MRC e-Science featuring PsyGrid

www.psygrid.org
John Ainsworth, University of Manchester
john.ainsworth@manchester.ac.uk
MRC e-Science Projects

- NeuroGrid
  - Image storage, retrieval and analysis
- CLEF
  - Extraction, visualisation of routine medical data
- VOTES
  - Virtual Organisations for Trials and Epidemiological Studies
- CancerGrid
  - Clinical trial design and meta-data management
- … and PsyGrid
PsyGrid Aims

- **Aim 1 (Technological):** Develop applications and middleware to support studies and clinical trials in mental health
  - Software
  - A supported operational system

- **Aim 2 (Clinical):** Use outputs of ‘Aim 1’ to assemble longitudinal cohort study of patients with First Episode Psychosis over 2 years and study epidemiology of schizophrenia
UK Mental Health Research Network

North West Hub
led by University of Manchester & University of Liverpool

Heart of England Hub
led by University of Birmingham

West Hub
Universities of Bristol, Bath, West of England, and Peninsula Medical School

North East Hub
Universities of Durham & Newcastle

Nottingham & Lincolnshire Hub
Universities of Nottingham & Lincoln

Cambridgeshire and Norfolk Hub
Universities of Cambridge and East Anglia

North London Hub
Imperial College & University College London

South London Hub
led by Institute of Psychiatry and South London & Maudsley NHS Trust

Mental Health Research Network
PsyGrid clinical science progress: first episode psychosis recruitment
Informatics Initial Focus

- Development of a stable, reliable data collection system for longitudinal studies and clinical trials
- Meet the immediate needs of the clinical team
- Operational and supported
- Data quality and usability
- Create a solid foundation on which to build
Design Principles

- **Generic**
  - Nothing specific to psychiatry except configuration

- **Modular**
  - Reuse all or part of the system

- **Secure**
  - Communications, access
  - What you see is what you can access (WYSIWYCA) paradigm

- **Reliable**
  - No single point of failure

- **Customisable**
  - Per project security, dataset, by clinical researchers

- **Easy to use**
  - Uncluttered, intuitive user interface

- **Open Source**
Components

- **Security** - PsyGrid employs a role-based access control system based around the SAML standard. All communications are secure using PKI. Services are provided for authentication, privilege management and policy enforcement.

- **Data Repository** - a central database for storing clinical information, uses the Hibernate ORM layer and can be used with most RDBMS systems.

- **Electronic Screening Log (ESL)** - a separate database used for storing patient identifiable information, which can be configured with an independent security policy.

- **CoCoA** - a Java application interprets the data set and renders the data collection interface to the user.
Components

- ESL Web Interface - provides browser based access to the ESL for looking up patient study numbers from identifiable details or vice-versa. Only provides access to the results of randomisation for non-blind users.
- Randomisation - a service which implements a configurable multi-site, stratified, permuted random block length randomiser.
- Data Management Web Interface - provides browser based access to data export, reporting and document viewer functions.
- Security Management - a graphical tool which simplifies user account and privilege management
Features

- Off-line and on-line data entry
- Email notification of follow-ups due
- PKI secured with role-based access control
- Trend analysis reports
- Support for multiple independent projects
- Data anonymisation on import or export
- Model driven data set definition
- Data import from file or URL
- Data entry forms dynamically rendered from data set definition
- Open Source

- Data set versioning, publication and resynchronisation
- Data review and approve capability
- Recruitment reports
- Multi-centre remote data entry
- Audit trail and data provenance
- NHS N3 compatible
- Call out to 3rd party application data processors
- UKCRN accrual reports
- Conversion of documents to PDF format for printing
- Fully customisable data set definition including data elements, validation rules and scheduling
Perfect timing

- **New models of scientific cooperation promoted:**
  - Clinical science ↔ informatics ↔ computer science
  - Multi-centre national and international collaborations

- **Input into new NHS R&D informatics programme**
  - UK CRN; NIHR

- **New methods of clinical research data archiving**
  - UK CRN portfolio tracking
  - MRC Data Curation initiative

- **Leveraging further research support**
  - MRC; DH programme grants; FP7; US NIH
Directions

Data Collection Core

Modelling, Prediction, Decision Support

E-Epidemiology

Clinical Trials Platform
E-epidemiology

Scientist builds on previous data and findings to identify new research questions.

Identify existing assessment instruments.
Create new instruments.
Randomisation scheme.
Study protocol.
Identify research uses.
Privacy and confidentiality issues.
Submit to MREC/COREC for approval.
Statistical workflows used for epidemiological analysis.
Output to visualisation tools.
Data set placed into the public domain and preserved for future use.
Results placed into public domain through existing peer review mechanism.
Experimental provenance and workflow recorded for repeatability.

Use PsyGrid trial designer to deploy data set, randomisation and study protocol. Install users and allocate privileges.
PsyGrid data collection system supports recruitment, consent, randomise, treat assess and record data cycle.
PsyGrid supports workflows for data cleaning through heuristic guide user interaction and federation of same schema data sources.
PsyGrid provides virtualisation of disparate data sources to provide a unique view for specified research purpose.
Using imaging for prediction

1. Patient referred by clinician
2. Baseline Structural MRI Scan
   - Store on PsyGrid
3. 12 Month follow-up structural MRI Scan
   - Store on PsyGrid
4. Retrieve baseline scans
5. Submit scans for morphometric analysis to PsyGrid analysis
6. If longitudinal changes detected
7. Retrieve demographically matched non-psychotic patient data
   - Compare with patient - diagnostic indicator
8. Retrieve demographically matched psychotic patient data with medium term outcome data
   - Compare with patient - prognosis indicator and treatment guide
9. Feedback diagnostic and prognosis indicators to clinician
Web of PsyGrid

Usability

Data Set Designer

Security Manager

Element Library

Other Health Data

E-Epidemiology

Decision Support

Data Quality

Socio-economic data

Imaging Data

Genetic Data

Data Linkage

Ontology

Query Builder

Decision Support

Clinical Trials

Data Collection

Data Storage

Data Retrieval

Randomisation

Geographic Mapping

Installer

Trial Designer

Installer

Installer

 Installer
PsyGrid

John Ainsworth, The University of Manchester
john.ainsworth@manchester.ac.uk

http://www.psygrid.org/technology
https://developers.psygrid.org

http://atisha.smb.man.ac.uk:8181/psygrid-data-client
http://atisha.smb.man.ac.uk:8080/eslweb
user “DRN”, Password “sausages”