TITLE: Transforming patient care: automated mapping of hospital patient episodes using process mining with sequential probability analysis

Introduction

Hospitals seek excellence in the quality of care with efficiency by understanding their activities and the variations within. Quantitatively but effectively representing and comparing complex clinical episodes across a large and diverse patient population is challenging. Usual methods for analysing clinical pathways rely on manual data curation, domain-specific knowledge, and may not adequately capture clinical practice variability. Therefore, there is the potential to develop an automated system to address the above-mentioned challenges for a busy tertiary hospital. This study utilises a novel automated system for mining and mapping versatile patient care services during their episode of care, ultimately assisting in identifying the most efficient, high-quality, and affordable pathways.

Methods

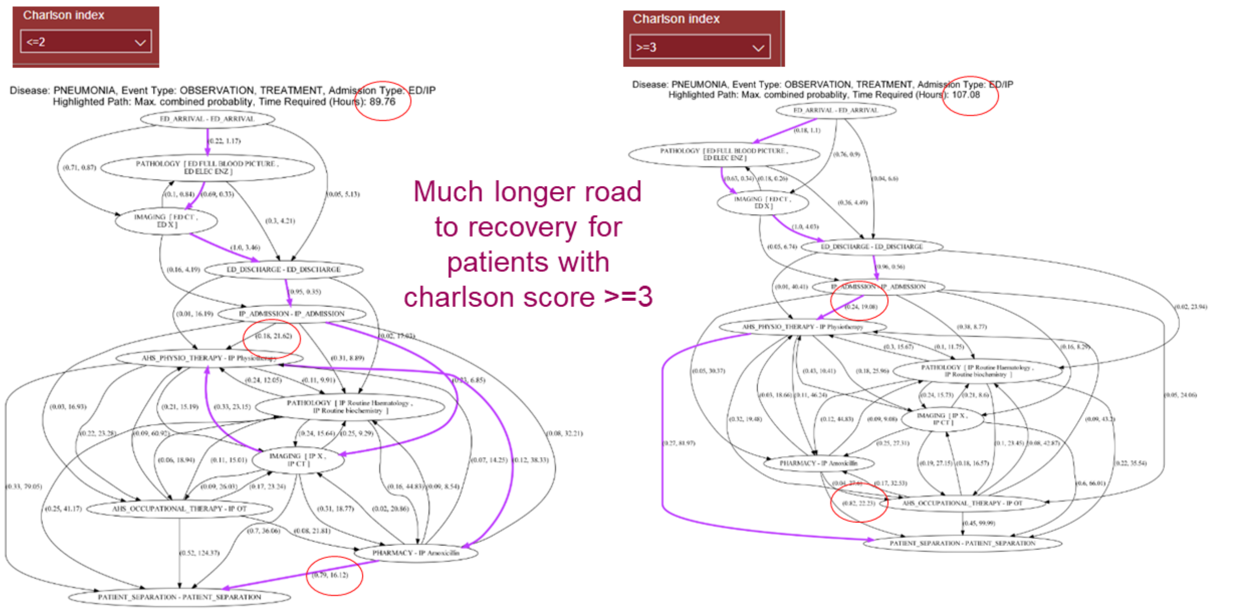
This study utilised a longitudinal dataset of 83251 hospital episodes from 53304 patients discharged from Sir Charles Gairdner Hospital, WA, Australia, over four fiscal years, 2019-23. The data included timestamps for pathology and imaging, theatre interventions, ward, responsible team movements, and selected pharmacy items. R scripts were used for data preprocessing (cleaning, sequencing, and grouping), while Python scripts enabled service mining and pathway mapping. An in-house built process mining algorithm categorised episodes by diagnosis, comorbidity, and complexity, iteratively clustering services based on service pair probability and time difference within each case-mix group. These clustered services are represented as multidirectional graphs, pruned to identify the most probable patient treatment pathways based on available disease groups and other attributes such as age, gender, admission time, length of stay, comorbidities, and hospital-acquired complications.

Results

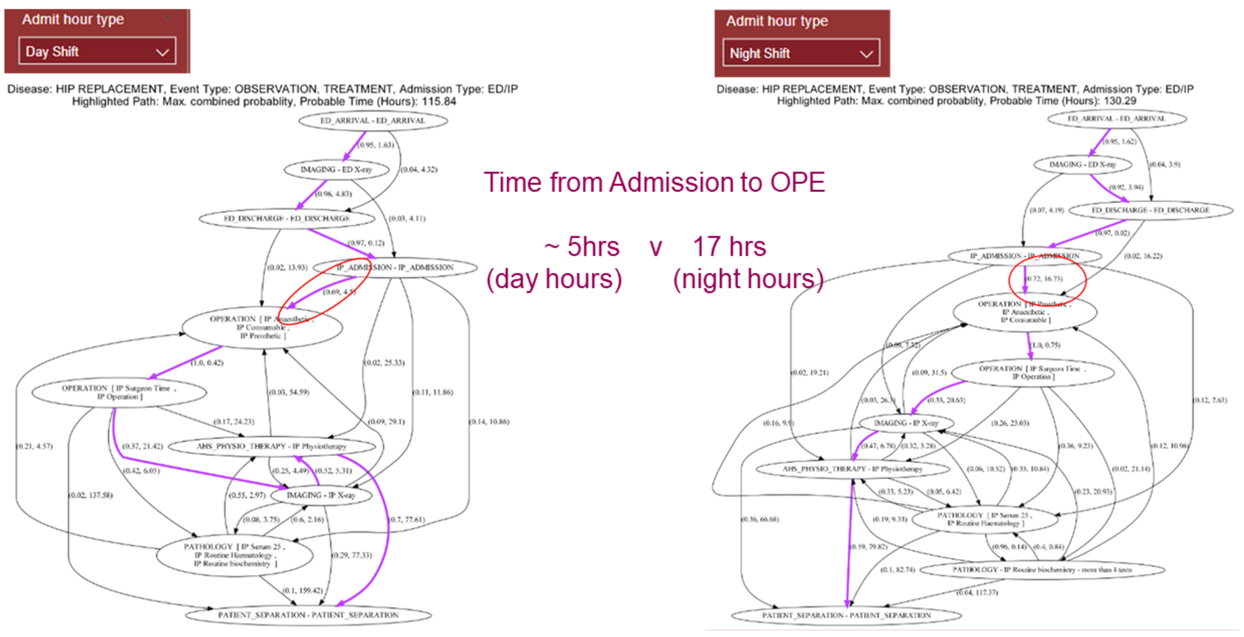
The system initially generates visuals of all available service pairs with numerous connections/paths, providing a granular view of service interactions. Then, it further refines the visual comprising the clusters of these service pairs into larger service groups (e.g., pathology, imaging, ward, emergency), offering a higher-level perspective of patient pathways (Figure 1). This visual map highlights key metrics, including maximum sequential probability, maximum combined probability, and shortest time between service pairs, allowing for identifying common and efficient care sequences. Using the same data modelling architecture, a separate visual also maps patient locations during their episode of care to identify efficient bed utilisation within a high-demand setting.

Discussion/Conclusion

This study demonstrates the feasibility of an automated system for the surrogation and visual mapping of complex patient episode data. The resulting visuals represent service relationships and impact, enabling clinicians and executives to observe clinical pathways, assess the impact of service sequence changes for various disease groups, and subsequently identify the variations in the clinical pathway. By providing data-driven insights into clinical workflows and treatment demand, this approach can support care quality improvement, resource allocation, clinical decision-making, and cost analysis by patient-reported and post-discharge outcomes. Future research would advance more congruent data presentation to clinicians and managers and integrate this system with electronic health records for other healthcare settings.



(a)



(b)

Figure 1: Comparable maps of clinical pathways based on (a) Charlson score and (b) admission hours type. Service pairs have directional pair probability and required time (in hours). The maximum combined pair-probability pathway is highlighted in purple.

Keywords: Automated system, Clinical pathways, Healthcare activity mapping, Hospital costing data, Hospital service optimisation