

# Open Biomedical Ontologies Applied to Prostate Cancer

**James A. Overton, Cesare Romagnoli, Rethy Chhem**

Department of Medical Imaging  
London Health Sciences Centre

Joseph L. Rotman Institute of Science and Values  
University of Western Ontario

International Conference on Biomedical Ontology  
July 24, 2009

## Interdisciplinary Prostate Ontology Project

- ▶ develop expertise
  - ▶ practical applications
  - ▶ improve communication
  - ▶ improve patient outcomes
- 
- ▶ reporting, search, and analysis
  - ▶ radiology, surgery, pathology, oncology, anatomy

## TOP First Stage

- ▶ creating database-backed website
- ▶ annotating textual reports
- ▶ looking for good fits
- ▶ looking for mismatches and gaps

- ▶ Controlled Vocabularies
- ▶ OBO Ontologies
- ▶ Annotating Reports
- ▶ Gaps in OBO Ontologies
- ▶ Filling Gaps
- ▶ Conclusions

## Controlled Vocabularies

## Systematized Nomenclature of Medicine – Clinical Terms

## Digital Imaging and Communications in Medicine



## Radiology Society of North America's Radiology Lexicon

## OBO Ontologies

## OBO Ontologies

- ▶ interoperable network
- ▶ division of labour
- ▶ shared best practices
- ▶ shared Basic Formal Ontology
- ▶ permissive licenses
- ▶ open source approach

## Foundational Model of Anatomy

- ▶ human anatomy
- ▶ FMA:9600 “prostate”
- ▶ parts of the prostate
- ▶ neighbouring organs

## Disease Ontology

- ▶ human disease
- ▶ parallels FMA
- ▶ DOID:47 “prostate disease”
- ▶ DOID:514 “prostatic neoplasms”
- ▶ DOID:8634 “carcinoma in situ of prostate”

## Protein Ontology

- ▶ proteins and their relations
- ▶ prostate specific antigen (PSA)

## Gene Ontology

- ▶ cellular component, biological process, molecular function
- ▶ processes related to PSA
- ▶ GO:0004252 “serine-type endopeptidase activity”
- ▶ GO:0016525 “negative regulation of angiogenesis”

## Phenotypic Quality Ontology

- ▶ labelling animal phenotypes
- ▶ works for many qualitative descriptions
- ▶ PATO:0000014 “color”
- ▶ PATO:0000060 “spatial pattern”
- ▶ PATO:0000701 “smooth”



## Units of Measurement Ontology

- ▶ organizes International System of Units (SI)
- ▶ adds terms such as UO:0000190 “ratio”

## Annotating Reports

## Radiology Report Sample

“**Peripheral Zone:** This zone is relatively homogeneous with a **smooth contour** although it is compressed by a large **transition zone.**”

## Radiology Report Sample

- ▶ “peripheral zone” corresponds to FMA:19587 “peripheral zone of prostate”
- ▶ “smooth” is PATO:0000701
- ▶ “contour” roughly corresponds to PATO:0000052 “shape”
- ▶ “transition zone” corresponds to FMA:45721 “transition zone of prostate”

## Surgery Report Sample

“Once the **prostate** was mobilized in a cephalad direction, I could see **Denonvilliers fascia**. This was opened in the midline. We then dissected out the **ampulla of Vater**, which were clipped and divided. The **seminal vesicles** were dissected off in their entirety quite easily using clips for **hemostasis**.”

## Surgery Report Sample

- ▶ “prostate” is FMA:9600
- ▶ “Denonvilliers fascia” is a synonym for FMA:19933  
“rectovesical septum”
- ▶ “ampulla of Vater” is a synonym for FMA:15076  
“hepatopancreatic ampulla”
  - ▶ did the author intend FMA:19259 “ampulla of deferent duct”?
- ▶ “seminal vesicle” is FMA:19386
- ▶ “hemostasis” is GO:0007599

## Pathology Report Sample

“The specimen consist of 2 cores of **pale tan tissue**, the larger measures 1.3 **cm** and the smaller measures 1.1 **cm**. All **tissue** is submitted in one cassette.”

## Pathology Report Sample

- ▶ “pale tan” is a close synonym of PATO:0001268 “desaturated brown”
- ▶ “tissue” corresponds to FMA:9637 “portion of tissue”
- ▶ “cm” is UO:0000015



## Gaps in OBO Ontologies

## Radiology Report Gaps

“Peripheral Zone: This zone is **relatively homogeneous** with a smooth contour although it is **compressed** by a **large** transition zone.”

## Surgery Report Sample

“Once the prostate was **mobilized** in a **cephalad direction**, I could **see** Denonvilliers fascia. This was **opened** in the **midline**. We then **dissected** out the ampulla of Vater, which were **clipped** and **divided**. The seminal vesicles were **dissected** off in their **entirety quite easily** using **clips** for hemostasis.”

## Pathology Report Gaps

“The **specimen consist** of 2 **cores** of pale tan tissue, the **larger measures** 1.3 cm and the **smaller measures** 1.1 cm. All tissue is **submitted** in one **cassette**.”

## Near Synonyms

- ▶ term  $T$  is not in any ontology
  - ▶ a synonym  $S$  exists in some ontology
  - ▶ but  $T$  is not listed as a synonym of  $S$
  - ▶ **example:** “pale tan” is not in PATO
- 
- ▶ relatively easy for humans to detect
  - ▶ difficult for machines to detect
  - ▶ either add to ontology
  - ▶ or change practice

## Missing Composites

- ▶ terms  $S$  and  $T$  are in the ontology
- ▶ but composite term  $ST$  is not
- ▶ **example:** “nanogram per millilitre” is not in UO
  
- ▶ either add  $ST$  to the ontology
- ▶ or build the composite  $S + T$  using a relation

## Conflicting Fiat Boundaries

- ▶ *bona fide* boundaries are easy to agree upon
- ▶ *fiat* boundaries vary by convention and application
- ▶ **example:**
  - ▶ FMA lobes of prostate: anterior, posterior, right lateral, left lateral
  - ▶ RadLex divisions: outer and inner glands; peripheral, central, and transition zones
  - ▶ RadLex divisions are more useful in locating prostate tumours
- ▶ difficult to resolve

## Medical Procedures

- ▶ missing many terms for medical procedures
- ▶ not covered by biology-oriented ontologies
- ▶ **examples:**
  - ▶ digital rectal exam, transrectal ultrasound
  - ▶ surgery, mobilize, dissect, clips
  - ▶ biopsy, specimen, core, fragment, cassette
- ▶ some will be covered by the Ontology for Biomedical Investigations (OBI)
- ▶ others will have to be created



## Filling Gaps

## Medical Imaging Modalities

- ▶ Medical Image
  - ▶ Magnetic Resonance Imaging (MRI) Image
  - ▶ Ultrasound (US) Image
  - ▶ X-Ray Image
  - ▶ Nuclear Medicine Image

## Medical Imaging Modalities

- ▶ Magnetic Resonance Imaging (MRI) Image
  - ▶ T1 Weighted MRI Image
    - ▶ MRI Image without Contrast
    - ▶ MRI Image with Contrast
  - ▶ T2 Weighted MRI Image
  - ▶ Proton Density Weighted MRI Image

## Radiology Qualities

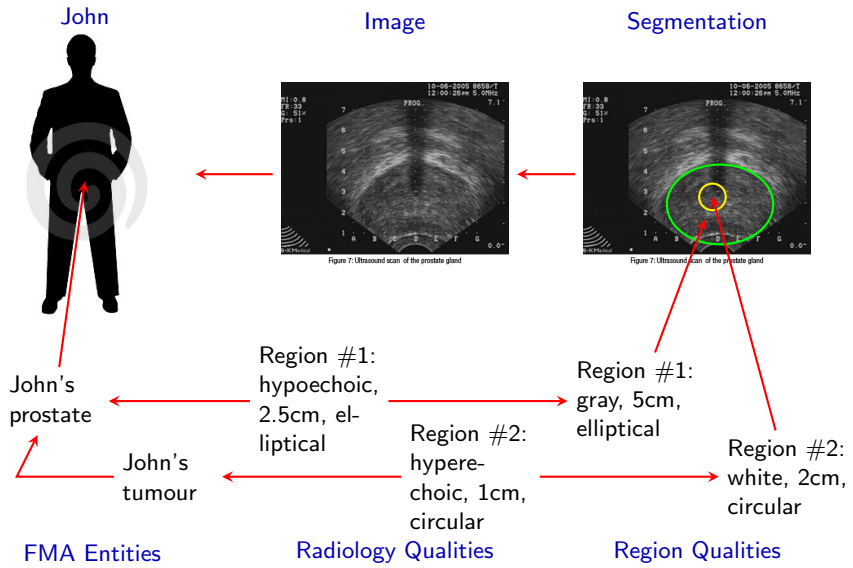
- ▶ intensity (MRI)
- ▶ echogenicity (ultrasound)
- ▶ transparency (X-Ray)
- ▶ density (CT)

## Relativity of Qualities

- ▶ nearby tissue
- ▶ change in time
- ▶ change with contrast enhancement
- ▶ change with sequence type

## Radiology Reporting

- ▶ clinical information – EHR standards, HL7?
- ▶ technique of examination – DICOM, OBI, Image Ontology?
- ▶ description of findings – FMA, DO, Image Ontology?
- ▶ conclusions – FMA, DO, OBI



## Conclusions



## Conclusions

- ▶ Annotation is very labour intensive, which limits application.
- ▶ Many reporting terms are already in OBO, but many are not.
- ▶ Current OBO ontologies focus on biomedical research, not medical practice.

## Conclusions

We see a need for ...

- ▶ more ontologies focused on medical practice
- ▶ more efficient annotation tools
- ▶ more work on annotating images