Utilizing OpenClinica for in-depth Phenotyping in Rare Disease Research

Vera Matser
vam38@cam.ac.uk

Clinical Engagement and Training Lead
Rare Disease Translational Research Collaboration
University of Cambridge
Rare Diseases

- 6000 – 8000 Rare Diseases (RDs)
  - Fewer than 5 in 10,000
  - 1 in 17 or 7% of the population
  - 3.5 million people in UK and 30 million across Europe
  - Up to 30,000, majority far less, handful or single patient

- Rare diseases are often chronic and life-threatening

- 75% of Rare Diseases affect children
  - 30% of RD patients die before 5th Birthday

www.raredisease.org.uk/
Rare Diseases

- 80% of Rare Diseases have genetic origins
  - single gene, multifactorial, chromosomal
  - Causative gene of 3,200 RDs known

- 5 new Rare Diseases described each week

- include rare cancers such as childhood cancers, and cystic fibrosis and Huntington’s disease

www.raredisease.org.uk/
Rare Diseases

- Interest from Industry and funding bodies
- Collaborative & international environment
- RDs are gateways to understand common diseases
  - Familial Hypercholesterolemia (FH)
    - mutations in the LDL Receptor gene
    - Statins
  - Cystic Fibrosis
    - biallelic inheritance of the ΔF508 mutation
    - single copy of ΔF508 ablates CFTR’s binding site and confers resistance to pox

www.bhdsyndrome.org/
Genomics Research

- Sequencing Holy Grail for medical research?
  - Combination of genomic and clinical data

- Genetic testing

- Exome sequencing
  - Part of the genome that codes for proteins
  - Overtaken by Whole Genome Sequencing

- Whole Genome Sequencing (WGS)
  - Chromosomal & Mitochondrial
WGS & search for causal variant

- Genotype to Phenotype
- Link between Genetic variant (e.g. mutations) and Disease
  - Genetic heterogeneity
  - What is “normal” variation & what is causing disease
  - Recruitment of trio’s & relatives
  - Detailed (sub)phenotyping to recover lost power
In-depth phenotyping ~ deep phenotyping

- “Phenotype”
  - Medical: deviation from normal morphology, physiology, or behaviour

- Deep phenotyping:
  - P.N. Robinson (2012): *the precise and comprehensive analysis of phenotypic abnormalities in which the individual components of the phenotype are observed and described*

- Phenotype of a patient “Clinical data”
  - Medical history, physical examination, diagnostic imaging, blood tests etc
NHS & NIHR

NHS - National Health Service
- “Unique” patient identifier
- Free at point of care
- 1 source of data

NIHR – National Institute for Health Research
- Research arm of the NHS
Rare Diseases Clinical Infrastructure team

- NIHR Bioresource Rare Disease
  - BRIDGE projects
  - In-depth phenotyping & WGS
- NIHR Rare Disease Translation Research Collaboration (RD-TRC)
  - In-depth phenotyping
- East of England Genomics Medicine Centre (GMC)
  - 100.000 Genomes project
- Other Rare Disease projects
Rare Diseases Clinical Infrastructure team

Establishing a national federated rare disease patient database

- Physician-entered
- patient-entered (long-term)
A challenge

- "you cannot construct a database with scale, functionality and security because if you design a large system for ease of access it becomes insecure, while if you make it watertight it becomes impossible to use."

Ross Anderson
The ideal system

Universally **accessible**\* electronic system which allows clinicians, researchers and patients to **easily** enter data describing diseases in a **controlled** manner, so that it may be computed upon later.

\* With data privacy caveats, obviously
Confidential data – system setup

- Patient Details [CiviCRM]
- Patient Identifiable [OC Secure]
- Link Tables [OC]
- Pseudonymised Data sets
  - Location
  - People
  - Activities
- Good security standard
- IG Toolkit standard

- N3 - NHS private network
- EPIC - CUH Hospital System

VPN

Portal
Access to OCsecure

- Sophos Access Systems
- 2-factor Authentication
Different Access Scenarios

- Portal – HTML5 VPN Portal
  - Locked down environment
- Remote Desktop Connection (RDC)
  - Hybrid
- Virtual Private Network (VPN)
  - Maintain ability to print, cut & paste
- No public URLs
- Trusted local URL
OpenClinica Instances

**Training**
1. OpenClinica Play
   - Available for new study data managers to familiarise and play with OpenClinica prior to starting the design process.
2. OpenClinica Train
   - RDCIT project teams clean training environment

**Design**
3. OpenClinica Design
   - An OpenClinica environment available for studies to build their study, sites, rules etc.

**Live**
4. OpenClinica Anonymous (interim)
   - Submit collected data
   - Interim live environment to allow studies to get started pre integration with the NHS systems
   - Completion estimated Summer 2015
   - Will be replaced with
5. OpenClinica Secure (final)
   - Final secure live OpenClinica environment
   - Civicrm lookup
   - NHS systems

6. OpenClinica Local
   - Optional if studies do not want to use the RDCIT hosted OpenClinica
Confidential data – OpenClinica setup

- Heavily rely on site concept in OpenClinica
- Central office (RDCIT) are business administrators
  - Everyone else User
  - PI & Clinical Research Fellows (Data Coordinators) at Study
  - All data entry at site level
  - Site per Consultant Group e.g. Addenbrookes_Jones
  - Best Practice compulsory
Best Practice Guide

Best Practice - OpenClinica

Purpose of this document: To provide best practice guidelines for each of the Rare Diseases studies to conform to common standards which will allow the data to be stored in a secure, accessible, extendable and sustainable way. This document's focus is OpenClinica study setup.

- Conventions & standards for the use of OpenClinica
  - to ensure the OC native security is preserved
  - to enforce the minimum data set and clinicians details, site information etc
  - to guarantee record set compatibility between projects
  - to allow the cross linking of data and datasets

- Applicable to all the federated OC instances
Output (in development)

Data Warehouse environment (based on L. Stevens scripts)
- Data Structure modelled on i2b2
- Anticipate using Labkey and/or i2b2 for visualisation
- Remote Desktop
  - Limit capability to extract data

Look, touch but don’t take
Challenge-led development
SPEED
Specialist Pathology Evaluating Exomes in Diagnostics

- Two main areas of recruitment:
  - inherited retinal disease
  - paediatric neurology
- Initial recruitment target of 1000
- 7 sites in England

Strategy

- Basic phenotype captured: 3-10 terms
- Free text
- Low barrier to recruitment
Need for medical coding

Need for an easy way to look up dictionary terms

Why code?

Data quality beyond the project

- SNOMED CT
- Human Phenotype Ontology (HPO)
- ICD10
Clinical coding (1) – Camfetch

- RDCIT developed Camfetch
- List of predefined javascript (jQuery) functions that can be included in CRFs
- User only has to Copy paste the code
- Ability to change script behaviour without changing all CRFs
- Lookup service
  - Ontologies & dictionaries
  - Panogram integration
Clinical coding (1) – Camfetch

SNOMED-CT

Human Phenotype Ontology (HPO)
BPD - Bleeding & Platelet Disorders

- Recruited individuals: 1029
- 23 sites in the UK + 8 international
- Data items: 120 +

- Pedigree data is essential
  - Select WGS candidate
  - Co-segregation

- Genetic counselling
Need to capture family relationships

Black: Affected by a condition
Grey: unknown

Thrombocytopenia = Low platelet count
Excessive bruising, prolonged Bleeding
Causal Variant of Low Platelet

Thrombocytopenia
Low platelet count
Bleeding

Sequencing
Sequencing

Candidate Gene List
1. CBL
2. THPO
3. KALRN
4. TUBB1
5. TPM1
6. TAOK1
7. PTPN11

• Which genes do you already know?
• Which genes come up in multiple analyses
• What is the patient phenotype telling you?
Confirming the causal variant

Co-segregation Analysis

- Candidate variant in affected
- Different variant in unaffected
Panogram – Camfetch integration

- Stand-alone version of phenotips
- The ability to create multiple pedigree trees identified with a family ID
- The ability to load existing pedigree information to Panogram
- The ability to save the modified pedigree
- The ability to pass clinical data from Panogram to a CRF item (under development)
- Pedigrees are stored with the current study name
Panogram - Pedigree Drawing Tool

Enter a Family ID to create pedigree.

Family ID: TRN001

Family ID: TRN003

View Pedigree

Create Pedigree
Panogram - Pedigree Drawing Tool

Currently Firefox only!

OpenClinica
Panogram – finished pedigree

- Anakin Skywalker [DarthVader]
- Padme Amidala
- Luke Skywalker
- Han Solo
- Leila Organa
- Jacen Solo
- Jaina Solo
Panogram

- Link to OpenClinica via external ID
- Export in OC does not show the pedigree
- External ID needs to be filled in manually
  - No check on Gender
- Carrier status
  - Unknown not an option
- Limited options to add genomics information
  - Heterozygous individual
OC Participate - Excited about this!

Patient entered bleeding diary

- Type of Bleeding
- Location
- provoked / unprovoked
- How long did the episode last?
- Intervention needed to resolve episode?
  - Consulted with Dr (phone)
  - Extra medication needed
  - Local haemostasis
PID – Primary Immune Disorders

- Genetic causes of severe immune disorders
  - e.g. Common Variable Immunodeficiency (CVID)

- 900 participants recruited

- 15 UK sites and 3 international sites

- Challenging data entry
  - Lab data
  - Low/Normal/High
Importing Legacy Data

- Shift to OpenClinica mid-study
- Data from different sources

- OpenClinica Data Importer Web application (ODIN)
Select study

OpenClinica Data Importer web application

Select a study: Upload data file > Validate data file > Import subjects > Schedule events > Upload mapping file > Define CRF versions > Mapping > Create XML > Import

Select a study:

Please choose a file:

Upload File

ODIN Beta - ver 0.4 Development
## Upload Data file

<table>
<thead>
<tr>
<th></th>
<th>SubjectId</th>
<th>SecondaryId</th>
<th>EnrollmentDate</th>
<th>PersonId</th>
<th>Gender</th>
<th>DOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DV001</td>
<td></td>
<td>2015-04-01</td>
<td>darthvader</td>
<td>M</td>
<td>1950-02-21</td>
</tr>
<tr>
<td>2</td>
<td>TT001</td>
<td></td>
<td>2015-04-01</td>
<td>theterminator</td>
<td>M</td>
<td>1951-07-28</td>
</tr>
<tr>
<td>3</td>
<td>PLO001</td>
<td></td>
<td>2015-04-01</td>
<td>princessleia</td>
<td>F</td>
<td>1980-11-02</td>
</tr>
<tr>
<td>4</td>
<td>HP001</td>
<td></td>
<td>2015-04-01</td>
<td>harrypotter</td>
<td>M</td>
<td>1999-01-08</td>
</tr>
<tr>
<td>5</td>
<td>GTW001</td>
<td></td>
<td>2015-04-01</td>
<td>gandalfwhite</td>
<td>M</td>
<td>1910-09-29</td>
</tr>
</tbody>
</table>

---

**OpenClinica Data Importer web application**

*Home*  |  *Profile*  |  *Templates*  |  Logged in as: **ch686** (Logout)

Select a study > Upload data file > Validate data file > Import subjects > Schedule events > Upload mapping file > Define CRF versions > Mapping > Create XML > Import

The file moviecharactersdata.csv has been uploaded.

- Go back
- Continue to validation

---

**ODIN Beta - ver 0.4 Developed by Csaba Halmagyi@RDCIT**
Validating the data file - errors

OADIM is the Data Importer web application.

Select a study > Upload data file > Validate data file > Import subjects > Schedule events > Upload mapping file > Define CRF versions > Mapping > Create XML > Import

Current study: MovieCharactersHealthStudy

Datafile structure must be the following:
SubjectId SecondaryId EnrollmentDate PersonId Gender DOB + followed by the data fields.

The following error(s) occurred:
SubjectId must never be empty!
Wrong enrollment date format! 010/04/2015
Incorrect DOB format! 29/09/19100
Missing person id!
Missing gender or invalid format (must be m or f)!

Go back

ODIN Beta - ver 0.4 Developed by Csaba Halmagyi@RDCIT
Successful validation

Current study: MovieCharactersHealthStudy

Datafile structure must be the following:
SubjectId SecondaryId EnrollmentDate PersonId Gender DOB + followed by the data fields.

There were no errors in the data file.
Continue to import subjects
Importing Study Subjects

createSubject: Success
Subject name in xlsx: DV001
SCID = SS_DV001 label = DV001

createSubject: Success
Subject name in xlsx: TT001
SCID = SS_TT001 label = TT001

createSubject: Success
Subject name in xlsx: PL001
SCID = SS_PL001 label = PL001

createSubject: Success
Subject name in xlsx: HF001
SCID = SS_HF001 label = HF001

createSubject: Success
Subject name in xlsx: GTW001
SCID = SS_GTW001 label = GTW001

createSubject: Success
Subject name in xlsx: CJ001
SCID = SS_CJ001 label = CJ001

createSubject: Success
Subject name in xlsx: IJ001
SCID = SS_IJ001 label = IJ001

Subjects import finished.
Successful imports: 7
New subjects: 0
Errors: 0

Continue to scheduling events
Subject Matrix after importing subjects
Schedule events for subjects

OpenClinica Data Importer web application

Select a study > Upload data file > Validate data file > Import subjects > Schedule events > Upload mapping file > Define CRF versions > Mapping > Create XML > Import

<table>
<thead>
<tr>
<th>Event name</th>
<th>Schedule?</th>
<th>Date</th>
<th>Time</th>
<th>Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>BasicQuestions</td>
<td>✓</td>
<td>2015-04-29</td>
<td>10:20</td>
<td>• CharBasicForm - v1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• CharBasicForm - v1.0</td>
</tr>
<tr>
<td>BloodPressure</td>
<td>✓</td>
<td>2015-04-29</td>
<td>10:20</td>
<td>• CharBPForm - v1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• CharBPForm - v1.0</td>
</tr>
</tbody>
</table>

Schedule selected events!
Define CRF version

<table>
<thead>
<tr>
<th>Events</th>
<th>Default CRF for the Event</th>
<th>Check</th>
<th>Version of CRF to use for import</th>
<th>Items in study</th>
</tr>
</thead>
<tbody>
<tr>
<td>BasicQuestions</td>
<td>CharBasicForm - v1.1</td>
<td>☑</td>
<td>CharBasicForm - v1.1</td>
<td>GenHealth, Visit_GP, Have_Asthma</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Comments, PrevIncidents, IncidentCode</td>
</tr>
<tr>
<td>BloodPressure</td>
<td>CharBPForm - v1.1</td>
<td>☑</td>
<td>CharBPForm - v1.1</td>
<td>Systolic, Diastolic, HeartRate</td>
</tr>
</tbody>
</table>

Continue to mapping!
Mapping

- BasicQuestions
  - GenHealth: GeneralHealth
  - Visit_GP: GP visits
  - Have_Asthma: HaveAsthma
  - Comments: Comments
  - PrevIncidents: Incidents
  - IncidentCode: ICD10 code

- BloodPressure
  - Systolic
  - Diastolic
  - HeartRate

Headers from csv
- Comments
- Systolic
- Diastolic

Skip empty cells in datafile
Create ODM XML

© OC15Europe
Creating CDISC ODM XML

![OpenClinica Data Importer web application](image)

Select a study > Upload data file > Validate data file > Import subjects > Schedule events > Upload mapping file > Define CRF versions > Mapping > Create XML > Import

<table>
<thead>
<tr>
<th>Subject name</th>
<th>Event id</th>
<th>O#</th>
<th>Form id</th>
<th>Item id</th>
<th>Value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV001</td>
<td>SE_BASICQUESTIONS</td>
<td>1</td>
<td>F_CHARBASICFOR_V11</td>
<td>I_CHARB_GENHEALTH</td>
<td>1</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BASICQUESTIONS</td>
<td>1</td>
<td>F_CHARBASICFOR_V11</td>
<td>I_CHARB_VISIT_GP</td>
<td>0</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BASICQUESTIONS</td>
<td>1</td>
<td>F_CHARBASICFOR_V11</td>
<td>I_CHARB_HAVE_ASTHMA</td>
<td>1</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BASICQUESTIONS</td>
<td>1</td>
<td>F_CHARBASICFOR_V11</td>
<td>I_CHARB_COMMENTS</td>
<td>Luke! I am your father!</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BASICQUESTIONS</td>
<td>1</td>
<td>F_CHARBASICFOR_V11</td>
<td>I_CHARB_PREVINCIDENTS</td>
<td>Burns of multiple regions::Victim of lightning::War operations involving explosion of marine weapons</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BASICQUESTIONS</td>
<td>1</td>
<td>F_CHARBASICFOR_V11</td>
<td>I_CHARB_INCIDENTCODE</td>
<td>::T29.3::X33::Y36.0</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BLOODPRESSURE</td>
<td>1</td>
<td>F_CHARBPFORM_V11</td>
<td>I_CHARB_SYSTOLIC</td>
<td>180</td>
<td>Done!</td>
</tr>
<tr>
<td>DV001</td>
<td>SE_BLOODPRESSURE</td>
<td>1</td>
<td>F_CHARBPFORM_V11</td>
<td>I_CHARB_DIASTOLIC</td>
<td>120</td>
<td>Done!</td>
</tr>
</tbody>
</table>

**XML import file created successfully.**

*Download XML* (Right click, save as)

- Import data from this XML
- Start a new import
The final CDISC ODM XML

```xml
<ODM>
  <ClinicalData StudyOID="S_Moviecha" MetaDataVersionOID="1">
    <SubjectData SubjectKey="SS_DV001">
      <StudyEventData StudyEventOID="SE_Basicquestions" StudyEventRepeatKey="1">
        <FormData FormOID="F_Charbasicfor_v11">
          <ItemGroupData ItemGroupOID="IG_Charb_basicquestions" ItemGroupRepeatKey="1" TransactionType="Insert">
            <ItemData ItemOID="I_Charb_genhealth" Value="1" />
            <ItemData ItemOID="I_Charb_visit_gp" Value="0" />
            <ItemData ItemOID="I_Charb_have_asthma" Value="1" />
            <ItemData ItemOID="I_Charb_comments" Value="Luke! I am your father!" />
          </ItemGroupData>
          <ItemGroupData ItemGroupOID="IG_Charb_ICD10terms" ItemGroupRepeatKey="1" TransactionType="Insert">
            <ItemData ItemOID="I_Charb_previncidents" Value="Burns of multiple regions" />
            <ItemData ItemOID="I_Charb_incidentcode" Value="T29.3" />
          </ItemGroupData>
          <ItemGroupData ItemGroupOID="IG_Charb_ICD10terms" ItemGroupRepeatKey="2" TransactionType="Insert">
            <ItemData ItemOID="I_Charb_previncidents" Value="Victim of lightning" />
            <ItemData ItemOID="I_Charb_incidentcode" Value="X33" />
          </ItemGroupData>
          <ItemGroupData ItemGroupOID="IG_Charb_ICD10terms" ItemGroupRepeatKey="3" TransactionType="Insert">
            <ItemData ItemOID="I_Charb_previncidents" Value="War operations involving explosion of marine weapons" />
            <ItemData ItemOID="I_Charb_incidentcode" Value="Y36.0" />
          </ItemGroupData>
        </FormData>
      </StudyEventData>
    </SubjectData>
  </ClinicalData>
</ODM>
```
Importing XML into OpenClinica
Subject Matrix after importing

Subject Matrix for Movie Characters' Health Study

<table>
<thead>
<tr>
<th>Study Subject ID</th>
<th>BasicQuestions</th>
<th>BloodPressure</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CJS001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>DV001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>GTW001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>HP001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>IJ001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>PLO001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>TT001</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
</tbody>
</table>

Results 1 - 7 of 7.
Profile & Mapping files

OpenClinica Data Importer web application

Select a study > Upload data file > Validate data file > Import subjects > Schedule events > Upload mapping file > Define CRF versions > Map > Create XML

User: ch686
First name: Csaba
Last name: Halmagyi
Email: ch686@medschl.cam.ac.uk
Current Import ID: 554a01941e113

<table>
<thead>
<tr>
<th>Mapping file name (right click, save as)</th>
<th>Date created</th>
<th>Time created</th>
<th>XML Data file</th>
</tr>
</thead>
<tbody>
<tr>
<td>map_55422dd7b38f5_.csv</td>
<td>2015-Apr-30</td>
<td>14:47</td>
<td>import_55422dd7b38f5.xml</td>
</tr>
<tr>
<td>map_5540b59fe97d9_.csv</td>
<td>2015-Apr-29</td>
<td>11:51</td>
<td>import_5540b59fe97d9.xml</td>
</tr>
<tr>
<td>map_553798833043e_.csv</td>
<td>2015-Apr-22</td>
<td>13:53</td>
<td>import_553798833043e.xml</td>
</tr>
<tr>
<td>map_552bc2562a036_.csv</td>
<td>2015-Apr-13</td>
<td>14:33</td>
<td>import_552bc2562a036.xml</td>
</tr>
</tbody>
</table>

Go back
ODIN Future development

- Validation module
- Import module
- User documentation
- Best Practise
  - Scheduled import
  - Importing into OCdesign
Want to try ODIN?

Option 1:

- Email vam38@cam.ac.uk for OCplay account
- openclinica-testing.medschl.cam.ac.uk/odinbeta

*User documentation on the website*

Option 2:

- Download ODIN from rdcit.org
- Add your own OC instance to settings file (instructions on the website)
Clinical Coding (2) - Decode

- Blog – demystify decode

- Camfetch lookup – unknown clinical codes
  - Clinical knowledge required

- Known clinical codes are incorporated into the CRF using Decode

- Invisible to data entry person
Clinical Coding (2) - Decode

2. Test Result
3. SNOMED code test was performed
4. HPO code for abnormality
Clinical Coding (2) - Decode

Training implications

- Error messages
- CRF far more complicated
  - CRF workshops
  - Webinars
  - Chat sessions for advise
Bringing in existing projects

- Manual recoding required
- Non-destructive method
- Clinician checking clinical coding
Bringing in existing projects

<table>
<thead>
<tr>
<th>A</th>
<th>DATA FIELD NAME</th>
<th>DATA ITEM / USER NAME</th>
<th>OPTIONS/UNITS</th>
<th>CURRENT CODES</th>
<th>CODED TERM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DATE_REG</td>
<td>Date Registered</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>NHS_NO</td>
<td>NHS number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NHS_NO_TYPE</td>
<td>Hospital number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SEX</td>
<td>Sex</td>
<td>Male/Female</td>
<td>1,2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ETHNIC_GP</td>
<td>Ethnic group</td>
<td>map to NHS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CONSENT</td>
<td>Signed consent form</td>
<td>If yes</td>
<td></td>
<td>Consented</td>
</tr>
<tr>
<td>7</td>
<td>RENAL_UNIT</td>
<td>Renal unit</td>
<td>Site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>STATUS</td>
<td>Status</td>
<td>Current</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>DATE_DIAG</td>
<td>Date of Diagnosis</td>
<td>Date</td>
<td>Date of diagnosis (observable entity)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>DIAG</td>
<td>Diagnosis</td>
<td>MPGN</td>
<td>Membranoproliferative glomerulonephritis</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>DIAG_TXT</td>
<td>Diagnosis text</td>
<td>C3 nephropathy</td>
<td>C3 nephropathy</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>CLIN_PRES</td>
<td>Clinical presentation A</td>
<td>Neprotic</td>
<td>Neprotic range proteinuria</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>CLIN_PRES_B</td>
<td>Clinical presentation B</td>
<td>Neprotic</td>
<td>Neprotic range proteinuria</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>AGE_AT_DIAG</td>
<td>Age at diagnosis</td>
<td>Hematuria</td>
<td>Hematuria</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>WEIGHT</td>
<td>Weight</td>
<td>Proteinuria</td>
<td>Proteinuria</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>SYS_BP</td>
<td></td>
<td>Hematuria</td>
<td>Hematuria</td>
<td></td>
</tr>
</tbody>
</table>

© OC15Europe

OpenClinica
Bringing in existing projects

<table>
<thead>
<tr>
<th>Data Item</th>
<th>Response</th>
<th>Ontology Type</th>
<th>Ontology Term</th>
<th>Ontology Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

DATA PLAN - METHOD - Excel

OpenClinica #OC15Europe
100,000 Genomes Project

- Genomics England was set up in 2013 to deliver the 100,000 Genomes Project
  - Company owned by the Department of Health

- NHS Genomic Medicine Centres (GMCs)
  - Recruitment, Consent

- Genomics England Clinical Interpretation Partnership (GeCIP)
  - Analysis
Ethics & Consent

Pilot study consent form

100,000 Genome Project consent form

Copyright ©

OpenClinica

#OC15Europe
Ethics & Consent

Pilot Information Leaflet

100,000 Genome Project Information Leaflet
Rare Diseases Minimum Dataset

- Participant (19)
- Referral (6)
- Consent (10)
- Pedigree (4)
- Current Diagnosis (5) *
- Phenotype (7) *
- Participant History (5) *
- Family History (8) *
- Genetic Testing (10) *
- Clinical Testing (5) *
- Imaging (6) *
- Medication History (7) *
- Surgery (8) *
- Non-pharmaceutical therapies (8) *
- Death (8)

Data Tariff: between 100 and (est) 200 data points

Legend: * indicates a repeating group, such as medication history
7 data points PER medication

EDC = OpenClinica
Training & support

- Hands-on Courses
- Online material
  - User documentation
  - Best Practise
  - Standard NHS CRF templates (FBC, SF-36)
  - Glossary of how to pages
- Helpdesk / scheduled chat help sessions
- CRF workshops
- Webinars

- recently hired new training officer
Training & support

Hands-on OpenClinica courses

- OpenClinica User Training & CRF design (2-day)
  - Clinical Research Fellows (PhD students)
  - Research nurses
  - Clinical Trial Coordinator
  - Data Coordinators

- Data Import & Export Course (1-day)
  - Clinical Research Fellows (PhD students)
  - Data Coordinators

- Data Entry module (2-3 hours) - next priority

- Output module
Training & support

- OpenClinica User Training & CRF design (2-day)

- Knowledge exchange
  - Clinical Coding & Data management
  - Clinical knowledge
Resources

- @RDCIT
- Announce news and events

www.rdcit.org
- User documentation
- Training material
- News

- RDCIT

helpdesk@rdcit.org
OC Training Material

RDCIT Tools

Tools developed by the RDCIT are open source and are available to download on this page.

**ODIN:**

During our work in the Rare Diseases research domain we have come across the problem of importing large amounts of legacy and laboratory data from projects. We have developed a data mapping and importing tool OpenClinica Data Importer web ApplicationN "ODIN" which comes with a user-friendly web interface. ODIN uses the OpenClinica SOAP web services to import subjects and their clinical data into OpenClinica, leaving the OpenClinica audit trail intact. A interface makes the importer easy to use. Please download the source files from [HERE](#).

**Panogram - OC integration:**

Within the area of Rare Diseases collecting accurate pedigree information is essential as 80% of Rare Diseases are genetic in origin. We have integrated Panogram, a standalone Pedigree drawing tool based on the Phenotips platform, into OpenClinica to allow clinicians to draw pedigrees online.

**Camfetch:**

Camfetch is a list of predefined javascript (jQuery) functions that can be included in a CRF. We use camfetch to call our custom services such as Panogram or different lookup services (ontology,
Acknowledgements

- RDCIT Team
  Csaba Halmagyi
  Ping Yu
  Catherine Titterton
  Julie von Ziegenweidt
  Coleen McJannet
  Roger James

- Ex-Team members
  Sarah Hunter

- BRIDGE Cambridge
  Matthew Brown (PID)
  Eleanor Dewhurst (GeL)
  Marie Erwood (SPEED)
  Sofia Papadia (BPD)
  Suthesh Sivapalaratnam (BPD)

- RD-TRC Team
NIHR Bioresource - Rare Diseases

- Identify genetic causes of Rare Diseases
- Improve rates of diagnosis
- Develop and validate treatments

Improve Care for people with Rare Disease & their families

BRIDGE studies
  - In-depth phenotyping
  - Sequencing (exome, WGS)
Rare Disease Translation Research Collaboration (RD-TRC)

- Focus on the causes, impacts and treatment of Rare Diseases
- Establishing a national federated rare disease patient database (RDCIT)
  - OpenClinica
- Clinical Research Fellows & Research Projects
  - Funding for in-depth Phenotyping
- Link phenotype data to genomics data to understand the mechanisms underlying Rare Diseases