Exploring Rehabilitation Care Pathways from Electronic Health Records of Patients with COVID-19

Konstantin Georgiev[1], Jacques Fleuriot[2], Susan D Shenkin[3], Joanne McPeake[4], Atul Anand[1]

1. *BHF Centre for Cardiovascular Science, University of Edinburgh, UK;*
2. *Advanced Care Research Centre, Usher Institute, University of Edinburgh, UK*
3. *Artificial Intelligence and its Applications Institute, School of Informatics, University of Edinburgh, UK*
4. *The Healthcare Improvement Studies Institute, University of Cambridge, UK*

SCAN ME

R code available on GitHub

INTRODUCTION

COVID-19 rehabilitation continues to face many challenges:

- Early rehabilitation is a necessity in patients recovering from **acute** COVID-19 symptoms[1].
- The pandemic has been severely **disruptive** towards services in the rehabilitation ward.
- We are yet to understand the evidence base for optimum treatments.
- A more **patient-centred** approach is needed to understand how to plan treatment for complex needs.

The increasing use of Electronic Health Records (**EHR**) and big data in modern healthcare systems[2] in the COVID-19 domain can allow us to effectively:

- collect and analyse detailed hospital episodes in COVID-19 patients
- develop Process Mining (**PM**) models to explore interactions with Allied Healthcare Professionals (**AHPs**) and other providers
- evaluate changes in practice between different stages of the pandemic (**Wave 1 and Wave 2**).

METHODS

Our case study consists of three main stages:

- **Data collection and preprocessing**: developing the COVID-19 cohort based on positive episodes captured across three acute hospitals in Edinburgh, Scotland (part of NHS Lothian).
- **Process mapping**: we built an event log from timestamped and coded AHP contacts, with detailed time-frequency sequence maps.
- **Summary of rehabilitation needs**: we reported the common trace combinations and differences in time between nodes.

We defined the cutoff for the two COVID-19 cohorts as follows:

- **Wave 1** (between 1st March 2020 and 31st May 2020)
- **Wave 2** (between 1st September 2020 and 31st March 2021)

We used the **bupaR**[3] framework (version 0.5.2, R 4.2.0) to produce the event logs and process maps individually for each COVID-19 cohort. Sequences of **AHP** visits were described by order of occurrence, and each edge in the graph was weighed by a significance metric relative to **time** or **likelihood** of that instance. This representation is similar to that of the **Fuzzy Miner** process discovery algorithm with a directly-follows graph (DFG) notation[4]. We used **frequency-based** filters to eliminate the **infrequent** flows (MINF and MAF in figures) and simplify the maps. R code to reproduce this workflow with dummy data is provided on **GitHub** (scan QR code).

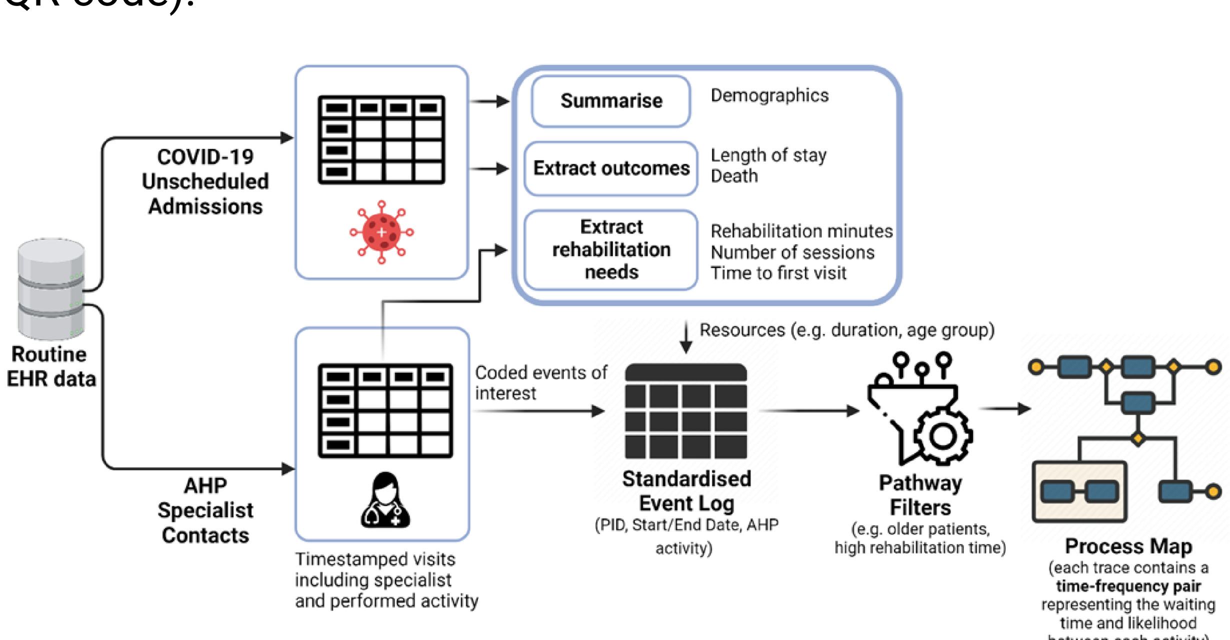

Fig. 1. Conceptual flowchart describing the steps required to transform EHR data into a COVID-19 rehabilitation pathway map.

CONCLUSION

- Our **PM** approach offers an alternative way of understanding COVID-19 hospital interactions with AHPs and healthcare providers.
- Identifying common rehabilitation processes can aid in providing **clearer** descriptions of rehabilitation requirements.
- We recognise the limitations stemming from capturing a fragmented patient journey only linked to rehabilitation contacts.
- In future work, we plan to collect and integrate events from a wide range of **multidisciplinary** care settings and validate this approach using qualitative measures from healthcare providers.

RESULTS

Coded inpatient teams included:

- **NURSE** (care delivered by a specialist nurse)
- **PT** (physiotherapist)
- **OT** (occupational therapist)
- **DT** (dietician providing nutritional support)
- **SPL** (speech and language therapist)
- **CP** (clinical pharmacist)
- **PC** (specialist in palliative care)
- **INFSV** (infection prevention service)
- **OTHER** (unidentified care provider)

Characteristic	All (n=1,000)	Wave 1 (n=259)	Wave 2 (n=741)	p
Age, years (median, IQR)	78 (67, 86)	79 (66, 86)	78 (67, 86)	0.85
Sex	..	..	..	0.01*
Male	495 (50%)	146 (56%)	349 (47%)	..
Female	505 (50%)	113 (44%)	392 (53%)	..
SIMD in quintiles	..	..	..	0.28
1 (most deprived)	135 (14%)	26 (10%)	109 (15%)	..
2	277 (28%)	68 (26%)	209 (28%)	..
3	146 (15%)	42 (16%)	104 (14%)	..
4	172 (17%)	50 (19%)	122 (16%)	..
5 (least deprived)	270 (27%)	73 (28%)	197 (27%)	..
Outcomes	..	..	..	..
Length of stay, days (median, IQR)	30.9 (18.3, 56.9)	30 (17, 49)	31 (19, 59)	0.19
Death at 1 year	406 (41%)	93 (36%)	313 (42%)	0.07
Death prior to discharge	272 (27%)	65 (25%)	207 (28%)	0.38
Rehabilitation needs	..	..	..	..
Minutes of rehabilitation care (median, IQR)	232 (125, 436)	280 (117, 492)	215 (125, 420)	0.08
Total number of contacts (median, IQR)	7 (4, 14)	**9 (4, 17)**	**7 (4, 13)**	0.02*
Out-of-hours contacts (mean, SD)	2.47 (3.04)	2.55 (3.12)	2.44 (3.02)	0.95
Nursing care (mean, SD)	2.92 (3.36)	3.03 (3.51)	2.88 (3.3)	0.96
Physiotherapy (mean, SD)	6.58 (6.91)	**7.55 (7.35)**	**6.24 (6.72)**	0.003**
Occupational therapy (mean, SD)	0.32 (0.8)	0.32 (0.81)	0.32 (0.79)	0.83
Dietetics (mean, SD)	0.92 (1.87)	**1.1 (1.94)**	**0.85 (1.84)**	0.02*
Speech and Language (mean, SD)	0.35 (1.22)	**0.48 (1.42)**	**0.31 (1.14)**	0.006*
Minutes of rehabilitation per day (median, IQR)	8 (5, 11)	**9 (5, 13)**	**8 (5, 11)**	0.002**
Time to first contact, days (median, IQR)	3 (2, 6)	4 (2, 6)	3 (2, 6)	0.12
Time to second contact, days (median, IQR)	6 (4, 9)	6 (4, 8)	6 (4, 9)	0.34

Table 1. Baseline characteristics of the two COVID-19 cohorts. Values are in proportion of patients (%) unless stated otherwise. SIMD – Scottish Index of Multiple Deprivation; Significance: * $p < 0.05$; ** $p < 0.005$;

Fig. 2. Filtered specialist process map containing the most frequent cases over both waves. The weight of the arrows is adjusted by the probability of the interaction.

Fig. 3. Precedence matrices representing the relative trace frequency between Wave 1 and Wave 2.

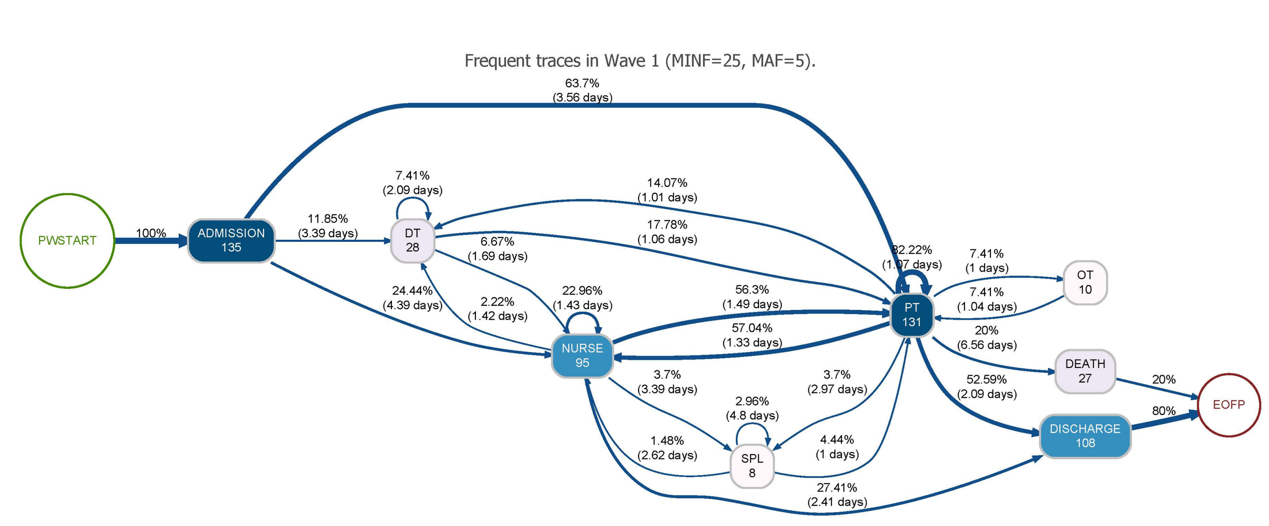

Fig.2 (a) Common traces in the Wave 1 event log (MINF=25, MAF=5).

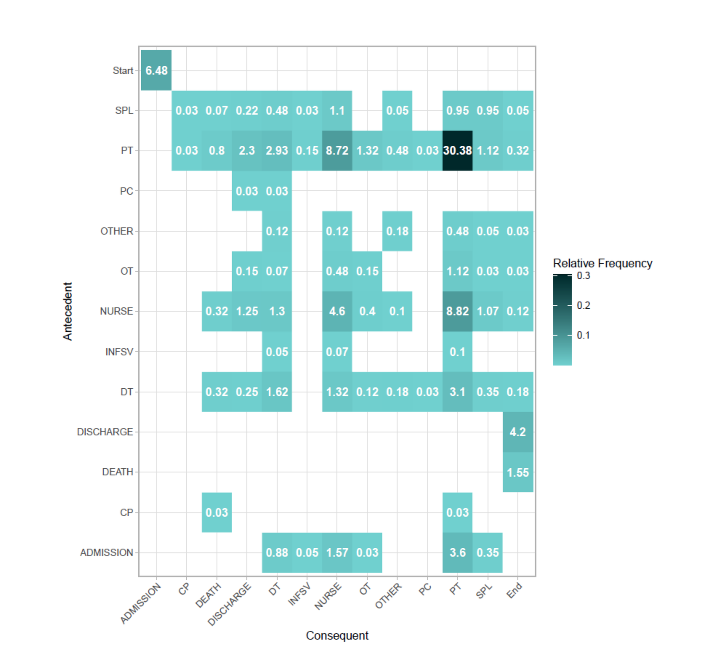

Fig. 3 (a) Precedence matrix (Wave 1)

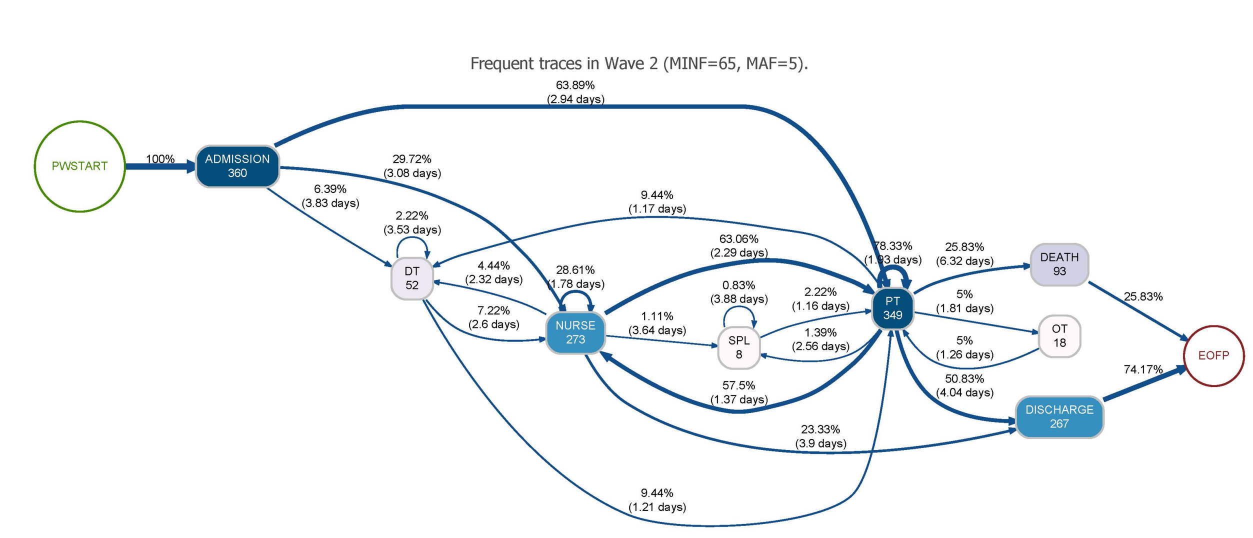

Fig. 2 (b) Common traces in the Wave 2 event log (MINF=65, MAF=5).

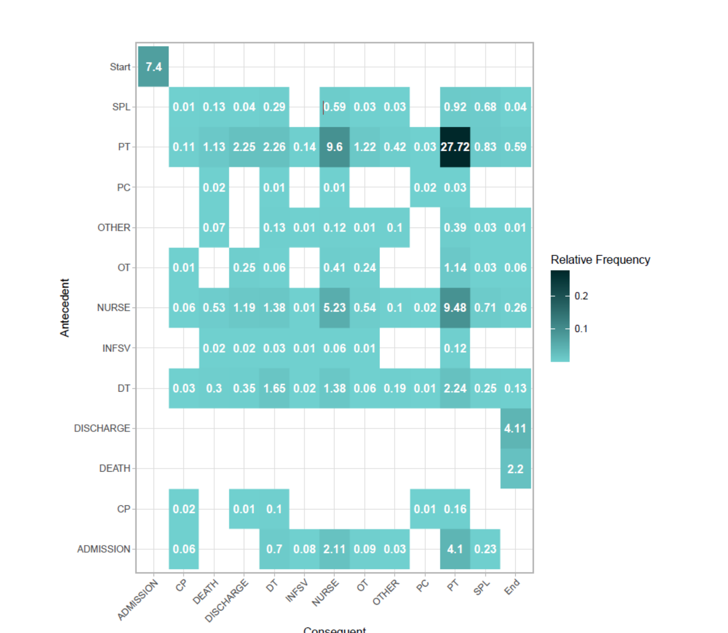

Fig. 3 (b) Precedence matrix (Wave 2)

Fig. 4. Cumulative time plots representing throughput time (total execution time of the pathway) and idle time (total hospital time with no related AHP activity) by age and COVID-19 Wave.

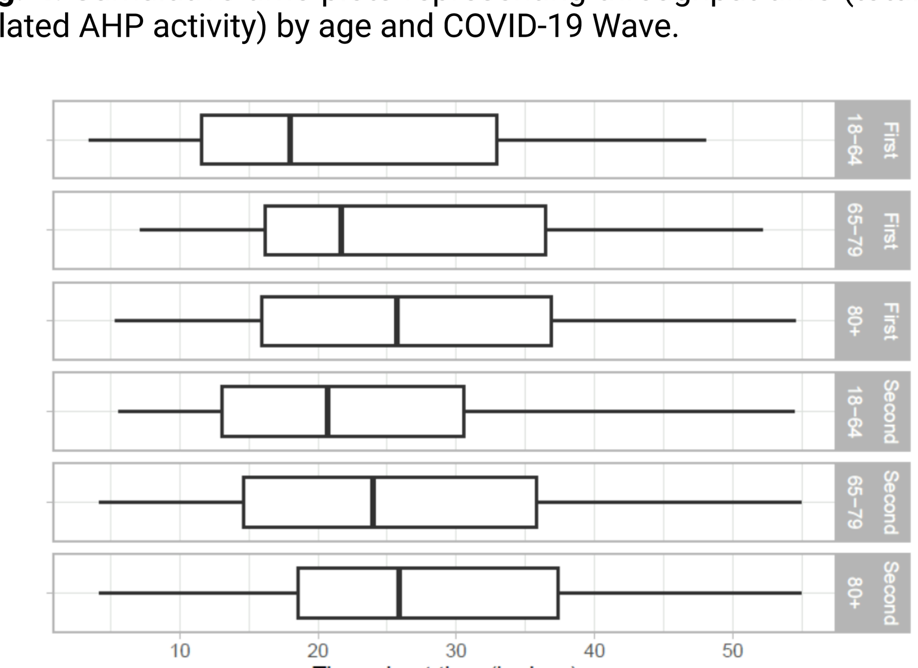

Fig. 4 (a) Throughput time in days over the complete event log

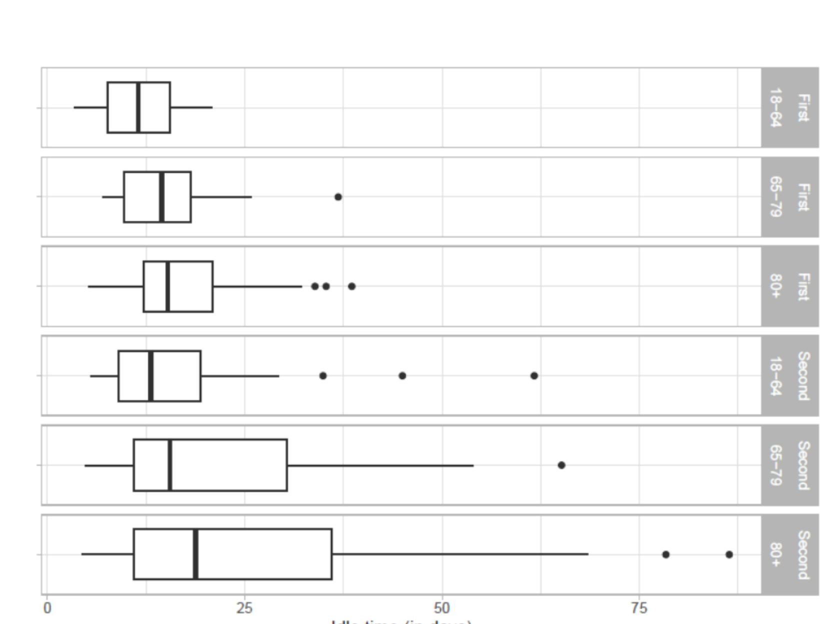

Fig. 4 (b) Idle time in days over the complete event log

References:

1. Udina, C. *et al.* Rehabilitation in adult post-COVID-19 patients in post-acute care with Therapeutic Exercise. *J Frailty Aging* **10**, 297–300 (2021).
2. Mehta, N., Shukla, S.: Pandemic Analytics: How Countries are Leveraging Big Data Analytics and Artificial Intelligence to Fight COVID-19? SN Comput Sci. 3, 54 (2022). https://doi.org/10.1007/s42979-021-00923-y.
3. Janssenswillen, G., Depaire, B.: bupaR: Business Process Analysis in R. https://bupaverse.github.io/docs/.
4. Günther, C.W., van der Aalst, W.M.P.: Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. Springer, Berlin, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75183-0_24.

KG is supported by a PhD Fellowship award from the Sir Jules Thorn Charitable Trust (21/01PhD) as part of the University of Edinburgh's Precision Medicine PhD programme.