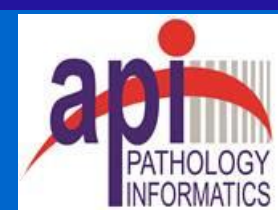




# *National Cancer Policy Forum* *February 13<sup>th</sup>, 2018*



## Computational Pathology, Artificial Intelligence and Informatics to Enhance Cancer Diagnostics



Michael J. Becich, MD PhD - [becich@pitt.edu](mailto:becich@pitt.edu)  
Chairman and Distinguished University Professor,  
Department of Biomedical Informatics  
<http://www.dbmi.pitt.edu>  
University of Pittsburgh School of Medicine



# Disclosures of COI for 2017 for MJB

## Startup/Public Companies Royalties/Licensing, Equity):

- De-ID Data Corp – de-identification software (licensing agreement)  
<http://www.de-idata.com/>
- SpIntellx – Spatial Intelligence for Cancer Diagnostics (founder equity)

## Consultancy (honoraria)

- Cancer Center Consulting – Baylor, CINJ, U Colorado, U NM
- CTSA Consulting – MCW, Northwestern, UC Davis, U Chicago, U IN, UC Davis, U NM, U WI
- Biomedical Informatics Consulting – MUSC, Northwestern, UC Davis, U Chicago, U FL

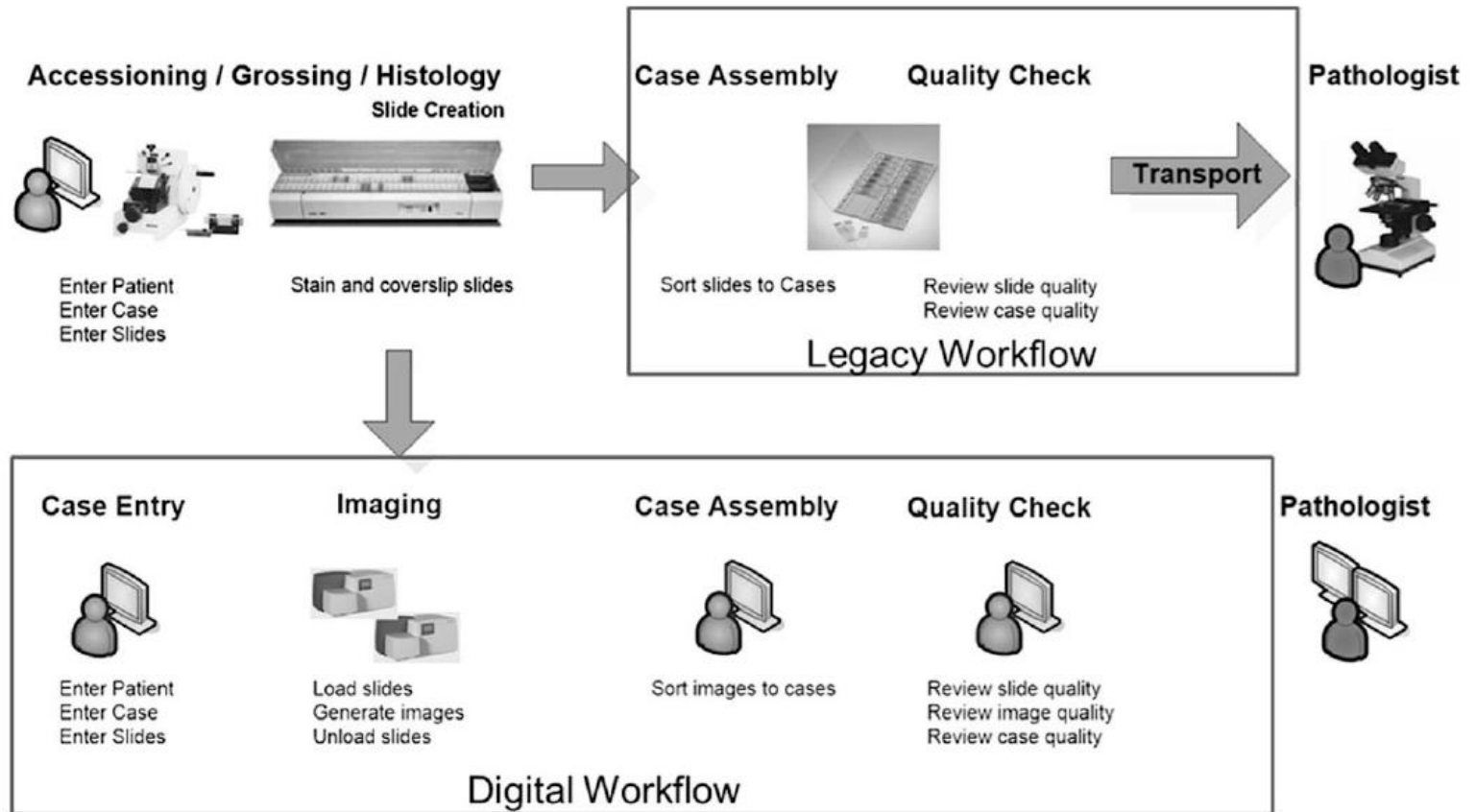
## Federal Grants

- CDC NIOSH – National Mesothelioma Virtual Bank
- NCI - CCSG, ITCR and SPORE
- NCATS - CTSA and ACT
- NHGRI - BD2K
- NLM - BMI TP
- PCORI - CDRN

# Take Home Messages


- **FDA approval of whole slide imaging is driving Computational Pathology and AI in Pathology**
- **Informatics is key to implementing Computational Pathology and AI in clinical practice**
- **Pathology and Radiology partnership can be enabled through:**
  - TIES – Text Information Extraction System
  - TCRN – TIES Cancer Research Network
- **AI will be enabled through image & data sharing via TCRN and NCI's Cancer Imaging Archive**

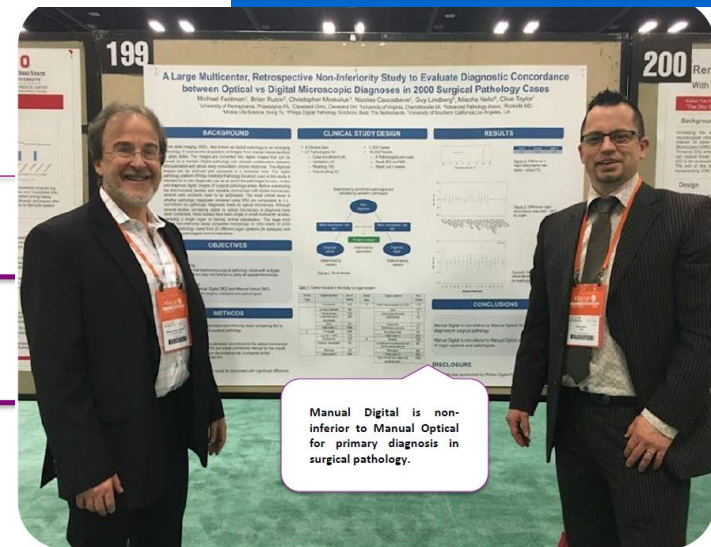
# Whole Slide Imaging (WSI) Workflow



# FDA Approval for Whole Slide Imaging

## WSI Regulatory Timeline

Date	Event	Consequence
2000	Commercial WSI devices	Digital Pathology trend started Validation of diagnostic applications
2009	FDA advisory panel 	High risk (class III) device Non-clinical use cases expanded Non-US regulatory approval DPA & select vendor discussions with FDA
2015	TPA* guideline	Assures manufacturers follow same standards
2017	FDA approval of WSI	For primary diagnosis in surgical pathology



Feldman M et al. A Large Multicenter, Retrospective Non-Inferiority Study to Evaluate Diagnostic Concordance between Optical vs Digital Microscopic Diagnoses in 2000 Surgical Pathology Cases. *Mod Pathol* 2017; 30(S2):395A.

\*TPA = Technical Performance Assessment

*From Pantanowitz, Path Info Summit, 2017*



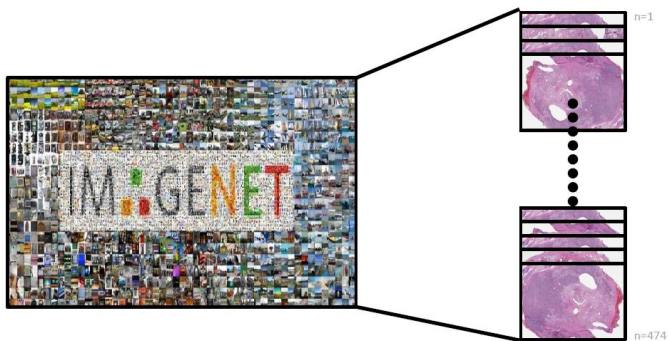
# Computational Pathology – What is it?

## Computational Path = Big Data Science Meets Pathology

- Massive Increases in Volume of Digital Data Generated from WSI PLUS genomic sequencing data –
- Heralds the rise of computational pathology
- Critical for Personalized Medicine, Learning Health Systems, Basic Research and “Big Data/Data Science”

Dataset Sizes: Computer Vision vs. Computational Pathology

*From Fuchs, 2017*

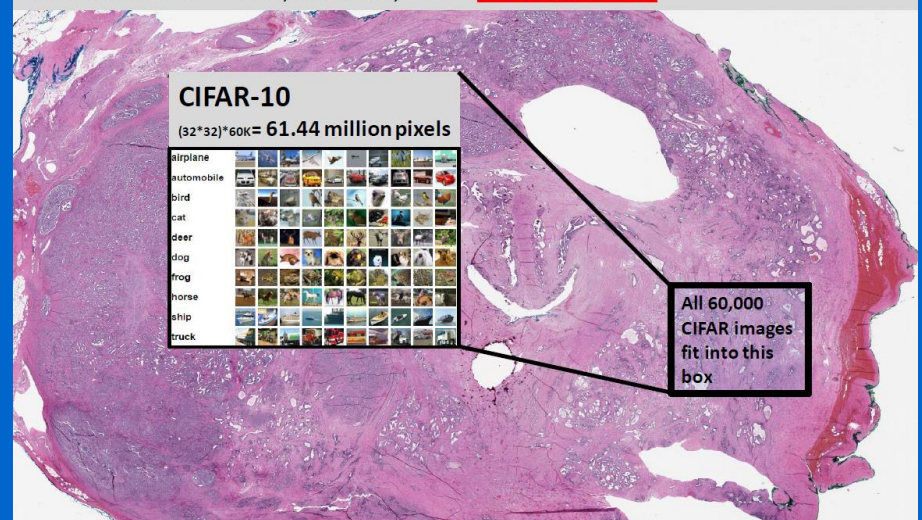


**All of ImageNet**  
482 x 415 \* 14,197,122  
= 2.8 trillion pixels

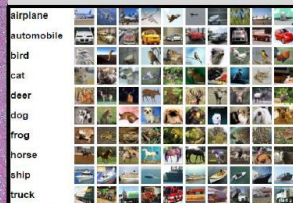
**474 Whole Slides**  
100,000 x 60,000 \* 474  
= 2.8 trillion pixels

Dataset Sizes: Computer Vision vs. Computational Pathology

1 Whole Slide = 100,000 x 60,000 = **6 billion pixels**



**CIFAR-10**  
(32\*32)\*60K = 61.44 million pixels




All 60,000  
CIFAR images  
fit into this  
box

# Definition of Computational Pathology

- An approach to diagnosis that incorporates multiple sources of data (H&E, IHC, IF & genomic data)
- Presents clinically actionable knowledge (big data to knowledge)
- Advanced decision support for precision (personalized) medicine
- Helps to redefine Pathology from an observational to knowledge engineering discipline hence critical to healthcare data science

(Louis et al Arch Path Lab Med 2014)

# Computational Pathology and Informatics Science

 A Research Data Warehouse (RDW) and robust Biorepository Supported by Informatics Science are key enablers!!!

- This will require biomedical informatics expertise:
  - Expertise in **database design and query** – *PCORnet and PaTH*
  - **Natural language processing** of text (H&P, Consults, Discharge, Pathology Reports, etc..) – *TIES/TCRN and CDP*
  - Structured capture of key medical data will require controlled **vocabularies** and implementation of **ontologies** - *TIES/TCRN & CDP*
  - **De-identification** of text for sharing with researchers (De-ID Data Corp) – *TIES/TCRN & CDP*
  - Implementation of **Bayesian algorithms** to make genomic data “actionable” **via Causal Modeling and Discovery** – *BD2K – Center for Causal Modeling and Discovery*
    - *Pitt just funded for four years and \$11M with CMU, UPMC, Yale*
- Personalized Medicine critically requires **biorepositories\***
  - **TCRN**

\*(U Pitt is #1 contributor to The Cancer Genome Atlas – see subsequent slides)





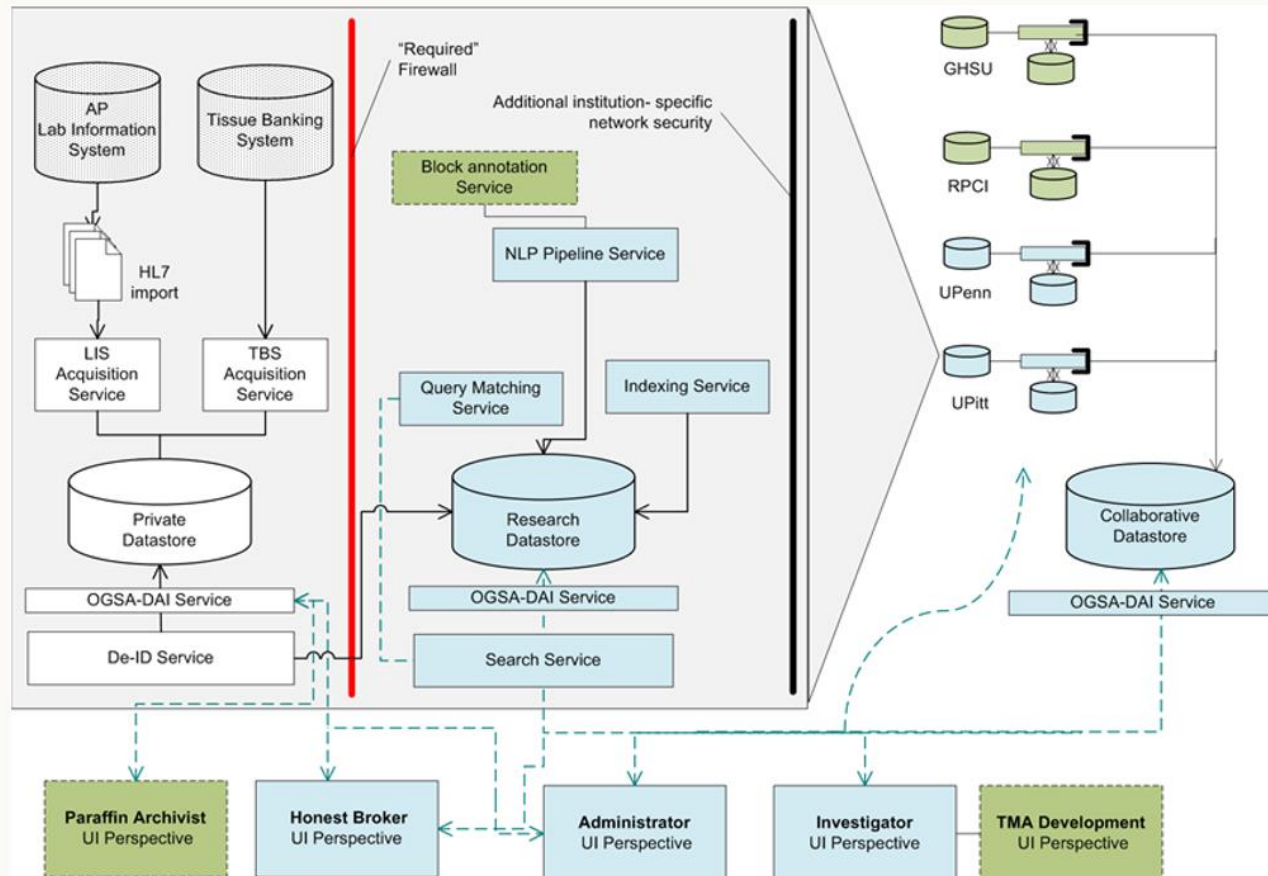
# What is TIES?

- An NLP and Information Retrieval system for de-identifying, annotating, storing and retrieving **pathology and radiology** documents
- A system for indexing research resources (FFPE, FF, WSI) 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, biospecimen and both Pathology and Radiology images** for sharing among networks of cancer centers and other institutions

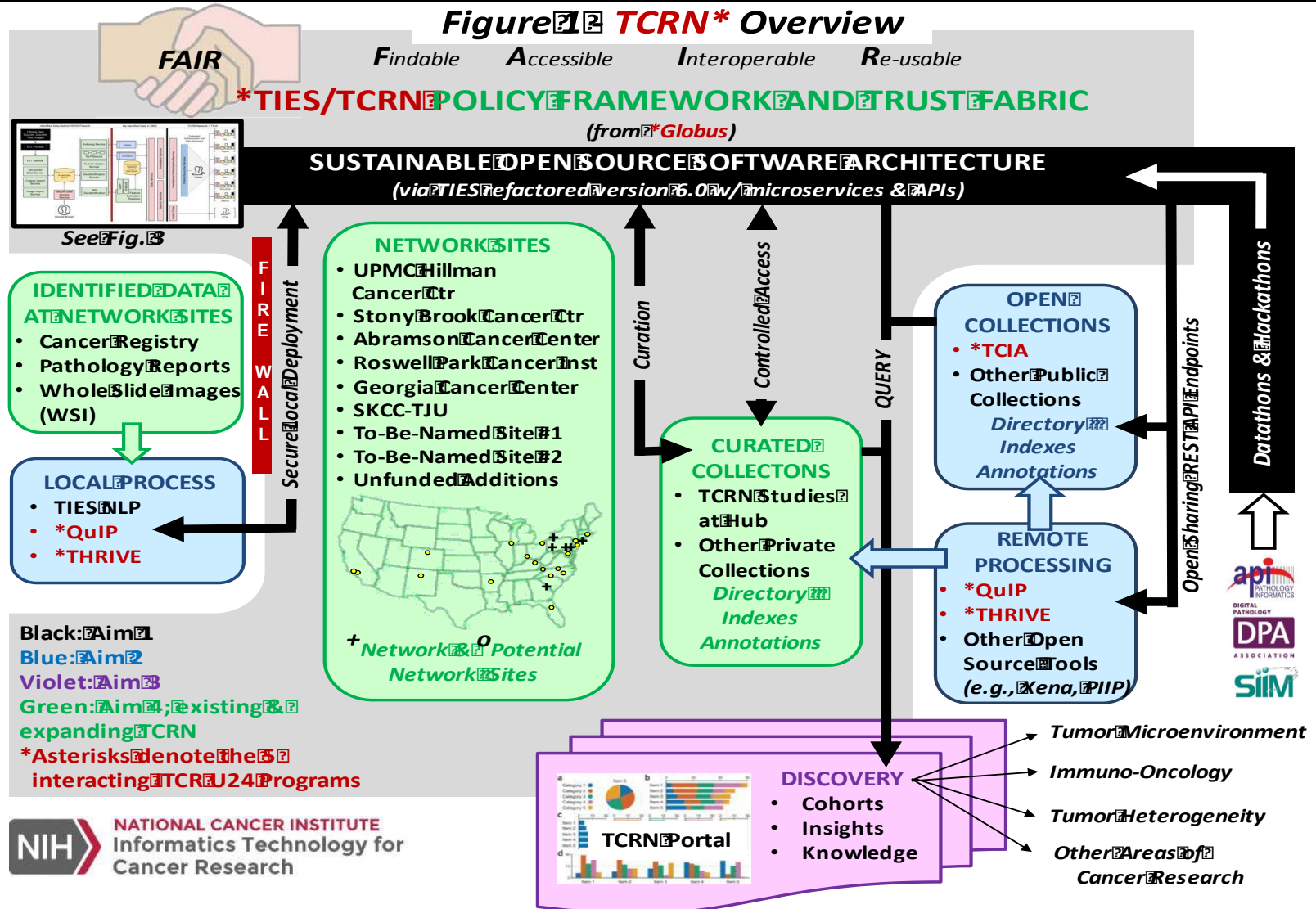
<http://ties.upmc.com/>

# Text Information Extraction System (TIES)


## TIES System Architecture



# TIES Cancer Research Network (TCRN)



# Computational Pathology and Informatics Science

 A Research Data Warehouse (RDW) and robust Biorepository Supported by Informatics Science are key enablers!!!

- This will require biomedical informatics expertise:
  - Expertise in **database design and query** – *PCORnet and PaTH*
  - **Natural language processing** of text (H&P, Consults, Discharge, Pathology Reports, etc..) – *TIES/TCRN and CDP*
  - Structured capture of key medical data will require controlled **vocabularies** and implementation of **ontologies** - *TIES/TCRN & CDP*
  - **De-identification** of text for sharing with researchers (De-ID Data Corp) – *TIES/TCRN & CDP*
  - Implementation of **Bayesian algorithms** to make genomic data “actionable” **via Causal Modeling and Discovery** – *BD2K – Center for Causal Modeling and Discovery*
    - *Pitt just funded for four years and \$11M with CMU, UPMC, Yale*
- Personalized Medicine critically requires **biorepositories\***
  - **TCRN**

\*(U Pitt is #1 contributor to The Cancer Genome Atlas – see subsequent slides)



# Pitt Center for Causal Modeling and Discovery

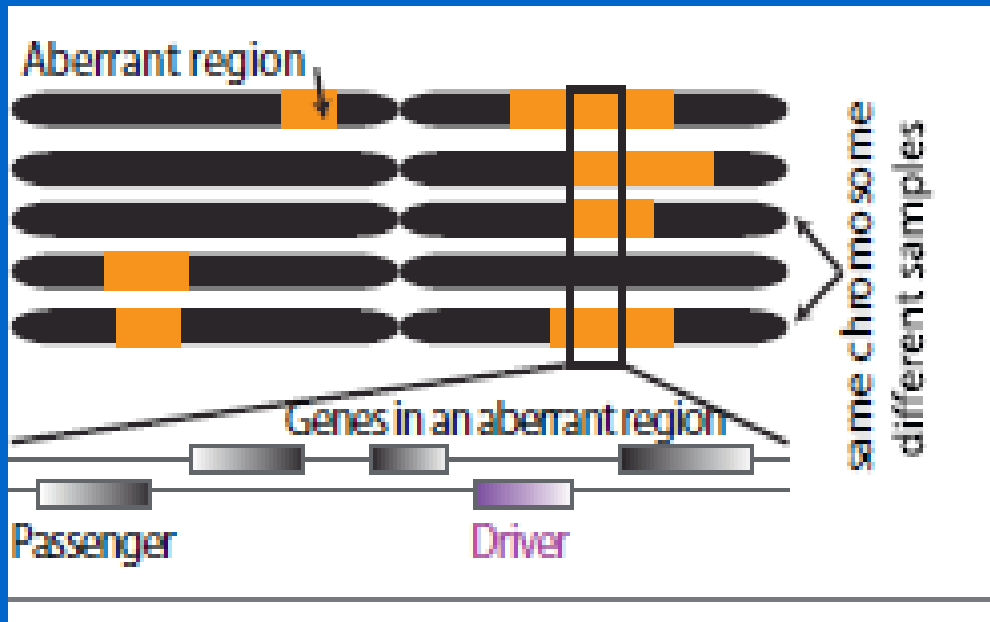
- Led by Greg Cooper, MD PhD (Biomedical Informatics), Ivet Bahar, PhD (Computational and Systems Biology) and Jeremy Berg, PhD (Director of Institute for Personalized Medicine) involves Becich/Crowley – TCGA/PGRR/Computable Phenotypes
- Theme: Modeling and discovery of causal networks from genome and phenome (patient EHR/i2b2 RDW) biomedical datasets including those from clinical trials and other controlled biospecimen based efforts
- Aims
  - **Represent** causal knowledge within a unified, formal framework
  - **Discover** causal knowledge from biomedical data (both observational and experimental) and background knowledge (e.g., from the literature) using efficient algorithms
  - **Apply** causal knowledge to support browsing, answering causal queries, simulating causal processes, and designing experiments to resolve causal uncertainties
- Driving Biological Problem areas: Signaling Pathways in Cancer/TCGA (**Computational Pathology and Genomics/TCGA**), Idiopathic Pulmonary Fibrosis (**Computational Pathology Imaging**) and “Connectome” via fMRI brain imaging



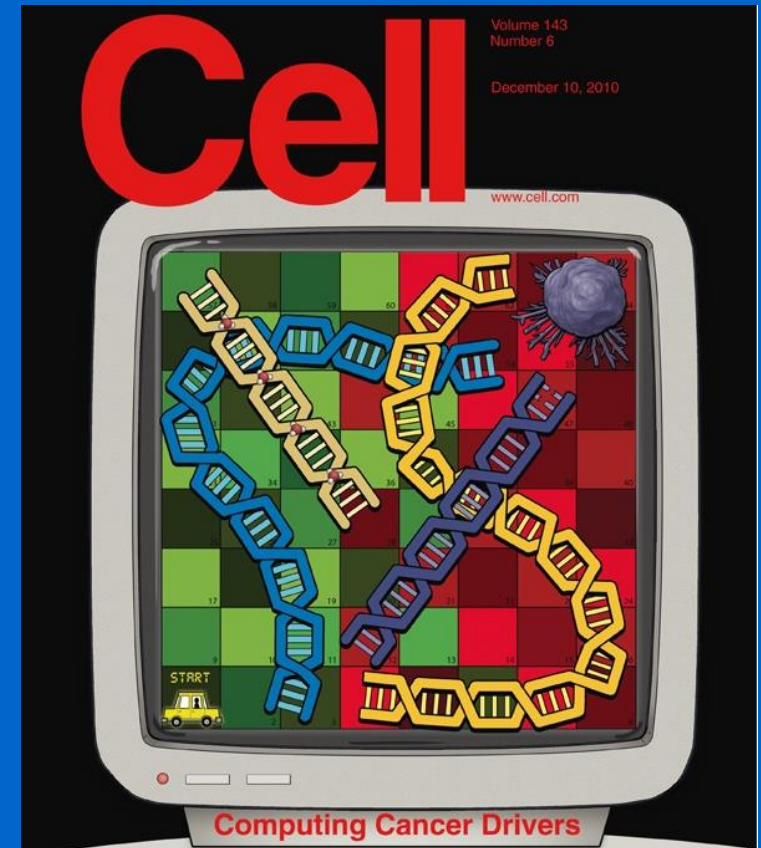


# Causal Network Discovery = Computational Pathology

A probabilistic network approach to uncover genetic drivers of melanoma using data on copy number variation and gene expression\*



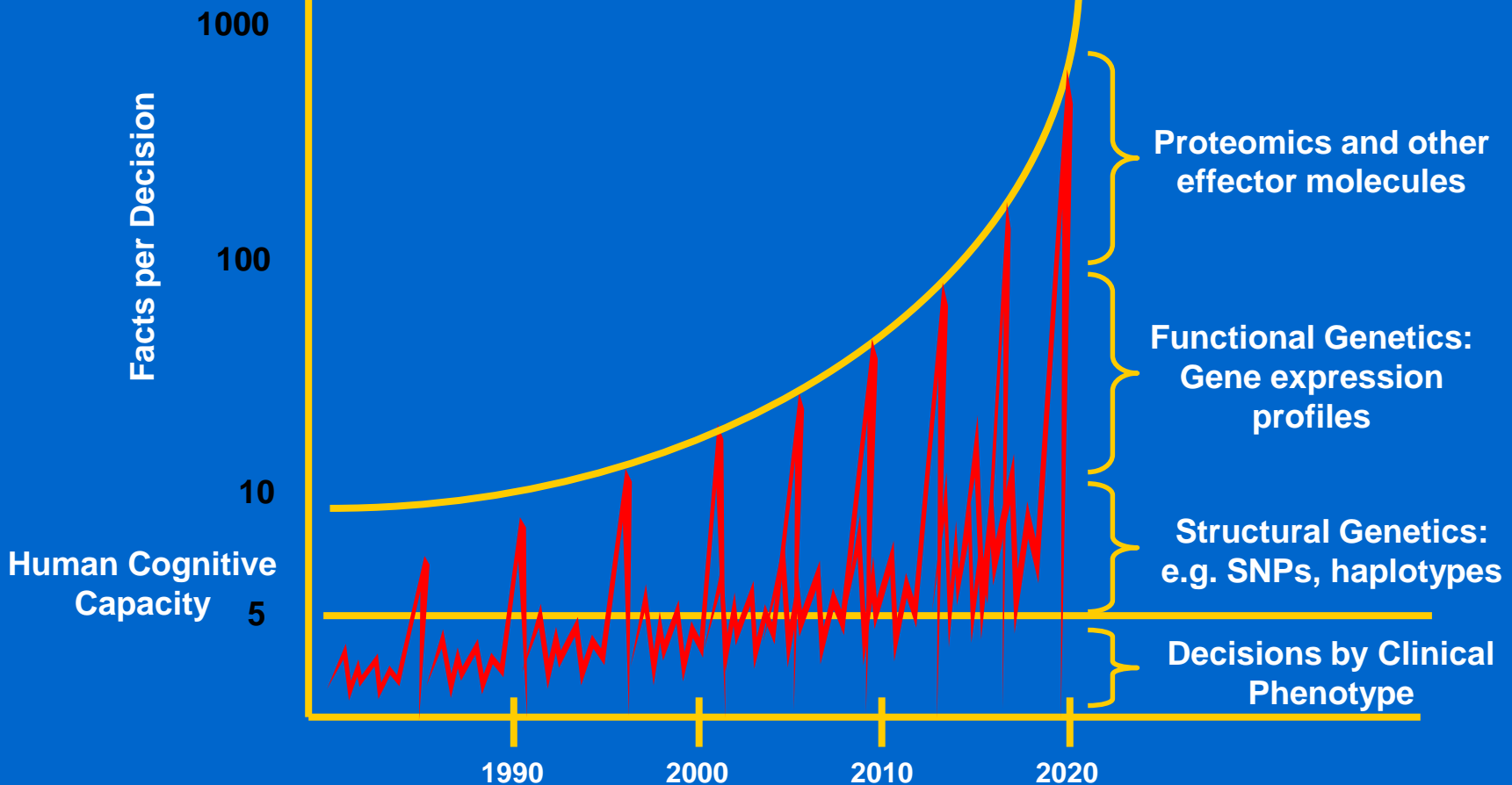
Akavia UD, et al. *Cell* 143 (2010) 1005-1017.  
(The figure above appears in this paper)



**10 December, 2010 Volume 143, Issue 6**

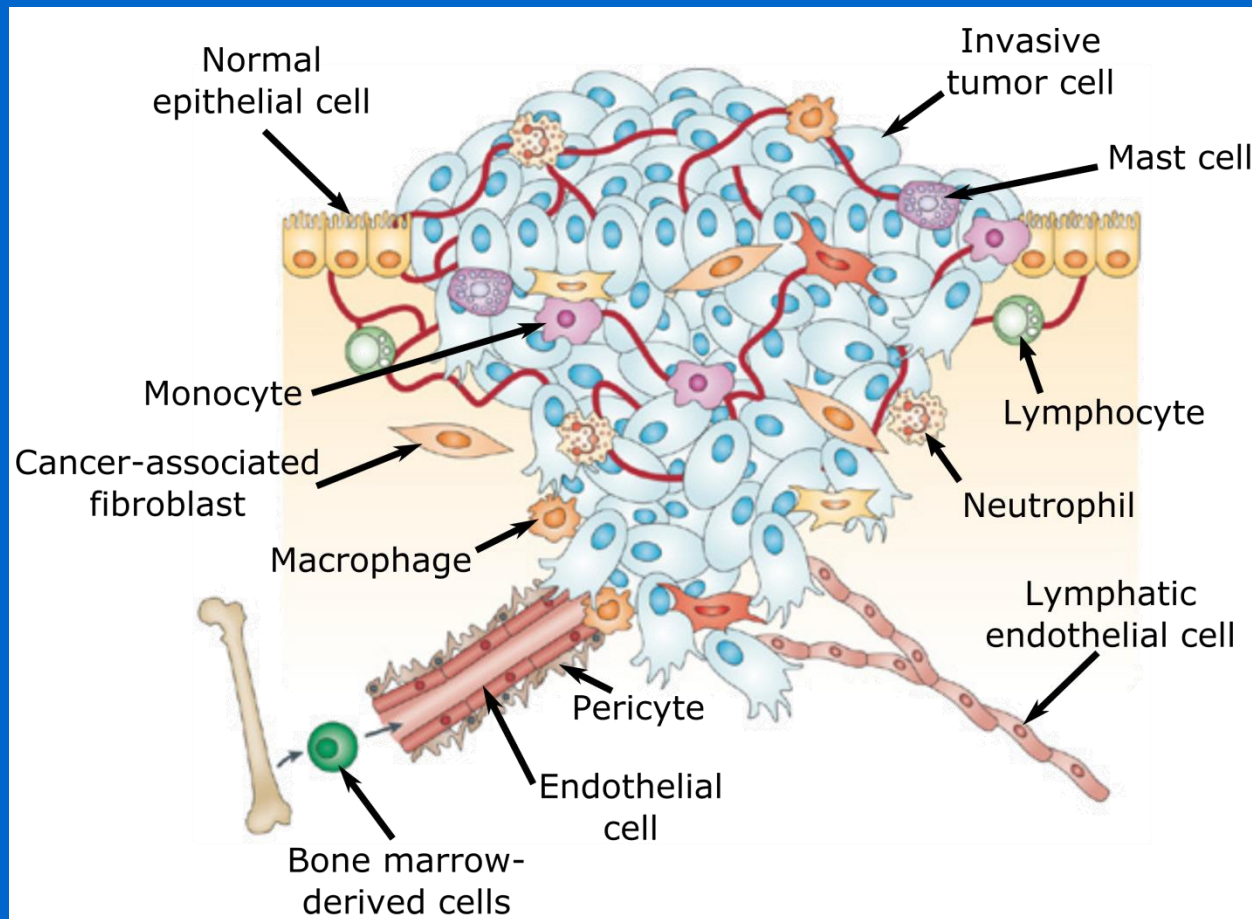
On the cover: In cancer, heterogeneous genetic aberrations frequently result in similar phenotypic outcomes. In this issue, Akavia et al. (pp. 1005–1017) report a computational algorithm that identifies driving mutations and links them with their downstream transcriptional effects. The approach thus allows the authors to begin disentangling the complex mechanisms by which genetic aberrations drive transformation. The cover shows a driver with many DNA-mediated paths to a tumor.

# Computational Pathology Algorithms like those in CCD for imaging and genomics are key enablers!!!

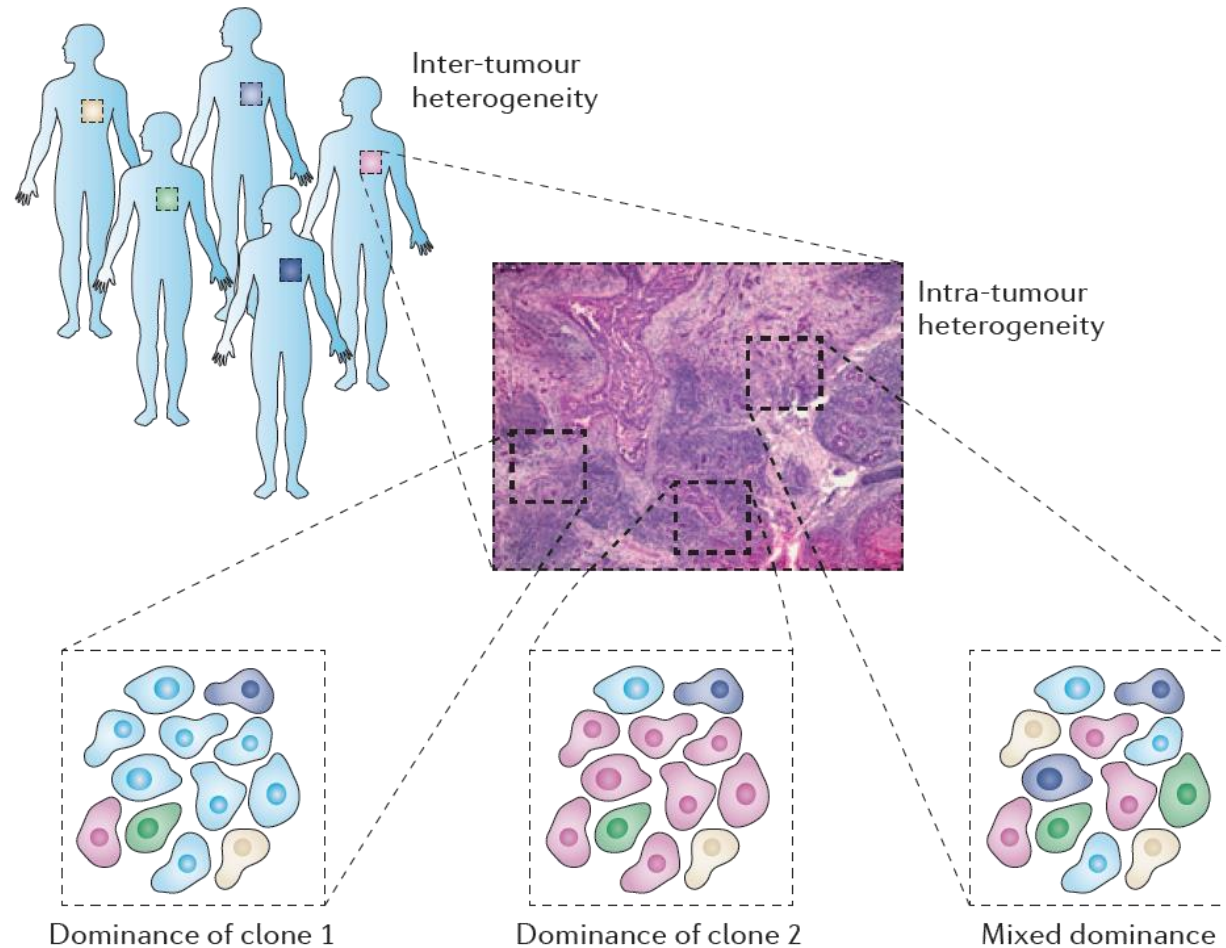


From William Stead: [http://courses.mbl.edu/mi/2009/presentations\\_fall/SteadV1.ppt](http://courses.mbl.edu/mi/2009/presentations_fall/SteadV1.ppt)  
& [http://www.mbl.edu/education/courses/special\\_topics/pdf/med\\_sched09\\_fall.pdf](http://www.mbl.edu/education/courses/special_topics/pdf/med_sched09_fall.pdf)

# 1. Intra-tumoral spatial heterogeneity

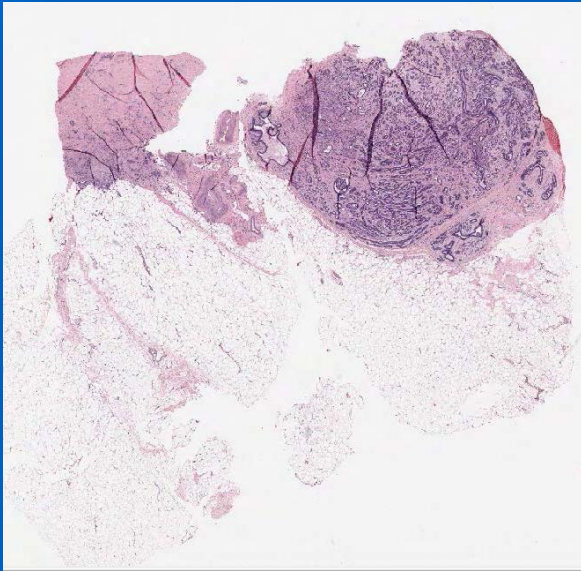


# 1. Quantify Heterogeneity for Diagnostics, Prognostics & Immunotherapy



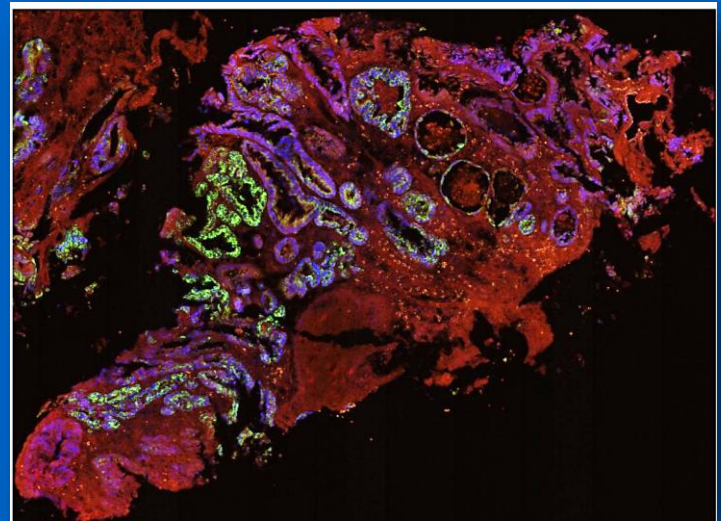
## 2. Integration of H&E, IHC and IF

### H&E in transmitted light



- Universal method
- Limited molecular measurements in transmitted (IHC)
- Complex tissue “scenes”

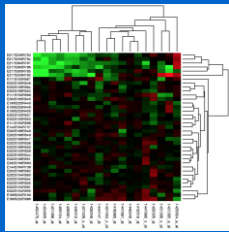
### Multi to hyperplexed



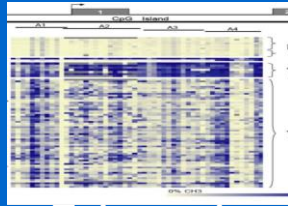
- Emerging method with potential to measure many DNA/RNA/proteins in the same tissue section/TMA
- Structural biomarkers
- Quantify biomarker expression levels and tissue based spatial relationships.



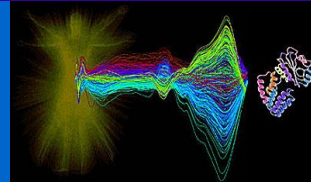
# 3. Histopathonomics: Spatial ITH



Genomics



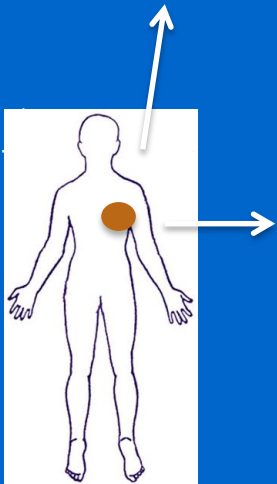
Epigenomics



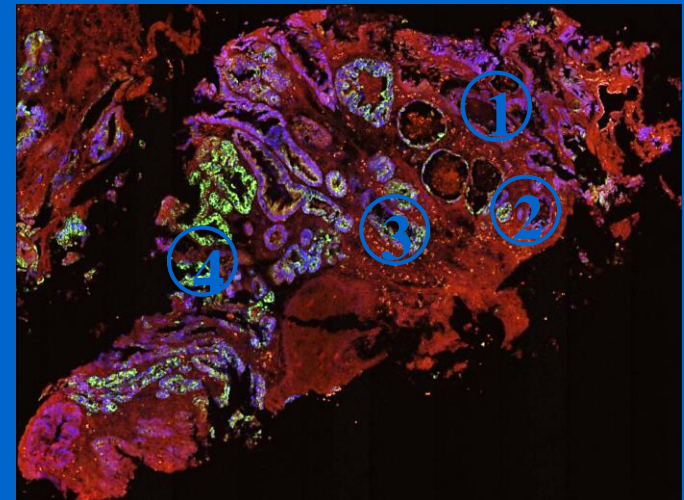
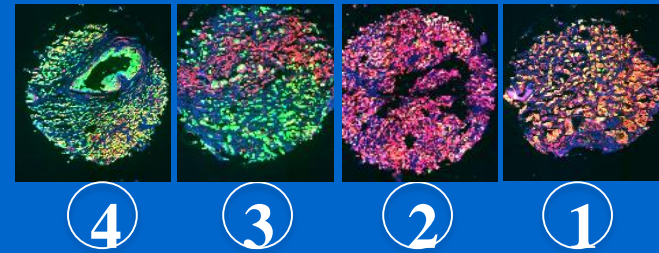
Proteomics



Metabolomics



H&E stained whole tissue section  
from FFPE tumor sample



Multi to hyperplexed fluorescence imaging of whole  
section for higher spatial resolution and tissue context

## 4. Ground-truth & Annotation

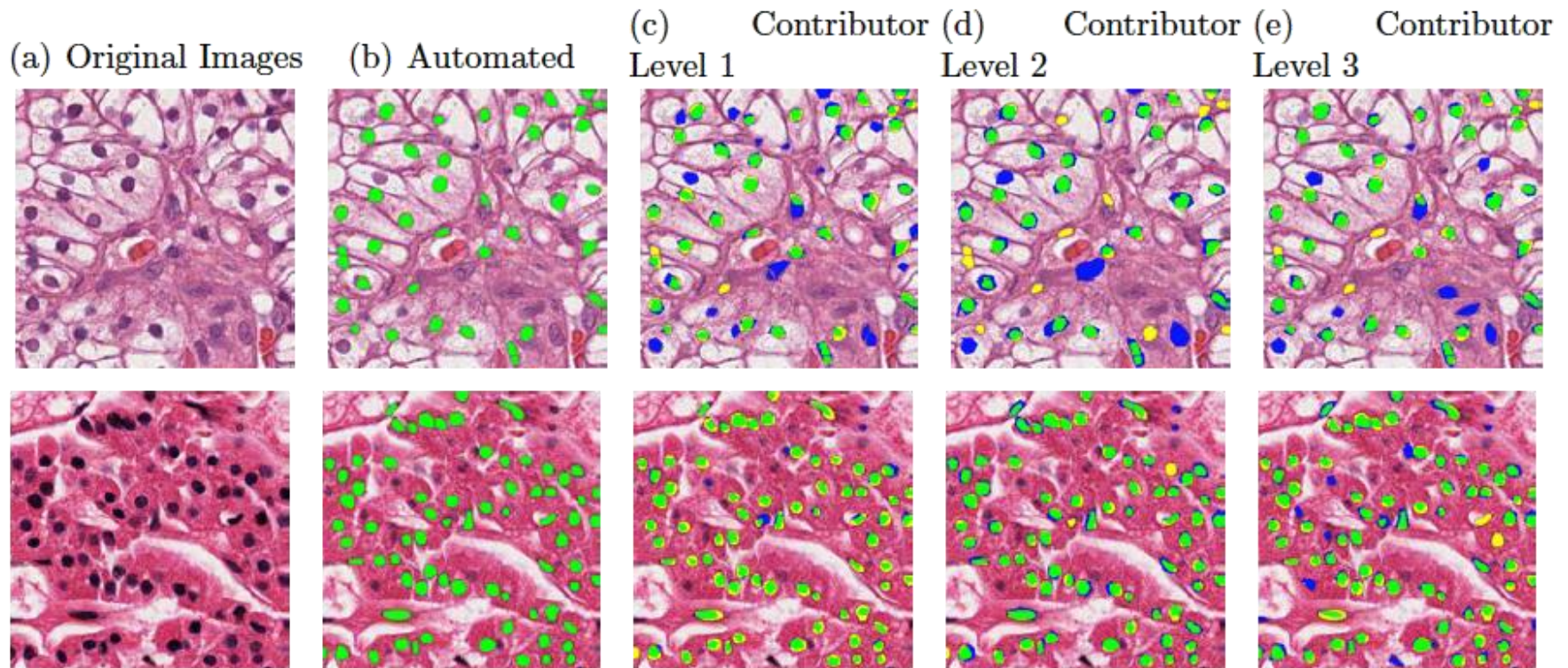
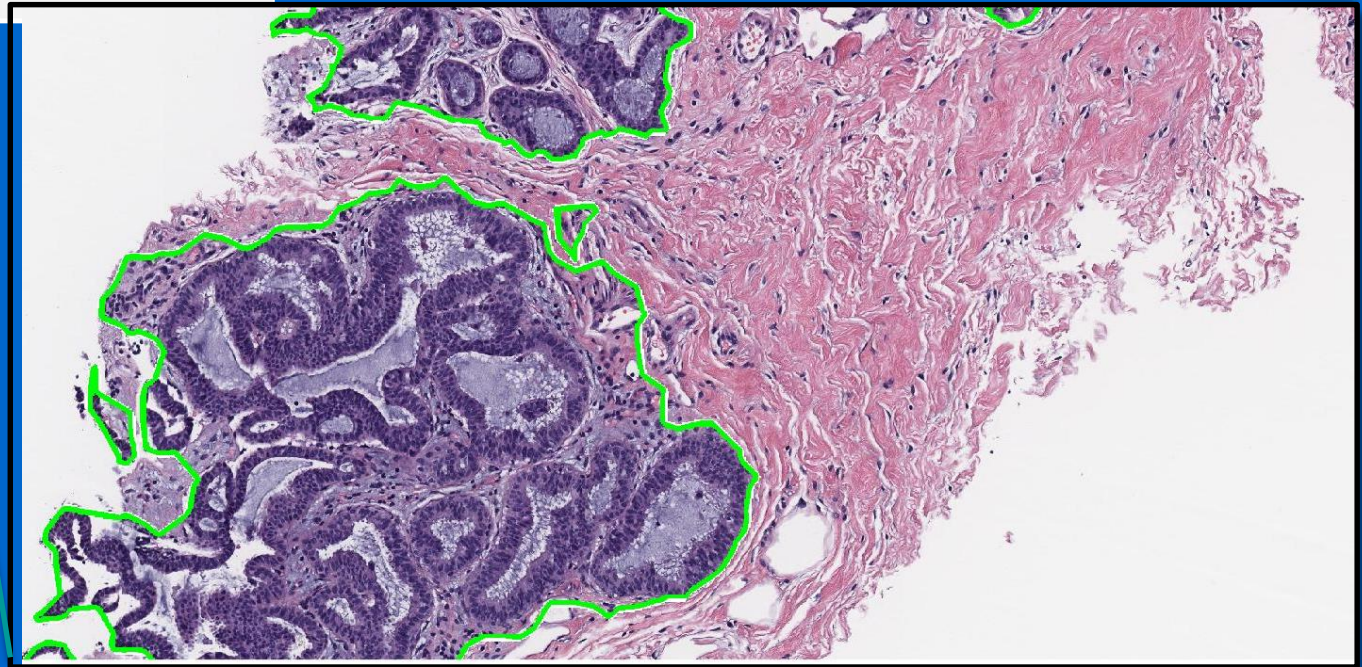
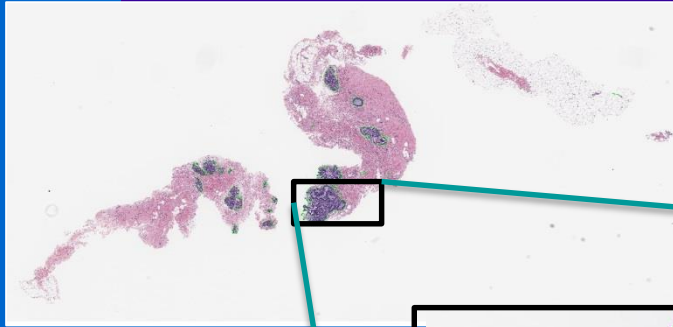


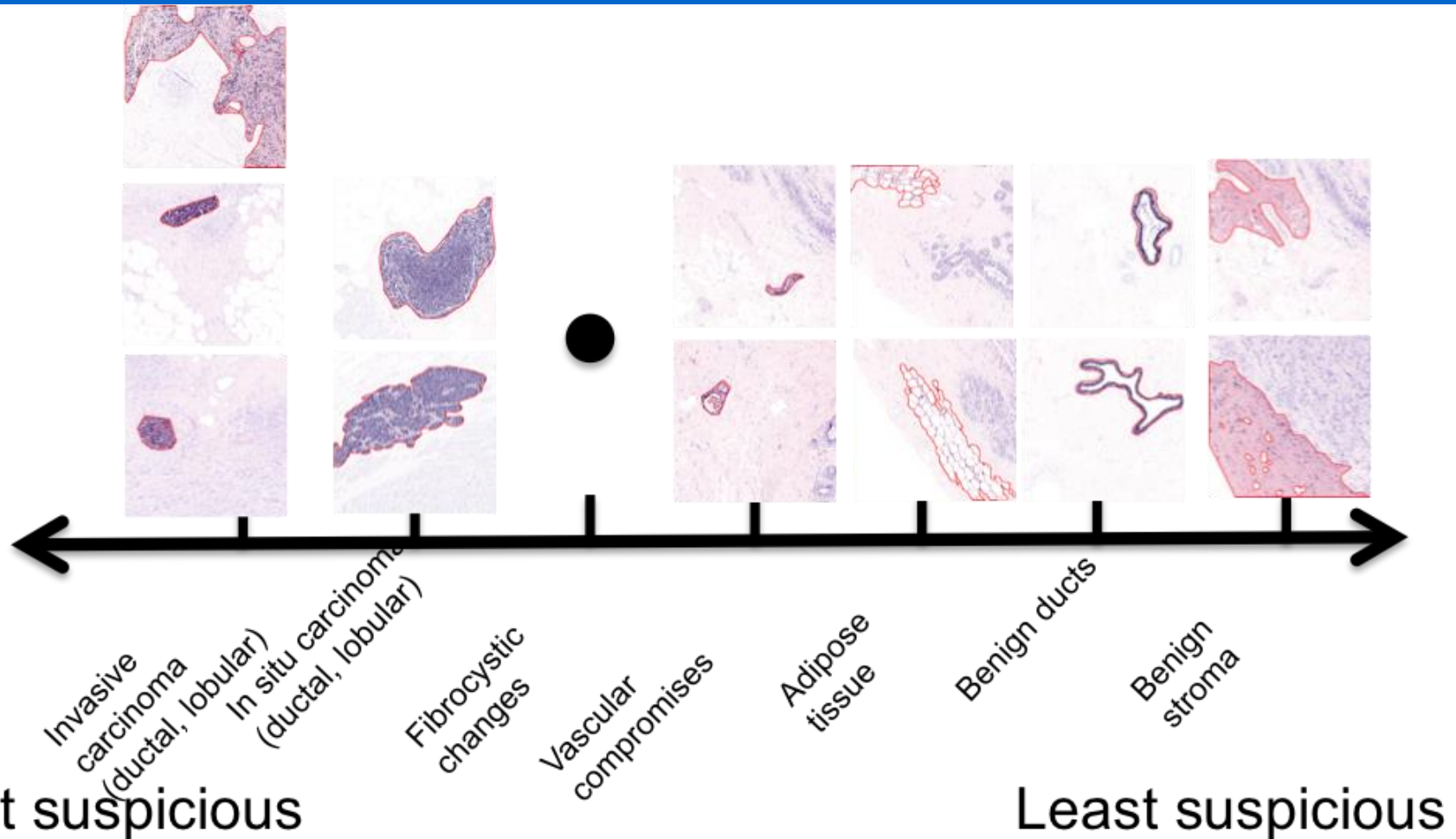
Fig. 2. Examples of nuclear segmentation using an automated method and increasing contributor skill level, ranging from 1 to 3. (Green region indicates TP region, yellow region indicates FN region and blue region indicates FP region). The automated nuclei segmentation used as ground truth.



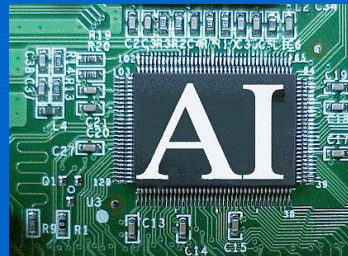
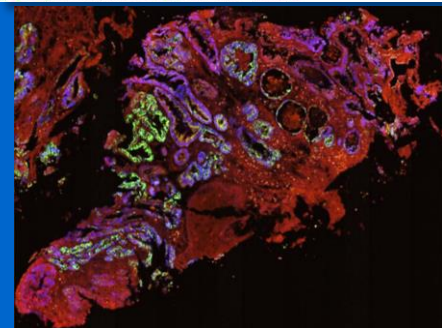
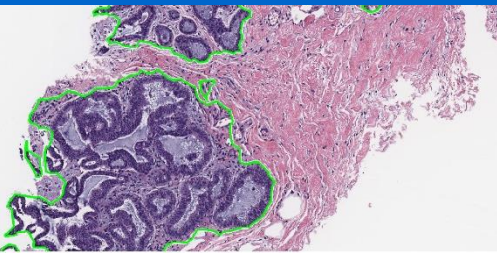
## 4. "Callable" ID of Histologic Features



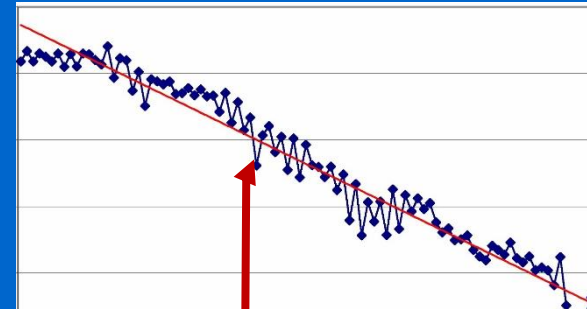
## 5. Triaging ROIs



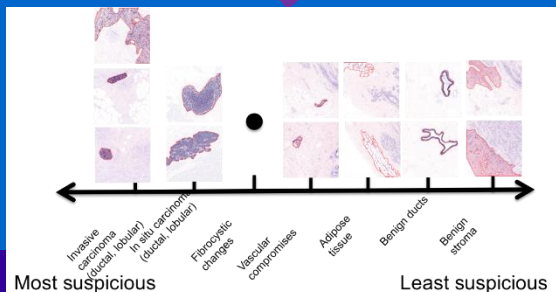
# 6. Computational Pathology and AI



Error over time



Intervention

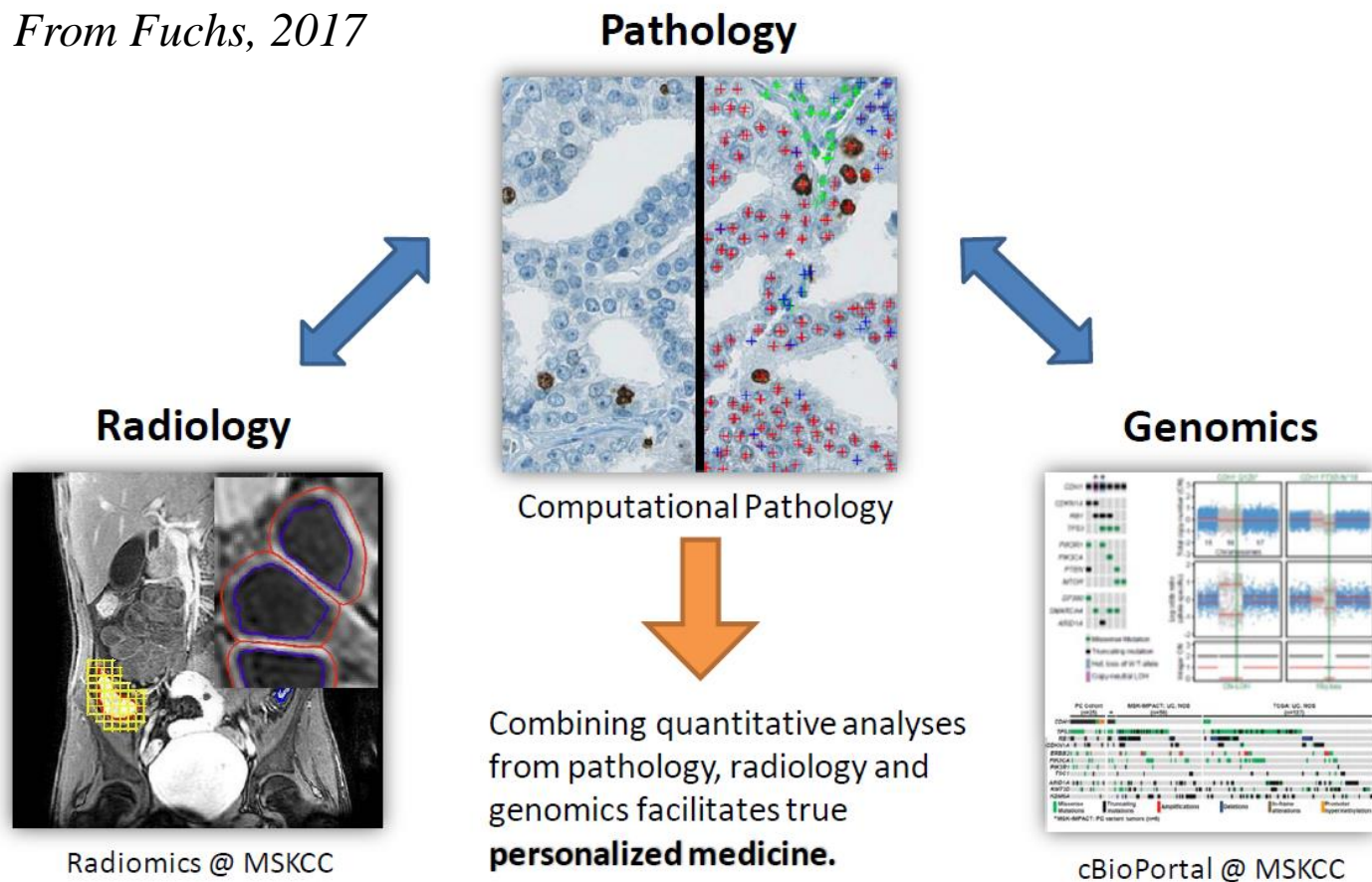


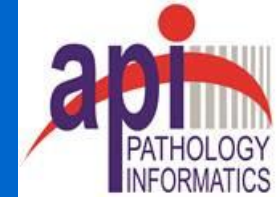


# Pathology and Radiology Must Partner!

## A Joint Effort for Personalized Medicine

*From Fuchs, 2017*





- 
- 
- End of Talk – e-mail me at [becich@pitt.edu](mailto:becich@pitt.edu) if you have
- questions/clarifications not covered in the discussion.

NOTE: E-mail me if you want PDFs of articles or presentation.



**Thank you for organizing NCPF**




# Association for Pathology Informatics (API)

<http://www.pathologyinformatics.org>




*“...to advance  
the field of  
pathology  
informatics as an  
academic and a  
clinical  
subspecialty of  
pathology...”*

Slide 26



## Association for Pathology Informatics




[Home](#) [Membership](#) [Membership Benefits](#) [Meetings](#) [Journal](#) [Awards](#) [Training](#) [Donate](#) [Resources](#) [Sponsors](#) [Contact Us](#) [LIS Toolkit](#)




### Job Opportunities

Fellowship Posts  
Other Job Posts


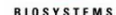

### 2015 Conference Diamond Level Sponsor



### Corporate Sponsors



PHOTON IS OUR BUSINESS



### Pathology Informatics Summit - May 5 - 8, 2015 in Pittsburgh


**ALL ONLINE REGISTRATION ENDS WEDNESDAY, APRIL 29, 2015**

Register now:  
<http://pathologyinformatics.com/content/registration>

Deadline for reservations at the meeting discount is Monday, April 20, 2015

Make your reservations at the Wyndham [here](#)


2015 Schedule:  
<http://www.pathologyinformatics.com/content/pathology-informatics-summit-2015-schedule>



[Click here for the PDF.](#)

### New Webinar: Digital Pathology Meets Surgical Pathology

The Association for Pathology Informatics (API) and Sunquest Information Systems sponsored a webinar titled, “Digital Pathology Meets Surgical Pathology” on Thursday, March 5, 2015, at 1 PM EDT/10 AM PDT, presented by Stephen M. Hewitt, MD, PhD, FCAP, FASCP, Clinical Investigator in the Laboratory of Pathology, Center for Cancer Research, National Cancer Institute.




[View the Webinar](#)


### New Webinar: 10 Years of Direct Access Genetics: What Have We Learned?

The Association for Pathology Informatics (API) and Sunquest Information Systems sponsored a webinar titled, “10 Years of Direct Access Genetics: What Have We Learned?” on Thursday, February 19, 2015, at 12 PM EDT/9 AM PDT, presented by Jill Hagenkord, MD, FCAP, Chief Medical Officer of 23andMe.


### Teaching Institution Sponsors




Albert Einstein College of Medicine  
OF YESHIVA UNIVERSITY




Beth Israel Deaconess  
Medical Center




Cleveland Clinic




COLUMBIA UNIVERSITY  
IN THE CITY OF NEW YORK




DALHOUSIE  
UNIVERSITY  
*Inspiring Minds*




Duke University  
School of Medicine




EMORY UNIVERSITY  
Department of  
Biomedical Informatics  
Emory University School of Medicine




EMORY  
UNIVERSITY  
SCHOOL OF  
MEDICINE



Henry Ford  
HEALTH SYSTEM



Icahn School of Medicine at  
Mount Sinai



ROBERT WOOD JOHNSON  
MEDICAL SCHOOL

# Journal of Pathology Informatics

Co-Editors Liron Pantanowitz, MD PhD and Anil Parwani, MD PhD

Please support JPI,  
API and  
Pathology  
Informatics as the  
Home for Digital  
Pathology -  
Great Academic  
and Strategic  
Partnership with  
Multiple  
Benefits!!!



The screenshot shows the homepage of the Journal of Pathology Informatics. At the top, the journal's name is displayed in a large, serif font, with the subtitle "Publication of Association for Pathology Informatics, USA" below it. A navigation bar contains links for JOURNAL, EDITORIAL BOARD, BROWSE ARTICLES, INSTRUCTIONS, SUBMISSION, SUBSCRIBE, and ADVERTISE. To the right of the navigation bar are social media icons and a "Contact us" link. Below the navigation bar, there is a section for "Latest published articles" featuring three articles. Each article includes a title, authors, a background summary, the journal issue information, and links to the abstract, full text, PDF, and EPub versions. The first article is "Distance reporting in digital pathology: A study on 950 cases" by Aleksandar Vodovnik. The second is "Imaging file management to support international telepathology" by Liron Pantanowitz, Jeffrey McHugh, William Cable, Chengquan Zhao, and Anil V Parwani. The third is "Default settings of computerized physician order entry system order sets drive ordering habits" by Jordan Olson, Christopher Hollenbeak, Keri Donaldson, Thomas Abendroth, and William Castellani. At the bottom of the page, there is a footer with a sitemap, contact information, and various accreditation logos including COC, Open Access, ROME, Dublin Core, W3C HTML 4.1, W3C CSS, ACAP, and ENABLED. The footer also includes the ISSN number and a link to the mobile site.

Journal of Pathology Informatics  
Publication of Association for Pathology Informatics, USA

JOURNAL EDITORIAL BOARD BROWSE ARTICLES INSTRUCTIONS SUBMISSION SUBSCRIBE ADVERTISE

Contact us | Home | Login | Users Online: 4

<http://www.jpathinformatics.org>

**Latest published articles**

**Distance reporting in digital pathology: A study on 950 cases**  
Aleksandar Vodovnik  
**Background:** Increased workload, case complexity, financial constraints, and staffing shortages justify wider implementations of digital pathology. One of its main advantages is distance reporti...  
*J Pathol Inform* 2015, 6:18 (30 April 2015)  
[\[Abstract\]](#) | [\[HTML Full text\]](#) | [\[PDF\]](#) | [\[Mobile HTML Full text\]](#) | [\[EPub\]](#)

**Imaging file management to support international telepathology**  
Liron Pantanowitz, Jeffrey McHugh, William Cable, Chengquan Zhao, Anil V Parwani  
**Background:** Telepathology practice across international borders has become increasingly popular. Our telepathology consultation service with a laboratory in China was hampered by latency issues...  
*J Pathol Inform* 2015, 6:17 (24 March 2015)  
[\[Abstract\]](#) | [\[HTML Full text\]](#) | [\[PDF\]](#) | [\[Mobile HTML Full text\]](#) | [\[EPub\]](#)

**Default settings of computerized physician order entry system order sets drive ordering habits**  
Jordan Olson, Christopher Hollenbeak, Keri Donaldson, Thomas Abendroth, William Castellani  
**Background:** Computerized physician order entry (CPOE) systems are quickly becoming ubiquitous, and groups of orders ("order sets") to allow for easy order input are a common feature. This provi...  
*J Pathol Inform* 2015, 6:16 (24 March 2015)  
[\[Abstract\]](#) | [\[HTML Full text\]](#) | [\[PDF\]](#) | [\[Mobile HTML Full text\]](#) | [\[EPub\]](#)

[Read more....](#)

Sitemap | What's New | Feedback | Disclaimer |  
© Journal of Pathology Informatics | Published by Wolters Kluwer - Medknow  
Online since 10<sup>th</sup> March, 2010  
Editorial and Ethics Policies

CC BY-NC-SA Open Access ROME Dublin Core Metadata W3C HTML 4.1 W3C CSS ACAP ENABLED View mobile site

ISSN: Print -2229-5089, Online - 2153-3539



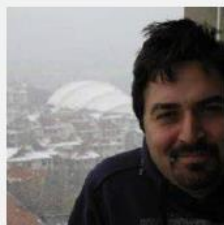
# Thanx to the Pitt team!



Dan Spagnolo



Dr. Shikhar Uttam



Dr. Burak Tosun



Lương Nguyễn



Maurice Marx



Dr. Om Choudhary



Dr. Filippo Pullara



Akash Parvatikar



Nathan Ong



Chakra



Alex Andonian



Tammy Ma

- UPMC & Center for Commercial Applications of Healthcare Data Grant #711077
- NIH-NCI U01 CA204826
- NHGRI BD2K U54HG008540 (Lung DBP)

Drug Discovery Institute/CSB

**Dr. Lansing Taylor**

Dr. Andy Stern

Dr. Bert Gough

Dr. Tim Lezon

UPMC/Magee/Pathology

**Dr. Jeffrey Fine**

Center for Commercial Applications of  
Health Care Data/DBM

**Dr. Michael Becich**

**Dr. Don Taylor**

GE Global Research

**Brion Sarachan**

**Dr. Yousef Al-Kofahi**

**Dr. Fiona Ginty**

Magee-Womens Research Institute

Dr. Adrian Lee

Dr. Rekha Gyanchandani

CSB

Dr. Takis Benos

**Dr. Dimitrios Manatakis**

**Dr. Joe Ayoob**

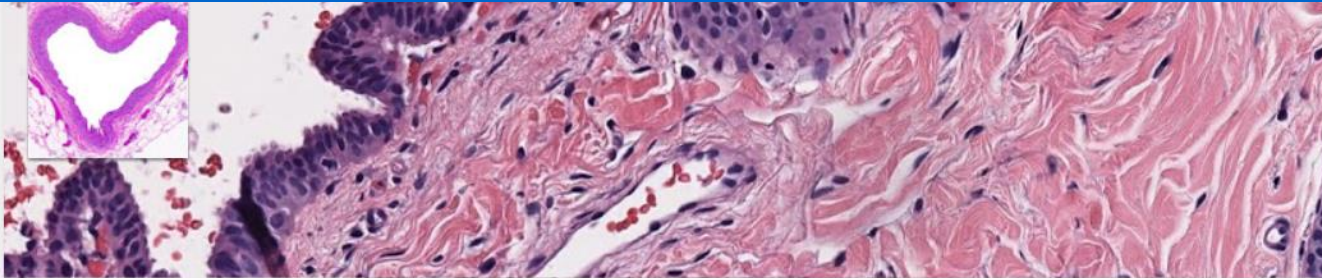


•  
•  
•  
**Thanx to Chakra Chennubhotla, PhD & Burak Tosun, PhD**



## Computational Pathology Interest Group and Lecture Series

<http://www.csb.pitt.edu/comppath/>  
Youtube Channel for Lectures



### Pittsburgh CompPath Lecture Series

Pittsburgh Computational Pathology Interest Group and Lecture Series is a special group formed to bring focus on developing basic compute... [Show more](#)

<https://www.youtube.com/watch?v=Xrvxc0YNcAM&feature=youtu.be>

• • • • • • • • • •

# Comp Path Team Profile - Pittsburgh

- **Special Interest Group Members (n=93):**
  - Bioengineering (5%)
  - Biomedical Informatics (10%)
  - Computational and Systems Biology (10%)
  - Computer Science and Machine Learning (20%)
  - Industry and Entrepreneurs (20%)
  - Medicine (10%)
  - Pathology Informatics (10%)
  - UPMC (15%)

Participation from Carnegie Mellon & Duquesne Universities