Development of a Specialist Medical Vocabulary For Radiation Oncology

Dr Andrew Miller
Senior Staff Specialist in Radiation Oncologist
Illawarra Cancer Care Centre
The Wollongong Hospital

Dr Ping Yu
Health Informatics Research Centre
School of Information and Computer Science
University of Wollongong
Australia

illawarra

Specialist Medical Vocabulary For Radiation Oncology

- Radiation Oncologists (RO)
  - medical doctors
  - cancer specialists
  - all use radiation to treat cancer conditions
- Oncology Information Systems (OIS)
  - verifies & records radiation machine use
  - also has attached Electronic Medical Record
    - » The 'OIS-EMR'
- Electronic Health Record (EHR) perspective
  - our data needs to be represented
  - our information will be transferred
  - our meaning needs to be preserved

illawarra

Miller, A.A. & Yu, P.
Wollongong, NSW Australia

Specialist Medical Vocabulary For Radiation Oncology

- <u>Informatics requirements</u>
  - Limited subspecialty area
    - Affects 25-30% of population
  - Specific data needs
    - Date-specific events
      - Necessary for reporting
    - Decisions based on a defined finite dataset
      - » DIAGNOSIS*
      - » STAGING
      - Informs discrete treatment decision
      - Produces idiosyncratic therapy sequences
    - Focused on Outcomes
      - Measured side effects of therapy & control of disease
      - Informs prognostication & therapy choices

illawarra CANCER CARE CENTRE

* Most medical specialties start with a Problem or Complaint

Miller, A.A. & Yu, P.
Wollongong, NSW Australia

illawarra

Specialist Medical Vocabulary For Radiation Oncology

- The "competent" OIS-EMR
  - should be based on representation of medical knowledge of Radiation Oncologists
    - Vocabulary (the right words)
    - Nomenclature (the right connections)
    - Ontology (the right framework)
  - all should be seen in ERD construction of databases

- BUT …. no vocabulary has been defined!

  - **Problem: Just ask the 'experts', or find an objective text source/corpus for Vocabulary?**
  - **Solution: objective source for Vocabulary is better**
  - Later: Is this Vocabulary embedded in existing Nomenclatures? Is a new Ontology required?

illawarra

Miller, A.A. & Yu, P.
Wollongong, NSW Australia

Specialist Medical Vocabulary For Radiation Oncology

- There is an objective corpus of text
  - the RO literature
    - Read by all ROs
    - Reviewed by many ROs
    - introduces and propagates new concepts
    - Publication defines the contained data as 'important'
  - Define time period
    - Content analysis
      - "text deconstruction"
        - » find data elements used in literature construction
        - » list according to workflow
        - » Domain-specific knowledge required

illawarra

Miller, A.A. & Yu, P.
Wollongong, NSW Australia

Specialist Medical Vocabulary For Radiation Oncology

text deconstruction 1.Select article 2.Read 3.Analyse text 4.collate

PATIENTS AND METHODS

Trial

From 1990 to 1995, 918 patients with HNSCC (laryngeal, pharyngeal, nasopharyngeal and oral cavity tumors) were randomly assigned between

ORIGINAL REPORT

of Two Separate Hypoxia Response
ducible factor 2 alpha and carbonic
ociated With Radiotherapy Failure in
er Patients Recruited in the CHART

PATIENTS AND METHODS

Trial

From 1990 to 1995, 918 patients with HNSCC (laryngeal, pharyngeal, nasopharyngeal and oral cavity tumors) were randomly assigned between

And so on …

Protocol_Type
Protocol_Name
Treatment_Intent
Therapy_Modality
Modality_Intent
Total_Radiation_Dose
Dose_Per_Fraction
RT_Start_Date
RT_End_Date …..

all classified in ICD-10
e.g. larynx – C32
"Head & Neck squamous cell carcinoma"

Deconstruct & categorize
- **Diagnosis CANCER SITE category**
- **Diagnosis_classification MORPHOLOGY type**

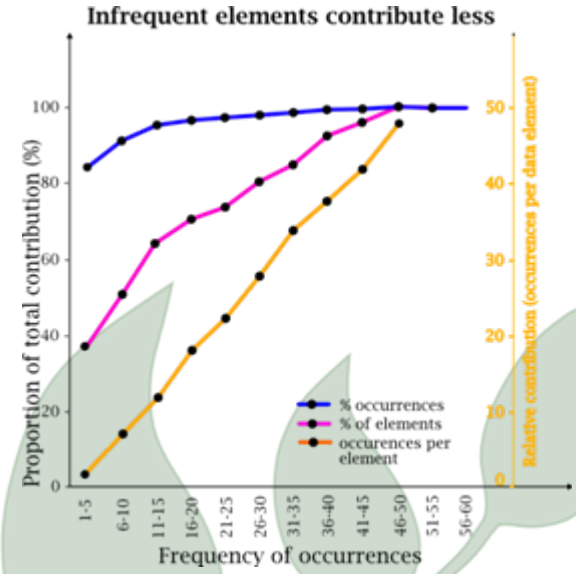
Categorical Data Elements
- **Diagnosis**
- **Morphology confirmation**

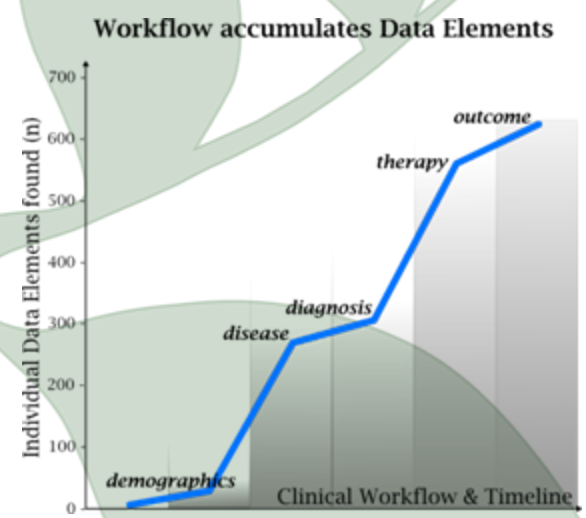
illawarra

Miller, A.A. & Yu, P.
Wollongong, NSW Australia

Specialist Medical Vocabulary For Radiation Oncology

- Content Analysis
  - 97 articles accessed
    - 17 unsuitable (no patient data)
  - 622 data elements found
  - 2392 occurrences of data elements found
    - many infrequent elements
      - Minor contribution to occurrences
    - few frequent elements
      - Major contribution to occurrences
  - Data elements describing patient ("Patient Milieu")
    - 54.8% of data elements
    - 1.69 per element (mean)
  - Data element accumulation is a continuous process





Miller, A.A. & Yu, P.
Wollongong, NSW Australia

illawarra

Conclusions

- This specialist medical vocabulary (SMV) defined manually from the literature can act as a gold standard for later automated efforts
- The SMV of clinical notes should be compared
- The SMV should be mapped to 'standard' nomenclatures and ontologies (e.g., SNOMED-CT)
- OIS-EMRs should be constructed to hold this SMV, <u>and</u> respect their relationships, and allow units to reproduce the same data collection as reported in the literature
  - Reports submitted for publication might be more useful if they listed the data elements used

Miller, A.A. & Yu, P.
Wollongong, NSW Australia

- Further Work
  - discover SMV used in clinical notes/letters
  - map SMV into SNOMED-CT, DICOM-RT & UMLS
  - integrate into medical ontologies

Miller, A.A. & Yu, P.
Wollongong, NSW Australia