

# Supporting trials and evaluative research through data linkage in Wales

Professor Ronan Lyons

Centre for Health Information, Research and Evaluation (CHIRAL)

Institute of Life Science

School of Medicine

Swansea University

# Health Information Research Unit (HIRU)

---

- Core funded by NISCHR
- Additional funding from many bodies for project work
- Created the Secure Anonymised Information Linkage (SAIL) system

## The essence of HIRU's work . .

---

**Linking anonymised data, at individual and ecological levels, across multiple datasets,** drawn from operational systems in health services, national databases, clinical datasets and databases in social care, education, housing, etc.

# Anonymised Data Linkage

---

- Use split files and different organisations to ensure confidentiality is protected
- Use an NHS body to link identities to unique numbers
- Use state of art encryption to produce and protect anonymity
- NHS linkage body knows nothing about data content
- HIRU knows nothing about identity
- Can anonymise both at individual and household levels
- Data do not leave SAIL unless informed consent
- Can provide data to cohorts and trials

# Some datasets; individual and ecological

---

## Individual level - NHS:

Population (NHSAR)

Inpatients -PEDW (HES)

Births

Deaths

Outpatients

ED

GP Data

Laboratory systems

Out of Hours Services

Child Health Database Wales

NHS Direct Wales

Radiology- Imaging

## Individual level – non-NHS:

Social Services

Educational Attainment

## Clinically rich databases:

Cancer

Screening (multiple conditions)

Congenital Anomalies

Diabetes

Arthropathies

Etc.

## Ecological datasets (many are GIS):

Census - small areas

Ordnance Survey - Mastermap

Environmental Health

# Types of research supported by data linkage

---

1. Observational Epidemiology/HSR
2. Pure electronic, hybrid and enhanced cohorts
3. **Clinical and Public Health Trials**
4. Evaluation of Policy Initiatives/ Natural experiments

# Examples of project funding involving HIRU

---

- DECIPHer - UK Public Health Centre of Excellence (UKCRC)
- WISERD (ESRC/HEFCW)
- Free School breakfast Initiative: data augmentation (NPRI)
- Ankylosing Spondylitis Cohort (MRC)
- INTEGRIS – linking A+E and inpatient data in Europe (EU, FP7)
- ALSPAC PEARL Study (Wellcome Trust)
- MS Register UK pilot (MS Society)
- WECC - Wales Electronic Cohort for Children (NISCHR)
- SAFER1, SAFER2, CONSTRUCT trials (HTA)

# E-Health and Trials

---

E-Health can contribute to a variety of aspects of trial design, conduct, and efficiency

- Feasibility
- Recruitment
- Outcomes

# Trial Feasibility (1)

---

## Factitious diabetes trial (Jeff Stephens)

- Drug Z vs therapy with Metformin, Acarbose, Glibenclamide or Rosiglitazone in drug-naïve patients with type 2 DM
- 250,086 individuals (Swansea)
  - 10,205 with type2 DM
  - 711 potential participants

Brooks CJ, Stephens JW, Price DE, Ford DV, **Lyons RA**, Prior SL, Bain SC. Use of a patient linked data warehouse to facilitate diabetes trial recruitment from primary care Primary Care Diabetes 2009 online publication

<http://dx.doi.org/10.1016/j.pcd.2009.06.004>

# Trial Feasibility (2): utility of GP data

---

- Not all pathology data are in GP records
- Comparison of
  - GP diagnosis only
  - GP diagnosis and GP pathology
  - GP diagnosis and Hospital pathology
- GP Dx/Path has sensitivity of 80-87% and false +ve rate of 19-25%

Brooks CJ, Stephens JW, Price DE, Ford DV, Lyons RA, Prior SL, Bain SC.

Use of a patient linked data warehouse to facilitate diabetes trial recruitment from primary care

Primary Care Diabetes 2009 online publication

<http://dx.doi.org/10.1016/j.pcd.2009.06.004>

# Trial Feasibility (3): HITE

---

Health informatics Trial Enhancement Project (HITE)

- Using routine primary care data to attempt to identify potential participants for a depression trial
- Construction of an algorithm with validation against expert opinion
- Sensitivity and specificity both  $> 0.9$
- Proof of concept demonstrated

Mcgregor J, Brooks C, Chalasani P, Chukwuma J, Hutchings H, Lyons RA, Lloyd K. The Health Informatics Trial Enhancement Project (HITE): Using routinely collected primary care data to identify potential participants for a depression trial.

Trials 2010, 11:39. doi:10.1186/1745-6215-11-39 <http://www.trialsjournal.com/content/pdf/1745-6215-11-39.pdf>

# More reliable outcome measures?

---

Two similar large scale (n>3,000) Vitamin D fracture prevention trials in similar populations from residential homes, but very different fracture rates/power

▪ Lyons et al: 7.4/100

▪ Law et al: 3.9/100

▪ Lyons RA, Johansen A, Brophy S, Newcombe RG, Phillips CJ, Lervy B et al. Preventing fractures among older people living in institutional care: a pragmatic randomised double blind placebo controlled trial of vitamin D supplementation. *Osteoporos Int* 2007 June;18(6):811-8.

▪ Law M, Withers H, Morris J, Anderson F. Vitamin D supplementation and the prevention of fractures and falls: results of a randomised trial in elderly people in residential accommodation. *Age Ageing* 2006 September;35(5):482-6.

# Unusual and longer term outcomes

---

- Free School Breakfast Initiative Data Augmentation and Analysis (NPRI funded)
- Originally a cluster RCT of breakfast vs not
- Now more, including relationship between breakfast and school performance
- 5750/5758 (99.9%) cohort matched to NHS numbers and subsequently to education data (97.3%)

# Wales Electronic Cohort for Children (WECC)

---

- Anonymised data from 730,000 children
- Demonstration project and platform for translating information into child health population policy:
  - What factors determine the future health service need for individuals that are vulnerable at birth?
  - What is the influence of the social and physical environment on childhood obesity?

# UK Biobank longitudinal follow up

---

- Pilot study of retrospective data on 15,133 participants
  - 9,694 with admissions
  - 11,055 with outpatient appointments
  - 1,259 cancer registrations
  - 21,500 breast screens
  - 39,960 cervical screens
- SAIL system will be used prospectively with additional data sources

# Residential Anonymised Linkage Fields (RALFs)

---

Objective: to support household level observational and interventional studies

- Central address/RALF table held securely within HSW
  - Assign each ALF to a RALF,
  - Track movements of ALFs between RALFs over time (never knowing who or where)
- Conduct research:
  - Household exposures/interventions and health (e.g. air quality, heating, etc)
  - Model infection disease spreads through household contacts
  - Link survey data with environmental exposure and health outcomes

# SAFER1 Trial

---

- Support and Assessment for Fall Emergency Referrals
- Pragmatic cluster RCT of computer decision support technology for on-scene decision making in ambulance service (Wales and England)
- Outcomes: further falls, **deaths, ED attendance, admissions**, QoL..
- difficulties with ability to consent (opt out) and follow up data

Snooks S, Cheung W-Y, Close J, Dale J, Gaze S, Humphreys I, **Lyons R**, Mason S, Merali Y, Peconi J, Phillips C, Phillips J, Roberts S, Russell I, Sánchez A, Wani M, Wells B, Whitfield R. Computerised on-scene decision support for emergency ambulance staff to assess and plan care for older people who have fallen: evaluation of costs and benefits using pragmatic cluster randomised trial: Support and Assessment for Fall Emergency Referrals (SAFER 1) trial protocol

BMC Emergency Medicine 2010;10:2. doi:10.1186/1471-227X-10-2.

<http://www.biomedcentral.com/1471-227X/10/2>

overview



Clinical and cost effectiveness of infliximab and ciclosporin in acute severe colitis

Commissioned by the HTA Programme

Pragmatic RCT, using mixed methods

40 hospitals; 240 patients in each arm

Primary outcome HRQoL @ 3, 6, 12 & 24 mo

Secondary outcomes **clinical & administrative**

Health economics

Patient views

Email: [construct@swansea.ac.uk](mailto:construct@swansea.ac.uk)

# Summary

---

- SAIL data linkage system working well for retrospective and prospective routine data
- Task is the same whether cohort or trial involved
- Ability to link to environmental data and movements (through anonymised household data adds additional functionality for public health focussed trials
- Task of defining variables is quite resource intensive
- Growth in GP and Pathology data will be particularly beneficial

# Relevant methodology papers

---

- Ford DV, Jones KH, Verplancke J-P, Lyons RA, John G, Brown G, Brooks C, Bodger O, Couch T and Leake K. **The SAIL Databank: building a national architecture for e-health research and evaluation.** *BMC Health Services Research* 2009, 9:157 (4 September 2009)
- Lyons RA, Jones KH, John G, Brooks CJ, Verplancke J-P, Ford DV, Brown G and Leake K. **The SAIL databank: linking multiple health and social care datasets.** *BMC Medical Informatics and Decision Making* 2009, 9:3 (16 January 2009)
- Rodgers SE, Lyons RA, Dsilva R, Jones KH, Brooks CJ, Ford DV, John G and Verplancke J-P. **Residential Anonymous Linking Fields (RALFs): A Novel Information Infrastructure to Study the Interaction between the Environment and Individuals' Health.** *Journal of Public Health*, 2009, pp. 1-7.