TIES Cancer Research Network (TCRN)
ITCR 2017 Meeting – Santa Cruz
U24 CA 180921

May 31st 2017

Mike Becich, MD PhD – New TCRN PI

Becich Conflicts of Interest (Disclaimer)

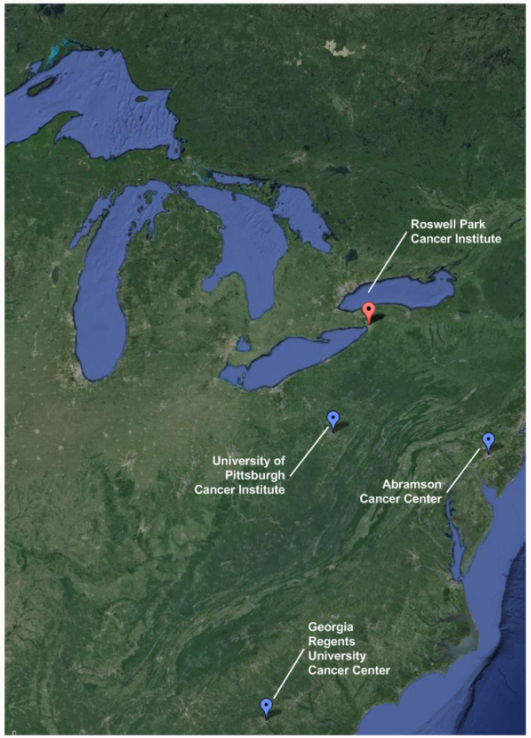
- SPDx (founder and stock) – computational pathology company
- Cancer Center Consultancies and EABs – Baylor, University of Colorado, University of Michigan and Wake Forest
- CTSA Consultancies and EABs – numerous (not a conflict for this talk except possibly for U of Chicago Institute for Translational Medicine)

Disclaimer – I am a member of NCI's Board of Scientific Advisors

What is TIES?

- An **NLP and IR system** for de-identifying, annotating, storing and retrieving pathology and radiology documents
- A system for **indexing research resources** (FFPE, FF, images) with document annotations
- An **GUI for querying** large repository of annotated documents and obtaining resources locally, using an honest broker model
- A platform to support **data and biospecimen sharing** among networks of cancer centers and other institutions

TIES Cancer Research Network



- **University Pittsburgh Cancer Institute**
- **Abramson Cancer Center (Penn)**
- **Roswell Park Cancer Institute**
- **Georgia Regents Cancer Center**
- **New partners are getting ready to join**
- **Network Trust Agreements**
  - IRBs agree that use of data for investigators is NHSR, no need for an additional IRB protocol even to access record level de-id data
  - Governance
  - Agreement to abide by SOPs
  - Instrument of Adherence

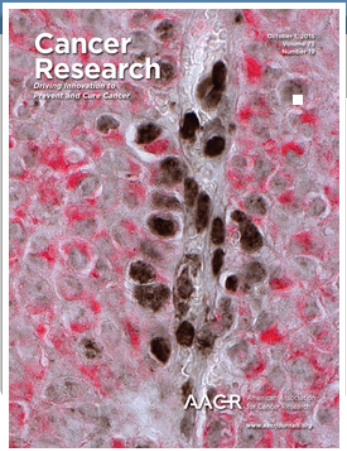
Table 2. TCRN case statistics for numbers of patients and cases (A) and the number of cases of rare tumors (B) and common cancer categories (C) based on final diagnosis

	GRU	RPCI	ACC	UPCI	Total
A. Case statistics					
Patients	76,404	72,376	465,717	1,840,156	2,454,653
Pathology cases	157,316	156,555	857,681	4,588,017	5,759,569
B. Rare tumors					
Adenoid cystic carcinoma	41	88	404	509	1,042
Adrenocortical carcinoma	5	20	59	63	147
Alveolar soft part sarcoma	3	15	10	25	53
Angioimmunoblastic lymphadenopathy	12	35	58	84	189
Chordoma	5	14	124	245	388
Follicular dendritic cell sarcoma	2	2	8	13	25
Merkel cell carcinoma	9	72	165	196	442
Ovarian granulosa cell tumor	4	10	23	34	71
Phaeochromocytoma	15	38	272	164	489
Pleomorphic xanthoastrocytoma	2	5	12	53	72
Pseudomyxoma peritonei	6	36	46	129	217
Rhabdomyosarcoma	34	70	86	270	460
Sebaceous adenocarcinoma	13	33	26	94	166
Sinonasal undifferentiated carcinoma	2	6	31	27	66
Thymoma	13	45	433	210	701
C. Common cancer categories					
Bladder carcinoma	345	1,618	3,873	6,711	12,547
Breast carcinoma	1,143	9,605	28,262	37,691	76,701
Colorectal carcinoma	465	2,530	6,898	11,608	21,501
Endometrial carcinoma	394	1,815	3,707	7,706	13,622
Esophageal carcinoma	63	1,477	2,452	3,514	7,506
Hepatic carcinoma	153	633	2,912	5,720	9,418
Lung carcinoma	820	4,264	10,208	17,955	33,247
Lymphoma	1,387	6,795	10,605	15,689	34,476
Malignant glial neoplasm	242	292	2,198	4,943	7,675
Malignant melanoma	335	2,675	5,180	7,068	15,258
Ovarian carcinoma	503	2,872	4,659	6,446	14,480
Pancreatic carcinoma	162	740	1,866	3,622	6,390
Prostate carcinoma	903	3,612	18,867	19,445	42,827
Renal cell carcinoma	364	1,319	3,183	10,950	15,816
Thyroid carcinoma	474	1,236	7,681	12,387	21,778

Use in Tissue Bank

- Honest Broker functionality is the key
- Order biospecimens and images from within TIES, or export manifest for another system
- Tags, Structured Data can be used to import info from LIMS, enabling search from within TIES
- Whole Slide Images
- Next phase of deliverables includes Computational Pathology Support

Cancer Research
The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)



• View the new Impact Factor
• View the Most-Cited Articles of *Cancer Research*

Cancer Research

Resource

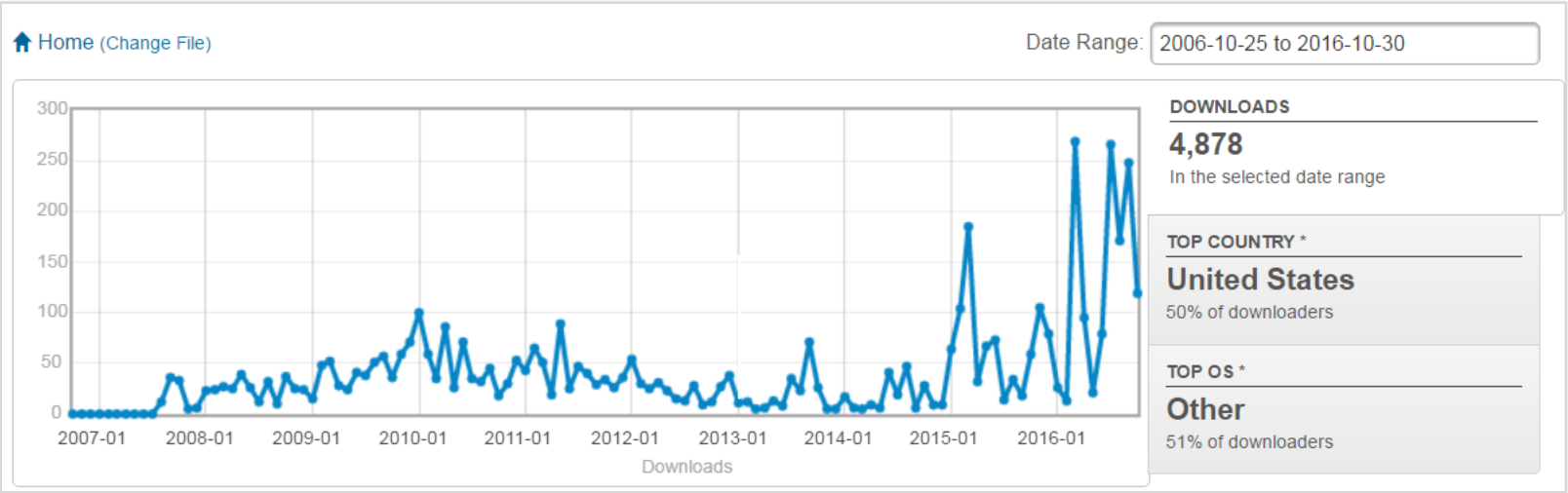
A Federated Network for Translational Cancer Research Using Clinical Data and Biospecimens

Rebecca S. Jacobson[1], Michael J. Becich[1], Roni J. Bollag[2], Girish Chavan[1], Julia Corrigan[1], Rajiv Dhir[1], Michael D. Feldman[3], Carmelo Gaudioso[4], Elizabeth Legowski[1], Nita J. Maihle[2], Kevin Mitchell[1], Monica Murphy[4], Mayurapriyan Sakthivel[4], Eugene Tseytlin[1], and JoEllen Weaver[3]

Abstract

Advances in cancer research and personalized medicine will require significant new bridging infrastructures, including more robust biorepositories that link human tissue to clinical phenotypes and outcomes. In order to meet that challenge, four cancer centers formed the Text Information Extraction System (TIES) policies, and procedures, enable regulatory compliance. The TIES Cancer Research Network now provides integrated access to investigators at all member institutions, where multiple investigator-driven pilot projects are underway. Examples of federated search across the network illustrate the potential impact on

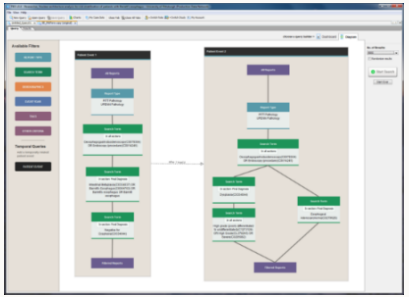
TIES Downloads

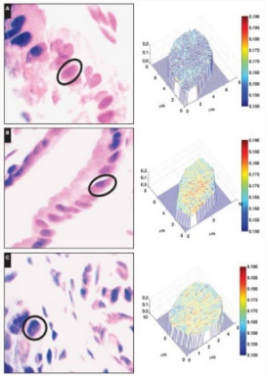


ITCR

Example of TCRN Pilot Project

UPCI Investigator
Yang Liu, PhD





Published OnlineFirst September 17, 2015; DOI: 10.1158/0008-5472.CAN-15-1274

Integrated Systems and Technologies

Cancer Research

Early Prediction of Cancer Progression by Depth-Resolved Nanoscale Mapping of Nuclear Architecture from Unstained Tissue Specimens

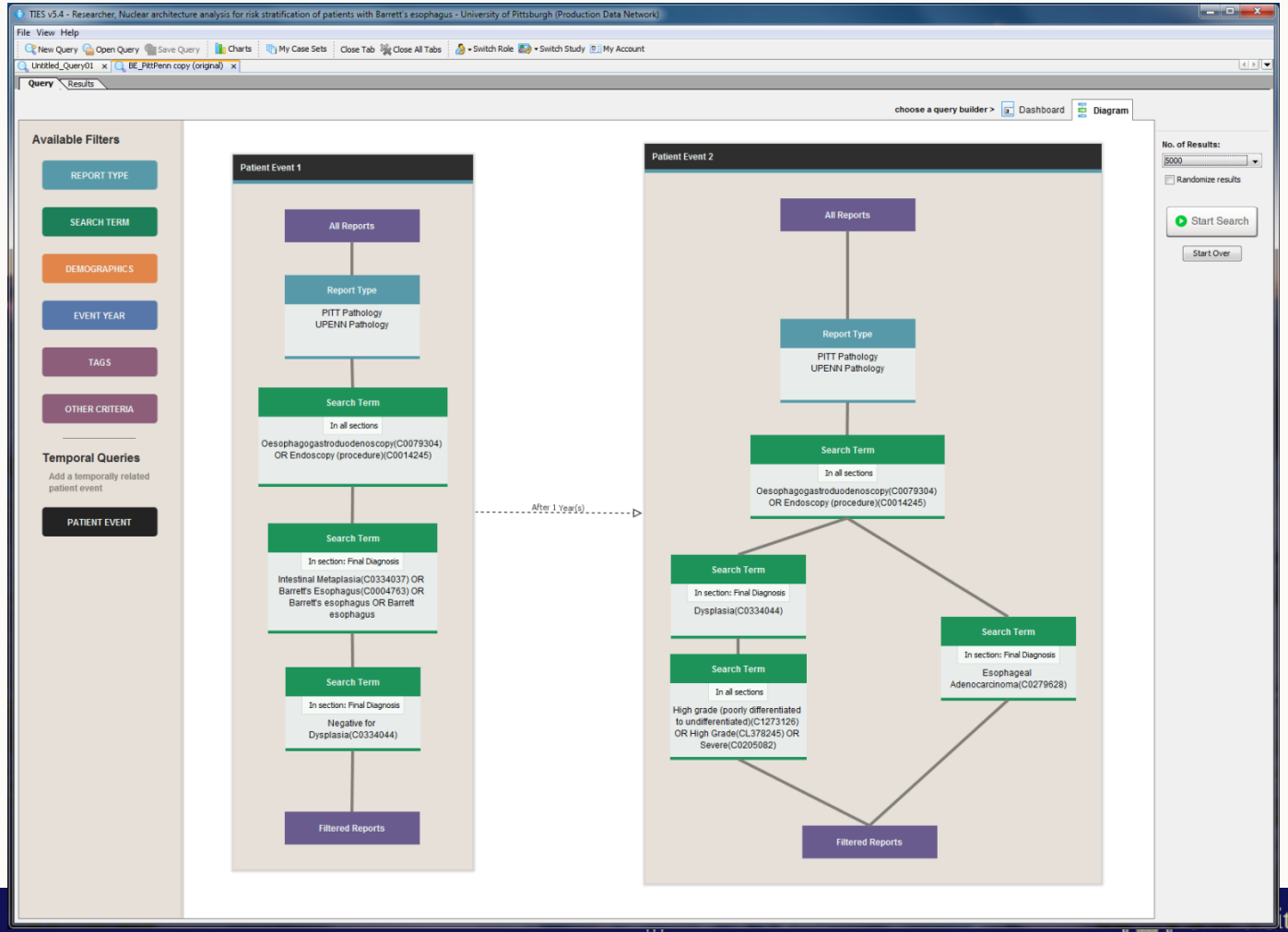
Shikhar Uttam[1], Hoa V. Pham[1], Justin LaFace[1], Brian Leibowitz[2,3], Jian Yu[2,3], Randall E. Brand[4], Douglas J. Hartman[2], and Yang Liu[1,3,4]

Abstract

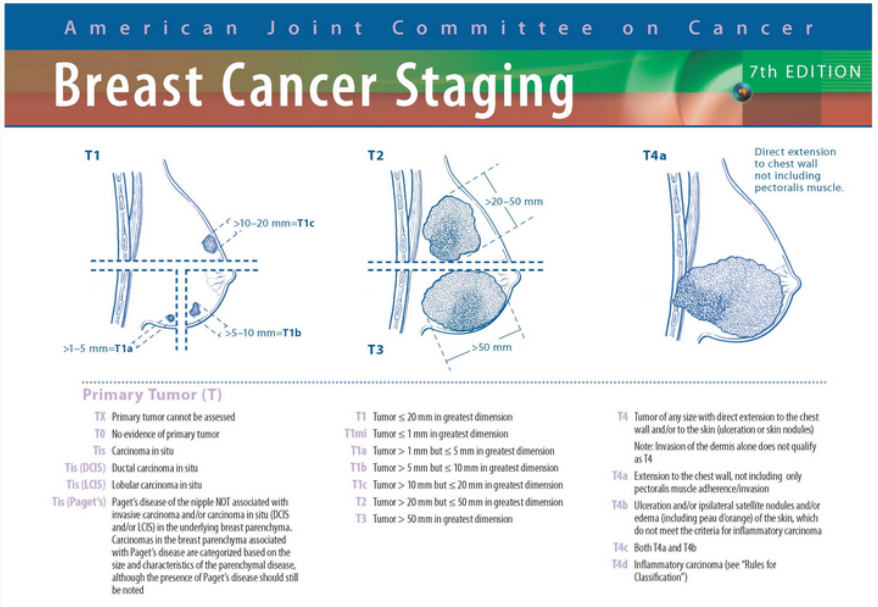
Early cancer detection currently relies on screening the entire at-risk population, as with colonoscopy and mammography. Therefore, frequent, invasive surveillance of patients at risk for developing cancer carries financial, physical, and emotional burdens because clinicians lack tools to accurately predict which patients will actually progress into malignancy. Here, we present a new method to predict cancer progression risk via nanoscale nuclear architecture mapping (nanoNAM) of unstained tissue sections based on the intrinsic density alteration of nuclear structure rather than the amount of stain uptake. We demonstrate that nanoNAM detects a gradual increase in the density alteration of nuclear architecture during malignant transformation in animal models of colon carcinogenesis and in human patients with ulcerative colitis, even in tissue that appears histologically normal according to pathologists. We evaluated the ability of nanoNAM to predict "future" cancer progression in patients with ulcerative colitis who did and did not develop colon cancer up to 13 years after their initial colonoscopy. NanoNAM of the initial biopsies correctly classified 12 of 15 patients who eventually developed colon cancer and 15 of 18 who did not, with an overall accuracy of 85%. Taken together, our findings demonstrate great potential for nanoNAM in predicting cancer progression risk and suggest that further validation in a multicenter study with larger cohorts may eventually advance this method to become a routine clinical test. *Cancer Res; 75(22); 4718–27. ©2015 AACR.*

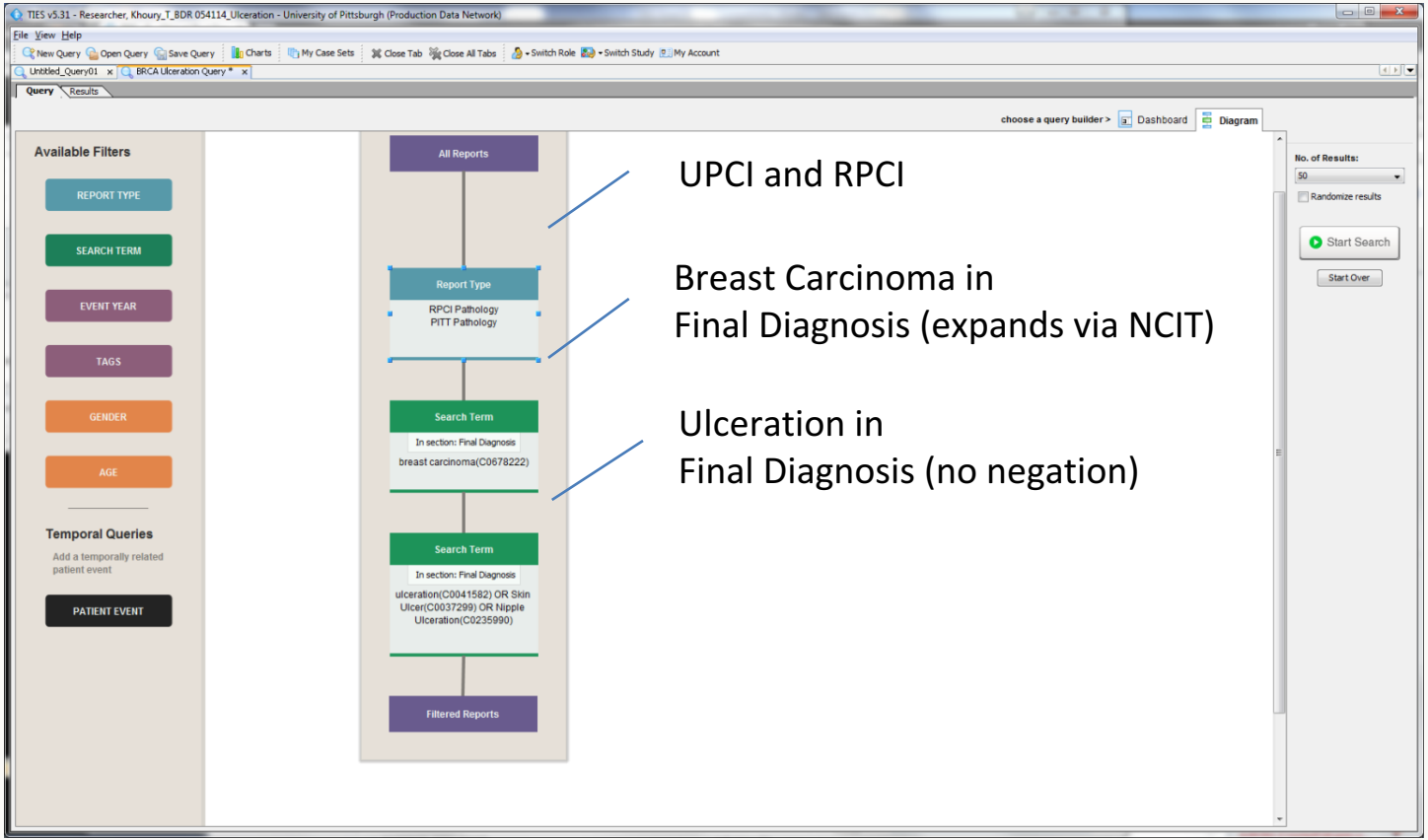
Doubled Study N using TCRN

UPMC: 46
Penn: 44


TIES v5.4 - Researcher, Nuclear architecture analysis for risk stratification of patients with Barrett's esophagus - University of Pittsburgh (Production Data Network)
File View Help
New Query
Open Query
Save Query
Charts
My Case Sets
Close Tab
Close All Tabs
Switch Role
Switch Study
My Account
Untitled_Query01
BE_PittPenn copy (original)
Query
Results
choose a query builder >
Dashboard
Diagram
Available Filters
REPORT TYPE
SEARCH TERM
DEMOGRAPHICS
EVENT YEAR
TAGS
OTHER CRITERIA
Temporal Queries
Add a temporally related patient event
PATIENT EVENT
Patient Event 1
All Reports
Report Type
PITT Pathology
UPENN Pathology
Search Term
In all sections
Oesophagogastroduodenoscopy(C0079304) OR Endoscopy (procedure)(C0014245)
Search Term
In section: Final Diagnosis
Intestinal Metaplasia(C0334037) OR Barrett's Esophagus(C0004763) OR Barrett's esophagus OR Barrett esophagus
Search Term
In section: Final Diagnosis
Negative for Dysplasia(C0334044)
Filtered Reports
After 1 Year(s)
Patient Event 2
All Reports
Report Type
PITT Pathology
UPENN Pathology
Search Term
In all sections
Oesophagogastroduodenoscopy(C0079304) OR Endoscopy (procedure)(C0014245)
Search Term
In section: Final Diagnosis
Dysplasia(C0334044)
Search Term
In section: Final Diagnosis
Esophageal Adenocarcinoma(C0279628)
Search Term
In all sections
High grade (poorly differentiated to undifferentiated)(C1273126) OR High Grade(CL378245) OR Severe(C0205082)
Filtered Reports
No. of Results:
5000
Randomize results
Start Search
Start Over

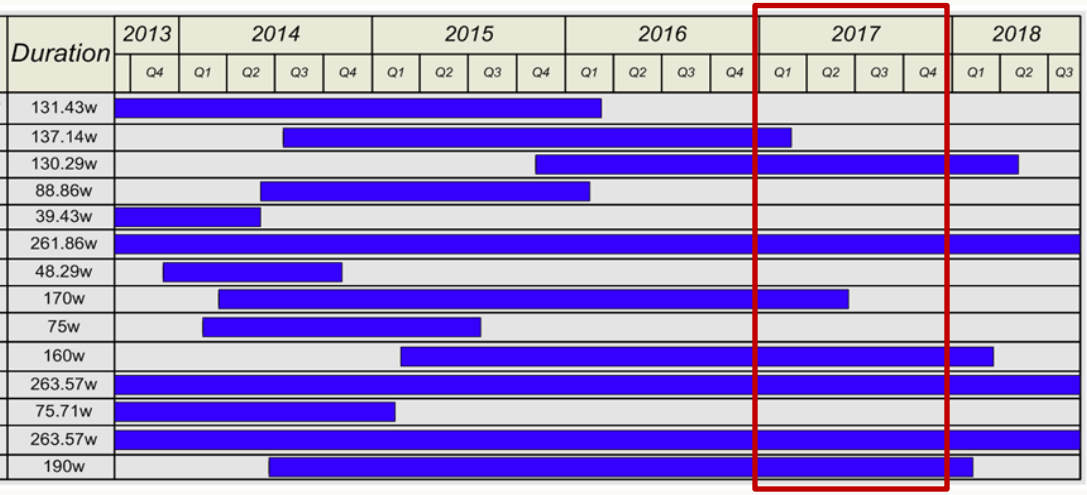
Examples of Research Project
Thaer Khoury, MD
Roswell Park Cancer Institute




TIES v5.31 - Researcher, Khoury_T_BDR 054114_Ulceration - University of Pittsburgh (Production Data Network)
File View Help
New Query
Open Query
Save Query
Charts
My Case Sets
Close Tab
Close All Tabs
Switch Role
Switch Study
My Account
Untitled_Query01
BRCA Ulceration Query *
Query
Results
choose a query builder >
Dashboard
Diagram
Available Filters
REPORT TYPE
SEARCH TERM
EVENT YEAR
TAGS
GENDER
AGE
Temporal Queries
Add a temporally related patient event
PATIENT EVENT
All Reports
Report Type
RPCI Pathology
PITT Pathology
Search Term
In section: Final Diagnosis
breast carcinoma(C0678222)
Search Term
In section: Final Diagnosis
ulceration(C0041582) OR Skin Ulcer(C0037299) OR Nipple Ulceration(C0235990)
Filtered Reports
No. of Results:
50
Randomize results
Start Search
Start Over
UPCI and RPCI
Breast Carcinoma in
Final Diagnosis (expands via NCIT)
Ulceration in
Final Diagnosis (no negation)

Plans for Y4

ID	Specific Aim/Sub Aim	Duration
1	1.1 Implement “trust framework”; support regulatory workflow	131.43w
2	1.2 Paraffin registry development functionali	137.14w
3	1.3 TMA development functonality	130.29w
4	1.4 Virtual slide visualization	88.86w
5	2.1 Install TIES 5.2 at all sites	39.43w
6	2.2 Establish and maintain governance structure	261.86w
7	2.3 Execute trust agreements	48.29w
8	2.4 Develop and refine policies and processes	170w
9	3.1 Initial set of Pilot Projects	75w
10	3.2 Additional Pilot Projects	160w
11	4.1 Software releases (at least 2/year)	263.57w
12	4.2 Develop deployment and connection blueprints	75.71w
13	4.3 Dissemination activities	263.57w
14	4.3 Measure impact of TCRN	190w



Work to complete

Help with LIMS Integration
Cancer Registry Integration
New NobleCoder Integration
Image Annotation Tools
Computational Pathology framework
Optional additional de-identifier

Other goals

Publications from users
Dissemination at all sites
Additional adoptions at other institutions

Adding Cancer Registry Data to TIES

- Identified as a high value development target from users
- We have secured additional funding from or Institution for Precision Medicine in Pittsburgh
- Senior Developer Mike Davis leads this effort.
- Starting with Breast Cancer first
- Work that we do here can immediately be leveraged by all of you to similarly add CR data to your TIES instances
- Result = deeper patient annotation and outcomes data

Data Elements

Demographics	Primary	Treatment	Outcome
Race	Primary Site	Surgery	Vital Status
Gender	Histology	Chemotherapy	Cancer Status
Age @ Diagnosis	Grade	BRM	Recurrence
Smoking	Path TNM	Hormonal	Cause of Death
Alcohol	Clinical TNM	Immunotherapy	
	Prognostic Factors (including site specific)	Rad Onc	

Version Release 5.7 Coming Soon

Winter 2016 Release – v5.7 – Anticipated June. 2017

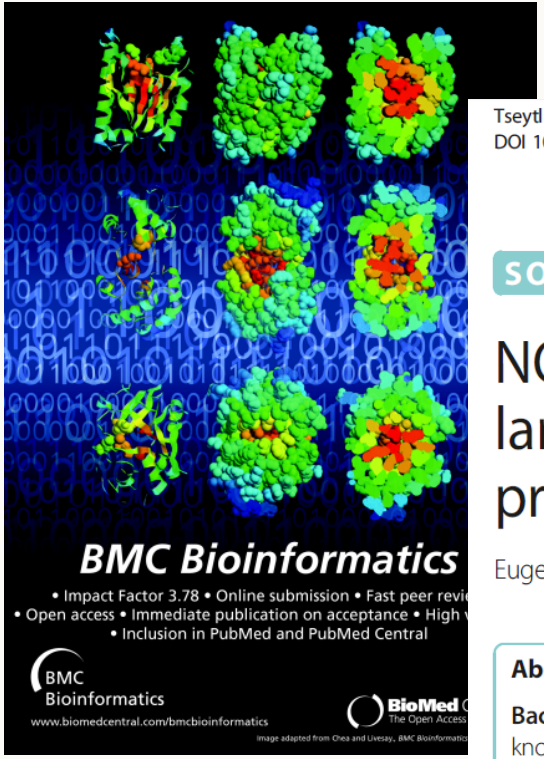
- Improved structured data import to support large batches and grouped data elements.

Backlog

- TMA support
- Expose uncertainty and temporality in search
- Standardize MetriQ data across sites for loading into TIES
- Free alternative to De-ID
- Patient level visualization
- Show patient MRN in selection tree for HBs for case sets
- Remove outside consult from cases

Development Plans

- New Coding Pipeline
  - Integrate NobleCoder v1.1. More accurate coding, faster coding. Uncertainty, polarity, experiencer and temporality annotations.
  - Latest NCIM terminology with more fine tuned sources.
- Cancer Registry data integration
- Email based management of account review and approvals.
- Patient level search index and visualization
- Manual Annotation Tool Enhancements
  - Link report text annotations to data in form fields.
  - Intelligent auto-highlighting and filling of form fields.
  - Library of forms to choose from, making it easy to share and reuse previously created forms.



Tseytlin *et al. BMC Bioinformatics* (2016) 17:32
DOI 10.1186/s12859-015-0871-y

BMC Bioinformatics

SOFTWARE Open Access

NOBLE – Flexible concept recognition for large-scale biomedical natural language processing

Eugene Tseytlin, Kevin Mitchell, Elizabeth Legowski, Julia Corrigan, Girish Chavan and Rebecca S. Jacobson*

Abstract

Background: Natural language processing (NLP) applications are increasingly important in biomedical data analysis, knowledge engineering, and decision support. Concept recognition is an important component task for NLP pipelines, and can be either general-purpose or domain-specific. We describe a novel, flexible, and general-purpose concept recognition component for NLP pipelines, and compare its speed and accuracy against five commonly used alternatives on both a biological and clinical corpus.

NOBLE Coder implements a general algorithm for matching terms to concepts from an arbitrary vocabulary set. The system's *matching options* can be configured individually or in combination to yield specific system behavior for a

TIES and the TIES Cancer Research Network

TIES Team

Girish Chavan
Eugene Tseytlin
Kevin Mitchell
Julia Corrigan
Liz Legowski

Adi Nemlekar
Yining Zhao
Vanessa Benkovich
Liron Pantanowitz
Rajiv Dhir

Roswell Park
Carmelo Gaudioso
Monica Murphy
Mayurapriyan Sakthivel
Amanda Rundell

GRU
Roni Bollag
Samir Khleif
Jennifer Carrick
Nita Maihle
And more.....

Penn
Michael Feldman
Nate DiGiorgio
Tara McSherry
Joellen Weaver

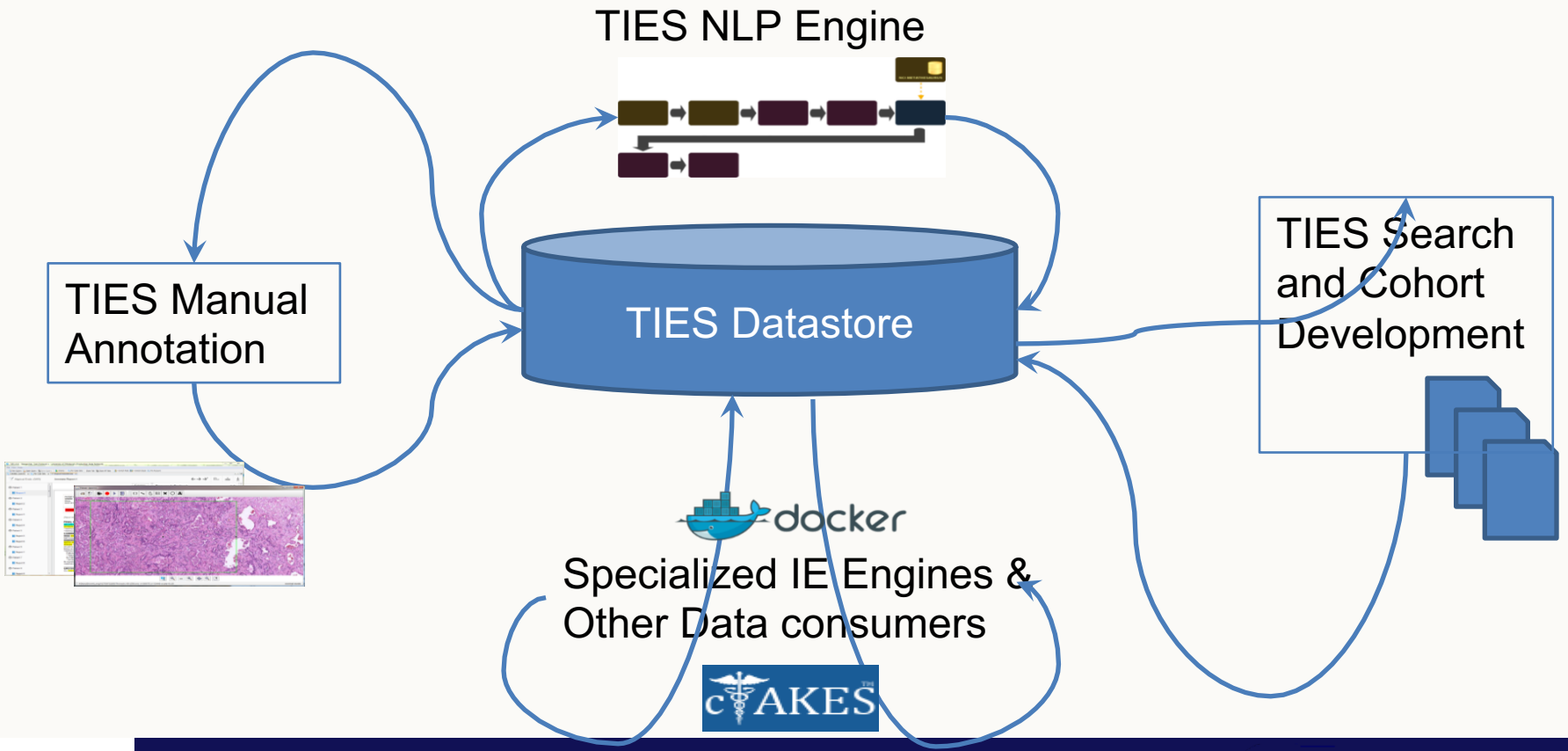
Funding

NCI U24 CA180921 Enhanced Development of TIES

Led by
Rebecca Jacobson, MD MSIS

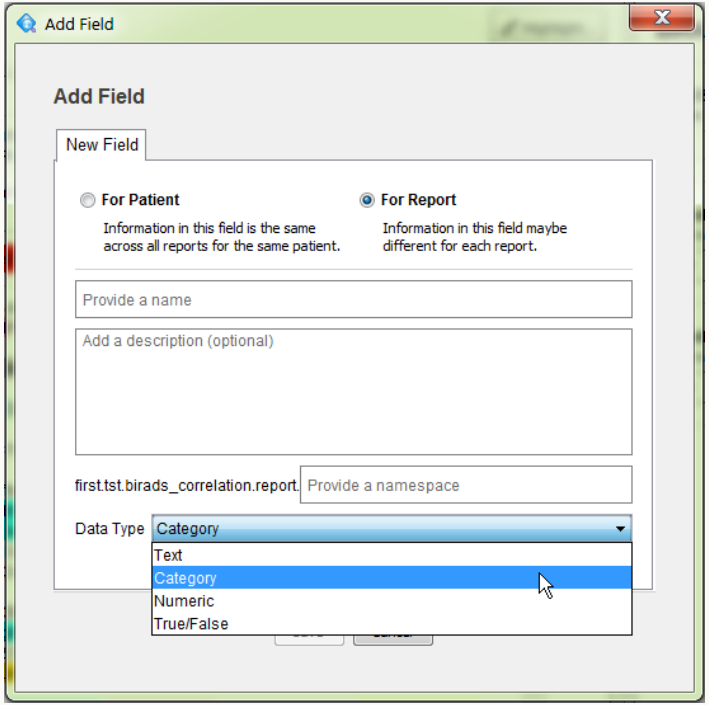
I have some
pretty big shoes to fill...

Annotation workflows

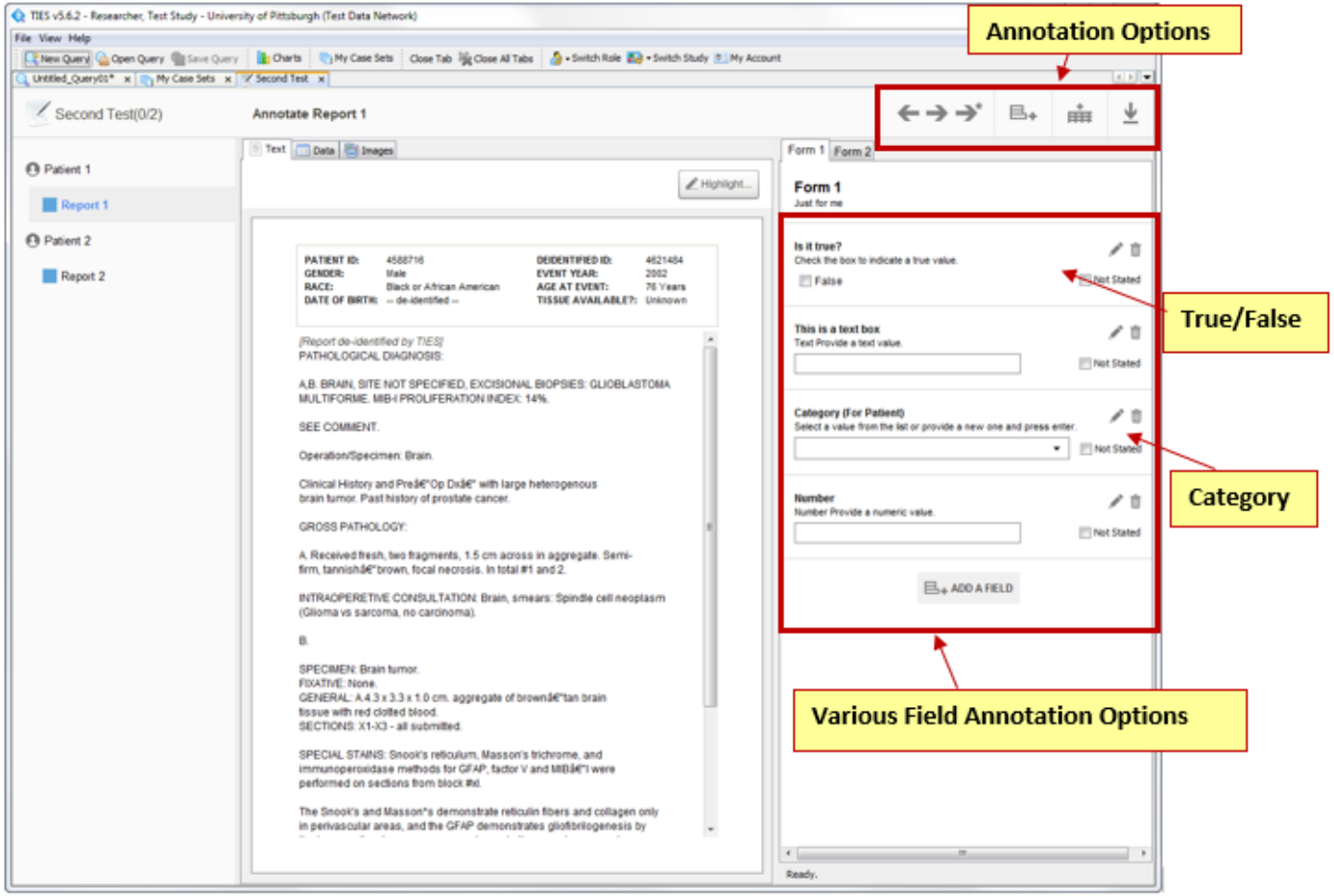


Manual Annotation Tool

- Allows you to manually enter structured data associated with case sets.
- Eliminates the need to store it in a separate spreadsheet as the expert reviews the reports.
- Data organized by forms and fields. Forms are study specific and can be shared with other study members or made public.
- Fields can be of Text, Number, Boolean and Category data types.
- Data is exported to Excel with each field stored in a separate column and a row for each report.
- Access the tool under the My Case Sets tab. Click Annotate from the Available Tasks menu under the name of the Case Set.



Manual Annotation Tool



Summary by Body System and Age at Diagnosis Report

UPMC Network Cancer Registry - Common Patients/Primaries Merged

Primary Site		Total %		0 - 29 %		30 - 39 %		40 - 49 %		50 - 59 %		60 - 69 %		70 - 79 %		80 - 89 %		90+ %
Peritoneum, Omentum & Mesentery	392	0.2%	2	0.0%	4	0.1%	33	0.2%	73	0.2%	120	0.2%	105	0.2%	53	0.2%	2	0.1%
Other Digestive Organs	111	0.1%	0	0.0%	2	0.0%	9	0.0%	14	0.0%	24	0.0%	35	0.1%	23	0.1%	4	0.1%
RESPIRATORY SYSTEM	27,939	14.0%	69	1.3%	206	3.0%	1,328	6.6%	4,392	11.2%	8,110	16.0%	9,263	18.8%	4,223	16.5%	348	11.3%
Nose, Nasal Cavity & Middle Ear	487	0.2%	15	0.3%	32	0.5%	61	0.3%	120	0.3%	105	0.2%	89	0.2%	57	0.2%	8	0.3%
Larynx	2,203	1.1%	5	0.1%	33	0.5%	179	0.9%	524	1.3%	694	1.4%	537	1.1%	214	0.8%	17	0.6%
Lung & Bronchus	25,137	12.6%	26	0.5%	133	1.9%	1,075	5.4%	3,734	9.6%	7,294	14.4%	8,609	17.5%	3,944	15.4%	322	10.4%
Pleura	18	0.0%	1	0.0%	0	0.0%	1	0.0%	2	0.0%	3	0.0%	5	0.0%	5	0.0%	1	0.0%
Trachea, Mediastinum & Other Respira	94	0.0%	22	0.4%	8	0.1%	12	0.1%	12	0.0%	14	0.0%	23	0.0%	3	0.0%	0	0.0%
BONES & JOINTS	711	0.4%	245	4.6%	61	0.9%	74	0.4%	101	0.3%	94	0.2%	85	0.2%	49	0.2%	2	0.1%
Bones & Joints	711	0.4%	245	4.6%	61	0.9%	74	0.4%	101	0.3%	94	0.2%	85	0.2%	49	0.2%	2	0.1%
SOFT TISSUE	1,762	0.9%	224	4.2%	114	1.6%	247	1.2%	286	0.7%	312	0.6%	335	0.7%	208	0.8%	36	1.2%
Soft Tissue (including Heart)	1,762	0.9%	224	4.2%	114	1.6%	247	1.2%	286	0.7%	312	0.6%	335	0.7%	208	0.8%	36	1.2%
SKIN EXCLUDING BASAL & SQUAM(	7,074	3.5%	498	9.3%	659	9.5%	1,095	5.5%	1,551	4.0%	1,321	2.6%	1,226	2.5%	652	2.5%	72	2.3%
Melanoma -- Skin	6,695	3.3%	481	9.0%	641	9.2%	1,063	5.3%	1,491	3.8%	1,252	2.5%	1,129	2.3%	583	2.3%	55	1.8%
Other Non-Epithelial Skin	379	0.2%	17	0.3%	18	0.3%	32	0.2%	60	0.2%	69	0.1%	97	0.2%	69	0.3%	17	0.6%
BASAL & SQUAMOUS SKIN	2	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	1	0.0%	1	0.0%	0	0.0%
Basal/Squamous cell carcinomas of Sk	2	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	1	0.0%	1	0.0%	0	0.0%
BREAST	35,359	17.7%	145	2.7%	1,424	20.4%	6,304	31.5%	8,710	22.3%	8,417	16.6%	6,743	13.7%	3,233	12.6%	383	12.4%
Breast	35,359	17.7%	145	2.7%	1,424	20.4%	6,304	31.5%	8,710	22.3%	8,417	16.6%	6,743	13.7%	3,233	12.6%	383	12.4%
FEMALE GENITAL SYSTEM	14,435	7.2%	206	3.8%	729	10.5%	1,860	9.3%	3,464	8.9%	3,745	7.4%	2,816	5.7%	1,442	5.6%	173	5.6%
Cervix Uteri	1,851	0.9%	83	1.6%	353	5.1%	522	2.6%	356	0.9%	272	0.5%	182	0.4%	74	0.3%	9	0.3%
Corpus & Uterus, NOS	7,552	3.8%	17	0.3%	170	2.4%	749	3.7%	2,042	5.2%	2,311	4.6%	1,517	3.1%	683	2.7%	63	2.0%
Corpus Uteri	7,426		16		168		735		2,014		2,282		1,485		665		61	
Uterus, NOS	126		1		2		14		28		29		32		18		2	
Ovary	3,448	1.7%	84	1.6%	126	1.8%	420	2.1%	788	2.0%	825	1.6%	761	1.5%	392	1.5%	52	1.7%
Vagina	188	0.1%	0	0.0%	8	0.1%	17	0.1%	32	0.1%	41	0.1%	40	0.1%	40	0.2%	10	0.3%
Vulva	1,052	0.5%	16	0.3%	64	0.9%	131	0.7%	178	0.5%	180	0.4%	232	0.5%	216	0.8%	35	1.1%
Other Female Genital Organs	344	0.2%	6	0.1%	8	0.1%	21	0.1%	68	0.2%	116	0.2%	84	0.2%	37	0.1%	4	0.1%
MALE GENITAL SYSTEM	19,091	9.5%	325	6.1%	297	4.3%	684	3.4%	4,017	10.3%	7,095	14.0%	5,134	10.4%	1,427	5.6%	112	3.6%
Prostate	18,068	9.0%	6	0.1%	9	0.1%	498	2.5%	3,923	10.0%	7,047	13.9%	5,084	10.3%	1,395	5.5%	106	3.4%
Testis	871	0.4%	314	5.9%	282	4.0%	176	0.9%	74	0.2%	15	0.0%	6	0.0%	4	0.0%	0	0.0%
Penis	127	0.1%	1	0.0%	4	0.1%	4	0.0%	16	0.0%	30	0.1%	41	0.1%	26	0.1%	5	0.2%
Other Male Genital Organs	25	0.0%	4	0.1%	2	0.0%	6	0.0%	4	0.0%	3	0.0%	3	0.0%	2	0.0%	1	0.0%
URINARY SYSTEM	14,056	7.0%	133	2.5%	243	3.5%	923	4.6%	2,224	5.7%	3,565	7.0%	4,175	8.5%	2,471	9.7%	322	10.4%
Urinary Bladder	7,384	3.7%	22	0.4%	48	0.7%	282	1.4%	924	2.4%	1,759	3.5%	2,438	5.0%	1,665	6.5%	246	8.0%
Kidney & Renal Pelvis	6,144	3.1%	110	2.1%	193	2.8%	632	3.2%	1,257	3.2%	1,678	3.3%	1,554	3.2%	658	2.6%	62	2.0%
Ureter	426	0.2%	0	0.0%	2	0.0%	5	0.0%	30	0.1%	109	0.2%	151	0.3%	118	0.5%	11	0.4%
Other Urinary Organs	102	0.1%	1	0.0%	0	0.0%	4	0.0%	13	0.0%	19	0.0%	32	0.1%	30	0.1%	3	0.1%
EYE & ORBIT	361	0.2%	28	0.5%	11	0.2%	40	0.2%	46	0.1%	86	0.2%	103	0.2%	43	0.2%	4	0.1%
Eye & Orbit	361	0.2%	28	0.5%	11	0.2%	40	0.2%	46	0.1%	86	0.2%	103	0.2%	43	0.2%	4	0.1%
BRAIN & OTHER NERVOUS SYSTEM	6,703	3.4%	862	16.1%	477	6.8%	869	4.3%	1,323	3.4%	1,274	2.5%	1,137	2.3%	671	2.6%	90	2.9%