IRDiRC

INTERNATIONAL
RARE
DISEASES
RESEARCH
CONSORTIUM

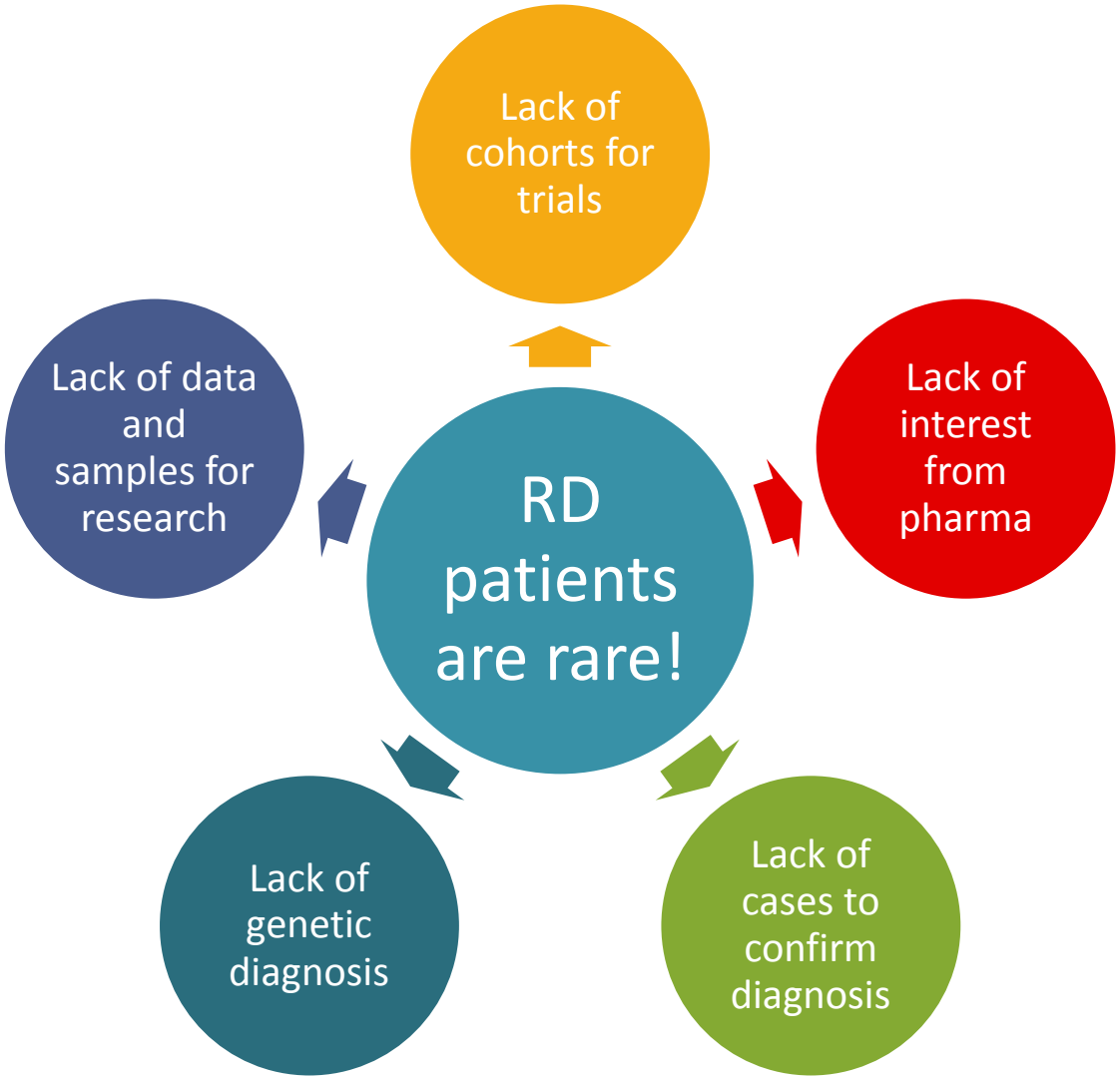
Cross-cutting bottlenecks and solutions in rare disease research

Hanns Lochmüller, Newcastle University

Chair (outgoing), Interdisciplinary Scientific Committee

8 February 2017

Bottlenecks in rare disease research



IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

Many bottlenecks are cross-cutting

…across diseases and across research domains

- A lot of them come down to data
- Not just scarcity of data, but lack of options to reuse the data that does exist
  - Privacy protection issues, particularly across borders
  - Lack of infrastructure for data sharing
  - Lack of standards and interoperability
  - Reluctance to share unpublished data
  - Lack of capacity to analyse large amounts of data
  - Challenges of linking different datasets in different places

Many European and global initiatives are working on solutions

…IRDiRC works with all relevant initiatives through its scientific committees and task forces

IRDiRC scientific committees

…bring together international scientific expertise to address the bottlenecks

- Diagnostics committee
  (Kym Boycott, Gareth Baynam)
  - Bottlenecks in diagnosis of RD

- Therapeutics committee (Diego Ardigò)
  - Bottlenecks in RD therapy development

- Interdisciplinary committee
  (Hanns Lochmüller, Petra Kaufmann)
  - "Everything else?"

IRDiRC
INTERNATIONAL
RARE DISEASES RESEARCH
CONSORTIUM

Interdisciplinary Scientific Committee (ISC)

Hanns Lochmüller (Chair, outgoing)

•University Newcastle upon Tyne (UK)

Petra Kaufmann (Chair, incoming)

•NCATS/ORDR, NIH (USA)

Angel Carracedo

•University of Santiago de Compostela (Spain)

Gema Chicano

•EURORDIS, AADE (Spain)

Jack Goldblatt

•Genetic Services and the Familial Cancer Program of Western Australia (Australia)

Steven Groft

•ORDR/ NCATS, NIH (USA)

Ken Ishii

•Japan Agency for Medical Research and Development (Japan)

Edmund Jessop

- NHS England, UK

Bartha Maria Knoppers

•McGill University (Canada)

Jeffrey Krischer

•University of South Florida (USA)

Samantha Parker

•Lysogene (France)

Rumen Stefanov

•Medical University of Plovdiv (Bulgaria)

Domenica Taruscio

•Italian National Centre for Rare Diseases (Italy)

IRDiRC task forces

…bring together nominated experts from different backgrounds, affiliations and geographical areas on an ad hoc basis

- To work on selected, actionable topics/research areas and push forward policy change
- Collaborate through teleconferences and workshops
  - Production and dissemination of reports
  - Implementation of outcomes
  - Publication in peer-reviewed journals
  - Presentation at conferences

IRDiRC
INTERNATIONAL
RARE DISEASES RESEARCH
CONSORTIUM

Task force:
Privacy-preserving record linkage

- Development of mechanisms to enable linkage of records belonging to the same individual to enable research data sharing across multiple projects and institutions
- Product: Guidelines on the technical and ethical-legal requirements of patient identifiers in Rare Disease Research; recommendations for the most practical, streamlined and minimalistic approach that maximises uptake whilst complying with relevant legal regulations.
- Joint IRDiRC-GA4GH collaboration

IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

Global Alliance for Genomics & Health

Task force:
Automatable Discovery and Access

- Associate clinical data with the scope of consent given for use of that data
- Develop standardized and computer-readable data use types for use in consent forms
- Product: The Automatable Discovery and Access Matrix (ADA-M) provides a standardized way to unambiguously represent consent and other conditions of use, making such information computer-readable and hence directly available for digital communication, searching and automation activities.
- Joint IRDiRC-GA4GH collaboration

IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

Global Alliance for Genomics & Health

IRDiRC policies and guidelines

…mandate that researchers funded under IRDiRC address data sharing and reuse

- "RD research should be collaborative. Resources, data and results should be shared among IRDiRC research projects and made publicly available to the broader community, and duplication should be avoided."
- "RD research should involve patients and/or their representatives in all relevant aspects of the research."
- "International, national, regional and local legislation/ regulations need to be adhered to with respect to data protection and ethical approvals."
- Data producers acknowledge their responsibilities to release data rapidly and to publish initial analyses in a timely manner. IRDiRC members will encourage and facilitate rapid data release"

IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

http://www.irdirc.org/reports-guidelines/policies-guidelines/

Sharing: Benefits

- Overcoming the “rare disease problem”
  - Cohort size
  - Powering trials
  - Finding confirmatory cases
- Reducing costs
- Reducing duplication of effort
- Facilitating validation of results
- Enabling engagement with experts and the patient community

IRDiRC
INTERNATIONAL
RARE DISEASES RESEARCH
CONSORTIUM

Sharing: What?

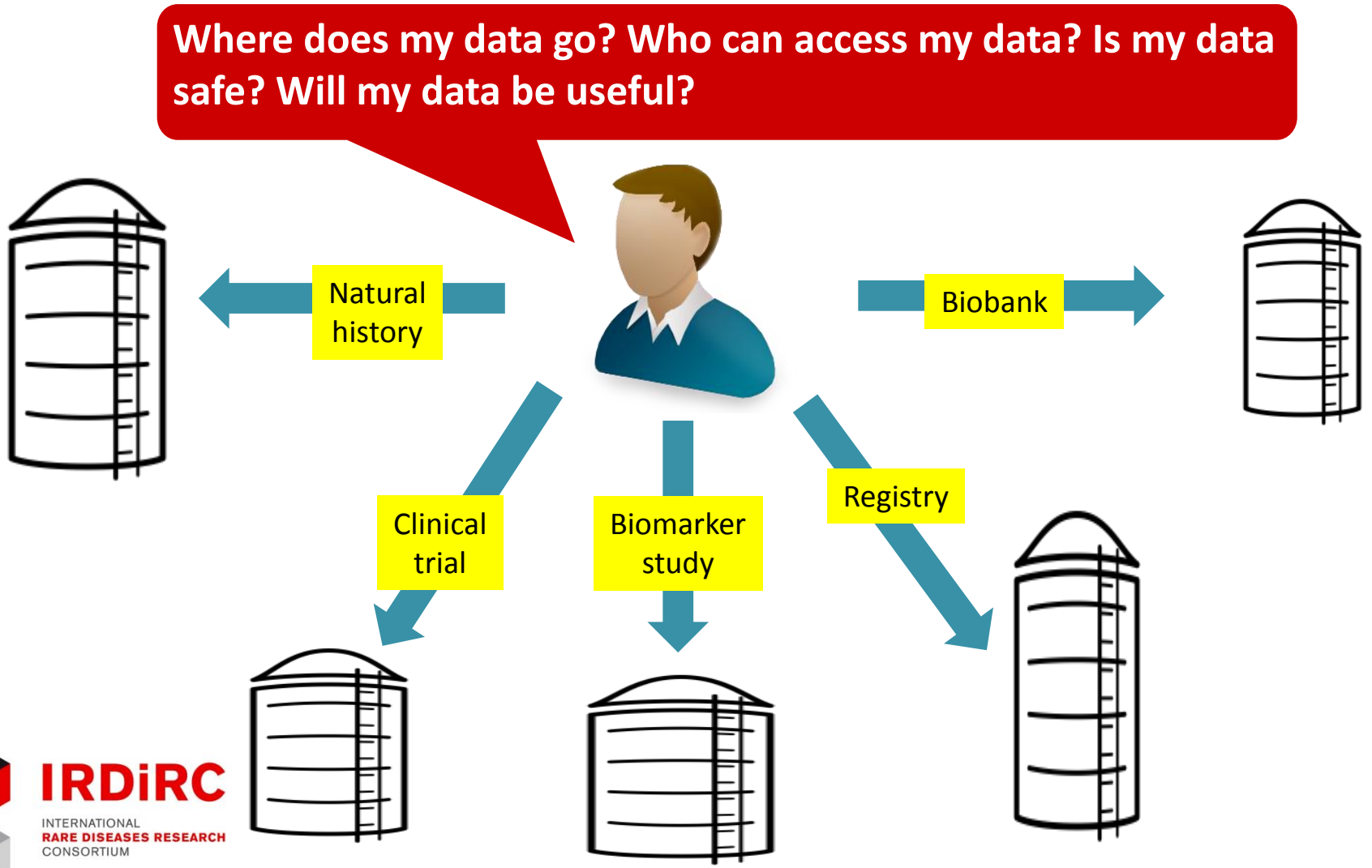
- Raw data from all types of studies
  - Genomic data
  - Phenotypic data
  - Natural history data
  - Clinical trial data
- Biosamples (blood, DNA, tissue samples, cell lines…)
- Linked data and samples
- Brokering access to patients
- …

IRDiRC
INTERNATIONAL
RARE DISEASES RESEARCH
CONSORTIUM

Sharing: Barriers

- General
  - Privacy protection issues: “do I have the patient’s permission?”
  - Lack of infrastructure: “I want to share data but where do I put it?”
  - Lack of standards and interoperability
- Academia
  - Culture of protecting research results: “someone else might scoop my publication!”
  - Lack of incentives for sharing
- Industry
  - IP issues/competition (when sharing own data)
  - Concerns over data quality, regulatory compliance (when reusing data from academia or patients)

What happens to the data of a rare disease patient?



IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

RD-Connect: Infrastructure for RD research data sharing

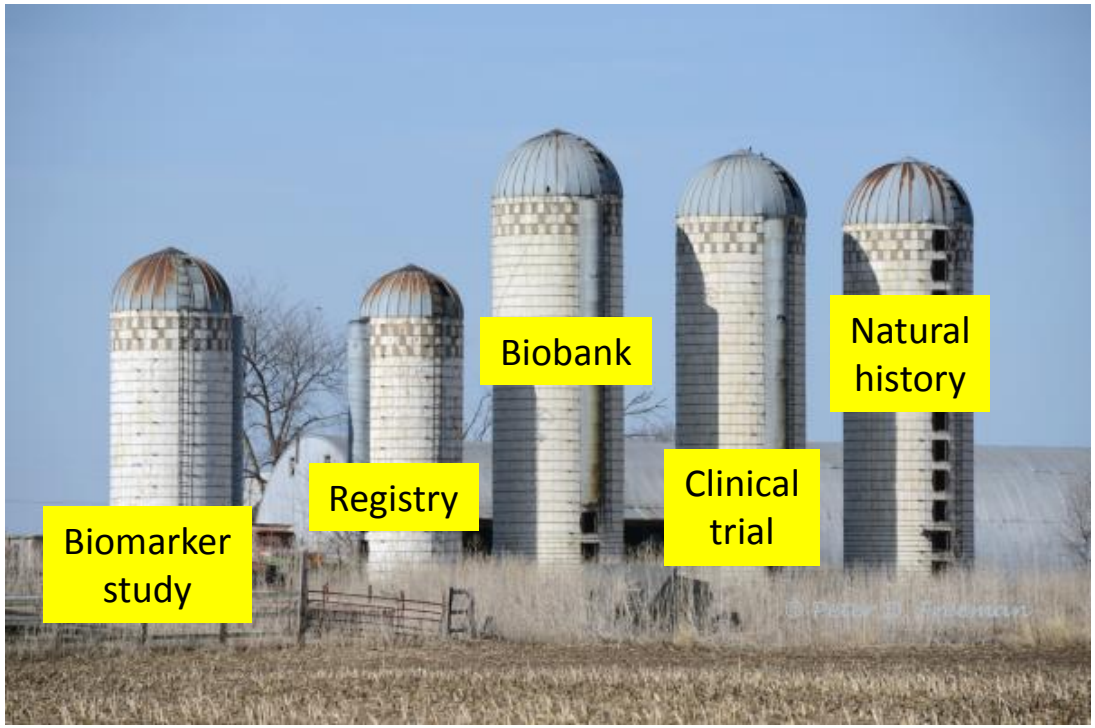
An integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research

Overarching objectives:

- Contribution to the IRDiRC objectives of delivering 200 new therapies for rare diseases and means to diagnose most rare diseases by the year 2020
- Development of an integrated, quality-assured and comprehensive platform in which complete clinical profiles are combined with -omics data and sample availability for rare disease research, in particular IRDiRC-funded research.

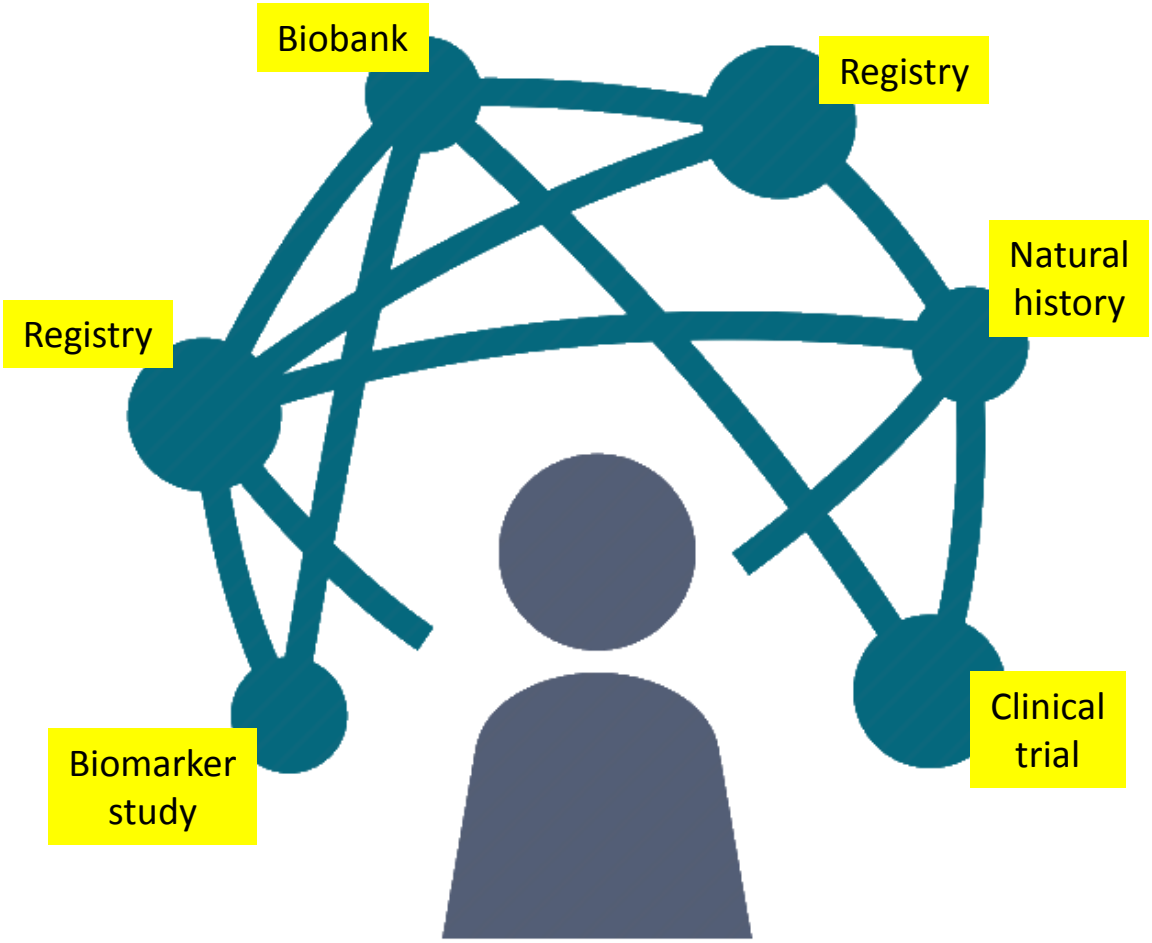
RD-Connect's guiding principle

Overcoming silos



Data sharing for research and better data analysis

Enabling data linkage and reuse



RD-Connect's main aims

- Creation of central system and repository for **reprocessing**, **storing** and **analysing** omics data
  - Raw data hosted at European Genome-phenome Archive (EGA)
  - Raw data reprocessed through standard analysis pipeline for consistency
  - Reprocessed data accessible via National Centre for Genomic Analysis (CNAG) Barcelona platform with user-friendly online analysis interface
- Integration of phenotypic data
- Integration of biosample data
- Development of new bioinformatic tools
- Ethical and legal considerations for data sharing
- Patient input
- Outreach and impact: interaction with rare disease community

Data in the RD-Connect platform

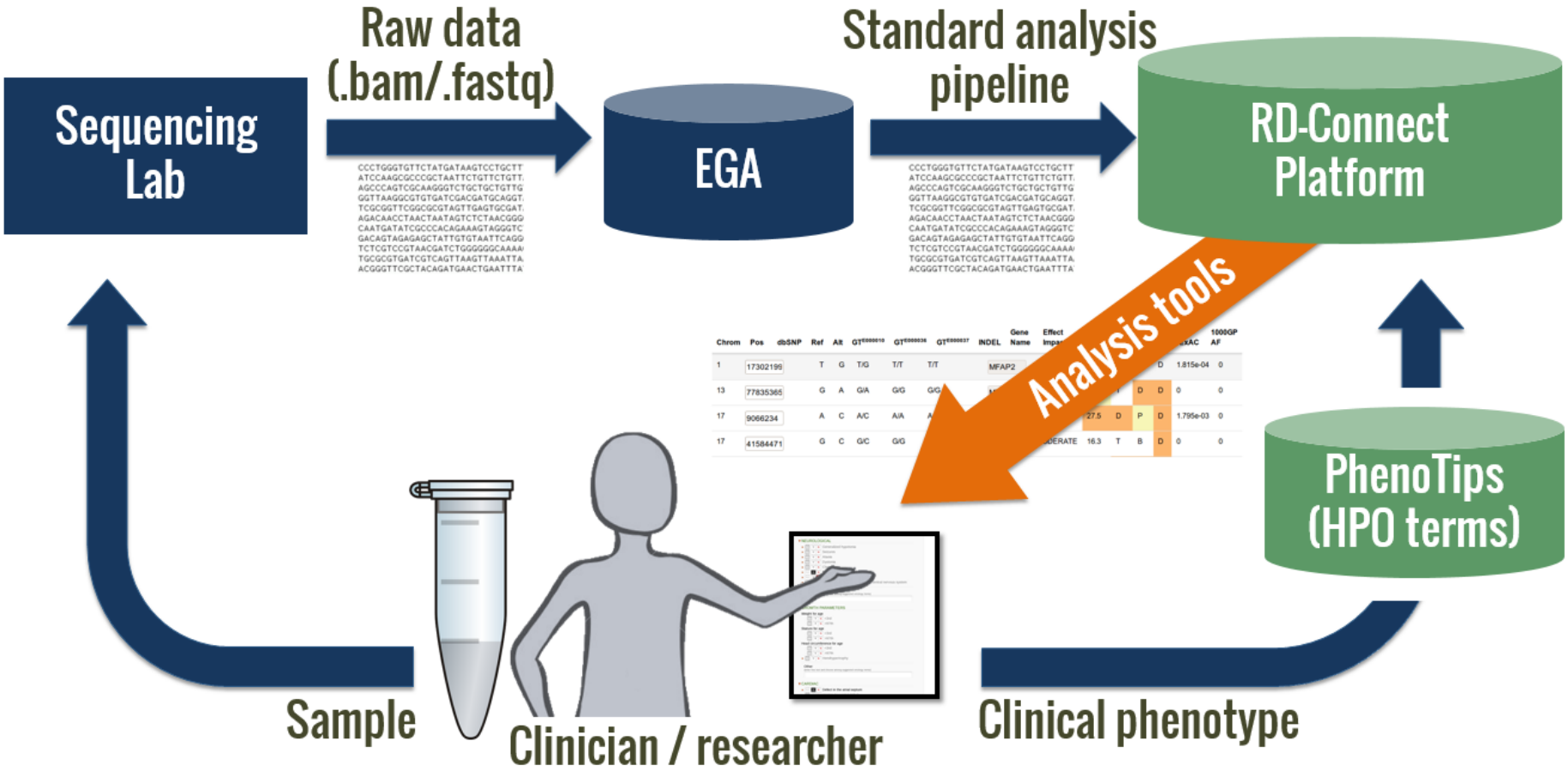
Sample data
(biobank databases)

Clinical data
(registries, and phenotypic databases)

Genomic data
(WES, WGS)

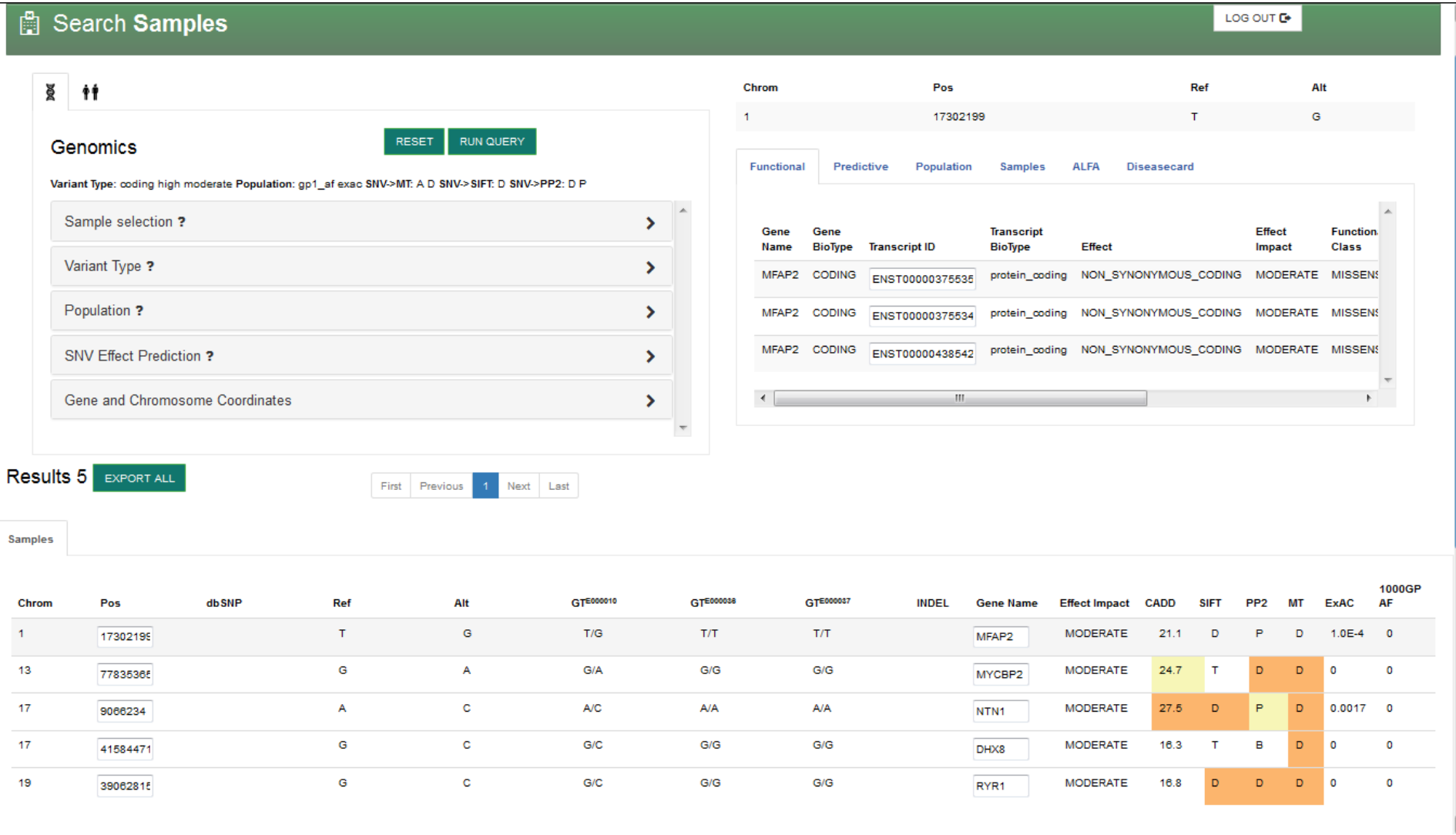
Other omics data
(transcriptomics, metabolomics, proteomics …)

Genomic data in RD-Connect



Almost 2000 WES/WGS to date; 5000 expected in 2017

RD-Connect genomic analysis platform



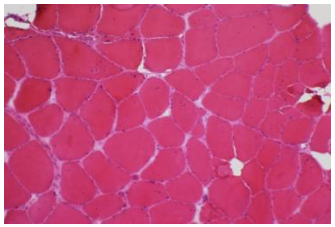
RD Connect

Unique system for researchers to analyse their own data online

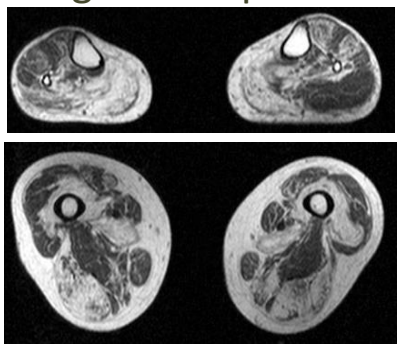
Exome sequencing and data sharing: new congenital myopathy gene

Newcastle case

- Childhood onset
- Proximal muscle weakness, mainly lower limbs
- Slow progression
- CK: normal or mildly elevated
- Muscle biopsy: internal nuclei, fibre splitting and fibre type 1 predominance
- Pattern resembling DNM2 patients



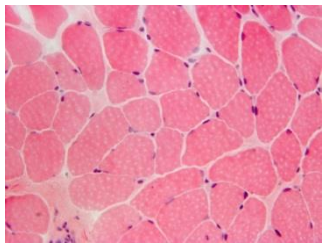
51 years



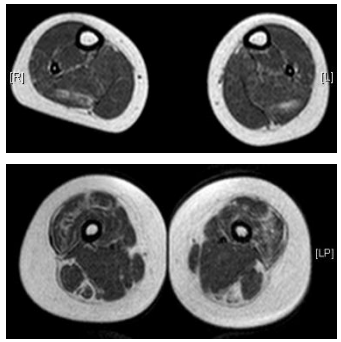
65 years

London case

- Antenatal onset with reduced foetal movement
- Proximal muscle weakness, mainly lower limbs
- Axial weakness
- Joint laxity of hands and ankles
- Slow improvement
- Muscle biopsy: minicores, central cores and some internal nuclei



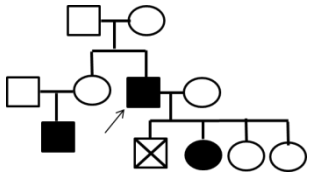
4 years



4 years

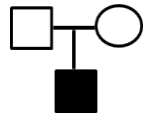
Exome sequencing and data sharing: new congenital myopathy gene

Newcastle case



- Stop gain
- novel (absent from 62k)
- chrX:153049846 G>A; p.Trp415Ter

London case



- Essential Splice Site
- novel (absent from 62k)
- chrX:153050629 G>A

SRPK3

Ana Töpf

- Serine/arginine protein kinase
- Muscle specific, regulated by myocyte enhancer factor 2 (MEF2)
- Known to regulate mRNA splicing and nuclear lamina proteins
- KO mice develop centronuclear myopathy (Nakagawa et al 2005)
- Preliminary data in zebrafish morpholino knockdown shows slow movement and muscle disorganization (unpublished)
- Four new mutations found (manuscript in preparation)

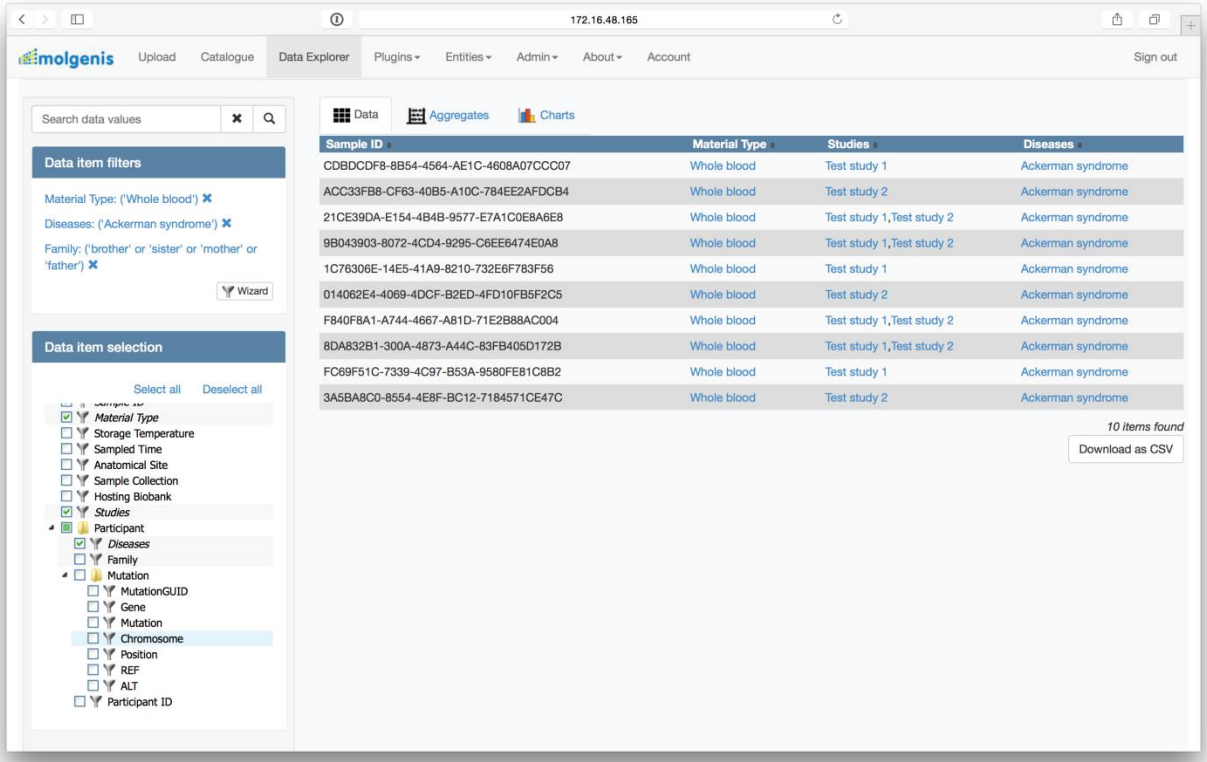
Room 106 from 16:00: European Perspective on Sharing –Omics Data for Personalized Medicine in Rare Diseases
Ivo Gut, Centro Nacional de Análisis Genómico (CNAG), Spain

Biosample sharing

- **(1) Cataloguing** and registration of rare disease biobanks
  - Biobanks can sign up and give details of their biobank in an “ID card”
  - Allows biobanks to participate in RD-Connect infrastructure and research
  - Standardised assessment procedure, MTAs etc.
- **(2)** Sharing **sample-level data** in a common database
  - Not just sample numbers but drill-down right to individual samples
    - Researchers can find the samples they need for their research
    - Allows data from omics experiments to be traced back to the sample it came from for further research

Biosample database

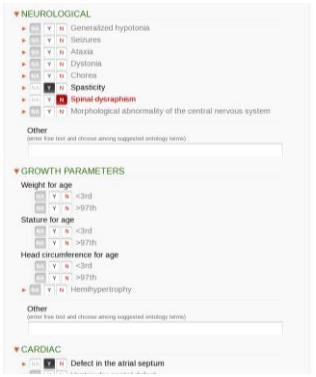
- RD-Connect biosample database contains sample-level data from all participating biobanks



Clinical and phenotypic data

- Phenotype is **more important than ever** in the context of clinical outcome measures and next-generation sequencing analysis
- Requires transformation into a “computable” form
- Requires linkage from different sources (multiple registries, phenotypic databases…)

➡ **FAIR DATA**



RD Connect

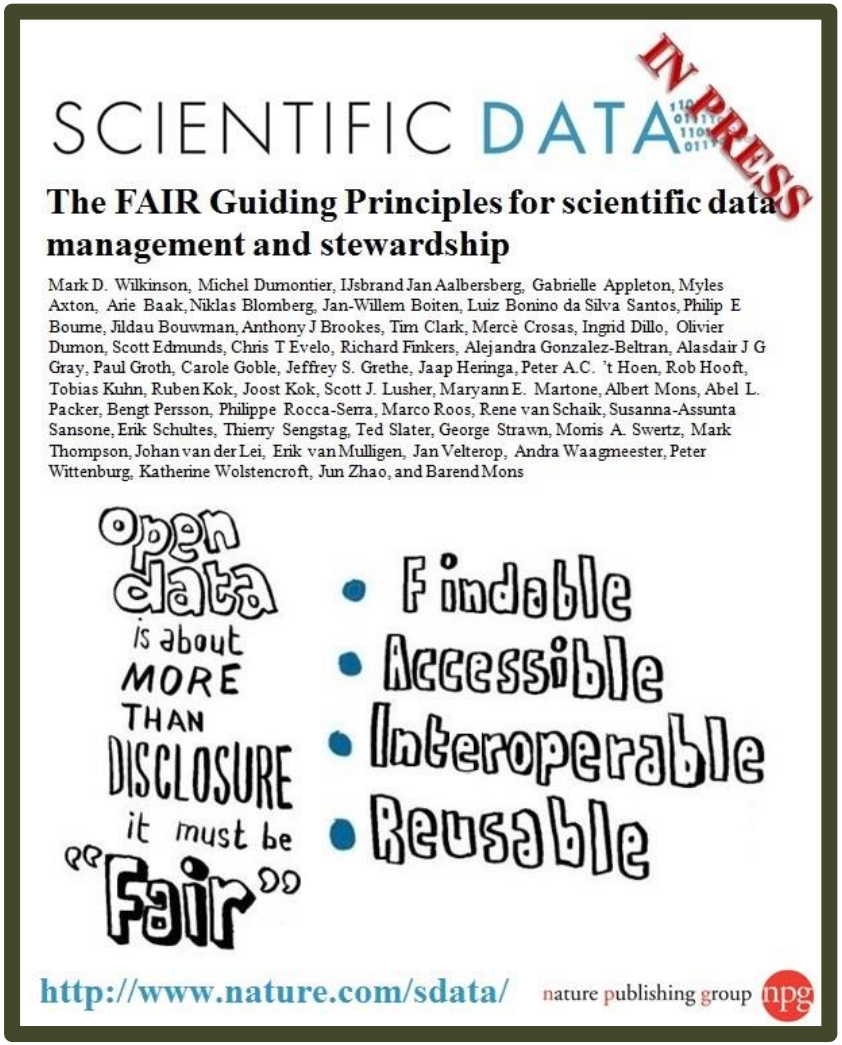
What is FAIR data?

Findable - (meta)data is uniquely and persistently identifiable. Should have basic machine readable descriptive metadata.

Accessible - data is reachable and accessible by humans and machines using standard formats and protocols.

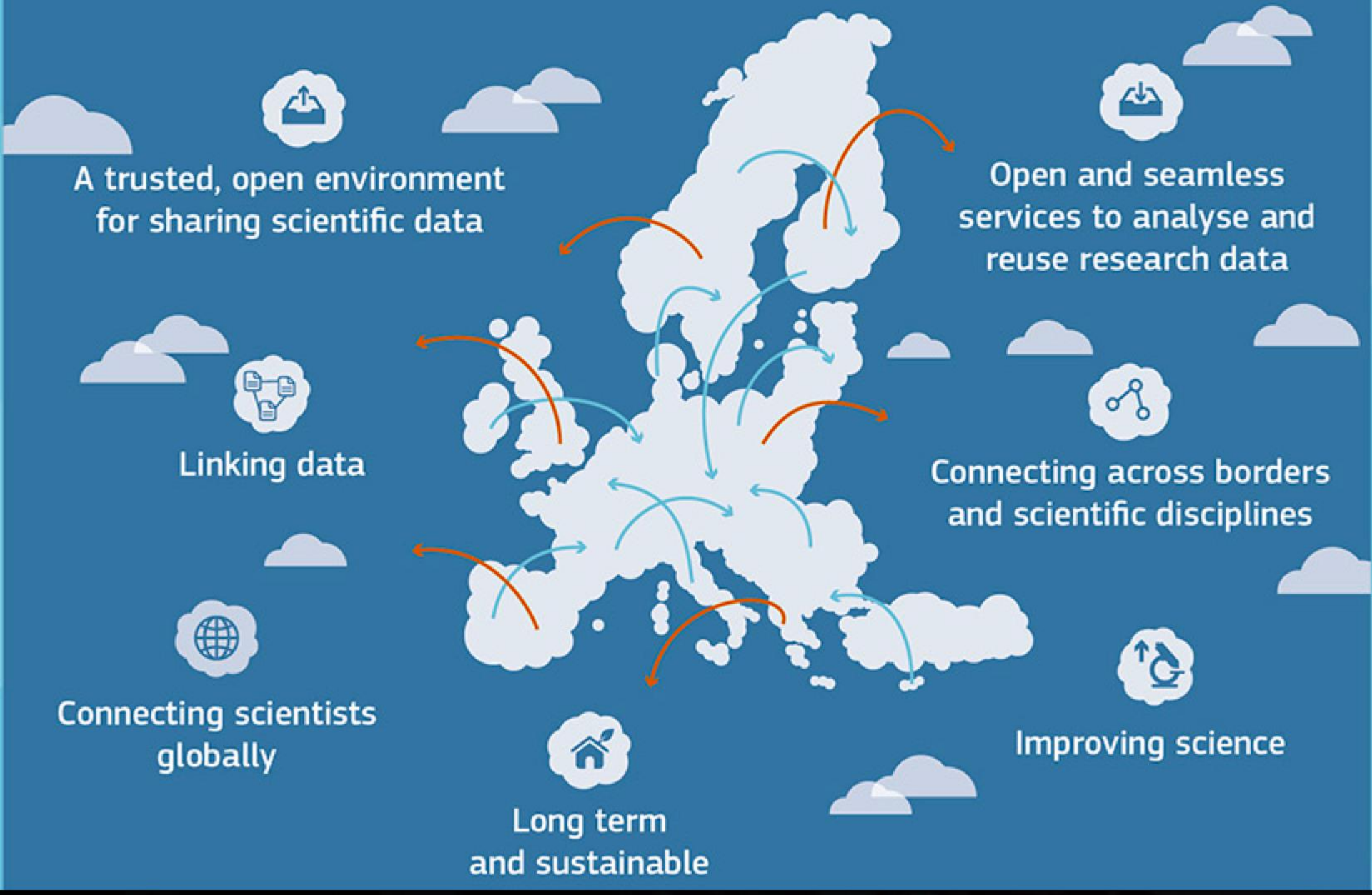
Interoperable - (meta)data is machine readable and annotated with resolvable vocabularies/ontologies.

Reusable - (meta)data is sufficiently well-described to allow (semi)automated integration with other compatible data sources.



EUROPEAN OPEN SCIENCE CLOUD

BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES



Prerequisites for sharing and reuse

- Standards and infrastructure for biosample sharing and reuse

EuroBioBank BBMRI-ERIC Biobanking and BioMolecular resources Research Infrastructure RD-Connect

- Standards and infrastructure for data sharing and reuse

RD-Connect elixir H P O orphanet Inserm EUROPEAN GENOME-PHENOME ARCHIVE

- Patient involvement and consent and privacy issues

EURORDIS RARE DISEASES EUROPE BBMRI-ERIC Biobanking and BioMolecular resources Research Infrastructure RD-Connect

When consent is inadequate…

SARAH ZHANG SCIENCE 04.11.16 11:00 AM

GENETIC SUPERHEROES WALK AMONG US. BUT SHHH! NO ONE

natur biote

Attempted recontact of candidate resilient individuals

We were unable to recontact any of the 13 candidate resilient individuals identified in this study, often due to the absence of a recontact clause in the original informed consent forms used for the studies from which these individuals were identified.

Hanlin Zhou[3], Lifeng
Hardik Shah[1,2]
Ja

The genetic superh
mutations but not the disea

of disease-causing mutations in afflicted
to identify healthy individuals resilient to highly
ensive screen of 874 genes in 589,306 genomes led to the
for 8 severe Mendelian conditions, with no reported clinical manifestation of
demonstrate the promise of broadening genetic studies to systematically search for well
ing the effects of rare, highly penetrant, deleterious mutations. They also indicate that incomplete
Mendelian diseases is likely more common than previously believed. The identification of resilient individuals
may provide a first step toward uncovering protective genetic variants that could help elucidate the mechanisms of Mendelian
diseases and new therapeutic strategies.

Data sharing charter

EJHG Open

European Journal of Human Genetics (2015) 23, 721–728

www.nature.com/ejhg

POLICY

International Charter of principles for sharing bio-specimens and data

Deborah Mascalzoni[*,1,2], Edward S Dove[3], Yaffa Rubinstein[4], Hugh JS Dawkins[5,6,7,8], Anna Kole[9], Pauline McCormack[10], Simon Woods[10], Olaf Riess[11], Franz Schaefer[12,13], Hanns Lochmüller[10], Bartha M Knoppers[3] and Mats Hansson[1]

IRDiRC RECOGNIZED RESOURCES

- Ethical foundations for data and bio-specimen sharing
- Material and Data Transfer Agreement to ensure uniformity of access across projects and countries

Room 106 from 16:00: How to Promote Data Sharing in Rare Disease while Protecting Privacy
Mats Hansson, Uppsala University, Sweden

RD Connect

The General Data Protection Regulation

- GDPR is a regulation under EU law, with direct applicability. It will enter into force on 25 May 2018.
- The GDPR provides that the free movement of personal data within the European Union shall be neither restricted nor prohibited for reasons connected with the protection of natural persons with regard to the processing of personal data.
- The GDPR allows Member States to maintain or introduce further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health. However, this should not hamper the free flow of personal data within the Union.

Thanks to Petra Wilson

How do we address this?

- Large-scale initiative to create a code of conduct for research data sharing in compliance with the GDPR

BBMRI-ERIC
Biobanking and BioMolecular resources Research Infrastructure

RD Connect

elixir

eatris

ECRIN
EUROPEAN CLINICAL RESEARCH INFRASTRUCTURE NETWORK

GENETIC ALLIANCE UK

EURORDIS
RARE DISEASES EUROPE


nature
International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue

Archive > Volume 541 > Issue 7638 > Column: World View > Article

NATURE | COLUMN: WORLD VIEW

BBMRI-ERIC

We must urgently clarify data-sharing rules

Scientists have worked hard to ensure that Europe's new data laws do not harm science, but one last push is needed, says Jan-Eric Litton.

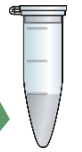
A paradigm for best practice in data sharing in genomic studies?

bbmri-lpc
BBMRI – Large Prospective Cohorts

RD Connect

EuroBioBank

CONSENT



Sample

Phenotypic record

Biobank

EuroBioBank

INFRASTRUCTURE FOR SAMPLE SHARING

Sample can be reused for further research

PHENOTYPIC DATA STANDARDS

Diagnosis

INFRASTRUCTURE FOR DATA ANALYSIS

Sequencing

Online analysis

BBMRI-LPC clinician/ researcher

RD-Connect Platform

STANDARDIZED ALIGNMENT, CALLING & ANNOTATION

Data can be reused for further research

Combined genome-phenome database

INFRASTRUCTURE FOR DATA SHARING

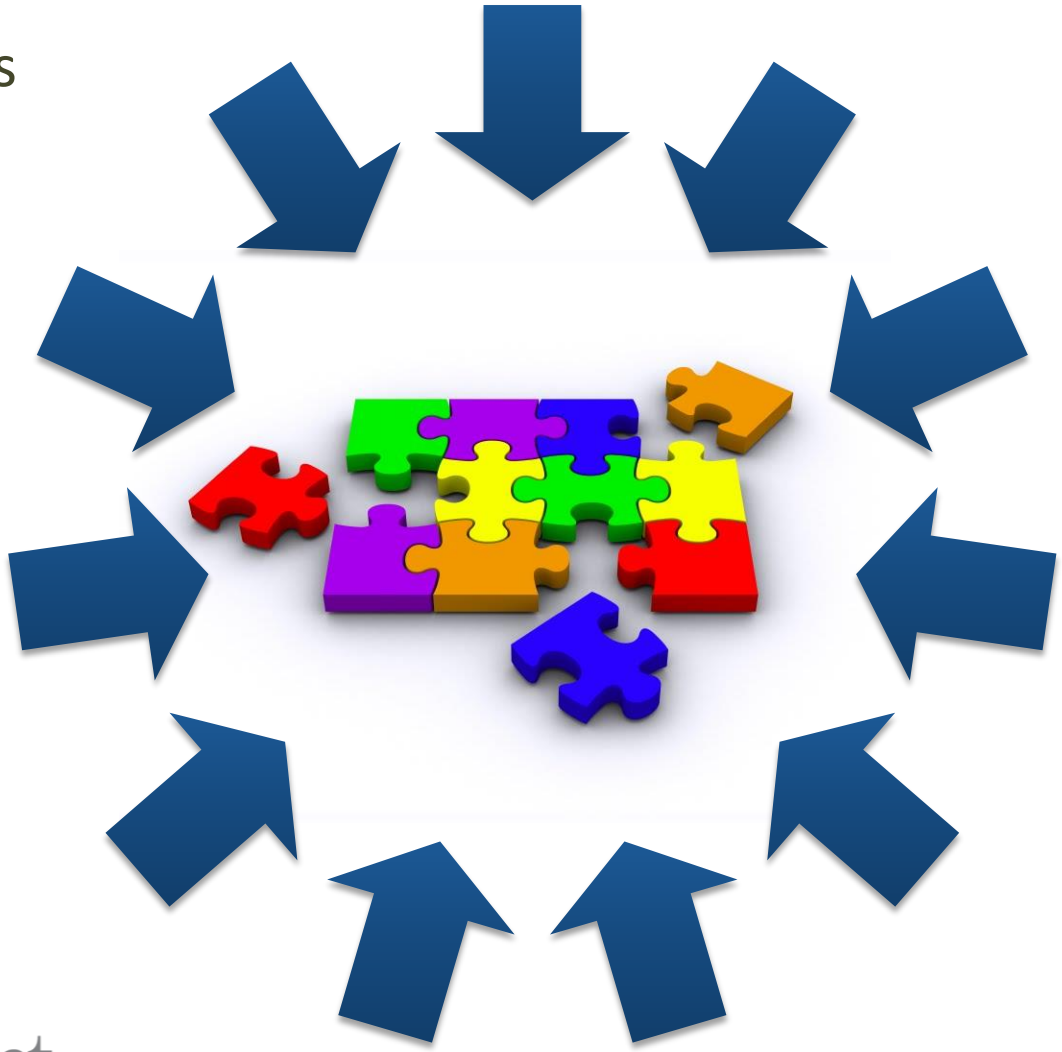
EGA

European Reference Networks

- A mechanism for implementing cross-border healthcare established by the European Commission under Directive 2011/24/EU of 9 March 2011
- Networks of healthcare providers that have been recognised by their own member state as centres of expertise for a particular disease area
- *"European reference networks for rare diseases should serve as research and knowledge centres, updating and contributing to the latest scientific findings, treating patients from other Member States and ensuring the availability of subsequent treatment facilities where necessary."*
- December 2016: networks approved under 23 different disease domains (24 expected soon)
- Formal kickoff March 2017

ERN data in RD-Connect

RD-Connect has agreed with ERN coordinators to include data from all ERNs (genomic, biosample, phenotypic/ registry data)



Supports ERNs in their research and diagnostic goals

Enables data sharing for research

Scientists are invited to….

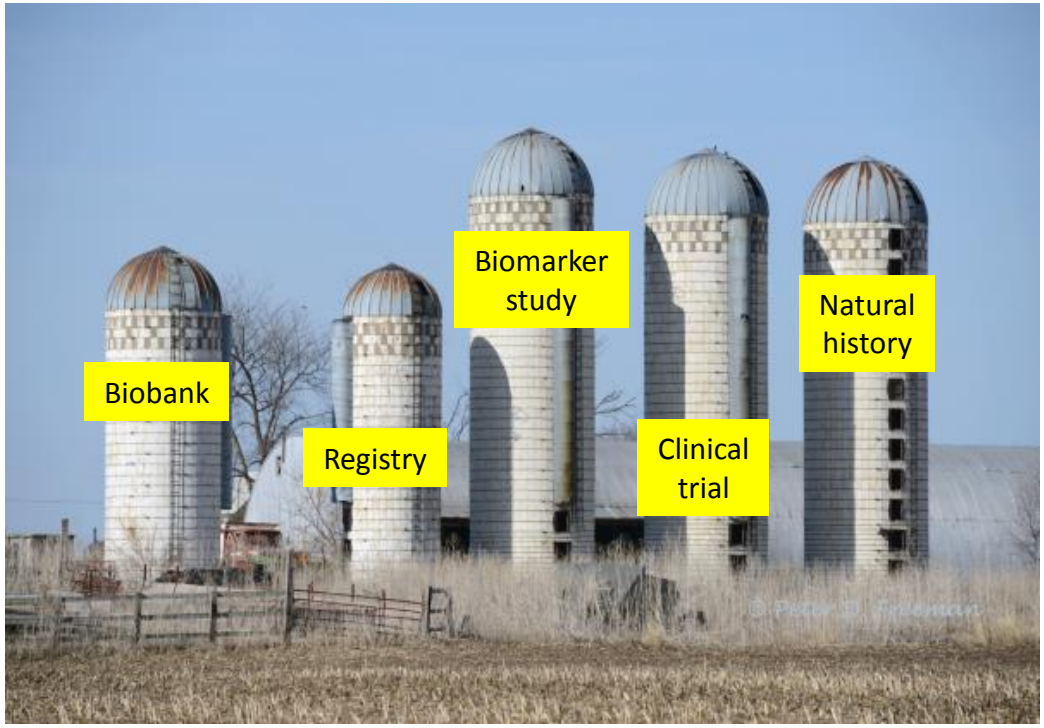
- Deposit WES/WGS/panel raw data from sequencing projects for integration into the RD-Connect platform
- Request an account to look at genomic data in the platform
- Use the RD-Connect platform for gene discovery
- Let our developers know if you want the functionality of the platform improved/adapted for your research
- Add a registry/biobank to the RD-Connect catalogue (ID card)
- Participate in multi-omics user groups, co-develop the functionality that you are interested in (use cases)
- Make sure future projects are fit for sharing (consent)
- Use the RD-Connect impact and dissemination channels

Clinicians are invited to…

- Use appropriate consent for NGS and data sharing
- Use standardized ontologies to describe a patient's phenotype
- Collect samples from patients in a standardized way and deposit them in appropriate biobanks
- Get involved with the interpretation of genome/exome/panel results (genome rounds)
- Feed back results to your patients and point them towards lay-friendly information, encourage their participation
- Include patients in other research (registries, cohorts, natural history, clinical trials, etc)
- Use the RD-Connect impact and dissemination channels

The bottom line: data sharing benefits research

- Data sharing enhances research and data analysis
- Gene and modifier discovery
- Samples for further research
- Genotype-phenotype correlation
- Biomarkers, therapeutic targets…



Learn more in Berlin

http://meeting.rd-neuromics.eu

Joint meeting and outreach day for RD-Connect NeurOmics, EURenOmics and E-RARE

- RD-Connect: 1-2 May
- Outreach day: 3 May
- NeurOmics and EURenOmics: 4-5 May

Questions/feedback

Coordinator:

Hanns Lochmüller – hanns.lochmuller@ncl.ac.uk

Project manager:

Libby Wood – elizabeth.wood2@ncl.ac.uk

Data helpdesk – personalised support!

John Dawson – john.dawson@ncl.ac.uk

Technical questions:

Sergi Beltran – sergi.beltran@cnag.crg.eu

Research questions:

Rachel Thompson – rachel.thompson@ncl.ac.uk

THE John Walton MUSCULAR DYSTROPHY RESEARCH CENTRE

Newcastle University

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

RD Connect
NeurOmics

RD ACTION
DATA AND POLICIES FOR RARE DISEASES

MRC Centre for Neuromuscular Diseases

TREAT-NMD Neuromuscular Network

Thanks to:

The Newcastle JWMDRC team

All our international collaborators

The IRDiRC scientific committee chairs

The interdisciplinary committee

The IRDiRC secretariat

@ConnectRD
@RareDiseasesEU

IRDiRC
INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM