RD Connect

SEMANA GLOBAL ENFERMEDADES RARAS Ciudad de México, ISSSTE Buenavista 2015

ICORD International Conference on Rare Diseases & Orphan Drugs

Instituto de Salud Carlos III

# USEFULNESS OF ONTOLOGIES FOR RARE DISEASES

Manuel Posada, Verónica Alonso and Estrella López Martín

Institute of Rare Diseases Research (IIER-ISCIII)

Madrid, Spain

IRDiRC INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

SEVENTH FRAMEWORK PROGRAMME

RD-Connect Workshop Data Linkage and Ontologies
ISS, Rome, 24-25th September 2015

- No conflict of interest is linked to this presentation


Research enhancement

Better knowledge of RD

New therapies

Improvement of patients life's quality

# INTEROPERABILITY

*"Capability of two or more systems or components to exchange information and to use the exchanged information"*

RD Connect

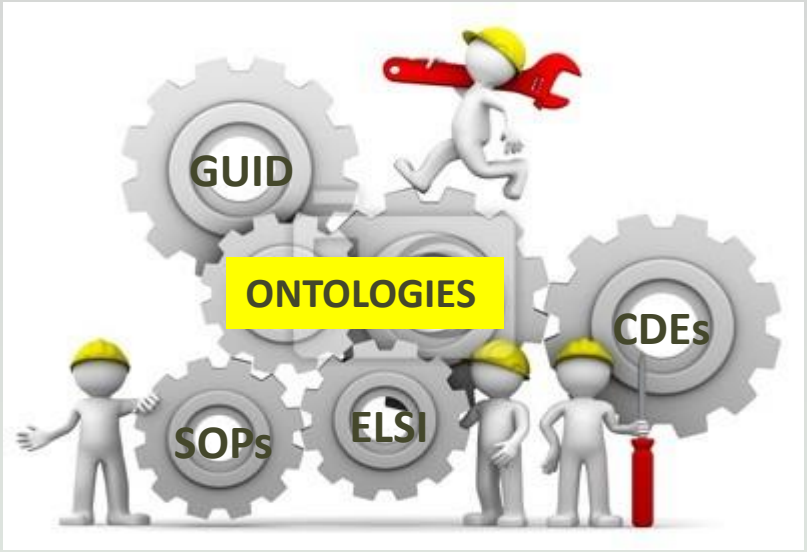
Instituto de Salud Carlos III

# BIOBANKS

## (Sample data)

DATA FLOW

DATA FLOW




## PATIENTS REGISTRIES (Clinical data)

## INTEROPERABILITY TOOLS

## REPOSITORIES FOR –OMICS DATA (Genomic & other -omics data)

DATA FLOW

RD Connect

Instituto de Salud Carlos III

# EUROPEAN RESOURCES

DATA FLOW


DATA FLOW

## INTEROPERABILITY TOOLS

## USA RESOURCES

## AUSTRALIAN RESOURCES

DATA FLOW

# ...WHAT ARE THE ONTOLOGIES?

# ONTOLOGY

- An ontology describes the concepts in some knowledge domain, some of their properties and how the concepts relate to each other.

- Each concept (theoretical item) is represented by one unique reference (URI, real item), which is computer-readable. The references are associated to human-readable terms, definitions and synonyms.

RD Connect

Instituto de Salud Carlos III

## HUMAN READABLE

MULTIPLE POSSIBILITIES¡¡

**Protein HDAC1 *interacts with* gene Parvb**

**or**

**La proteína HDAC1 *interacciona con* el gen Parvb**

**or**

(different synonyms, languages, etc)

## COMPUTER READABLE

UNIQUE REFERENCES ¡¡

**http://purl.uniprot.org/uniprot/Q13547**

***http://conceptwiki.org/index.php/Concept:e6559...***

***http://bio2rdf.org/geneid:29780***

RD Connect

Instituto de Salud Carlos III

## …WHAT IS THE USEFULNESS OF ONTOLOGIES IN THE RARE DISEASES FIELD?

Every term has a unique reference provided by ontologies (standardization)

+

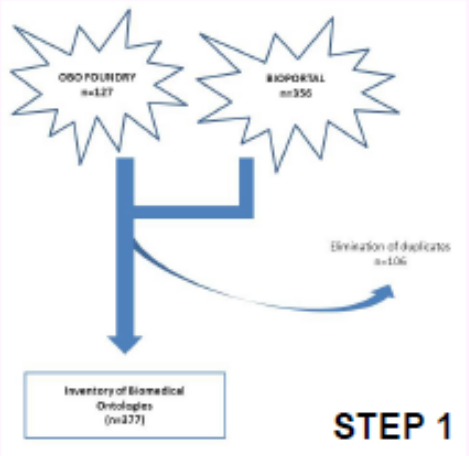
Mappings specify the relation between datasets (linked data)

INTEGRATION OF LINKABLE DATA

# CLASSIFICATION OF BIOMEDICAL ONTOLOGIES USEFUL IN THE RD FIELD

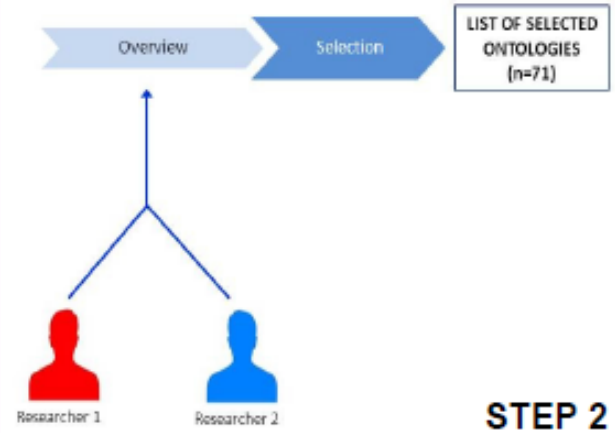
## STRATEGY FOR THE SEARCHING OF ONTOLOGIES

**It consisted of a multi-stage process:**




**STEP 1**

The two leading repositories of biomedical ontologies, the Bioportal and the OBO Foundry, were consulted for collecting a preliminary inventory of existing biomedical ontologies, which was composed by 377 ontologies.




**STEP 2**

After the revision of the inventory by two different researchers, a list formed by the selected biomedical ontologies was obtained. Our searching process was focused on phenotype ontologies, but we also included some other types of ontologies as support tools. This list contained 71 records.


**STEP 3**

Afterwards, the selected biomedical ontologies were classified into the following groups: phenotype, clinic, classifications & nomenclatures, disorder/disease-specific, anatomy, patients-related data, pharmacy, epidemiology, health indicators, miscellaneous and discarded.

# CLASSIFICATION OF BIOMEDICAL ONTOLOGIES USEFUL IN THE RD FIELD

- **PHENOTYPE**
- CLINIC
- CLASSIFICATIONS AND NOMENCLATURES
- DISORDER/DISEASE-SPECIFIC ONTOLOGIES
- ANATOMY
- PATIENTS-RELATED DATA
- PHARMACY
- EPIDEMIOLOGY
- HEALTH INDICATORS
- OTHER

# PHENOTYPE ONTOLOGIES USEFUL IN THE RD FIELD

- Human Phenotype Ontology (HPO): The HPO contains more than 10,000 terms and provides a structured and controlled vocabulary for the phenotypic features encountered in human hereditary and other diseases. The HPO itself does not describe individual disease entities but, rather, the phenotypic abnormalities associated with them.

- Orphanet Ontology (ORDO): The ORDO is an ontology of rare diseases whose content attempts disorders, subtypes, clinical signs and genes. It also includes epidemiology data related to rare diseases in Europe.

- In practical terms, HPO and ORDO cover different aspects and their use is perfectly complementary.

RD Connect

Instituto de Salud Carlos III

**EXAMPLE 1**

# ATAXIA_USE_CASE_ISCIII

- Woman, 49 years old (age at diagnosis).
- Diagnosis: Spinocerebellar Ataxia Type 3 (SCA3), also known as Machado-Joseph Disease.
- Patient case registered in the Spanish National Rare Diseases Registry.

# EXAMPLE 1

## ATAXIA & FAMILIAL SPASTIC PARAPARESIS PATIENT REGISTRY

- Personal Information
- Family History
- Diagnosis-related Issues
- Phenotypic Traits
- Genetic Data

# EXAMPLE 1

| PHENOTYPIC TERMS (REGISTRY) | HPO CODE | PHENOTYPIC TERMS (HPO) |
|---|---|---|
| Cerebellar system affected | HP:0001317 | Abnormality of the cerebellum |
| Pyramidal system affected | HP:0007256 | Abnormality of pyramidal motor function |
| Brain stem affected | HP:0002363 | Abnormality of brainstem morphology |
| Sensory central system affected | HP:0011730 | Abnormality of central sensory function |
| Polyneuropathy | HP:0001271 | Polyneuropathy |
| Demyelinating polyneuropathy | HP:0011402 | Demyelinating sensory neuropathy |
| | HP:0007220 | Demyelinating motor neuropathy |
| Axonal polyneuropathy | HP:0003390 | Sensory axonal neuropathy |
| | HP:0007002 | Motor axonal neuropathy |
| Mixed polyneuropathy | HP:0007327 | Mixed demyelinating and axonal polyneuropathy |
| Second motor neuron signs | HP:0002366 | Abnormality of the lower motor neuron |
| Extrapyramidal system affected | HP:0002071 | Abnormality of extrapyramidal motor function |
| Visual system affected | HP:0000504 | Abnormality of vision |
| Hearing system affected | HP:0000364 | Hearing abnormality |
| Cognitive/Behavioural impairment | HP:0100543 | Cognitive impairment |
| | HP:0000708 | Behavioural abnormality |
| Epilepsia | HP:0011097 | Epileptic spasms |
| Walk | HP:0002355 | Difficulty walking |
| Autonomous walk | Not found | Not found |
| Use of walking stick | Not found | Not found |
| Use of wheelchair | Not found | Not found |
| Bedridden | Not found | Not found |
| Oculomotor signs | HP:0006860 | Abnormality of eye movements |
| Other signs (free text) | - | - |

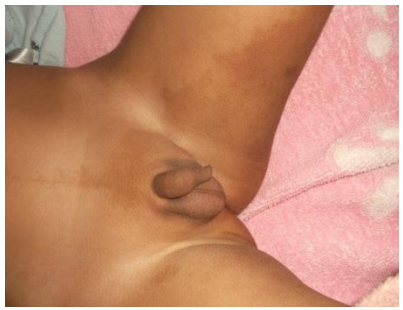
# EXAMPLE 2

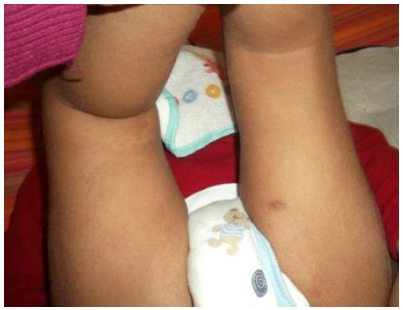
**Five years old child (non diagnosed case) with the following phenotype:**

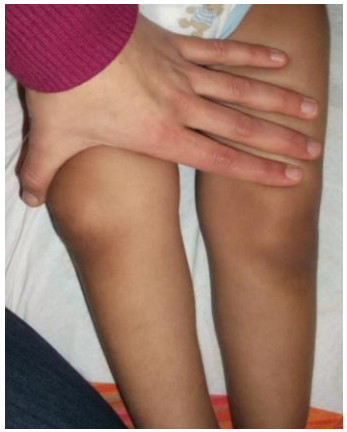
- **Hypothyroidism**
- **Hyperpigmentation**
- **Intellectual disability**
- **Delayed speech and language development**
- **Delayed motor development**
- **Short stature**
- **Microcephaly**
- **Microretrognathia**
- **Trigonocephaly**
- **Macrotia**
- **Low IGF-1 blood levels**
- **Low $Fe^{2+}$ blood levels**

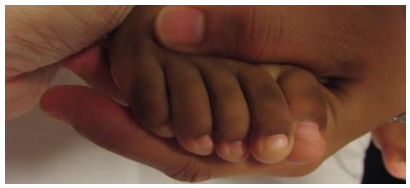
# SpainRDR: Issues regarding the use of ontologies in registry data curation

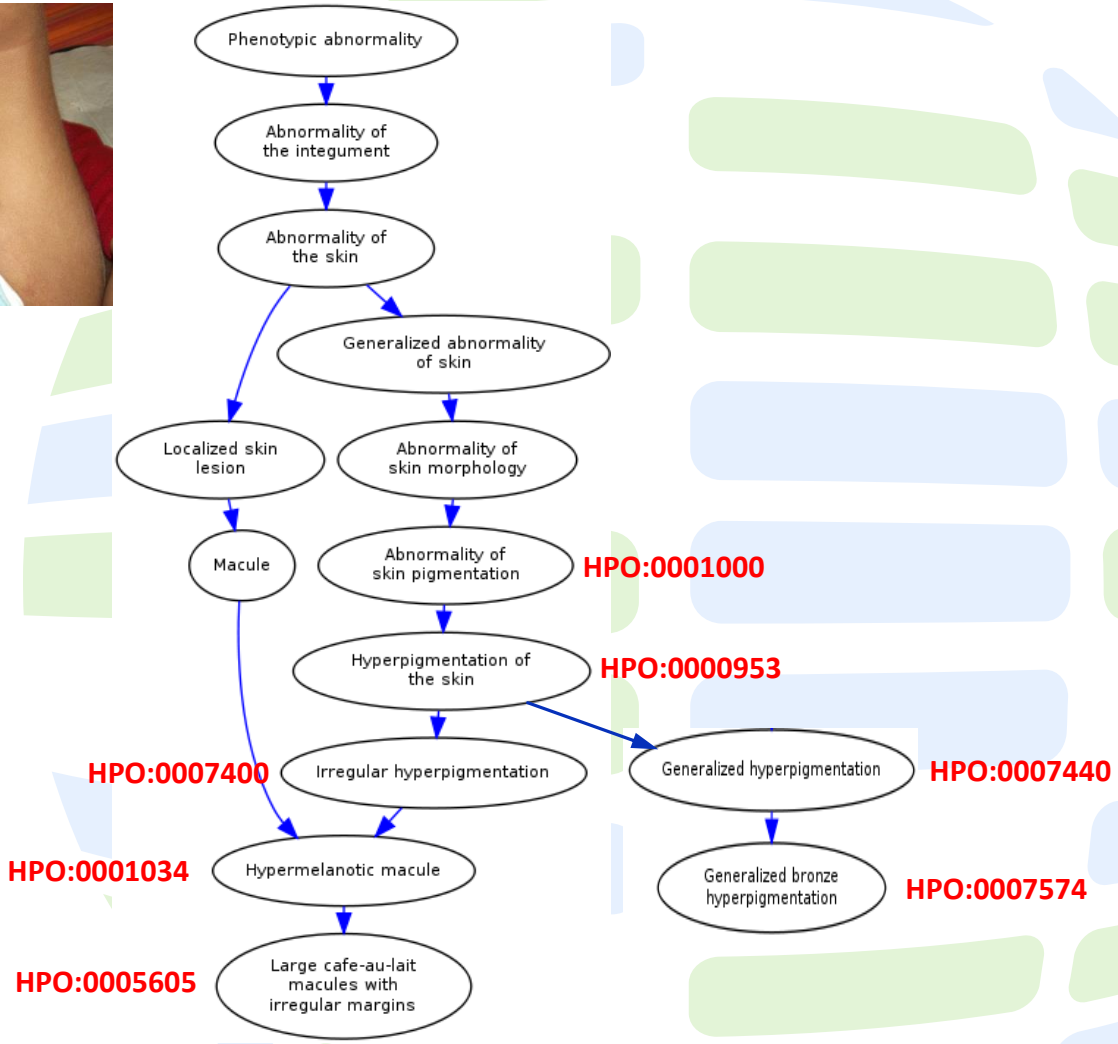













RD Connect

# SpainRDR: Issues regarding the use of ontologies in registry data curation



- Limiting factors for the integration of phenotype ontologies into the RD registries architecture: **Practical examples** extracted from a real case of the IIER Undiagnosed Disease Program (IIER-UDP)

- Five years old child with the following phenotype:
  - ✓ Hypothyroidism
  - ✓ Intellectual disability
  - ✓ Delayed speech and language development
  - ✓ Delayed motor development
  - ✓ Short stature
  - ✓ Hyperpigmentation
  - ✓ Microcephaly
  - ✓ Microretrognathia
  - ✓ Trigonocephaly
  - ✓ Macrotia
  - ✓ Low IGF-1 blood levels
  - ✓ Low $Fe^{2+}$ blood levels

Temporal evolution of a phenotype (limiting factor)

EXAMPLE 1

Accuracy in descriptions of phenotypic features (limiting factor)

EXAMPLE 2

# Useful Resources

# Undiagnosed Cases Program – SpainUDP, IIER: Software to help to the diagnosis

Instituto de Salud Carlos III

Phenolyzer

PHENOTIPS


PhenomeCentral

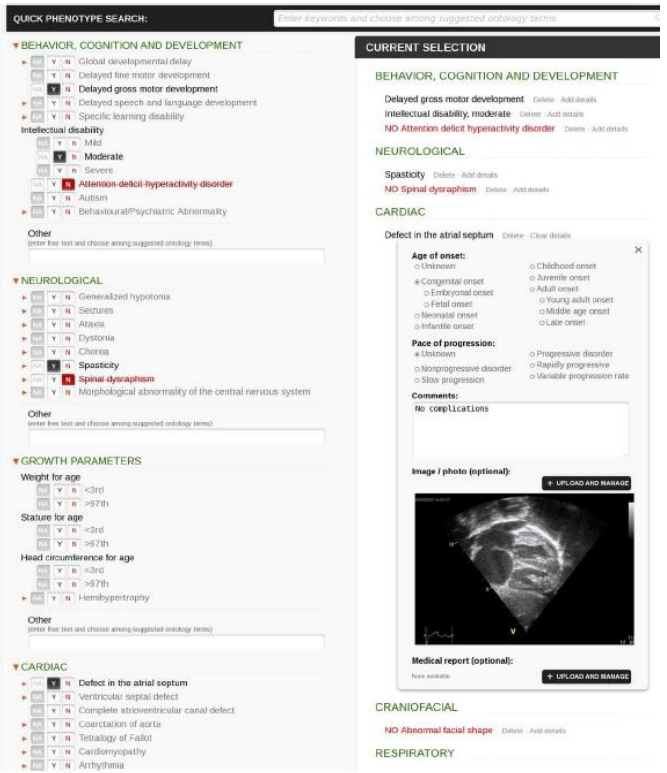
Phenomizer

Others…

# Undiagnosed Cases Program
# SpainUDP - IIER

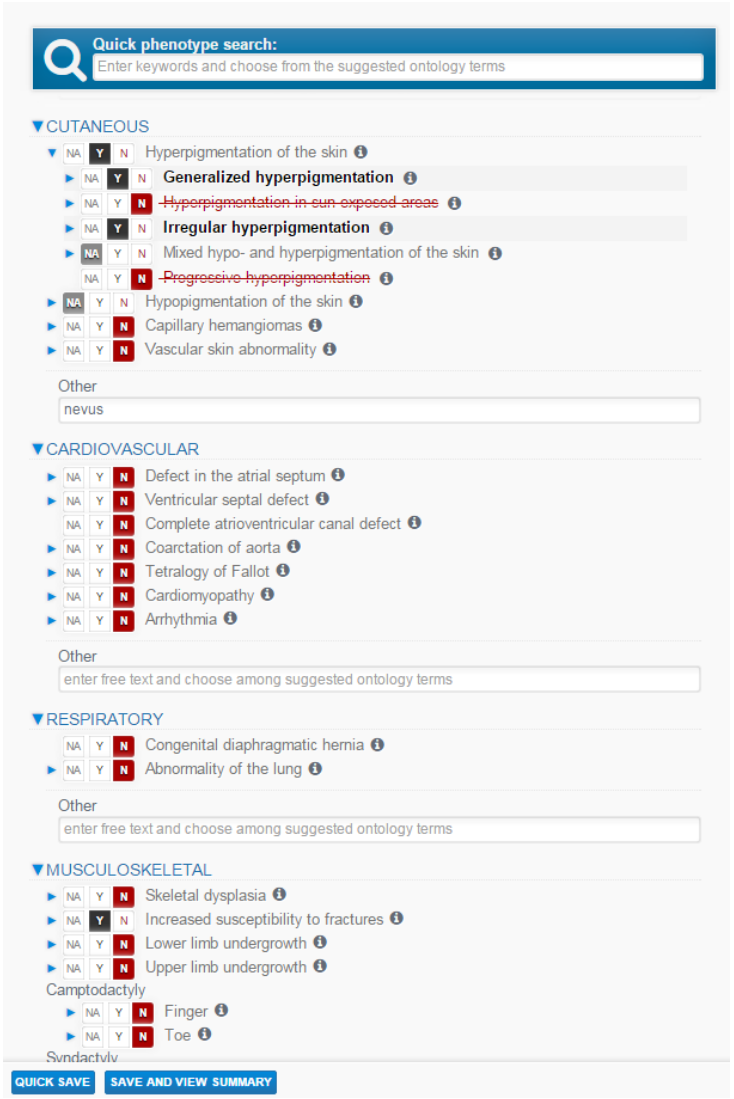
## EQUIPO MULTIDISCIPLINAR

- Clínica
- Genética
- Biología celular y molecular
- Anomalías congénitas
- Bioinformática



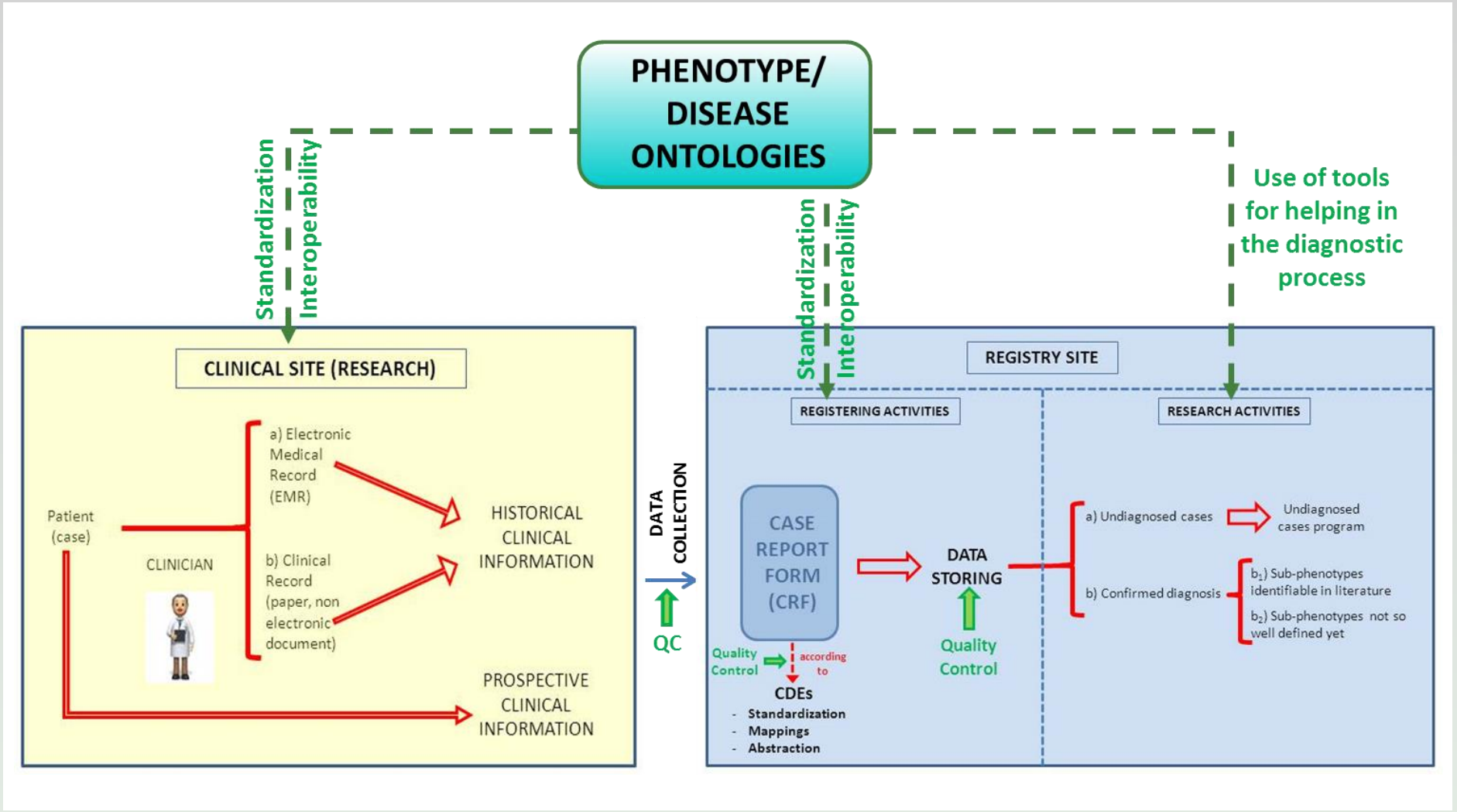

## SOFTWARES DE AYUDA AL DIAGNÓSTICO

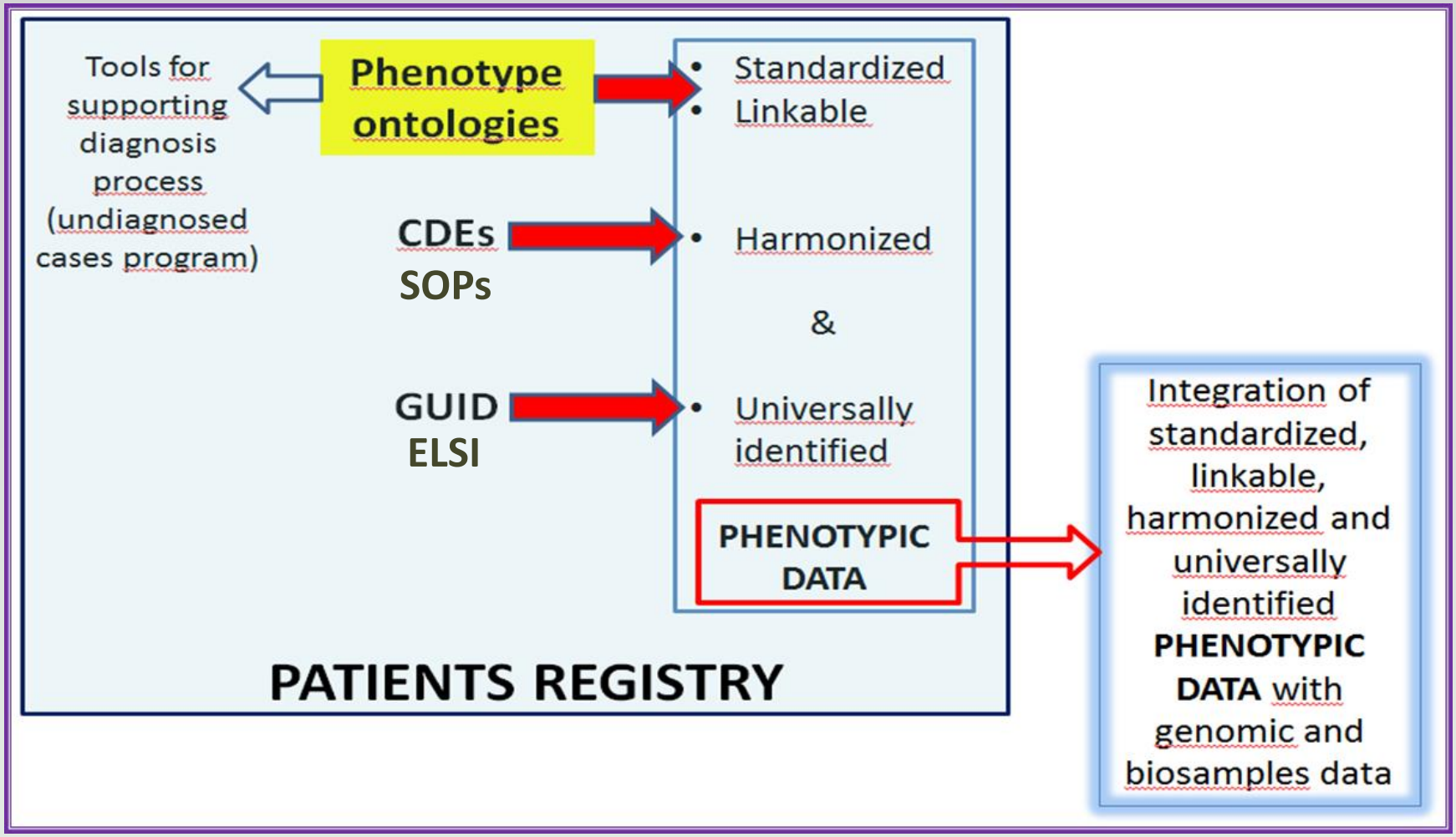
# Present vs Absent Features




CUTANEOUS

Hyperpigmentation of the skin
NO Capillary hemangiomas
NO Vascular skin abnormality
Generalized hyperpigmentation
Irregular hyperpigmentation
Nevus
NO Absent skin pigmentation
NO Deep-set nails
NO Thin nail
NO Woolly hair
NO Concave nail
NO Fragile nails
NO Hyperconvex thumb nails
NO Partial albinism
NO Lumbosacral hirsutism
NO Lack of skin elasticity
NO Piebaldism
NO Hyperpigmentation in sun-exposed areas
NO Epidermal nevus
NO Ichthyosis
NO Abnormal blistering of the skin
NO Progressive hyperpigmentation
NO Generalized hypopigmentation


PHENOTYPE/
DISEASE
ONTOLOGIES
Standardization
Interoperability
Standardization
Interoperability
Use of tools
for helping in
the diagnostic
process
CLINICAL SITE (RESEARCH)
Patient
(case)
CLINICIAN
a) Electronic
Medical
Record
(EMR)
b) Clinical
Record
(paper, non
electronic
document)
HISTORICAL
CLINICAL
INFORMATION
PROSPECTIVE
CLINICAL
INFORMATION
DATA
COLLECTION
QC
REGISTRY SITE
REGISTERING ACTIVITIES
RESEARCH ACTIVITIES
CASE
REPORT
FORM
(CRF)
Quality
Control
according
to
CDEs
- Standardization
- Mappings
- Abstraction
DATA
STORING
Quality
Control
a) Undiagnosed cases
Undiagnosed
cases program
b) Confirmed diagnosis
b1) Sub-phenotypes
identifiable in literature
b2) Sub-phenotypes not so
well defined yet


Tools for supporting diagnosis process (undiagnosed cases program)
Phenotype ontologies
Standardized
Linkable
CDEs
SOPs
Harmonized
&
GUID
ELSI
Universally identified
PHENOTYPIC DATA
PATIENTS REGISTRY
Integration of standardized, linkable, harmonized and universally identified PHENOTYPIC DATA with genomic and biosamples data

RD Connect

Instituto de Salud Carlos III

# *Thank you for your attention*