

Real time data linkage for clinical care and research – experience from Leeds

Dr Rick Jones

+44 (0)113 3430853

r.g.jones@leeds.ac.uk

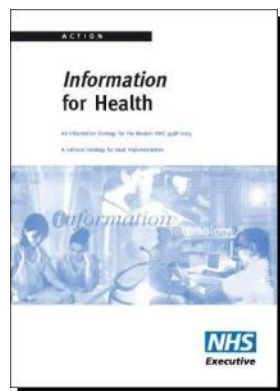


www.ychi.leeds.ac.uk

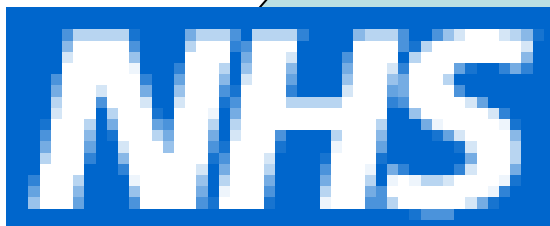
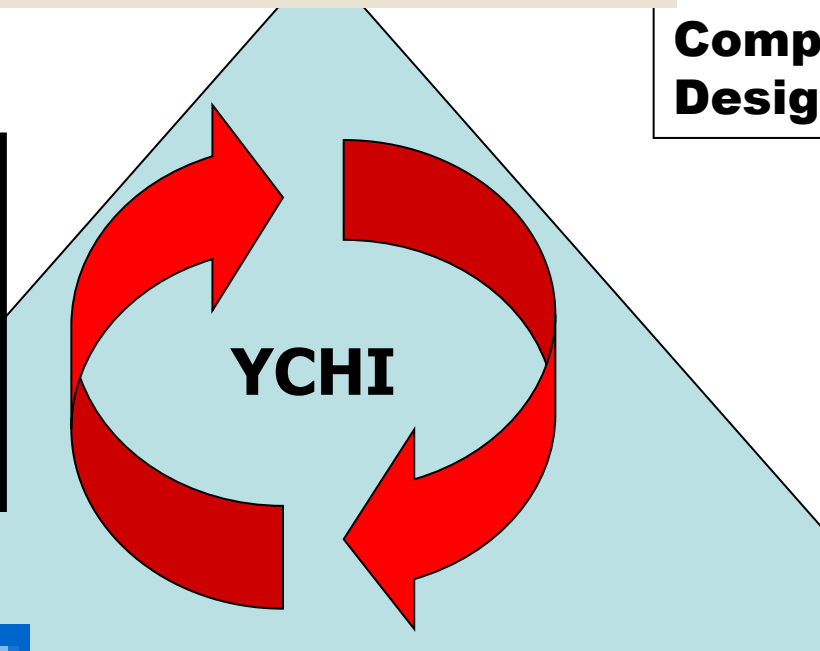
YCHI Formed in 1997



**Health
Computer Science
Design**



**DH
CfH
SHA
Trusts**



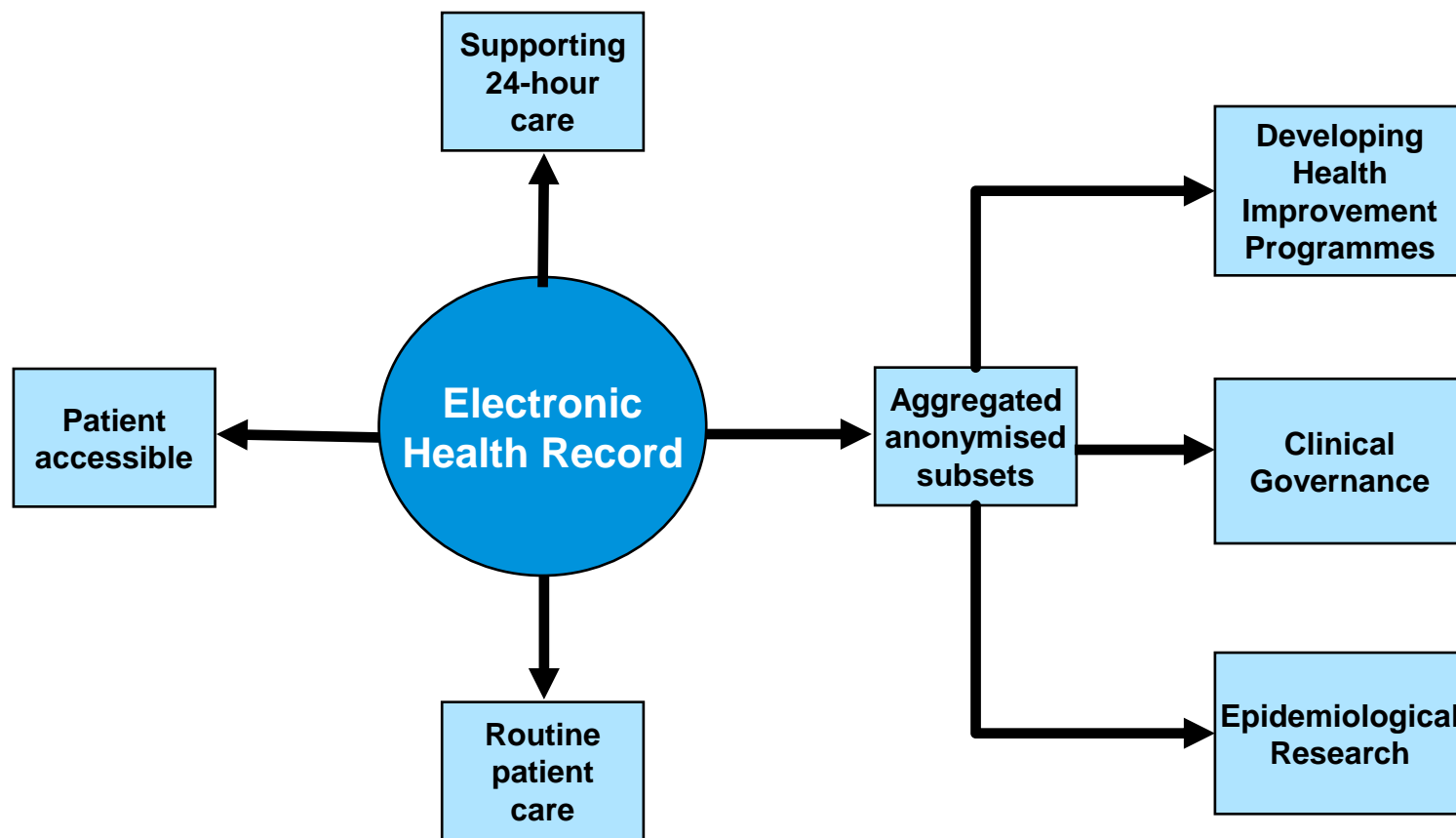
**Diagnostics
Pharma
Informatics**



Plus ca change...



UNIVERSITY OF LEEDS



Information for Health, DH, 1998.

- Teaching
 - Undergraduate medicine – informatics in the curriculum - ICE
 - Undergraduate CS - Computing for Medicine, Intercalated BSc
 - MSc Health Informatics
- Research
 - Decision support
 - Health systems
 - Health outcomes
 - Screening
- Knowledge Transfer / Innovation
 - Industrial links
 - NHS Policy & Innovation
 - Spin outs & Proof of concept
 - CPD
 - Masterclass programme for regional informatics directors

Wellcome / MRC funded study designed and validated training courses for clinical research, industrial and NHS informatics staff to aid the exploitation of digital information resources.

Research Questions

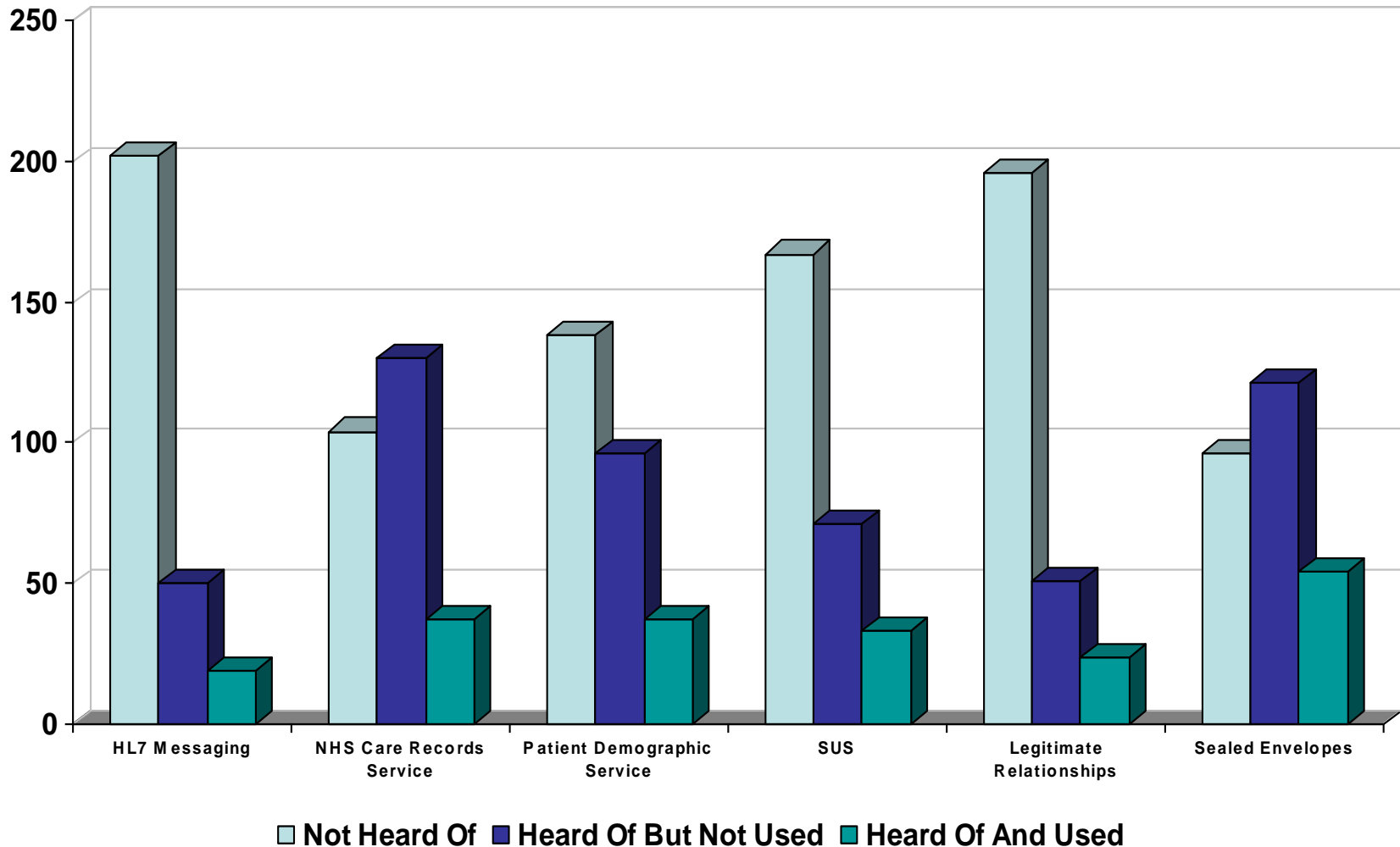
1. Can focussed training of academic and NHS database developers/administrators help:
 - record alignment between research databases and clinical systems
 - support strategic research
2. How can awareness of best practice in information systems development and management be best increased within the research and clinical IT communities?



Confirmation of Knowledge Gap



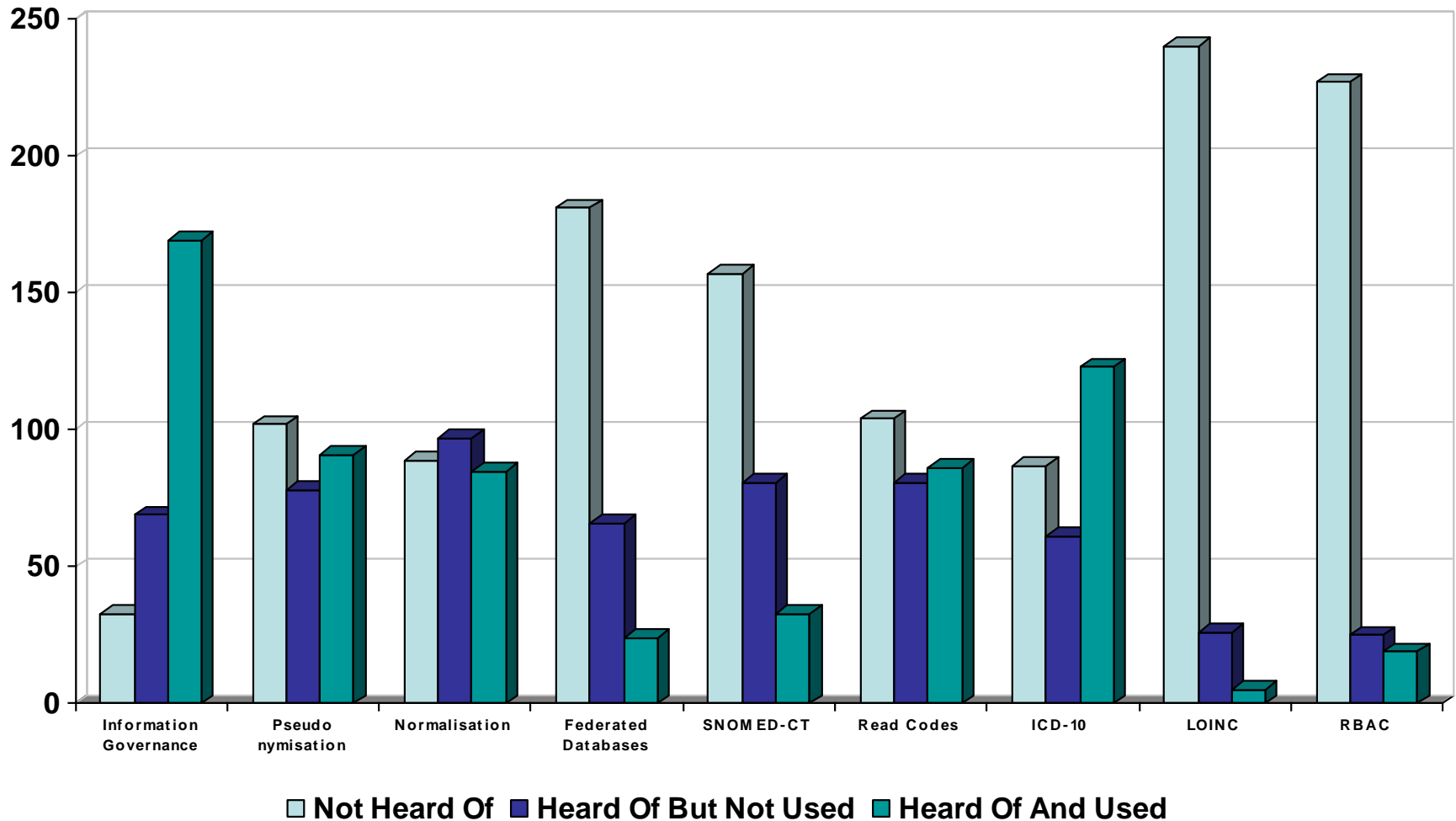
UNIVERSITY OF LEEDS



Confirmation of Knowledge Gap



UNIVERSITY OF LEEDS





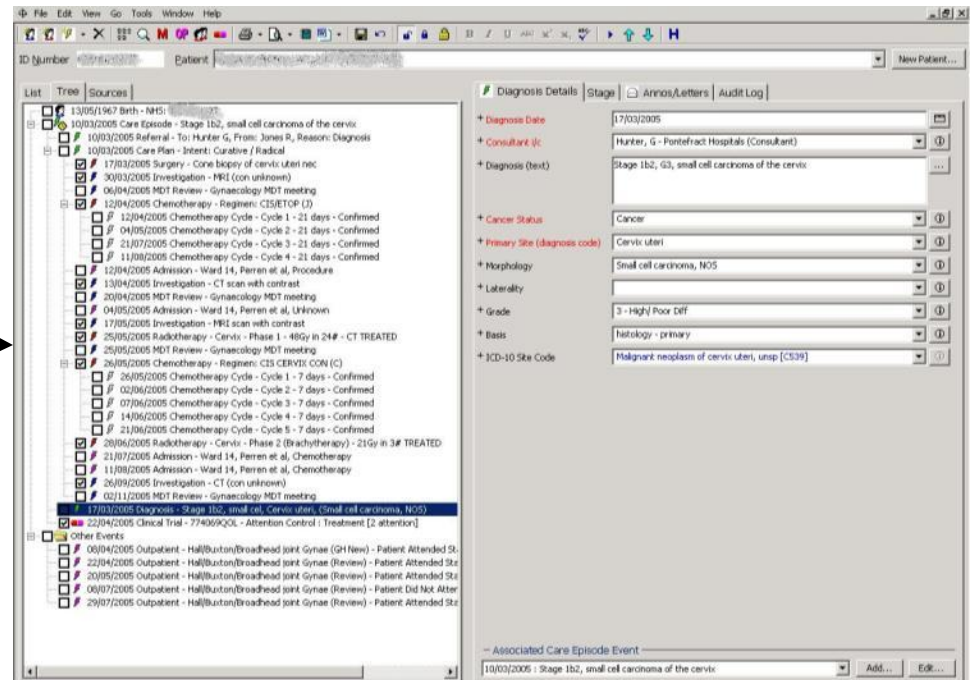
What makes Leeds unique?

- Very large health economy
 - Leeds Trust 1% of NHS turnover
 - Regional centre for Cancer / Cardiology / MSK
- Major DH / CfH footprint
 - CfH / DHID
 - NHS IC
- Large University
 - Largest medical school in UK - 1250 Undergraduates
 - Pre-eminent computer science
 - Medical engineering – hip joints, sensors
- Large e-Health industrial base
 - EMIS / TPP / CSC

Patient Pathway Manager (PPM): EPR and Coded database

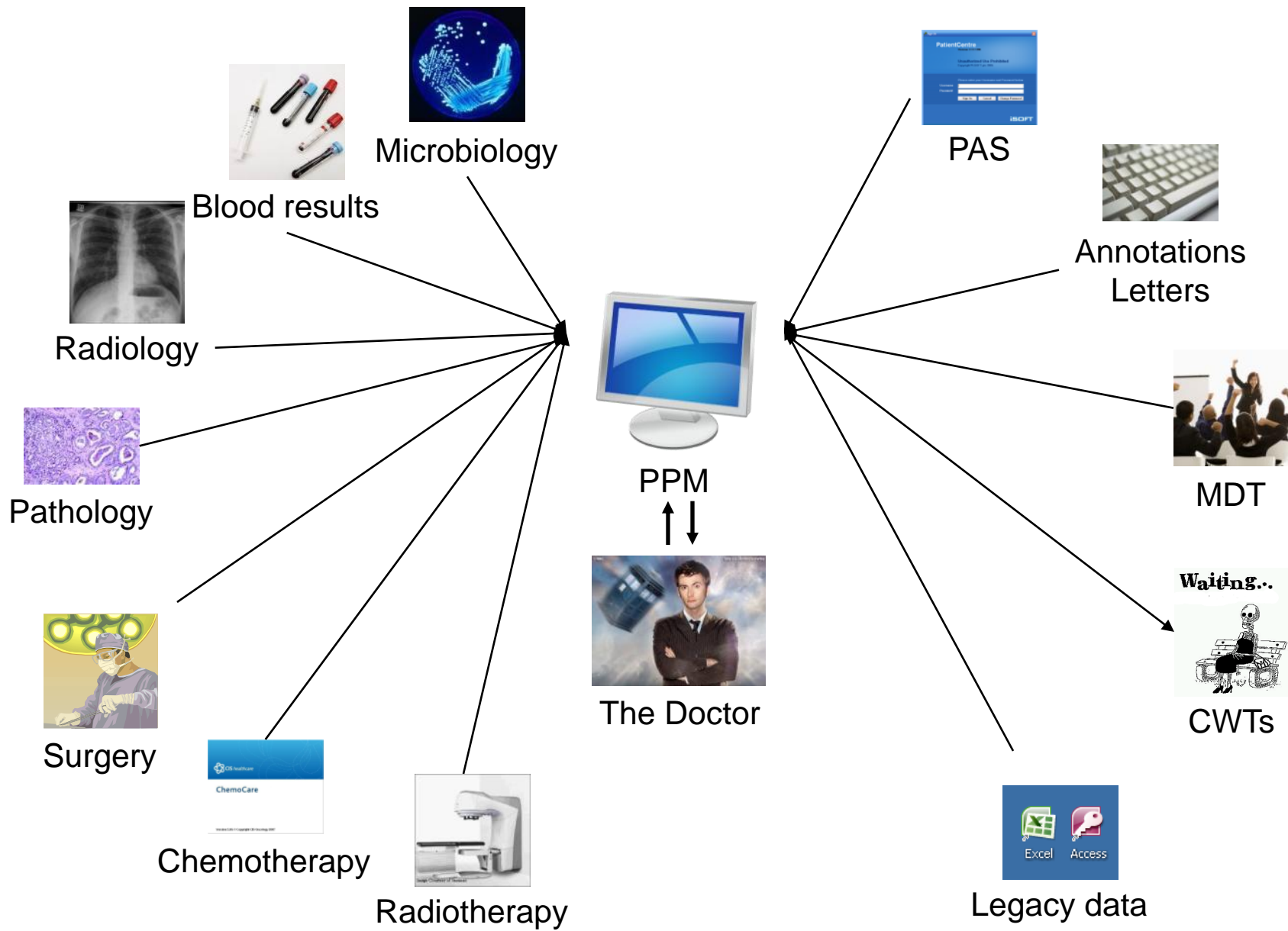


UNIVERSITY OF LEEDS



Research driven development built to support trials
University funded agile development – max 2 wte developers

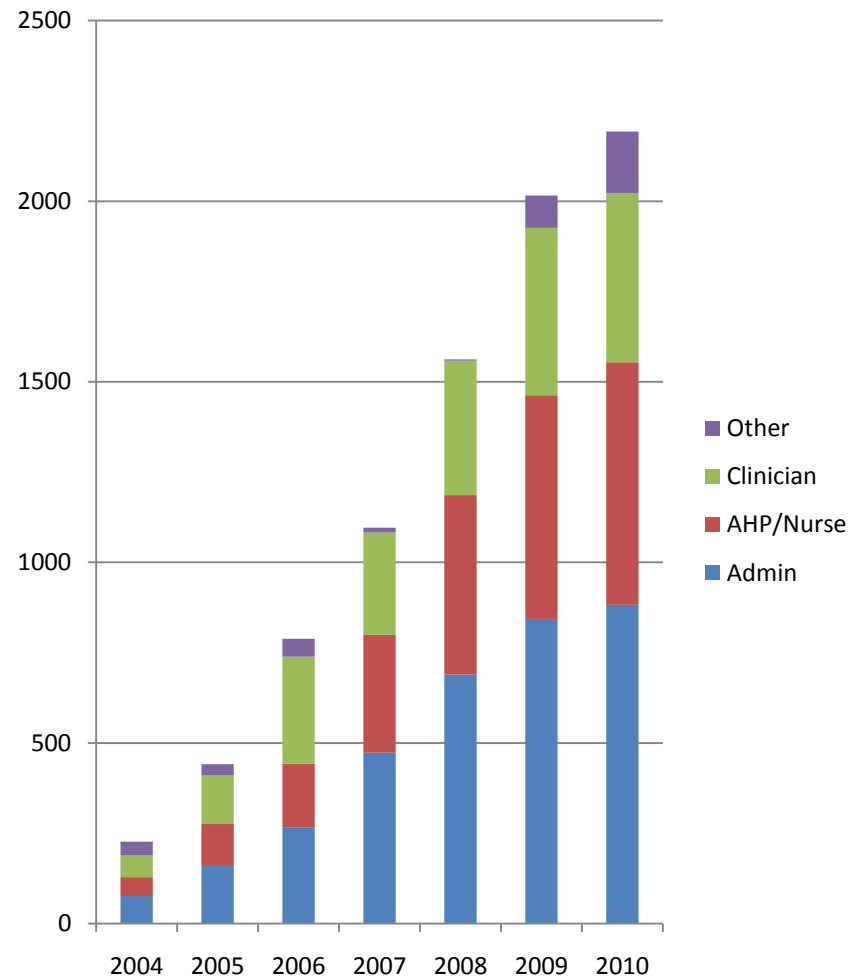
UNDERSTANDING WHAT SUCCESS IN HEALTH INFORMATION SYSTEMS LOOKS LIKE: THE PATIENT PATHWAY MANAGEMENT (PPM) SYSTEM AT LEEDS, *Owen A. Johnson, Sofel E Abiodun: 2011.*



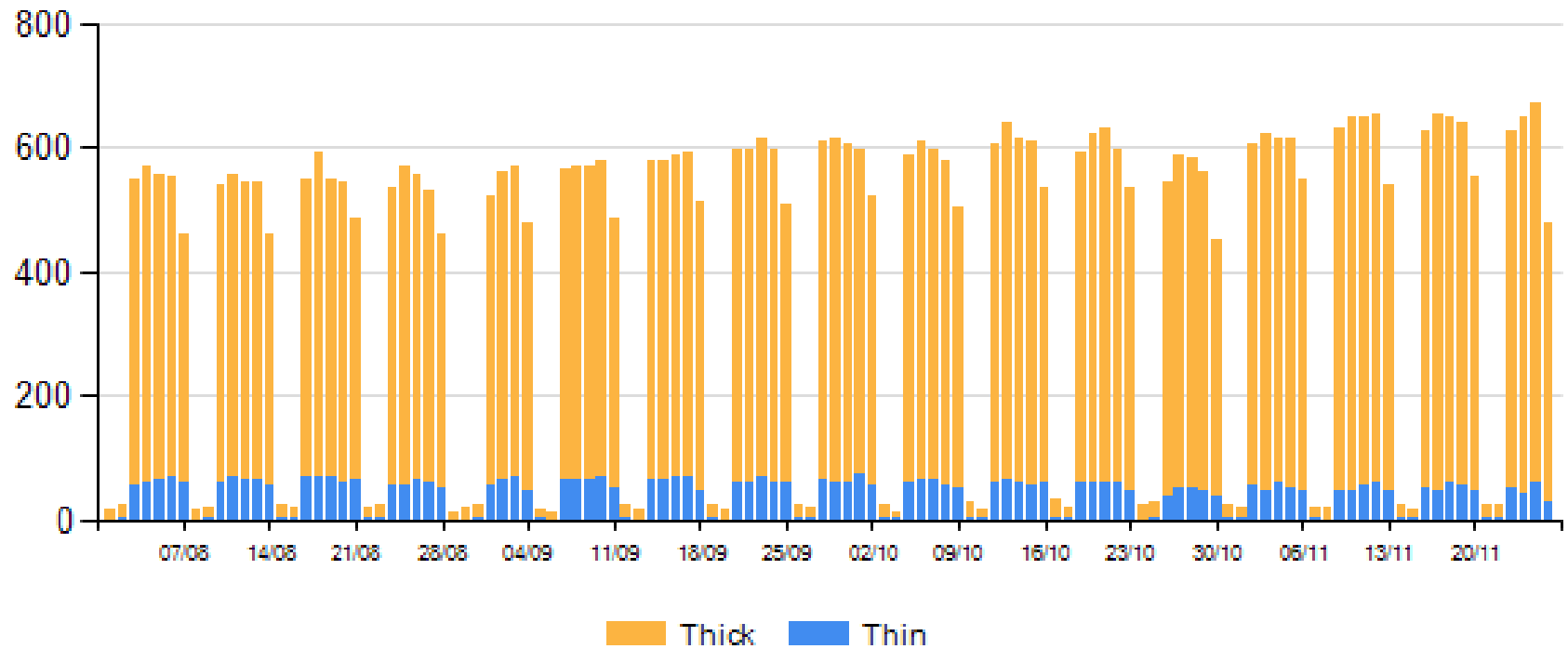
PPM - Numbers



- 3000 + active users
 - 800 users per day
 - 4 hospitals (full)
 - 9 hospitals (read-only)
- 800,000 patients
 - 1.2 million annotations
 - 16.9 million events



Daily LTHT Instance Users by Connection Type



Currently upgrading to LTHT portal using agile / open source methods



UNIVERSITY OF LEEDS

EHEALTH OPEN SOURCE



search here ...

Go


[Home](#) [About](#) [Case Studies](#) [Codeforge](#) [Contact](#) [HL7 Materials](#) [Knowledge Centre](#) [Webinar](#)

Welcome to the new home of eHealthOpenSource

A Partnership Between...



UNIVERSITY OF LEEDS

The Leeds Teaching Hospitals 
NHS Trust

THE UNIVERSITY *of* York


Open Source Community for Healthcare

Since 1975, eHealthOpenSource has been committed to ethically producing the highest quality service in the world. Today with more than 5000 employees worldwide eHealthOpenSource is still committed to same producing quality.

[VISIT OUR CODEFORGE](#)

Recent Posts

[Moorfields OpenEyes Case Study](#)
[eHealthOpenSource Is Currently Under Transition Process](#)

 eHealthOpenSource Feed

[eHealthOpenSource.org Is Currently Under Transition Process](#)
[Cough up the hairball and make NHS technology open](#)

Useful Links

[Open Source Initiative](#)
[OS Health Informatics UK](#)



PPM Portal
went live 17th
March 2012

- “TPP aims to connect different healthcare organisations through comprehensive IT solutions. Founded in 1999 in West Yorkshire.”
- “TPP is the company that produces SystemOne clinical software. SystemOne fully supports the NHS vision for a ‘one patient, one record’ model of healthcare. Professionals should be able to access a single source of information, detailing a patient’s contact with the health service across a lifetime. This record should be accessible whatever the care setting and available so any health professional can enter information. It should document every appointment, every medication, every allergy and every contact the patient has ever had.”



NME TPP SystemOne Site Deployments Per Module and SHA

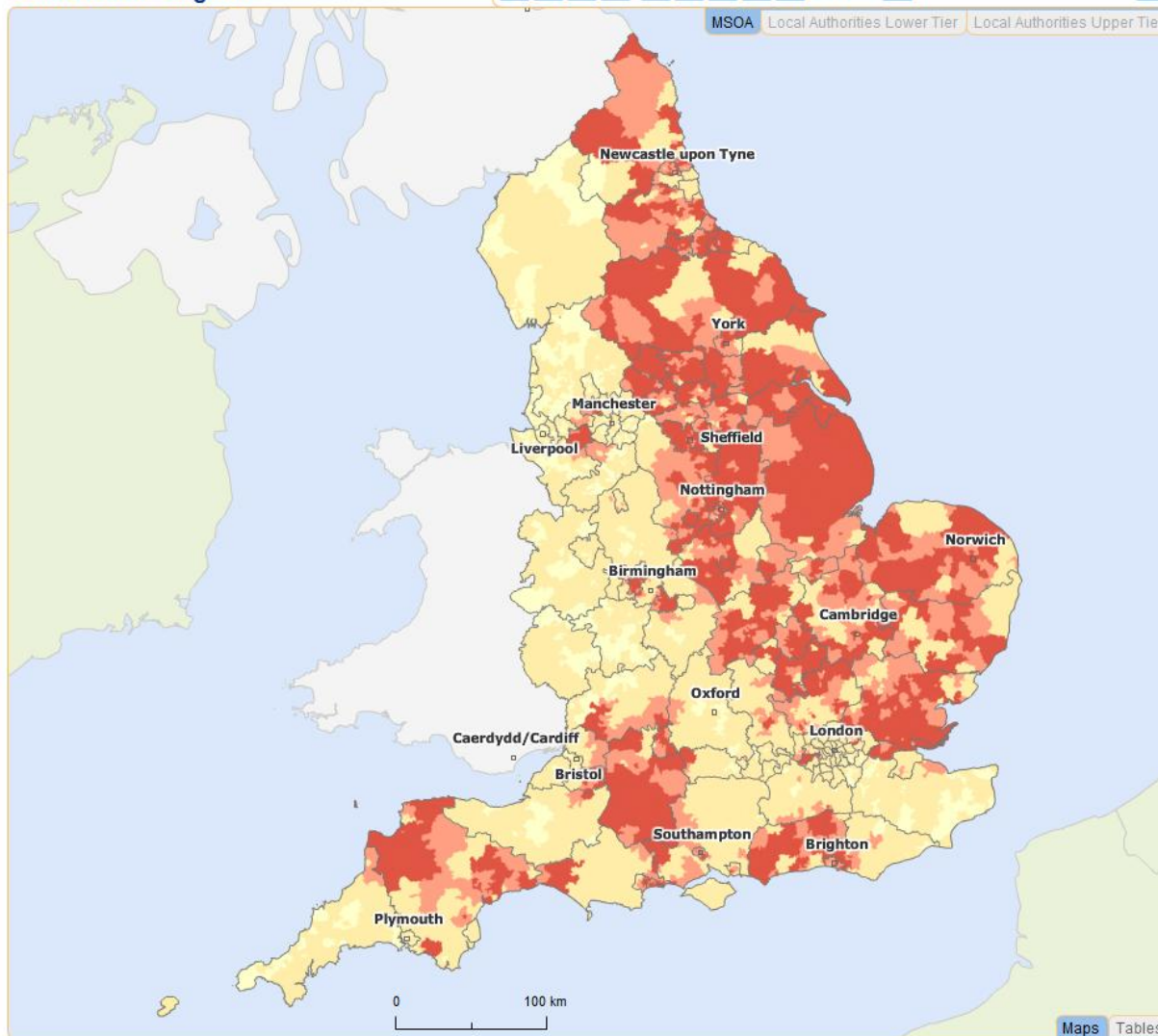
SystemOne Module	East Midlands	East of England	West Midlands	North East	North West	Yorkshire & the Humber	Totals
GP	398	390	79	131	36	553	1587
Community#	9	22	2	11	0	21	65
Out of Hours	18	11	2	11	0	12	54
Minor Injury Unit	3	1	0	0	0	5	9
Child Health#	10	19	2	9	0	18	58
Hospice	4	7	10	2	3	24	50
Prisons*	17	15	12	8	10	16	78





TPP SystemOne - England

Local Health - England



Analysis with colour shading

Imported or computed data

Patient Count

Patient Count

- 3,000 to 20,926 (1,711)
- 3,000 to 3,000
- 500 to 3,000 (1,006)
- 500 to 500
- 0 to 500 (3,341)
- 0 (723)

source : ...

Symbolic Analysis

Select a theme...

Select an indicator...

4 colours, from lightest to darkest, represent the number of current GP registered patients in that area:

- 0
- 1-500
- 501-3000
- >3000

OS Openspace Street Maps

England Zooms

Navigation

Zoom in: 100 %



Outputs



Leeds: Exploring new paradigms



UNIVERSITY OF LEEDS

- This coverage of primary and secondary care populations is unique in UK.
- Individually and collectively it opens up new approaches to health research
 - More efficient trial endpoint harvesting
 - Data mining of existing data
 - New trial paradigms
 - Data linkage in real time

PPM – research / audit

Ovarian Cancer



UNIVERSITY OF LEEDS

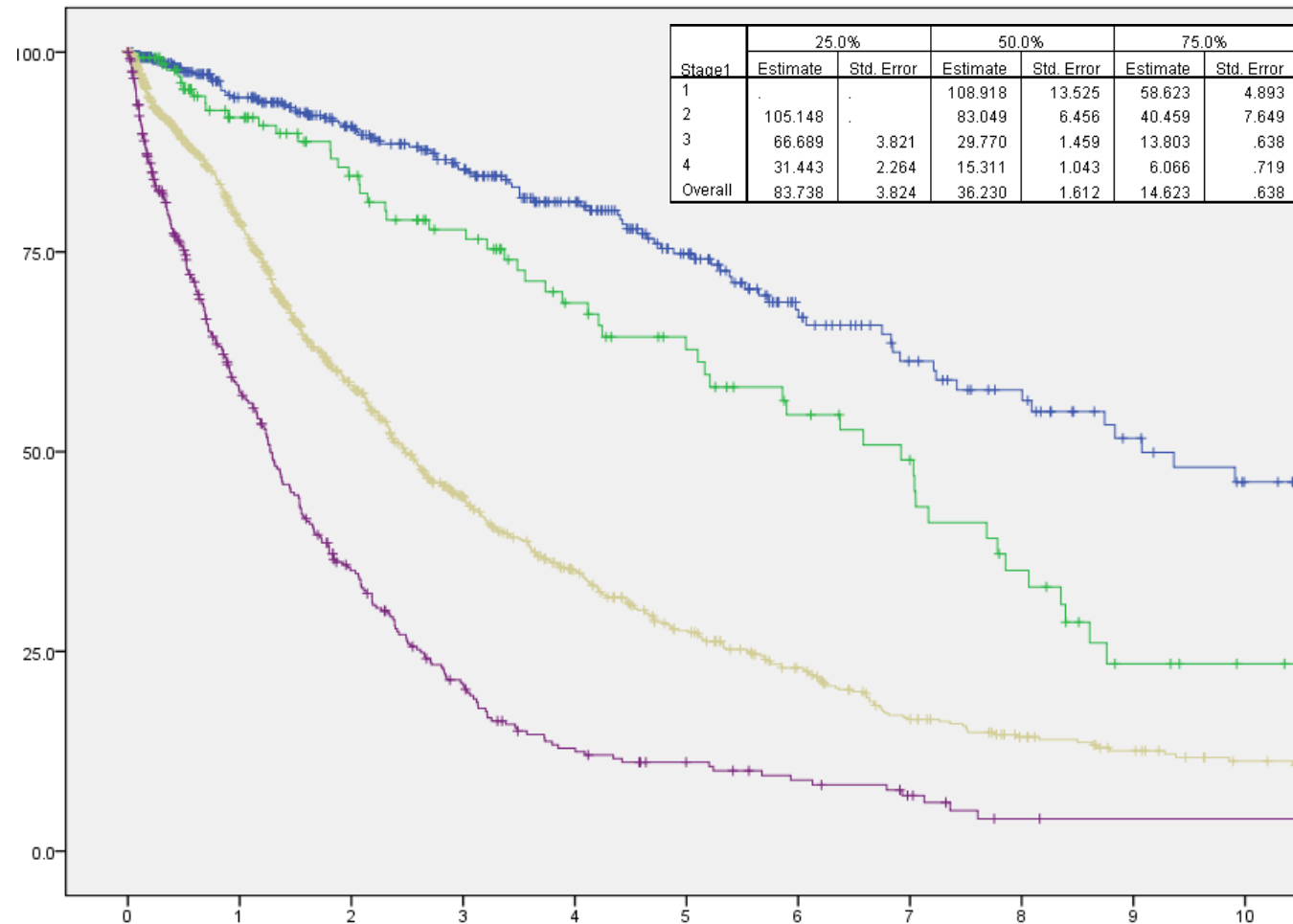
100094197	Null	September 13, 1936	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	67
100094471	Wakefield PCT	March 15, 1934	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	16
100094845	Wakefield PCT	February 6, 1934	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	46
100095129	Leeds PCT	June 3, 1931	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	111
100095638	Leeds PCT	December 5, 1933	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	64
100095943	Wakefield PCT	February 27, 1923	Ovary	3	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	13
100096283	Calderdale PCT	October 2, 1945	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	90
100096284	Leeds PCT	October 16, 1941	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	12
100096314	Wakefield PCT	June 26, 1952	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	37
100096510	Leeds PCT	March 14, 1944	Ovary	4	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	True	52
100096648	Leeds PCT	August 20, 1950	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	13
100097306	Kirklees PCT	July 7, 1935	Ovary	3c	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	True	28
100097737	Leeds PCT	August 14, 1949	Peritoneum - Mullerian origin	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	31
100098281	Leeds PCT	June 10, 1950	Ovary	4	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	False	99
100098385	Wakefield PCT	October 15, 1940	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	33
100098436	Leeds PCT	June 27, 1955	Ovary	2c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	63
100098566	Other PCT	October 30, 1947	Ovary	3c	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	True	46
100098677	Leeds PCT	April 21, 1943	Ovary	3a	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	116
100098841	Leeds PCT	December 19, 1941	Ovary	3b	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	79
100099417	Leeds PCT	January 2, 1944	Ovary	2b	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	97
100099745	York PCT	December 19, 1943	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	34
100099780	Leeds PCT	September 11, 1921	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	88
100100006	Leeds PCT	September 28, 1933	Ovary	2c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	94
100100163	Leeds PCT	June 16, 1945	Peritoneum - Mullerian origin	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	32
100100247	Kirklees PCT	June 22, 1952	Peritoneum, NOS	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	51
100100779	Wakefield PCT	November 1, 1944	Peritoneum - Mullerian origin	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	94
100101247	Wakefield PCT	August 18, 1942	Ovary	3c	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	True	51
100102077	Leeds PCT	May 29, 1930	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	13
100102741	Leeds PCT	September 2, 1930	Ovary	1a	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	54
100102762	Wakefield PCT	March 19, 1931	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	14
100102850	Leeds PCT	March 22, 1929	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	37
100103170	York PCT	June 13, 1931	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	86
100103243	Kirklees PCT	January 13, 1936	Peritoneum, NOS	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	91
100103282	Leeds PCT	January 22, 1938	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	50
100104102	Kirklees PCT	December 19, 1939	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	False	82
100104179	Leeds PCT	September 30, 1928	Ovary	4	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	15
100104190	York PCT	December 5, 1938	Peritoneum - Mullerian origin	3c	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	True	35
100104251	Leeds PCT	January 21, 1942	Ovary	4	3 - High/ Poor Diff	Papillary serous cystadenocarcinoma	True	3
100104258	Wakefield PCT	September 1, 1944	Ovary	3c	3 - High/ Poor Diff	Serous cystadenocarcinoma, NOS	True	31

Clinical Outcomes

Ovarian - stage specific survival



UNIVERSITY OF LEEDS

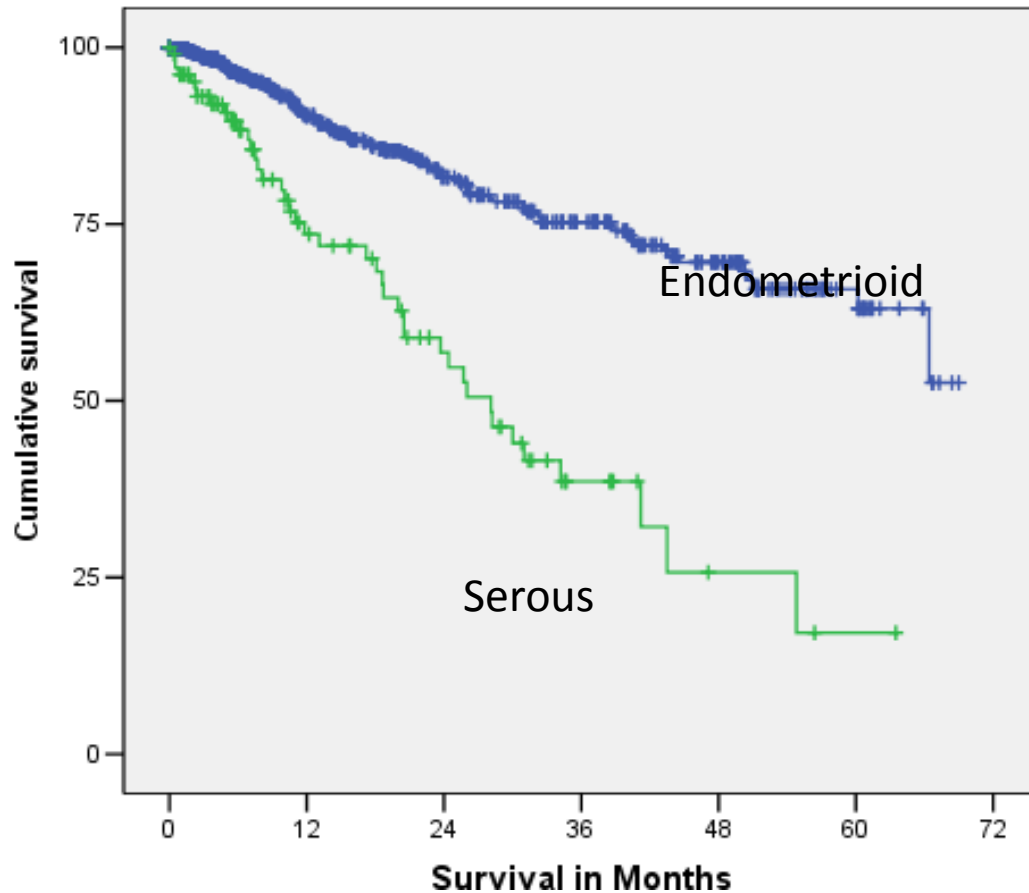


Clinical Outcomes

Endometrial - morphology specific survival



UNIVERSITY OF LEEDS



- 919 cases
 - 813 endometrioid
 - 106 serous
- Median overall survival
 - Overall = 66. months
 - Endometrioid > 6 years
 - Serous = 28.1 months

- Advanced disease at diagnosis well documented
 - 100% collection of stage at presentation
- Recurrent disease hard to assess
 - No 'code' for advancing / progressive disease
 - Difficult to re-stage for EVERY treatment delivered
- Hence potential analysis of historical data
 1. Analysis of tumour markers
 2. PPM Event Profiling
 3. Natural language processing

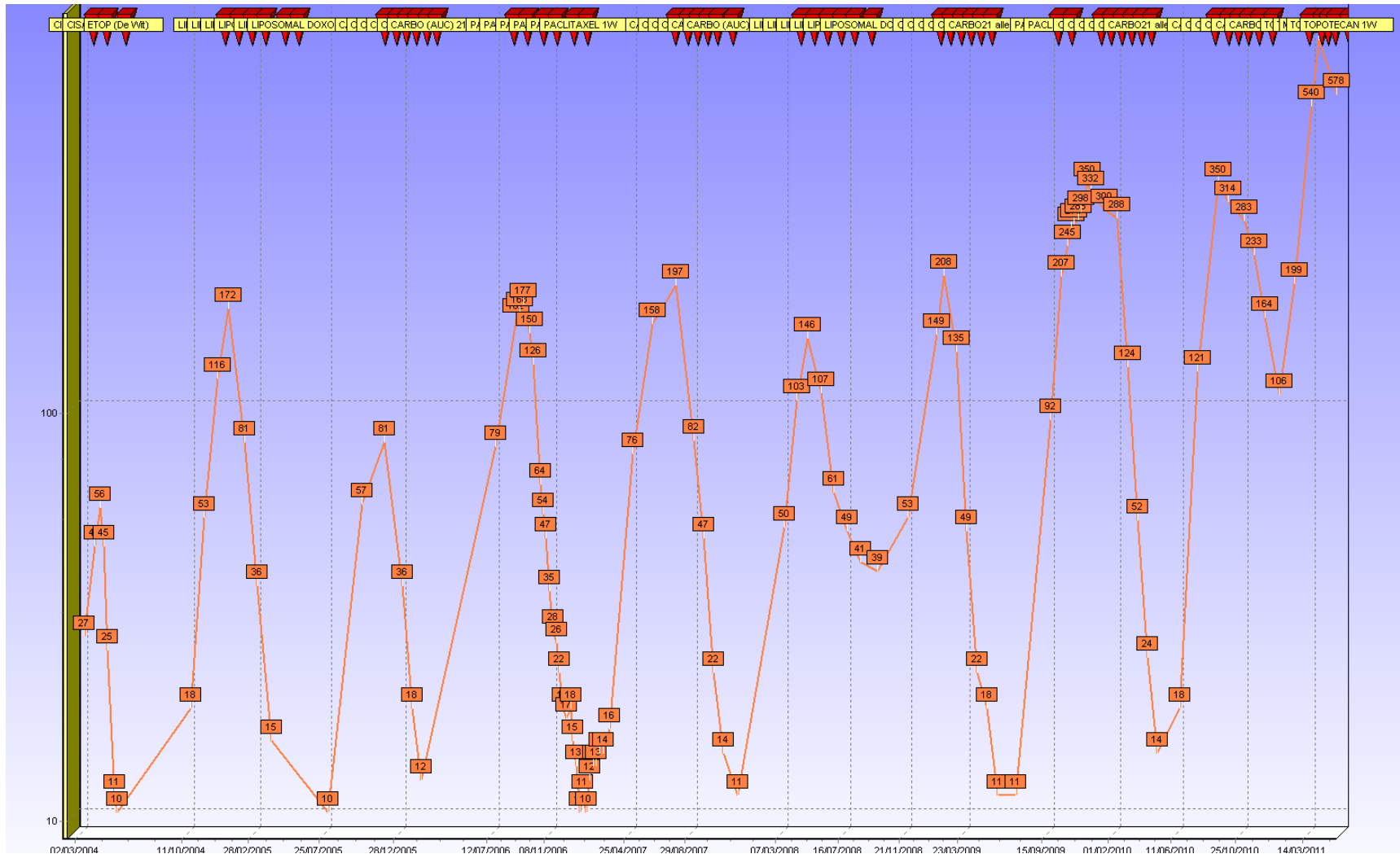
Drivers

- Better indicators of clinical need
- Trial recruitment e.g. IMPACCT – end of life pain management

Identification of advanced disease



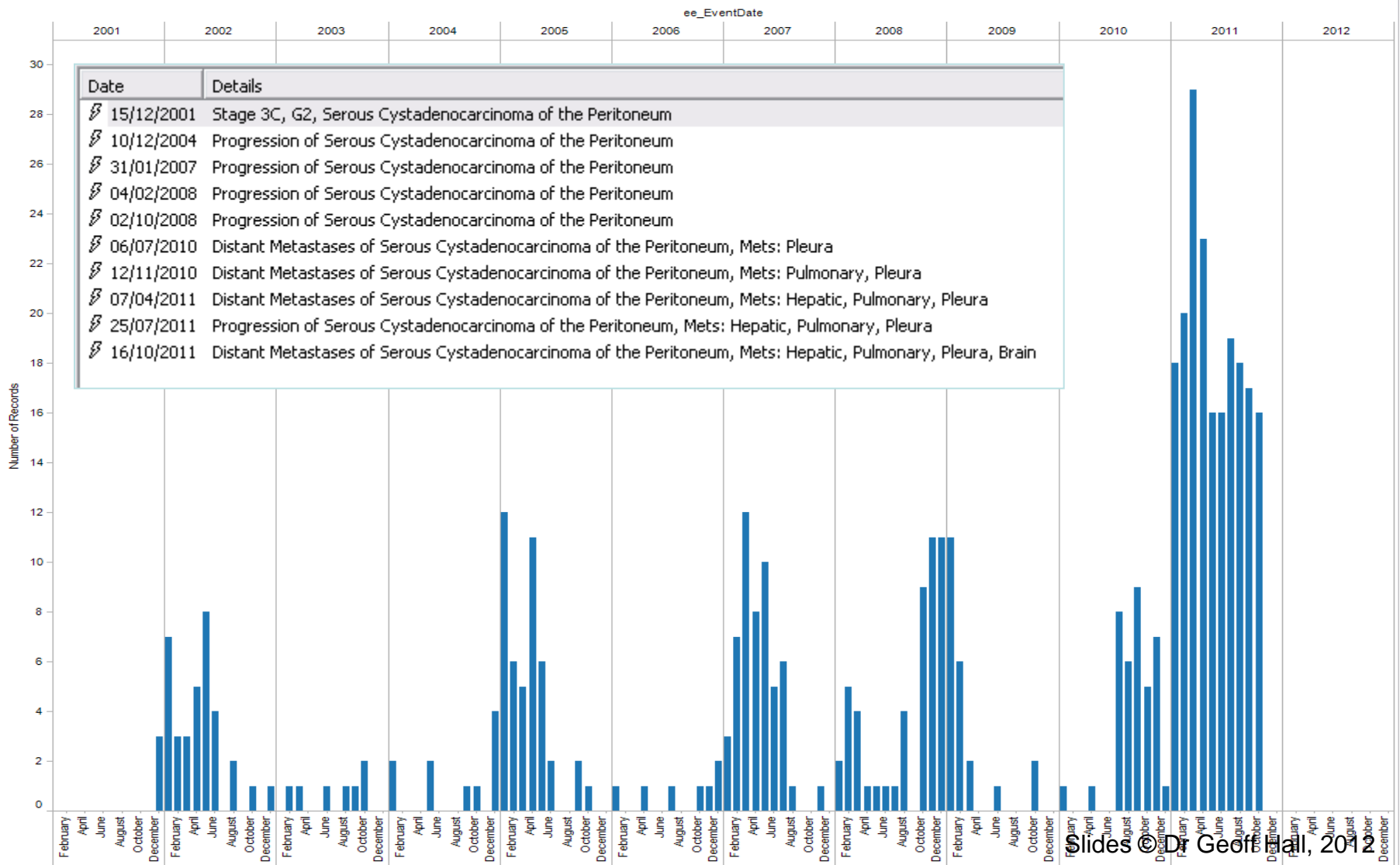
UNIVERSITY OF LEEDS



Identification of advanced disease: PPM Event Profiling



UNIVERSITY OF LEEDS



Exploiting Unstructured Data

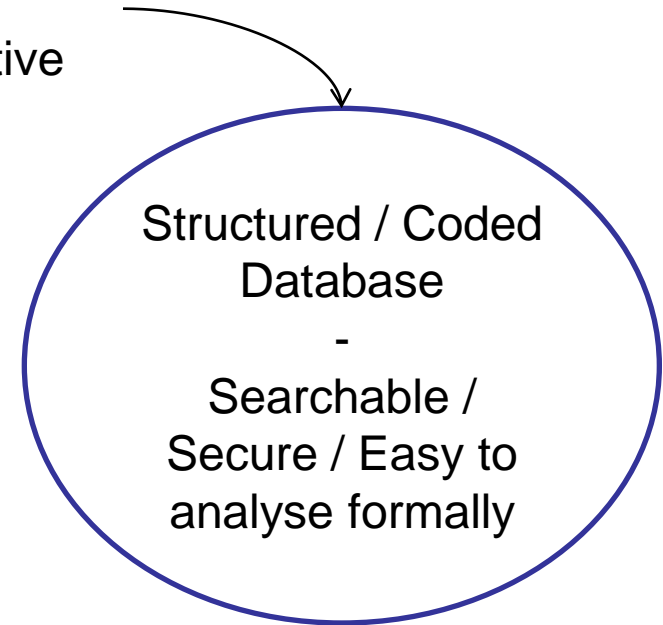
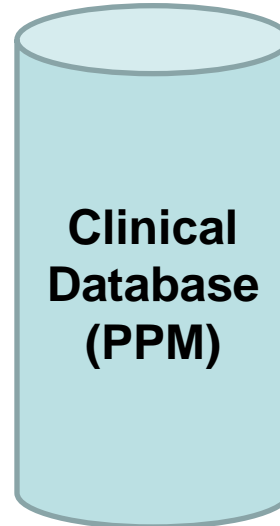
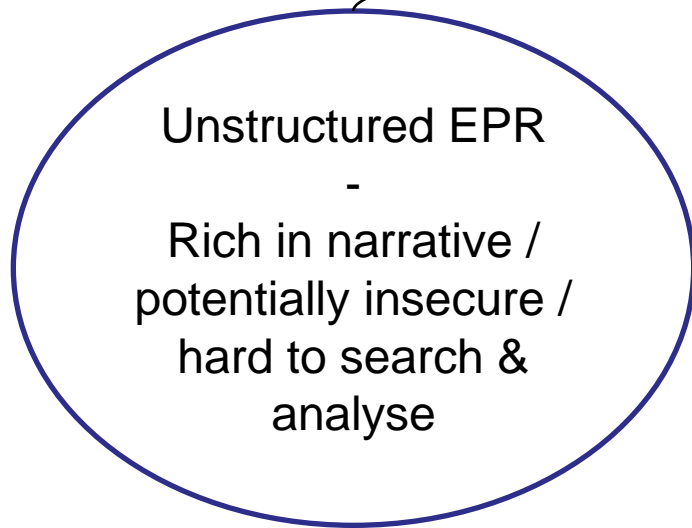
Natural language processing



UNIVERSITY OF LEEDS

Most information on progression is in the narrative account

extract and impose
structure on the narrative
NLP



**Supports
Clinical
“Business as
usual”**

**Supports
Enriched
Research
Datasets**

Identifying breast cancer recurrence, Atwell, Johnson and Hall, CancerResearchUK, 2012



Technology Strategy Board Knowledge Transfer Partnership



- Owen Johnson, Senior Fellow from the University of Leeds says **“The KTP will be jointly supervised by academics from both the School of Computing and Health Informatics, which is part of the School of Medicine. The key challenge is bringing these different disciplines together and realising that the greatest innovation comes from combining fresh perspectives. The big challenge in health, perhaps the biggest challenge there is, is using what we already know (locked inside our computer systems and databases) to create new and better understandings of how to treat patients and improve their lives.”**



Technology Strategy Board Knowledge Transfer Partnership



- Developed protocol for an ethical research database ethical research
- TPP ResearchOne – launched May 2012
- Ethical approval with NIGB



Hospital Episode Statistics (HES)

**National
Information
Governance
Board**



NHS
National Patient Safety Agency
National Research Ethics Service



Technology Strategy Board Knowledge Transfer Partnership



- 92% SUS-sample patients have a SystemOne record
- 68% SystemOne patients have a recent (2/5yr) BMI
- 23% have smoking status entered in last 5yr (11% in 2yr)
- 0.6% consultations have implausible dates
- All have sex, postcode and DoB
- Multiple imputations introduce random variability into predictions of missing values that are based on existing values from other variables





Technology Strategy Board Knowledge Transfer Partnership



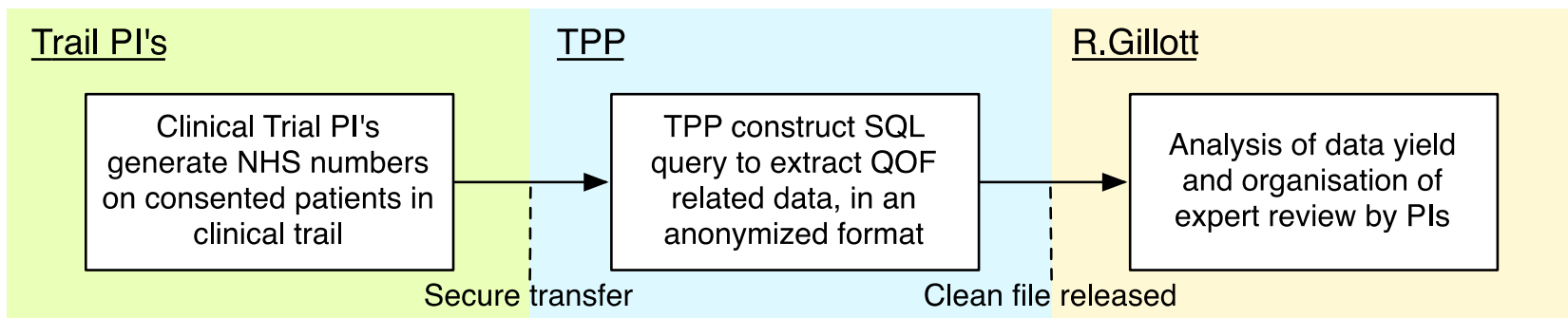
- ResearchOne -
 - eHIRC partnership (MRC bid)
- Leeds Patient Pathway Manager (PPM)
 - Primary and tertiary care linkage
 - Assists the DH 'strategy for cancer'
 - tracking primary care access
 - Cancer patients with diabetes
- IMPROVE-PC
 - Improving Prevention Of Vascular disease in Primary Care
- Syndromic Surveillance with HPA
 - Weekly/daily counts of symptoms recorded by GPs
 - Programme to extract the feeds is written
- Electronic notification of infectious diseases
 - Successful tender to develop and launch a system-agnostic eNOID though automatic and clinician-led data reporting

Harvesting clinical trial endpoints from TPP Pilot trial (R Gillott, 2011)

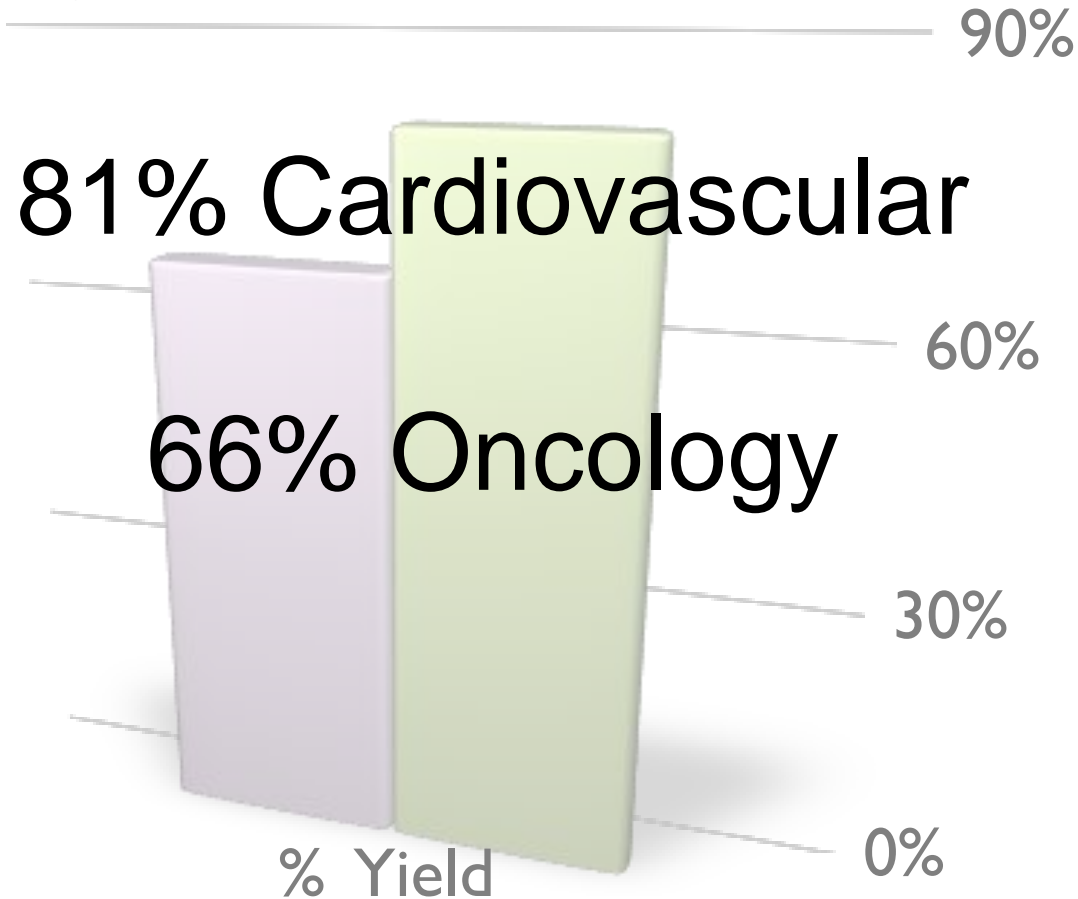


UNIVERSITY OF LEEDS

- An MRC funded studentship looked at 5000 research subjects in existing trials in oncology and cardiology.
- The pilot aimed to demonstrate the value of the information contained in the patient record, and prove whether the data was sufficient in its coverage of the population and its completeness



Subjects in TPP database

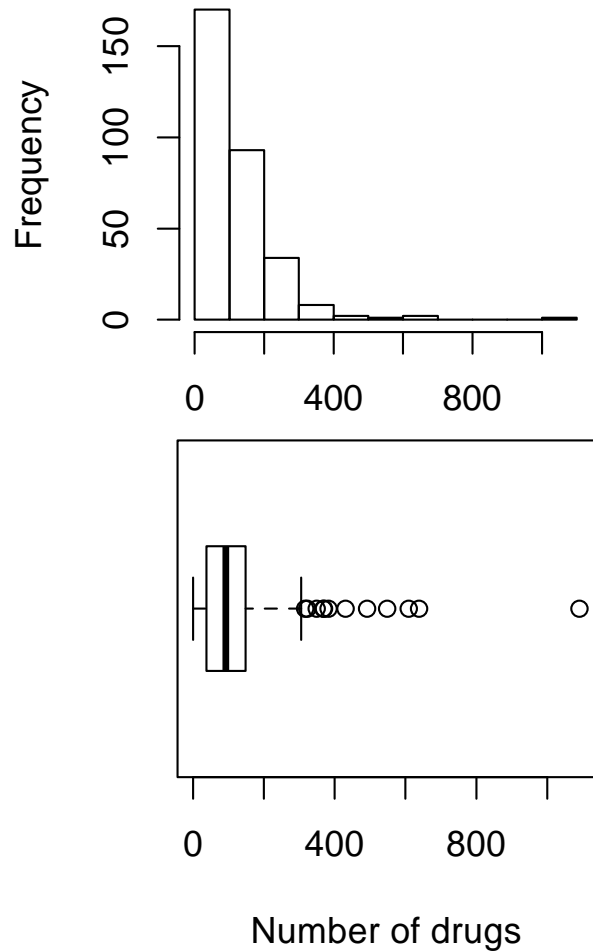


Results: Data Quality

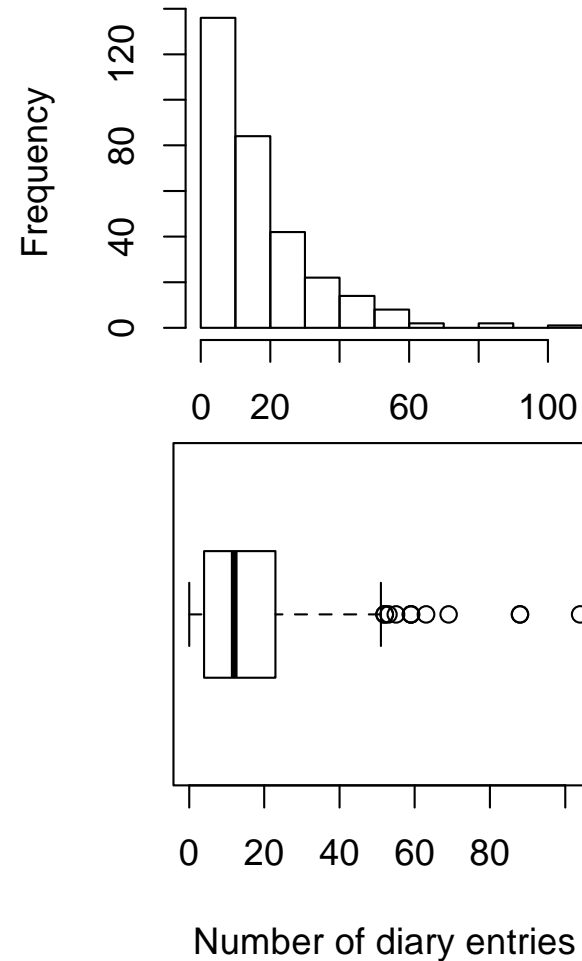


UNIVERSITY OF LEEDS

Frequency of drugs for cardiology patients



Frequency of diary entries for cardiology patients



New opportunities TSB KTP with Philips Healthcare



UNIVERSITY OF LEEDS

Question – could home testing provide better chemotherapy management?

- Reduced adverse events
- More tailored protocols
- Improved quality of life



Example of ‘teaching driven research’ –
Result of summer project by 2
undergraduates (Karl Baker – CS, Rohan
Goel Med) underpinned the KTP
application.

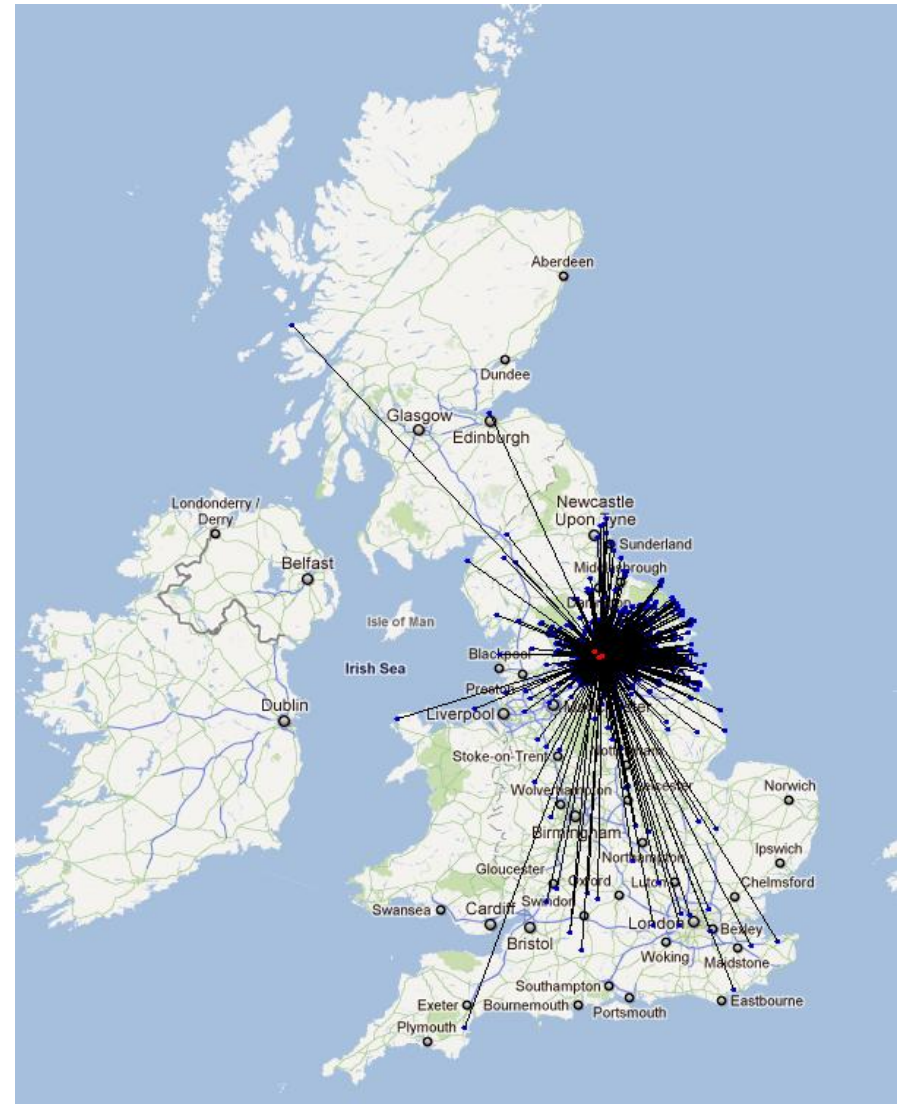
PPM Datamining

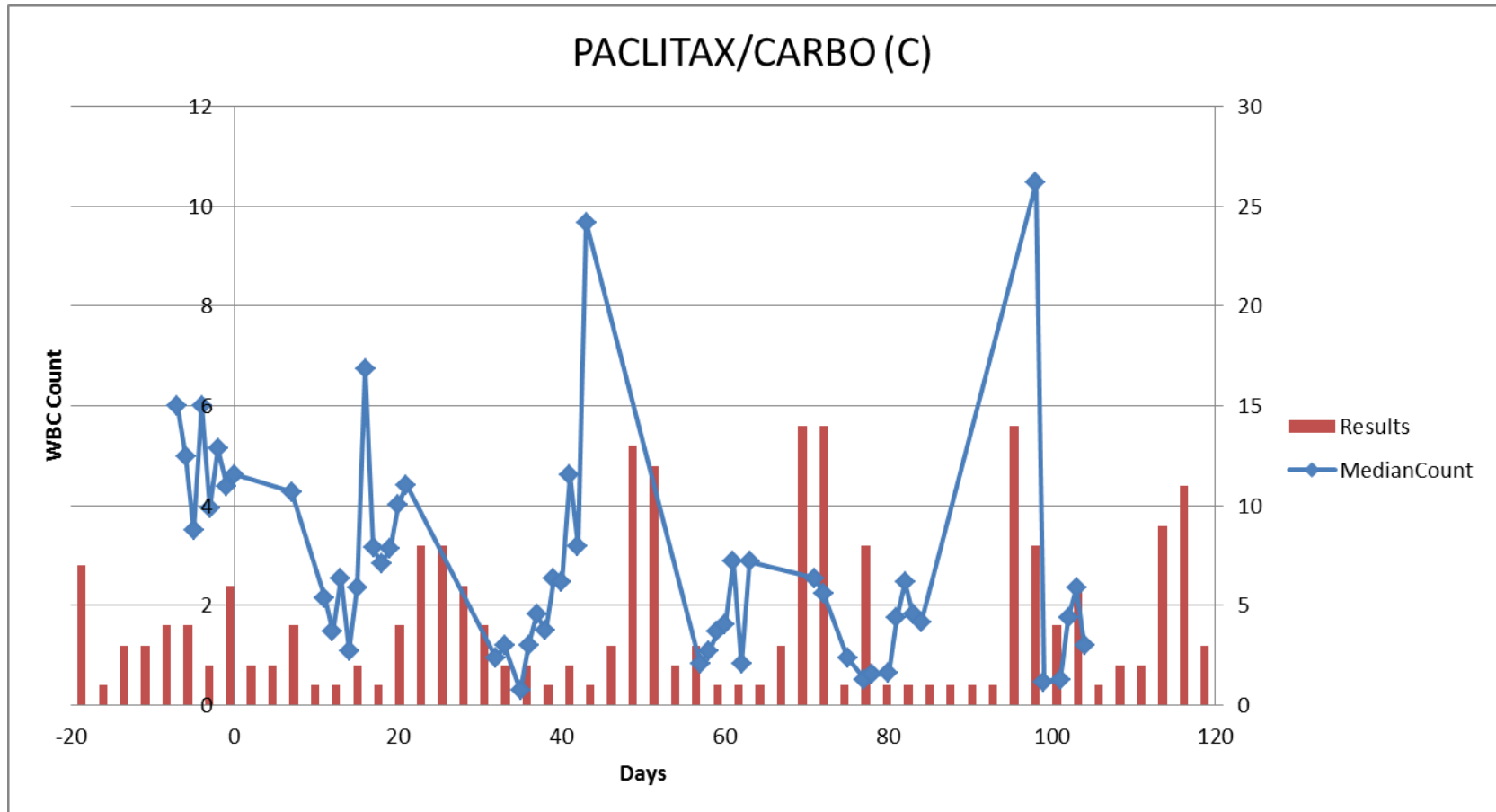
Travel distances for Chemotherapy



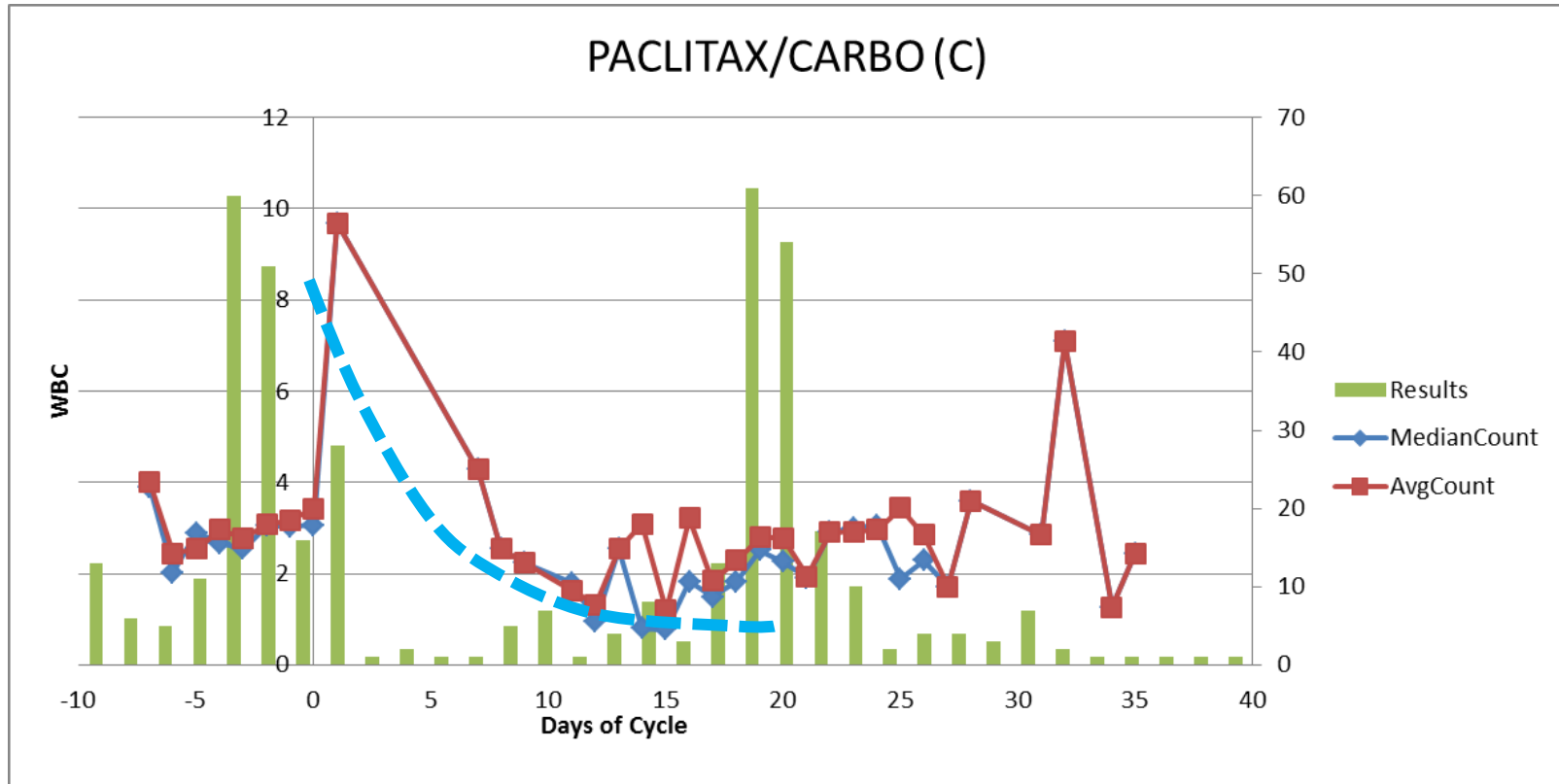
UNIVERSITY OF LEEDS

- Data mining and Google mash-up of 5000 patient episodes.
- Average distance 32kms per chemotherapy visit
- Some long distances are possible artefacts of temporary residence
- Shows potential of travel avoidance / carbon effect





Progressive decline in haemopoetic reserve over repeat cycles



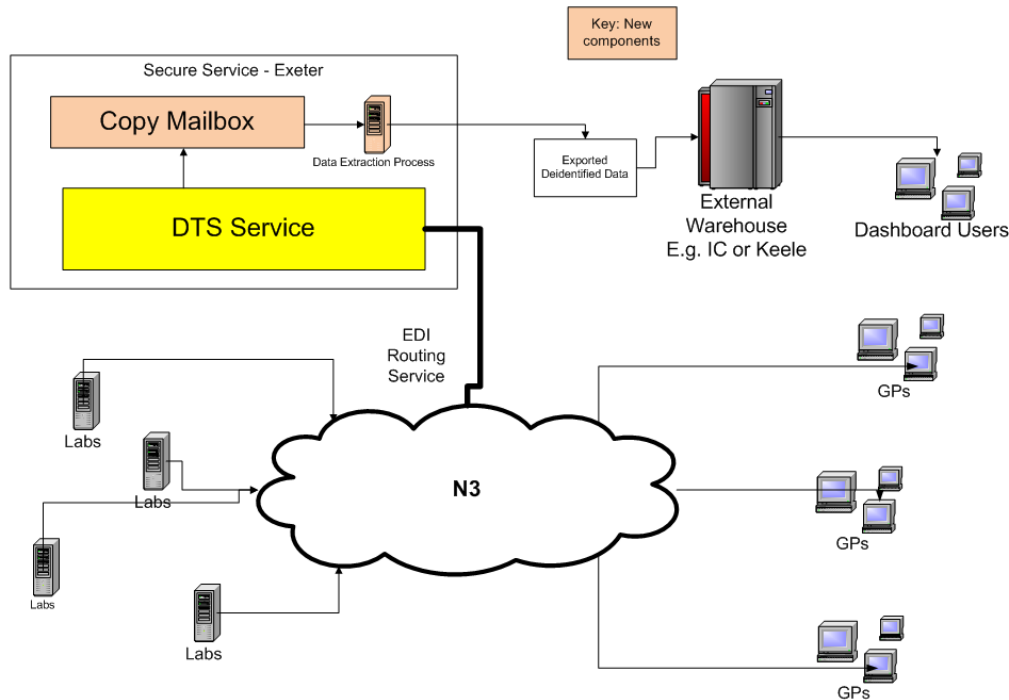
Aim is to identify the 'Luker' curve to predict within cycle response to allow individualised neutropenic risk and potential dose intensification. TSB SBRI Grant in preparation.

Capturing Operational Pathology Data at National Level



UNIVERSITY OF LEEDS

PMIP Data Extract Model – “On the fly” option



All GP pathology messages pass through a mailbox service called DTS.

A copy message is extracted.

Messages are Edifact translated and de-identified. Data in the agreed fields extracted and posted to a database file.

This data is then included in a data warehouse where it is linked to other datasets using the organisation and test codes.

This can be audited against emergent standards.

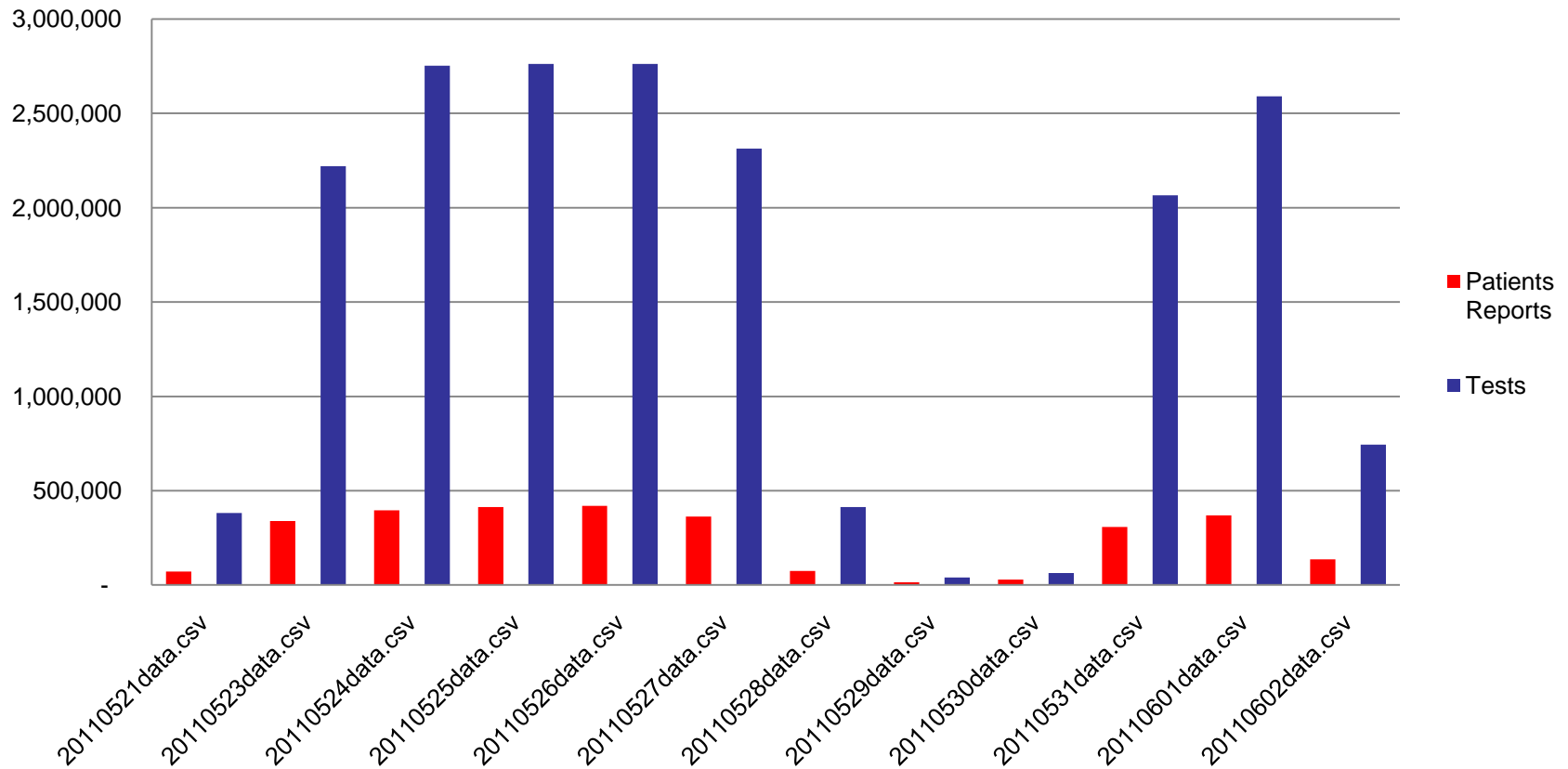
Work conducted as part of safety audit of data standards:
See Potential Clinical Errors Arising from Pathology Result Combination on Clinical Systems, Jones 2010.

- Laboratory:
 - Test code (local / READ)
 - Patient age
 - Patient gender
 - Lab ID (Org Code)
 - GP Practice ID (Org Code)
 - Date/Time of request
 - Date/Time of report
- QoF:
 - List size
 - Disease prevalence

From these we can derive:

- Workload
- Turnaround times
- Tests per patient per condition

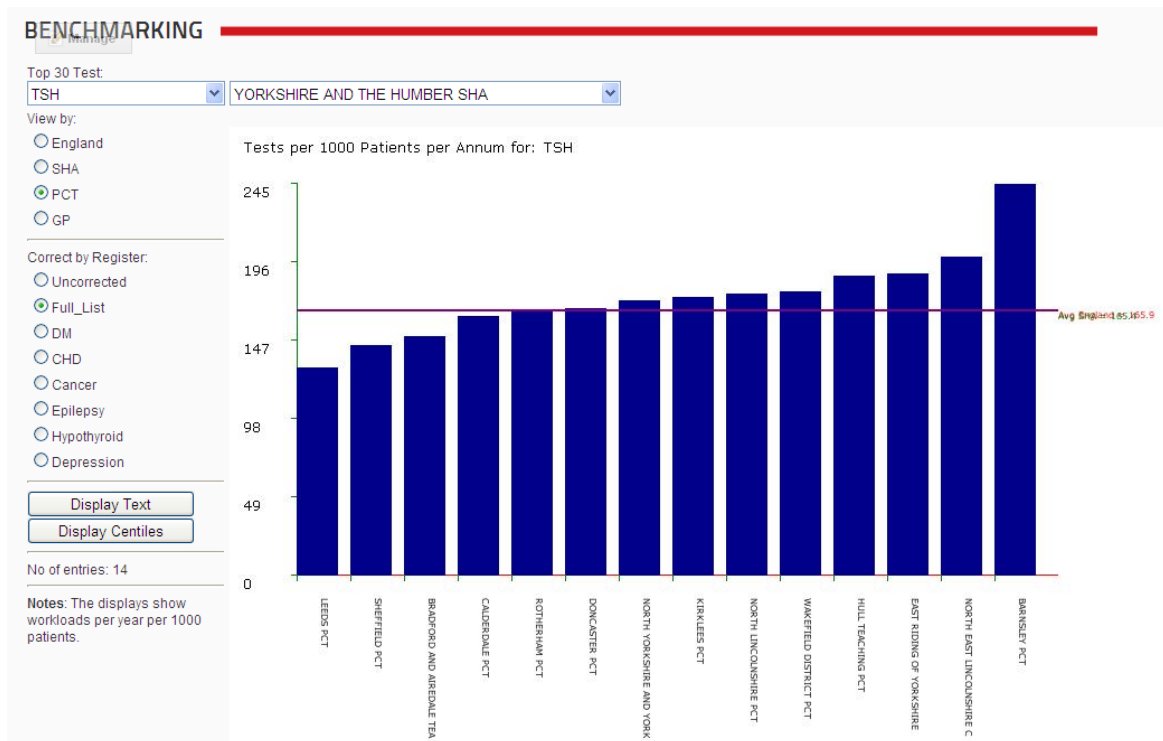
Patient Reports & Tests per Day



Capability to drill down from SHA to GP Practice



UNIVERSITY OF LEEDS



- Access to web dashboard which provides simple drill down to practice level.
- Normalisation options by QOF returns

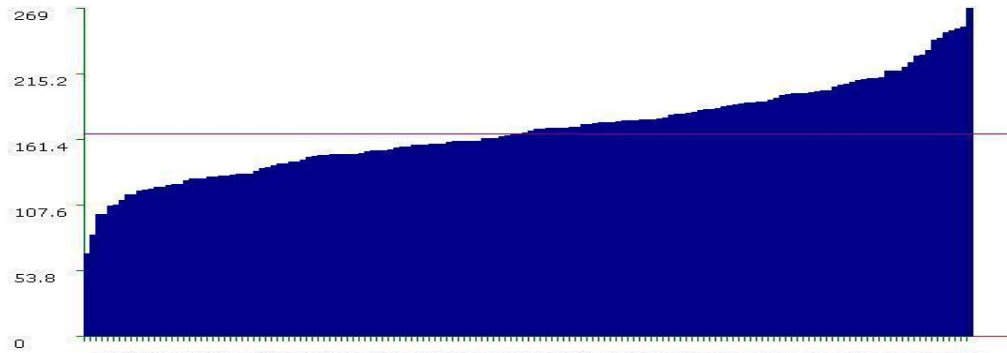
<http://www.ychi.leeds.ac.uk/pmipunits>

Capability to drill down from SHA to GP Practice

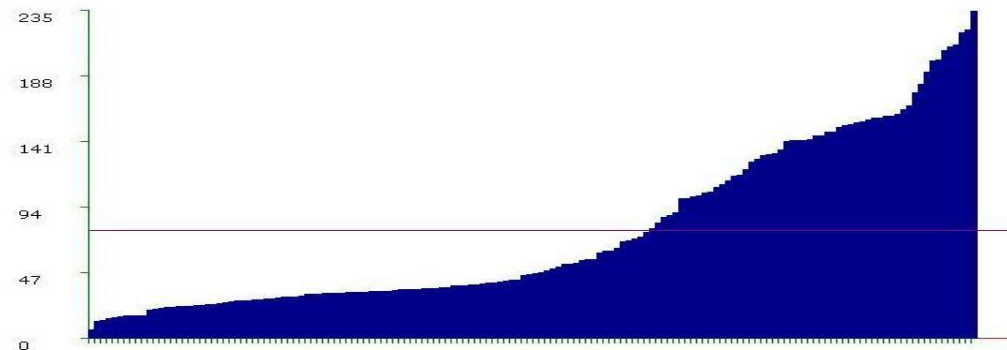


UNIVERSITY OF LEEDS

Tests per 1000 Patients per Annum for: TSH



Tests per 1000 Patients per Annum for: T4 (free)



- Gives insight into variable investigation policies
- Compliance with best practice

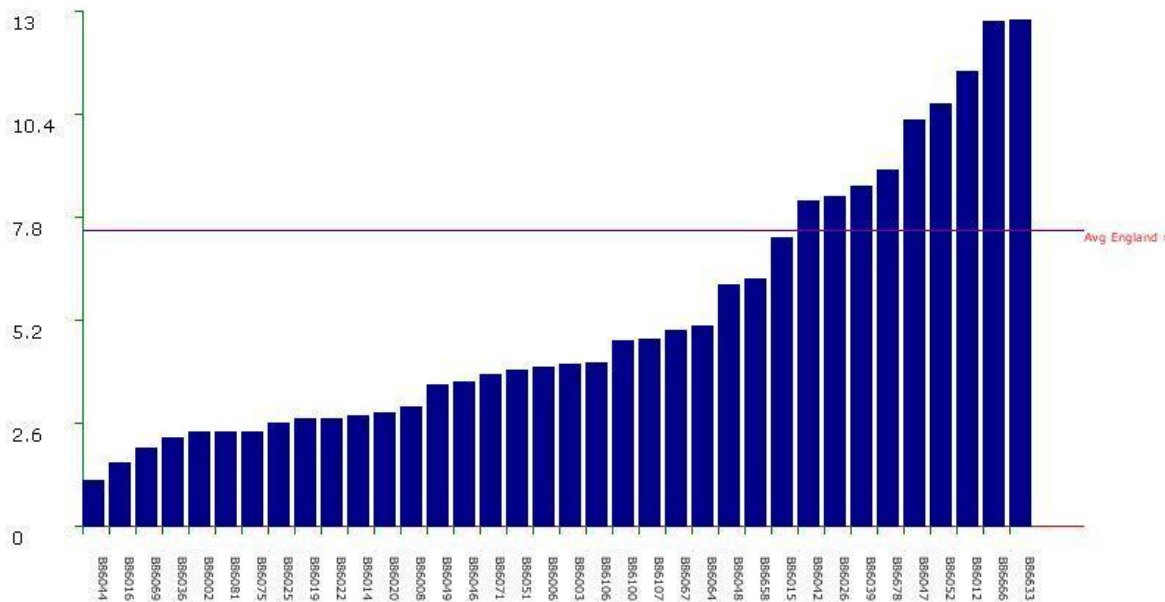
<http://www.ychi.leeds.ac.uk/pmipunits>

CA125 (Ovarian cancer) ?Link to PPM diagnostic date



UNIVERSITY OF LEEDS

Tests per 1000 Patients per Annum for: CA125



- Potentially identify practices with low / late diagnostic interventions
- Compliance with best practice

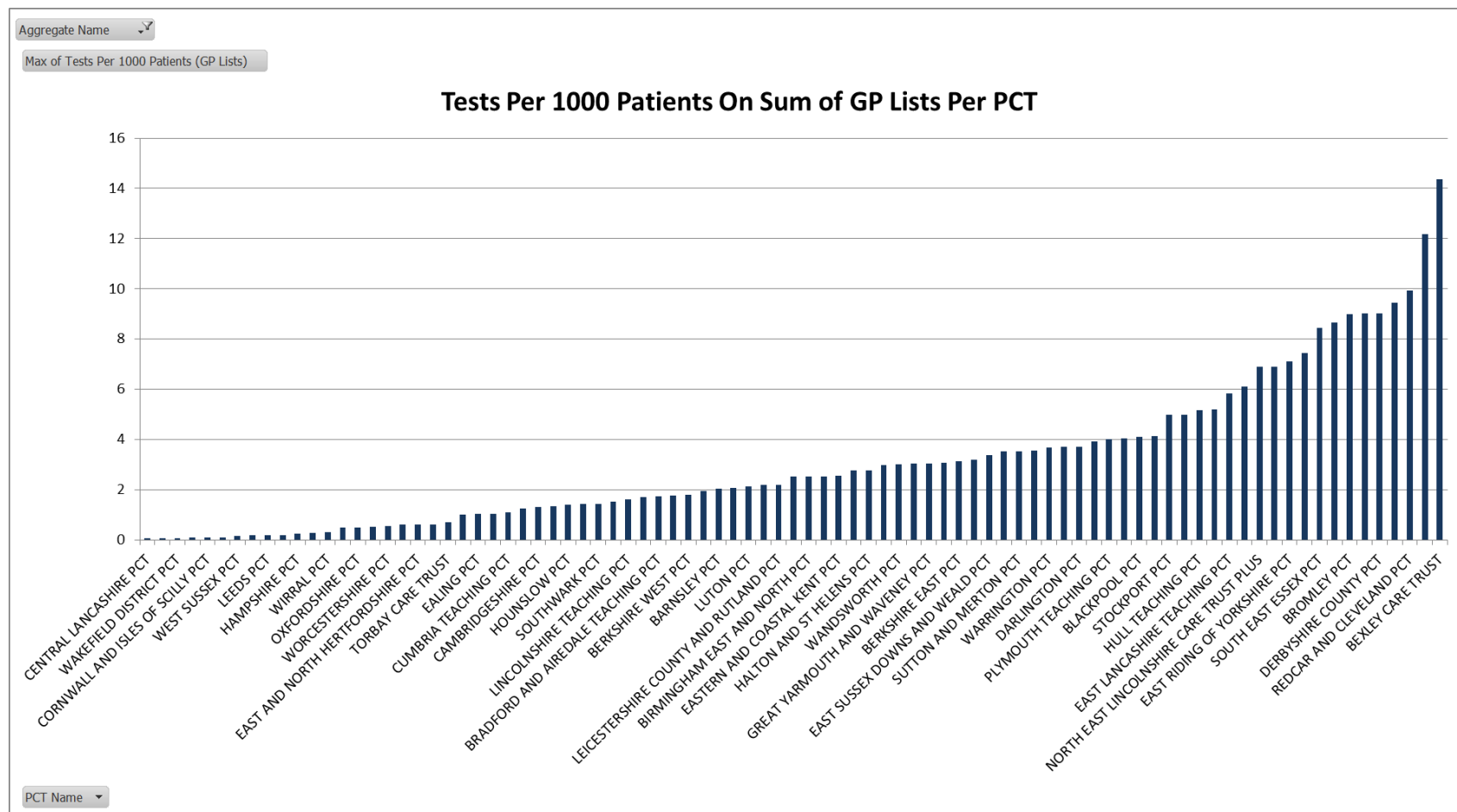
<http://www.ychi.leeds.ac.uk/pmipunits>

Brain Naturetic Peptide (CCF)

Lack of compliance with NICE



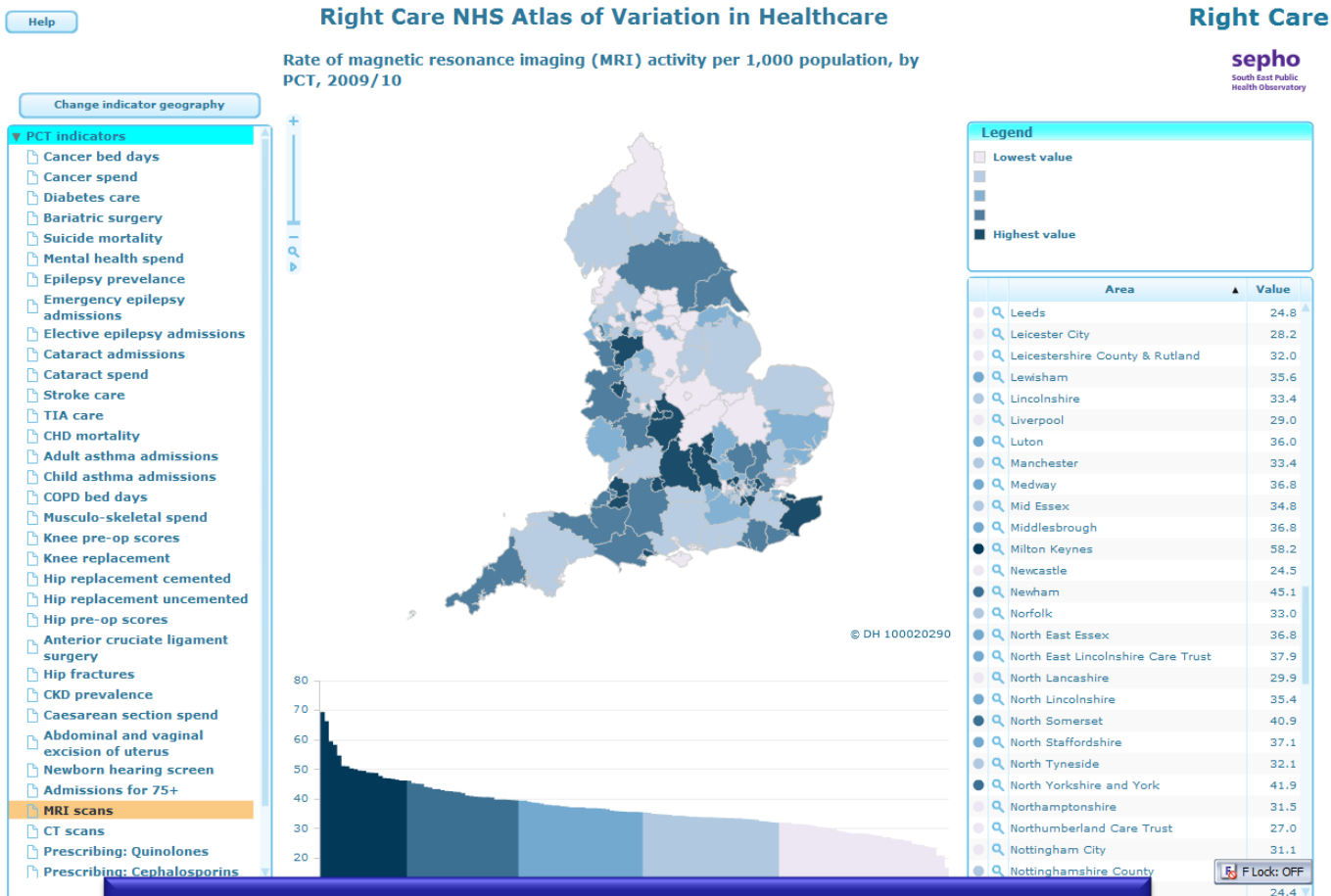
UNIVERSITY OF LEEDS



GIS Display – live and online! Diagnostic Atlas – Q3 2012



UNIVERSITY OF LEEDS



<http://www.rightcare.nhs.uk/index.php/nhs-atlas/>

- Move to real-time continuous capture
 - Safety – compliance with standards
 - Harmonisation of practice
 - Interoperability
- Interest in:
 - Feeding commissioning work flows
 - Providing data feeds to Cancer Registry – section 251 already cleared for such data
 - Early warning reports on outbreaks – HTA
 - Patient accessible data – diabetes, thyroid, LTCs
 - Feeding CPRD – part of current MRC eHIRC potential



The White Rose Consortium

NHS
Connecting for Health

- 3 leading Schools of Computing
- 3 Medical Schools
- A Major NHS research-led hospital
- A nationally recognised e-health SME
- The NHS Information Centre



The
Information
Centre

for health and social care



THE UNIVERSITY *of York*

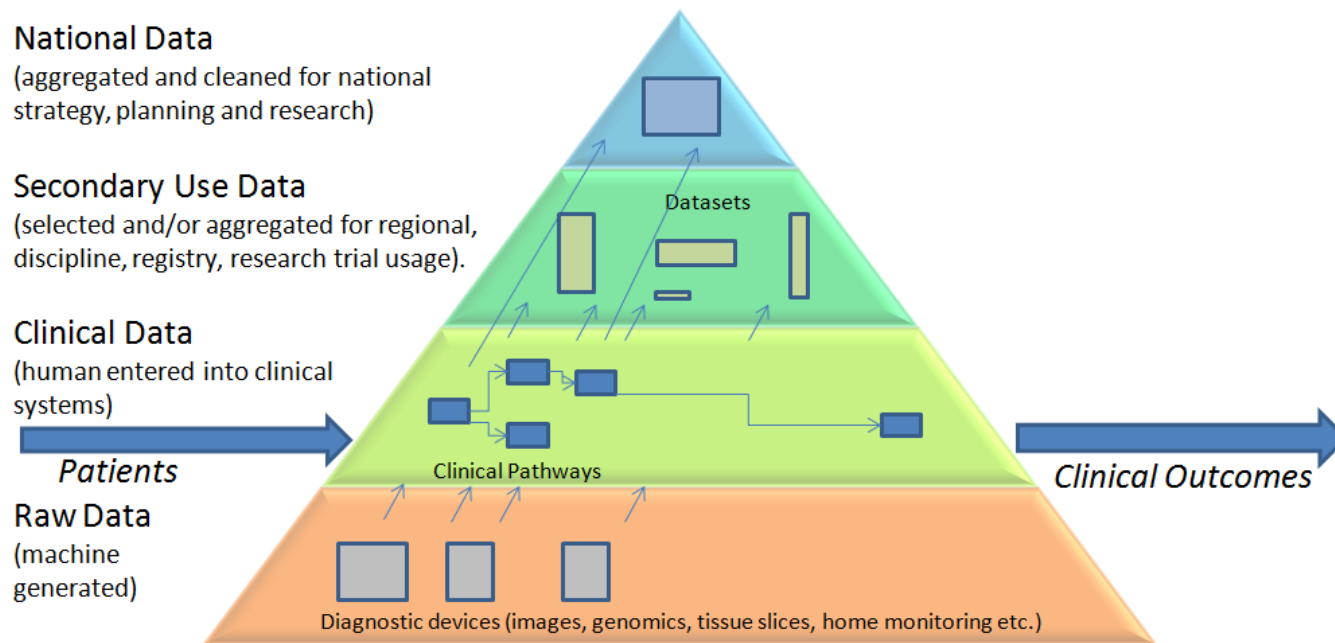


UNIVERSITY OF LEEDS



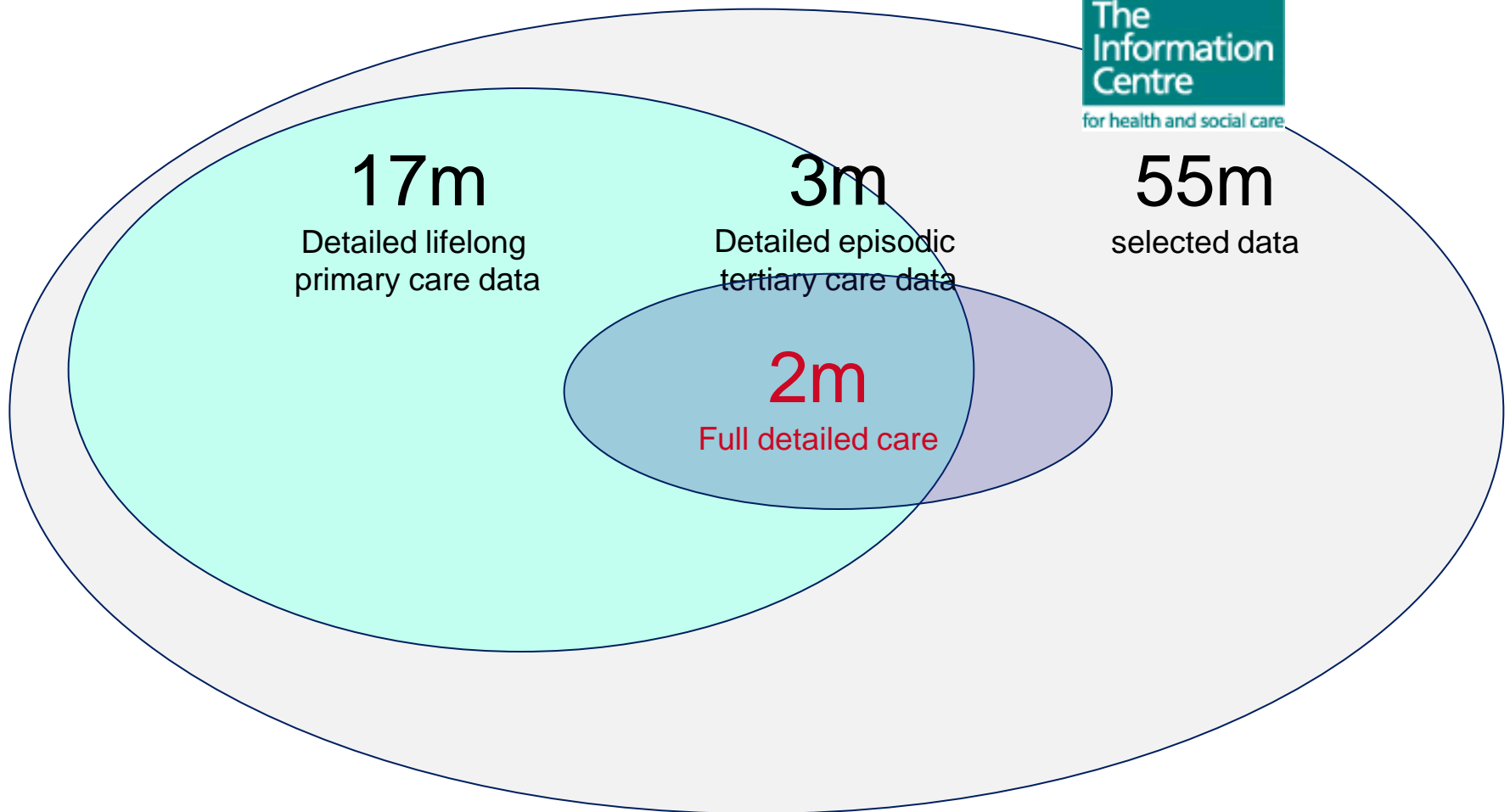
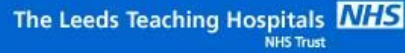
Electronic Point of Care (EPOC) paradigm in e-Health research

Extending the principles of supply chain automation and electronic point of sale (EPOS). Each interaction or transaction with a customer (patient) should be recorded once, correctly and electronically but the complexity of data capture should be hidden from the end user by presenting simple, easy-to-use interfaces. Data should flow both up and down the value chain (care pathway) to support local management and data should be automatically aggregated to inform tactical and strategic decision making - specifically supporting service design, public health and research. Underlying the model is data linkage, secure access, rigorous standards and the ability to work with massive, complex, human authored data sets.

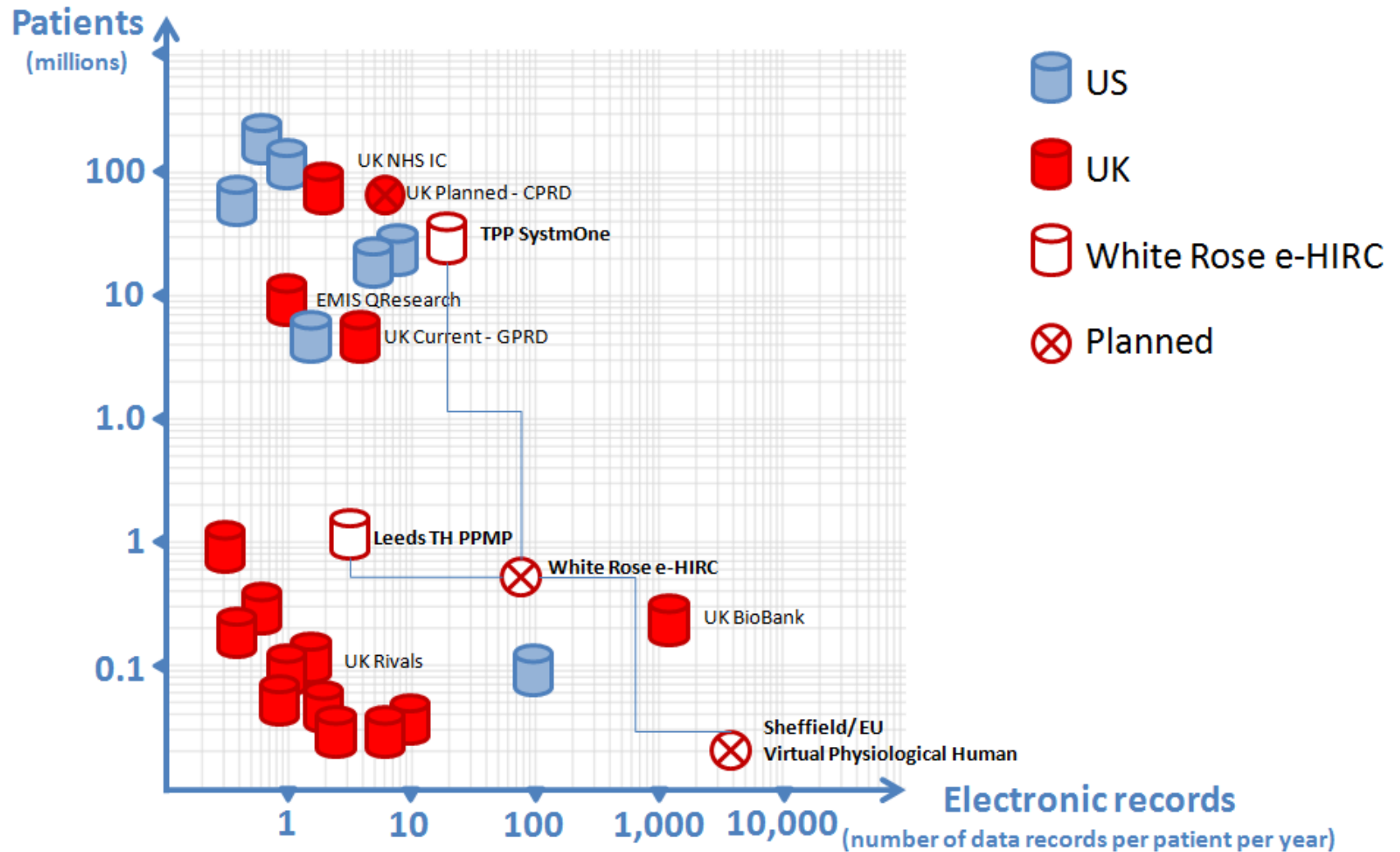


E-Health Records Linkage

Understanding the Scale



The world's largest EPR?



Source: Clamp, Johnson, Magare – *paper in development*

White Rose MRC e-HIRC Technical Architecture

PUBLIC ACCESS

- Via NHS IC
- Hi-Viz tooling (public)
- Aggregated views

NHS Information Centre

(Safe Haven – 50m+ patients, administrative and outcomes data)

JISC Grant won 2012 to build prototype very large-scale safe haven.

ResearchOne

(Safe Haven – 20m+ patients, lifelong primary care record)

?

(other sources added as required)

PPM2/R2

(Safe Haven – 0.9m+ patients, detailed episodes of hospital care, patient reported outcomes)

Enterprise Service Bus (Data Mash-ups)

ETHICAL RESEARCH ACCESS

- Via MRC e-HIRC hubs
- Hi-Viz tooling (research)
- Pseudonymised views
- CPRD

National Grid Service (NGS)

White Rose Grid

- Health data of sufficient volume and quality now exists nationally to reconsider the paradigms for research
- The new techniques will not supplant clinical trials but will enhance them
 - Reduced costs
 - Faster and better recruitment
 - Shorter times to completion
 - Efficient and extended outcome data capture
- New opportunities will arise demanding new techniques
 - Data / text mining
 - Statistical – modelling, missing data
 - Cluster randomisations
 - New techniques of benefits analysis for rapidly evolving technology

Yorkshire Centre for Health Informatics

Director Dr Susan Clamp

+44 (0)113 343 4960

s.clamp@leeds.ac.uk



www.ychi.leeds.ac.uk