Using Semantic Web Technologies to Build a Community-driven Knowledge Curation for the Skeletal Dysplasia Domain (SKELETOME)

Tudor Groza, Andreas Zankl, Yuan-Fang Li and Jane Hunter

27 Oct. 2011
Bonn, Germany
Skeletal dysplasias …

… from the knowledge representation, curation and discovery perspectives …

- Challenges & Requirements
- The SKELETOME Platform
- Next steps
Skeletal dysplasias

Background

- Group of heterogeneous genetic disorders
- 487 types – 40 groups
- Rare (< 1 : 10,000)
- Sparse distribution
Skeletal dysplasias
Background (II)
Skeletal dysplasias
Challenges

- Rare Disorders
- Few established guidelines
- Growing domain knowledge
- Sparseness
- Heterogeneity

Decision support
Integration
Skeletal dysplasias
Current status

ISDS Nosology
(Warman et al, 2011)

European Skeletal Dysplasia Network

International Skeletal Dysplasia Registry

SKELNET

4 years publishing cycle

- Groups
- Bone Dysplasias
  - OMIM
- Genes
  - Locus
- Proteins
Challenges / Requirements
Common terminology

Gorlin/Hall
- Postaxial polydactyly of the hands and feet
- Short long bones with smooth rounded metaphyses
- ‘Drumstick’ appearance of the radius and ulna

ISDR/Lachman
- Feet – Polydactyly – postaxial – bilateral
- Femur – Metaphysis Proximal – rounded
- Femur – Metaphysis distal – rounded
- Radius – Metaphysis (Distal) – widened
Challenges / Requirements
Knowledge transfer and evolution

Filamin Group

Melnick-Needles syndrome
- Distinct facial appearance
- Short stature
- Hypertension
- Progressive hearing loss
- Pronounced bowing of long bones

Osteolysis Group

Hajdu-Cheney syndrome
- Generalized osteoporosis
- Short stature
- Serpentine fibulae
- Cystic kidneys
- Truncating mutation in exon 34 of NOTCH2

(Gray et al., 2011)
SKELETOME
SKELETOME
Overview

- Ontology-based ...
- ... community-driven knowledge curation platform
- Domain knowledge
- Patient repository
SKELETOME
Aims

Knowledge / Data Integration

Decision Support

Privacy and access control

Provenance and expertise

Common terminology

Knowledge transfer and sustained knowledge evolution
SKELETOME
Knowledge engineering cycle

Collaborative Knowledge Curation
Iterative Data Exploration
Collaborative Diagnosis

Decision Support Framework
Entity Extraction
Expertise Modeling
Ontology Evolution

Bone Dysplasia Ontology
++
Annotation Vocabulary

Patient Ontology
Context Ontology

ISWC 2011, October 2011
SKELETOME
Ontology set

(Common terminology & Knowledge / Data integration)
SKELETOME
Ontology set

Bone Dysplasia Ontology
- Bone Dysplasias
- Genotype
- Phenotype

Patient Ontology
- Patients
- Diagnoses
- Investigations

Context Ontology
- Provenance
- Temporality
SKELETOME
Ontology set

Bone Dysplasia Ontology

1,200+ concepts

Bone Dysplasia
Perlecan Group
characterized_by

Phenotype Information
Phenotypic Characteristic
NCI ... HPO

Genotype information
Gene Mutation Gene Protein

characterized_by
SKELETOME
Ontology set

Bone Dysplasia Ontology
SKELETOME
Ontology set

Bone Dysplasia Ontology

Diagram showing relationships between Gene, Protein, Bone Dysplasia, and Mutation.

Properties:
- accession_no
- omim_no
- umls_cui
- uniprot_id
- locus
- entrezgene_id
- mesh_id
- ref_seq

Relationships:
- Gene is_encoded_by Protein
- Gene involves Mutation
- Bone Dysplasia characterised_by Gene
- Mutation mutation_type NCI: Mutation Abnormality
SKELETOME
Ontology set

Bone Dysplasia
Ontology

Diagram showing relationships between Bone Dysplasia, Phenotypic Characteristic, REAMS: Abnormality, PATO: Quality, and FMA: Anatomical Entity.

Has quality
Describes
Characterised by
Unit of

UO: Unit
REAMS: Abnormality
PATO: Quality
FMA: Anatomical Entity
HP: Phenotypic Abnormality
Class: **Achondroplasia**

SubClassOf:

- characterized_by only (GM000001 or GM000361 or HP_0000238 or HP_0002938)

SubClassOf:

- mode_of_inheritance only HP_0000006 and
- mode_of_inheritance some HP_0000006
SKELETOME
Ontology set

Bone Dysplasia Ontology

Patient Ontology

- Bone Dysplasia
- Patient
- Investigation
- Gene Mutation
- Diagnosis
- Observation
- Phenotypic Characteristic
- asserts
- carries
- exhibit
- has
- shows
SKELETOME
Ontology set

Bone Dysplasia Ontology

Patient Ontology

Context Ontology

Provenance

State Context

Investigation

Gene Mutation

Bone Dysplasia

Patient

Diagnosis

Observation

has

asserts

shows

carries

exhibits

asserts
SKELETOME

Community features

(Knowledge transfer and evolution)
A technical perspective
Contributions

- NCBO Annotator Wrapper
- Fully customizable

NCBO Services

- Vocabulary import
- Ontology 2 taxonomy
- Modularization

Bone Dysplasia Ontology

- Knowledge import
- Ontology 2 content

- Ontology evolution
- RDF backend

Bone Dysplasia Ontology

Vocabulary import

Ontology 2 taxonomy

Modularization

Ontology evolution

RDF backend
A technical perspective
RDF in Drupal 7 ++

- Core Drupal 7
- Drupal content to RDF via instances
- SPARQL support
- Start from an existing ontology
- Ontology content to Drupal
- Concept-based mapping
- Relations vs. class axioms
Collaborative knowledge curation

- Wiki-style editing

Achondroplasia

**Disease characteristics.** Achondroplasia is characterized by abnormal bone growth, with disproportionately short arms and legs, a large head, and characteristic nose and midface hypoplasia. In infancy, hypotonia is typical, and acquisition of skills often delayed. Intelligence and life span are usually normal, although some have upper airway obstruction increases the risk of death in infancy.

**Diagnosis/testing.** Achondroplasia can be diagnosed by characteristic clinical features and radiographic findings. Molecular genetic testing can be used to detect a muta

detected in most affected individuals and is available in clinical laboratories.

**Management.** Recommendations for management of children with achondroplasia include measures to avoid obesity, MRI or CT for early detection of spinal cord compression, adenotonsillectomy, continuous positive airway pressure mask, and tracheostomy to correct obstructive sleep apnea; suboccipital decompression; lower-limb hyperreflexia or clonus and central hypopnea; surgery to correct scoliosis and support in socialization and school adjustment.
Collaborative knowledge curation

- Wiki-style editing
- Micro-contributions

Achondroplasia

**Disease characteristics.** Achondroplasia is characterized by abnormal bone growth, particularly with disproportionately short arms and legs, a large head, and characteristic facial features. In infancy, hypotonia is typical, and acquisition of motor skills is often delayed. Intelligence and life span are usually normal, although complications such as upper airway obstruction increases the risk of death in infancy.

**Diagnosis/testing.** Achondroplasia can be diagnosed by characteristic clinical features and is most often confirmed by radiological evaluation of the long bones, skull and spine, as well as through genetic testing. Molecular genetic testing can be used to detect a mutation in the fibroblast growth factor receptor 3 (FGFR3) gene.

**Management.** Recommendations for management of children with achondroplasia include weight and height monitoring, measures to avoid obesity, MRI or CT for signs of spinal cord compression, adenoidectomy, continuous positive airway pressure (CPAP) or a mask and tracheostomy to correct obstructive sleep apnea; suboccipital decompression, lower-limb hyperreflexia or clonus and central hypoventilation; surgery to correct scoliosis, support in socialization and school adjustment.

**Investigation**

Cervical spine MRI with CSF flow studies is the best investigation to assess symptomatic craniocervical junction compression in children with Achondroplasia.
Collaborative knowledge curation

- Wiki-style editing
- Micro-contributions
- Direct domain knowledge “encoding”
Collaborative knowledge curation

- Wiki-style editing
- Micro-contributions
- Direct domain knowledge “encoding”

- Alter the ontology structure
- Alter class axioms

Bone Dysplasia Ontology

Investigation

Cervical spine MRI with CSF flow studies is the best investigation to assess symptomatic craniocervical junction compression in children with Achondroplasia.

Posted By admin
Knowledge transfer

Bone Dysplasia Ontology

Annotation Vocabulary

Context Ontology

Patient Ontology

Decision support

Micro-contributions

Collaborative Diagnosis

Time
Semantic annotation

Create Clinical summary

<table>
<thead>
<tr>
<th>Title</th>
<th>Clinical summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tags</td>
<td>Craniosynostosis, Hypercalciuria, Nephrocalcinosis, Delayed skeletal maturation, Limited elbow movement, Pes planus, Tapered fingers</td>
</tr>
</tbody>
</table>

She has a combination of craniosynostosis, hypercalciuria, nephrocalcinosis, osteopenia, delayed bone age, and unusual hands and feet. She has limited elbow mobility, but hyperextensibility of the digits apart from the thumbs and flat feet. I have attached a number of x-rays but particularly note the cone epiphyses and the short proximal fifth metacarpals. She has quite long tapering fingers with digitalised thumbs. She is unable to flex the left thumb.

HP:HP_0001363
HP:HP_0002150
HP:HP_0001182
Semantic annotation

Edit image Image 24

TITLE
Image 24

DESCRIPTION

Path: p

TEXT FORMAT
- Web page addresses and e-mail addresses turn into links automatically.
- Lines and paragraphs break automatically.

TAGS
Short long bones Flared metaphyses

HP:HP_0001363
HP:HP_0002150
HP:HP_0001182
HP:HP_0003026

ISWC 2011, October 2011
Collaborative diagnosis

- Open discussions
- 5-star rating / voting
- Multiple diagnoses
- Case owner driven
Collaborative diagnosis

- Open discussions
- 5-star rating / voting
- Multiple diagnoses
- Case owner driven
Access control

- Authentication

- Group-based
  - Intra-group sharing
  - Inter-group sharing
Access control

- Authentication
- Group-based
- Individual and role-based

- Sensitive data
- Domain knowledge
Access control

- Authentication
- Group-based
- Individual and role-based
- Data filtration

- Patient data ➔ patient descriptions
Decision support

Aims

A ranked list of features given a particular disorder

A probabilistic correlation ranking of a set of features given a disorder

A ranked list of disorders given a particular feature

A probabilistic correlation ranking of a set of disorders given a set of features
(HP:Lordosis, **HP:Arachnodactyly**) → (BDO:Achondroplasia, 65%)
(HP:Hydrocephalus, **HP:Tapered Fingers**) → (BDO:Achondroplasia, 80%)

**Extract Association Rules**

**SKELETOME Domain Knowledge**

**Exploit Domain Ontologies**

**Ranked Correlations**

HP:Arachnodactyly subClassOf HP:Slender Fingers
HP:Slender Fingers subClassOf HP:Abnormality of the Fingers
HP:Tapered Fingers subClassOf HP:Abnormality of the Fingers

BDO:Achondroplasia characterizedBy HP:Trident Abnormality
SKELETOME

Future work
Future work
Ontologies

- Bone Dysplasia Ontology
  - Protein structures
- Ontology population
  - Mining genotype-phenotype associations
  - Mining phenotype-disease associations
- Linked Data
  - Proper publishing
Future work

Community

- **Integration with existing resources**
  - International Skeletal Dysplasia Registry
  - European Skeletal Dysplasia Network
Future work
Development

- X-Ray tagging & clinical summaries annotation
  - Mining composite tag structures
- Advanced ontology evolution support
  - Protégé 4 as back-end

Very short and tapered middle phalanges of the thumb

Anatomical entity part of Anatomical entity

Quality Quality
Summary

Thank you!

Existing domain knowledge

Collaboratively acquired patient cases
Collaborative diagnosis

Knowledge evolution
Decision support
Iterative data exploration