How to solve surgery waiting times: R, Drake and Operational Analytics for a data-driven future

Dr William Bryant

NHS-R Community Conference - 11th November 2020

Background

Digital Research Environment (DRE) @ GOSH

- Research support
- Analytics platform powered by R
- Data pipelines (and engineers) for data provision from EPR
- Analytics team to leverage big datasets

Motivation

- Hundreds of elective surgeries cancelled due to COVID-19
- No capacity for additional theatres
- Operational efficiency required for catch-up
- Rapid, agile, projects needed to show value of data-driven methodology

Project

Objectives

- Swift PoC with reusable code/workflow
- Predict surgery length
- Show (and quantify) potential efficiency saving using scheduling

Methods

- Patient clinical data from GOSH EPR
- Drake + Gitlab
- Surgery Scheduling: mixed integer linear programming
  - `ROI`, `ROI.plugin.glpk`, `ompr`, `ompr.roi`
- Surgery Duration
  - `survival`, `quantreg`, `flexmix`, `GAMLSS`
  - `XGBOOST`

Development flow - Drake

Advantages

- Functionalise all code (reuse)
- One function per file (version control)
- Standardised plan file (collaboration)
- (Very useful error messages)

Drake targets:

```
## How many lab tests were performed for each patient?
df_lre_present_counts = wrangle_lre_present_counts(df_lre_joined),
plt_lre_missing_values_density = plot_lre_missing_values_histogram(
  df_lre_present_counts),
```

Scheduling Optimisation - monthly

- Monthly cardiac theatre data (retrospective)
- Known durations => 'perfect data' => upper limit on efficiency
- Max efficiency gain = 11.3%

surgery_month	surgery_count	available_days	minimum_days_opt	result
2019-05-01	23	17	14	Feasible
2019-06-01	30	20	18	Feasible
2019-07-01	36	23	19	Feasible
2019-08-01	32	19	18	Feasible
2019-09-01	31	20	17	Feasible
2019-10-01	34	22	20	Feasible
2019-11-01	33	22	20	Feasible
2019-12-01	17	13	14	Minimum feasible
2020-01-01	26	21	17	Feasible

Healthcare data challenges - incomplete features

- What fraction of patients have each test prior to surgery?

Figure 1: Fraction of patients taking a test

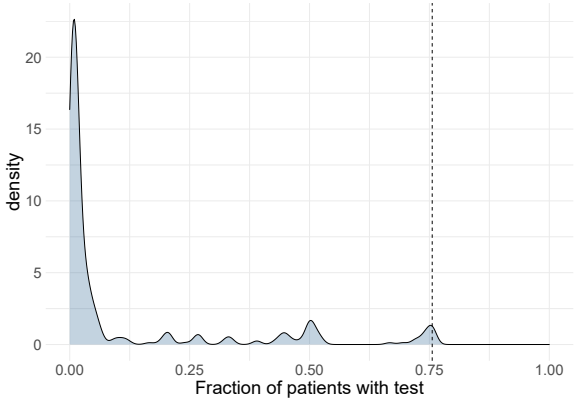
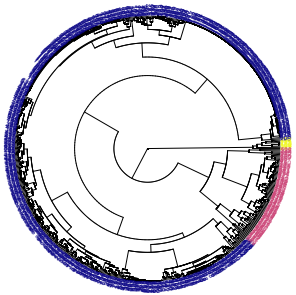


Figure 2: Phylogeny of patient tests profile



Discussion

Surgery duration

- Initially exploring point-of-surgery prediction
- Cardiac data amounting to 693 cases
- 12 features
  - demographics,
  - theatre list,
  - primary procedure,
  - RACHS score (cardiac-specific complexity measure)

Surgery scheduling

- Shows optimisation possible in principle
- Needs full hospital-side engagement

Outlook

Conclusions

- A lot of potential for data-driven efficiency improvements
- More work needed to show PoC for point-of-booking predictions
- More engagement required to push work forwards

Future work

- More data!
- Estimate per-surgery uncertainty
  - Quantile loss in XGBOOST
- Integrate with scheduling

Acknowledgements

Modelling work

- Dr Lydia Briggs

Data engineering

- John Booth
- Mohsin Shah

Clinical expertise

- Dr Richard Issitt

Thanks

- Dr Anastassia Spiridou (DRE lead)
- Prof. Neil Sebire (CRIO)