Early Report on OCRe:
Ontology of Clinical Research

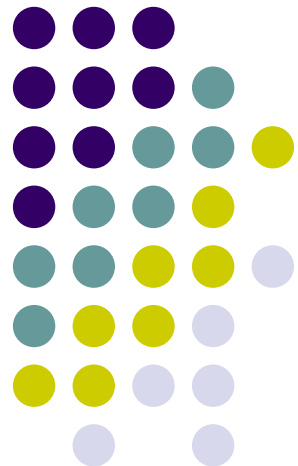
CTSA Symposium

September 28, 2007

I Sim[1], S Carini[1], S Harris[2], P Maccallum[2], A Rector[3], I Toujilov[2], S Tu[4]

[1]UCSF, [2] UK CancerGrid, [3]U Manchester, [4] Stanford

Funded in part by LM06780, MRC-G0100852

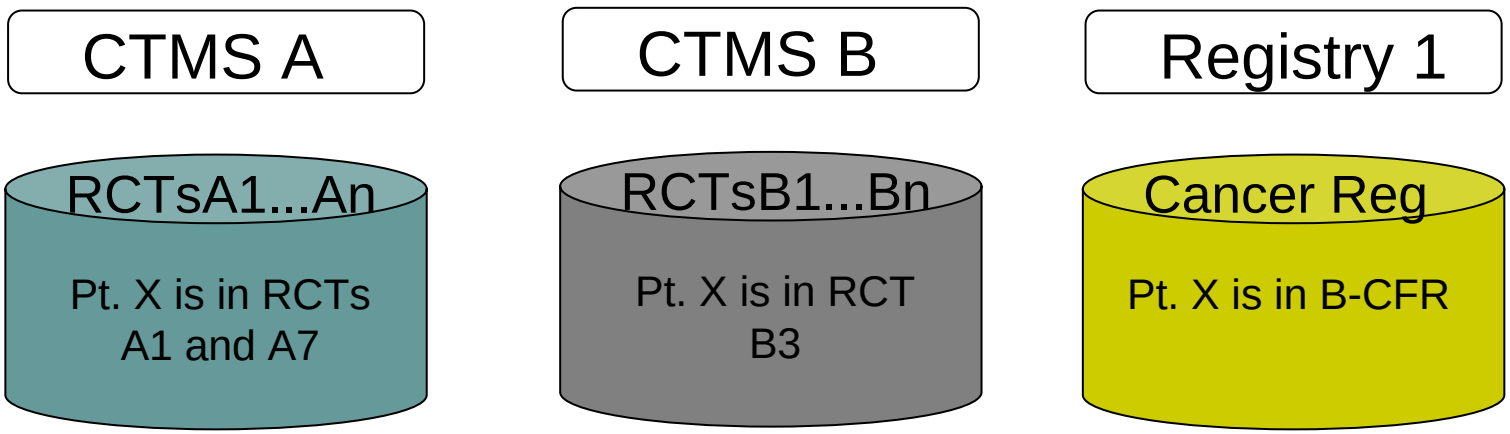


Outline

- Need for Clinical Research Ontology
- Ground Principles
- Scope & Model
- Example: NeoTango Trial
- Conclusion

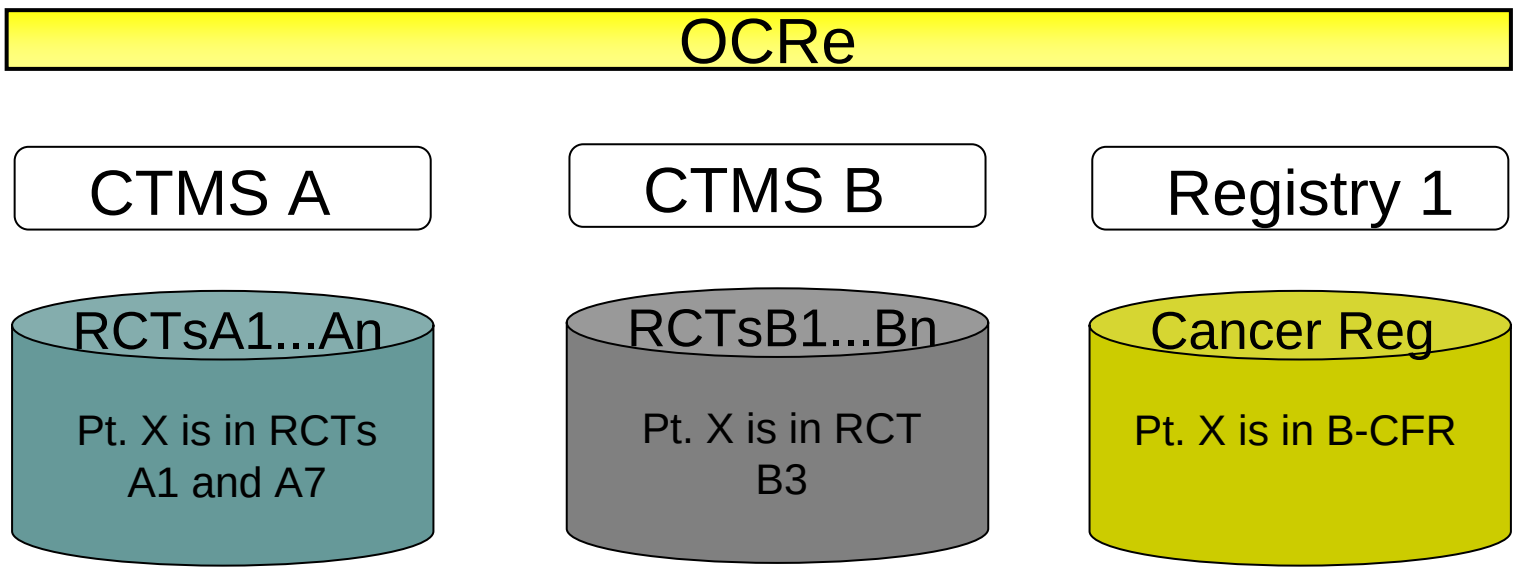
Translational Research Use Cases

- Within CTMS A studies, which randomized placebo-controlled trials testing “glitazones” analyzed liver expression profiles?
- Across all HIV RCTs in CTMSs A&B, what were overall rates of reported adverse events in Study Site A compared to Study Site B?
- Across all trials and the B-CFR Registry, how many Stage III breast CA patients have tissue samples from before and after chemotherapy? who are the PIs for those studies?



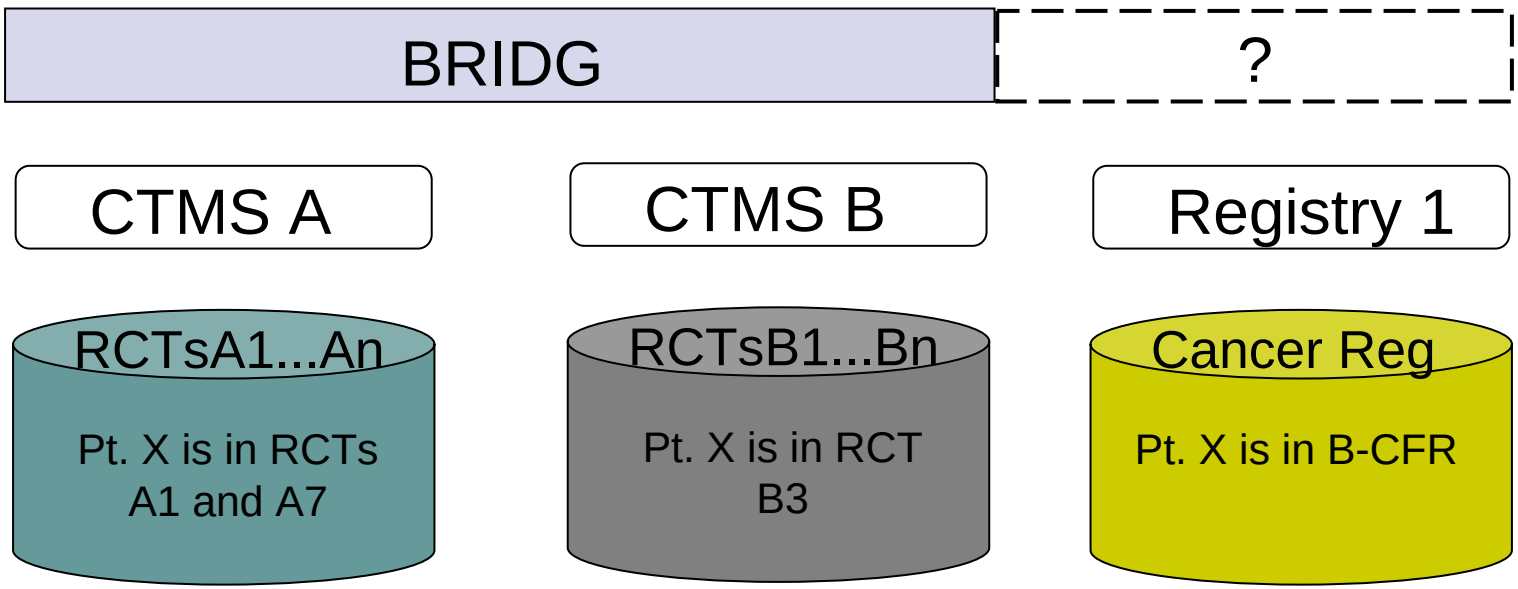
OCRe for Semantic Interoperation

- Systems often use different information models
- For these use cases, we need a terminology and semantics for describing human experiments, independent of information model used



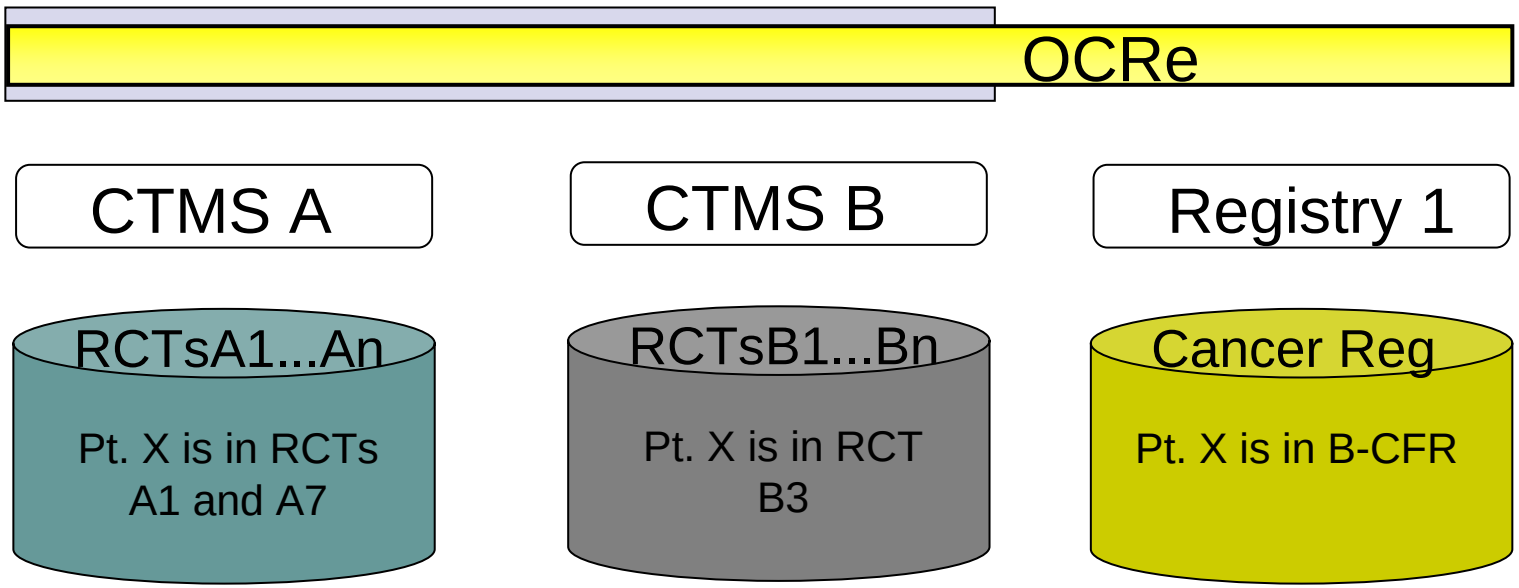
Our Understanding of BRIDG

- For syntactic and semantic interoperation of CTMSs mapped to or built from BRIDG information model
  - primarily for executing and exchanging information about protocol-driven research, modeled at participant level



OCRe and BRIDG Complementary

- OCRe is for semantically integrating data schemas
  - of different studies
  - being executed or reported in different systems
    - that may or may not be BRIDG-compliant



Semantics of OCRe and BRIDG

OCRe

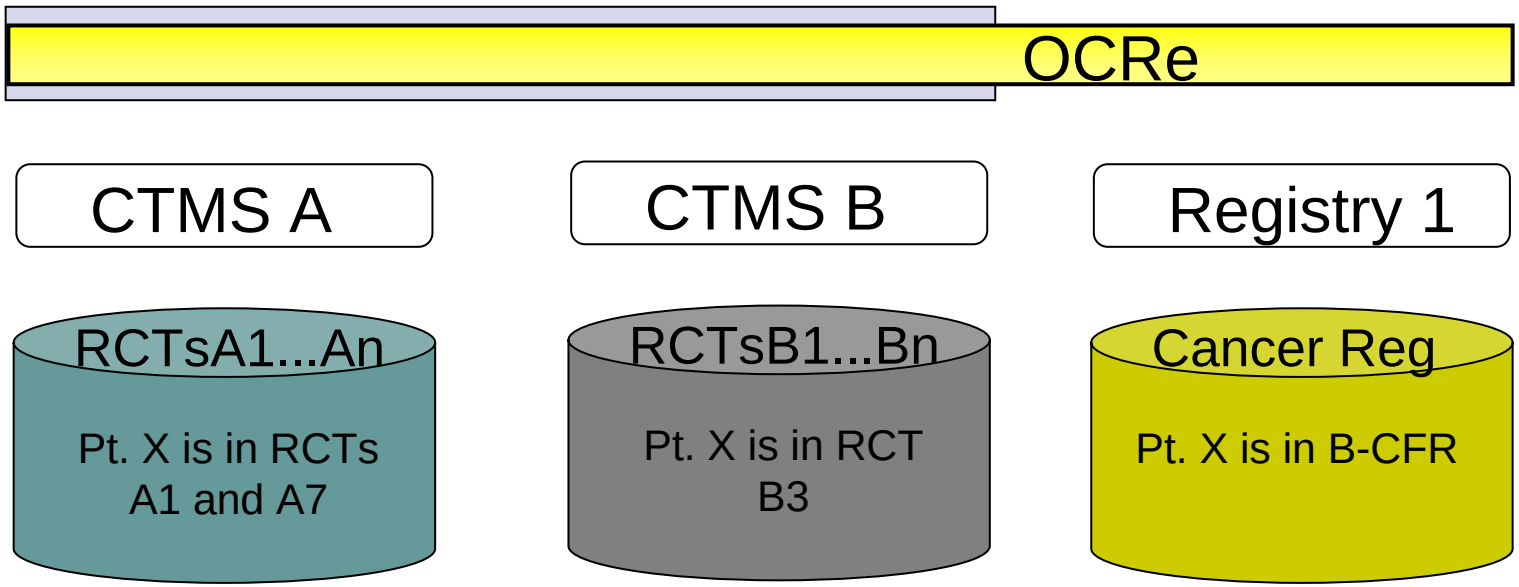
- Primarily for describing studies
  - for indexing, discovering, reasoning & inferencing
- Relationships and constraints
  - formal, in OWL-1.1
  - logical reasoner “for free”
- Independent of information models

BRIDG

- Primarily for running studies
  - for operational interoperation
- Relationships and constraints
  - semi-formal, in text annotations & documentation
  - logical relationships not formalized
- Intertwined with information model

OCRe and BRIDG Complementary

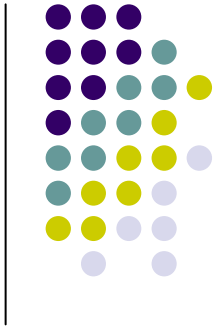
- OCRe provides a formal defined semantics for describing human studies stored in different information models



Outline

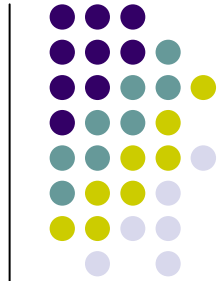
- Need for Clinical Research Ontology
- Ground Principles
- Scope & Model
- Example: NeoTango Trial
- Conclusion

Design Principles



- Basic strategy is “modularize & bind”
  - a core set of modular, related ontologies about the design and methods of human investigations
  - provide defined interfaces for binding to external ontologies and vocabularies as needed
  - provide mappings to external information models (e.g, BRIDG)
- Represent trials as planned *and* as executed
  - i.e., real-world trials with protocol violations, etc.

Development Principles



- Follow best practices and “normalization” principles for ontology construction (e.g., Smith, Ceusters, Rector)
- Iterate ontology based on usefulness/usability for practical applications
- Priority on providing documentation and examples to make OCRe widely useful

Outline

- Need for Clinical Research Ontology
- Ground Principles
- Scope & Model
- Example: NeoTango Trial
- Conclusion

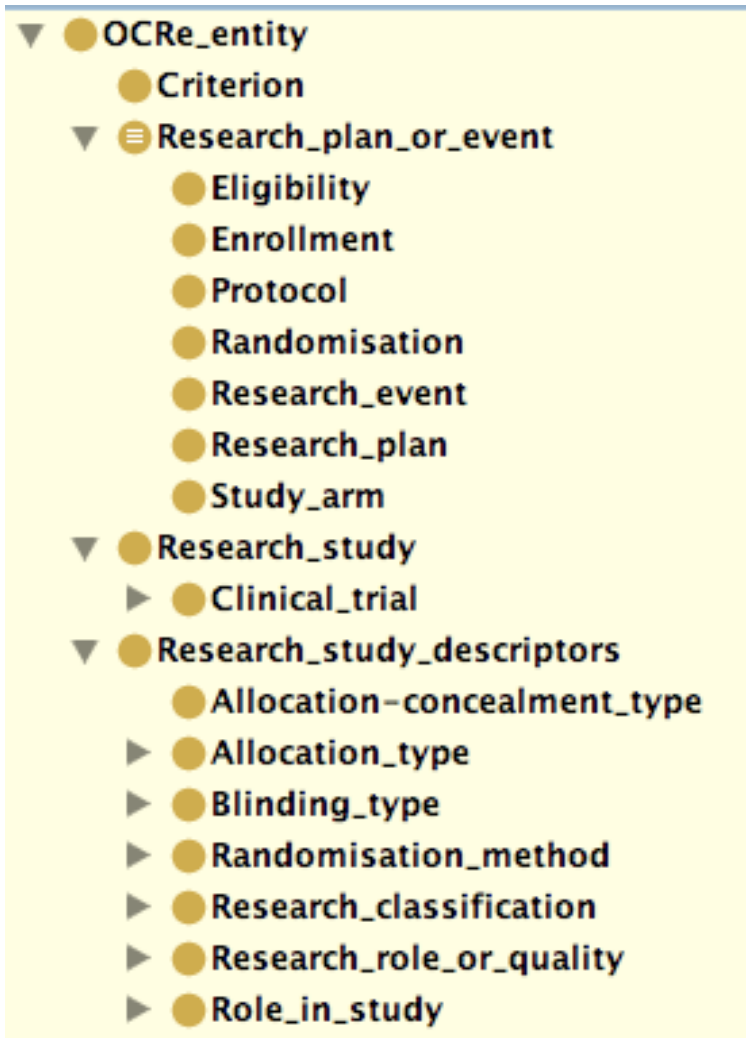
Scope

- All investigations of human subjects
  - interventional studies
    - all phases, randomized or not, multi-arm, crossover, etc.
  - observational studies
    - case control, cross-sectional and longitudinal cohort, etc.
  - any “intent”
    - therapeutic, diagnostic, preventive, etc.
- Any clinical domain
  - cancer, cardiology, endocrinology, etc.
- Any type of data
  - quantitative, qualitative, imaging, genomics, proteomics, etc.

OCRe Core

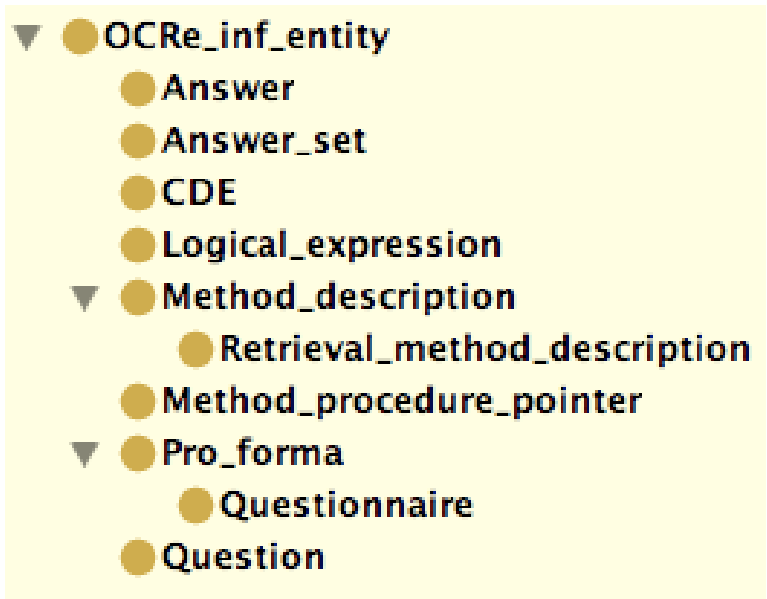
- OCRe has 3 core ontologies, representing
  - human studies, study-level structure and features
    - e.g., randomization, details of control intervention, definition and timepoints of primary and secondary outcomes
  - information artifacts by which information is collected according to the study design
    - e.g., documents, forms, CDEs, questions, answer sets
  - categories of clinical entities
    - linking points for importing clinical entities from external ontologies/vocabularies

Initial Human Studies Ontology



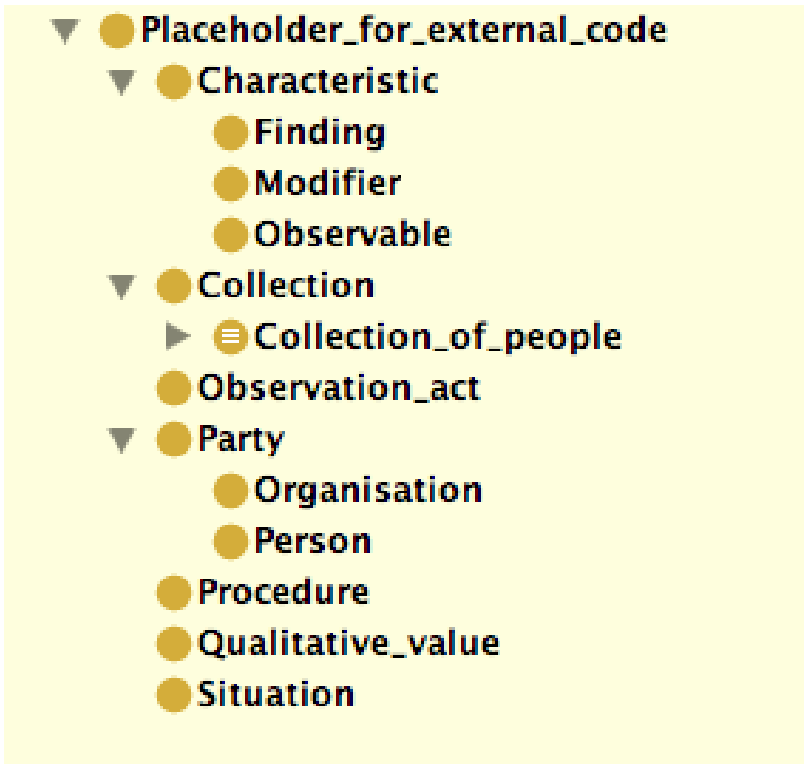
- Entities concerning the design and methods of human experiments
  - criteria (e.g., eligibility)
  - plans & events
  - descriptors
- Will map to upper ontologies (e.g., BFO/OBI, DOLCE) later

Initial Information Artifacts Ontology



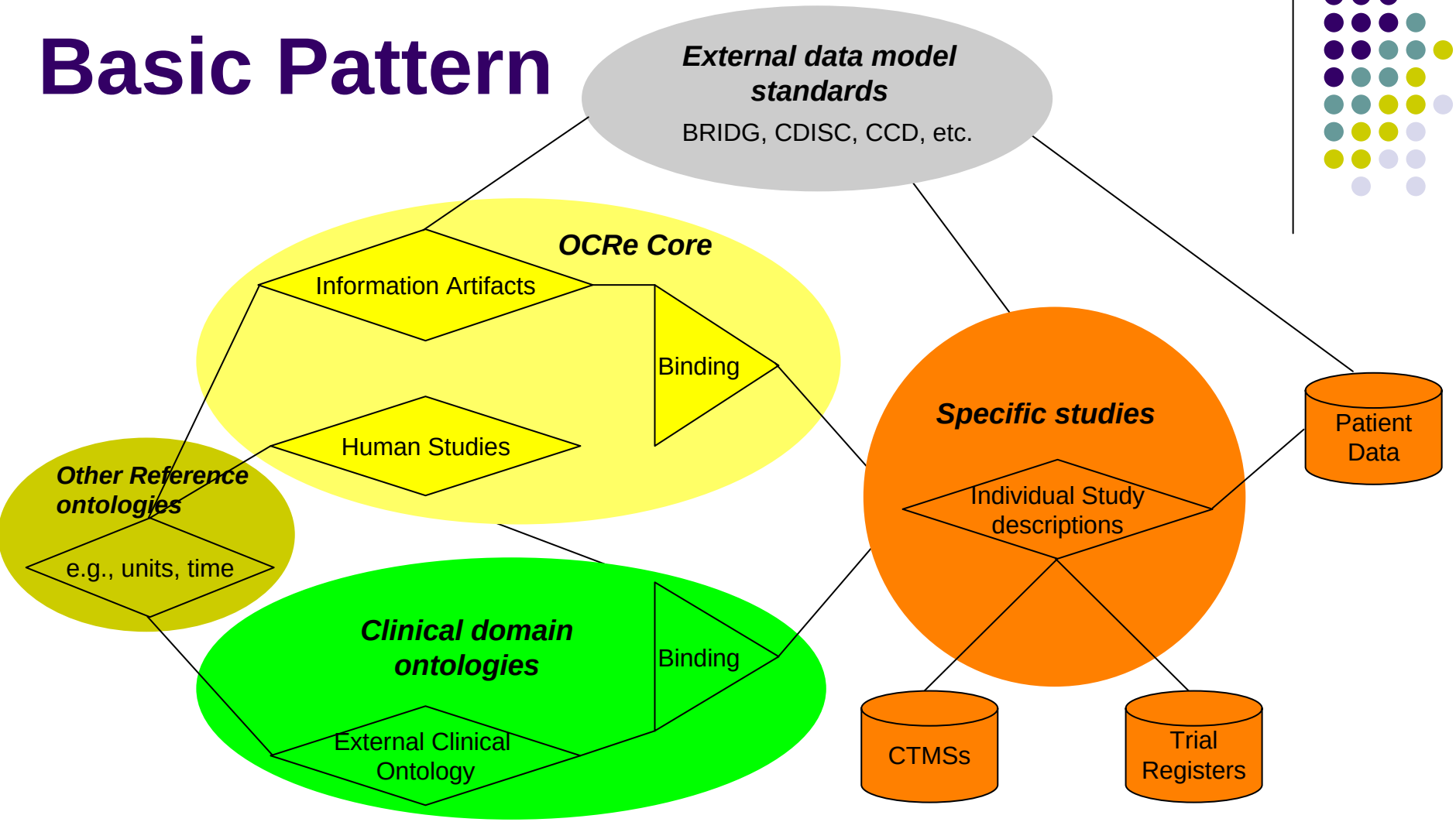
- Entities concerning the information artifacts used in implementing the study
  - pointers to external documents, CDEs, etc.
- …bare beginnings so far…

Initial Interface Ontology



- For importing external clinical ontologies/vocabs
  - e.g., NCIT, SNOMED
- Includes categories of clinical entities useful for OCRe (to be refined)
  - e.g., ocre_interface:Finding
- Using “axioms”
  - e.g. ncit:Neoplasia subclassOf ocre_interface:Finding

Basic Pattern

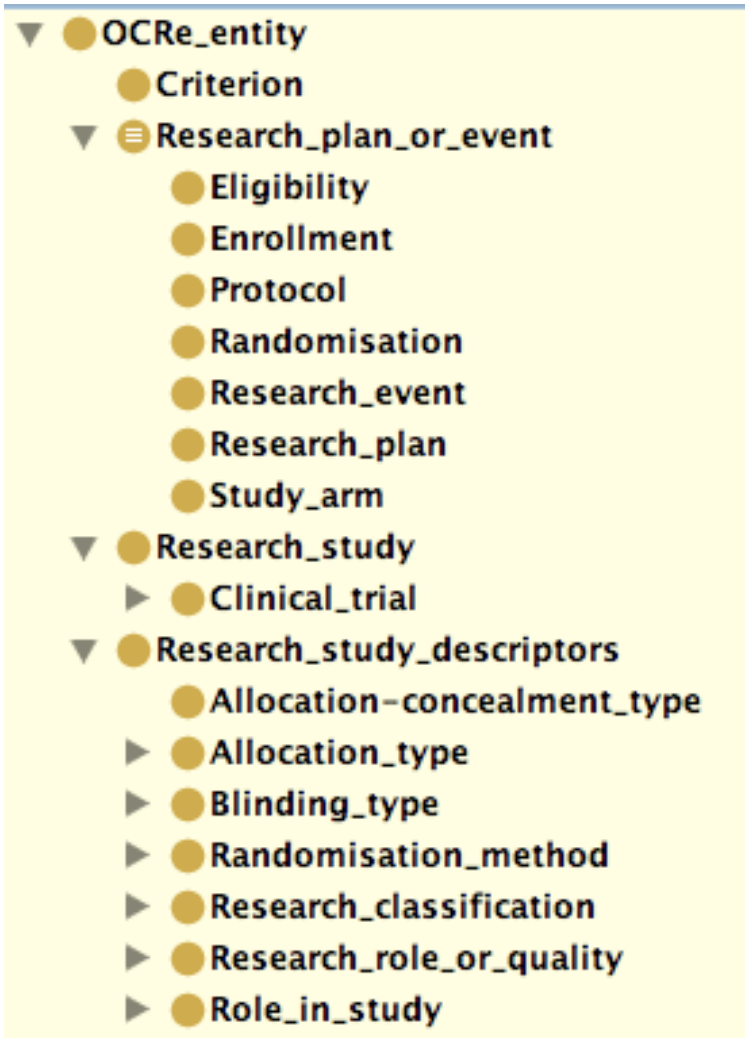


- Different studies may bind to different external ontologies/vocabs
  - e.g., NCIT for breast CA trial, SNOMED for renal trial, ICD9 for HSR study....

Outline

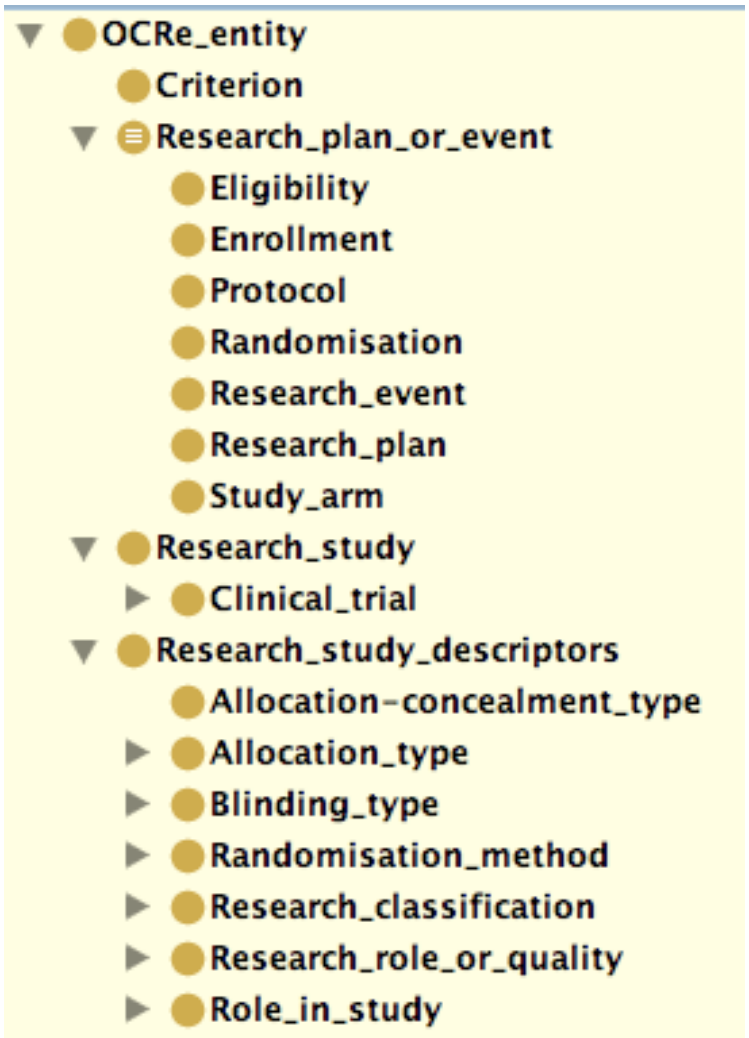
- Need for Clinical Research Ontology
- Ground Principles
- Scope & Model
- Example: NeoTango Trial
- Conclusion

NeoTango Example



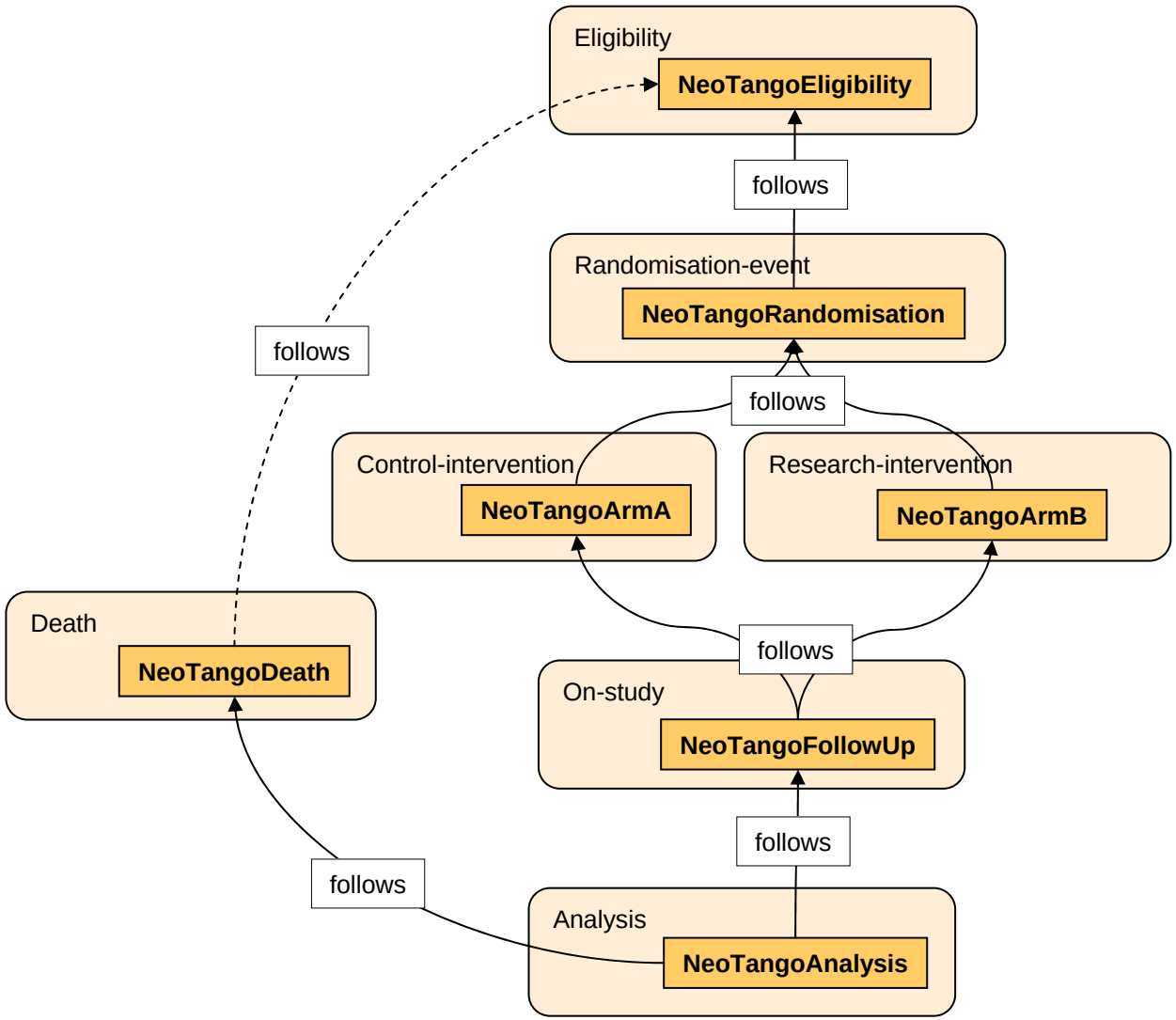
- Studies have topics
  - study:NeoTango
    ocre:has_topic SOME
    ncit:Breast_carcinoma
    ocre:has_control_intervention SOME
    ncit:Adjuvant_chemotherapy
- ...have methods
  - study:NeoTango *ocre:has_method* ocre:Randomization_method
- ...have *intended* plans & *actual* events, which may not match
  - study:NeoTango *ocre:has_intended_step* VALUE study:NeoTango_randomisation

NeoTango Example (cont.)



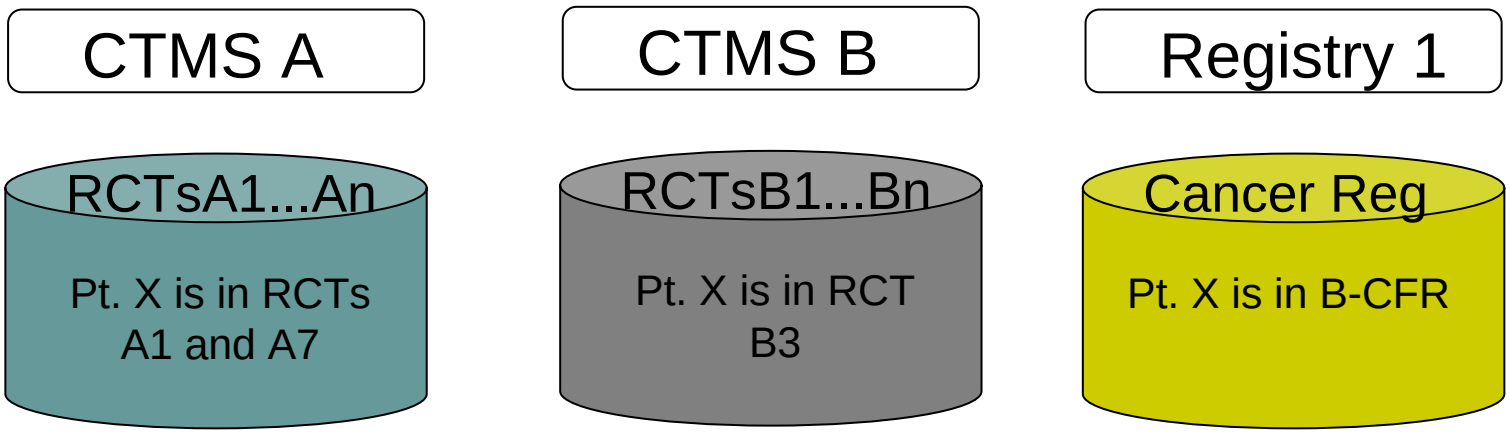
- Criteria have questions, methods, rules, etc.
  - neotango:t2_tumour_criterion *ocre:has_question* VALUE neotango:question2_pointer
- Various situations can be described, e.g. for SOME Person
  - (ocre:Situation THAT *ocre:at_event* ocre:Enrollment_event) *includes* SOME (ncit:Breast_carcinoma THAT *ocre:at_stage* ncit:pT1_Stage_Finding)

NeoTango Flowchart

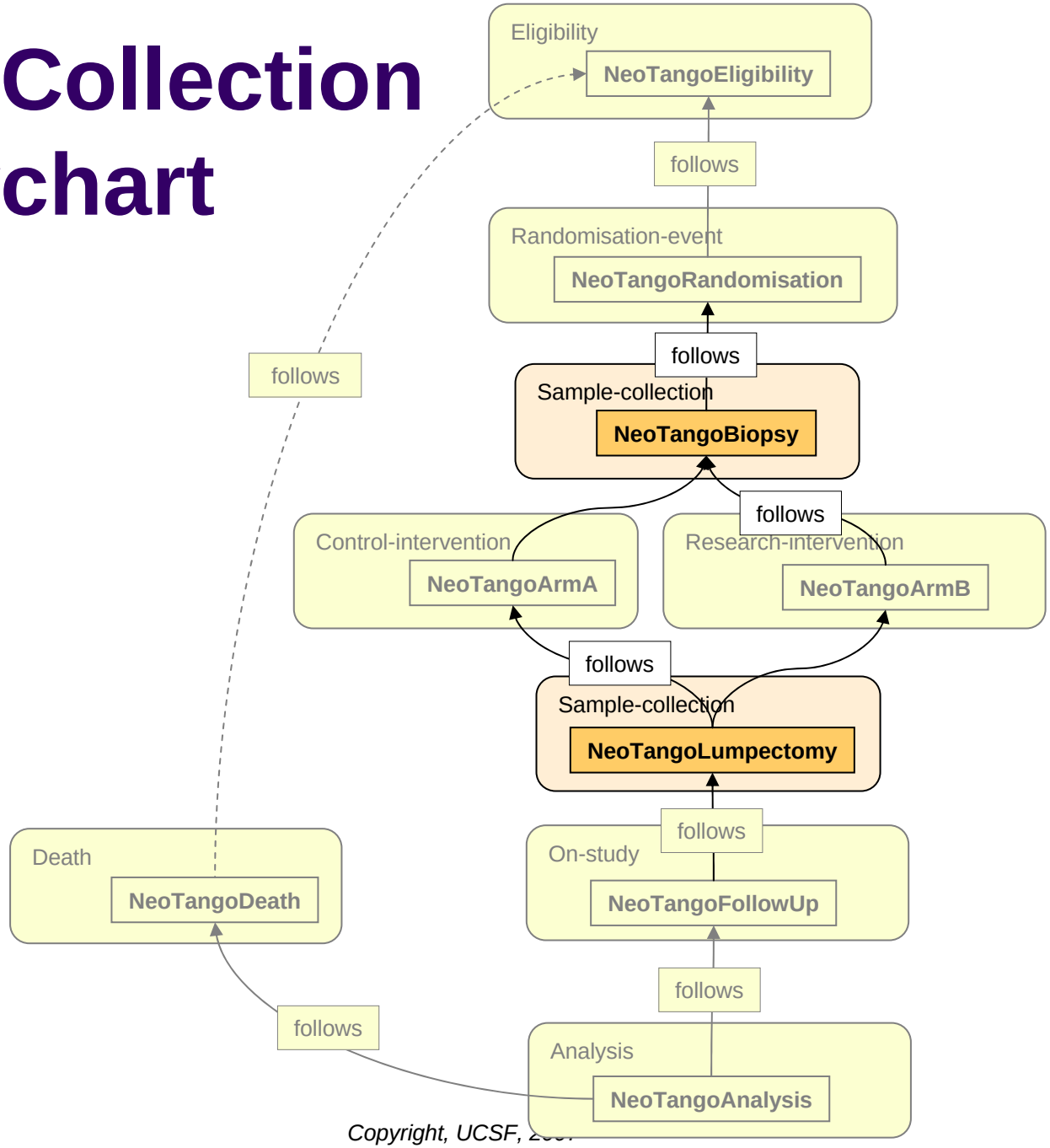


Translational Research Use Cases

- Within CTMS A studies, which randomized placebo-controlled trials testing "glitazones" analyzed liver expression profiles?
- Across all HIV RCTs in CTMSs A&B, what were overall rates of reported adverse events in Study Site A compared to Study Site B?
- Across all trials and the B-CFR Registry, how many Stage III breast CA patients have tissue samples from before and after chemotherapy? who are the PIs for those studies?



Tissue Collection in Flowchart



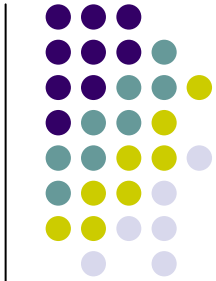
Outline

- Need for Clinical Research Ontology
- Ground Principles
- Scope & Model
- Example: NeoTango Trial
- Conclusion

Where We Are

- Initial core ontologies, factoring, and import structure
- Preliminary testing using NeoTango study
- Goals for end of 2007
  - first-cut representation of 4 UK CancerGrid studies (3 interventional, 1 observational), with binding to
    - (a fraction of) NCIT
    - CancerGrid CDEs
  - alpha release for community use and critique
    - e.g., via National Center for Biomedical Ontology's BioPortal
- 2008 and beyond
  - wider testing
  - mapping to BRIDG, upper ontologies, comparison to OCI, etc.

Summary and Opportunity



- OCRe provides reference semantics for talking about designs and methods of human experiments
  - Provides a consistent vocabulary across information models => supports integration of disparate information models
  - Supports explicit binding to external ontologies and information resources => supports rich indexing and resource discovery
- Of value for translational use cases in CTSA...
  - Within or across CTSA institutions
  - With clinical research management systems with/without BRIDG
  - For testing in non-cancer domains

Thanks! and Contacts

- Peter Maccallum (peter.maccallum@cancer.org.uk)
- Alan Rector (rector@cs.man.ac.uk)
- Ida Sim (ida.sim@ucsf.edu)