Using routinely collected clinical data to support Clinical Trials: a view from Scotland

Professor Colin McCowan
Robertson Centre for Biostatistics and Glasgow Clinical Trials Unit
From birth to death...

Health Care Data - Hospital, Primary Care, Laboratory, Imaging, Prescribing

Potential data for linkage:
- 1939 register - occupation
- Pre & neo-natal exams
- Health Visitor
- Scottish Mental Survey 1932 & 1947 - IQ
- Registration of death

Social Care Data for older people

Existing collections:
- Neo-natal & Maternal Databank
- Growing Up In Scotland
- Scottish Longitudinal Study - Census - employment, household, chronic illness etc., Education, Hospital episodes, Vital events - births, deaths and marriages

Disease Registries for:
- Multiple sclerosis
- Asthma
- Chronic kidney disease
- Parkinsons
- Arthritis
Electronic Data Research and Innovation Service (eDRIS)

Joined up data for better decisions: Guiding Principles for Data Linkage
http://www.scotland.gov.uk/Publications/2012/11/9015
National Data Catalogue
National Datasets

You are: Home | National Datasets

National Dataset Menu
- National Datasets Home
- Audits / Registers
- Cancer
- Child Health
- Costs
- Diseases
- Hospital Activity Statistics
- Lifestyle
- NRS - Population
- Other Datasets
- Screening
- Waiting Times
- Workforce
- Datasets A-Z

National Datasets

ISD holds NHS health and health related data for over 5 million people in Scotland, which in some cases span an individual's whole life: from before birth, with the mothers antenatal records, through to that individual's death registration record. This is a wealth of information which can be linked, summarised and analysed to support research studies.

The NDC details the full list of datasets that are held by ISD. Further information on each of the datasets is available by selecting the dataset name using either the A-Z or the dataset groupings in the left hand bar.

(please note that due to the nature of the groupings, some datasets may appear under more than one topic area)

An interactive map (link at the top right of screen) will also aid users to navigate the various information that is available for all of the datasets listed.

http://www.ndc.scot.nhs.uk/
<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vital statistics</td>
<td>Birth certificates</td>
</tr>
<tr>
<td></td>
<td>Marriage certificates</td>
</tr>
<tr>
<td></td>
<td>Death certificates</td>
</tr>
<tr>
<td>Administrative data</td>
<td>Outpatient attendances (SMR00)</td>
</tr>
<tr>
<td></td>
<td>Hospital admissions: acute (SMR01), psychiatric (SMR04)</td>
</tr>
<tr>
<td></td>
<td>Maternity (SMR02)</td>
</tr>
<tr>
<td></td>
<td>Primary care</td>
</tr>
<tr>
<td></td>
<td>Prescribing Information System (PIS)</td>
</tr>
<tr>
<td></td>
<td>Scottish Immunisation Recall System (SIRS)</td>
</tr>
<tr>
<td></td>
<td>Child Health Systems Programme (CHSP): preschool, school</td>
</tr>
<tr>
<td></td>
<td>Support Needs System</td>
</tr>
<tr>
<td>Disease registries</td>
<td>Cancer registrations (SMR06)</td>
</tr>
<tr>
<td></td>
<td>Bowel screening system (BoSS)</td>
</tr>
<tr>
<td></td>
<td>Scottish Renal Registry (SRR)</td>
</tr>
<tr>
<td></td>
<td>Scottish Stillbirth &amp; Infant Death Enquiry (SSIDE)</td>
</tr>
<tr>
<td></td>
<td>Scottish Stroke Care Audit</td>
</tr>
<tr>
<td></td>
<td>SCI-CHD</td>
</tr>
<tr>
<td></td>
<td>SCI-Diabetes</td>
</tr>
<tr>
<td>Population surveys</td>
<td>Scottish Health Surveys (SHS)</td>
</tr>
<tr>
<td></td>
<td>Scottish Schools Adolescent Lifestyle and Substance Use Survey (SALSUS)</td>
</tr>
<tr>
<td></td>
<td>Growing Up in Scotland (GUS)</td>
</tr>
<tr>
<td></td>
<td>Health Behaviour in Scottish Children (HBSC)</td>
</tr>
<tr>
<td></td>
<td>Scottish Longitudinal Study (SLS)</td>
</tr>
</tbody>
</table>
- Collaboration with Greater Glasgow & Clyde Health Board
- Develop a virtual EHR for research for population of ~1.3 million
- Develop governance systems to allow access to data for approved projects for approved researchers
- Promote clinical research using routine data and facilitate research through collaboration with RCB statistics and IT group
- Develop national and international collaborations using Glasgow Safe Haven data
Current Datasets

• Datasets in Safe Haven
  – SMR00 - Outpatient Attendance
  – SMR01 – Acute inpatient & Day Care
  – SMR02 – Maternity
  – SMR04 – Mental Health
  - CHI – GG&C patient population (1.3 million)
  - GRO – Births and deaths date for GG&C
  - ePrescribing – encashed prescriptions for Glasgow
  - GP (LES and Keep Well) 250 practices

• REC approval is to submit an amendment every time 6 new new databases are added

• In discussion to extend health data to other Boards in NRS West – Lanarkshire, A&A, D&G and Golden Jubilee

- Heart Failure – locally held national Heart Failure database
- Rheumatology – local clinical database
- SCI DC – GGC population of national Diabetic database
- SCI Store results for GGC
- Parkinson – local clinical database for Movement disorders
- Weight Management
- PsyCIS – schizophrenia database
- Clozapine database
SHIP is an ambitious, Scotland-wide research platform for the collation, management, dissemination and analysis of anonymised Electronic Patient Records (EPRs). Find out more.
Creating safe settings and data security

• SHIP Blueprint highlights:

Safe Projects
Safe People
Safe Data
Safe Settings
Safe Outputs
• **Safe Projects**
  – Ethics approved
  – Caldicott Guardian approved
  – Privacy Advisory committee approved
  – Peer reviewed

• **Safe People**
  – Accredited organisation
  – Accredited researchers
  – Information Governance Training
Governance approvals from clinical dataset to research extract

- NHS Clinical dataset
- Safe Haven dataset
- Approved research extract
- Caldicott / LPAC approval
- Data Access Committee / Data Controller approval
Thomas and Walport Data Sharing Review (2008):

“environments for population based research and statistical analysis in which the risk of identifying individuals is minimised”
• Disclosure
  – Disclosure occurs when information released reveals the identity or information about a respondent not already in the public domain.

• Statistical Disclosure Control
  – Assessing the risk of disclosure from a potential release and taking measures, if appropriate, to lower that risk by modifying the data, changing the design of the statistic, or a combination of these.

(Office of National Statistics, 2006)
Research Using Routinely Collected Data

- **Epidemiology**
  - Identify patients with specific disease or taking a specific treatment
  - Use hospitalisations and death data to build-up clinical histories and investigate outcomes

- **Support other research methods**
  - Cross-sectoral linkage
  - Act as sampling frame for qualitative studies

- **Support clinical trials**
  - Support feasibility studies and patient recruitment
  - Passive follow-up of subjects: cheaper, maximises follow-up
Women with breast cancer using endocrine therapy were missing tablets.

Study: Linking prescribing records to hospital discharge, cancer registry, cancer clinic data and mortality records.

Results: Low adherence was associated with poorer survival (HR=1.20; 95% CI=1.03–1.40, P=0.019).
• Lower uptake of bowel screening in women than for breast or cervical screening

• Study: Link screening records and examine difference in characteristics for uptake. *Invite women for interview to explain why they use other cancer screening services but not bowel*

National Awareness and Early Diagnosis Initiative (NAEDI)
• Landmark trial of statin use in patients with no previous CHD

• Study: Used hospital discharge and mortality records to identify endpoints for adjudication

• Results: 31% relative risk reduction in nonfatal myocardial infarction or death from coronary heart disease

• Cost £20m

WOSCOPS follow-up

- Fifteen year follow-up using record linkage

- Study: Use record linkage of trial data to hospital discharge, cancer registry, and mortality records

CHD death or CHD event

- Results:

- Cost: £15k

WOSCOPS follow-up

- Twenty year follow-up using record linkage
### Adjudicated endpoints vs record linkage

<table>
<thead>
<tr>
<th>WOSCOPS cause of death</th>
<th>Record Linkage cause of death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CHD</td>
</tr>
<tr>
<td>Coronary Heart Disease</td>
<td>96</td>
</tr>
<tr>
<td>Other cardiac</td>
<td>0</td>
</tr>
<tr>
<td>Other vascular</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
</tr>
</tbody>
</table>

- **PLOS One: Barry et al (2013)**
### Adjudicated endpoints vs record linkage

<table>
<thead>
<tr>
<th>Record Linkage events</th>
<th>First WOSCOPS MI event</th>
<th>Subsequent WOSCOPS event</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Myocardial Infarction</td>
<td>217 (81.0%)</td>
<td>4 (11.1%)</td>
<td>221 (72.7%)</td>
</tr>
<tr>
<td>Subsequent MI</td>
<td>1 (0.4%)</td>
<td>17 (47.2%)</td>
<td>18 (5.9%)</td>
</tr>
<tr>
<td>Coronary Heart Disease</td>
<td>22 (8.2%)</td>
<td>12 (33.3%)</td>
<td>34 (11.2%)</td>
</tr>
<tr>
<td>Other Cardiovascular Disease</td>
<td>3 (1.1%)</td>
<td>0 (0.0%)</td>
<td>3 (1.0%)</td>
</tr>
<tr>
<td>Non-cardiac Chest Pain</td>
<td>15 (5.6%)</td>
<td>2 (5.6%)</td>
<td>17 (5.6%)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (0.7%)</td>
<td>0 (0.0%)</td>
<td>2 (0.6%)</td>
</tr>
<tr>
<td>Non-Match</td>
<td>8 (3.0%)</td>
<td>1 (2.8%)</td>
<td>9 (3.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>268</td>
<td>36</td>
<td>304</td>
</tr>
</tbody>
</table>

### PLOS One: Barry et al (2013)
The EHR4CR Project
Electronic Health Records for Clinical Research
EHR4CR brings together multiple stakeholder groups to achieve something that individual groups cannot realise alone.

35 participants

- 11 hospital sites (data providers)
- 10 Pharma Companies (data users)
- 14 other: Academia, small and medium-sized enterprises, patient associations, public authorities

Electronic Health Records for Clinical Research

UKCRC Registered Clinical Trials Units
Problems with clinical trials

- The percentage of studies that complete enrolment on time: 18% in Europe, 7% in the US
- 1/3 of protocol amendments are avoidable, at a cost of $0.5m
- Almost 50% of all trial delays caused by patient recruitment problems
- Each day a drug is delayed from market, sponsors lose up to $8m
- 50% of today's clinical trials fail to achieve the target recruitment

Potential of Real World Data

NHS
UKCRC Registered Clinical Trials Units

Greater Glasgow and Clyde
EHR4CR services will remove bottlenecks and drive a paradigm shift in the way we conduct clinical research in Europe

By unlocking the potential of information stored in EHRs to
- Optimise protocol design
- Accelerate patient recruitment
- Streamline trial execution by reducing redundant data entry and data transcription errors

Clinical research bottlenecks in figures

34% of protocol amendments are avoidable
43% of protocol amendments take place before first patient first dose
27% of suitable trial candidates are missed by a manual search process compared with electronic
50% of today’s clinical trials fail to achieve the target recruitment rate
70% of clinical trial data are estimated by investigational sites to be duplicated between EHR and clinical trial systems
This will happen within a sustainable and multi-faceted environment for trustworthy re-use of health data.

**Pre-consent**
- **Consolidation**: Only aggregated patient numbers leave the hospital prior to consent. Fuzzing protects low numbers.
- **De-identification**: Individual patient is anonymous.

**Pre- and Post-Consent**
- **Segregation**: A one-way transfer of limited EHR data is separated from EHR.
- **Security**: Leading edge security features meet ethical, privacy, regulatory and legal standards.
- **Site Control**: Access to data and link to patient identifiers under exclusive site control.

**Post-consent**
- **Governance**: Institute provides independent oversight and control to ensure data are accessed in a trustworthy way.
- **Investigator Control**: Investigator selects and approves anonymised data to be included in eCRF.

---

**Electronic Health Records for Clinical Research**

---

**NHS** 
**British Columbia**