June 17, 2020

How to apply R in a hospital environment on standard available hospital-wide data

Mieke Deschepper

UZ GENT

UNIVERSITEIT GENT

Data science / statistical analysis in hospitals mostly performed on pathology specific deseases / research context

- Example of usage:
  - Research
  - University hospitals
  - Pathology specific
- No / less usage in
  - Non-research hospitals
  - Hospital-wide topics

Current approach whithin (Belgian) hospitals non-research

- Management reporting performed in a Business Intelligence (BI) cel using Excel or BI tools (e.g. Cognos, Qlik)

- Support available
- Platform to distribute results
- Quick start with standard reports available
- No IT skills required

- Not flexible enough: data and reporting
- No statistical analysis possible
- Very cost full (license and support)
- Documentation / problem solving not free available (no stack overflow questions).
- Advanced reporting still require IT insights

- Statistical analysis on hospital level are rarely performed
  - No research budget
  - Publications are to costly in non-academic field

Use R as alternative in a hospital-wide setting

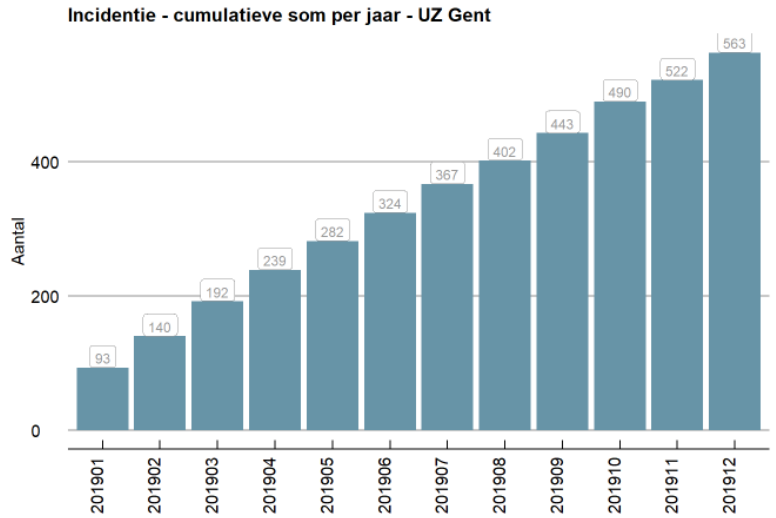
- R as a reporting tool
- R as a statisical tool
- R as a data scientist tool

R as a reporting tool

- Problem:
  - Not al data available within the BI datawarehouse
  - Setup new ETL is too costly
  - Boxplots or other user specific graphs not available
  - Excel is no valid alternative!
- Reporting with R
  - using R and markdown as a tool for management reporting
  - using R for data handling (ETL)
  - Shiny applications as alternative for dashboarding
  - Disadvantage: no distribution platform without IT support

Example R as a reporting tool
Incidence of decubitus

- Problem: the calculation of the incidence of decubitus is too complex to perform in the BI tools (e.g. performance)
- Solution: set up a markdown document to generate the analysis



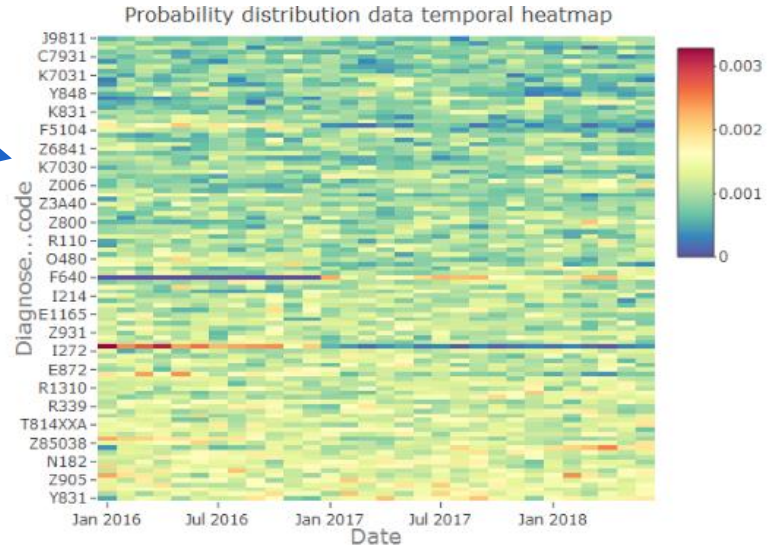
- Not solved: we have no access to distribute this on a hospital-wide platform

Use R as alternative in a hospital-wide setting

- R as a reporting tool
- R as a statisical tool
- R as a data scientist tool

R as a statistical tool

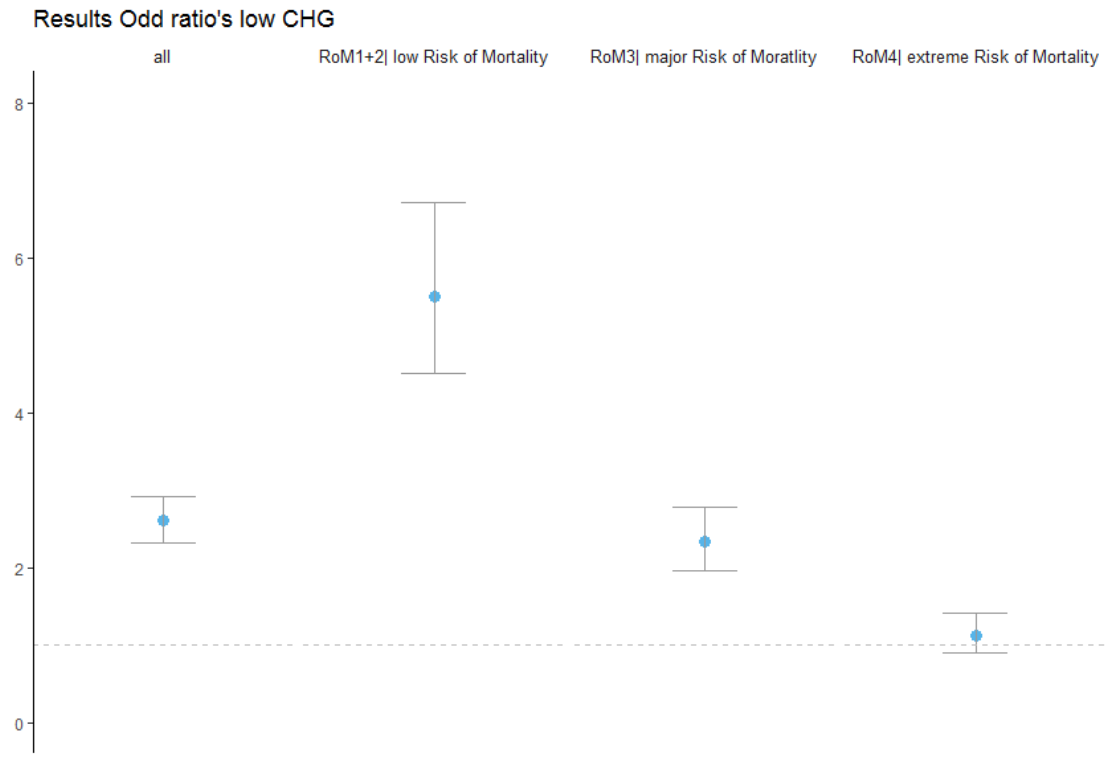
- Problems:
  - BI tools only have limited statistical tools (e.g. trend line)
- Wide range of statistical analysis:
  - Most common: Regression analysis → correlations
    - E.g. Effect of chlorhexidine gluconate oral care on in-hospital mortality
  - Data quality (e.g. heatmaps)



- Reporting / Visualising important
- Example statiscal analysis AND reporting
  - Covid19-pandemic

Example R as statistical tool
Effect of chlorhexidine gluconate oral care on in-hospital mortality

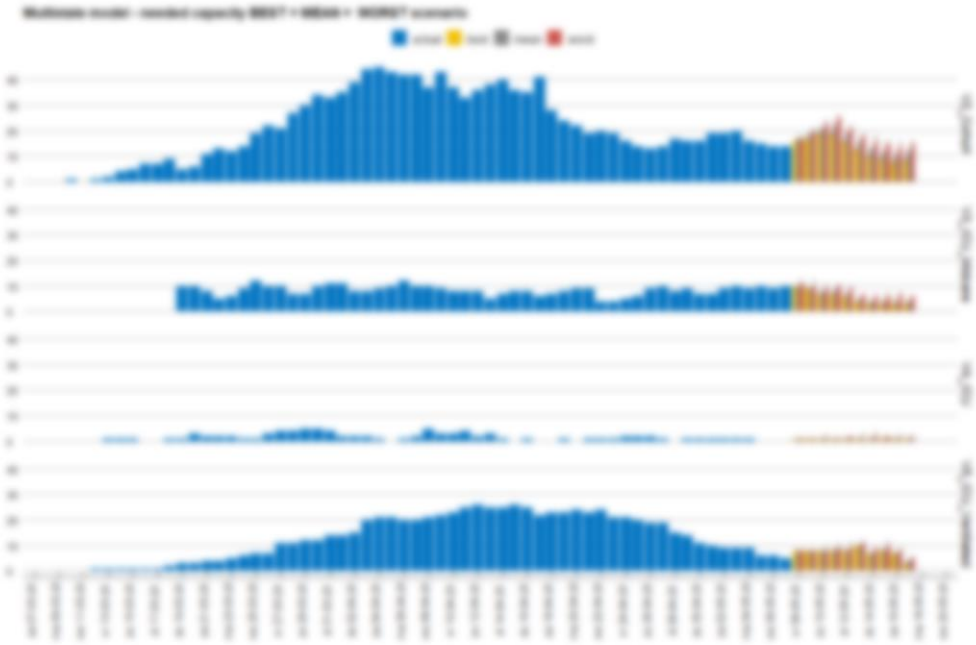
- Why:
  - Suspect correlation between chlorhexidine gluconate oral care and in-hospital mortality
- Method:
  - Logistic Regression (with stepwise)
- Data:
  - 3 years of billling data
- Conclusions:
  - Number needed to harm = 1 out of 47
  - higher risk for low risk mortality groups

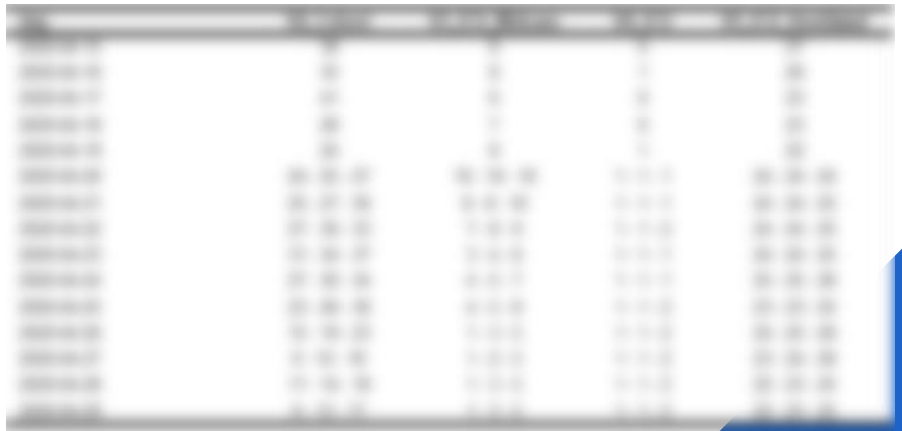


Deschepper et. al 2018 Intensive care Medicine

Example R as a statisical tool (and reporting)
Planning tool bedcapacity during Covid19-pandemic

- Goal: planning tool for task force UZ Gent
- Development time: as narrow as possible < 2 weeks
- Frequency: weekly reporting
- Statistical analysis: Multistate analysis and Poisson modelling
- Output:
  - Small group experts: full analysis
    - .html output (`Rmarkdown`)
  - Large Task force: 3 slides with results (table and graph)
    - .ppt output (`Rmarkdown`) with use of corporate idenity template



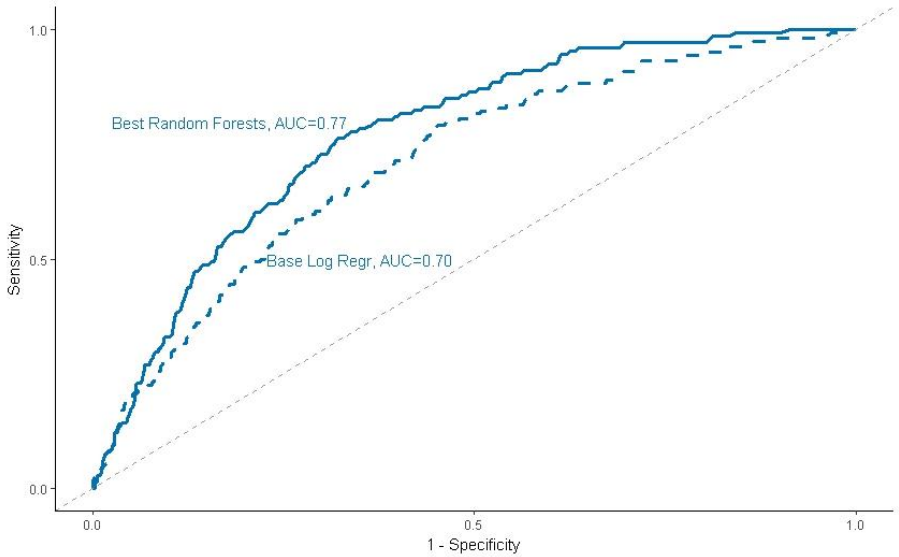


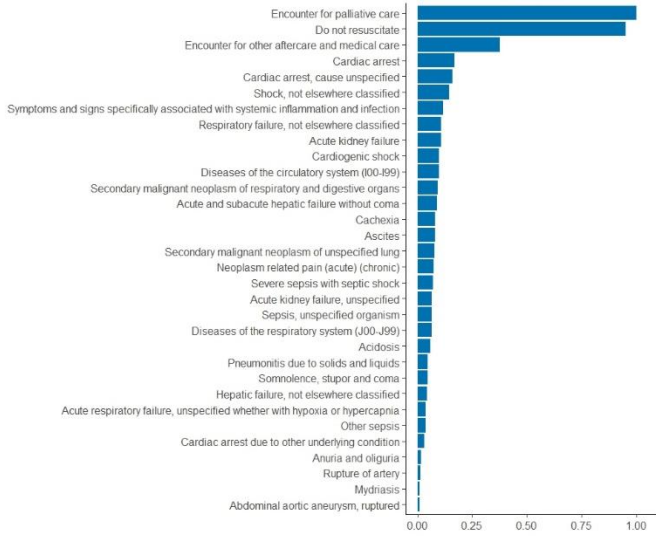
Use R as alternative in a hospital-wide setting

- R as a reporting tool
- R as a statisical tool
- R as a data scientist tool

R as a data science tool

- PhD 2019: “Using standard available hospital-wide data in the interpretation and prediction of outcome indicators”
  - Part 2: prediction of outcome indicators
  - Goal:
    - Predict unplanned readmissions at discharge
    - Predict in-hospital mortality at admission
- Data wrangling in R: high dimensional data
- Using Machine Learning classification algorithms:
  - Random Forests (`h2o`)
  - Gradient Boosting (`xgboost`)
  - …





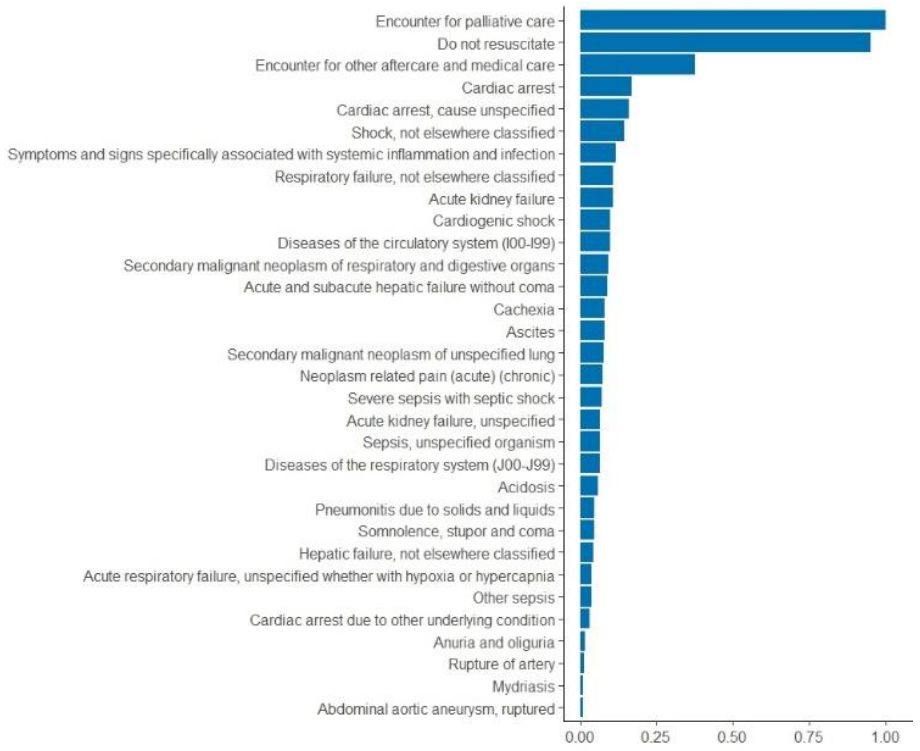
Example R as a data science tool

Using structured pathology data to predict hospital-wide mortality at admission

- Goal:
  - use individual diagnosis codes instead of aggregated measures to predict in-hospital mortlity at admission
  - Effect of Do Not Ressusicate & pallative care codes
- Methods: Random Forests (`h2o`)
- Results:

AUCROC	# predictors	All	Without DNR & palliative care
CCI	1	0.7435	0.7015
RoM	4	0.8797	0.8601
ICD	4743	0.9477	0.8791

- Conclusions:
  - ICD codes (= individual diagnosis codes) outperform CCI and RoM
  - DNR & pallative care code have hight impact on model



Deschepper et. al 2020 Plos One

Use R as alternative in a hospital-wide setting

- R as a reporting tool
  - E.g. using Rmarkdown for adhoc analysis or recurrent analysis
  - E.g. Incidence of decubitus
- R as a statisical tool
  - E.g. performing logistic regressions
    - E.g. Effect of chlorhexidine gluconate oral care on in-hospital mortality
  - E.g. set up planning tool to predict capacity during Covid-19
- R as a data scientist tool
  - E.g. Machine learning tools as Random Forest
  - E.g. predict unplanned readmissions on basis of structured pathology data

R is an alround alternative
Reporting – Statistics - Data science - …

- Some advantages
  - Low licence cost
  - Super flexible
  - Custom graphs (e.g. combine barchart and boxplot)
  - Easy combing multiple sources
- Current problems:
  - No clear RoI for the management
  - No profiles who can use this / set up
  - Hard to set up distributed without IT resources (access)
- Next steps
  - Start a data science team
  - Set up shiny applications as alternative for (non existing) dashboards

MIEKE DESCHEPPER
mieke.deschepper@uzgent.be
Strategic Policy Cel

Universitair Ziekenhuis Gent
C. Heymanslaan 10 | B 9000 Gent
T +32 (0)9 332 21 11
E info@uzgent.be

www.uzgent.be

Volg ons op

UZ GENT

UNIVERSITEIT GENT