

Big Data Platforms and Analysis in Radiotherapy Trials

Rushil Patel

National Radiotherapy Trials QA Group, Mount Vernon Cancer Centre

Learning Objectives

- Introduction to Randomised Clinical Trials (RCTs)
- Why RCTs are important
- What data we collect
- Challenges analysing this data
- Big Data platforms
- RTTQA approach

Randomised Clinical Trials

What do we use RCTs for?

- Clinical trials can change practice and help set standards of care in the UK and international
- Trials can also be used to introduce new techniques and technologies

Primary Questions

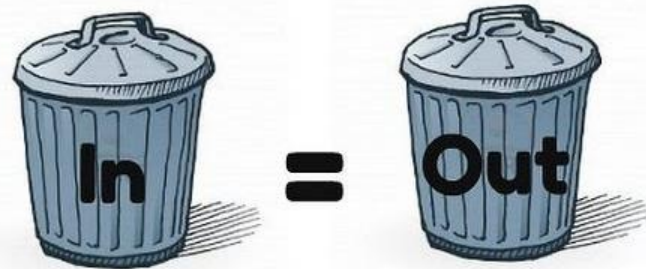
- Overall Survival
- Progression Free Survival
- Toxicity
- Secondary Outcomes...

Practice Changing Trials (Breast Radiotherapy)

- Traditional Breast Radiotherapy
 - 50Gy in 25 fractions over 5 weeks
- Start trial (1999-2002)
 - 40Gy in 15 fractions over 3 weeks
- Fast Forward (2011- ongoing)
 - 27/26Gy in 5 fractions over 1 week

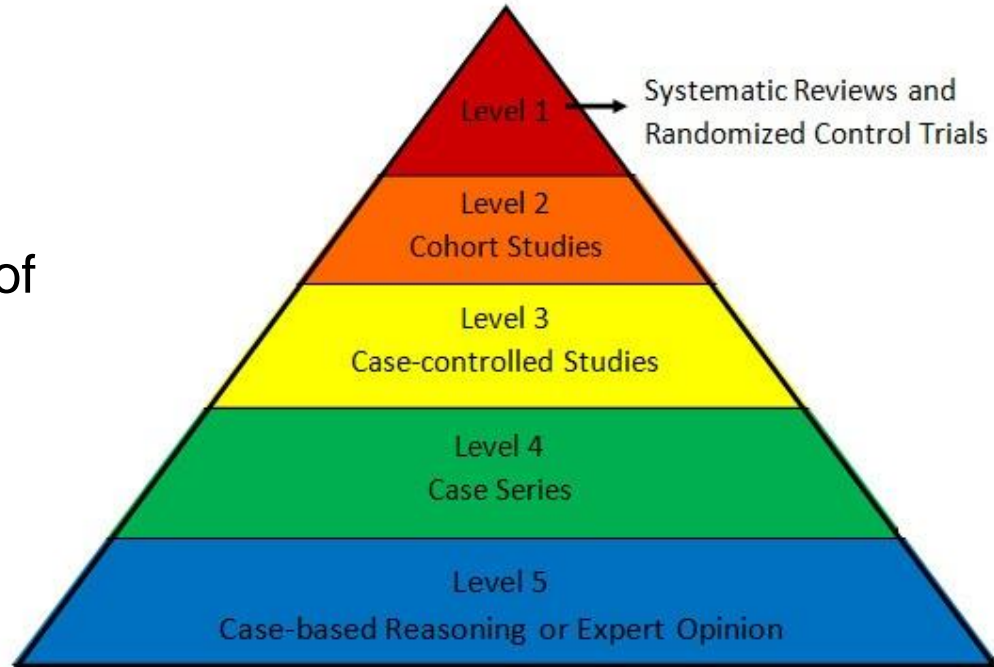
Trial Analysis

- How do we know if we improved outcomes?
 - Look at survival and recurrence of disease
- How do we know if we improved side effects?
 - Measure Serious Adverse Events and Serious Adverse Reactions
- Perform statistical analysis which helps us interpret the results of the trial
- Garbage in, Garbage out



Why is data from RCTs valuable?

Oxford centre for
Evidence Based
Medicine – Levels of
Evidence



Big Data and Trials

- Big Data refers to the collection and analysis of large sets of data elements and relationships that are difficult to process using traditional methods



Challenges of Data Analysis

- Data Quality - Bias
- Patient Privacy
- Geographical spread
- Interoperability

Big Data in Radiotherapy

Residual setup errors towards the heart after image guidance linked with poorer survival in lung cancer patients: do we need stricter IGRT protocols?

(Residual setup errors towards the heart linked to poorer survival)

Corinne N. Johnson^{1,2}, MSc, Gareth J. Price^{1,2}, PhD, Corinne Faivre-Finn^{1,2}, MD, PhD, Marianne C. Aznar^{1,2}, PhD and Marcel van Herk^{1,2}, PhD.

¹Manchester Cancer Research Centre, Division of Molecular and Clinical Cancer Science, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester Academic Health Sciences Centre, UK

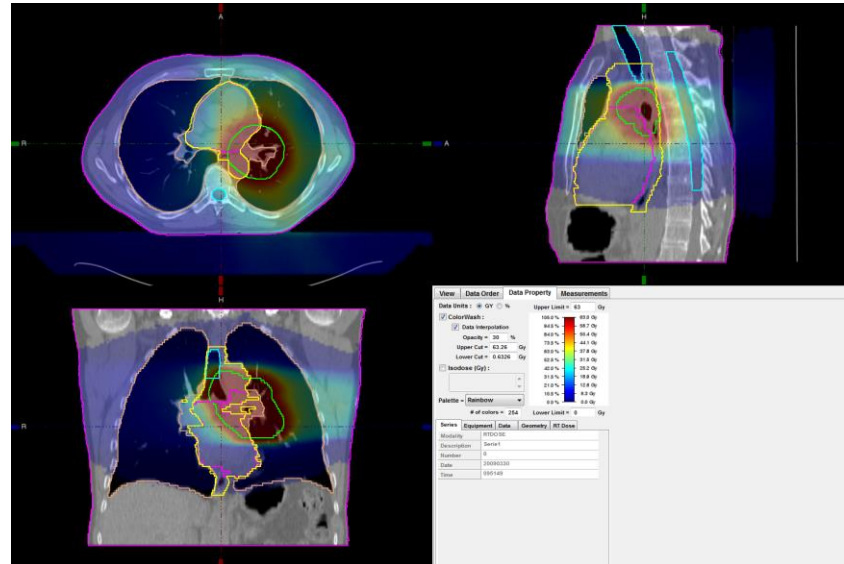
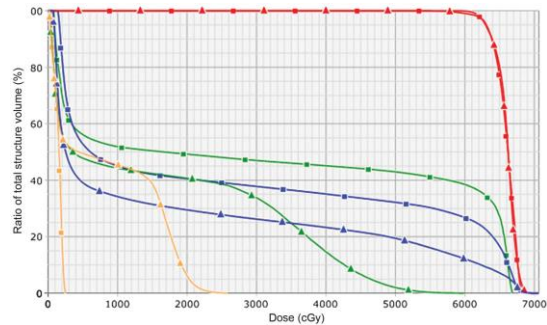
²The Christie NHS Foundation Trust, Manchester Academic Health Sciences Centre, UK

Why are trials good for big data

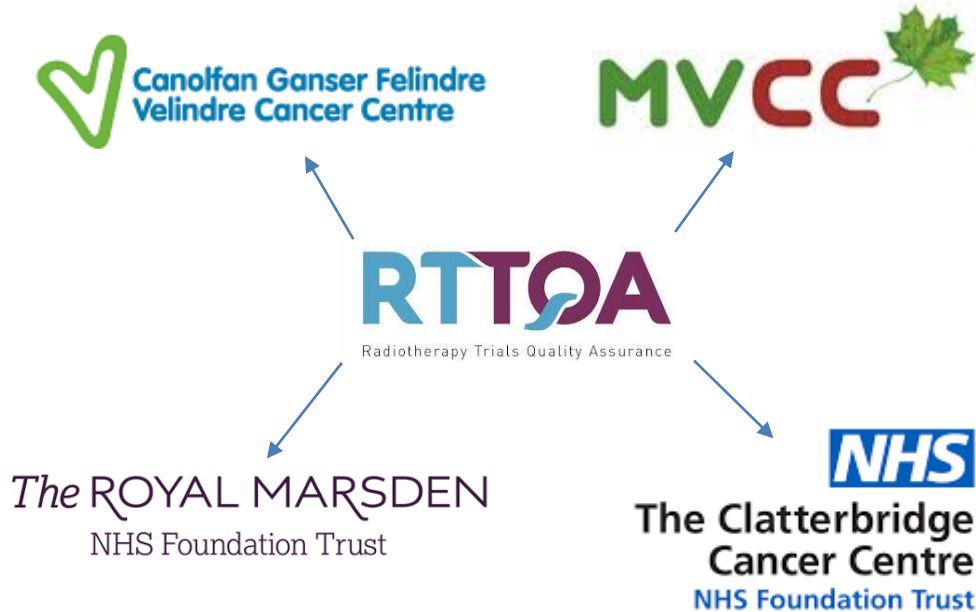
- Data Quality – High Quality & Low Bias
- Patient Privacy – Consent & Ethics
- Geographical spread – Data centrally collected
- Interoperability – Standardisation

Data Collection for Trials – Treatment Data

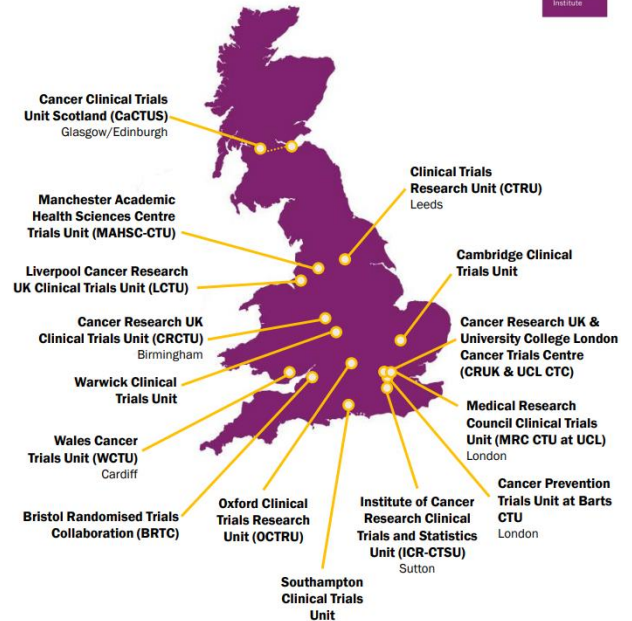
- Standardised DICOM format (mostly)
- Patient imaging – CT scans, MRI, PET.....
- Outlining- Clinician Contouring
- Plan files
- Dose files



UK Perspective



Members of the NCRI Cancer CTU Group, 2015



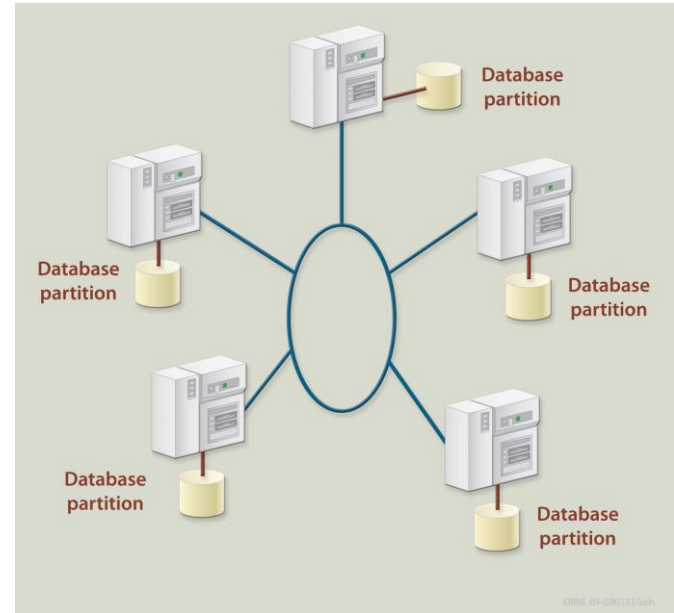
Big Data – Central Database

- Pros
 - Standardised sharing systems
 - Highly curated data
 - Data in as single system
- Cons
 - Data may not be accessible
 - Small patient numbers
 - May not represent all patients



Federated Databases

- Pros
 - No need to share data
 - Develop models and test elsewhere
 - Standard libraries
- Cost
 - High admin/setup cost
 - Ethics/access may be complex



RTTQA - XNAT System

- Built by KCL and UCL
- Database for treatment information
- Centralised data storage
- Accessible using internet for NHS/Non-NHS
- Ensures data is clean and validated
- Provides automation
- Facilitates data analysis
- Low Cost
- Does not contain clinical data, but it could



Summary

- RCTs are a source of high quality data
- This data is underutilised
- Systems to allow Big Data analysis of RCT data under development