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

DH
CfH
SHA
Trusts

UNIVERSITY OF LEEDS

YCHI

Computer Science
Design

Health

Diagnostics
Pharma
Informatics

NHS
Plus ça change…

Electronic Health Record

Supporting 24-hour care

Patient accessible

Routine patient care

Aggregated anonymised subsets

Developing Health Improvement Programmes

Clinical Governance

Epidemiological Research

YCHI - Core Activities

• 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?

www.ychi.leeds.ac.uk/eprresearch
Confirmation of Knowledge Gap

- HL7 Messaging: Not Heard Of, Heard Of But Not Used, Heard Of And Used
- NHS Care Records Service: Not Heard Of, Heard Of But Not Used, Heard Of And Used
- Patient Demographic Service: Not Heard Of, Heard Of But Not Used, Heard Of And Used
- SUS: Not Heard Of, Heard Of But Not Used, Heard Of And Used
- Legitimate Relationships: Not Heard Of, Heard Of But Not Used, Heard Of And Used
- Sealed Envelopes: Not Heard Of, Heard Of But Not Used, Heard Of And Used
Confirmation of Knowledge Gap

- Information Governance
- Pseudonymisation
- Normalisation
- Federated Databases
- SNOMED-CT
- Read Codes
- ICD-10
- LOINC
- RBAC

- Not Heard Of
- Heard Of But Not Used
- Heard Of And Used
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

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


Slides © Dr Geoff Hall, 2012
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
Currently upgrading to LTHT portal using agile / open source methods

Welcome to the new home of eHealthOpenSource

A Partnership Between...

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

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 SystmOne clinical software. SystmOne 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 SystmOne Site Deployments Per Module and SHA

<table>
<thead>
<tr>
<th>SystmOne Module</th>
<th>East Midlands</th>
<th>East of England</th>
<th>West Midlands</th>
<th>North East</th>
<th>North West</th>
<th>Yorkshire &amp; the Humber</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP</td>
<td>398</td>
<td>390</td>
<td>79</td>
<td>131</td>
<td>36</td>
<td>553</td>
<td>1587</td>
</tr>
<tr>
<td>Community#</td>
<td>9</td>
<td>22</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>21</td>
<td>65</td>
</tr>
<tr>
<td>Out of Hours</td>
<td>18</td>
<td>11</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>12</td>
<td>54</td>
</tr>
<tr>
<td>Minor Injury Unit</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Child Health#</td>
<td>10</td>
<td>19</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>18</td>
<td>58</td>
</tr>
<tr>
<td>Hospice</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>3</td>
<td>24</td>
<td>50</td>
</tr>
<tr>
<td>Prisons*</td>
<td>17</td>
<td>15</td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>78</td>
</tr>
</tbody>
</table>
Leeds: Exploring new paradigms

• 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
<table>
<thead>
<tr>
<th>ID</th>
<th>Tumor Location</th>
<th>Ovary</th>
<th>Date</th>
<th>Diagnosis</th>
<th>Grade</th>
<th>Diff</th>
<th>Diagnosis</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>100094107</td>
<td>Null</td>
<td>Ovary</td>
<td>September 13, 1936</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>False</td>
</tr>
<tr>
<td>100094471</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>March 16, 1934</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100094456</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>February 6, 1934</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096129</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>June 2, 1931</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>False</td>
</tr>
<tr>
<td>100096520</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>December 5, 1933</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096943</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>February 27, 1933</td>
<td>3</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100098233</td>
<td>Calderdale PCT</td>
<td>Ovary</td>
<td>October 2, 1945</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100099624</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>October 2, 1941</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096314</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>June 26, 1962</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096510</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>March 14, 1944</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Papillary serous cystadenocarcinoma</td>
<td>True</td>
</tr>
<tr>
<td>100096548</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>August 20, 1960</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100097306</td>
<td>Kirklees PCT</td>
<td>Ovary</td>
<td>July 7, 1935</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100097737</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>August 14, 1949</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100098231</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>June 10, 1960</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100098335</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>October 16, 1940</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100098438</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>June 27, 1955</td>
<td>2c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096666</td>
<td>Other PCT</td>
<td>Ovary</td>
<td>October 30, 1947</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Papillary serous cystadenocarcinoma</td>
<td>True</td>
</tr>
<tr>
<td>100096677</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>April 21, 1943</td>
<td>3a</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>False</td>
</tr>
<tr>
<td>100096541</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>December 19, 1941</td>
<td>3b</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096447</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>January 2, 1944</td>
<td>2b</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096515</td>
<td>York PCT</td>
<td>Ovary</td>
<td>December 19, 1940</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100096310</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>September 11, 1921</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010008</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>September 28, 1933</td>
<td>2c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010163</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>June 16, 1945</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010147</td>
<td>Kirklees PCT</td>
<td>Ovary</td>
<td>June 22, 1952</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010579</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>November 1, 1944</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>False</td>
</tr>
<tr>
<td>100010127</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>August 18, 1942</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Papillary serous cystadenocarcinoma</td>
<td>True</td>
</tr>
<tr>
<td>100010207</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>May 29, 1930</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010271</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>September 2, 1935</td>
<td>1a</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010262</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>March 19, 1931</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010290</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>March 22, 1929</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010370</td>
<td>York PCT</td>
<td>Ovary</td>
<td>June 13, 1931</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010524</td>
<td>Kirklees PCT</td>
<td>Ovary</td>
<td>January 13, 1936</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010328</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>January 22, 1938</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010102</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>December 19, 1939</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>False</td>
</tr>
<tr>
<td>100010119</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>September 30, 1928</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010450</td>
<td>York PCT</td>
<td>Ovary</td>
<td>December 9, 1938</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
<tr>
<td>100010451</td>
<td>Leeds PCT</td>
<td>Ovary</td>
<td>January 21, 1942</td>
<td>4</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Papillary serous cystadenocarcinoma</td>
<td>True</td>
</tr>
<tr>
<td>100010456</td>
<td>Wakefield PCT</td>
<td>Ovary</td>
<td>September 1, 1944</td>
<td>3c</td>
<td>3</td>
<td>High/Poor Diff</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>True</td>
</tr>
</tbody>
</table>
Clinical Outcomes
Ovarian - stage specific survival

<table>
<thead>
<tr>
<th>Stage</th>
<th>25.0% Estimate</th>
<th>25.0% Std. Error</th>
<th>50.0% Estimate</th>
<th>50.0% Std. Error</th>
<th>75.0% Estimate</th>
<th>75.0% Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>108.918</td>
<td>13.525</td>
<td>58.323</td>
<td>4.893</td>
</tr>
<tr>
<td>2</td>
<td>1.051.45</td>
<td>1.051.45</td>
<td>83.804</td>
<td>8.408</td>
<td>40.450</td>
<td>7.840</td>
</tr>
<tr>
<td>4</td>
<td>31.443</td>
<td>2.264</td>
<td>15.311</td>
<td>1.043</td>
<td>6.066</td>
<td>7.19</td>
</tr>
<tr>
<td>Overall</td>
<td>83.736</td>
<td>3.824</td>
<td>36.230</td>
<td>1.617</td>
<td>14.523</td>
<td>6.38</td>
</tr>
</tbody>
</table>

Slides © Dr Geoff Hall, 2012
Clinical Outcomes
Endometrial - morphology specific survival

- 919 cases
  - 813 endometrioid
  - 106 serous
- Median overall survival
  - Overall = 66. months
  - Endometrioid > 6 years
  - Serous = 28.1 months
New paradigms: Identification of advanced disease

- 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
## Identification of advanced disease: PPM Event Profiling

<table>
<thead>
<tr>
<th>Date</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>15/12/2001</td>
<td>Stage 3C, G2, Serous Cystadenocarcinoma of the Peritoneum</td>
</tr>
<tr>
<td>10/12/2004</td>
<td>Progression of Serous Cystadenocarcinoma of the Peritoneum</td>
</tr>
<tr>
<td>31/01/2007</td>
<td>Progression of Serous Cystadenocarcinoma of the Peritoneum</td>
</tr>
<tr>
<td>04/02/2008</td>
<td>Progression of Serous Cystadenocarcinoma of the Peritoneum</td>
</tr>
<tr>
<td>02/10/2008</td>
<td>Progression of Serous Cystadenocarcinoma of the Peritoneum</td>
</tr>
<tr>
<td>06/07/2010</td>
<td>Distant Metastases of Serous Cystadenocarcinoma of the Peritoneum, Mets: Pleura</td>
</tr>
<tr>
<td>12/11/2010</td>
<td>Distant Metastases of Serous Cystadenocarcinoma of the Peritoneum, Mets: Pulmonary, Pleura</td>
</tr>
<tr>
<td>07/04/2011</td>
<td>Distant Metastases of Serous Cystadenocarcinoma of the Peritoneum, Mets: Hepatic, Pulmonary, Pleura</td>
</tr>
<tr>
<td>25/07/2011</td>
<td>Progression of Serous Cystadenocarcinoma of the Peritoneum, Mets: Hepatic, Pulmonary, Pleura</td>
</tr>
<tr>
<td>16/10/2011</td>
<td>Distant Metastases of Serous Cystadenocarcinoma of the Peritoneum, Mets: Hepatic, Pulmonary, Pleura, Brain</td>
</tr>
</tbody>
</table>
Exploiting Unstructured Data
Natural language processing

Unstructured EPR
- Rich in narrative / potentially insecure / hard to search & analyse

Structured / Coded Database
- Searchable / Secure / Easy to analyse formally

Clinical Database (PPM)

Most information on progression is in the narrative account

Supports Clinical “Business as usual”

Supports Enriched Research Datasets

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

Slides © Dr Geoff Hall, 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.”
• Developed protocol for an ethical research database for ethical research
• TPP ResearchOne – launched May 2012
• Ethical approval with NIGB

Hospital Episode Statistics (HES)
- 92% SUS-sample patients have a SystmOne record
- 68% SystmOne 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)

- 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.
Results: Yields of Records

Subjects in TPP database

- 81% Cardiovascular
- 66% Oncology
Results: Data Quality

Frequency of drugs for cardiology patients

Frequency of diary entries for cardiology patients
New opportunities  
TSB KTP with Philips Healthcare

**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

- 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
Modelling neutropenic responses

Progressive decline in haemopoietic 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

All GP pathology messages is 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.
Pathology Messaging Datasets

- **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
Test volumes across the NHS

Patient Reports & Tests per Day

- Patients
- Reports
- Tests
• 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

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

http://www.ychi.leeds.ac.uk/pmipunits
CA125 (Ovarian cancer)

- 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
Right Care NHS Atlas of Variation in Healthcare

Rate of magnetic resonance imaging (MRI) activity per 1,000 population, by PCT, 2009/10

Legend
- Lowest value
- 2nd lowest value
- 2nd highest value
- Highest value

<table>
<thead>
<tr>
<th>Area</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leeds</td>
<td>24.0</td>
</tr>
<tr>
<td>Leicester City</td>
<td>28.2</td>
</tr>
<tr>
<td>Leicestershire County &amp; Rutland</td>
<td>32.0</td>
</tr>
<tr>
<td>Luton</td>
<td>35.5</td>
</tr>
<tr>
<td>Lincolnshire</td>
<td>33.4</td>
</tr>
<tr>
<td>Liverpool</td>
<td>25.0</td>
</tr>
<tr>
<td>Luton</td>
<td>36.0</td>
</tr>
<tr>
<td>Manchester</td>
<td>23.4</td>
</tr>
<tr>
<td>Monmouth</td>
<td>36.5</td>
</tr>
<tr>
<td>Mid Sussex</td>
<td>24.3</td>
</tr>
<tr>
<td>Middlesbrough</td>
<td>38.5</td>
</tr>
<tr>
<td>Milton Keynes</td>
<td>58.2</td>
</tr>
<tr>
<td>Newcastle</td>
<td>24.5</td>
</tr>
<tr>
<td>Newham</td>
<td>45.1</td>
</tr>
<tr>
<td>Rotherham</td>
<td>23.0</td>
</tr>
<tr>
<td>North East Essex</td>
<td>36.8</td>
</tr>
<tr>
<td>North East Lincolnshire Care Trust</td>
<td>37.9</td>
</tr>
<tr>
<td>North Lancashire</td>
<td>29.9</td>
</tr>
<tr>
<td>North Lincolnshire</td>
<td>35.4</td>
</tr>
<tr>
<td>North Somerset</td>
<td>40.5</td>
</tr>
<tr>
<td>North Staffordshire</td>
<td>27.1</td>
</tr>
<tr>
<td>North Tyneside</td>
<td>22.1</td>
</tr>
<tr>
<td>North Yorkshire and York</td>
<td>41.9</td>
</tr>
<tr>
<td>Northampton</td>
<td>31.5</td>
</tr>
<tr>
<td>Northumberland Care Trust</td>
<td>29.0</td>
</tr>
<tr>
<td>Nottingham City</td>
<td>31.1</td>
</tr>
<tr>
<td>Northumberland County</td>
<td>24.4</td>
</tr>
</tbody>
</table>

Future possibilities

- 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

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 University of York

UNIVERSITY OF LEEDS
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

17m
Detailed lifelong primary care data

3m
Detailed episodic tertiary care data

55m
selected data

2m
Full detailed care
The world’s largest EPR?

Source: Clamp, Johnson, Magare – *paper in development*
PUBLIC ACCESS
- Via NHS IC
- Hi-Viz tooling (public)
- Aggregated views

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

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

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

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

PPM2/R2
(Safe Haven – 0.9m+ patients, detailed episodes of hospital care, patient reported outcomes)

Enterprise Service Bus (Data Mash-ups)

National Grid Service (NGS)

White Rose Grid
Conclusions

• 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