EJP RD
European Joint Programme on Rare Diseases

H2020-SC1-2018-Single-Stage-RTD
SC1-BHC-04-2018
Rare Disease European Joint Programme Cofund

Grant agreement number 825575

Del 12.1
"Report on core set of unified FAIR data standards"

Organisation name of lead beneficiary for this deliverable:
Partner 50 – ULEIC, Anthony Brookes (lead)
Partner 45 - CNAG-CRG, Sergi Beltran (co-lead)

Key contributors (alphabetical):
Haitham Abaza, Olamidipupo Ajigboye, Jumamurat Bayjanov, Nirupama Benis, Eva Bermejo Sánchez, César Bernabé, Christophe Béroud, Ronald Cornet, Bruna Dos SantosVieira, Huong Dang, Francesca Frexia, Viviana Giannuzzi, Matthias Haimel, Marc Hanauer, Annika Jacobsen, Rajaram Kaliyaperumal, Annalisa Landi, Yanis Mimouni, Heimo Müller, Michael Nitzlnader, Manuel Posada, Nuria Queralt-Rosinach, Marco Roos, David Salgado, Roberto Santos, Gary Saunders, Günter Schreier, Esther van Enckevort

Due date of deliverable: month 18

Dissemination level:
Public

Table of Contents

1. INTRODUCTION

The complex, multi-site data system required for Rare Disease (RD) and being created by EJP RD must be based upon standards if it is to meet all its objectives. Such standards contribute fundamentally towards making RD resources increasingly Findable, Accessible, Interoperable, or Reusable (FAIR). To create the envisioned data, knowledge and analytics ecosystem comprising many cross-referencing and interoperating parts (i.e., the 'Virtual Platform' or 'VP'), EJP RD Pillar-2 must therefore employ and prioritise standards as one of its cornerstones. These standards must be intelligently chosen and deployed, so that different resources can operate in concert in the most appropriate ways - i.e., ranging from simple cross-referencing through to comprehensive integration as if they were a single database. Standards are key to achieving this, whilst not undermining the uniqueness or independence of any one resource - at all times respecting patient privacy and allowing for the needs and requirements of the resource operators.

Obviously, FAIR challenges cannot be tackled in an 'all or nothing' manner, but as a continuum. Hence, a project-wide approach must be based around different 'degrees' of standardisation. Not everything needs to be fully standardised or universally based on standards, whilst other activities critically depend on a completely unified approach. Therefore, decisions need to be made (and constantly updated, reviewed and adjusted) regarding what standard topics to prioritise, what existing standards to adopt or adjust, and where major gaps exist that require new standards development. Furthermore, for the myriad different things that could theoretically be standardised, there will oftentimes be several different standards development efforts underway somewhere across the globe. These standards initiatives may often be competing with each other, and/or being adopted in parallel in different places, meaning that 'cross-mapping' and 'harmonisation' work becomes extremely important in order for EJP RD resources to interoperate within and beyond its own scope of work.

1.1. Establishing EJP RD Standards

In line with the above, standards work in EJP RD Pillar-2 is conducted strategically according to the formal project workplan. This involves extensive interplay between many partners working in dedicated taskforces formed around each core standards topic. The approach exploits the proven 'middle-out' philosophy whereby we bring together all relevant expertise (in some cases collaboratively with external groups), to undertake iterative consideration of user needs and to standards testing. The overall result is the progressive emergence of standards

solutions that are able to underpin the Virtual Platform and ensure it is maximally FAIR.

Operationally, all this work is based on a series of guiding principles, namely:

1) **Collaboration and synergy**, within and beyond the project, allowing for some diversity so that existing activities can all progress rapidly whilst seeking harmonisation and increasing alignment in the future.
2) **Iterative testing of standards** as they evolve; and feeding back practical experiences so as to guide standards development towards optimisation for specific areas of use. This may be about enhancing utility to make assets Findable, Accessible, Interoperable, or Reusable (FAIR), or ideally all of these.
3) **Organising activities** around broad concepts of:
    - ***Semantic interoperability*** - covering ontologies, terminologies, definitions, data labels, classifications, nomenclature, and coding systems, plus mapping between different solutions.
    - ***Technical (syntactic) interoperability*** - concerning how data and metadata are structured and managed, plus implementation choices over the scope and relationships between data elements, ID systems, input/output formats and file types.
    - ***Ethical, regulatory, legal and organisational interoperability*** - relating to data governance and operational policies, as well as data security and safety, patient privacy, design and capture of consent, rules and objectives for sharing, quality assurance/metrics, and compliance with regulatory requirements. This includes the basis for Authenticating, Authorising users (AAI).

Gathering relevant standards as a basis for establishing the VP is the principle objective of Task 12.1 ("Compare and align core interoperability standards of relevance") which itself is divided into five distinct sub-tasks (see below). These activities interplay with the other work of Work Package (WP) 12 (Task 12.2 "Multi-team software development to enable the FAIR ecosystem" and Task 12.3 "Combining expertise for practical FAIRification support") and synergise with tasks within WP11 and WP13. All of this is facilitated by critical guidance and direction recommended by Overall Architecture (OA) considerations. The additional needs of WP11 and WP13 can include specialised standards requirements (e.g., metadata for resource description), and so such topics are led by those Task teams.

Work Focus groups ("WF") exist in EJP RD Pillar -2, in order to bridge WP Tasks and ensure synergy and optimisation of the efforts. These are obviously key in influencing the adoption and adaptation of standards. This is particularly the case for WF groups concerned with "Interoperability for registries", "Overall

architecture", "Distributed & federated consent control", "Personal data linkage service", "Query builder", "Metadata model & alignment service" and "FAIRification".

2. EJP RD STANDARDS

EJP RD has concentrated on standards challenges relevant to the Virtual Platform goals that are first on the roadmap. To this end, we have identified: (a) existing standards that can be used/piloted directly; and (b) standards where we need to aggregate and/or map between and/or extend existing standards before we can begin using them.

We did not identify any situations where completely new standards will have to be developed from scratch.

2.1. Existing Usable Standards

The following is a list of standards that already exist, and which have been committed to and are actively being employed within EJP RD.

- **Common Data Elements (CDE)** (https://eu-rd-platform.jrc.ec.europa.eu/set-of-common-data-elements_en): This core data standard for RD patients is an important building block of the European Platform on Rare Disease Registration (EU RD Platform), led from the JRC (Joint Research Centre). Released as the first practical instrument for increasing interoperability of RD registries, it endorses 16 data elements perceived to be essential for RD research, thus should ideally be employed by all RD registries across Europe. It provides a major cornerstone for the field. Towards meeting FAIR principles, it recommends using a number of other standards to denote CDEs, including ORPHAcodes, HGVS-nomenclature (Human Genome Variation Society), HGNC (Hugo Gene Nomenclature Committee), OMIM (Online Mendelian Inheritance in Man), and ICF (International Classification of Functioning and Disability). The EJP RD works on recommendations to further increase the FAIRness of CDEs in RD registries.

- **EUropean Patient ID (EUPID)**: The EUPID Services and its supporting technological implementation, are provided in league with the European Rare Disease Registry Infrastructure (ERDRI), the component of the EU RD Platform aimed at making RD registries searchable and findable. It provides a unifying way of managing RD subject unique identifiers across EJP RD. Its design ensures that duplicate entries for the same patient in one or more datasets can potentially be made apparent, whilst maximising patient privacy. Dedicated EUPID pseudonyms are generated upon patient registration for specific "contexts", making sure that datasets from different

contexts cannot be linked without employing the Privacy-Preserving Record Linkage (PPRL) services, which provides mechanisms to ensure that ethical and regulatory requirements are met before a linkage can take place. Also, the EUPID Services provide mechanisms for patient re-identification, if needed. Initially, the EUPID Services were developed in the Paediatric Oncology community with its adoption now being extended to, and presently implemented, for the rare disease community as a whole.

- **Consent and Use Conditions**: This complex domain must be digitally structured to facilitate unambiguous discovery, representation communication, and interpretation of such information. Two existing standards ('Data Use Ontology' (DUO) and 'Automatable Discovery and Access Matrix' (ADA-M) are now mature enough to now be tested for suitability in RD biobank and registry/research situations. Furthermore, both are still evolving under leadership of global standards bodies IRDiRC and GA4GH, with EJP RD partners centrally involved in this. EJP RD will employ these standards to manage consent and use conditions - the formulation of which will be advised by RD stakeholders and will often differ between sources and between countries (notwithstanding the EU standards and requirements). Capturing and representing this information in a standardised way will also help certain imperfections come to light, so that Ethical, Legal and Social Implications (ELSI) experts and stakeholders can use to improve current practices.

- **Disease and Phenotype Ontologies:** These have been a core RD standardisation topic for many years. Widely used, effective solutions now exist, not least the Human Phenotype Ontology (HPO), the Orphanet Rare Disease Ontology (ORDO) and the ontological module linking rare disease to their phenotypes (HOOM). Our use of these will also lead to insights that can inform their further improvement within EJP RD.

- **Phenotypic Data Exchange by PhenoPackets** (https://phenopackets-schema.readthedocs.io/en/latest/index.html): This standard (ISO/ WD 4454) enables the exchange of clinical phenotype related information between information systems within EJP RD and with other project (such as Solve-RD), not least to/from ERN registries, RD-NEXUS discovery tools, Linked Data Platforms, and GPAP. Its scope includes data on individuals and family/pedigree information. It is compatible with ontologies such as the HPO, ORDO and OMIM and with genetic data. A semantic model of PhenoPackets is currently being developed in collaboration with WP13 partners.

- **Omics Data Standards:** Many genomic data standards are provided by the Global Alliance for Genomics and Health (GA4GH), and these are used extensively within EJP RD. These agnostic include SAM/BAM (Sequence Alignment Map/Binary Alignment Map) and CRAM (Compressed Reference-oriented Alignment Map) for sequence alignments, VCF/BCF (Variant Call Format/Binary Call Format) for variant listing, and crypt4GH for storage of genomic data in an encrypted and authenticated state. Additional mature genomic standards in this area include BED, GFF, FASTA and FASTQ. Mature standards also exist for other omics data types. Transcriptomic information consists of positional intensity/signal and feature expression information. Positional data are managed using the bed/bigbed/bigwig standards, while feature data in the form of a matrix employ a tab delimited format or the LOOM file format (based on HDF5). Genomics file formats, such as SAM/BAM, BED or FASTQ, can also be used for transcriptomics data. For proteomics, the ProteomeXchange Consortium (http://www.proteomexchange.org) provides data storage options for mass spectrometry (MS) using recognised standards such as mzML (an open, XML-based format for mass spectrometer output files) and mzIdentML (an open, XML-based data standard for peptide and protein identifications). Finally, relevant here are ISA-Tab or ISA-JSON (https://isa-tools.org/format/specification.html) as multi-omics data exchange formats, along with the Omics Markup Language (OML, ISO/DIS 21393).

- **Data Discovery APIs**: To enable federated querying and subsequent access of data, GA4GH has produced refget (for reference sequences), htsget (for alignment files), RNAget (for data from RNA experiments) and Beacon-1 (for simple variant data), which has now been extended to Beacon-2 (for data on genes, genomics, mutations, phenotypes, clinical information, diseases, demographics, biosamples, and data use conditions). The Beacon-2 API is expected to become a formally approved GA4GH standard later in 2020, but is available to EJP RD already given our role as a GA4GH Driver Project and our (CNAG-CRG and ULEIC) leadership of this international standard development team. The Matchmaker Exchange (https://www.matchmakerexchange.org/) API is a related standard, which enables discovery of similar patients between a few central sites, based on phenotype and/or candidate gene data. Additionally, generic data query protocols, such as W3C's SPARQL protocol, can be applied to query any ontologized data, including patient record-level data. All the above will be integrated and used in concert within the EJP RD VP and its federated elements.

- **Domain-agnostic Interoperability Standards**: Standards that are designed for specific domains or purposes (e.g., regarding distinct data types such as

'genes', 'symptoms') are typically underpinned by a smaller number of more generic standards (e.g., regarding 'identifiers' and 'protocols'), which themselves depend on even fewer basal standards (e.g., 'digital objects', 'resources'). As illustrated by the protocol stack of the internet (https://dl.acm.org/doi/10.1145/3274770), this layering approach is essential for wider interoperability and machine processing. Standards of this type will therefore be needed by EJP RD, and include such things as the Data Catalogue Vocabulary (DCAT), RDF/JSON-LD, Web Ontology Language (OWL), the Linked Data Platform, JSON-related formats, the Digital Object Interface Protocol (DOIP), and its 'FAIR' derivative the FAIR digital object. Furthermore, for some use cases (not least automated machine reasoning), we will benefit from creating semantic models for other standards using the W3C web technology stack, so that they are interoperable at the level of the generic data model provided by the resource description framework (RDF). Also, relevant here are Schema.org and Bioschemas. Schema.org is a specification to mark-up data resources on the web with basic semantic terms to improve findability by search engines such as Google. It describes 'types' of information (items which can be described), which then have 'properties' (the descriptions). The Bioschemas community (sponsored by ELIXIR) proposes new types and properties (e.g. for 'BioChemEntity', and 'Taxon') to Schema.org – including types such as 'gene' or 'molecular entity'. Types can also be given 'Profiles' to state which properties must be used (minimum), should be used (recommended), and could be used (optional).

- **The Open Biological and Biomedical Ontology (OBO) Foundry**: Domain-specific ontologies can be usefully represented in terms of generic foundational ontologies that provide domain-independent definitions of the world (e.g., 'objects', 'events'). This facilitates broad semantic interoperability which enables one to compute over ontologies correctly (e.g., ensuring that "biosample" as a physical entity is not confused with the concept of "taking a biopsy"). The OBO Foundry is a collective of ontology developers that are committed to developing a family of interoperable ontologies that are both logically well-formed and scientifically accurate. To this end, OBO Foundry participants adhere and contribute to an evolving set of principles including open use, collaborative development, non-overlapping and strictly-scoped content, and common syntax and relations. Underpinning all this is the Basic Formal Ontology (BFO http://www.obofoundry.org/ontology/bfo.html) that all OBO Foundry ontologies follow so that they can provide additional semantic links that connect data sets at higher levels of abstraction.

2.2. Specialised Standards Development

Given the specialised topic of EJP RD, and the requirement that the Pillar 2 VP must interact with systems built by ERNs and other associated groups, there is a need to develop specialisations/extensions of a range of existing standards to underpin system interoperability. The starting point for this work will, in some but not all cases, be the existing deployable standards described in Section 2.1. Additionally, the project in ensuring such work is synergistic with activities of initiatives such as IRDiRC, ELIXIR, BBMRI-ERIC, GA4GH, Orphanet, GO FAIR, EuroBioBank, and EU projects that work on solutions for implementing FAIR principles.

This development work is organised under five subtasks within WP12 Task 12.1 ("Compare and align core interoperability standards of relevance"). These subtasks address standards that relate to RD genotype-phenotype databases, RD biobanks, RD patient registries, European Reference Networks (ERNs), and ultimately clinical and digital health data, PCOM/PROM information, natural history and environmental data. To this end, EJP RD has surveyed the standards that already exist, established how/where they are being used, and thereby defined overlaps, conflicts and gaps compared to what is needed by the project.

The five Sub-tasks of Task 12.1 address five overarching areas of need. Using these subtasks as grouping structures, the following list describes the specialised standards now under development:

- ***Subtask 12.1.1 (lead: LUMC) Metadata and data models***

For this subtask the objective is to establish what metadata and data models exist and how they can be improved and utilised, while also assessing interoperability needs of different RD resources. This also underpins interoperability between dataset level (WP11) and record level (WP12) systems. Progress made under this subtask includes:

- ***Creating a metadata model to represent catalogues***

This model must support a core discovery service in the project's VP, namely a federated approach to querying catalogues to find RD resources. Project partners reviewed and assessed a range of existing metadata models, not least the Data Catalogue Vocabulary (DCAT: https://www.w3.org/TR/vocab-dcat-2), the ISO/IEC 11179 Metadata Registry (MDR) standard, the Re3Data schema (https://www.re3data.org/schema), minimal data models like the 'Dublin Core', the Minimum Information About Biobank data Sharing model (MIABIS), the bioCADDIE Data Tag Suite (DATS), and real world catalogue solutions employed by Orphanet, ERDRI, and BBMRI-ERIC. Given that DCAT is built on top of existing vocabularies such as DCMI Metadata Terms

(https://www.dublincore.org/specifications/dublin-core/dcmi-terms/) and W3C's PROV Ontology (https://www.w3.org/TR/prov-o/), it helps address the 'I' and 'R' of FAIR. As such, this provided a good basis for facilitating catalogue interoperability by capturing resource descriptions in a machine processable model.

A hybrid metadata model was devised that heavily leveraged DCAT and included elements of other considered models, to meet the needs of the EJP RD VP. This EJP RD metadata model has been described fully in Deliverable D11.1 ("First Ontological model of resources metadata"). In brief, this model comprises a vocabulary and schema sufficient to define various catalogued resources (e.g., patient datasets, biobanks, bio.tools, disease codes, disease code systems) as minimum metadata representations supported by the VP technologies in EJP RD. Semantic relationships among these resources are captured using slims of ontologies taken from OMIABIS, ORDO, EFO, DUO, HP, NCIT, SIO, IAO, EDAM, so as to enable semantic interoperability. To test the model, example datasets were collected and entered into a resource publishing specification (also known as the 'FAIR Data Point') that was previously built in the RD-connect project and as part of the ELIXIR interoperability platform in a manner that conforms to the REST architectural style. This was successful, and thereby validated the metadata model for describing a dataset in a FAIR manner (i.e., for both machines and humans). EJP RD is now refining the model based on our experience of using it, and updating the FAIR Data Point specification to the new version of DCAT. Extensions to the model vocabulary will additionally enable its use for describing documents, information and resources in other Pillars of EJP RD.

- ***Extending the Common Data Elements (CDE) model***

The basic CDE model is being extended in two ways for use within EJP RD: (i) adding domain specific elements; and (ii) conversion to a sematic web version.

A) *Domain-Specific Common Data Elements (DCDEs)* are needed to increase the power and expressivity of the model, and hence support interactions between ERNs. This is to be achieved via workshops that involve ERN and EJP RD technical teams, being organised in collaboration with JRC. A list of suggested domains has now been created, and ERN 'responsible persons' plus curation teams are being formed around each domain. The goal of the workshops is to obtain lists of DCDEs that are considered essential for every identified ERN domain. These lists will be published on the RDRI website and added to ERDRI's metadata repository (MDR) upon harmonization, thus setting the standard for new registries to include both CDEs and DCDEs. The lists will also be incorporated into EJP RD's registry codebook (https://decor.nictiz.nl/art-decor/decor-project--vasca), thereby facilitating their inclusion in new registries.

B) *A semantic web compatible model of CDE* (and ultimately of DCDE also) is being developed as a priority target for FAIR data standards in EJP RD. The goal here is to reach the limit of 'complete FAIR ness', i.e., to enable CDE based data to be described and used in a fully computer-readable manner. This CDE semantic model will also enable linkage to the EJP RD metadata model that describes data resources. Progress on this semantic standard is currently published as *Version 0.1.0* intended for internal use and testing purposes. It contains two layers. The first captures the features of the Common Data Elements as free-form entity-relationship diagrams [https://github.com/ejp-rd-vp/ERN-common-data-elements/blob/master/images/complete_data_model_v2.0.png]. The second maps and organizes these features through the SemanticScience Integrated Ontology framework [https://github.com/MaastrichtU-IDS/semanticscience].

Work on the semantic model has 4 tiers. Tier 1 structures information about CDE according to the standard data element description, as specified by the ISO/IEC 11179 standard. Tier 2 provides a minimal 'core' ontological model [https://github.com/EJP-RD-vp/ERN-common-data-elements/wiki/Core-model-SIO] that considers each CDE value to be the result of some type of measurement process. This echoes previous developments in this domain [doi:10.1002/humu.22070], but adds features that build on the W3C technology stack to facilitate computational querying. The first released version employs the Semantic science Integrated Ontology (SIO) [doi.org/10.1186/2041-1480-5-14]. Tier 3 concerns defining ontological modules for each common data element, based on an appropriate set of ontology terms given the type of measurement process and the type of result thereby generated [https://github.com/ejp-rd-vp/CDE-semantic-model/wiki]. This is based on W3C recommendations to ensure that a connection can be found in the hierarchy of the ontology or via semantic relations between concepts (e.g. between a disease and a phenotype). Data elements can thus be meaningfully interlinked, even between databases. The first release of Tier 3 uses several well-known ontologies, such as the Laboratory Observation Identifier Names and Codes (LOINC), and the National Cancer Institute Thesaurus (NCIT).

A possible Tier 4 (not yet fully developed, see subtask 12.1.5) would entail mappings to other commonly used semantic modelling frameworks, such as CDISC, ICD, FHIR, OMOP, and others, to achieve interoperability between models. Such mappings typically pertain to semantic correspondence ('sameness') between concepts in different models but could also pertain to defining, and sharing, known semantic relations between different types of data elements to prepare for federated analysis over heterogeneous data in FDA/EMA compliant registries or EHRs.

- ***A patient module within the RD3 data model***

The 'Rare Disease Data about Data' (RD3) [https://github.com/molgenis/RD3_database] concept model was devised in the EU Solve-RD project to record the content and purpose of each data file and patient case represented in that project's cloud-based RD data and analytics service. RD3 initially comprised just one module for the representation of datasets, wherein the patient entity was represented merely as a link to a placeholder module for patient characteristics. That patient module was fleshed out in EJP RD to be compatible with the EU RD Platform's CDE fields and also the genotype-phenotype query API standards emerging in GA4GH. Via a CDE workshop, the module was presented and further aligned with the existing CDEs. This will now be used to pilot federated discovery services that bridge multiple ERN registries and refined as necessary.

- ***Subtask 12.1.2 (lead: BBMRI-ERIC, UMCG) Federation architectures and network registries***

For this subtask the objective is to contribute to the *overall architecture* design, which will underpin a growing federated arrangement of RD resources. If substantial work on developing new/specialised standards (rather than adopting existing solutions) is required under this sub-task, this will be undertaken in the second half of the EJP RD funding period, as data resources and services begin to be placed online and need to be federated. However, it is currently anticipated that we will principally employ global standards (e.g., a common LifeScience AAI, GDPR rules) to establish our federation architectures, as described in Section 2.1 (Existing Standards).

- ***Subtask 12.1.3 (lead: ULEIC) Automated Accessibility and Reusability elements***

For this subtask the overall objective is to address machine-readable consent and use conditions, governance policies, patient/sample/data identifiers, and anonymization/pseudonymization methods. Key progress made under this subtask includes:

- ***Representing Consent and Use Conditions***

The GA4GH 'Data use ontology' (DUO) has been evaluated and deemed a good starting point, whilst the 'Informed Consent Ontology' (ICO) appears less suitable for our needs as it has far fewer relevant terms, the majority of which are also covered by DUO. In combination these two standards do not, however, provide all the elements we need, and neither provides a standardised data structure into which consent and use conditions can be recorded. That latter requirement is the focus of the "Automatable Discovery and Access Matrix"

(ADA-M) standard which defines pure/atomic concepts of use conditions. Many, but not all, ADA-M concepts have been adopted by DUO. Plus, some concepts of use (e.g., for biobanking) are not present in ADA-M. Thus, EJP RD has begun to make use of both DUO and ADA-M (as noted in Section 2.1) as separate entities; but going forward we need to bring them together into one solution and extend their scope. To this end, EJP RD specific examples of data use consent, data use licenses, and institutional/researcher requirements, are now being placed into a combined ADA-M + DUO model (for now called 'ADA-M v2'). A draft of this model has been created, software has been created to enable its testing, and a successful bid was submitted to IRDiRC to have this work proceed under their auspices. Outreach to other global researchers will now take place, to optimise and establish this combined standard as a universal solution. As part of the development work so far, decisions have been made to leverage the 'schemablocks' (modular JSON elements) concept of GA4GH to describe this model, and to achieve a degree of simplification by combining two previously separate "Permissions" and "Terms" sections of the original ADA-M design. As and when ADA-M v2 is deployed to automate aspects of enabling data access, it will be beneficial to also integrate the GA4GH "Passport" standard (whereby access decisions are managed according to use conditions and user profiles) and the related "Visas" concept (explicit access permissions), in conjunction with a single-sign-on approach (LifeScience AAI).

- ***Ethical and Regulatory Underpinnings***

In the context of enabling federated data management, access conditions will need a formal ontological underpinning guided by collaboration between ELSI experts and EJP RD modellers. This has been initiated, and also engages researchers outside of EJP RD who are working on ontological models to automate federated data discovery and analysis (e.g. investigators of the 'Personal Health Train' paradigm). All such ethical and regulatory oversight of informed consent related matters is led jointly by partners FGB and ISCIII to ensure a unified, harmonised, and (wherever possible) standardised approach in compliance with ethics requirements and regulatory provisions and allied support services. Additionally, under Task 10.3 ("Coordination of technical GDPR implementation") oversight and guidance is provided by BBMRI-ERIC on requirements for safe processing of personal and special categories of data in compliance with the GDPR, not least regarding risk assessment and assistance on creating Data Protection Impact Assessments (DPIAs).

- ***Subtask 12.1.4 (lead: ULEIC) Query languages and data exchange formats***

For this subtask the objective is to provide a 'lingua franca' (plus translators) for databases and data discovery platforms to employ for federated query and

response communications, so that they can (with suitable permissions) increasingly merge their record-level query capabilities. Key progress made under this subtask includes:

- ***Interoperable APIs for record-level and resource-level querying***

As starting points for a generic API for querying record-level data, resource-level metadata, and both levels simultaneously, we have the semantic web underpinnings of the Linked Data Platform, the proven (but not yet finalised) 'Case Discovery' API created by GA4GH, the GA4GH 'Search' middleware, and the international 'Beacon-2' API (encompassing all variant types, RD patient diseases/phenotypes, clinical and biosample information). Beacon-2 developments internationally are led by EJP RD partners (CNAG-CRG and ULEIC), and hence take full account of the needs of EJP RD.

Work is now underway to design and pilot Beacon-2 (for inter-ERN matchmaking) and RDF/ SPARQL (for case counting) queries for record-level discovery across ERN registries. Both can exploit the HPO, ORDO and HOOM ontologies, and the latter can ultimately embed mappings to ICD, OMIM, UMLS, MeSH, MedDRA for cross domain querying. Assuming those pilots are successful, the next step will be to share experiences so that the two systems and potentially others can be adapted (including by making translators) to ensure interoperability across the complete VP.

Similarly, work is underway to test both SPARQL Queries and Beacon-2 based approaches for resource-level querying. The previous Linked Data Platform work to this end was described in Section 2.1, along with indications over how the underlying metadata model and platform technology are being advanced further. An API design that uses the Beacon-2 framework but imposes the EJP RD metadata model is now under construction. Just as for record-level querying, these parallel approaches will subsequently be brought together. Buy having this concordance of interoperable approaches across both record-level and resource-level querying, those two different levels can also ultimately be brought together for a seamless query capability that discovers resources and then (with permissions) the records within them.

- ***Subtask 12.1.5 (lead: AMC) Ontologies & semantics for record-level information***

For this subtask the objective is to provide standards which will steadily improve our ability to feed complex medical data into the research world. Key progress made under this subtask includes:

- ***Orphanet Rare Diseases Ontology (ORDO)***

As an IRDiRC recognized resource, ORDO references several terminologies (MeSH MedDRA, UMLS, ICD-10, OMIM), and maps to clinical entities/genes plus

external references (Genatlas, HGNC, OMIM, IUPHAR, Reactome, Ensembl, Uniprot). It likewise maintains mappings between ORPHAcodes (clinical entities concepts) and the terminologies, with a recorded degree of precision (Exact match or BTNT/NTBT). As such, work has to be continuously applied to keep these relationships complete and accurate, to provide a 'semantic interoperability backbone' for EJP RD and beyond. Work is also undertaken (in the Orphanet Knowledgebase) to leverage HPO terms for clinical entity annotations based on deep phenotyping – resulting in the 'HPO-ORDO Ontological Module' (HOOM) based on the generic association model 'Open Biomedical AssociatioN' (OBAN). Both ORDO and HOOM are available in the standardised OWL format. This ongoing RD-specific semantic standards development work by EJP RD is central to the needs of the RD field. The collection of Orphanet's curated mappings is an example resource to be FAIR itself in support of FAIR data stewardship in a community (i.e. to support reuse instead of redundant mapping efforts). Therefore, we work with EOSC-Life and FAIRsFAIR projects, for which defining FAIR-supporting registries and repositories is a goal.

More recently, efforts have been made to combine relevant standards and formalisms used in clinical and in research environments – towards a proof-of-concept openEHR-based 'Clinical Document Repository'. This standard would enable users to store and exchange genomic information along with related clinical data. The approach leverages the Phenopacket data schema as an openEHR archetype, and its demonstrator application engages with the Phenopacket platform and related HL7-FHIR resources.

In order for these ontologies to support computational reasoning, it is necessary that the URI's used can be resolved to extract precise meanings. To verify whether this is the case for given data sets, methods are being developed to automatically assess the correctness of used OWL-ontologies, in terms of syntax and resolvability. As part of this, we have applied ontology matching tools to ORDO along with SNOMED CT and the NCI Thesaurus, to see if the standards are good enough to support automated alignment of datasets based on those ontologies. Performance was found to be reasonable (measured against UMLS Metathesaurus and BioPortal), with 'F1' measures ranging from .42 to .78 for alignment based on complete ontologies, and ranging from .25 to .64 for alignment based on star-modules. This experience will now be fed back to the ORDO developers to help them improve this standard.

2.3. Advising ERNs on Standards for Interoperability

RD patient registries are organised information systems that use observational study methods to collect data (medical and other, entered by clinicians and/or patients) on patients and/or a population defined by a particular disease, condition, or exposure, and serves one or more predetermined scientific, clinical, or policy purposes.

In 2019, based on the Council Recommendation on RD (2009) and on the EUCERD recommendations on patient registration (2013), the EC launched a funding call to help up to 19 ERNs set up high-quality RD patient registries. This added to funding provided through the Annual Work Plan 2016 of the Health Programme, which financed the creation of 5 ERN registries, namely ERKReg (ERKNet), ERN-LUNG registry, EuRRECa (Endo-ERN), PARTNER (PaedCan) and U-IMD (MetabERN), with those projects being initiated in April 2018.

Those funding calls were issued against a backdrop of extreme fragmentation of the field, in that there are 600-1000 RD registries across Europe, within which data are typically not collected EU-wide and there is limited use of shared standards. Therefore, standards that can drive uniformity and interoperability must be strongly emphasised, and submitted proposals had to align with JRC standards. To this end, applicants must work with the European Platform on Rare Disease Registration (https://eu-rd-platform.jrc.ec.europa.eu) and employ CDEs.

To facilitate standards adoption within RD registries, EJP RD worked with the ERN community to create a guidance document [https://www.ejprarediseases.org/wp-content/uploads/2019/09/EJPRD-Pillar2-proposal-for-ERN-registries.pdf] that describes ways to approach standardisation in RD registry design, construction and operation. EJP RD is not funded to assist ERNs with the actual construction and running of such registries, but the project can and does offer advice and services to help ERNs maximise the FAIRness and hence the utility, impact and sustainability of their registries.

Principal topics covered in this guidance document include:

1. Interoperability is not all or nothing
2. Collaboration and synergy are valuable
3. Practical interoperability considerations (Semantic, Technical, and Legal + organisational considerations)
4. EJP RD support for FAIRification of registries
5. Basic standards that could be employed now/soon

2.4. Synergies with Other Standards Efforts

Standards will have greatest impact if they are widely adopted. To help ensure this, standards are best developed in a manner that considers the needs and preferences of many stakeholders, whilst avoiding having many different groups working on similar and competing standards. To this end, EJP RD concentrates its efforts on standard topics it particularly needs and synergises wherever feasible in the context of the overall architecture approach with related efforts elsewhere. The main situations where this occurs are described below.

- **Global Alliance for Genomics and Health (GA4GH)**

This forefront global standards body invites applications from major projects to act as 'Driver Projects', whose role is to identify and recommend challenges for which standards are needed, combine efforts within standards development teams, and test out standards as they progressively mature. At the start of the EJP RD funding period we successfully applied to be a GA4GH Driver Project.

Anthony Brookes (ULEIC) and Sergi Beltran (CNAG-CRG) were selected as 'Champions' for this relationship, based on the longstanding prior involvement in GA4GH activities. In this capacity, they informed the EJP RD consortium about the GA4GH structure and activities, gathered EJP RD partner volunteers for involvement in GA4GH standards teams, identified EJP RD leaders for different GA4GH 'workstreams', and thereby can progressively deliver the latest GA4GH standards to the EJP RD program for early testing and adoption.

Key examples of our partnered activities with GA4GH that directly deliver on the EJP RD missions include:

(i) Query API standards

Having previously developed the globally adopted 'Beacon-1' API (that supports federated querying for data about single base position in the human genome), this team moved on to devise the 'Beacon-2' API (as summarised in Section 2.1). This initiative is co-lead by two EJP RD partners (ULEIC and CNAG-CRG), which ensures the needs of EJP RD are fully considered. Beacon-2 will allow queries to be based on many more parameters (e.g., mutation type, genome regions, individual demographics, sample details, phenotypes, disease diagnoses, gene names, data use conditions, etc.), along with a far wider set of response options (e.g., Yes/No, counts, summary statistics, subject IDs, primary data, web-links, and 'handoffs' to data sites). Beacon subgroups ('Scouts') that work on 'filters', 'domain models', and 'structural Variants', are also led by EJP RD partners. Additionally, our VP could provide one of the early 'implementation demonstrators' required for GA4GH to formally register Beacon-2 as one of their standards. Having this in place will allow us to further federate our patient-based discovery services with other major global projects such as Genomics England Ltd (GEL), and the EU 1M Genomes Project.

(ii) Ethics and Regulatory standards

Viviana Giannuzzi, Annalisa Landi (FGB) and Eva Bermejo (ISCIII) are active EJP RD contributors to the Regulatory and Ethics Work Stream (REWS) of GA4GH. They participate in all REWS teleconferences and have helped revise and optimize the following GA4GH policy standards: Ethics Review Recognition Policy; Machine Readable Consent Clauses; Responsible Data Sharing to Respond to the COVID-19 Pandemic; Model Research Consent Clauses. Recently, Annalisa Landi joined the "Data Access Committee (DAC) Review Standards" group in charge of developing a deliverable aiming at identifying "areas of best practices for procedural standards" to drive consistency and robust reviews for data access requests to genomic and health-related data.

- **ELIXIR**

ELIXIR unites Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated in publicly funded research. It coordinates and integrates bioinformatics activities across its member states and helps users in academia and industry access data, tools, standards, compute and training services to support their research. To date, 22 countries plus EMBL have signed the ELIXIR Consortium Agreement: Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Israel, Luxembourg, the Netherlands, Norway, Portugal, Slovenia, Spain, Sweden, Switzerland, UK and Greece with Cyprus as an observer. Governments and ministries of ELIXIR Member States are responsible for contributing funding for the ELIXIR Hub and coordinating the scientific community in their country into a national Node that provides services to the ELIXIR community. ELIXIR is organised into five 'platforms' (Data, Tools, Compute, Interoperability, Training) and several user communities, including the rare disease community.

Many ELIXIR institutes are represented in EJP RD as partners, or as linked third parties. Key activities that ELIXIR and EJP RD partner on include:

(i) EU-wide strategy for GA4GH interactions

ELIXIR has a long-standing focus on the generation, accreditation, and implementation of globally approved standards. Rather than centring on any single technological solution, the focus has been on the interfaces that enable technically different backends and databases be interoperable. To this end, in 2019, ELIXIR and GA4GH announced a Strategic Partnership that provides technical standards and regulatory frameworks to facilitate responsible sharing of genomic data across national borders. This drives the creation of a secure infrastructure to help projects like EJP RD store, access, and analyse sensitive genomic and health data from participants across and beyond Europe.

EJP RD is an important member (Driver Project) of the ELIXIR/GA4GH Strategic Partnership. Our tight relationship here enables RD related portfolio gaps, bottlenecks, and/or technical issues to be especially highlighted to GA4GH leadership. This stronger voice than any one Driver Project alone could muster, helps to ensure that the suite of GA4GH products meets the needs of the European RD community, and that products are interoperable on a global scale.

Additionally, ELIXIR coordinates the Beyond One Million Genomes (B1MG) 'Coordination and Support Action' funded by the EU as the vehicle for the 1+Million Genomes Member States initiative. The generation, approval, and implementation of the GA4GH portfolio of products is a key driver of the B1MG project. Many EJP RD members and linked third parties are supported in the B1MG project, permitting additional strategic influence and direction on the generation of the standards and regulatory frameworks to enable cross-border access to sensitive human data.

(ii) ELIXIR Rare disease community and commissioned services

This RD community contributes to the development and use of the ELIXIR infrastructure – initially as a use case (in the ELIXIR-EXCELERATE project), and later as one of ELIXIR's Human Data Communities. It focuses on three RD-related topics: (i) Data analysis platforms; (ii) FAIR data; and (iii) Training. The ELIXIR rare disease community collaborates with ELIXIR's Interoperability and Training Platforms; and has thereby helped co-developed FAIR services such as the FAIR data point, rare disease specific FAIR metrics, and the Bring Your Own Data workshop as a model for training workshops. The ELIXIR Rare Diseases Infrastructure project is implementing RD-specific FAIR metrics in a FAIR assessment tool. These currently include standards recognised by IRDiRC and FAIR solutions being used by groups such as ERN or patient organisations. The procedure for implementation of standards in this tool was tested by EJP RD partners in the annual 'Bring Your Own Data' workshop (BYOD) in Rome.

(iii) CNV solutions

Copy Number Variation (CNV) is the most frequently occurring form of genetic mutation, but gaps remain in our ability to share these mutation data and their associated phenotypes. To this end, within ELIXIR, EJP RD partner INSERM-AMU has previously developed the Achropuce French Network of diagnostic laboratories and the BANCCO database that contains more than 31,696 CNV from 20,634 patients. This system is now being further developed in partnership with EJP RD, by leveraging ELIXIR standards work (hCNV community and GA4GH) to ensure proper description of CNVs; and establish ways to share this information with others under a controlled access environment for sensitive data.

(iv) ELIXIR Interoperability Platform

The ELIXIR Interoperability Platform work to assist and encourage the life science community to adopt standardised file formats, metadata, vocabularies and identifiers, and hence is an important platform for ELIXIR's contribution to standards and FAIR principles. ELIXIR nodes in collaboration with rare disease projects (including EJP RD) have co-developed the concept of the Bring Your Own Data workshops for registries (lead nodes: ELIXIR-IT/SI/NL), and the FAIR data point specification (lead node: ELIXIR-NL). The platform further developed criteria for recommended interoperability resources that facilitate FAIR-supporting activities in scientific research (https://elixir-europe.org/platforms/interoperability/rirs). Such services will support the uptake of FAIR data standards by data stewards and FAIRification stewards, coordinated not least by EJP RD.

- **BBMRI-ERIC**

As a European research infrastructure for biobanking, BBMRI-ERIC brings together all the main players in the biobanking field (researchers, biobankers, industry, and patients) to boost biomedical research. To that end, BBMRI-ERIC offers quality management services, support with ethical, legal and societal issues, and a number of online tools and software solutions. BBMRI-ERIC teams are also represented directly as EJP RD partners. In this capacity, their standards work naturally takes account of BBMRI-ERIC activities, so as to maximally synergise the two.

Some key examples of our partnered activities with BBMRI-ERIC that directly deliver on the EJP RD missions include:

(i) The Minimum Information About Biobank data Sharing (MIABIS)

This standard aims to define a core set of data elements needed to describe biobanks, research on samples and associated data. The MIABIS Community Standards apply on several granularity levels; and aim to support interoperability between biobanks sharing their data. General attributes that target the aggregated/metadata level are defined in MIABIS Core 2.0, as described by Roxana Merino-Martinez et al. in 2016 (doi:10.1089/bio.2015.0070). MIABIS modules for describing samples and sample donors at the individual level have been approved by BBMRI-ERIC and have been described by Eklund et al. in 2020 (doi:10.1089/bio.2019.0129). Currently, two main streams of further MIABIS work are underway that will proceed in collaboration with EJP RD partners: (a) final discussion around the release of MIABIS Core 3.0; and (b) data model early discussions.

(ii) Migrating the RD-Connect Products to become BBMRI-ERIC RD resources

The previous EU project 'RD-Connect' produced a "Registry and Biobank Finder" and a "Sample Catalogue", both of which are now to become BBMRI-ERIC resources in the RD domain. These catalogues employ a standardized metadata model which will be mapped and made interoperable with the EJP RD metadata model, and already utilize MIABIS plus other recommended ontologies (i.e., ORDO and HPO). This will allow registries, biobanks, data repositories, and other resources and tools to all be findable by EJP RD based queries. Placing the RD-Connect catalogue in the BBMRI-ERIC framework will also entail extending the BBMRI-ERIC Negotiator, to support RD-specific metadata in its access request pipeline.

- **GO FAIR Implementation Networks**

Global Open FAIR (GO FAIR) implementation networks comprise groups of volunteers who join forces to promote FAIR principles, by building services, fostering culture change, or training (or combinations thereof). The Rare diseases network (RDs GO FAIR) is a GO-CHANGE network that aims to foster and strategically oversee implementation of FAIR principles towards a critical mass of FAIR RD data. It follows RD initiatives, including the EJP RD, for choosing, defining and adopting FAIR standards, guidelines, infrastructures and tools (https://www.go-fair.org/implementation-networks/overview/rare-diseases/). It supports the overarching GO FAIR initiative towards convergence on FAIR standards (https://doi.org/10.1162/dint_a_00038), contributing to generic and specific FAIR standards that are relevant for the RD community.

APPENDIX – I: GA4GH interactions & WorkStream membership

Clinical & Phenotypic Data Capture WorkStream

EJP RD Lead: Alain Verloes

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Clinical Data Exchange	
Alain Verloes	Alessandra Renieri
Ana Rath	Annalisa Landi
Brane Leskošek	Birute Tumiene
Erwin Brosens	Christophe Béroud
Estrella López-Martín	David Salgado
Fabian Prasser	Eva Bermejo-Sánchez
Francesca Frexia	Manila Boarini
Manuel Posada-de-la-Paz	Marco Roos
Marc Hanauer	Marina Mordenti
Petr Holub	Matthias Haimel
Ronald Cornet	Sergi Beltran
Product: Pedigree Representation	
Alain Verloes	David Salgado
Birute Tumiene	Estrella López-Martín
Erwin Brosens	Eva Bermejo-Sánchez
Francesca Frexia	Manila Boarini
Alessandra Renieri	Marina Mordenti
Brane Leskošek	Matthias Haimel
Christophe Béroud	Sergi Beltran
Product: Phenotype Representation for Genomic Medicine and Research	
Alain Verloes	Alessandra Renieri
Brane Leskošek	Ana Rath
Erwin Brosens	Lydie Lane
Estrella López-Martín	Manila Boarini
Eva Bermejo-Sánchez	Marina Mordenti
Francesca Frexia	Montse Gustems
Marco Roos	

Cloud WorkStream

EJP RD Lead: Morris Swertz

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Cloud Testbed Interoperability Demonstration	
Erwin Brosens	Estrella López-Martín Francesca Frexia Matthias Haimel
Product: Data Repository Service (DRS) API	
Erwin Brosens	Francesca Frexia
Product: Task Execution Service (TES)	
Erwin Brosens	Francesca Frexia
Product: Workflow Execution Service (WES) API	
Erwin Brosens	Francesca Frexia Matthias Haimel

Discovery WorkStream

EJP RD Lead: Tony Brookes & Sergi Beltran

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Beacon API	
Alessandra Renieri Brane Leskošek Christophe Béroud David Salgado Erwin Brosens Fabian Prasser Petr Holub Sergi Beltran Tony Brookes (co-lead) Jordi Rambla (lead)	Alain Verloes Marc Hanauer Pablo Gómez-del-Arco
Product: Search API	
Alessandra Renieri Beatriz Martínez-Delgado Christophe Béroud David Salgado Erwin Brosens Fabian Prasser Lydie Lane Petr Holub Sergi Beltran Tony Brookes Olamidipupo Ajigboye	Alain Verloes Brane Leskošek Marc Hanauer
Product: Service Registry Prototype	
Erwin Brosens Eva Bermejo-Sánchez	Alain Verloes Estrella López-Martín

Data Security WorkStream

EJP RD Lead: Heimo Müller

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Authentication and Authorization Infrastructure (AAI)	
Heimo Müller	Erwin Brosens
Annalisa Landi	Francesca Frexia
Fabian Prasser	Manila Boarini
Michal Prochazka	Marc Hanauer
Petr Holub	Marco Roos
	Marina Mordenti
	Matthias Haimel
	Sergi Beltran
	Tony Brookes
Product: Breach Response Protocol	
	Erwin Brosens
	Manila Boarini
	Marina Mordenti
	Sergi Beltran

Data Use & Researcher Identities WorkStream

EJP RD Lead: Heimo Muller

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: DUO + ADA-M Approaches	
Heimo Müller	Annalisa Landi
Erwin Brosens	Francesca Frexia
Fabian Prasser	Hélène Dollfus
Marco Roos	Jean Muller
Nancy Mah	Manila Boarini
Petr Holub	Marc Hanauer
Ronald Cornet	Marina Mordenti
Tony Brookes	Esther van Enckevort
Roberto Santos	
Product: Researcher Identity and Bona Fide Status	
Erwin Brosens	Francesca Frexia
Mary Wang	Manila Boarini
Tony Brookes	Marco Roos
	Marina Mordenti
	Sergi Beltran

Genomic Knowledge Standards WorkStream

EJP RD Lead: Pending

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Variant Annotation: Data Model	
Hélène Dollfus	Alessandra Renieri
Jean Muller	Birute Tumiene
Lydie Lane	Christophe Béroud
Pablo Gómez-del-Arco	David Salgado
Sarah Hunt	Erwin Brosens
	Estrella López-Martín
	Francesca Frexia
	Manila Boarini
	Marco Roos
	Marina Mordenti
	Sergi Beltran
	Tony Brookes
Product: Variant Representation: Data Model/Specification	
Hélène Dollfus	Alessandra Renieri
Jean Muller	Christophe Béroud
Lydie Lane	David Salgado
Sarah Hunt	Erwin Brosens
	Estrella López-Martín
	Francesca Frexia
	Manila Boarini
	Marco Roos
	Marina Mordenti
	Sergi Beltran
	Tony Brookes

Large Scale Genomics WorkStream

EJP RD Lead: David Salgado

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Crypt4GH	
	Christophe Béroud David Salgado Erwin Brosens Lydie Lane
Product: Genetic Variation File Formats (VCF/BCF)	
Christophe Béroud David Salgado	Alessandra Renieri Erwin Brosens Hélène Dollfus Jean Muller Marc Hanauer Marina Mordenti Sergi Beltran Tony Brookes
Product: htsget Streaming API	
Pablo Gómez-del-Arco	Christophe Béroud David Salgado Erwin Brosens Lydie Lane Matthias Haimel Sergi Beltran
Product: Read File Formats (SAM/BAM/CRAM)	
	Alessandra Renieri Christophe Béroud David Salgado Erwin Brosens Hélène Dollfus Jean Muller Marc Hanauer Marina Mordenti Sergi Beltran Tony Brookes
Product: refget API	
	Christophe Béroud David Salgado Erwin Brosens Lydie Lane Christophe Béroud David Salgado Erwin Brosens Lydie Lane
Product: RNASeq Expression Matrix	
Erwin Brosens	Christophe Béroud David Salgado Lydie Lane

Regulatory & Ethics WorkStream

EJP RD Lead: Annalisa Landi, Eva Bermejo-Sánchez, Viviana Giannuzzi

CONTRIBUTOR (and/or ADOPTION)	INFORMED (and/or ADOPTION)
Product: Data Protection	
Annalisa Landi	Alain Verloes
Fabian Prasser	Brane Leskošek
Eva Bermejo-Sánchez	Erwin Brosens
Viviana Giannuzzi	Francesca Frexia
	Manila Boarini
	Marco Roos
	Marina Mordenti
	Tony Brookes
Product: Intellectual Property Licensing Policy	
	Alain Verloes
	Annalisa Landi
	Erwin Brosens
	Manila Boarini
	Marina Mordenti
Product: International Participant Values Survey	
Eva Bermejo-Sánchez	Manila Boarini
	Manuel Posada-de-la-Paz
Product: Paediatric Pharmacogenomic Initiative	
	Annalisa Landi
	Erwin Brosens
	Manila Boarini
	Marco Roos
	Marina Mordenti
Product: Return of Results Policy	
Erwin Brosens	Alain Verloes
	Eva Bermejo-Sánchez
	Manila Boarini
	Manuel Posada-de-la-Paz
	Marina Mordenti