LOMA LINDA UNIVERSITY
HEALTH

An evaluation of feasibility of implementation of computable cancer phenotyping with pathology records from Cerner® EHR System

Huan Mo, MD (hmo@llu.edu), George W. Saukel, MD

Dept. Pathology, Loma Linda University Health, CA

Context: Electronic health record (EHR)-driven phenotyping has been widely used in biomedical and translational research as well as in precision medicine. Hochheiser et al (2016) proposed the DeepPhe information model for cancer phenotyping. As many academic institutes, the Cerner EHR system in Loma Linda University Health (LLUH) is mainly used for patient care and billing purposes. Here, we evaluated the availability and structure of Cerner pathology data for implementing the DeepPhe database model.

Design: We surveyed recent neoplastic pathology reports from LLUH Cerner system (n=30), and manually mapped Cerner pathology information to data elements in the DeepPhe.

Results: The correct diagnosis codes are available in 22 reports (73%) as structured data automatically extracted by Cerner as Systematized Nomenclature of Medicine (SNOMED) Morphology (M) code. However, average 0.83 incorrect M-codes per case are also captured by Cerner (mostly "benign lymph nodes" as "tumor, benign"). Body sites are captured as SNOMED topography (T) codes in 28 reports (93%), but body sites and diagnoses are not paired in cases with multiple specimens. Synoptic summaries (whenever available) provide paired and structured tumor diagnosis, body sites, and other tumor parameters. "Observations" for immunophenotyping are reported as either templated tables or free texts. On Level 3 of DeepPhe (episodes), current pathology records are capable of representing at least three scenarios: 1) treatment and re-biopsy; 2) precancerous to cancer; 3) leukemia follow-ups.

Conclusion: For future use of DeepPhe database in research, the required information is mostly available and structured in the current Cerner EHR pathology data system.

References:

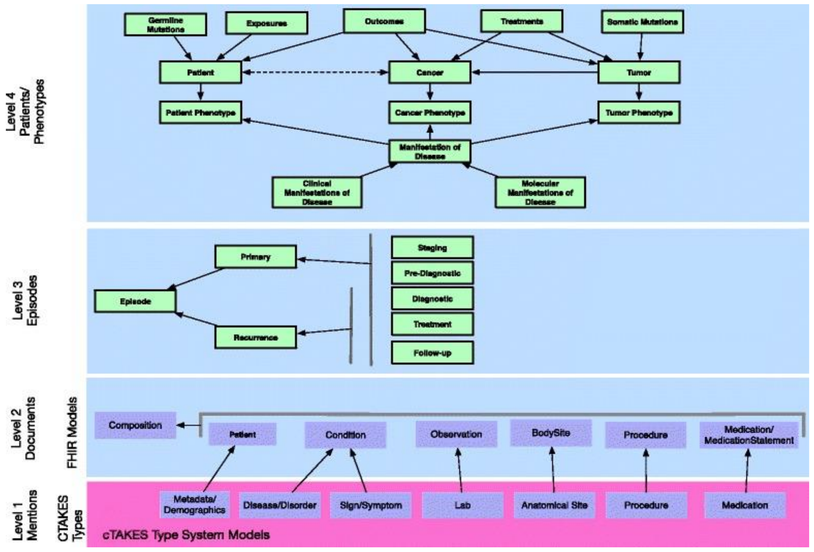
1. Hochheiser et al., An information model for computable cancer phenotypes., BMC Medical Informatics and Decision Making, 2016; 16:121. (DeepPhe model)
2. Mo et al., Desiderata for computable representations of electronic health records-driven phenotype algorithms., J Am Med Inform Assoc. 2015; 22(6):1220-30. (EHR-based phenotyping)

Acknowledgement:

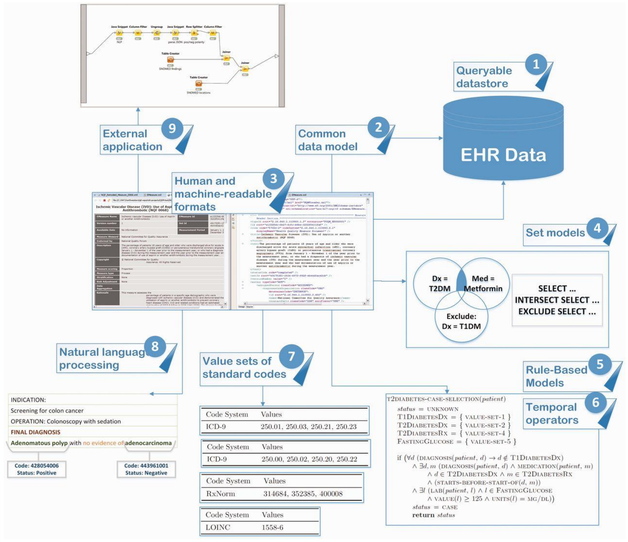
- Dr. Paul Herrmann, Dr. W. Tait Stevens, and Dr. Camilla Cobb have provided guidance.
- This work is based on knowledge from the Phenotype Execution Modeling Architecture (PhEMA, projectphema.org) project.

Phenotype level	Objects	Cerner Data	Comments
L1 (mentions) and L2 (Documents)	Disease (cTAKES), Conditions (FHIR)	SNOMED/M codes, diagnosis lines, Synoptics	
	Lab (cTAKES), Observation (FHIR)	IHC results, flow cytometry results, molecular studies	Mostly templated tables for IHC; structured data for flow cytometry
	Body site (FHIR)	SNOMED/T codes, diagnosis lines, Synoptics	
	Medication (FHIR)	Mostly none	Clinical EHR required
L3 (Episodes)	Primary	Resection report	
	Recurrence	Re-biopsy, follow-up bone marrows	
L4 (Phenotypes)	Cancer phenotype	Example: A p16 (+) SCC was at first diagnosed in neck lymph node biopsy, and later found in tonsil resection; it responses well after radiation therapy.	Complex queries are required
	Patient Phenotype	Example: A female was diagnosed endometrial adenocarcinoma that suggests DNA mismatch repair deficiency. Later, she has a colon polyp that shows high-grade dysplasia.	Complex queries are required

Table: Cerner® EHR system for pathology reports and the DeepPhe database model for cancer phenotyping. EHR: electronic health record. cTAKES: clinical Text Analysis Knowledge Extraction System (a natural language processing [NLP] system for extraction of information from EHR clinical free-text, http://ctakes.apache.org/). FHIR: Fast Healthcare Interoperability Resources (hl7.org/fhir), a standard for exchanging healthcare information electronically. SNOMED: Systematized Nomenclature of Medicine (M: Morphology; T: Topography). IHC: immunohistochemistry. SCC: squamous cell carcinoma.



Ref 1 (Hochheiser et al., 2016): DeepPhe Model



Ref 2 (Mo et al., 2015): Desiderata for computable EHR-driven phenotyping.

A
Sigmoid colon (sigmoidectomy):
Invasive moderately differentiated adenocarcinoma
Size: 3 cm
Extent: tumor invades into muscularis mucosa
Margins: uninvolved
Lymph-vascular invasion: not identified
Thirty-three lymph nodes, negative for metastatic tumor (0/33)
Tattoo ink identified
See the separate Synoptic Summary report for the complete listing of the tumor parameters

Grouping	Code	Description	Assignment M...
1	T-59470	Sigmoid colon	SYSTEM
1	M-81403	Adenocarcinoma, NOS	SYSTEM
1	M-80000	Tumor, benign	SYSTEM
1	T-1A512	Muscularis	SYSTEM
1	T-00400	Mucosa	SYSTEM
1	T-C6020	Lymph, NOS	SYSTEM
1	T-40000	Vascular	SYSTEM
1	D0-70150	Tattoo	SYSTEM
1	T-59300	Colon, NOS	SYSTEM

Figure: An example of Loma Linda University Health Pathology Report and the diagnostic codes automatically extracted by Cerner®.

CAP 2017 The Pathologists' Meeting, Poster No. 93, Session No. 400