



Registries and Collections of cases: practical and ethical issues

Ségolène Aymé

Orphanet / INSERM

Paris, France



Registries in the EC Communication

- Collaborative efforts to establish and maintain data collection should be supported, providing that these resources are accessible through agreed upon rules. Many research and public health networks financially supported by DG RTD and by DG SANCO have put in place such shared infrastructures, which have been proven to be very efficient tools in improving knowledge and organising clinical trials.

Registries in EC Communication (2)

- Areas to be supported by the MS and the European Commission include: quality standards, including development of strategies and tools for periodical monitoring of the quality of databases and for database upkeep; a minimum common set of data to be collected for epidemiological and public health purposes; attention to user-friendliness, transparency and connectivity of databases; intellectual property, communication between databases/registries (genetic, more generically diagnostic, clinical, surveillance-driven, etc). Importance should be given to linking international (European) databases to national and/or regional databases, when existing.



Context and Rationale

- Very small number of patients affected by a specific rare disease: need to collect data at international level
- Promising achievements but research and development gaps and bottlenecks
 - Limited knowledge on natural history
 - Early market authorisations for Orphan Drugs
- Availability of information technology
 - Easier to store and retrieve data
 - Online solutions for collaborative efforts

Context and Rationale (2)



- Many patient “registries” already in place
- Funding at national and EU level
- Registries for research purpose
 - mainly run by physicians and biologists with no training in epidemiology
 - Discontinuity of funding after completion of the research project
- Registries run by companies
 - Independence of clinicians and researchers ?
 - Rules unclear
- More and more legal constraints

purposes

- Assess the feasibility of trials
- Clinical heterogeneity
- Natural course / quality of life / time to diagnosis
- Treatment outcome / pharmacovigilance
- Assessment tools
- Quality of care
- Establishment of biological markers
- Genotype/phenotype correlation
- Comparison between similar disorders
- Epidemiology of the disease
- Advocacy
- Planning services / advice health authorities
- Appropriate use of treatments
- Basis for cohorts
- Facilitating recruitment
- Support to issue guidelines
- Benchmarking tool
- Build up a community



Clinical research: the needs

- Incidence and prevalence of specific phenotypes
- Incidence and prevalence of specific mutations
- Genotype/phenotype correlation
- Natural course of the disease
- Course of the disease under treatment



Clinical research: requirements

- Establish collections of clinical data
- Establish collections of biological material
- Establish retrieval systems
- Work on a collaborative basis
 - Multidisciplinary
 - All interested parties
 - International

Specificities of RD patient registries

- most RD are genetic in origin and a large proportion of them are familial, which implies that family related cases have to be identifiable;
- The scarcity of cases imposes a large geographical coverage of the data collection which implies multiple collaborations and exchanges of data, ideally transnational; language barriers
- The cost of establishing and maintaining a PR is nearly equal for a prevalent disease as it is for a RD, although budgets are more difficult to obtain for the latter. Legacies are less frequent.
- Intrinsic high motivation of clinicians and researchers
- Active patients organisations as partners to set up the registry and to contribute to awareness and acceptance



Issue 1

Typology of data collections
Importance to select the right type

Clinical research: the tools

- Patient information management systems
- Repository of cases
- Ad hoc observational studies:
 - Transversal studies
 - Longitudinal studies: Cohorts
- Registries

Patient information management systems

- Already in place: no cost to collect the data but cost to access them
- Death certificates, hospital discharges, prescribing, drug consumption, disability certificates
- Not designed for research but for management
- Have to be used knowing the limitations of the data
- Better use as one source among others for capture/recapture studies
- Prerequisite for use should be made easier
- Encourage use of electronic data collection systems and the interoperability of the systems

Ad Hoc observational case studies

- Designed to serve specific purposes
- Clear protocol adapted to analysis
- Prospective and retrospective
- Limited in time
- Less expensive
- Very powerful / +++research
- Necessity of a repository of data



Hospital based registries

- Catalog of cases and of data
- Most common type of registration for RD
- No systematic outreach of additional cases
- Not adapted to establish incidence and prevalence
- Adapted to study the natural history of diseases with the limit due to selection biases: mild cases are missed / poor/ early stages
- Permanent registration
- Lower cost than population registries



Population based Registries

- Permanent collection of cases and of data
- population defined capture area
 - Excludes hospital-based collections of cases
 - Complete ascertainment of cases
- Defined protocol
- Ready to use or basis for complementary ad hoc studies
- Very expensive
- Too stable set of data

Clinical research: goals/tools

- Incidence and prevalence
- Course of the disease
- Ad hoc study design
- Population registries
- Cohorts
- Registries with repeated data collection
- Ad hoc studies with retrospective data



Issue 2

Points to consider for establishing
a data collection or funding it

Goal and tools

- Justification for the systematic permanent collection of data vs ad hoc study
- Case definition well documented
- Clinical objectives
 - Research questions
- Design
 - Study population, study in time
 - Method of case ascertainment, data sources
 - Voluntary or mandatory
- List of variables

Definition of data

- Agreement between experts / health authorities / patients organisations
- Foresee the use when defining the data
- Minimum data set vs maximum data set
- Temptation to collect every thing
 - Time and money consuming
 - Discourage participation
- Difficulties of data definitions
 - Diagnostic criteria
 - Qualitative variables /scores/ quantitative
 - To be adapted to future analysis

Handling of familial cases / genetic data

- Possibility to link cases within a family
- Need to identify individuals recorded several times at different places and within centres (in general)
- Mechanism:
 - family ID
 - Link between individuals
 - pedigree
- Complexity of representation of genetic data

Collection of data



- Data Format
 - Ready to use for statistical analysis
 - Problems of missing data
 - Validation process (good practice)
- Data Support
 - Paper files / questionnaires
 - Computerized database (PC)
 - Shared database: on-line access



Pooling of data

- Model 1: minimum investment
 - Paper files are sent to a unique place
- Model 2: easier for legal reasons and not dependent on Internet connection
 - Data are computerised and stored locally
 - Data are transmitted to a central place from time to time as flat files (Eurocat)
- Model 3: expensive to establish/legal issues
 - Data are stored centrally and collected online

Collection of data

- Data collection: responsible person
 - Clinician in charge of the patient at clinics
 - Research assistant if based on hospital files
 - Patients organisation
- Registrar scope of responsibilities
 - Quality control, safety, reporting, controlling access to data, documenting changes, archiving
- Data storage level
 - By the clinician in direct contact with the patient
 - By a national coordinator: anonymized data
 - By an International coordinator

Softwares

- Need for a software for database management
 - Several commercial products for off line databases
 - Online systems have to be developed
 - To be customised easily
 - Storage of data
 - Management of the storage process
 - Restricted access
 - Data security
 - Logical verification
 - User friendly
- Need for a software for data analysis
 - Any software is usable as the data can be extracted



Issue 3

legal and ethical issues
Intellectual property rights

Status of collections

- Anonymous collections
 - Impossible to link the data with names by any mean as the collection itself was anonymous
- Anonymized collections
 - Stripped of all identifiers but originally collected with names
- Pseudoanonymised collections/ coded collections
 - Data are unidentified for research purpose
 - Can be linked to their sources through the use of a code



Consent requirements

- Informed consent is required for all types of collection
 - Potential benefits justify the establishment of data collections
 - Possibility of misuse imposes a responsibility of
 - appropriate management, including continuity of collections
 - Protection of the subjects' interest



Consent requirements

- IRBs oversight is required to ensure a proper balance between risks and benefits
- Different from one country to another
- Written consent / information to the patient
- Information with respect to the types of research that might be carried out
- Patient Right to withdraw



Collection of data: difficulties

- Responsible body
 - Clinician /researcher
 - Professional organization
 - Patients organization
 - Industry
- Funding body
 - Academic funds
 - Industry support
- Conflicts of interest

Rules about the access to data base and about the use of data

- Rules for the members of the network
 - Member = someone who provides data according to an agreed protocol
 - Every member accesses and uses freely its own data: responsible for the quality.
 - Every member can analyse the whole set of data if agreement of the other members
 - On a project basis (submission process to be established by the consortium)

Rules about the access to data base and the use of data

- Rules for third parties
 - Open access policy. Third parties should be entitled to apply to access the data to maximize the output of the data collection
 - Every member should have the right to accept or refuse the collaboration, either directly or through an ad hoc committee (review board)
 - Authorship rules have to be established
 - Authors = scientists with an added value + collaboration of the network mentioned

Ownership



- Agreements about ownership should be determined by multi-party contracts
- The subject is the primary controller of its own data
- The payor is the owner of the database
- The institution of the researcher at the origin of the data is the owner of the aggregated data
- When processed, the data becomes research data

Ownership (2)

- The principal investigator is the custodian of the data
 - Has to take all appropriate steps to protect the data, its storage, use and access
- Intellectual property to the researcher with due consideration for benefit sharing
- Use by third parties but no transfer of ownership

Conclusion



- Ethical imperative to promote access and exchange of information
- Provided that confidentiality is protected
- Implementation of security mechanisms to ensure
 - Security
 - Long term conservation
 - Long term funding



Conclusion (2)

- Mechanisms in place before start of data collection
- Appropriate funding
- Policy statement about collaborative research
- Written protocol describing the rights and obligations of all parties