



Driving R adoption in an NHS information service: barriers and solutions

Chris Mainey

Intelligence Analyst
University Hospitals Birmingham NHS Foundation Trust

chris.mainey@uhb.nhs.uk

















University Hospitals Birmingham NHS FT



- Second largest Acute-Hospital Trust
- Formed by merger of former UHB and Heart of England trusts
- Four hospitals seeing 440,000 in-patients per year
- Large informatics team(s):
 - Core Information
 - SQL / Dashboard developers
 - Analytics
 - Research
 - Commercial









Benchmarking Platforms

Healthcare Evaluation Data (HED)

www.hed.nhs.uk

- Online hospital benchmarking system
- Statistical models and analysis tools
- Training and support
- Used by ~60 NHS and other organisations







Data analysis tools in the NHS



- Excel and Access were baseline 10-years ago
- 'Complicated' Database work conducted by IT department
- Some Trusts sent regular extracts from system suppliers
- Increasing use of Qlik, PowerBI and Tableau recently

UHB:

- · Recruited SQL developers and slowly moved department to SQL Server
- Excel, SQL Server Reporting Services and TIBCO Spotfire used for outputs







- R was used only for rare 'statistical' calculations
- A few keen medical student or researchers
- · Conserved pattern at other Trusts
- The case in our team...





One keen user...









Allowed space to develop skills:



- My own PhD project, funded by UHB, on statistical modelling in NHS.
- Chose to use R from the outset, rather than SAS or Stata.
- Python is good candidate too, but harder to find training when I started
- Developed statistical, visualisation and data manipulation skills
- Moved some HED models to R
- · Supported other team members to use it
- · This generated a demand for basic training







Followed established patterns:



- Set up a local user group
- Trained senior team
- Encouraged new starters to use it
 - Plan work that requires it
- External relationships:
 - Twitter has been key
 - NHS-R Community: https://nhsrcommunity.com/
 - External R user groups
 - Regional relationships: Other Trusts, PHE, City Council
 - Advice from Mango, RStudio and others







Barriers



- Open Source fears: IT teams sometimes reluctant to install
- Wrestling with default settings:
 - Mapped user folder that aren't writable
 - Can't write to package library, or temp files
- Poor hardware
 - · Mostly desktop users at hospitals
 - Defaults can be low on disk space with minimal RAM
- Lack of understanding of fundamentals:
 - Networking/file locations
 - Suspicions around security
 - R coding or application to problems







Solutions



- Permissions: Explicitly identify default settings
 - Set a specific library, using R_libs_user variable
 - Set temporary directories using .Renviron file
- Hardware:
 - Make the case for more RAM! It doesn't cost much
 - Consider server approaches: VMs or 'Desktop' on a 'Server'
 - RStudio Server / RStudio Connect
- Skill/Awareness:
 - Allow space to develop knowledge
 - Windows users often don't know file paths, SSH etc.
 - Central guidance on R/Python implementation







How do we use it?



- 1. Regression, and related models (mgcv, lme4, glmmTMB, glmnet)
- 2. Machine Learning technique (caret, randomForest, gbm)
- 3. Optimisation problems (solving equations using nloptr)
- 4. Summary stats and Exploration
- 5. RMarkdown Reports and Presentation
- 6. API interaction (NHS ODS service, CQC register organisations)

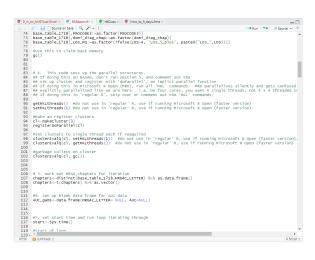








Initial adoption:



- Collections of R scripts saved to directory
 - Required lead analyst to write
 - Required team skills to run and/or debug
- · Constant debugging:
 - Data changing
 - Littered with comment, hard to decipher





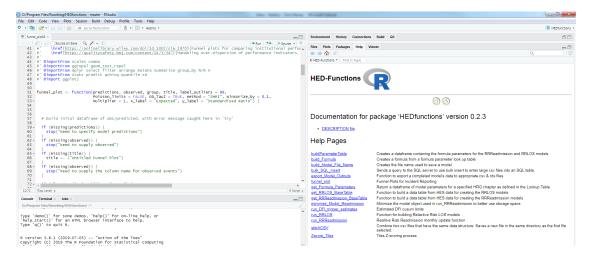


Creating a local package



Wrapped functions in to an R package structure

- · Built binary and save to file structure
- Use roxygen2 to write your support material as you define functions









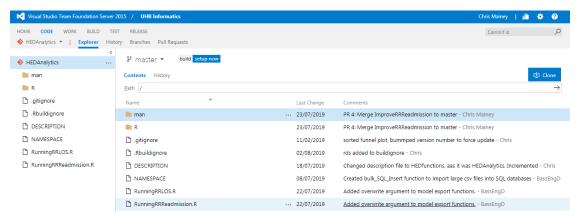


Source control

- Moved to source control
 - · Install git for windows
 - Git learning curve
 - Happy Git with R!
 - https://happygitwithr.com/

- Git repo on TFS
 - Figure out https connections
 - Stef Locke's tfsr

https://github.com/lockedata/tfsR









NHS-R community



- Funded by grant from the Health Foundation
- Aimed to support and promote R use
- Training provision
- Blog posts
- Discussion and dissemination of best practice
- · Development of tools
- National Conference











What's next?

- Continuing to develop analyst skills in R
- Contributing to, or building Open Source tools with NHS-R community
 - FunnelPlotR: https://github.com/chrismainey/FunnelPlotR
 - NHSRdatasets: https://github.com/nhs-r-community/NHSRdatasets
- Wider adoption of Git/Git Hub
- Pushing further into ML techniques for prediction
 - Improving current models: Ensemble techniques etc.
 - Unsupervised techniques for cluster analysis









Thanks for your time!

- chris.mainey@uhb.nhs.uk
- http://www.hed.nhs.uk
- **y** <u>@chrismainey</u>
- chrismainey
- http://www.mainard.co.uk
- b https://orcid.org/0000-0002-3018-6171



