



ODYSSEUS DATA SERVICES

THE HEALTHCARE DATA EXPERTS



CAP eCC & Nebraska Lexikon

Vlad Korsik, Dmitry Dymshyts, Michael Kallfelz

- CAP provides their protocols in electronic format as eCC files
- Ingestion of breast cancer related files (5) has been performed
- CAP release 2020_02 was used
- CAP also provides a CAP unique key (Ckey) to ICDO3 and SNOMED mapping table for selected items (~1700)
 - DCIS, Bx: 4 / DCIS, Rx: 4 / Invasive, Bx: 19 / Invasive, Rx: 21, Biomarkers: 0

CAP content

```
<Question name="Q_38125" order="147" ID="38125.100004300" title="Nuclear Pleomorphism">
<ListField name="lf_38125_1" order="148">
<List name="lst_38125_2" order="149">
<ListItem name="LI_16354" order="150" ID="16354.100004300" title="Score 1 (Nuclei small with little increase in size in comparison with normal breast epithelial cells, regular outlines, uniform nuclear chromatin, little variation in size)">
<Property name="p_rptTxt_16354_1" order="151" propName="reportText" val="Score 1" />
</ListItem>
<ListItem name="LI_16355" order="152" ID="16355.100004300" title="Score 2 (Cells larger than normal with open vesicular nuclei, visible nucleoli, and moderate variability in both size and shape)">
<Property name="p_rptTxt_16355_1" order="153" propName="reportText" val="Score 2" />
</ListItem>
<ListItem name="LI_16356" order="154" ID="16356.100004300" title="Score 3 (Vesicular nuclei, often with prominent nucleoli, exhibiting marked variation in size and shape, occasionally with very large and bizarre forms)">
<Property name="p_rptTxt_16356_1" order="155" propName="reportText" val="Score 3" />
</ListItem>
<ListItem name="LI_16357" order="156" ID="16357.100004300" title="Only microinvasion present (not graded)" />
<ListItem name="LI_16359" order="157" ID="16359.100004300" title="Score cannot be determined">
<ListItemResponseField name="lirf_16359_1" order="158">
<Response name="rsp_16359_2" order="159">
<string name="str_16359_3" order="160" />
</Response>
</ListItemResponseField>
</ListItem>
</List>
</ListField>
</Question>
```

A413					DCIS of the Breast: Biopsy		
	A	B	C	D	E	F	G
	Template Name	Item Ckey	Item Title	ConceptID	FSN	SNOMED Match	Status
1							
413	DCIS of the Breast: Biopsy	59744,1	Ductal carcinoma in situ (DCIS)	86616005	Intraductal carcinoma, noninfiltrating, no Internati	Semantic Match	Unchanged
414	DCIS of the Breast: Biopsy	40594,1	Paget disease	2985005	Paget's disease, mammary (morphologic abnormali	Semantic Match	Unchanged
415	DCIS of the Breast: Biopsy	44736,1	Encapsulated papillary carcinoma without invasive carcinom	703545003	Encapsulated papillary carcinoma (morphologic ab	Semantic Match	Unchanged
416	DCIS of the Breast: Biopsy	36008,1	Solid papillary carcinoma without invasive carcinoma	703546002	Solid papillary carcinoma in situ (morphologic abnor	Semantic Match	Unchanged
417	DCIS of the Breast: Resection, AJCC 8th ed	6454,100004	Ductal carcinoma in situ	86616005	Intraductal carcinoma, noninfiltrating, no Internati	Semantic Match	Unchanged
418	DCIS of the Breast: Resection, AJCC 8th ed	49894,1	Paget disease	2985005	Paget's disease, mammary (morphologic abnormali	Semantic Match	Unchanged
419	DCIS of the Breast: Resection, AJCC 8th ed	48537,1	Encapsulated papillary carcinoma without invasive carcinom	703545003	Encapsulated papillary carcinoma (morphologic ab	Semantic Match	Unchanged
420	DCIS of the Breast: Resection, AJCC 8th ed	45964,1	Solid papillary carcinoma without invasive carcinoma	703546002	Solid papillary carcinoma in situ (morphologic abnor	Semantic Match	Unchanged
531	Invasive Carcinoma of the Breast: Biopsy	45629,1	Invasive carcinoma of no special type (ductal)	82711006	Infiltrating duct carcinoma (morphologic abnormali	Semantic Match	Visible text change
532	Invasive Carcinoma of the Breast: Biopsy	46502,1	Invasive lobular carcinoma	89740008	Lobular carcinoma (morphologic abnormality)	Semantic Match	Unchanged
533	Invasive Carcinoma of the Breast: Biopsy	50830,1	Invasive carcinoma with mixed ductal and lobular features	444057000	Infiltrating carcinoma with ductal and lobular featu	Semantic Match	Visible text change
534	Invasive Carcinoma of the Breast: Biopsy	49791,1	Mucinous carcinoma	72495009	Mucinous adenocarcinoma (morphologic abnormali	Semantic Match	Unchanged
535	Invasive Carcinoma of the Breast: Biopsy	46153,1	Tubular carcinoma	4631006	Tubular adenocarcinoma (morphologic abnormality)	Semantic Match	Unchanged



Some numbers for CAP (breast cancer)

Concept_class_id	CKeys (%)
CAP Header	28 (2.3)
CAP Protocol	10 (0.82)*
CAP Value	861 (70.7)
CAP Variable	319 (26.2)
Total	1218

- 1218 CKey entries created in CAP vocabulary
*(10 de novo for protocol name)

- Nebraska Lexicon provides a file download for UMLS user
- Ingestion of all Nebraska Lexicon items was performed
 - Result is a mix of duplicates of SNOMED CT codes and SNOMED extension codes by Nebraska Lexicon
- Nebraska Lexicon version 2019 / 08 was made available
- Scott Campbell provided additional mapping material
 - Breast / invasive XML RefSet with CKey to SCT mapping
 - Master Breast Spreadsheet with CAP terms to SNOMED terms (and codes)

Nebraska material

nuclear pleomorphism							
ID ▼	CODE ▼	NAME ▼	CLASS ▼	CONCEPT ▼	VALIDITY ▼	DOMAIN ▼	VOCAB ▼
3451806	373371003	Nuclear pleomorphism finding	Clinical Finding	Non-standard	Valid	Condition	Nebraska Lexicon
3304035	371471007	Nuclear pleomorphism score	Observable Entity	Non-standard	Valid	Observation	Nebraska Lexicon
3436747	384735004	Nuclear pleomorphism: small regular nuclei	Clinical Finding	Non-standard	Valid	Condition	Nebraska Lexicon
3469722	373399007	Nuclear pleomorphism, grade 2: neither nuclear grade 1 nor nuclear grade 3	Clinical Finding	Non-standard	Valid	Condition	Nebraska Lexicon
3470959	373400000	Nuclear pleomorphism, grade 3: Markedly pleomorphic nuclei, usually >2.5 red blood cell diameters, with coarse chromatin and multiple nucleoli	Clinical Finding	Non-standard	Valid	Condition	Nebraska Lexicon
3433619	369782002	Ductal carcinoma in situ nuclear pleomorphism, grade 2: neither nuclear grade 1 nor nuclear grade 3	Clinical Finding	Non-standard	Valid	Condition	Nebraska Lexicon
3187111	4520001000004105	Nottingham nuclear pleomorphism score of primary malignant neoplasm of breast	Observable Entity	Standard	Valid	Observation	Nebraska Lexicon

<Question order="9370" ID="38125.100004300" SCT="4520001000004105" title="Nuclear Pleomorphism">
<ListField>
<List>

5				value)	e in size in comparison with normal
6	Nuclear Pleomorphism	4520001000004105	Nottingham nuclear pleomorphism score of excised malignant breast neoplasm (observable entity)		
7			Score 1	38112003 One (qualifier value)	en vesicular nuclei, visible nucleol
8			Score 2	19338005 Two (qualifier value)	
9			Score 3	79605009 Three (qualifier value)	minent nucleoli, exhibiting marked v
10			Score cannot be determined	82334004 Indeterminate (qualifier value)	
11			Other	169429731000004101 Histologic grade (qualifier value)	

</ListItemResponseField>
</ListItem>
</List>
</ListField>
</Question>

Target vocabulary distribution

Target vocabulary	# C-keys (%)
LOINC	29 (2.38)
NAACCR	6 (0.49)
Nebraska Lexicon	1024 (84.07)
SNOMED	18 (1.47)
Null	141 (11.57)*
Total	1218 (100)*

Overall

Target vocabulary	%
LOINC	2.52
Nebraska Lexicon	91.28
SNOMED	0.78
NULL	5.43

Target vocabulary	%
LOINC	0.72
Nebraska Lexicon	95.68
SNOMED	0.72
NULL	2.88

Invasive resection
Invasive biopsy

Target vocabulary	%
LOINC	0.36
Nebraska Lexicon	89.05
SNOMED	0.73
NULL	9.85

Target vocabulary	%
Nebraska Lexicon	92.75
SNOMED	1.45
NULL	5.8

DCIS resection
DCIS biopsy

Target vocabulary	%
LOINC	6.67
NAACCR	2.86
Nebraska Lexicon	53.33
SNOMED	4.76
NULL	32.38

Biomarkers

Nebraska Lexicon unique matches

Uniqueness Category	# codes (%)
NL_only	72 (23.1)
SNOMED Int	240 (76.9)
Total	312 (100)

Uniqueness Category	Mappings provided by Nebraska	# codes (% per category)
NL_only	yes	48 (66.7)
NL_only	no	24 (33.3)
SNOMED Int	yes	93 (38.75)
SNOMED Int	no	147 (61.25)

- 72 out of 312 concepts used exist only in Nebraska
- Mapping was provided for 141 out of 312

Status of CKey mapping

- Direct mapping was provided for the *Invasive Carcinoma of the Breast Protocol* by an XML Schema
 - 133/516 CKeys were missing in Schema (newer CAP version, not complete)
 - 383/516 CKeys were present in Schema
 - 137 have a mapping to Nebraska, 30 of these were remapped manually
 - 246/383 have no mapping and were mapped manually

Issue type: loss of context

Only DCIS is present after presurgical (neoadjuvant) therapy Ductal Carcinoma In Situ (DCIS) (Note G) Present Ductal Carcinoma In Situ (DCIS) (Note G) TUMOR	16313	373176000	Ductal carcinoma in situ
--	-------	-----------	--------------------------

- missing C-Key - SNOMED relation in Lexicon
- *?Number of Invasive Tumor Cells Counted|HER2 by in situ Hybridization (Note C)|Test(s) Performed (Note A)|TEST(S) PERFORMED*
=> tumour cells, malignant

Issue type: not a real concept

- administrative information (header)
- free text (comments) => eventually augment concept with header information to preserve meaning



Issue type: needs a precoordinated concept

Distance of DCIS from Other Margins (specify millimeter distance including greater than if appropriate) Positive for DCIS ?DCIS Margins MARGINS (Note I)	43941	384822004	Distance of tumour from margin
Parental items and mappings for 43941			
Positive for DCIS ?DCIS Margins MARGINS (Note I)	16658	370109009	Surgical margin involved by tumour
?DCIS Margins MARGINS (Note I)	16638	2130001000004106	Presence of ductal carcinoma in situ at surgical margin in specimen excised from breast
MARGINS (Note I)	16567	<i>The source item classified as CAP header, not a real concept on our understanding.</i>	

- hierarchy + statement: Distance of DCIS from Other Margins

Issue type: poor data modeling

- fields with missing relations (Other margin)

