discharge type, admission type and registration as states

discharge activities to actions and the 'ER Registration' and 'return ER' activities as states. This mapping allows us to find whether the admission to IC or NC and the discharge types have any influence on the patient returning to the hospital or not. The concrete mapping is listed in Table 2. In the following sections, we discuss the results found.

4.3. Continuous monitoring & substance administration

As mentioned previously, one of the challenges in the ER room is to find out how doctors can influence the trajectories of patients. In that respect, we developed a log where we are interested in the states: ward type and discharge type. *Ward type* can be either normal ward (Admission NC) or intensive care (Admission IC). *Discharge type* can be one of five types categorized as release A - release E. Finally, there can be patients still in the ER at the time of data collection; this should result in a nan (No return to ER). Next to the states we are interested in the *actions* that can impact the state transitions: CRP, Leucocytes, Lactic acid, IV liquid, and IV antibiotics.

The first three actions are all decisions of the medical team to (*continuously*) *monitor* a patient's status through blood tests. The blood tests (CRP, leucocytes, and lactic acid) are performed almost by default when a patient is admitted to the ER and has a slight fever.

These tests are also performed regularly after a patient has been admitted to the hospital to check whether an applied treatment is working or not and may serve as a grounding to discharge patients. The last two actions regarding an IV refer to the decision of the medical team to *administer substances* to the patient. We can see in Figure 4 that these patient pathways exist, but we are interested in determining the *impact* of taking these blood tests and administering substances. The results from our approach can be split into two parts: (1) the relations visible in the graph, and (2) the lack of expected relations.

In Figure 5 we can see the statistically significant relations between the actions and the states. What we observe in the figure is that 656 cases are admitted to the normal care ward, roughly 60% of the population. In line with expectations, we can observe that two of the continuous monitoring actions are significantly related to the further trajectory of the patient: CRP and leucocytes. The CRP test refers to a c-reactive protein test; the leucocytes test is a white blood cell test. These are tests to check if the patient is reacting to an inflammation.

In Figure 5 we can observe that performing these tests are related to a higher probability of release type A (normal discharge) and release type D (transfer to other care facility). Especially release type A is common, we observe that this follows CRP and leucocytes in 459 cases where we would only expect 372 cases based on the statistical tests. In addition, performing these two tests is related to a lower probability of readmission to the normal care ward, this is observed 121 times where we would expect 192 cases. In short, the results show that monitoring a patient's blood values is related to a higher probability of a normal discharge of that patient, which is a desired outcome for all parties involved. This is not so surprising, as the results of these tests are used to determine whether a patient can be discharged or not.

What cannot be observed from the figure is that there are a number of relations that are expected, but were *not* found. First, we would expect to see a similar pattern for all blood tests and substance administration. Interestingly, the other blood test (lactic acid) and the administration of liquid and/or antibiotics through an IV do not have a significant impact on the patient pathway. These three actions are performed when the medical team estimates that the patient's condition is severe. Therefore, we would expect it to have a relation with the admission to the IC and/or to other types of discharges. However, no such relation is found in the data. It seems worthwhile to collaborate with domain experts to further investigate the effect of these actions on patient well-being. The goal of such a project could be to

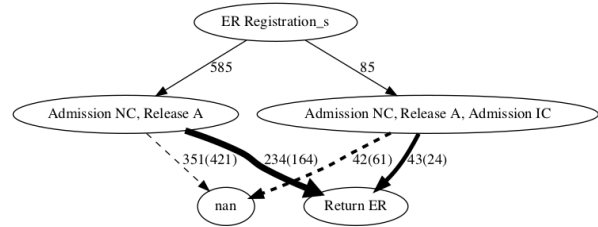


Figure 6. Sepsis analysis when we use registration and return as states, and admission type and discharge type as actions

optimize the cost-benefit balance for the care for these patients.

4.4. Medical team decisions

The second challenge relates to the return of patients after they leave the ER. Previous work could not identify rules to determine what factors play a role in the return of patients [21]. To follow up on this challenge, we investigate whether the decisions made by the medical team to admit a patient to a normal or intensive care unit and how the patient is discharged has an impact on the return of patients. As such, the *states* for this challenge are: registration and return to ER or nan (no return to ER). In turn, the *actions* are: admission type and discharge type. From the original process model in Figure 4 we can see that the return of patients happens in almost a third of the cases (27.8 %), but we cannot infer what might cause a patient to return.

In Figure 6 we can see the statistically significant relations between the states and actions. What we can observe is that admitting the patient to the normal care ward and having a normal discharge (Release A) is a relatively frequent pathway for patients: it is observed for roughly half the patients (585 out of 1,050 cases). If we look at the impact of these actions, we see that they are related to a higher probability of a patient returning to the ER. We observe that this happens 234 times whereas we expect it to happen 164 times. In addition, these actions are related to a lower likelihood of patients not returning to the ER. This is observed 351 times versus the expected 421 times.

A similar pattern shows if a patient is not only admitted to the normal ward (Admission NC), but also the intensive care (Admission IC) followed by a normal discharge (Release A). The patient is, again, more likely to return to the ER and less likely to not return to the ER. This patient pathway is considerably less common, but is still observed for roughly 10% of the patient population.

This is an interesting finding as it indicates that there

is a group of patients that seem to be doing relatively well. This is indicated by the fact that they are admitted to a normal ward and discharged normally. However, a subgroup of this patient group might not be doing so well in reality as they return to the ER. This points to the fact that symptoms might be overlooked during the first admission that come into play at a later stage. Further research should look into this patient population to identify which patient characteristics are associated with this to lower the return of patients. Gaining insights into this can result in better care for the patients as their pathology can be improved. In turn, this reduces both costs for the hospital and patients.

5. Related Work

As discussed in the problem statement, causal process mining techniques related closely to the problem we set out to tackle in this paper. To discover *causal mechanisms* in process models often decision rule or root cause approaches are proposed [13, 24, 25]. The authors of [13] propose an exemplary case of such a decision rule mining technique. They develop a model in a financial context in which they use uplift trees to determine for which subgroups of clients a specified treatment has an effect. Although this technique produces valuable input, the technique heavily relies on the definition of subgroups. This is not always a suitable approach as it can be that there is a specific process for a (largely) uniform subgroup of patients. No insights can be generated into the effectiveness of treatments.

Another limitation that this approach poses becomes clear when we consider the output these techniques produce. The output of the decision rule mining techniques are often *declarative* in the sense that they produce a set of rules described in text format which serves as guideline to optimize decision making in the process (e.g. [13]). In other work on causal mining, the output of causal mining techniques is more imperative in the sense that the authors propose a more graphical style output (e.g. [12]). Existing research has shown benefits for each type of output, but conclude that a hybrid form is optimal for maximizing the understandability of a process [18]. In turn, this should increase the effectiveness of turning process mining results into actionable insights [4].

In the present work we present a novel approach that addresses these limitations. Our approach is generic in the sense that it is flexible in its formalization of state and action. In addition, our approach produces graphical representations (i.e. process models) with declarative elements (i.e. statistical relations and annotations) as output. This ensures that the approach is: (1) applicable

to a wider spectrum of problem scenarios, and (2) that the results that are produced better can be used to support the interpretation and communication of the findings. To the best of our knowledge, no such approach or techniques have been previously proposed.

6. Conclusion

In this paper we proposed a novel approach to discover the statistical relations between states and actions. This work combined two main approaches to propose an approach that is both generic and identifies complex relations in organizational processes. The approach is generic in the sense that it can deal with a wide range of scenarios, as long as we can identify states and actions. The generic applicability stems from the flexibility in the concepts of states and actions. To exemplify, an action can consist of a single or multiple activities. In the latter case, the action can be defined in different ways, for example, as a set or sequence of activities. The limitation of such a generic approach is that it requires manual work to define such actions and states. Ideally, this is done through a collaboration between process analyst and domain expert.

We used a case study for two purposes: (1) to test the applicability of our approach, and (2) to generate new insights into complex relations. The case study concerned a public data set on 1,050 patient with sepsis that were admitted to the emergency room in a Dutch hospital. We applied the approach to tackle two remaining questions regarding patient pathways and patient return rates. First, the approach aimed to identify those actions that have an influence and are related to a desired patient pathway. Thereby optimizing the cost-benefit balance in the care of patients. Second, the approach was used to determine what actions in the processes are related to higher or lower probabilities of patients returning to the ER at a later stage. Ultimately providing new insights into which patients are susceptible to early return to the ER room with sepsis symptoms.

There are two main limitations to the proposed approach: (1) domain knowledge is required to define actions and states, (2) the approach cannot confirm that the relations are causal relations. As for the first limitation, the technique is not automated in the sense that it can detect states and actions. Ideally, this would be the case as it would decrease the burden on the involvement of domain experts. A brute force approach is an alternative, here one would try and map the activities in the process in all possible combinations of states and actions. However, this would severely escalate the multiple comparisons problem.

The second limitation concerns the fact that the approach discovers potential causal relations. It shows that there is a relation between a state and an action, but cannot guarantee that this relation is causal. In order to do that, a number of others steps need to be taken. For example, possible confounding variables need to be checked and a number of other independence criteria need to be met (i.e. exchangeability, positivity, and consistency) [13].

In future work we will focus on the notion of causality by including additional tests and checks into the approach such that it can be automatically detected. In addition, we will address the problem of automatically detecting states and actions from a given event log.

Acknowledgments This research was supported by the NWO TACTICS project (628.011.004) and Lunet Zorg in the Netherlands.

References

- [1] W. M. P. Van der Aalst, "Data science in action," in *Process Mining*, pp. 3–23, Springer, 2016.
- [2] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, "Process mining in healthcare: A literature review," *Journal of biomedical informatics*, vol. 61, pp. 224–236, 2016.
- [3] R. S. Mans, W. M. P. Van der Aalst, and R. J. B. Vanwersch, *Process mining in healthcare: evaluating and exploiting operational healthcare processes*. Springer, 2015.
- [4] N. Martin, J. De Weerd, C. Fernández-Llatas, A. Gal, R. Gatta, G. Ibáñez, O. Johnson, F. Mannhardt, L. Marco-Ruiz, S. Mertens, *et al.*, "Recommendations for enhancing the usability and understandability of process mining in healthcare," *Artificial Intelligence in Medicine*, vol. 109, p. 101962, 2020.
- [5] R. S. Mans, M. H. Schonenberg, M. Song, W. M. P. van der Aalst, and P. J. M. Bakker, "Application of process mining in healthcare—a case study in a dutch hospital," in *International joint conference on biomedical engineering systems and technologies*, pp. 425–438, Springer, 2008.
- [6] M. Rovani, F. M. Maggi, M. De Leoni, and W. M. P. Van Der Aalst, "Declarative process mining in healthcare," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9236–9251, 2015.
- [7] A. J. Weijters and W. M. P. Van der Aalst, "Rediscovering workflow models from event-based data using little thumb," *Integrated Computer-Aided Engineering*, vol. 10, no. 2, pp. 151–162, 2003.
- [8] C. W. Günther and W. M. P. Van Der Aalst, "Fuzzy mining—adaptive process simplification based on multi-perspective metrics," in *International conference on business process management*, pp. 328–343, Springer, 2007.
- [9] S. J. Leemans, D. Fahland, and W. M. P. van der Aalst, "Discovering block-structured process models from event logs—a constructive approach," in *International conference on applications and theory of Petri nets and concurrency*, pp. 311–329, Springer, 2013.
- [10] M. Prodel, V. Augusto, X. Xie, B. Jouaneton, and L. Lamarsalle, "Discovery of patient pathways from a national hospital database using process mining and integer linear programming," in *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1409–1414, IEEE, 2015.
- [11] R. Gatta, M. Vallati, C. Fernandez-Llatas, A. Martinez-Millana, S. Orini, L. Sacchi, J. Lenkovicz, M. Marcos, J. Munoz-Gama, M. Cuendet, *et al.*, "Clinical guidelines: a crossroad of many research areas. challenges and opportunities in process mining for healthcare," in *International Conference on Business Process Management*, pp. 545–556, Springer, 2019.
- [12] J. J. Koorn, X. Lu, H. Leopold, and H. A. Reijers, "Looking for meaning: Discovering action-response-effect patterns in business processes," in *International Conference on Business Process Management*, pp. 167–183, Springer, 2020.
- [13] Z. D. Bozorgi, I. Teinemaa, M. Dumas, M. La Rosa, and A. Polyvyanyy, "Process mining meets causal machine learning: Discovering causal rules from event logs," in *2020 2nd International Conference on Process Mining (ICPM)*, pp. 129–136, IEEE, 2020.
- [14] M. L. van Eck, N. Sidorova, and W. M. P. van der Aalst, "Discovering and exploring state-based models for multi-perspective processes," in *International Conference on Business Process Management*, pp. 142–157, Springer, 2016.
- [15] W. G. Cochran, "The χ^2 test of goodness of fit," *The Annals of mathematical statistics*, pp. 315–345, 1952.
- [16] M. L. McHugh, "The chi-square test of independence," *Biochemia medica*, vol. 23, no. 2, pp. 143–149, 2013.
- [17] R. A. Fisher and F. Yates, *Statistical tables: For biological, agricultural and medical research*. Oliver and Boyd, 1938.
- [18] A. Agresti, *Categorical data analysis*, vol. 482. John Wiley & Sons, 2003.
- [19] R. G. J. Miller, *Simultaneous statistical inference*. Springer Series in Statistics, 1981.
- [20] W. Haynes, *Bonferroni Correction*, pp. 154–154. New York, NY: Springer New York, 2013.
- [21] F. Mannhardt and D. Blinde, "Analyzing the trajectories of patients with sepsis using process mining.," in *RADAR+ EMISA@ CAiSE*, pp. 72–80, 2017.
- [22] F. Mannhardt, M. De Leoni, H. A. Reijers, and W. M. P. Van Der Aalst, "Decision mining revisited—discovering overlapping rules," in *International Conference on Advanced Information Systems Engineering*, pp. 377–392, Springer, 2016.
- [23] M. De Leoni and W. M. P. van der Aalst, "Data-aware process mining: discovering decisions in processes using alignments," in *Proceedings of the 28th annual ACM symposium on applied computing*, pp. 1454–1461, 2013.
- [24] N. Gupta, K. Anand, and A. Sureka, "Pariket: Mining business process logs for root cause analysis of anomalous incidents," in *International Workshop on Databases in Networked Information Systems*, pp. 244–263, Springer, 2015.
- [25] S. Suriadi, C. Ouyang, W. M. P. van der Aalst, and A. H. M. ter Hofstede, "Root cause analysis with enriched process logs," in *International Conference on Business Process Management*, pp. 174–186, Springer, 2012.