Towards Understanding Aggressive Behavior in Residential Care Facilities Using Process Mining

Jelmer Koorn[1(✉)], Xixi Lu[1], Henrik Leopold[2,3], and Hajo A. Reijers[1]

[1] Utrecht University, Utrecht, The Netherlands
{j.j.koorn,x.lu,h.a.reijers}@uu.nl
[2] Kühne Logistics University, Hamburg, Germany
henrik.leopold@the-klu.org
[3] Hasso Plattner Institute, University of Potsdam, Potsdam, Germany

Abstract. People with intellectual disabilities often live in residential care facilities that aim to provide their clients with the best possible quality of life. Aggressive behavior forms a threat to not only the quality of life of the clients, but also the safety of the staff. This study aims to uncover the dynamics underlying the evolution of aggressive behavior in people with intellectual disabilities. We take a process mining approach to analyze patterns of aggressive behavior. More specifically, we analyze data from 1,115 clients from a Dutch residential care facility over a period of three years. Our results show that there are two different groups of cases: those exclusively showing the same type of aggressive behavior and those who show mixed types of aggressive behavior. What stands out is that physical aggression towards other people plays a key role in the patterns of aggressive behavior. The results were validated with a behavioral expert from the care organization.

Keywords: Process mining · Aggressive behavior · Healthcare

1 Introduction

Many people with intellectual disabilities live in residential care facilities, where they can get support for their daily needs. A central goal of these facilities is to make sure that their clients have an appropriate quality of life. While this primarily means the care facilities attend to the needs of their clients, it also requires them to deal with undesired behavior of clients, such as aggression.

Aggression is a complex phenomenon and has many facets. It involves physical aggression towards other people, verbal aggression, physical aggression towards objects, self-injurious behavior, and sexually inappropriate behavior [5]. In general, aggressive behavior is considered a threat to both staff and clients. Studies have shown that staff members who deal with aggressive behavior are more likely to experience stress and even have burnouts [7,8,13]. What is more, aggressors can be met with severe restrictive measures (e.g., seclusion) as well as

suffer physical injuries as a consequence of their aggressive behavior [3,14,18]. Hence, there is a desire to minimize the impact of aggression on staff and clients and, ultimately, prevent aggression incidents from happening altogether.

To do so, many researchers have investigated which factors may contribute to aggressive behavior. By mostly focusing on client characteristics, they found links to factors such as: age [4,19], severity of intellectual disability [4,5], and gender [4,5]. However, some researchers argue that focusing on client characteristics is too limited [12]. More recent research has followed up on this and started investigating how characteristics from the aggression incident itself, such as time, location, and trigger, can help to understand aggressive behavior [15]. In that light, observational data describing aggression incidents are increasingly digitally recorded in Information Systems.

Nevertheless, what all these studies have in common is that they take a static perspective. Due to the increasing digitization of behavioral data, new opportunities emerge to analyze this phenomenon from new dynamic perspectives. That is, to consider the changes in behavior of clients over time rather than for a single incident or client. Such a perspective has the potential to uncover, for instance, how aggression evolves over time and whether different aggression incidents are related. Process mining is such an emerging field that provides techniques to support the analyses of data from a causal perspective [20].

Therefore, we use this paper to study aggression incidents using the technology of process mining. We analyze data from 1,115 clients from a Dutch residential care facility over a time period of three years. We find that, on a high level, we can distinguish between cases exclusively showing the same type of aggressive behavior and cases showing a variety of types of aggressive behavior. Moreover, we find that although the division into both groups is useful, the repetition of the same type of aggressive behavior is the most frequently observed behavior. Lastly, physical aggression towards other people plays a key role as it occurs most often and usually follows after any other type of aggressive behavior.

The rest of the paper is organized as follows. First, the current status of process mining in healthcare is discussed in Sect. 2. Then, Sect. 3 provides a detailed description of the methodology of this paper. This is followed by Sect. 4 describing the results of the analyses. In Sect. 5 we discuss these results and the limitations of this study. Finally, in Sect. 6 the conclusions are presented.

2   Process Mining in Healthcare

Process mining is a family of data analysis techniques that aims to discover, monitor, and improve organizational processes by analyzing data from so-called event logs [20]. These event logs are generated by various information systems that are used in organizations and, therefore, capture how organizational processes are actually executed. Process mining has been applied in various healthcare settings [17]. Among others, process mining has been used to analyze patient care processes [6,9], dentistry processes [1,10], and cancer treatment processes [2].

Table 1. A snippet of the raw data as an example

Client ID	Timestamp	Timeslot	Target	Means	Consequences
L002eR	04/09/2015	16:00–17:00	Themselves	Teeth	Visible injuries
L002eR	11/09/2015	17:00–18:00	Themselves	Hands	Visible injuries
LHZ02	03/05/2016	11:00–12:00	Staff member	Hands	Pain 5 min
LHZ02	22/05/2016	10:00–11:00	Objects	Hands	Damage to property
LH88E3	26/12/2016	19:00–20:00	Staff member	Verbal	Felt threathened
H030E	02/02/2017	14:00–15:00	Objects	Feet	Damage to property

Despite the general potential of process mining in healthcare, the application of process mining in this domain is often associated with particular challenges [11]. One of the most common issues in this context is the absence of accurate timestamps. Healthcare information systems often only capture the day of an activity and not the exact point of time. As a result, the exact order of certain activities remains unclear. Another issue is the absence of a clear start and end point of a process. Often, the data entries for a single patient span several years. This, however, does not mean that all this data relates to the same treatment process. It could be that the patient received several treatments at the same time or had a recurrence after a couple of years. We have encountered both issues, and the steps taken to handle these issues are discussed in the following section.

3 Methodology

3.1 Data Extraction and Overview

For our paper, we acquired a data set from a Dutch healthcare organization that operates 54 residential care facilities in the Netherlands. It specializes in providing care for people with mild intellectual disabilities (IQ between 50 and 70), borderline intellectual functioning (IQ between 70 and 85), co-occurring psychiatric disorders, and physical disabilities. The total number of clients this organizations cares for is about 3,000 (the exact numbers varies over the time period we consider). Our data set covers all aggression incidents reported on all wards starting from the 1st of January 2015 until the 31st of December 2017. The total number of incidents in our data set is 21,706.

Table 1 shows an extract of the raw data we obtained. We can see that each entry about an aggression incident includes: a reference to the aggressor (client ID), the day (timestamp), an approximate point of time (timeslot), information about the target (e.g. objects), the means (e.g. hands), and the consequences (e.g. pain 5 min).

In order to apply process mining to our data set, a number of requirements need to be met. The starting point of every process mining analysis is a so-called event log. These event logs must contain at least three specific attributes:

(1) a unique case identifier (case ID), (2) an activity description, and (3) an appropriate timestamp. As illustrated by the extract shown in Table 1, the raw data set does not fulfill these criteria. Among others, a case ID is missing (note that a client ID is not a case ID since a single client might be associated with several cases) and there is no clear notion of an activity. Against this background, our methodology includes several preparatory steps represented in Fig. 1. The first step is described above in the section *Data Extraction and Overview.*
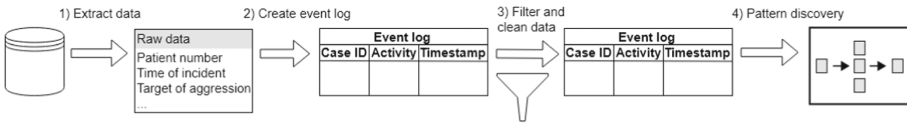

Fig. 1. Methodology

3.2   Event Log Creation

The second step was the creation of an event log that is suitable for a process mining analysis. To obtain such an event log, we had to introduce (1) a suitable activity attribute, (2) an appropriate case ID, and (3) a proper timestamp. Table 2 shows an extract of the final event log.

Activity Design. In order to introduce a suitable activity attribute, we designed a simple algorithm that categorized each incident into one out of four aggression categories: physical aggression towards other people (PP), self-injurious behavior (SIB), physical aggression towards objects (PO), and verbal aggression (VA). Due to privacy reasons we did not include data on sexually inappropriate behavior. We designed the algorithm using a rulebook approach based on [16]. The rules were adjusted based on the input from experts of the healthcare organization we collaborated with. It is important to note that a single incident can be associated with multiple types of aggressive behavior. In such cases only the most severe type of aggressive behavior is considered. This is in accordance with previous literature [16], who proposed a hierarchy of the severity of aggressive behavior. The four categories we consider are mentioned in their order of severity, that is, PP is most severe and VA is the least.

Case ID Design. Choosing the client ID as case ID may result in a process that spans over three years. To illustrate that this may lead to a distorted view on the data, consider the example of two clients. Client 1 has an SIB incident in early January 2015 and another in late December 2017. Client two has an SIB incident in early May 2016 and another in late May 2016. If the client ID is chosen as case ID, both clients would be considered to show the same pattern of aggressive behavior, although in the former case it is obviously hard to argue that the two incidents are related.

Based on a validation with experts, we decided to slice the data for each client in three ways: traces of 24 h, seven days, and one month. To this end, we created a case ID by combining the client ID with the year, month, week, and hour values from the timestamp. Thus, for example, one specific case for the client LHZ02 would be the case LHZ02-2015-01 including all incidents from January 2015. We found similar patterns using all three forms of data slicing. Thus, for consistency purposes we present our results on the basis of the month level.

Table 2. A snippet of the event log

Case ID	Activity	Timestamp
L002eR-2015-09	Auto mutilation (SIB)	04-09-2015 16:00
L002eR-2015-09	Auto mutilation (SIB)	11-09-2015 17:00
LHZ02-2016-05	Physical aggression towards others (PP)	03-05-2016 11:00
LHZ02-2016-05	Physical aggression towards objects (PO)	22-05-2016 10:00
LH88E3-2016-12	Verbal aggression (VA)	26-12-2016 19:00
H030E-2017-02	Physical aggression towards objects (PO)	02-02-2017 14:00

Timestamp Design. In our raw data set, the date (dd-mm-yyyy) of an incident was provided. To make the timestamp more precise we added available data about the timeslot and combined both into a single timestamp.

3.3   Data Cleaning and Filtering

The event log obtained after the preparatory steps outlined above contains a total of 21,706 aggression incidents related to 1,115 clients or 8,557 cases (client-months). To be able to detect relevant patterns using process mining, we applied three specific filters: (1) an event filter, (2) a time filter, and (3) a recurrence filter. After applying all three filters our final sample size used for process mining contained 16,794 incidents (78.2% of the total sample) spread over 822 clients included in 4,149 cases (48% of the total sample).

The *event filter* removed all incidents for which no aggression incident type could be determined (N = 322). The *time filter* was applied to exclude cases for which erroneous data was recorded for the considered time frame from 2015 to 2017 (N = 20). Finally, we applied a *recurrence filter* to exclude all cases that only contained one single activity (N = 4570). While this had a considerable impact, our analysis aimed to discover behavioral patterns. In order to detect a pattern at least two activities need to be included in a case.

3.4   Pattern Discovery

To discover relevant patterns, we used the commercial process mining tool Disco[1]. After we completed the exploratory analysis, it became clear that the data could be split into two large groups: cases that exclusively contain the same type of aggressive behavior (e.g., case L002eR-2015-09 in Table 2), and cases that contain a mix of aggressive behavior types (e.g., case LHZ02-2016-05 in Table 2). Based on this insight, we analyzed the patterns for each of these groups separately.

More specifically, we analyzed the data of each group from three angles: frequency, time, and variation. *Frequency* relates to the visual patterns and captures information about the number of incidents and transitions. *Time* captures details regarding duration between incidents. *Variation* looks at the various traces (i.e., specific orders of incidents) of which the patterns consist. When performing the analyses, we looked at both the grander schemes of all types of aggression combined per group of behavior as well as at each individual type of aggression within each group of behavior.

Finally, to make sure our understanding of the discovered patterns is correct, we validated our results with a behavioral expert from the care organization.
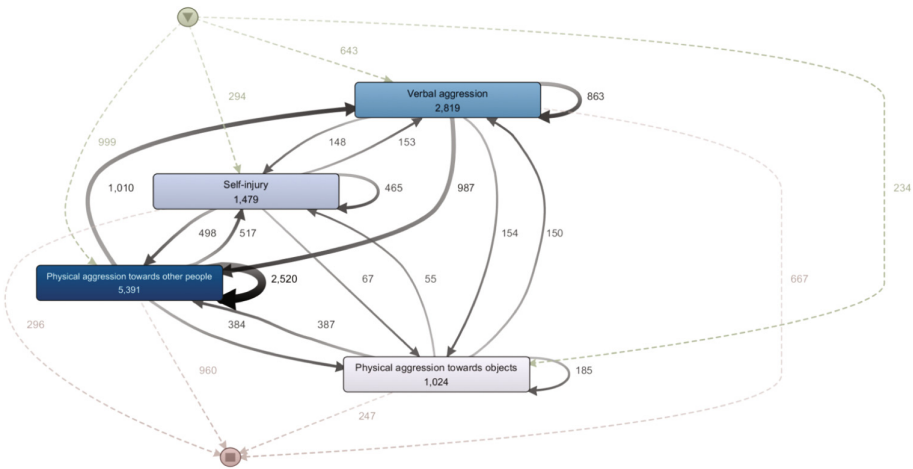

Fig. 2. The directly follows graph for the mixed behavior group, for a detailed account regarding the way this graph should be interpreted please refer to [20, p. 223].

---

[1] https://fluxicon.com/disco/.

4   Results

4.1   Mixed Behavior

The patterns resulting from all cases containing at least two types of aggressive behavior are summarized in Fig. 2. It shows the frequency of occurrence of each type of aggressive behavior as well as the transition frequencies (see respective arc label). In total, Fig. 2 captures 2,170 cases and 10,713 incidents.

Figure 2 illustrates that there are two types of aggressive behavior that are relatively frequent: physical aggression against other people (PP) and verbal aggression (VA). The other two types of aggression (physical aggression towards objects (PO) and self-injurious behavior (SIB)) account for less than 25% of all incidents. Furthermore, we can see that around half the cases start or end with an incident of PP. This indicates that PP serves as both enabler of and amplifier to other types of aggression.

Table 3. Relative transition frequencies

	VA	PP	PO	SIB
VA	**62.0%**	29.7%	4.3%	4.0%
PP	12.0%	**76.4%**	4.8%	6.8%
PO	15.9%	**44.1%**	35.2%	4.9%
SIB	7.2%	26.3%	3.0%	**63.5%**

Table 3 provides further insights into the underlying patterns by showing the relative transition frequencies (i.e. in how many percent of the cases we observe a transition from one type of behavior to another). The bold figures indicate the most frequent transitions and the underlined figures the second most frequent transitions for each type. The data allows for four interesting observations. First, the repetition of the same type of aggressive behavior is the most frequent pattern (see VA-VA, PP-PP, and SIB-SIB). Second, focusing on patterns without repetition, the most frequent transition is to PP (see VA-PP and SIB-PP). Third, PO shows the reverse of both previous points, its most frequent transition is to PP and its second most frequent transition is to itself (PO-PO). Fourth, PO and SIB are rarely followed by other types of aggressive behavior besides a repetition of the same behavior or a transition to PP.
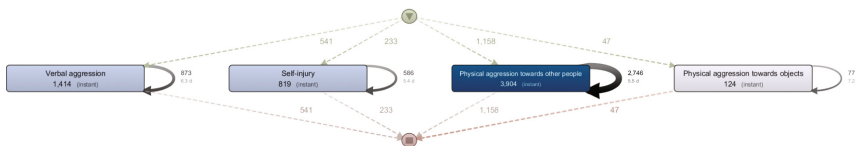

Fig. 3. The directly follows graph for the homogeneous behavior group.

4.2   Homogeneous Behavior

The results for the group with homogeneous behavior is summarized in Fig. 3. It shows the frequency of occurrence of each type of aggressive behavior as well as the transition frequencies (see respective arc label). In total, Fig. 3 captures 1,979 cases and 5,388 incidents.

The numbers in Fig. 3 show that PP is the most frequent type of repetitive behavior (60%) followed by VA (25%). Only about 10% of the incidents relate to SIB and 4% to PO incidents. The same distribution holds when we are looking at the number of cases rather than the number of incidents.

Looking into the average number of incidents per month allows for further insights. Interestingly, this number varies per type of aggression: SIB (3.5 incidents) and PO (3.4 incidents) have a higher average than PP and VA (both 2.6 incidents). In addition, when we consider the aspect of time, we observe that the median duration between two incidents is considerably longer for PO. Here the median between two PO incidents is the longest with 6 days, whereas the median for PP and VA is 4 days and SIB is 3 days. This shows that there is a negative relation between the average number of incidents per month and the median duration between two incidents. In other words, the higher the average number of incidents per month, the shorter the median duration between two incidents. However, PO is an exception to this since the average number of incidents per month is on the high end, but the median number of days between two incidents is high as well.

5   Discussion

Taking a holistic view at the results from our analysis, we identify three findings: (1) repetition of the same type of behavior occurs most frequently, (2) PP plays a central role, and (3) PO exhibits deviating patterns.

First, we see that *repeating* the same type of aggressive behavior is the most frequently observed transition between two incidents. Although the number of cases included in the mixed and in the homogeneous group are roughly equal, we can see that the most observed behavior within the mixed group is still a repetition of the same type of behavior. Through the validation with a behavioral expert we found that this can be explained by the fact that the behavior of clients is usually determined by their habits. As such, it is unexpected for clients to change their behavior. However, most clients also follow behavioral adaptation treatments aimed at changing this behavior. In this light, it is unexpected to see such a high percentage of repeating behavior. It might indicate that the treatments are often not successful.

Second, in both groups of aggressive behavior, PP is the most frequently observed behavior accounting for around 60% of the total number of incidents. Besides being the most frequently observed type of aggressive behavior, in Fig. 2, we see that PP in a mixed behavior environment is the epicenter of the system. There is relatively little interaction among the other forms of aggressive behavior except when accommodated through PP. Through the expert validation we found

that one possible explanation is that when an aggressive incident occurs, a nurse usually intervenes, thereby moving the target of aggression towards him/herself.

Third, PO represents an exception to the general patterns observed in both mixed and homogeneous behavior groups. We observed that cases showing this type of aggression follow a different evolution of behavior compared to the other types of aggression. We see for example that the pattern PO-PO is not as frequently observed as with the other types of aggression. In addition, if it is observed, the time in between two incidents is relatively long. This is interesting as it could indicate that PO is a fundamentally different type of aggression compared to the other types of aggression.

It is important to note that our findings are subject to a number of limitations. More specifically, we identify two main limitations to this study. First, our data set is not representative in a statistical sense. While our data set contains a considerable number of patients and incidents, we cannot extrapolate our findings to other care facilities. Second, our data set may contain different kinds of biases due to manual reporting. For example, nurses may report incidents in bulk, meaning multiple incidents are included in a single report at the end of the day or week in order to reduce administrative load. Although we know that this was not used very frequently, the existing bulk reports are counted as a single incident as there is no indication about how many incidents are reported in one bulk report. Another reporting bias relates to VA, SIB, and PO. From the discussion with experts we learned that these incidents are sometimes reported less frequently since they are perceived as less severe by staff.

6   Conclusion

Research looking into the aggressive behavior in clients with intellectual disabilities has exclusively investigated this phenomenon from a static perspective. In this paper we advanced on this by using a process mining approach to look into the evolution of aggressive behavior in clients with intellectual disabilities. This enabled us to obtain insights into the relations among different types of aggression and to infer patterns of aggressive behavior. We found that there are mainly two groups of clients: those with homogeneous and those with mixed aggressive behavior. Among others, we found that repetitive behavior is the most frequently observed behavior. In addition, results show that physical aggression towards other people plays a central role in a majority of the behavioral patterns.

With these insights, this research contributed to a better understanding of aggressive behavior aiding further development in this field. From a practitioner point of view, the discovered patterns can aid the development of prevention and treatment techniques. Despite the interesting findings, we are aware of the limitations of our study. Therefore, we plan to follow up on our work by further developing insights into the discovered patterns. We aim to identify causal relations between the static characteristics of incidents and the behavioral patterns, thereby uncovering more refined patterns of aggression. Furthermore, we plan to include additional care facilities to increase the external validity of our results.

References

1. Bakhshandeh, M., Schunselaar, D.M., Leopold, H., Reijers, H.A.: Predicting treatment repetitions in the implant denture therapy process. In: 2017 IEEE International Conference on Big Data (Big Data), pp. 1259–1264. IEEE (2017)
2. Binder, M., et al.: On analyzing process compliance in skin cancer treatment: an experience report from the evidence-based medical compliance cluster (EBMC^2). In: Ralyté, J., Franch, X., Brinkkemper, S., Wrycza, S. (eds.) CAiSE 2012. LNCS, vol. 7328, pp. 398–413. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-31095-9_26
3. van den Bogaard, K.J., Nijman, H.L., Palmstierna, T., Embregts, P.J.: Characteristics of aggressive behavior in people with mild to borderline intellectual disability and co-occurring psychopathology. J. Ment. Health Res. Intellect. Disabil. **11**(2), 124–142 (2018)
4. Cooper, S.A., et al.: Adults with intellectual disabilities: prevalence, incidence and remission of aggressive behaviour and related factors. J. Intellect. Disabil. Res. **53**(3), 217–232 (2009)
5. Crocker, A.G., Mercier, C., Lachapelle, Y., Brunet, A., Morin, D., Roy, M.E.: Prevalence and types of aggressive behaviour among adults with intellectual disabilities. J. Intellect. Disabil. Res. **50**(9), 652–661 (2006)
6. Fei, H., Meskens, N., et al.: Discovering patient care process models from event logs. In: Proceedings of the 8th International Conference on Modelling Simulation (MOSIM), pp. 10–12. Citeseer (2010)
7. Hensel, J.M., Lunsky, Y., Dewa, C.S.: Staff perception of aggressive behaviour in community services for adults with intellectual disabilities. Community Ment. Health J. **50**(6), 743–751 (2014)
8. Hensel, J., Lunsky, Y., Dewa, C.S.: Exposure to client aggression and burnout among community staff who support adults with intellectual disabilities in Ontario, Canada. J. Intellect. Disabil. Res. **56**(9), 910–915 (2012)
9. Kim, E., et al.: Discovery of outpatient care process of a tertiary university hospital using process mining. Healthc. Inform. Res. **19**(1), 42–49 (2013)
10. Mans, R., Reijers, H., van Genuchten, M., Wismeijer, D.: Mining processes in dentistry. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, pp. 379–388. ACM (2012)
11. Mans, R.S., van der Aalst, W.M.P., Vanwersch, R.J.B., Moleman, A.J.: Process mining in healthcare: data challenges when answering frequently posed questions. In: Lenz, R., Miksch, S., Peleg, M., Reichert, M., Riaño, D., ten Teije, A. (eds.) KR4HC/ProHealth -2012. LNCS (LNAI), vol. 7738, pp. 140–153. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36438-9_10
12. McClintock, K., Hall, S., Oliver, C.: Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. J. Intellect. Disabil. Res. **47**(6), 405–416 (2003)
13. Mills, S., Rose, J.: The relationship between challenging behaviour, burnout and cognitive variables in staff working with people who have intellectual disabilities. J. Intellect. Disabil. Res. **55**(9), 844–857 (2011)

14. Nieuwenhuis, J.G., Noorthoorn, E.O., Nijman, H.L.I., Naarding, P., Mulder, C.L.: A blind spot? Screening for mild intellectual disability and borderline intellectual functioning in admitted psychiatric patients: prevalence and associations with coercive measures. PLoS ONE **12**(2), e0168847 (2017)
15. Nijman, H., Palmstierna, T.: Measuring aggression with the staff observation aggression scale-revised. Acta Psychiatr. Scand. **106**, 101–102 (2002)
16. Nijman, H.L., et al.: The staff observation aggression scale-revised (SOAS-R). Aggress. Behav. Off. J. Int. Soc. Res. Aggress. **25**(3), 197–209 (1999)
17. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Processmining inhealthcare: a literature review. J. Biomed. Inform. **61**, 224–236 (2016). https://doi.org/10.1016/j.jbi.2016.04.007. http://www.sciencedirect.com/science/article/pii/S1532046416300296
18. Tenneij, N., Koot, H.M.: Incidence, types and characteristics of aggressive behaviour in treatment facilities for adults with mild intellectual disability and severe challenging behaviour. J. Intellect. Disabil. Res. **52**(2), 114–124 (2008)
19. Tyrer, F., et al.: Physical aggression towards others in adults with learning disabilities: prevalence and associated factors. J. Intellect. Disabil. Res. **50**(4), 295–304 (2006)
20. van der Aalst, W.: Data science in action. In: van der Aalst, W. (ed.) Process Mining, pp. 3–23. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49851-4_1