White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/144202/

Version: Accepted Version

Proceedings Paper:
Johnson, OA, Ba Dhafari, T, Kurniati, A et al. (2019) The ClearPath Method for Care Pathway Process Mining and Simulation. In: Lecture Notes in Business Information Processing. BPM 2018: Business Process Management Workshops, 09-14 Sep 2018, Sydney, Australia. Springer, Cham, Switzerland, pp. 239-250. ISBN: 9783030116408. ISSN: 1865-1348.

https://doi.org/10.1007/978-3-030-11641-5_19

© Springer Nature Switzerland AG 2019. This is an author produced version of a paper published in International Conference on Business Process Management. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

The ClearPath Method for Care Pathway Process Mining and Simulation

Owen A JOHNSON[1,2], Thamer BA DHAFARI[1], Angelina KURNIATI[1,3], Frank FOX[1] and Eric ROJAS[3]

[1] University of Leeds, Leeds, LS2 9JT, UK
[2] X-Lab Ltd, Leeds, LS3 1AB [3] Telkom University, Bandung, Indonesia
[4] Pontificia Universidad Católica de Chile, Santiago, Chile
o.a.johnson@leeds.ac.uk

Abstract. Process mining of routine electronic healthcare records can help inform the management of care pathways. Combining process mining with simulation creates a rich set of tools for care pathway improvement. Healthcare process mining creates insight into the reality of patients' journeys through care pathways while healthcare process simulation can help communicate those insights and explore "what if" options for improvement. In this paper, we outline the ClearPath method, which extends the PM^2 process mining method with a process simulation approach that address issues of poor quality and missing data and supports rich stakeholder engagement. We review the literature that informed the development of ClearPath and illustrate the method with case studies of pathways for alcohol-related illness, giant-cell arteritis and functional neurological symptoms. We designed an *evidence template* that we use to underpin the fidelity of our simulation models by tracing each model element back to literature sources, data and process mining outputs and insights from qualitative research. Our approach may be of benefit to others using process-oriented data science to improve healthcare.

Keywords: Healthcare, Care Pathways, Process Mining, Process Simulation.

1 Introduction

The provision of high quality healthcare involves such complex systems that even those involved in their organization and delivery can feel it is impossible to comprehend. The *care pathway* is one well established and useful concept for bringing much needed clarity [1]. A care pathway describes the sequence of care that is recommended for patients with similar conditions requiring similar treatment [2] and is analogous to a *de jure* business process. Process mining of routine electronic healthcare records (EHR) can provide insight into the *de facto* compliance with care pathways including measuring performance and outcomes [3]. Although EHRs are a rich data source they present significant challenges of data quality, veracity and complexity [4]. In reality, care providers support multiple, simultaneous, diverse pathways for patients with highly variable personal needs and many of the interactions, events and decisions occur "off the radar" of the electronic systems.

To have utility, the outputs of process mining efforts need to be iteratively refined with the assistance of domain experts and then presented in a form that makes them accessible to wider stakeholders. In previous work [5] we have found agent based models with discrete event simulation presented through an interactive graphical representation to have been effective in stakeholder engagement. We developed the NETIMIS software tool (www.netimis.com) to support healthcare process simulation and this is now a commercial product and available for academic research. Healthcare process mining presents opportunities for understanding some of the reality of real patients' journeys through care pathways while healthcare process simulation can help communicate these discoveries and explore "what if" options for improvement [6]. In our approach, we extend simulation models to *fill in the gaps* by adding process steps missing from the health record data, adding information such as costs and incorporate insights from domain experts and stakeholder feedback.

In this paper, we present the ClearPath method as an extension of the established PM^2 process mining method [11] to incorporate healthcare process simulation modeling. We illustrate the ClearPath method through three case studies within UK hospitals, which show the discovered pathways for alcohol-related illness, giant-cell arteritis and functional neurological symptoms. In each case, the disease pathway needs to fit within busy hospitals where pathways of care for many diseases are taking place simultaneously.

2 Background

2.1 Process Mining in Healthcare

There is growing interest in process mining in healthcare [7]. Process mining can help answer *frequently posed questions* from clinicians and medical specialists [8] from control-flow, performance, conformance, and organizational perspectives [9]. Frameworks for process mining include the L* life-cycle model [10] which describes the life-cycle of a typical process mining project and more recently the Process Mining Project Methodology (PM^2) which incorporates iterations and gives detailed descriptions for six project stages [11]. Bozkaya et al. [12] propose a methodology called Process Diagnostics Method (PDM) which has been adapted to Business Process Analysis in Healthcare environments (BPA-H) [13]. Mans et al [3] provides a comprehensive guide to process mining in healthcare including health reference models and pathways. Finally, a question driven methodology to answer frequently asked questions was developed for healthcare [14]. The methodologies share similar steps including extracting event data, applying tools and techniques, analyzing the resulting models and improving based on stakeholder feedback. Process mining has been combined with process simulation [15] including to discover models for simulation [16] and at least once in healthcare [6]. In our approach, we have also combined process mining with traditional business process analysis methods to build a richer model than could be achieved by process mining alone.

2.2 Process Simulation in Healthcare

Brailsford *et al* (2017) reports on 50 years of healthcare simulation [17] and there are recognized frameworks for good practice in developing simulation models for healthcare [18]. In [6], we described the use of NETIMIS, a discrete event simulation tool for care pathway models which includes aspects of agent based approaches (patient characteristics affecting probabilities), and notions of time, cost and simplicity. We linked this to process mining and found challenges of EHR data quality (veracity, missing events, and missing data) and process complexity which suggested a mixed methods approach was required. There are many sophisticated tools for simulation but in this paper, we report on the use of NETIMIS.

NETIMIS is a cloud-based online service accessed using a standard browser and used to draw, share, evaluate and refine models of care pathways as runnable simulations. A NETIMIS model (see Figure 1) consists of a network of directed edges and nodes. The edges represent activities that take place over a period of time. The nodes represent events such as a decision point or the start or end of an activity. Pathways are animated with multiple moving tokens representing patients (shown as colored dots that move along the edges at a speed consistent with the time of the activity).

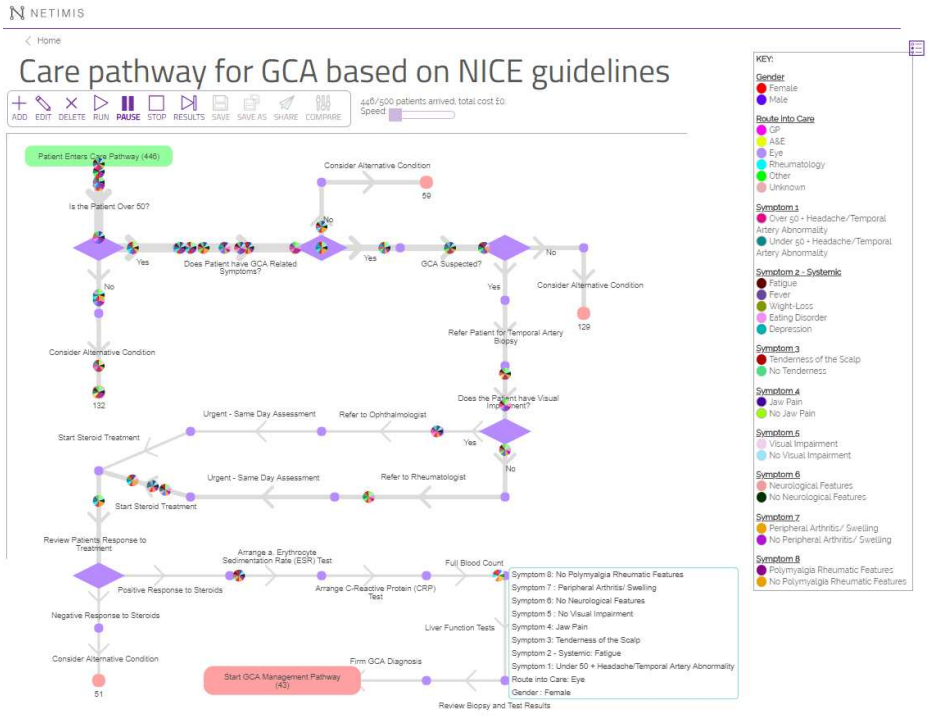

Figure 1: NETIMIS example showing a run of the Giant Cell Arteritis (GCA) care pathway derived from national guidelines. The simulation model can be run online at www.netimis.com/shared/5ad5fe6f7775761d4c5fd5ec

No patient-level data is needed for NETIMIS as the model is a simulation based on data from population-level analysis. Following agent-based approaches, patient tokens are randomized with attributes that reflect those of the base population and pathway junctions are given probabilities that are dependent on those attributes. Each patient token can be colored with a mini pie chart representing its attributes. The tool supports constraints that can lead to bottlenecks and probabilities that are affected by repetitions. Health outcomes are represented by pathway end nodes and each simulation run calculates total health economic costs and times based on the sum of individual costs and times assigned to each activity completed for all of the patients' care. Unlike Petri Nets, each token represents a single patient that cannot be "split" so there is no support for parallelism. Following the analogy of cars on roads, multiple patient tokens flow through care pathways to create a highly visual and engaging model. Users interact with the visualization through features including accelerate, pause, zoom, inspect, change, share and compare. "As is" and "to be" models can be run side by side so that differences can be explored visually.

2.3 Challenges using EHR data for process mining care pathways

EHR data is normally created for the purposes of patient treatment and administration and its secondary use for process mining of care pathways brings many challenges. Access to patient level data necessarily involves careful ethical, data protection and governance processes which can prove a significant administrative overhead. From the technical perspective, applying process mining to healthcare data is challenging due to its high volume and the diversity of the data types. Healthcare data ranges from administrative data such as admission times to machine generated vital signs, pathology results, diagnoses, and treatment procedures. Process mining all the available events in the EHR inevitably creates incomprehensible *spaghetti-like* models. Many EHRs are poorly designed to support easy, fast real-time use and with data being input by busy human beings doing demanding jobs it should come as no surprise that the data does not have the same provenance as clinical trials or registry data.

Data quality issues can be found at different levels. A missing field may only affect a single row whereas a large group of users who share a negative and hostile attitude towards their computer system might bias a complete data set. People, processes, organizational boundaries and cultures (and the EHR user interface) change over time and these changes will impact on the data. There is recognition that, the secondary use of EHR data for research demands validated, systematic methods of data quality assessment [19] and there is a correspondingly urgent need for process mining to incorporate techniques addressing these issues. Systematic logging techniques and the development of repair and analysis techniques should be in place and transparency around data cleaning and checking steps should be routine.

Four broad data quality dimensions for process mining of event logs were identified by [3]: missing, incorrect, imprecise and irrelevant. This adds 'irrelevant' to widely cited dimensions of EHR data quality that form the basis for the data quality assess-

ment method proposed by Weiskopf & Weng [19] and the valuable harmonized terminology produced by Kahn [20]. These dimensions were further detailed as 27 types of quality issues relating to the case, event and attribute levels of the data in an event log. The Process Mining Manifesto proposes a useful rating system for data quality ranging from 1-star to 5-star [10] and also emphasizes challenges of incompleteness, noise, granularity, event log complexity and concept drift. In [21] we describe the development of our *data quality management framework* to support the discovery, root-cause investigation, mitigation and careful documentation of these issues using software version control tools that are directly linked to the lines of software code in the Extract and Transform programs used to build the event log. The framework supports a close link between the design of individual process mining experiments and assessment of *fit-enough-for-use* quality.

3 The ClearPath Method

3.1 Rationale for an agile approach

Healthcare is a complex business and process mining and simulation in healthcare has some unique requirements. Clinicians work together across organizational and functional boundaries to meet the often highly individual needs of patients with complex conditions. We have found that domain experts such as clinical specialists can have quite limited views on the patient pathways beyond their specialism. Even a simple structured discussion with a number of specialists gathered around a whiteboard or process model has proved beneficial in improving pathways. We have used NETIMIS on multiple projects to structure these pathway discussions, elicit tacit knowledge and generate actionable insights including "what if" scenarios (www.netimis.co.uk/case-studies). Including patients in these discussions has proved incredibly powerful.

There is however a tension in healthcare improvement projects between the desire to drive radical change quickly and the demands of *evidence-based medicine* for detailed and careful reviews, particularly where adverse outcomes can be harmful and even fatal. Our approach has therefore been to adopt *agile methods* with time-boxed iterations which produce process simulation models of increasing fidelity that are backed by strong tooling (ProM, NETIMIS, data mining), traditional academic research methods (literature reviews, qualitative methods) and traditional business process analysis (observation, interviews, sample documents). We use the simulation model as the key output, and an evidence template to underpin the fidelity of model and present both to a Clinical Review Board at the end of each iteration.

3.2 Extending PM^2

The ClearPath method follows PM^2 with the following extensions.
Stage 1: Planning – *research questions* are often simply "what is the care pathway?" or "what does it look like?" and *composing project team* includes identifying a Clinical Review Board and pre-booking meetings so that iterations become time-boxed.

Stage 2: Extraction – the ethics of *extracting event data* when it is sensitive health data often mean long lead times so we make a data request and produce early iterations based on *transferring process knowledge* but with meticulous record keeping of artifacts (interview transcripts, whiteboard photos, journal references) from the investigation. In PM^2 business experts may be part of the project team but in healthcare these are often busy clinicians (or busy managers) so we engage them as interviewees within an iteration and/or in the Clinical Review Board at the end of each iteration.
Stage 3: Data Processing includes *filtering logs* to just include the patients of interest (those involved in the care pathway) and sometimes to *slice and dice* to examine subgroups of patients (e.g. frailty) and the pathway under different conditions (time of day, day of week) and in different locations. Healthcare reference models and coding systems such as SNOMED-CT and ICD-10 are used for *aggregating events*. We use our data quality framework to document data issues and software code solutions.
Stage 4: Mining and Analysis produces process models which are recreated in NETIMIS (currently by hand) with performance and compliance data added (e.g. mean durations, decision point probabilities) and documented in the evidence template. In Stage 4 we also add in details of the care pathway from our business process analysis, for example where activities are not recorded in the EPR and may also construct multiple models to examine different scenarios (e.g. weekends vs weekdays).
Stage 5: Evaluation includes *verify and validate* results against process insights from multiple reliable sources and root-cause investigations to *diagnose* anomalies. Stage 5 marks the end of each analysis iteration and takes the form of a Clinical Review Board (CRB) meeting where the evidence base, data quality management framework (assumptions, root cause analysis and mitigation decisions) are reviewed together with the latest "as is" and candidate "to be" NETIMIS model as runnable simulations. The CRB meetings are interactive and generally highly productive. The outcome of the meeting is to plan objectives for the next iteration.
Stage 6: Process Improvement and Support is marked by acceptance of the models and evidence by the CRB for *implementing improvements.* Models are published on NETIMIS and can be shared by other organizations and calibrated to local situations.

3.3 The Evidence Template

The ClearPath method focuses attention on the construction of simulation process models and the evidence template plays a key role in supporting early, low-fidelity models and an agile evidence building process. In the evidence template, each model element (patient agent, activity, decision point) and each model attribute (e.g. disease incidence, cost and duration, probability of next activity) are listed with references to the source material so that audit trail can be traced back to the literature sources, process mining outputs or investigation artifacts that were used. The modeler sets a Confidence Indicator (CI) to document their confidence in the evidence base for each element and attribute on a score of 1-5 with 5 being highest.

0 = No confidence/ not applicable/ system defaults
1 = Guess by Modeler
2 = Estimate from observation or domain expert interview

3 = Empirical evidence from process mining or published literature
4 = Confirmed from multiple reliable sources
5 = Confirmed through Clinical Review Board.

It is evident that a modeler can very quickly create a low fidelity model of the care pathway but will have to record CIs of mostly 0s and 1s. Conversely, a Clinical Review Board could review the evidence for every element in detail recording scores of 5 where they agree and leaving other elements as 3s or 4s where there is still uncertainty. The overall average CI therefore gives a rough indicator of overall confidence in the model and crucially, the modeling can stop when the Clinical Review Board believe they have enough evidence to make a process improvement decisions.

4 Case Studies

4.1 Use of the ClearPath Method

The ClearPath project aims to generate a method for combining process mining and simulation that is suitable for widespread use understanding and improving care pathways in the UK National Health Service. The case studies illustrate aspects of the method in use, some of the achievements and some of the unsolved challenges.

4.2 Case Study 1 Alcohol-related emergency admission pathway at Liverpool

In some parts of the UK hospital admissions for alcohol related illnesses are rising by 11% per year leading to chronic diseases such as ARLD (Alcohol Related Liver Disease) which has lower survival rates than most common cancers. In busy emergency departments, alcohol-related disruptive behavior may obscure a patient's serious advanced illness and also hamper treatment attempts [22]. There is growing recognition that clinicians need guidance on appropriate care pathways that can help them identifying and deal with alcohol related illness. For this case study, our project team worked with a data and pathway profiling team at the University of Liverpool who had created a data linkage framework based on EHR event data from emergency admissions from hospitals in the North West of England. Our approach consisted of embedding a member of our ClearPath team within the Liverpool team for up to two days per week over a three-month period. We resolved data governance issues by providing tool and analysis advice to the local team. In return, our analyst received sequence, aggregate and conformance data to populate a NETIMIS simulation model (see extract in Figure 2) and an evidence template (see extract in Figure 3).

Figure 3 illustrates how the evidence template was used to document the link between the percentages derived from the process data for ICU Disposal to the probability settings in the simulation model. Five iterations of the pathway model were developed starting from simple models from an initial workshop and enriched through investigation and reviews. In the final model these included age-banded probabilities extracted from the routine data and cost data sourced from standard activity tariffs. Several qualitative researchers had investigated the pathway through patient and clinician

interviews and their deep insights helped fill in the gaps and add paths that were not evident from the data. The model was calibrated so that the outcomes reflected published clinical outcomes for the region. The resulting model is being presented as a regional exemplar of data-driven care pathway improvement.

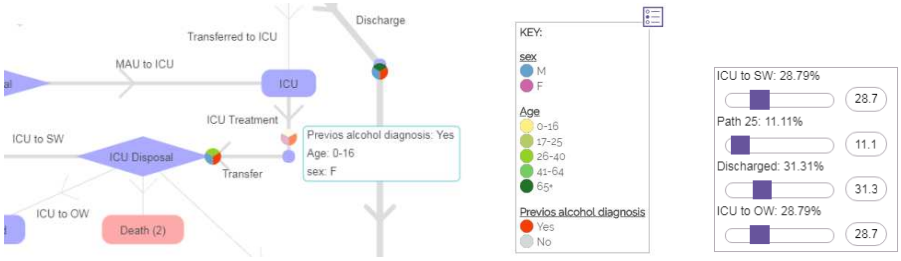

Figure 2: Close-up view of pathways in and out of the Intensive Care Unit (ICU) and (right) the probability settings for paths exiting ICU Disposal

Pathway Setting ICU Route				
Activity	Time (hours)	Cost (£)	Next Activity (Probability)	CI
ICU Treatment	36.2 [31]	2828.1 [32]	Transfer	3
Transfer	1 [33]		ICU Disposal	1
ICU Disposal			Other Ward 0 28.79%	2
			Specialist Ward – 28.79%	2
			Discharge – 31.31% [34]	3
			Death 11.11% [34]	3

Evidence references

...

31 Mean time from data mining exercise

32 Reference cost from list provided by domain expert

33 Unknown. This is a guess to ensure the model runs OK

34 From process mining of hospital data

Figure 3: Extract of the evidence template (left) and references (right) for the corresponding ICU elements

4.3 Case Study 2 - Giant cell arteritis (GCA) care pathway at Leeds

Giant cell arteritis (GCA) is a rare chronic inflammatory condition of blood vessels (vasculitis) that affects large and medium sized arteries. Symptoms include headaches, tenderness of the scalp, jaw aches and chewing problems and visual impairment. The symptoms displayed can often be mistaken for normal age-related symptoms or other diseases however if GCA is not diagnosed and treated quickly it can lead to visual loss, blindness, or in worst cases a stroke [24]. For this reason, patients are treated with steroids as soon as the diagnosis is considered but this creates other challenges as the steroids impact on the sensitivity of diagnostic tests. Our project team worked with the clinical specialists at Leeds Teaching Hospitals Trust (LTHT) with access to the national MRC-TARGET (Treatment According to Response in Giant cEll arteritis) consortium (https://lida.leeds.ac.uk/target). Figure 1 illustrates the *de jure* pathway for GCA drawn from the National Institute for Health and Care Excellence (NICE) repository (https://pathways.nice.org.uk).

Our objective here was to map the *de facto* pathways in a large and very busy teaching hospital. Our initial approach was to request anonymized data extracts from the hospital EHR for patients with suspected GCA which included time stamped information of the patient's journey starting from their original route of entry into care, through to discharge, or firm diagnosis of GCA. Data quality issues proved insur-

mountable. It was not possible to accurately identify patients of interest or enough relevant events within the hospital EPR to complete the envisaged process mining exercise. However, we were able to complete a detailed process model through both traditional business analysis investigation and produced five iterations with a Clinical Review Board consisting on the local TARGET consortium leads. We gathered sufficient data from clinical expert interviews and volume and time figures from the EPR and cost figures from hospital tariffs to build a robust working model (mean CI=4.1). Through documentary sources and interviews with other hospitals were able to produce models for other hospitals and develop a generic model that was a fit with the *de jure* guidelines and could be used to model *de facto* GCA pathways from small district hospitals to the complexity of LTHT. The study concluded with a costed model of the planned future model which identified points at which clinical pathways could be improved by recommending alternate diagnostic approaches. The simulation (Figure 4) indicate both significant improvements in patient outcomes and simultaneously reduced costs. The models have been presented to the national group. A second phase using process mining of the Leeds EHR data is planned.

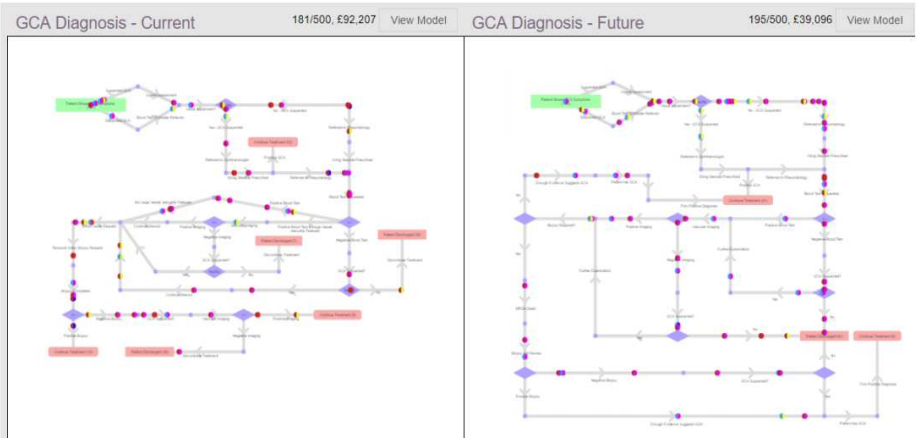

Figure 4: NETIMIS screenshot showing a side-by-side run of the current Leeds' GCA Diagnosis pathway against the proposed future pathway

4.4 Case Study 3 – Functional Neurological Symptoms (FNS) care at Leeds

Functional Neurological Symptoms (FNS) are a group of neurological symptoms which include weakness, abnormal movements and blackouts, they cause distress and dysfunction [25]. The symptoms are shared with neurological diseases such as epilepsy, multiple sclerosis or stroke but, in FNS patients, are caused by a brain dysfunction. As with GCA, diagnosis is challenging but studies have shown that 31% of patients attending Neurology outpatient clinics had FNS [26]. Many healthcare providers lack specific pathways or services for FNS patients, there are no NICE guidelines and there is little data on acute FNS to inform service improvement. Therefore, there was a need to see the current pathways for FNS patients, and to know how they

pass through the different healthcare services over time. Our first step was to generate visual models of the process from first presentation to diagnosis and referral to either psychology or psychiatry therapy.

Our aim with this project was to see whether process mining of the routine data was possible and we worked with a team of neurologists to understand the *de facto* care pathways at LTHT. Clinical inspection of the data quality of the EHR revealed data that was considered too unreliable to use for process mining. Issues included unrecorded events and observations, recording on letters and paper records rather than the EHR, mis-diagnosis and inappropriate referrals. Our alternative approach was to conduct a full audit of all the EHR data, clinical letters and discharge notes for each patient using all available sources (including phone interviews with treating clinicians to complete missing data). These resulting activities include diagnoses, emergency attendances, outpatient clinics, inpatient admissions and psychological / psychiatric referrals. The audit data was collated in the form of an event log which was used with a process mining tool (ProM). The initial results appeared disappointing - a spaghetti diagram with every single patient (n=205) having a unique and complex variant. It was however a shocking result for the clinical domain experts, the findings show a high healthcare burden, and slow or incomplete movement to appropriate care, with an urgent need for service improvements. The mean time from Presentation to Diagnosis was 22.1 months and a further 7.2 months to an appropriate Referral. These results were presented as a video at an Association of British Neurologist conference and are being used to make the case for clearer pathways for FNS.

5. Discussion and Conclusions

Our experience working with both process mining and process simulation has been that both approaches are complex, challenging and require considerable skill, domain knowledge and perseverance. Healthcare is a complex world and EHRs are not yet capturing sufficient detailed workflow for deep clinical insights. Given this state of affairs, combining as many valid techniques as possible would seem to be the most pragmatic approach for quickly generating insights. However healthcare also demands strong evidence and rigorous methods and the ClearPath method has helped us structure data-driven care pathway investigations that have yielded good results and maintained an audit trail of evidence that ensures the models are defensible.

All three of the case studies here are examples of special case pathways within more general processes such as emergency admissions. Many patients genuinely require variants that differ from the norm. Case Study 1 used a simulation model to combine process mining outputs and other sources to build a useful model backed by evidence that can be traced to its source. Both Case Study 2 and 3 illustrate how difficult it can be to obtain robust EHR data. In Case Study 2 we made do with other sources and in Case Study 3 a manually constructed event log revealed alarming variability in care.

Our methods are evolving through use but we expect to formalize the approach on PM^2, Clinical Review Boards, evidence templates and audit trails. Currently we use NETIMIS for care pathway simulation and there are many alternatives. With NETIMIS we expect to improve tool integration so that it can generate and consume event logs and learn branching probabilities and probability density functions for activity duration from the log. Parallel to this we plan to develop rules to automate the visualization layout, to simplify the task of analyzing complex processes and communicating the results to diverse stakeholders. Our approach has been well received in the UK and may be of benefit to the wider academic and healthcare community seeking to use process-oriented data science to improve healthcare.

Acknowledgment

This work was supported by the cYorkshire Connected Health Cities (CHC) project. The case studies were developed by Luke Naylor, Sahar Salimi Avval Bejestani, Clea Southall and Samantha Haley at the University of Leeds. Case study 1 was supported by Anna Jenkins and the University of Liverpool CHC. Case study 2 was supported by Prof Ann Morgan and the TARGET Consortium for GCA. Case study 3 was supported by Dr Stefan Williams. The third author would also like to thank the Indonesia Endowment Fund for Education (LPDP) for the support given during this research.

References

[1] K. Vanhaecht, M. Panella, and R. Van Zelm, "An overview on the history and concept of care pathways as complex interventions," *Int. J. Care Pathways*, vol. 14, no. 3, pp. 117–123, 2010.

[2] European Pathway Association, "Care Pathways." [Online]. Available: http://e-p-a.org/care-pathways/. [Accessed: 30-May-2018].

[3] R. S. Mans, W. M. P. van der Aalst, and R. J. B. Vanwersch, *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. Heidelberg: Springer, 2015.

[4] N. G. Weiskopf, S. Bakken, G. Hripcsak, and C. Weng, "A Data Quality Assessment Guideline for Electronic Health Record Data Reuse," *eGems (Generating Evid. Methods to Improv. Patient Outcomes)*, vol. 5, no. 1, p. 14, 2017.

[5] O. Johnson, P. S. Hall, and C. Hulme, "NETIMIS : Dynamic Simulation of Health Economics Outcomes using Big Data," *Pharmacoeconomics*, vol. 34, no. 2, pp. 107–114, 2015.

[6] R. Mans, H. Reijers, D. Wismeijer, and M. Van Genuchten, "A process-oriented methodology for evaluating the impact of IT: A proposal and an application in healthcare," *Inf. Syst.*, vol. 38, no. 8, pp. 1097–1115, 2013.

[7] E. Rojas and J. Munoz-Gama, "Process mining in healthcare: A literature review," *J. Biomed. Inform.*, vol. 61, pp. 224–236, 2016.

[8] R. S. Mans, W. M. P. van der Aalst, R. J. B. Vanwersch, and A. J. Moleman, "Process mining in healthcare: Data challenges when answering frequently posed questions,"

Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7738 LNAI, pp. 140–153, 2013.

[9] W. M. P. van der Aalst, "Process Mining: Data Science in Action," in *Process Mining*, 2nd ed., Springer-Verlag Berlin Heidelberg, 2016, pp. 3–23.

[10] W. M. P. van der Aalst, et al., "Process Mining Manifesto," *Bus. Process Manag. Work.*, vol. 99, pp. 169–194, 2011.

[11] M. L. Van Eck, X. Lu, S. Leemans, and W. M. P. Van Der Aalst, "PM2: A process mining project methodology," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9097, pp. 297–313, 2015.

[12] Bozkaya, M., Gabriels, J. and van der Werf, J.M., 2009, February. Process diagnostics: a method based on process mining. In Information, Process, and Knowledge Management, 2009. eKNOW'09. International Conference on (pp. 22-27). IEEE.

[13] A. Rebuge and D. Ferreira, "Business process analysis in healthcare environments: A methodology based on process mining," *Inf. Syst.*, vol. 37, pp. 99–116, 2012.

[14] E. Rojas, M. Sepúlveda, J. Munoz-Gama, D. Capurro, V. Traver, and C. Fernandez-Llataset al., "Question-Driven Methodology for Analyzing Emergency Room Processes Using Process Mining," *Appl. Sci.*, vol. 7, no. 3, p. 302, 2017.

[15] W. M. P. van der Aalst, "Business Process Simulation Revisited," *Enterp. Organ. Model. Simul.*, vol. 63, pp. 1–14, 2010.

[16] A. Rozinat, R. S. Mans, M. Song, and W. M. P. van der Aalst, "Discovering simulation models," *Inf. Syst.*, 2009.

[17] S. Brailsford, M. W. Carter, and S. H. Jacobson, "Five Decades of healthcare simulation," in *Winter Simulation Conference*, 2017, vol. 9, no. 6 2, pp. 365–384.

[18] D. A. Marshall, et al., "Selecting a dynamic simulation modeling method for health care delivery research - Part 2: Report of the ISPOR dynamic simulation modeling emerging good practices task force," *Value Heal.*, vol. 18, no. 2, pp. 147–160, 2015.

[19] N. G. Weiskopf and C. Weng, "Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research," *JAMIA*, vol. 20(1), pp. 144–151, 2013.

[20] M. G. Kahn, et al., "A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data.," *EGEMS (Washington, DC)*, vol. 4, no. 1, p. 1244, 2016.

[21] F. Fox, V. Aggarwal, H. Whelton, and O. Johnson, "A Data Quality Framework for Process Mining of Electronic Health Record Data," in *Proceedings of the Sixth IEEE ICHI*, 2018.

[22] Lekharaju, P., Thompson, E., Shawihdi, M., Pearson, M., Hood, S. and Bodger, K.et al., 2014. PTH-062 Emergency Admissions For Alcohol Related Conditions: Making Sense Of Routine Data. Gut, 63(Suppl 1), pp.A236-A236.

[23] Laskou, F., Fiona, C., Aung, T., Benerjee, S. and Dasgupta, B., 2018. 074 Fast track giant cell arteritis clinic and pathway for early management of suspected giant cell arteritis: an audit. Rheumatology, 57(suppl_3), pp.key075-298.

[24] S. Mobini, "Psychology of medically unexplained symptoms: A practical review," *Cogent Psychol.*, vol. 2, no. 1033876, 2015.

[25] J. Stone, A. Carson, et al, "Symptoms 'unexplained by organic disease' in 1144 new neurology out-patients," *Brain*, vol. 132, pp. 2878–2888, 2009.