Process Mining to Generate Healthcare Pathways

Alexander Coles
Paul Carroll
Martina Fonesca

June 2023

Contents

1 Introduction

The increased digitisation of events/records within process flows is facilitating their analysis for evaluation and improvement. Examples of events range from accepting a purchase order in a business, receiving an email in a chain and administering medication in a hospital. With all this information now being electronically recorded, the challenge is to harness it and use it to provide value, whether this be by identifying bottlenecks, protocol violations or streamlining processes.

Process Mining is the term given to a family of techniques that have been developed to analyse these recorded events and process flows Van Der Aalst et al. (2012). Process Mining has three main application areas, these are discovery, conformance checking and enhancement. In process discovery, the output is a fact-based process model based on the recorded events in the database. Discovery is the most common process mining investigation, where the discovered process model is often not expected by the supervisor of the process. In conformance checking we measure how well the recorded processes fit a pre-defined process model. Conformance-checking techniques can therefore be used to check whether certain rules and policies are being abided by within an organisation. Finally, enhancement is focused on improving the existing process models by analysing how additional attributes affect the throughput times and frequencies of activities in the process.

All process mining techniques require an event log to start. An event log is a sequential recording of the events or activities that happened to a particular case (the person or object travelling through the process) with the associated time stamps of the events. Additional attributes for cases or individual events such as the resource responsible for performing the activity on the case allow additional analysis of the effect the attributes have on the process but are not essential to start process mining.

An event log may seem like a simple requirement in the modern age where event data is electronically recorded however there are many challenges to the curation of an event log suitable for process mining. The first of which is that event logs are assumed to be complete for their purpose, with no missing events cataloguing a particular process, however, numerous data quality issues can and do cause missingness in event logs. Another assumption is that the event log is trustworthy and that all the recorded activities in an event log have indeed taken place. The granularity of data needed to answer certain questions also may not be available, or the data may be spread over a number of sources requiring data linkage. Finally, the amount of data required for a detailed event log and its storage location may raise security concerns due to the identifiability of people (e.g. patients in a healthcare investigation) within the event log.

With all these requirements and concerns in mind, a few methodologies have been created to aid the adoption and use of process mining within organisations. The most well-known are the L* and PM2 methodologies Aalst (2011), Van Eck et al. (2015). Both Methodologies describe a similar approach to conducting a process mining project.

1.1 The Process Mining Methodology

In both the L* and PM2 process inning methodologies, the first step for a process mining project is the planning stage Aalst (2011), Van Eck et al. (2015). The obvious steps to this stage are to identify the target process to investigate and formulate some research questions that can be further refined throughout the study. Once the target process has been selected the availability of the data should be assessed along with the quantity (number of cases, events and attributes) and quality (how much missingness is present in the event and attribute data) of the available data. An understated part of the planning stage is the organisation of a project team that not only includes the process analysts but also the business expert or in healthcare specifically a clinical expert who ideally is familiar with both the data to be investigated and first-hand experience with the process to be investigated. In process mining, it is common for the process analyst to initially be unfamiliar with the process and data to be investigated so the acquisition of an expert familiar with the data will enable the analyst to field questions and results to the expert throughout the project in regularly scheduled meetings. This provides the expert with immediate value from the investigation but should also help refine and progress the investigation faster should the analyst encounter anomalies that require their input.

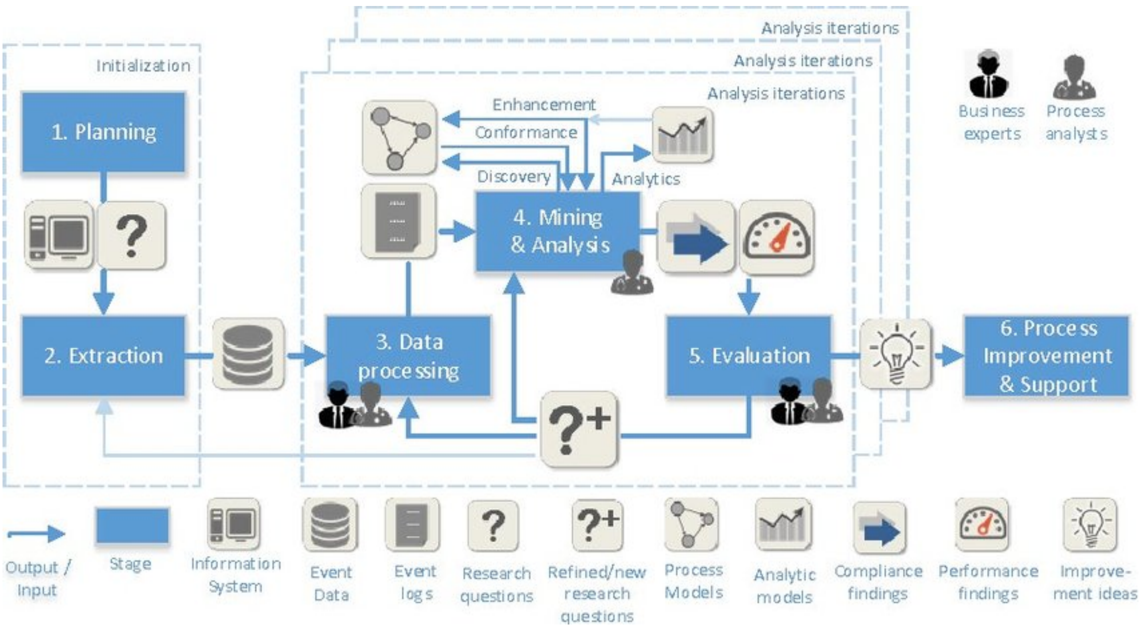

Figure 1: Diagram of the PM2 process mining pathway from Van Eck et al. (2015)

The second step in a process mining project is to extract the data from its original location, move it to a location suitable to perform analysis and construct an event log from the data. The challenges facing this stage include linking the relevant events of a process into one table from different sources and also including any attributes relevant to the research questions to be investigated. At this stage it is also a good opportunity to assess the scope of the data you have been able to extract and consider if its quality and quantity is sufficient to investigate the research questions discussed in the planning stage. If the data is deemed insufficient, the research questions may need to be changed or the viability of the project reconsidered. However, learning can already be fed back to improve the research potential of the data by advising on how it may be better recorded in the future (reducing the recording of unstructured data in favour of structured data where appropriate), reducing missingness and consolidating its collection into a location that is easier to access for prospective researchers.

Once an event log has been created. The process analysts can commence investigating the research questions using discovery, conformance and enhancement techniques. Depending on the research questions, the analyst may have to filter the event log using certain attributes contained within it. The granularity of the event log may be increased or reduced by splitting or regrouping specific activities. Clustering techniques can be brought into cluster processes with similar attributes together and create sub-logs. With each filtered log, Discovery techniques can visualise the processes contained within them, showing the frequency of commonly followed sequences of events (known as traces) or the time taken to get from one activity to another. Visualising these data-based/fact-based process models is often the goal in a process mining investigation, showing the true processes on what is commonly an assumed known process pathway. Conformance requires a process model either informed by the expert's understanding of the process pathway, informed from a prospective guideline for how the process should operate, or one derived from the data itself using various process mining algorithms. Once a process model has been agreed on various algorithms can be used to measure the traces within the event log to the process model. Low-fitting traces can then be further investigated to understand the root cause of their variation. This may advise alterations to the process or a re-imagining of the expected process to include the non-conforming pathways.

At every stage of the process mining investigation the advising clinical expert should be consulted on whether the findings are as expected, whether the research question should be changed to investigate a new anomaly or whether additional attributes can be added into the existing data to better understand their effect on the process. In the PM2 pathways, this happens at the evaluation stage Van Eck et al. (2015). However as stated, in practice it is ideal to be in regular contact with the advising expert to ensure the project is providing the best value to the organisation/team that implements the process. It is important to recognise that this exchange of ideas with the expert as new information is discovered throughout the process mining project commonly leads to process mining projects cycling through multiple iterations of data processing, discovery and conformance as the research further focuses on how certain attributes may affect the pathway.

The final step of a process mining project is the implementation of improvements to the process pathway. The findings of the process mining investigation can form part of the fact-based input to established process improvement frameworks such as Six Sigma Harmon & Trends (2010). Once changes have been made to the pathway, process mining can provide operational support to the new process, continuing to identify deviations from the norm and helping to inform future changes Van Eck et al. (2015).

1.2 Process Mining in Healthcare

With the increased use of electronic healthcare records that document patients' care pathways, process mining has begun being used to analyse healthcare pathways for research purposes. A systematic review by Batista & Solanas (2018) into the use of Process Mining in healthcare stated that oncology leads with the greatest number of published Process Mining articles, other healthcare areas that have seen process-mining investigations include diabetes, cardiology, radiology, dentistry, anaesthesia, chronic coughs, primary care, asthma, paediatrics, urology and the ambulance emergency care pathway.

The challenges facing process mining in healthcare were discussed in both the systematic review and papers dedicated to analysing these problems Batista & Solanas (2018), Munoz-Gama et al. (2022), Mans et al. (2012). One potential challenge is the temporal granularity of available event data, where many events may share the same date but do not have an associated time to place them in order Mans et al. (2012). The huge variation in healthcare pathways can also cause discovered processes to be uninterpretable, techniques to combat this include aggregating events together, clustering like traces together and filtering the event log by events or attributes Batista & Solanas (2018), Munoz-Gama et al. (2022). However, this may have the negative effect of removing infrequent behaviour that may have been of interest for the investigation. Acquiring research team members with relevant clinical knowledge to advise on the project is also recognised to be a challenge, Munoz-Gama et al. (2022) points out that nursing staff should also be included alongside physicians due to their deep involvement in healthcare pathways. The security concerns of storing and analysing potentially identifiable patient data are also highlighted by Batista & Solanas (2018), pre-processing may be required to de-identify the data set to an acceptable standard for research purposes. It should also be mentioned that healthcare data usually requires an ethics review prior to its use which can delay the research project, this time can be used to discuss the intended research project and familiarise the process analysts with the clinical process as advised by the clinical experts.

Some example process discovery studies in healthcare include an investigation into the clinical pathways of chemotherapy Baker et al. (2017), an exploration of the endometrial cancer pathways between their first referral to first treatment Kurniati et al. (2021) and a study to understand the looping pathways through the emergency department of patients with acute functional neurological disorders Williams et al. (2022). In Baker et al. (2017), only 5% of 535 breast cancer patients completed the planned six cycles of chemotherapy without having unplanned hospital contact. Over the six cycles, 31.6% of patients were found to have been admitted to hospital. Kurniati et al. (2021) found 50% of endometrial cancer patients took less than the expected 62 days to go from a referral to diagnosis. While Williams et al. (2022) found that people with acute functional neurological disorders that present to a UK city hospital tend to follow looping pathways through hospital healthcare centered around ED, with low rates of documented diagnosis and referral for psychological therapy. All of these studies primarily used process discovery techniques to present fact-based process models to clinical experts.

Examples of conformance-based research include an investigation into the concept drift of treatment pathways for cancer patients across a 15-year period Kurniati et al. (2020), and a study of the changes to the A&E and maternity pathways during the COVID-19 pandemic Puthur et al. (2022). In Kurniati et al. (2020), a general process model for the treatment of endometrial cancer was mined from the 15-year period between 2003-2017 and then compared against the data for each year within the timeframe. Changes to the general process path, and the frequency of individual events throughout those years were observed to detect changes to the treated pathway over time (known as concept drift). In Puthur et al. (2022), both A&E and maternity pathways were found to have reductions in the mean length of stay and a drop in the percentage of pathways conforming to normative during the COVID-19 pandemic,

2 Process Mining Case Study with East Midlands Ambulance Service

As part of an Internship Project with the NHS DART team for the use of Process Mining in healthcare, the following case study was enacted with the cooperation of the East Midlands Ambulance Service Trust (EMAS). This case study is split into the following 3 parts. An introduction as to why the UK Ambulance service is in crisis and a hypothesis as to how process mining can help in improving it, a description of the data set provided by EMAS and finally the project itself investigating the ambulance Job Cycle using EMAS data.

2.1 The UK Ambulance Service

In a report published by the House of Lords titled "Emergency healthcare a national; emergency" stated "patients are delayed at every stage of trying to access emergency healthcare" exacerbating existing health problems Public Services Committee (2023). A review by the Health Safety Investigation Branch stated that "Delays in the handover of patient care from ambulance crews to emergency departments (EDs) are causing harm to patients" Healthcare Safety Investigation Branch (2022). A review of the ambulance services in England from 2017-2018 "identified significant unwarranted variation in the proportion of patients that ambulance services take to hospitals across England" NHS England (2018). It is clear from these reports that there is room for improvement in the emergency healthcare services in England.

For context,NHS England (2018) reported that "Ten ambulance trusts respond to 10 million 999 calls every year across England" - "This costs around £1.8 billion, or £33 per person" - "60% of responses resulted in a patient being conveyed to AE" and " at any one time 40% of all patients in hospital beds in England will have been taken to hospital in an ambulance". The same report further hypothesises that reducing "avoidable conveyances to hospital" - "could release capacity equivalent to £300 million in the acute sector" NHS England (2018). There is therefore a strong financial incentive to ensure resources are being efficiently spent in the emergency healthcare service to reduce pressure on secondary care.

As part of NHS England (2018) recommendations for improvement they state that NHS Improvement and NHS England will "help trusts to identify opportunities for improvement by presenting their data in a comparable way" and advise trusts to "work together to improve performance and remove unwarranted variation". They also advise Ambulance trusts to ensure resources are used and monitored effectively by using "demand modelling software; optimising clinical support; improved rota and fleet management; adopting enabling technologies and implementing a make ready system" NHS England (2018).

In 2017, NHS England introduced the new performance standards designed to provide the most appropriate response for each patient based off of the severity of the patient's condition. Amongst these standards are the mean target response times for each call category as seen in figure 2. These categories are defined by the severity of the incident on a scale from 1-5 with a 1 being "life-threatening" and a 4 as "less urgent".

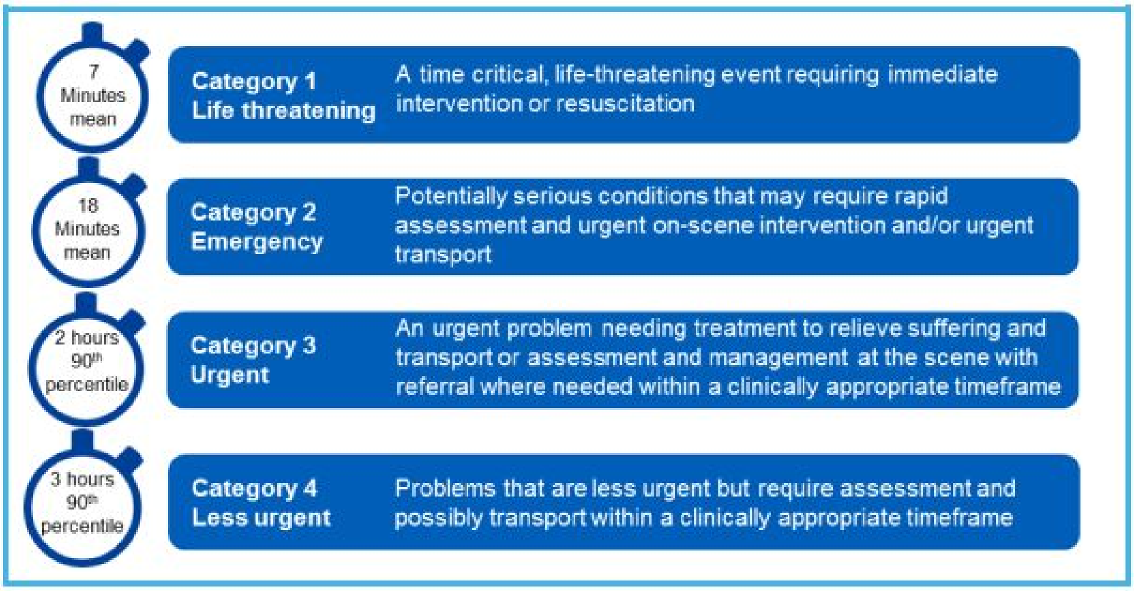

Figure 2: Ambulance service performance standards introduced by NHS England 2017 Public Services Committee (2023)

Across the Ambulance Trusts in England over half of all calls are triaged as category 2. The target mean target response time is set as 18 minutes. As the majority of all calls, improving the responses to category 2 and improving the treatment pathway is a major target for Ambulance Trusts.

The pathway or process an ambulance follows when responding to a 999 call is referred to as the Job Cycle. A breakdown of the Job Cycle can be seen in Figure 3. There are specific elements of the Job Cycle of great interest to ambulance trusts: these are the response time, the time spent on scene and Hospital handover time. Certain features that affect these times such as the geographical location of the incident which can obviously affect the response time are out of the control of the trusts. However, other features that affect these times can be operationally influenced. As the Job Cycle represents a process, it suits the application of process mining to discover a fact-based process model to measure the effect of certain attributes of the pathway and more specifically, investigate the previously described elements of interest and how certain attributes affect different steps in the Job Cycle.

2.2 East Midlands Ambulance Service Data

The East Midlands Ambulance Service is one such Ambulance Trust looking to increase the analysis of its recorded data documenting its Ambulance Job Cycles to help improve its efficiency. Having contacted the NHS DART team for consultation, the Job Cycle data was identified as a suitable candidate for an investigation

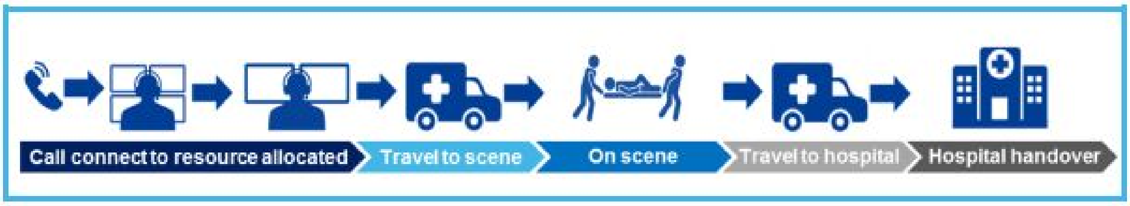

Figure 3: The stages of the Job Cycle Public Services Committee (2023)

using process mining techniques. As per the steps described in section 1.1 and the PM2 process mining methodology Van Eck et al. (2015), the investigation started by assembling a research team and assessing the available data and tooling required to perform the research. The assembled team included two process analysts and a number of clinical experts including a paramedic and two operations managers. The ambulance Job Cycle data is mostly automatically recorded into the Computer Aided Dispatch (CAD) system. Errors do occur in the recording of some events in the Job Cycle and where these are identified, efforts are made to manually correct these errors in recording. The data is not recorded in a typical event log format within the CAD system.

The research project was performed on an EMAS-provided laptop linked to their internal network due to data security concerns. Once this laptop was acquired we installed the required software to perform the analysis, namely Python, and the open-source process mining software PM4PY. Once the hardware and software were in place we were first provided with a dataset consisting of all the recorded Job Cycles for patients that were seen and treated and seen and conveyed between August and December 2022. This included 224,559 patients' Job Cycles. The proportion of each category can be seen in Table 1.

Table 1: Ambulance calls are assigned to a category based on the severity of the incident with a one being the most severe and a five being the least severe. This table shows the frequency of each call category received by East Midland Ambulance Service between the 1st of August and the 31st of December 2022.

Call Category	Number of Calls
1	41,493
2	152,525
3	28,177
4	458
5	1,906

As the original CAD data was not in an event log format suitable for process mining the next step, following the PM2 pathway, was to process the data into an event log. This step is always tailored to the original data being processed into an event log and is therefore difficult to make replicable for other datasets. However, the end result is an event log with three key columns, a case column, an event/activity column and a timestamp column. Stored within the CAD system for each Job Cycle were additional case and event attributes. A number of these were extracted at this stage to be incorporated into the event log. However, as described in section 1.1, it is common to iterate back through the process mining methodology and extract further attributes when additional research questions come about from further discussion with the clinical experts and the process analyst's increasing understanding of the process. Therefore, more attributes were extracted to the event later on in the research project. A complete list of the events in the event log can be seen in tabled 2 along with their frequency in the category call 2 data.

While constructing the event log, we were in touch with our clinical advisers on the project who expressed their interest in the category 2 call data above all the other categories, in line with what was discussed in in section 2.1. Therefore, we made the decision to filter down the data to category 2 call data prior to restructuring the data into an event log. This reduced the pre-processing time required to construct the event log. Having processed the CAD data into an event log we were now free to begin exploring the Category 2 Job Cycle data with process discovery.

Table 2:

Column Entry	Description	Frequency in Event Log
Patient ID	Patient Identifier of the Job Cycle	152,525
PCR Time	When the Patient Care Record detailing the Job Cycle was started for the incident. This usually occurs soon after the incident.	152,525
Incident Time	The time when the call operators receive the call for an ambulance.	152,525
Time Originated	The time at which the first ambulance crew is allocated to respond to the incident call.	152,257
Time Mobile	Time when the first ambulance starts moving towards the incident. Usually happens at the same time as Originated (Crew allocation) but can be soon after if the crew is not in the vehicle when receiving the incident allocation.	152,269
Arrive Scene Time	Time when the first vehicle arrives on the scene. Usually recorded automatically although errors can occur if the vehicle has to be parked far from the location of the incident.	152,206
Depart Scene Time	The time when the ambulance departs the scene. this can be with the patient if they are being conveyed to the hospital or without the patient if the patient is treated on scene and deemed ok to not attend hospital.	152,163
Arrive Destination Time	If the patient was conveyed to hospital, this is the time when the ambulance arrives at the hospital	98,404
Time of Care Transfer	Care transfer is when the patient has been handed over to the hospital staff and the ambulance staff begins to reset the ambulance ready to respond to another incident call.	98,668
RL Pre Alert Time	The ambulance staff can alert the Emergency Department that they are coming. This data point was badly recorded in the dataset and so was not included in the event log.	10,996
First NEWS Time Date	The first time that the National Early Warning Score was taken for the patient.	132,035
Last NEWS Time Date	The last time that the National Early Warning Score was taken for the patient.	81,776
First GCS Time Date	If the Glasgow Coma Scale was measured for the patient then this is the first time it was measured.	145,700
Last GCS Time Date	The last time that the Glasgow Coma Scale was taken for the patient.	123,108
IV Time	The first time any intravenous drugs were given to the patient.	17,486
IO Time	The first time any intraosseous drugs were given to the patient.	54
Drug Therapy	Details when any drugs were given to the patient and what the drug was.	91,305
DWInserted	A timestamp associated with the upload of the PCR to the CAD system.	152,525
DWUpdated	A timestamp associated with any updates to PCR in the CAD system.	152,525

2.3 Process Discovery

Perhaps the first thing to do when commencing process discovery is to produce what is called a directly follows graph (DFG) of the unaltered first event log. A DFG is a graph where every activity present in the log is shown with the traces navigating the activities represented by edges joining consecutive activities together. The frequency which each path or edge is followed is usually shown as an integer on the edge. DFGs do have limitations, an important one being that events with flexible ordering are usually shown in long looping paths and another being that DFGs can not show concurrency, where events can occur simultaneously. However, a DFG can provide an immediate understanding of how complex the process is, how frequent certain events are, which events are likely to follow others and, depending on the tool you have used to produce your DFG you can usually set it up to get an idea for the throughput time between consecutive events. Most likely, however, on a complex unfiltered process, the DFG may show what is called a spaghetti process model with some events able to occur in any order. This was the case for our initial DFG on the category 2 call data as shown in Figure 4. Certain events within the spaghetti figure 4 occurred in confusing sequences, (incidents occurring after the ambulance arrived on scene). These anomalous sequences indicated that some business rules were required to filter the data and enable the key pathways to become visible to the analysts.

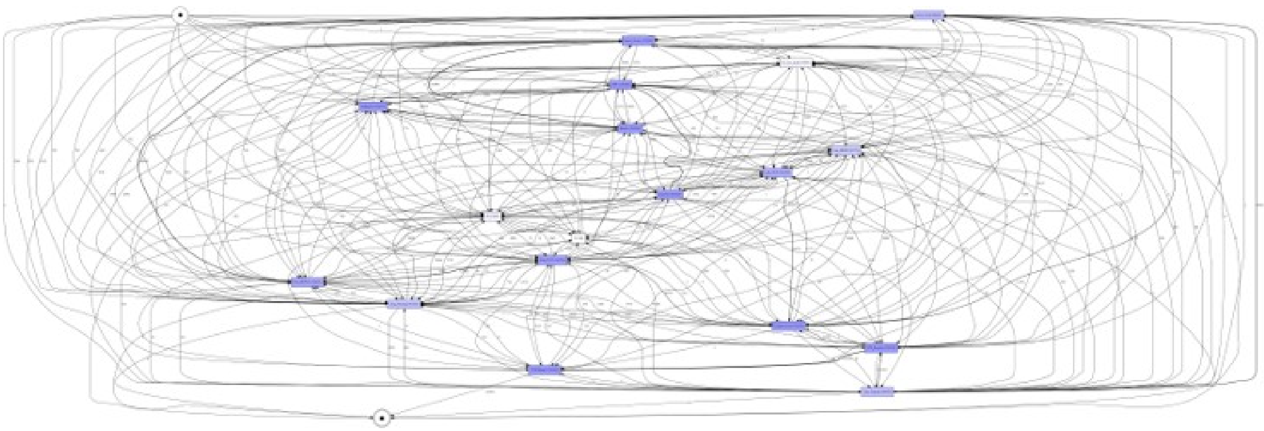

Figure 4: A directly follows graph of the unfiltered event log including 152,525 patients' Job Cycles with 19,754 variant traces.

The first filter applied focused on removing anomalous events. In the possible 18 activities within the event log, there were four activities that could occur at any time irrespective of the process of the ambulance Job Cycle, these events were first removed to prevent them interfering with further filtering. In the case of the category 2 ambulance data, our consultation with clinical experts revealed that certain key events of the ambulance Job Cycle should not be recorded in an order counter to the expected order. Following this, a minority of cases (528) were found to have these key events recorded out of order, the explanation from the clinical experts being that these events may have failed to be automatically recorded during the Job Cycle and have been manually inputted at some point after the Job Cycle with the wrong timestamp. These cases were dropped from the event log entirely to prevent these anomalies from influencing further analysis. This also brings into question how accurate the remaining traces were as these timestamps may still have been added retrospectively. However, the remaining data was the closest we had to a gold standard for the ambulance Job Cycles. The filtered DFG at this stage can be seen in 5.

Following the removal of anomalous traces and events, a final filter was applied to remove 7 "treatment" style events from the overall Job Cycle as they were able to occur at any point in the Job Cycle. While a focus on the treatment of a specific condition might benefit from the inclusion of these treatment events to observe finer patterns for these conditions, looking at all the conditions being treated in Category 2 meant that any sequence pattern from these events were clouded by the rest of the conditions being treated in the event log and so they were removed.

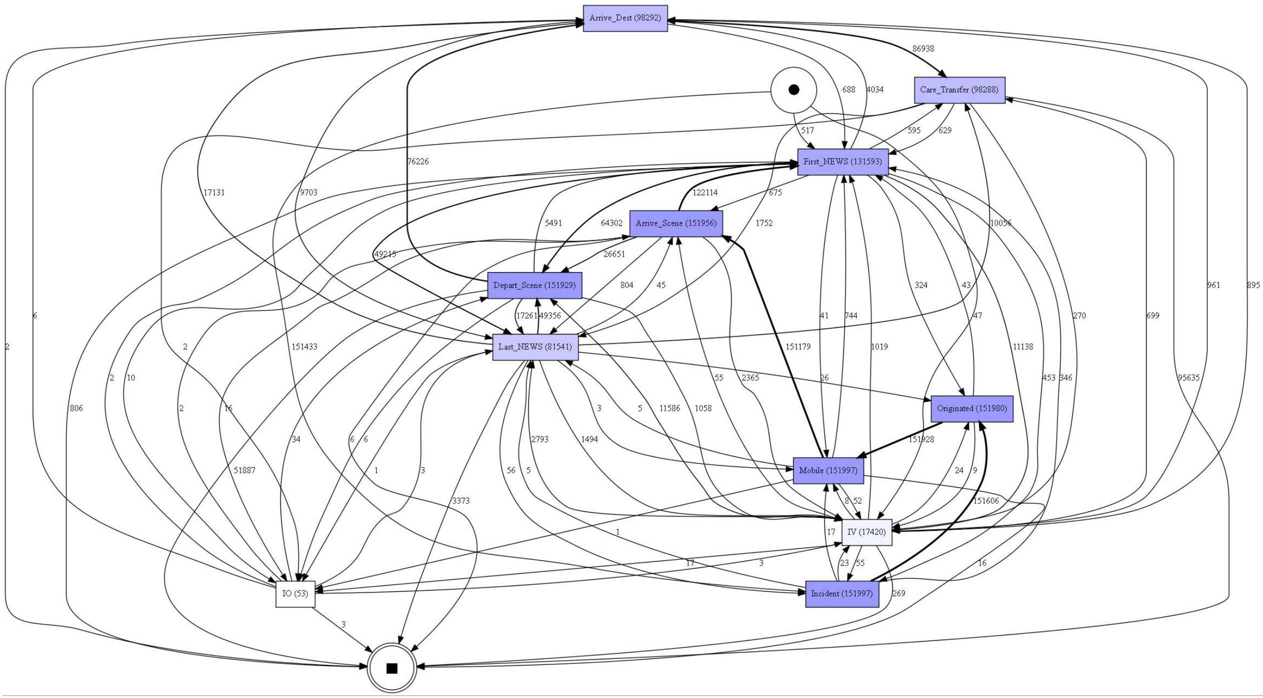

Figure 5: A directly follows the graph (DFG) of the event log with business rules applied to remove anomalous traces with events recorded in an order that could occur in the real world. This DFG shows 151,997 patient traces in 342 variants.

With these rules and filters applied we were left with 8 remaining trace variants however 2 of these variants accounted for 99 % of the log, these representing the patients that were treated and transported to the hospital and those patients that were treated without being transported. The 6 remaining variants were missing some of the key Job Cycle events with them having not been recorded. The 2 traces accounting for 99% of the of can be seen in figure 6.

Process discovery is all about finding a process model that can describe the process seen in the recorded data. While DFGs do provide a useful view of a process, other mining algorithms have been written to produce models that can show behaviour not clearly shown by DFGs. These mining algorithms usually present their mined process model in the form of a Petri net. New mining algorithms are being created constantly by the process mining community. However, most researchers and practitioners use established mining algorithms including the alpha miner, the inductive miner, the heuristic miner and fuzzy miner. Each miner has its strengths and limitations. The alhpa miner can not show loops in its discovered process model. The heuristic miner uses the frequency of sequential events to only include highly frequent sequences in its discovered process model which makes it useful in complex processes but it naturally excludes infrequent sequences that may be of interest to the analysts. The inductive miner guarantees that the discovered process is "sound" where all activities can be fired and all traces can reach the endpoint.

To showcase the ability of these models to show concurrency we applied the heuristic and inductive miners to the event log with the treatment events that can occur in any order included. The heuristically mined Petri net is shown in figure 7 while the inductive miner produced the petri net in figure 8. Neither of these models shows any more useful detail than that captured in the DFG, in the case of the ambulance Job Cycle it is evident that treatment actions can occur in any order while the key events noting the ambulance progression along the Job Cycle can only occur in a single order. The mined process models also have some clear flaws which will be discussed in the next section. However, these algorithms can be useful in other processes to discover concurrent activities that might not be expected. Unfortunately, there is no single way to choose the best miner and subsequent process model. Due to each miner's differences, it is best to evaluate the models and understand what behaviors they allow and remove. However, another method is to use conformance

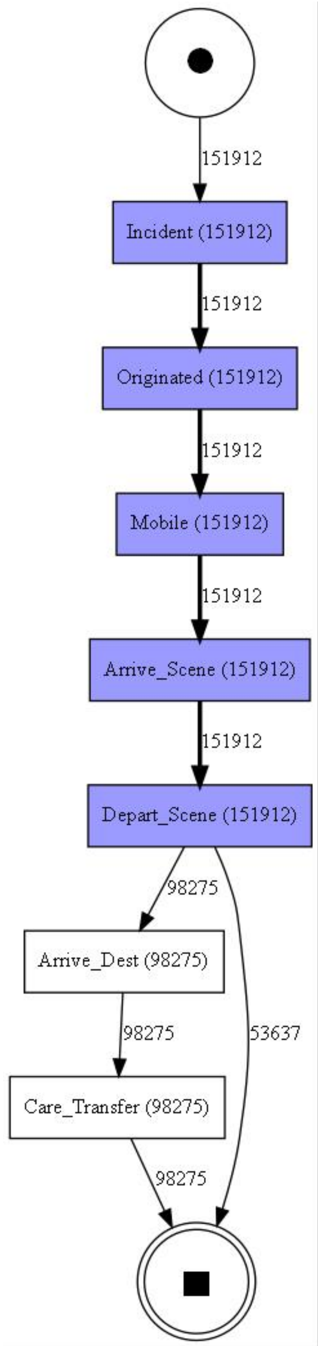

Figure 6: A Directly Follows Graph (DFG) showing the top 2 traces accounting for 99% of the Job Cycle traces. The 2 traces show patients who were treated and transported to the hospital and those patients who were only treated at the scene. The remaining 1% of traces not shown here were missing some of the activities shown in this DFG due to issues in their recording

techniques and measure metrics such as fitness, precision, and generalization of the process models against the event log they are modeling to assess the suitability of the mined model.

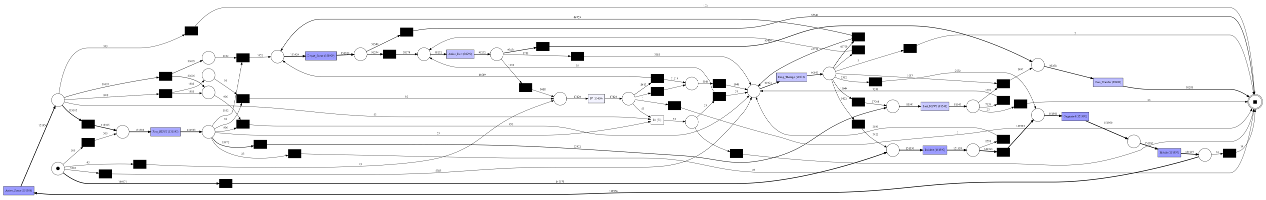

Figure 7: A petri-net process model produced using the Heuristic mining algorithm on the event log. The event log had patients with anomalously ordered events removed but the events recording treatments and vital signs measurements were left in so that the miners could show concurrency. This petri-net does allow both the First and Last NEWS to be measured prior to the Incident being called which is behavior not seen in the event log but it does ensure that the key events documenting the progression of the ambulance through the Job Cycle happen in the expected order.

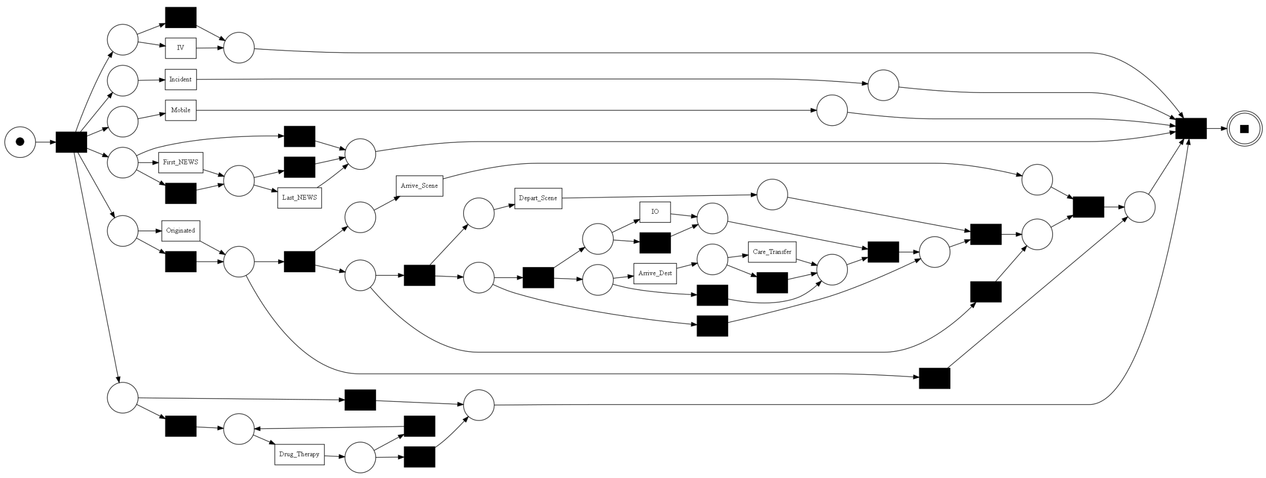

Figure 8: A petri-net process model produced using the Inductive mining algorithm on the event log. The event log had patients with anomalously ordered events removed but the events recording treatments and vital signs measurements were left in so that the miners could show concurrency. This petri-net allows behavior that cannot happen in the real world like the incident the ambulance is responding to happening at a later time than the patient being transferred to the hospital. Allowing a lot of behavior not seen in the event log makes this process model imprecise.

2.4 Conformance

Conformance checking refers to the measurement of how close an event log conforms to a defined process model. There are a few reasons for doing this, the most obvious is when a process model is set as an example or desired process model and you would like to know how close the real processes are aligned with the defined model. However, conformance techniques can also be used to understand how well a mined process model accounts for the behaviour in an event log. Another useful application of conformance checking is to observe changes in conformance to a process model over time, known as concept drift, by measuring conformance metrics across time intervals with the aim of identifying when the process changes and to understand why.

There are four key metrics used in conformance checking and process model evaluation. These are fitness, precision, generalisation and simplicity Buijs et al. (2012). Fitness is typically calculated using one of two algorithms, token-based replay or alignment-based replay. Fitness measures how well a process model can reproduce the behaviour seen in the log. Precision measures the amount of behaviour allowed by the log that

Table 3: The fitness, precision, generalisation and simplicity metrics measured from playing the event logs through process models mined from the event log using the heuristic and inductive miners shown in Figures 7 and 8 respectively.

Process Model	Fitness	Precision	Generalisation	Simplicity
Model mined by heuristic miner. Figure 7	0.91	0.96	0.87	0.53
Model mined by inductive miner. Figure 8	1.	0.59	0.95	0.62

is not seen in the event log. Generalisation quantifies the extent to which the process model will be able to show the future behaviour of the process. Finally, simplicity measures the number of activities that appear in the model compared to those in the event log and is often counter to the fitness and precision of a model where a simple model may not allow a lot of the behaviour seen in the log Buijs et al. (2012).

We can use these metrics to compare the heuristically mined process model in 7 to the Inductively mined petri net in figure 8. The metrics can be seen in table 3. The inductive miner reached a perfect fitness able to play all the traces seen in the event log perfectly. However, elements of the model look akin to a flower model where multiple activities can be played in any order. In addition, the model produced by the inductive miner is much less precise and while both allow behavior not seen in the event log, a lot of the behavior in the model produced by the inductive miner could not happen in the real world, like an ambulance arriving at the scene before an incident being called in. In this respect, while no model is perfect I believe the heuristic miner did a better job of displaying the process of the ambulance Job Cycle.

If we remove the treatment events from the model so we only include the key events showing the ambulance progression through the Job Cycle, the resulting heuristically mined process model can be seen in figure 9. If we measure the fitness of the wider event log without applying the business rules to ensure ordering but with treatment events still removed we get an average replay fitness for each trace of 99.89%. Again this doesn't show us any more information than was shown with the DFGs in figure 6. So in regards to the discovery process in ambulance Job Cycles, we understand that the key events showing an ambulance's progression through the Job Cycles should only happen in the expected order and the only way this can differ is through poor recording of the data. We also show that the treatment events happen in a flexible ordering to the patient at many points in the Job Cycle and there does not appear to be a clear trend when looking at all cases together, this indicates highly unique care pathways tailored to each patient's requirements as one might expect in an emergency situation.

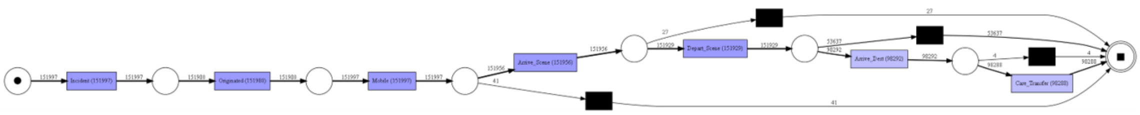

Figure 9: Process model created using the heuristic mining algorithm mined from the event log once the drug and vital sign measuring events were removed.

The final application of conformance checking we can show in this report is the observation of concept drift. Concept drift is where a process changes over time. We can observe this in the ambulance Job Cycles over the course of a year. Towards the end of this project we were provided additional category 2 data covering January to August 2023, pairing this with our previous data we can get a process model for the period from August 2022 - August 2023. Applying the heuristic miner to this data we measured the fitness of each month's data to the process model built from the year's data. Figure 10 shows the concept drift of the Job Cycle over time, with a clear drop in fitness during the winter months of 2022. As our process model was mined with the heuristic miner which includes frequent behavior and excludes infrequent behavior we can hypothesise that there is an increase in infrequent behavior during the winter months.

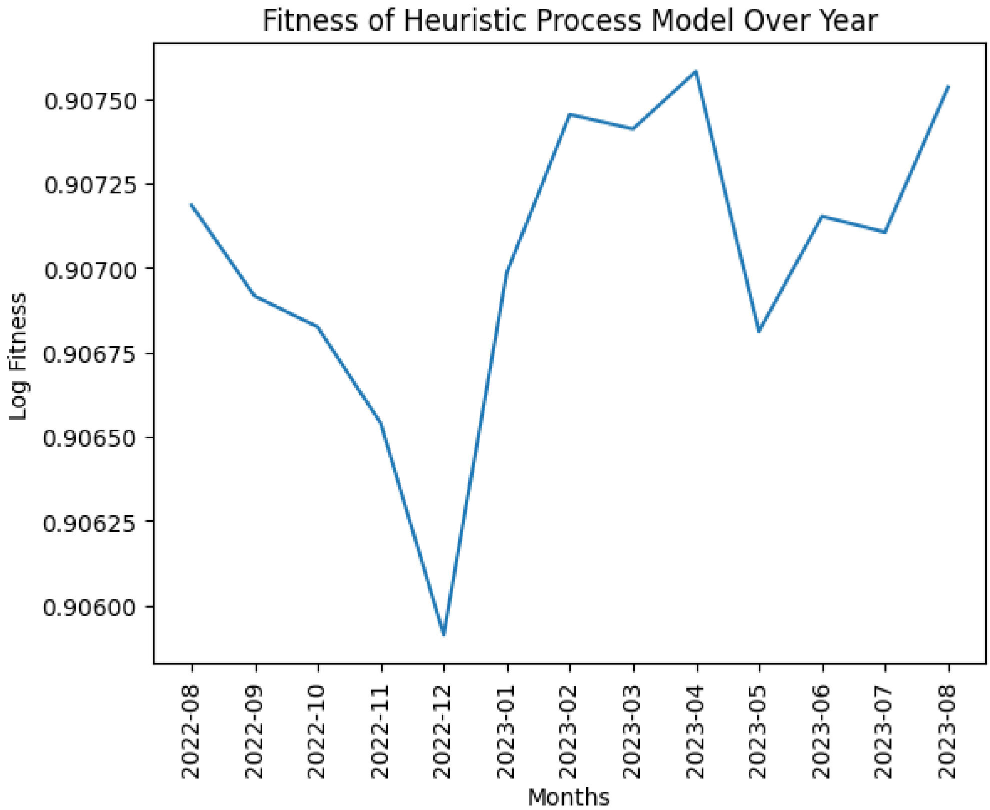

Figure 10: A graph showing concept drift or the change in a process over time. A process model was created from an event log of Job Cycle data for the entire year between August 2022 and August 2023 using the heuristic miner. This graph shows the fitness of each months data to the process model created from the entire year's data. The drops in fitness during the winter months of 2022 indicate an increase in behaviors not playable in the model.

2.5 Enhancement and Further Analysis

Enhancement is the last element in the trifecta of discovery, conformance, and enhancement that is process mining. It also seems to have the least focus in academic literature. Its most frequently used definition is “the extension or improvement of an existing process model using information about the actual process recorded in some event log” Yasmin et al. (2018), which could be argued to be unclear in its purpose. What is an ”extension” to the process model? In a literature review of studies that use enhancement, most studies were found to have looked at one of the three ”extensions”, resource performing the activity, timing of activity and times between activities, and other case attributes and their effect on the process model. Other studies looked ”repairing” at altering a mined or defined process model to better fit the activity in the log Yasmin et al. (2018).

In our case study with EMAS, our clinical experts were most interested in how the different attributes in the log affect different elements of the Job Cycle. The CAD data itself did have a number of attributes attached to the log already. However, we did add some attributes to enable further analysis. The complete list of attributes in the event log can be seen in the table 4.

From the available attributes, the Highest Qualification on Scene was the closest we had to a resource for an ambulance Job Cycle although they were involved with every activity on the Job Cycle. Many elements of the Job Cycle involve the ambulance traveling between locations which should not be greatly influenced by the qualifications of the health care workers in the ambulance. However, you might hypothesise that their level of expertise would affect the time spent on the scene treating the patient. We investigated this by producing a box plot showing the distribution of the times spent on the scene for each category of ”Highest Qualification on Scene”, this can be seen in figure 11. The majority of qualifications had a similar median time spent on the scene. However, there was an anomaly in the case of the Urgent Care Assistant (UCA).

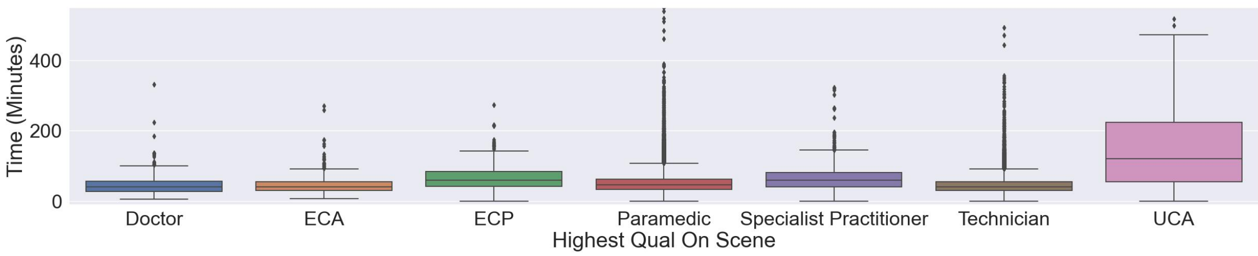

Figure 11: Boxplots showing the distribution in the time spent on the scene for category 2 calls with respect to the highest qualification of the health care team on scene. ECA and ECP stand for emergency care assistant and emergency care practitioner respectively while UCA stands for urgent care assistant. The anomaly of the UCAs’ increased time spent on scene was explained by EMAS stating that the UCAs often wait for additional support from other health care professionals.

At this point in a project, when finding an anomaly with how an attribute affects a certain part of a process, a wise step is to iterate back to process discovery with the cases only affected by this attribute. In the case of the Job Cycles where UCA was the highest qualification on scene there was no anomalous effect on the Job Cycle prior to the time spent on scene and we lacked any information after the patients were handed over to the emergency department at the hospital to see if this affected patient survival. However, we did raise this anomaly with EMAS and the explanation came back that these UCAs were likely waiting for additional assistance from other healthcare workers causing their time on scene to be extended.

Additional attributes of interest included the Geographical Barrier subdomain Decile (GBSD) and the clinical safety plan level. The GBSD is one of the Indices of Multiple Deprivation (IMD) which are values the UK government uses to quantify deprivation in its regions. These IMDs are published online and can be linked to Lower Layer Super Output Areas (LSOAs) and further to postcodes which allows us to add the IMDs at the postcode of the Job Cycle incident to the event log. We then measured the time between each of the activities of the Job Cycle for each IMD where these incidents occurred. The GBDs specifically quantify how far a location is from social services such as primary schools, post offices and hospitals and normalises the values between 1 and 10 where a 1 refers to an area far away from social services and a 10 is an area very close to

Table 4: A description of the attributes found and added to the event log.

Attribute	Description	Categorical or Continuous
Highest Qualification On Scene	The highest relevant qualification in the team who attended the scene of the incident.	Categorical
Gender of Patient	The gender includes Male, Female, Transgender and not known.	Categorical
Location of Incident	Either Home or Other. Over 95% of incidents occurred at Home.	Categorical
Week Day of Incident	The day of the week when the incident was called in.	Categorical
Clinical Category	The clinical category assigned to the call by the call handler from the description of the incident. Examples include Trauma, Obs/Gynae and Cardiovascular	Categorical.
Hour of each Activity in the Job Cycle	The hour in the day from 0-23 for each activity in the Job Cycle.	Continuous
Age of Patient	The age in years of the patient.	Continuous
NEWS	The First National Early Warning Score calculated from the first vital signs taken from the patient. NEWS take values between 1 and 20	Continuous
GCS	The Glasgow Coma Score takes values between 3 and 15	Continuous
Index of Multiple Deprivation (IMD) Decile	A value between 1 and 10 calculated from the weighted values of the other indices of multiple deprivation. We added the IMD deciles to the event log using the incidents recorded postcode in the CAD system ad the key.	Continuous
Income Index of Multiple Deprivation Decile	Measures the proportion of the population experiencing deprivation relating to low income. Takes values between 1 and 10.	Continuous
Employment Index of Multiple Deprivation Decile	Measures the proportion of the working-age population in an area involuntarily excluded from the labour market. Takes values between 1 and 10.	Continuous
Education Index of Multiple Deprivation Decile	Measures the lack of attainment and skills in the local population. Takes values between 1 and 10.	Continuous
Health Index of Multiple Deprivation Decile	Measures the risk of premature death and the impairment of quality of life through poor physical or mental health. Takes values between 1 and 10.	Continuous
Crime Index of Multiple Deprivation Decile	Measures the risk of personal and material victimisation. Takes values between 1 and 10.	Continuous
Barriers to Housing and Services Decile	Measures the physical and financial accessibility of housing and local services. Takes values between 1 and 10.	Continuous
Geographical Sub Domain Decile	Measures physical distance from services and facilities. Takes values between 1 and 10.	Continuous
Distance from hospital	Distance of the incident postcode from the receiving hospital postcode in meters calculated. We added this to the event log.	Continuous
Clinical Safety Plan Level	A value between 1 and 5 added to the event log using the hour of the day and date to identify the CSP level in a separate table.	Continuous

social services. We investigated how the response time of an ambulance (the time between the incident and the ambulance arriving at the scene) varied with the GBSD and plotted the distribution of response times in the boxplots shown in figure 12. The response time can be seen to follow a close to linear interrelationship with GBSD where incidents that happen closer to social services have faster response times.

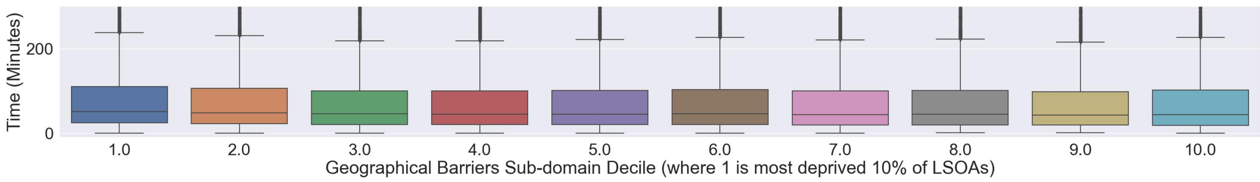

Figure 12: Boxplots showing the distribution in the response time of category 2 calls with respect to Geographical Barrier Sub Domain (GBSD) Decile of the incident postcode. The GBSD quantifies how far a postcode is from social services including schools, post offices and hospitals.The GBSD takes values between 1 and 10 where a value of 1 is most deprived.

The Clinically Saftey Plan (CSP) level attribute represents an internal quantity measured by EMAS to indicate the number of calls in any given hour requiring an ambulance. A CSP level of 1 represents a low number of calls that should not impact further incidents, a CSP level of 5 indicates severe pressure on the system which will impact further incidents. We measured the response time of Job Cycles against the CSP level for each call and plotted their distribution in the box plots in figure 13. An increase in CSP level can be seen to have a large impact on the response time for a call.

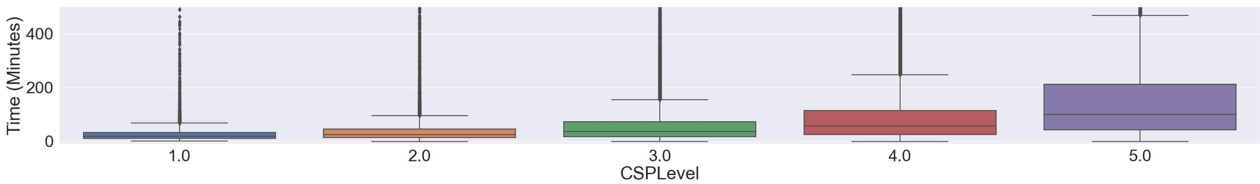

Figure 13: Boxplots showing the distribution in the response time of category 2 calls with respect to Clinical Safety Plan (CSP) level of the hour of the incident. The CSP level is a proxy for the number of calls requiring an attending ambulance in that hour. The CSP levele can take a value between 1 and 5 where 1 is defined as business as usual while a 5 indicates severe pressure on the system.

Finally, a useful overview of the time between events of interest can be seen by showing a DFG with the mean time between the events shown along the edges. Figure 14 shows that there is a bottleneck between the incident and the originated time stamp where crews are allocated to a job and a further bottleneck between the ambulance arriving at the hospital and actually handing the patient over the the emergency department when they are ready to receive the patient.

2.5.1 Outcome Prediction

At its core, an event log is a time series of data points. There are many different ways of investigating time series for various applications, with many making use of machine learning and AI. Bringing in machine learning allows for root cause analysis (trying to find out why a certain path might have been followed over another), investigating why a trace took as long as it did or even attempting to predict the end point of a patient's care. The PM4PY application in Python can easily be linked with the scikit-learn package to quickly apply machine learning to a problem.

In the case of the ambulance Job Cycle data we were interested in whether we could predict if a patient was to be transported to the hospital from early on in the ambulance's Job Cycle. Consulting with our clinical experts we asked when it would be most useful to make this prediction and what attributes we should include to make this prediction. It was agreed that the decision should be made after the first vital signs have been

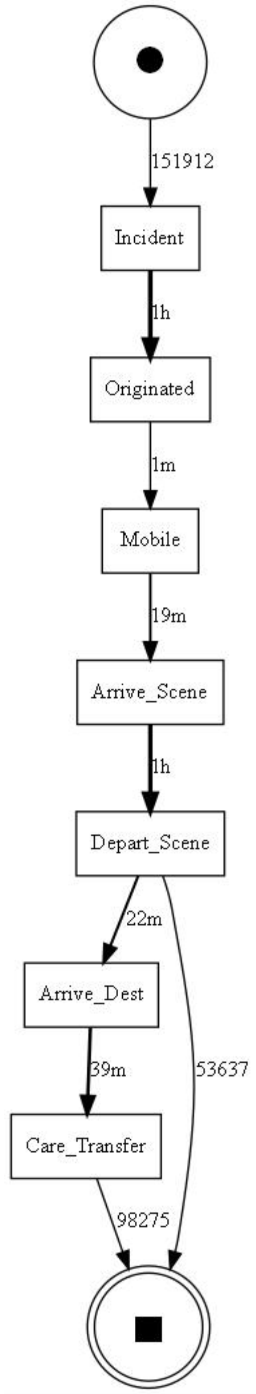

Figure 14: Directly Follows Graph showing the top 2 traces accounting for 99% of the Job Cycle traces. The 2 traces show patients who were treated and transported to the hospital and those patients that were only treated on the scene. The remaining 1% of traces not shown here were missing some of the activities shown in this DFG due to issues in their recording. A bottleneck can be observed between the call being received (Incident) and an ambulance crew being allocated (Originated) and another bottleneck between the ambulance arriving at the hospital (Arrive_Dest) and patient handover (Care_Transfer).

taken for the patient so this would include the first NEWS value of a patient and every event and attribute know up to that pint to predict whether a patient was to be transported to hospital or not.

We employed 3 machine learning models including a random forest, decision tree and logistic regression for this task. The data was split with 80% going into the training set and the remaining 20% set into a test set. We were immediately met with problems due to the imbalanced data set with 70% of the patients being transported to the hospital which led to all patients being predicted to have to go to the hospital. We undersampled this class in the training set to a 50:50 split and a second time to a 60:40 split with little improvement to results on the test set. However accuracy was not our priority in the task, we were interested in learning which attributes had the greatest influence on the classification. We were able to retrieve feature importances from the decision tree and random forest and the attributes coefficients from the logistic regression to better understand which attributes had the greatest influence on patients needing to be transported to the hospital. The feature importance for different attributes used by the decision tree can be seen in table 5.

Table 5: The sum of the importances for the different categories of input features used in a Decision Tree to predict whether a patient was to be taken to hospital using data available from early on in an ambulances Job Cycle.

Decision Trees Features	Feature Importance
Initial Clinical Category of Call	0.29
First Recorded National Early Warning Score (NEWS)	0.23
Age of Patient	0.17
Hour of Arrival on Scene	0.10
Index of Multiple Deprivation (IMD) Decile at Incident Location	0.075
Day of the Week	0.065
Highest Qualification on Scene	0.038
Gender of Patient	0.022
Location of Incident (Home/Other)	0.015
Whether specific Job Cycle activities have occurred prior to arrival on the scene	0.0013

As we can see, the initial clinical category of the patient inferred from the call to the ambulance has the greatest influence on the decision tree classifier when predicting if a patient was to be transported to hospital. This was followed by the first vital signs recorded from the patients which understandably indicates how severe a patient's condition is. It is interesting to see that the hour the ambulances arrive at the scene and the day of the week has a greater influence on the classifier than the highest qualification of the emergency care team on the scene. However, we hypothesise that there are additional features that influence the decision of the emergency care team on whether they transport a patient to hospital like the pressure that the destination emergency department is under and other aspects of the patient condition and care they are receiving that is not recorded in the CAD data. If these variables could be included in the model it may more accurately predict whether a patient needs to be transported to the hospital.

3 Discussion, Recommendations and Conclusion

We have presented a case study for the use of Process Mining techniques on an event log assembled from Category 2 Job Cycle data from the East Midlands ambulance service. Applying the business rules to the data and using Directly Follows Graphs enabled the observation of the two most common recorded traits followed by ambulances responding to category 2 data, those who were seen and treated and those who were treated and conveyed. Both the DFGs and the process models produced by process mining algorithms show that a number of events detailing drug treatments to the patient can occur at any time in the Job Cycle. Conformance techniques were used to measure how well the mined process models fitted the event log data. However, the variation in the treatment activities also caused the mined models to allow events including the call for an ambulance to happen in a sequence counterintuitive to real ambulance Job Cycles. Removing these treatment events and mining a new model only showed the paths already made visible in a DFG. Conformance techniques were also able to show a change in process during the winter months of 2022 compared to the process model mined over the whole August 2022 to August 2023 period. In our Enhancements investigation the timings between key elements of the ambulance Job Cycle were measured with respect to different attributes from the event log to infer their effect on the process, a DFG showing the timings between events was produced to observe bottlenecks and machine learning models were trained to predict if a patient was to be transported to the hospital with the secondary aim to understand the influence certain attributes had on the prediction.

The Process Discovery and Conformance Techniques used in this case study showed the expected path of an ambulance as it progressed through the Job Cycle while also showing that the order of the recorded treatment events can vary according to a patient's care. Given further time in this project we would like to have refined our process discovery investigation to individual clinical categories to see if the variation in the order of the treatment events reduced.

In a process like an ambulance Job Cycle which contains very few events that imply any structure to the pathway and some events that can happen in any order, we were only able to observe the extremely structured and expected pathways. From reading the literature, process discovery seems to be the most rewarding when used to investigate an unknown process or one where variation may be suspected but not known. Conformance can also help identify unexpected variations where a process model is known.

From an outcomes perspective, the most interesting part of this project has been observing the differences in the ambulance Job Cycles with respect to the different attributes. The enhancements side of process mining may also provide the most benefit for healthcare, identifying bottlenecks and attributes that might contribute to a patient's outcome may allow processes to be altered to improve a patient's outcome. This leads to our first recommendation of prospective process mining projects, if data on the survival or outcome of a patient is available, this should be included to enable analysis on whether certain pathways affect the outcome. If a patient's survival isn't relevant as an outcome to a pathway then a relevant outcome should be defined to compare pathways in the process and measure the pathways duration and attributes influence on the outcome.

As a second recommendation, the attributes in this event log enabled all of the investigations beyond simply discovering the process models, and given more time would enable the process discovery on the paths with specific attributes like clinical call categories. Therefore our second recommendation would be to assess the available attributes when first extracting an event log and to include as many as possible in the extraction. These attributes may enable investigations on their own or contribute to models investigating the route cause of certain paths being taken, the duration of paths, or the prediction of outcomes.

Our final recommendation would be to stress the advice set by Van Eck et al. (2015), to include business or rather clinical experts in the investigation. The Process analysts usually do not have exposure to the process, data or its method of recording prior to the investigation and so the insights that someone familiar with these can provide to the analysis not only provide the analyst with context to better understand the data but can also speed up the investigation and enable an exchange of ideas for the direction of research, potentially refining and improving the investigation of the process. Ideally, the process analysts and clinical experts would meet on a regularly scheduled basis where the analysts can share progress using the data, field questions and discuss any obstacles encountered with the data while discussing potential research avenues. When performing a data analysis project with healthcare data, this recommendation would likely benefit any project, not just process mining.

In conclusion, we followed the PM2 process methodology and used Discovery, Conformance, and Enhancements techniques on Category 2 Job Cycle data provided by East Midlands Ambulance Service. Data quality issues were tackled with business rules defined with help from clinical experts at EMAS. The 2 most common Job Cycle traces were observed in the data with bottlenecks in the processes identified. The effect on attributes recorded in the event log on the time between specific events were measured. Future work should focus on measuring how pathways, activities and attributes effect a patient's survival to ensure we are providing the best care to patients

References

Aalst, W. (2011), *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Springer.

Baker, K., Dunwoodie, E., Jones, R. G., Newsham, A., Johnson, O., Price, C. P., Wolstenholme, J., Leal, J., McGinley, P., Twelves, C. et al. (2017), 'Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy', *International journal of medical informatics* **103**, 32–41.

Batista, E. & Solanas, A. (2018), Process mining in healthcare: a systematic review, *in* '2018 9th international conference on information, intelligence, systems and applications (IISA)', IEEE, pp. 1–6.

Buijs, J. C., Van Dongen, B. F. & van Der Aalst, W. M. (2012), On the role of fitness, precision, generalization and simplicity in process discovery, *in* 'On the Move to Meaningful Internet Systems: OTM 2012: Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012, Rome, Italy, September 10-14, 2012. Proceedings, Part I', Springer, pp. 305–322.

Harmon, P. & Trends, B. P. (2010), *Business process change: A guide for business managers and BPM and Six Sigma professionals*, Elsevier.

Healthcare Safety Investigation Branch (2022), 'Interim bulletin: Harm caused by delays in transferring patients to the right place of care', https://hssib-ovd42x6f-media.s3.amazonaws.com/production-assets/documents/hsib-interim-report-harm-caused-by-delays-in-transferring-patients-to-the-righ_EPeMfuS.pdf[Accessed 17 October 2023].

Kurniati, A. P., McInerney, C., Zucker, K., Hall, G., Hogg, D. & Johnson, O. (2020), 'Using a multi-level process comparison for process change analysis in cancer pathways', *International Journal of Environmental Research and Public Health* **17**(19), 7210.

Kurniati, A. P., Rojas, E., Zucker, K., Hall, G., Hogg, D. & Johnson, O. (2021), 'Process mining to explore variations in endometrial cancer pathways from gp referral to first treatment'.

Mans, R. S., Van der Aalst, W. M., Vanwersch, R. J. & Moleman, A. J. (2012), Process mining in healthcare: Data challenges when answering frequently posed questions, *in* 'International Workshop on Process-oriented Information Systems in Healthcare', Springer, pp. 140–153.

Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O. A., Sepúlveda, M., Helm, E., Galvez-Yanjari, V., Rojas, E., Martinez-Millana, A., Aloini, D. et al. (2022), 'Process mining for healthcare: Characteristics and challenges', *Journal of Biomedical Informatics* **127**, 103994.

NHS England (2018), 'Operational productivity and performance in english nhs ambulance trusts: Unwarranted variations', https://www.england.nhs.uk/wp-content/uploads/2019/09/Operational_productivity_and_performance_NHS_Ambulance_Trusts_final.pdf[Accessed 17 October 2023].

Public Services Committee (2023), 'Emergency healthcare: a national emergency.', https://committees.parliament.uk/publications/33569/documents/187215/default/[Accessed 25 September 2023].

Puthur, C., Aljebreen, A., McInerney, C., Mebrahtu, T., Lawton, T. & Johnson, O. (2022), Measuring the impact of covid-19 on hospital care pathways, *in* 'International Conference on Process Mining', Springer, pp. 391–403.

Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., Van Den Brand, P., Brandtjen, R., Buijs, J. et al. (2012), Process mining manifesto, *in* 'Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I 9', Springer, pp. 169–194.

Van Eck, M. L., Lu, X., Leemans, S. J. & Van Der Aalst, W. M. (2015), Pm: a process mining project methodology, *in* 'International conference on advanced information systems engineering', Springer, pp. 297–313.

Williams, S., Southall, C., Haley, S., Dhafari, T. B., Kemp, S., Relton, S. D., Alty, J. E., Johnson, O., Graham,

C. D. & Maguire, M. (2022), ‘To the emergency room and back again: Circular healthcare pathways for acute functional neurological disorders’, *Journal of the Neurological Sciences* **437**, 120251.

Yasmin, F. A., Bukhsh, F. A. & Silva, P. D. A. (2018), Process enhancement in process mining: a literature review, *in* ‘CEUR workshop proceedings’, Vol. 2270, Rheinisch Westfälische Technische Hochschule, pp. 65–72.