Chart Abstraction on VINCI: tools and services

Olga Patterson
Dan Denhalter

Outline

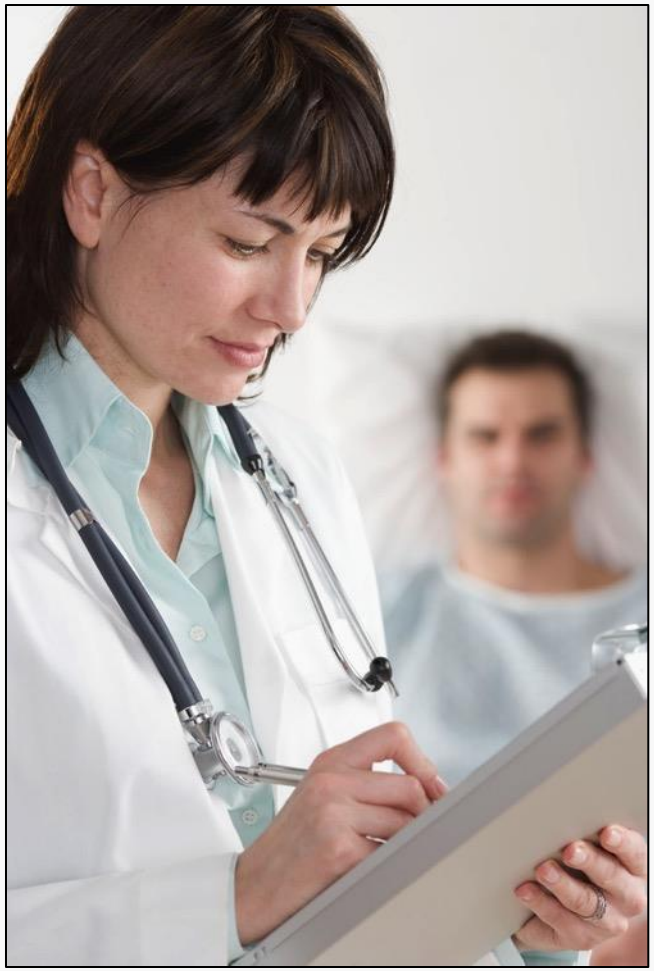
- Overview of chart abstraction
- Annotation project workflow
- Demonstration of tools:
  - eHost
  - Chart Review
- VINCI Annotation Services
- Questions and discussion

Overview

- Chart Abstraction
  - Chart review
  - Medical record review
  - Chart annotation
- A research methodology of data collection for retrospective investigation

How is EMR Created?

Creating patient documentation



Data Types in the EMR

An Electronic Medical Record contains information stored in both structured and unstructured formats.

- Structured data tables most often include objective information like labs and vital signs.
- Unstructured notes most often contain the patient's experience, the providers work-up and diagnosis, the treatment plan, and outcomes.

Structured Data

- Stored in database tables
- Include labs, meds, vitals, demographics, visit information, Codes, etc.

Patient ID	DateTime	LabID	TestName	Specimen	Result
101010	2012-10-06 1600	0409101	Potassium	Serum	4.9
101010	2012-10-06 1600	71368168	Sodium	Serum	136
202020	2012-10-06 0900	0409101	Potassium	Serum	3.6
202020	2012-10-06 0900	71368168	Sodium	Serum	145

Unstructured Data

- Stored in the database, but with text fields.
- Includes written or dictated text notes like progress notes, discharge summaries, and radiology notes.
- Also includes semi-structured information like templates and comment fields

Text Notes

SUBJECTIVE: The patient is in with several medical problems. He complains his mouth being sore since last week and also some "trouble with my eyes." He states that they feel "funny" but he is seeing okay. He denies any more diarrhea or abdominal pain. Bowels are working okay. He denies nausea or diarrhea. Eating is okay. He is emptying his bladder okay. He denies dysuria. His back is hurting worse. He complains of right shoulder pain and neck pain over the last week but denies any injury. He reports that his cough is about the same.

CURRENT MEDICATIONS: Metronidazole 250 mg q.i.d., Lortab 5/500 b.i.d., Allegra 180 mg daily, Levothroid 100 mcg daily, Lasix 20 mg daily, Flomax 0.4 mg at h.s., aspirin 81 mg daily, Celexa 40 mg daily, verapamil SR 180 mg one and a half tablet daily, Zetia 10 mg daily, Feosol b.i.d.

ALLERGIES: Lamisil, Equagesic, Bactrim, Dilatrate, cyclobenzaprine.

Vital Signs: His age is 66. Temperature: 97.7. Blood pressure: 134/80. Pulse: 88. Weight: 201 pounds.

HEENT: Head was normocephalic. Examination of the throat reveals it to be clear. He does have a few slight red patches on his upper inner lip consistent with yeast dermatitis.

Neck: Supple without adenopathy or thyromegaly.

Extremities: He has full range of motion of his shoulders but some tenderness to the trapezius over the right shoulder. Back has limited range of motion. He is nontender to his back. Deep tendon reflexes are 2+ bilaterally in lower extremities. Straight leg raising is positive for back pain on the right side at 90 degrees.

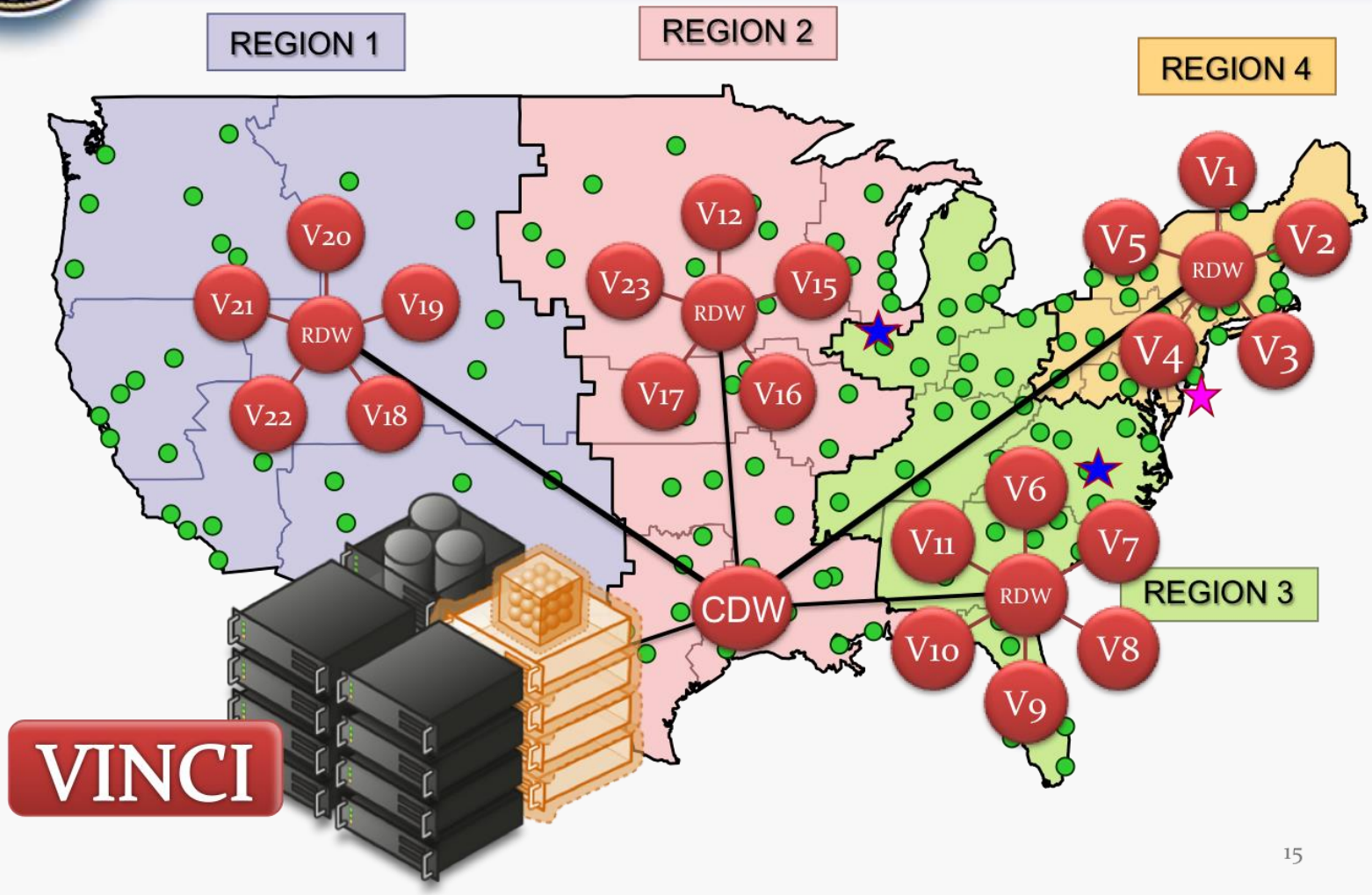
Abdomen: Soft, nontender without hepatosplenomegaly or mass. He has normal bowel sounds.

ASSESSMENT:

1. Clostridium difficile enteritis, improved.
2. Right shoulder pain.
3. Chronic low back pain.
4. Yeast thrush.
5. Coronary artery disease.
6. Urinary retention, which is doing better.

PLAN: I put him on Diflucan 200 mg daily for seven days. We will have him stop his metronidazole little earlier at his request. He can drop it down to t.i.d. until Friday of this week and then finish Fridayís dose and then stop the metronidazole and that will be more than a 10-day course. I ordered physical therapy to evaluate and treat his right shoulder and neck as indicated x 6 visits and he may see Dr. XYZ p.r.n. for his eye discomfort and his left eye pterygium which is noted on exam (minimal redness is noted to the conjunctiva on the left side but no mattering was seen.) Recheck with me in two to three weeks.

VA Informatics and Computing Infrastructure



15

Data in VINCI

21.6 million patients

6.9 billion lab tests

4.0 billion orders

2.9 billion procedures

2.7 billion clinical notes

2.1 billion diagnoses

2.0 billion medication fills

2.1 billion outpatient visits

14 million inpatient visits

Chart Abstraction

- Retrospective study
  - Utilizing clinical or administrative historical data that was originally collected for reasons other than research

- Chart abstraction focuses on unstructured data sources but can be combined with structured data

Benefit of Text

Text is where the majority of clinical information is stored.

- The patient experience is in text.
  - "Patient reports his knee hurts so bad he cannot sleep. He is also at risk of losing his job because he cannot work without sitting down"
- The type of illness, symptoms, and severity are in text.
  - "Diagnosed with relapsing remitting MS, currently mild tingling and weak grip."
- The timing of the episode is in text.
  - "The patient saw ENT last week and surg was scheduled. She was cleared by cardiology last Monday, labs yesterday were normal. Pt taken to OR at 3:00 PM for tonsillectomy, she was taken to PACU in good condition, returned to same day, and discharged at 8 pm."
- The disease course is detailed in text.
  - "The chest pain started at while at the gym running, but resolved with rest. The following day the patient again had CP while walking. Today the pain was constant and he presented to the ER. The pain resolved with aspirin, oxygen, nitro, MS."

Benefit of Text

- The treatment course is detailed in text.

  "The patient was started on albuterol, then changed to xop and spiriva after pulm consult and cardiology saw her."

- The outcomes are in text.

  "The patient was started on 48 weeks of peg interferon and ribavirin, but tx was stopped due to constant fatigue, anxiety, and concerns of his wife that he may harm himself."

- Even structured elements that are missing from the database are in text.

  "Patient is transferring care from university hospital. He is genotype 1, Vl 391,000, hep B immune, HIV negative."

- <u>The only thing not in text is what the provider failed to write.</u>

Challenges of Text

- EMR is written by providers for other providers
- Difficulty of document interpretation
  - Photocopies
  - Misspellings and grammar errors are pervasive
  - Terminology differs from non-clinical text
    - "patient endorses being verbally abused"
    - "patient status post spinal fusion"
    - "Angina, r/o MI"
  - Abbreviations and acronyms are common
    - "50 y/o pt c DM2, HTN, c/o SOB & CP. R/O MI"
- Missing data
- Incomplete and inconsistent documentation

Applications

- Retrospective clinical research
  - Case-control studies, case series
- Quality reporting
- Compliance auditing
- Guideline development
- Reference standard for computerized text processing

Chart Abstraction through Annotations

- Annotation = label that assigns meaning to data
  - Contain a pointer to start and stop points in a text
  - Can have class or attribute information with them
  - Generated by human, machine or human+machine.

"The CXR shows LLL consolidation."

Span: 15:31
Class: Finding
Assertion: Present

Annotation Project Workflow

1. Define concepts and variables
2. Select annotation tool
3. Document selection
4. Develop annotation guideline
5. Identify annotator qualifications
6. Train and manage annotators
7. Adjudication or Annotation quality measurement

Concept and variable definition

- Concept definition
  - Diagnosis, lab test, action, event…
- Variable definition – Values that the concepts can have:
  - Diagnosis: explicitly mentioned or inferred
  - Lab test: exact numeric value or range or direction
  - Action: planned or occurred
  - Event: explicitly mentioned or inferred

Concept and variable definition

Example of concept and variable definition for a study on quality of colonoscopy procedures

Concept	Variable definition	Source	Range of values
Bowel preparation	Explicitly stated quality of bowel preparation.	Colonoscopy report	Excellent Good Fair Poor Inadequate
Procedure indication	Explicitly stated indication for procedure.	Colonoscopy report	Screening High Risk/diagnostic Treatment

Concept and variable definition

Concept	Variable definition	Source	Range of values
Anemia	Any evidence of anemia	Any clinical note	Affirmed

Concept	Variable definition	Source	Range of values
Anemia	Explicitly stated diagnosis of aplastic anemia or pancytopenia consistent with ICD9 code 284, excluding non-specified anemia	TIU documents	Affirmed

Annotation tools

- Chart Review – on VINCI
- eHost - https://code.google.com/p/ehost/
- Knowtator – http://knowtator.sourceforge.net/
- RapTAT – https://code.google.com/p/raptat/
- BRAT - http://brat.nlplab.org/

- …..

Document selection

- Data sources in CDW
  - TIU Documents
  - Radiology notes
  - Comment and text fields in various tables
- Other sources

Sample size

Sample size for a descriptive study of a dichotomous variable 95-percent confidence interval

WIDTH OF THE CONFIDENCE INTERVAL (W)	*0.10*	*0.15*	*0.20*	*0.25*	*0.30*
Expected proportion (P)					
0.10	138	61			
0.15	196	87	49	31	
0.20	246	109	61	39	27
0.25	288	128	72	46	32
0.30	323	143	81	52	36
0.40	369	164	92	59	41
0.50	384	171	96	61	43

Adapted with permission from Hulley SB, et al. *Designing Clinical Research*, 3rd ed. Philadelphia: Wolters Kluwer Health; 2006:91.

Annotation guideline

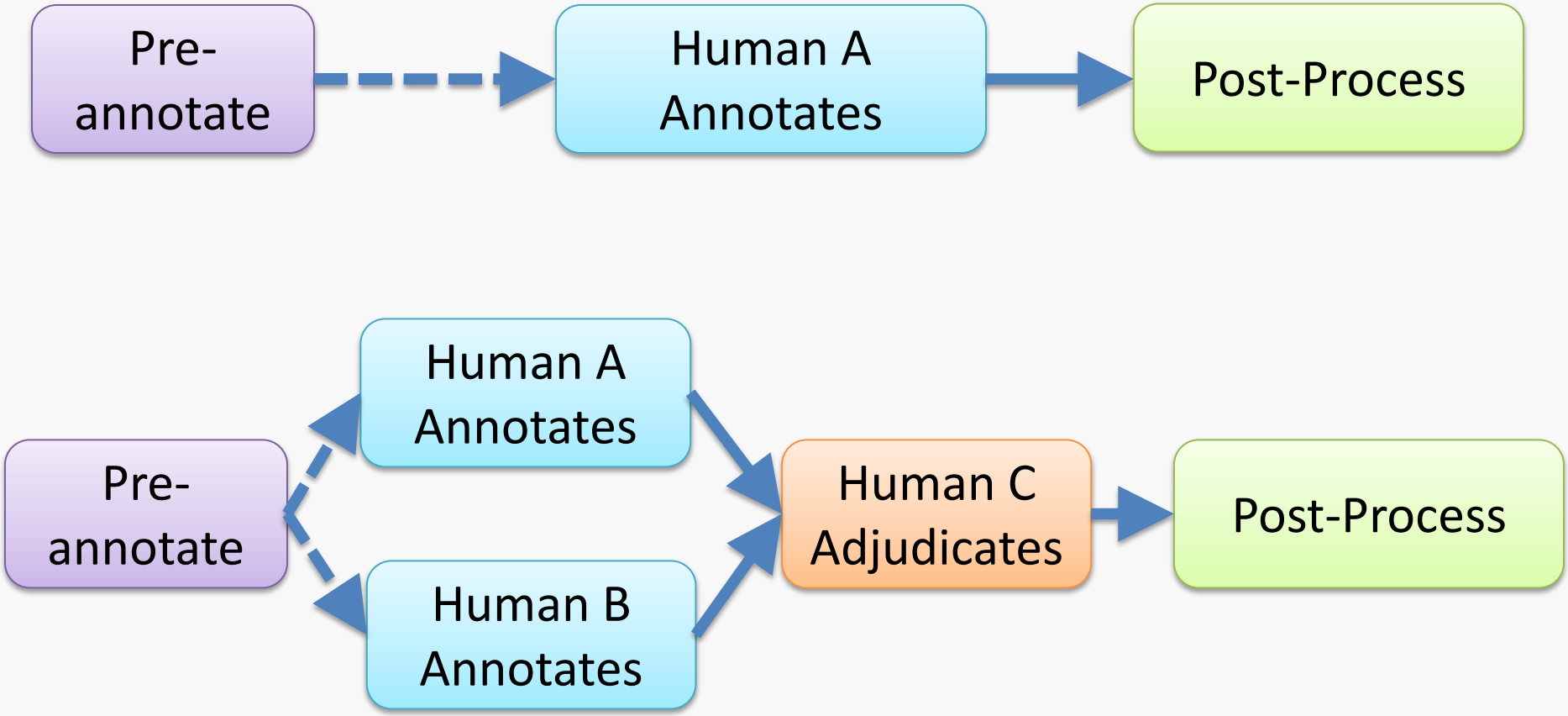
- Annotation schema
  - Annotation types = classes
  - Attributes
  - Relationships
- Formal step-by-step description of the annotation process
- Examples

Who Should Annotate?

- **Who**: depends on the clinical use case and information targets
- Domain expertise (physicians, nurses, nurse practitioners, pharmacists, physician assistants, coders, etc).
- Expensive domain experts are often not needed for specific tasks or portions of the task.

Human chart reviewer = annotator

Annotation Workflow



Annotation quality

– Inter-rater agreement

- Proportion of matched annotations to the number of all annotations performed by two people

$$\frac{Number\ matched\ annotations}{Total\ number\ of\ annotations}$$

– Cohen's Kappa

- p_o - observed proportionate agreement
- p_e - probability of random agreement

$$k = \frac{\rho_o - \rho_e}{1 - \rho_e}$$

VINCI Services Team

Jeff Scehnet

Kevin Malohi

VINCI@va.gov

VINCIServices@va.gov

- Concierge
- Data Provisioning
- Compliance
- Feasibility
- Recruitment
- **Annotation / Chart Review**
- Natural Language Processing
- Analytics and Data Services
- Application Development

VINCI Annotation Services

- Range of services
  - Education and training
  - Project definition and guideline development
  - Contracting annotators
  - Full chart abstraction project

VINCI Services Team

Jeff Scehnet

Kevin Malohi

VINCI@va.gov

VINCIServices@va.gov

- Concierge
- Data Provisioning
- Compliance
- Feasibility
- Recruitment
- **Annotation / Chart Review**
- Natural Language Processing
- Analytics and Data Services
- Application Development

Demo

eHost

Chart-Review

eHOST

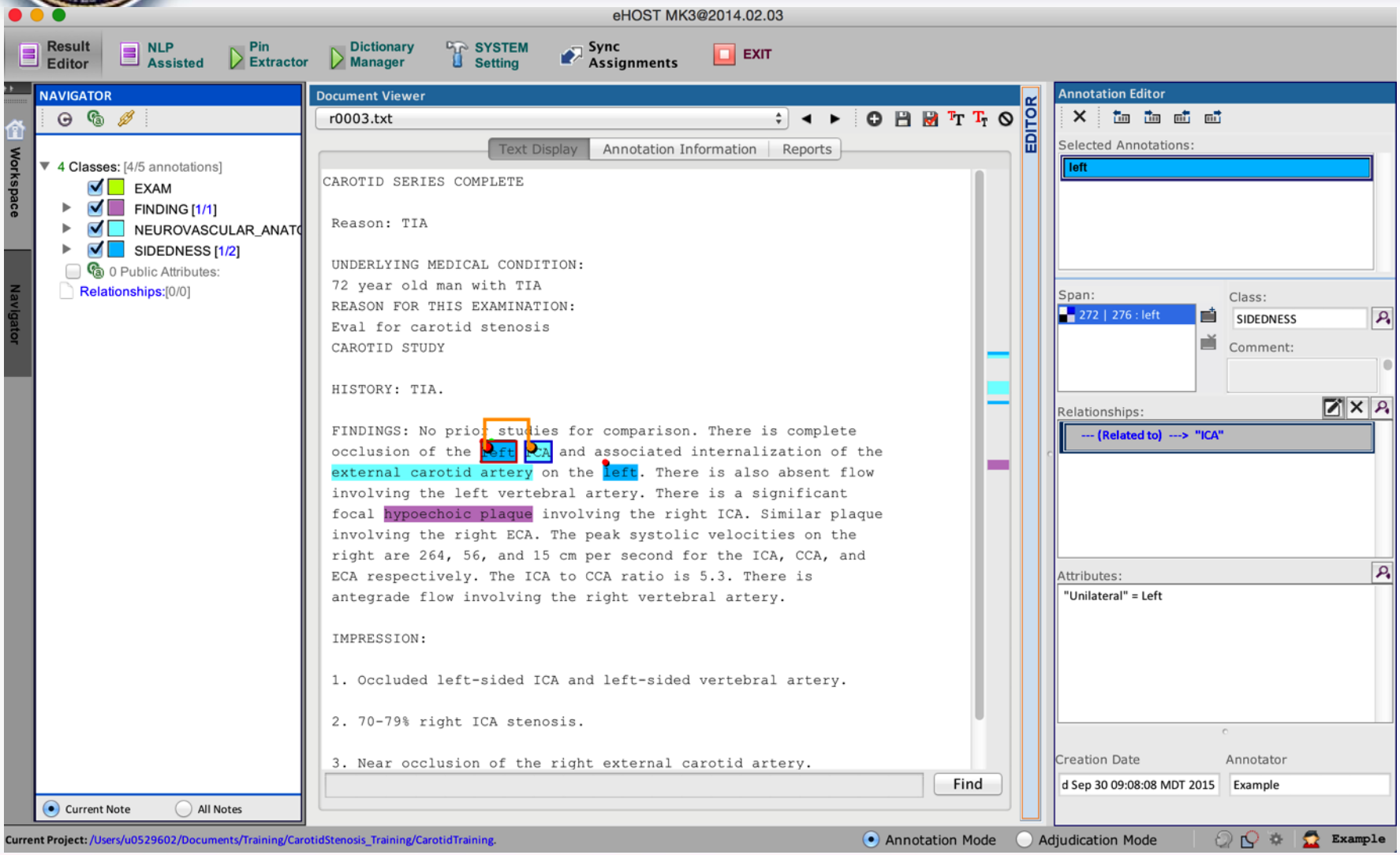

eHOST MK3@2014.02.03
Result Editor
NLP Assisted
Pin Extractor
Dictionary Manager
SYSTEM Setting
Sync Assignments
EXIT
NAVIGATOR
Workspace
Navigator
4 Classes: [4/5 annotations]
EXAM
FINDING [1/1]
NEUROVASCULAR_ANATC
SIDEDNESS [1/2]
0 Public Attributes:
Relationships:[0/0]
Current Note
All Notes
Document Viewer
r0003.txt
Text Display
Annotation Information
Reports
CAROTID SERIES COMPLETE
Reason: TIA
UNDERLYING MEDICAL CONDITION:
72 year old man with TIA
REASON FOR THIS EXAMINATION:
Eval for carotid stenosis
CAROTID STUDY
HISTORY: TIA.
FINDINGS: No prior studies for comparison. There is complete occlusion of the left ICA and associated internalization of the external carotid artery on the left. There is also absent flow involving the left vertebral artery. There is a significant focal hypoechoic plaque involving the right ICA. Similar plaque involving the right ECA. The peak systolic velocities on the right are 264, 56, and 15 cm per second for the ICA, CCA, and ECA respectively. The ICA to CCA ratio is 5.3. There is antegrade flow involving the right vertebral artery.
IMPRESSION:
1. Occluded left-sided ICA and left-sided vertebral artery.
2. 70-79% right ICA stenosis.
3. Near occlusion of the right external carotid artery.
Find
EDITOR
Annotation Editor
Selected Annotations:
left
Span:
272 | 276 : left
Class:
SIDEDNESS
Comment:
Relationships:
--- (Related to) ---> "ICA"
Attributes:
"Unilateral" = Left
Creation Date
d Sep 30 09:08:08 MDT 2015
Annotator
Example
Current Project: /Users/u0529602/Documents/Training/CarotidStenosis_Training/CarotidTraining.
Annotation Mode
Adjudication Mode
Example

Chart-Review: Configuration

ChartReview Home Admin Help Settings Logout (project1)

< Clinical Element Configuration List

Show Clinical Element Configuration

General Configuration | Columns | Content Template

Name	Lab Element
Description	Simple lab element based on example database.
Active	true
Title Field	LAB_NAME
Description Field	DESCRIPTION
Type	LIST
All Elements By Patient Id Query	select lab.lab_date, lab_test_lookup.lab_name, lab.result, concat(lab_test_lookup.lab_name, ' ', cast(lab.result as char)) as description, lab.id, lab.lab_performed_by from lab, lab_test_lookup, patient where lab.lab_performed_id = lab_test_lookup.id and lab.patient_id = patient.id and patient.id = ?
Single Element Query	select lab.lab_date, lab_test_lookup.lab_name, lab.result, concat(lab_test_lookup.lab_name, ' ', cast(lab.result as char)) as description, lab.id, lab.lab_performed_by from lab, lab_test_lookup, patient where lab.lab_performed_id = lab_test_lookup.id and lab.patient_id = patient.id and lab.id = ?
Created By	admin
Created Date	2014-03-05 16:28:16.0

Edit

Chart-Review: Process Creation

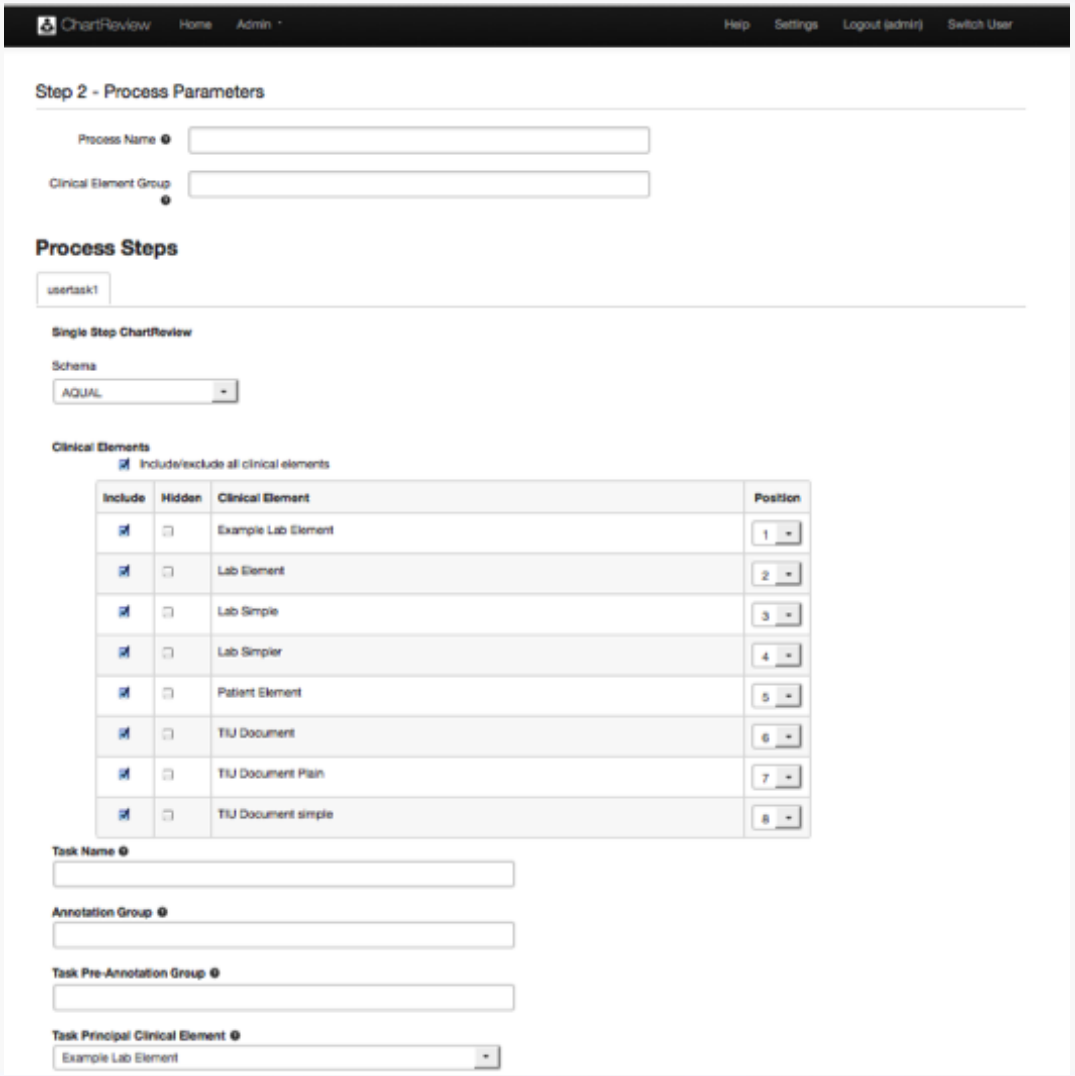


Chart-Review: Task Creation Query

Step 3 - Users and Clinical Elements In This Process

Task Creation Query

Enter the query to create process tasks.

```
select id from patient;
```

Users

- [ ] admin
- [ ] project1
- [ ] project2

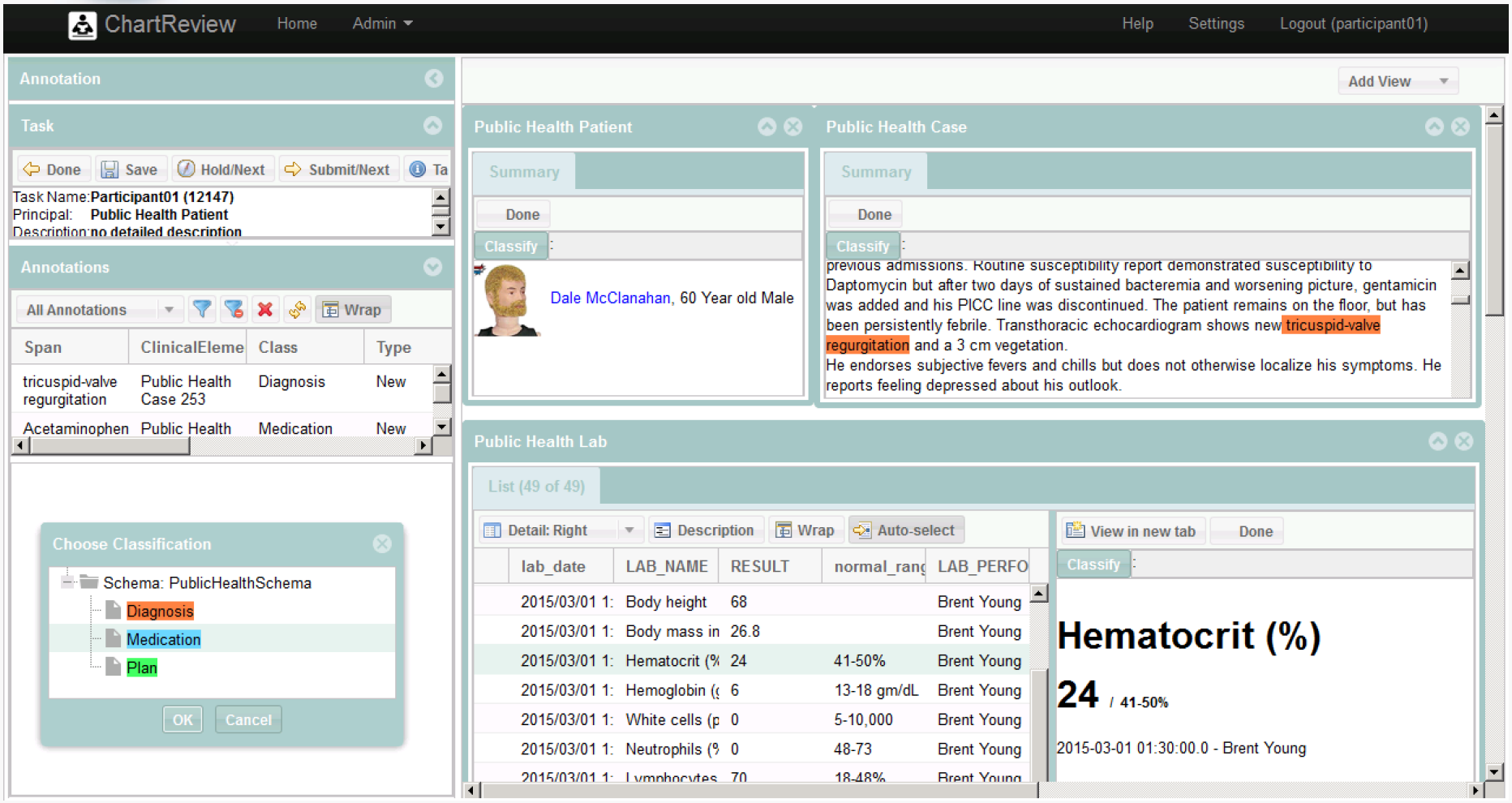
Other

Assignment style: By Process

Previous

Create Tasks and Finish

Chart-Review: User Interface



Acknowledgements

- Resources and Facilities
  - Veterans Affairs Salt Lake City Health Care System
  - Department of Epidemiology, University of Utah
- Funding Support
  - VA Informatics and Computing Infrastructure VA HSR HIR 08-204
  - VA Consortium for Health Informatics Research VA HSR HIR 08-374

Suggested reading

1. South BR, Shen S, Barrus R, DuVall SL, Uzuner O, Weir C. **Qualitative analysis of workflow modifications used to generate the reference standard for the 2010 i2b2/VA challenge**. AMIA Annu Symp Proc. 2011;2011:1243-1251.

2. Forbush TB, Shen S, South BR, Duvall SL. **What a catch! traits that define good annotators**. Stud Health Technol Inform. 2013;192:1213.

3. Warner JL, Anick P, Drews RE. **Physician inter-annotator agreement in the Quality Oncology Practice Initiative manual abstraction task**. J Oncol Pract. 2013;9(3):e96-e102.

4. Matt V, Matthew H. **The retrospective chart review: important methodological considerations**. J Educ Eval Health Prof. 2013;10:12.

5. Liu K, Chapman W, Hwa R, Crowley RS. **Heuristic Sample Selection to Minimize Reference Standard Training Set for a Part-Of-Speech Tagger**. J Am Med Informatics Assoc. 2007;14(5):641-650.

6. Ogren P V, Savova G, Chute CG. **Constructing Evaluation Corpora for Automated Clinical Named Entity Recognition**. In: Proceedings of the Sixth International Language Resources and Evaluation (LREC'08). Marrakech, Morocco; 2008.

7. Chapman WW, Dowling JN, Hripcsak G. **Evaluation of training with an annotation schema for manual annotation of clinical conditions from emergency department reports**. Int J Med Inf. 2008;77(2):107-113.