

# Algorithms, rapid analyses, and data integrity in clinical practice

Samuel Volchenboum, MD, PhD

University of Chicago

March 23, 2018



# Agenda

- Pediatric cancer
- Innovations in
  - data processing
  - clinical trials
  - data collection



New therapies are needed for pediatric cancer



# Pediatric cancer is rare

## Adult cancers annual incidence

All	1,688,780
Oral	49,670
GI	310,440
Lung	222,500
Skin	95,360
Breast	255,180
Ovary	22,440
Prostate	161,360
Urinary	146,650
Lymphoma	80,500
Myeloma	30,280
Leukemia	62,130

Source: cancer.org - 2017

## Pediatric cancers annual incidence

All	15,780
ALL	3,080
CNS	2,780
Hodgkin lymphoma	1,180
NHL	1,040
AML	730
Neuroblastoma	710
Bone	820
Thyroid	570
Wilms	510
Germ cell	540
Rhabdomyosarcoma	340
Retinoblastoma	280
Melanoma	310
Other	2,890

Source: CDC - 2014



# Children's Oncology Group

1955

Cooperative group system for clinical research

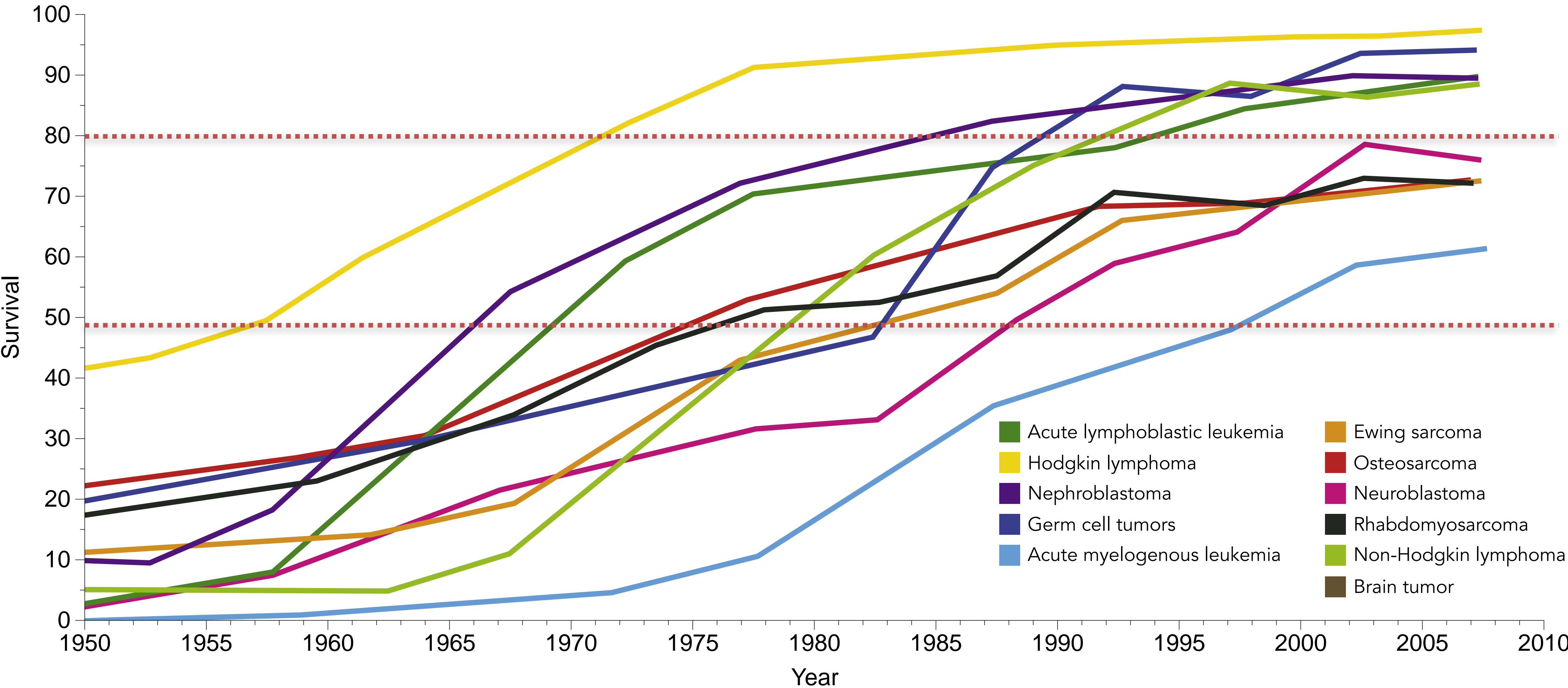
Pediatric Oncology Group (POG)  
Children's Cancer Group (CCG)  
National Wilms' Tumor Study Group (NWTSG)  
Intergroup Rhabdomyosarcoma Study Group (IRSG)

2000

Children's Oncology Group (COG)

200 centers in the United States, Canada, Switzerland,  
the Netherlands, Australia, and New Zealand

# Survival - Pediatric cancer



# The “long tail” of pediatric extracranial cancers

- PROFILE study (DFCI, BWH, BCH)
- All kids with suspected cancer
- DNA sequencing panel
- Used ICD-O for diagnoses
- Sequenced 338 patients

Courtesy: K. Janeway

doi:10.1172/jci.insight.87062



THE UNIVERSITY OF  
CHICAGO MEDICINE &  
BIOLOGICAL SCIENCES



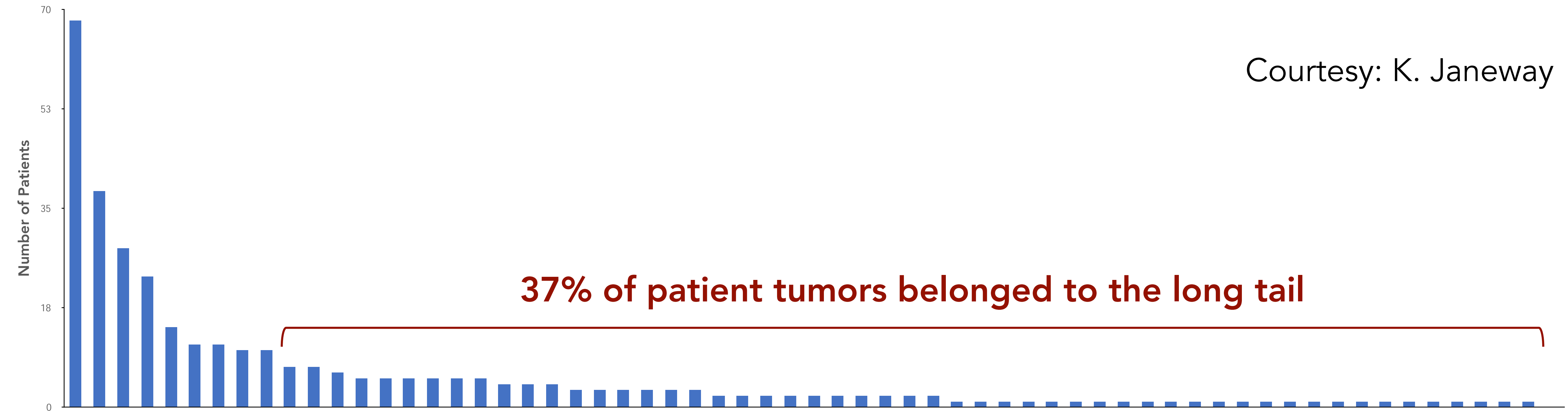
CENTER FOR  
**RESEARCH**  
INFORMATICS

@samvolchenboum

# There were many rare diagnoses

- 63% were from nine common pediatric diagnoses
- Remaining 37% represent a “long tail”

Courtesy: K. Janeway



37% of patient tumors belonged to the long tail



# Diagnoses grouped into categories

## Liver Tumors

Hepatoblastoma  
Hepatocellular carcinoma

## Germ Cell/Sex Cord Tumor

Choriocarcinoma  
Germinoma  
Granulosa cell tumor  
Malignant teratoma  
Seminoma  
Sertoli-Leydig cell tumor  
Yolk sac tumor

## Carcinoma

Acinar cell carcinoma  
Carcinoma, NOS  
Embryonal carcinoma  
Merkel cell carcinoma  
Neuroendocrine carcinoma  
Solid pseudopapillary carcinoma  
Squamous cell carcinoma in situ  
Adrenal cortical carcinoma  
Mucoepidermoid carcinoma  
Urothelial carcinoma  
Renal cell carcinoma

## Sarcoma (other)

Adenosarcoma (Mullerian)  
Alveolar soft part sarcoma  
Angiomatoid fibrous histiocytoma  
Chondrosarcoma  
Clear cell sarcoma of kidney  
Dermatofibrosarcoma  
Desmoplastic small round cell tumor  
Embryonal sarcoma  
Epithelioid sarcoma  
Infantile fibrosarcoma  
Leiomyosarcoma, NOS  
Mesenchymal chondrosarcoma  
Myxoid liposarcoma  
Round cell sarcoma  
Sarcoma, NOS  
Spindle cell sarcoma  
Synovial sarcoma

## Rhabdomyosarcoma

Embryonal  
Alveolar  
Spindle cell  
NOS

## Other

Carcinoid tumor  
Desmoid tumor  
Gastrointestinal stromal tumor  
Giant cell tumor of bone  
Glomus tumor, malignant  
Malignant melanoma  
Malignant peripheral nerve sheath tumor  
Malignant rhabdoid tumor  
Myofibroblastic tumor  
Neoplasm, NOS  
Neurofibromatosis  
Paraganglioma  
Pheochromocytoma  
Pigmented dermatofibrosarcoma protuberans  
Pleuropulmonary blastoma  
Retinoblastoma  
Thymoma



# Diagnoses grouped into categories

## Liver Tumors

Hepatoblastoma  
Hepatocellular carcinoma

## Germ Cell/Sex Cord Tumor

Choriocarcinoma  
Germinoma  
Granulosa cell tumor  
Malignant teratoma  
Seminoma  
Sertoli-Leydig cell tumor  
Yolk sac tumor

## Carcinoma

Acinar cell carcinoma  
Carcinoma, NOS  
Embryonal carcinoma  
Merkel cell carcinoma  
Neuroendocrine carcinoma  
Solid pseudopapillary carcinoma  
Squamous cell carcinoma in situ  
Adrenal cortical carcinoma  
Mucoepidermoid carcinoma  
Urothelial carcinoma  
Renal cell carcinoma

## Sarcoma (other)

Adenosarcoma (Mullerian)  
Alveolar soft part sarcoma  
Angiomatoid fibrous histiocytoma  
Chondrosarcoma  
Clear cell sarcoma of kidney  
Dermatofibrosarcoma  
Desmoplastic small round cell tumor  
Embryonal sarcoma  
Epithelioid sarcoma  
Infantile fibrosarcoma  
Leiomyosarcoma, NOS  
Mesenchymal chondrosarcoma  
Myxoid liposarcoma  
Round cell sarcoma  
Sarcoma, NOS  
Spindle cell sarcoma  
Synovial sarcoma

## Rhabdomyosarcoma

Embryonal  
Alveolar  
Spindle cell  
NOS

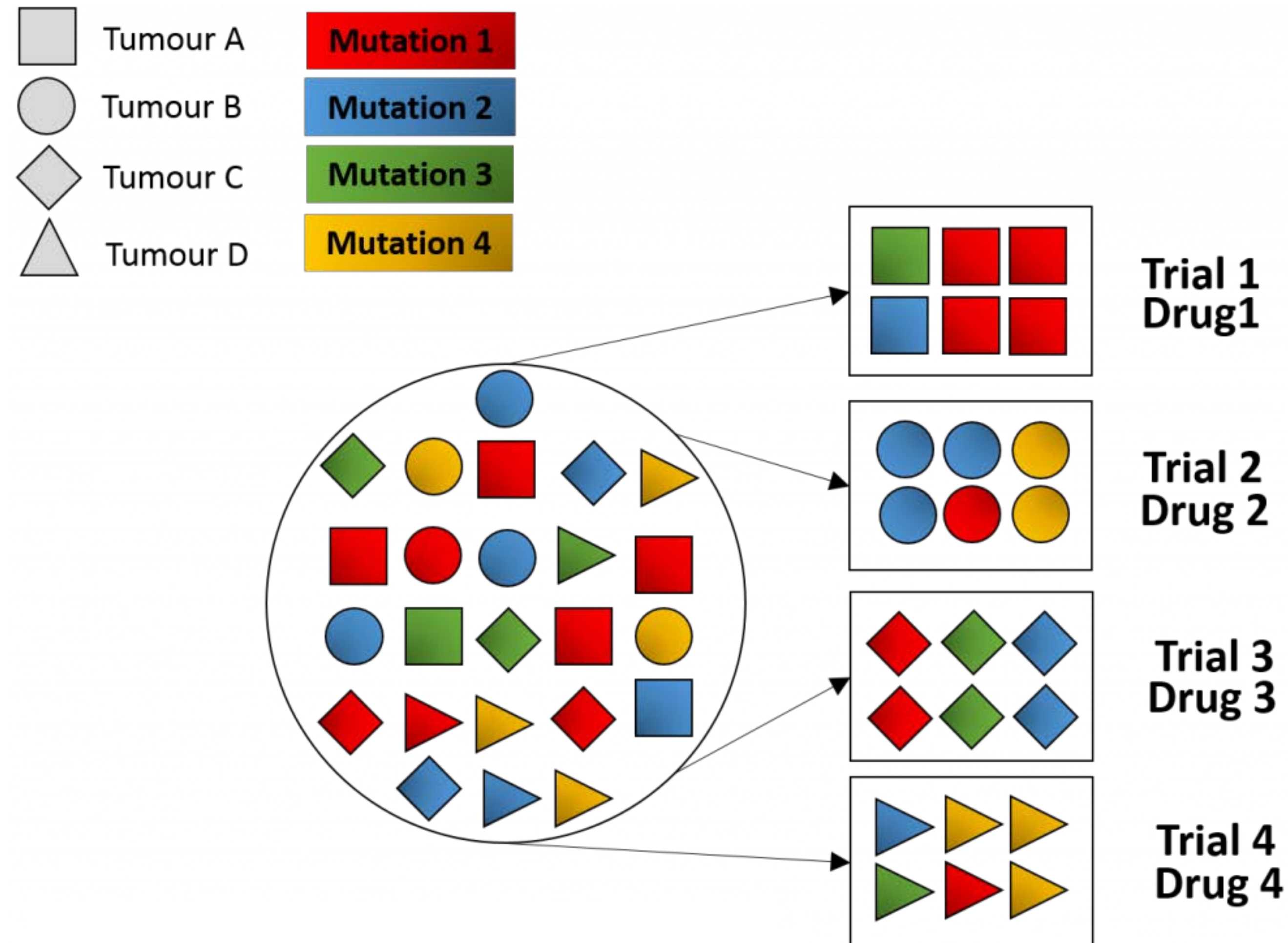
## Other

Carcinoid tumor  
Desmoid tumor  
Gastrointestinal stromal tumor  
Giant cell tumor of bone  
Glomus tumor, malignant  
Malignant melanoma  
Malignant peripheral nerve sheath tumor  
Malignant rhabdoid tumor  
Myofibroblastic tumor  
Neoplasm, NOS  
Neurofibromatosis  
Paraganglioma  
Pheochromocytoma  
Pigmented dermatofibrosarcoma protuberans  
Pleuropulmonary blastoma  
Retinoblastoma  
Thymoma

Need a new way to  
think about therapy

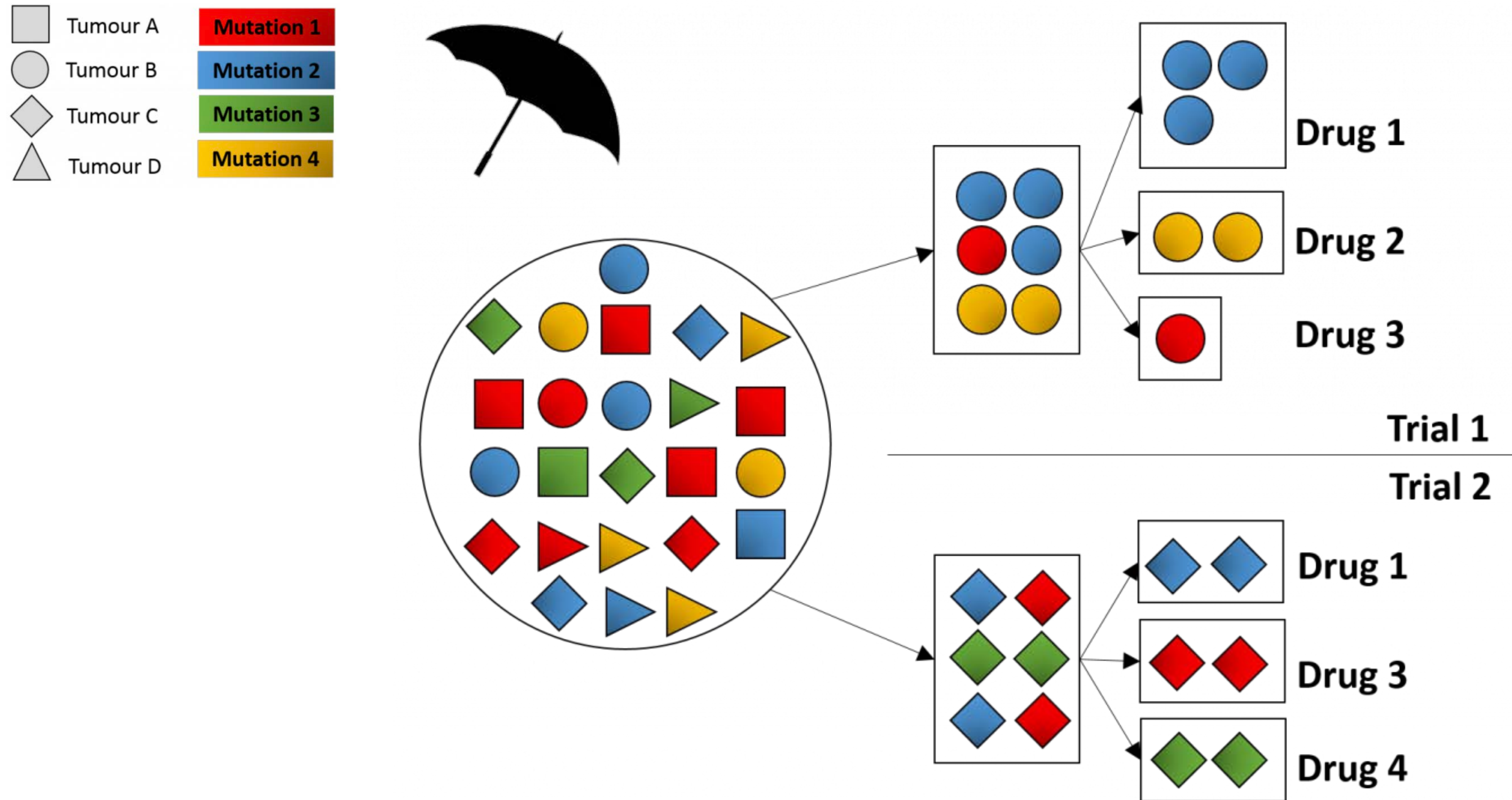


# Traditional trial based on histology



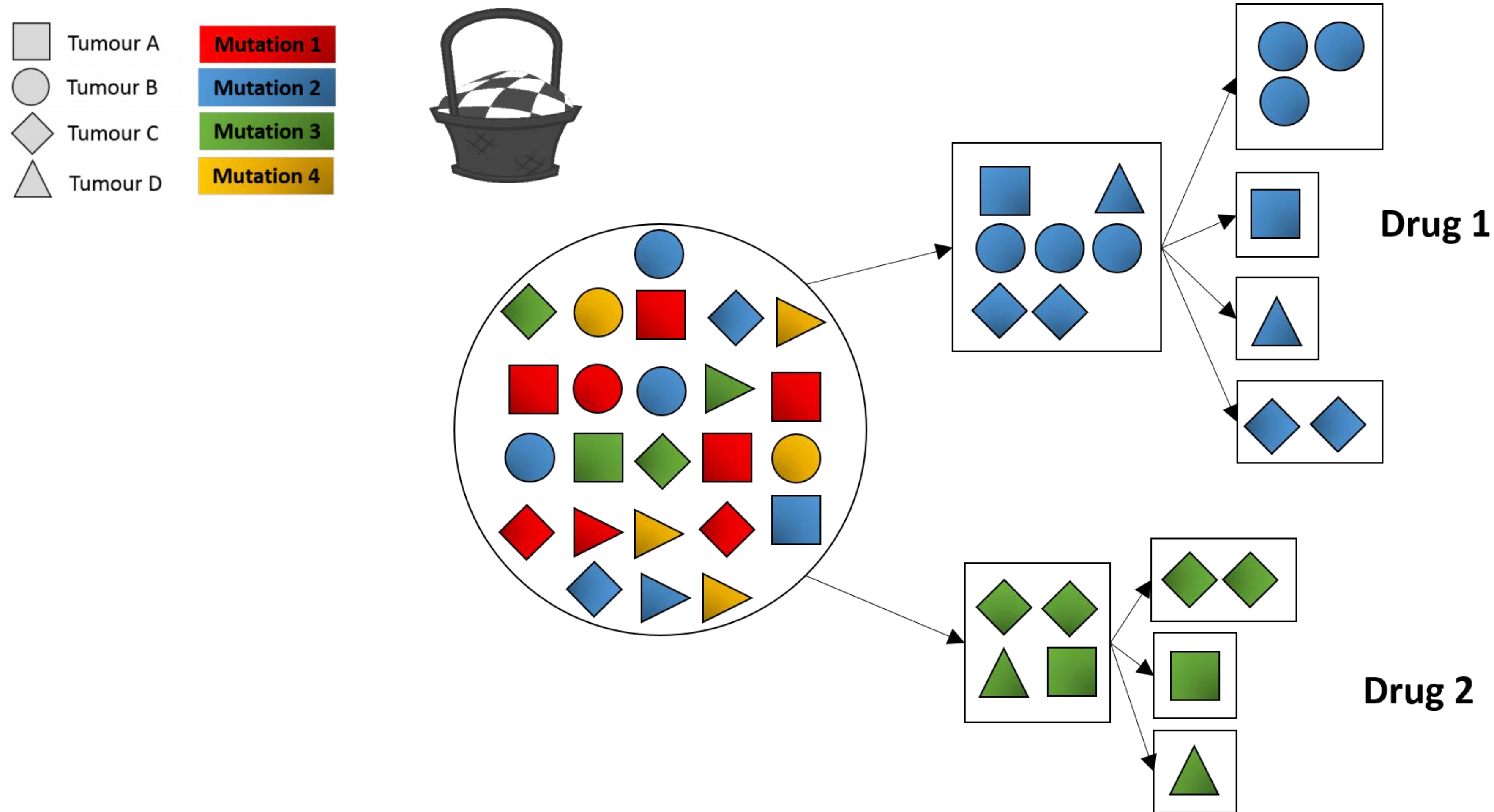
<https://www.bhdsyndrome.org/forum/bhd-research-blog/genetic-sequencing-approaches-to-cancer-clinical-trials/>

# Umbrella trial - based on tumor type and genotype



<https://www.bhdsyndrome.org/forum/bhd-research-blog/genetic-sequencing-approaches-to-cancer-clinical-trials/>

# Basket trial - based on genotype



<https://www.bhdsyndrome.org/forum/bhd-research-blog/genetic-sequencing-approaches-to-cancer-clinical-trials/>

# Precision trials in pediatric oncology

Trial type	Examples	Sponsor
Basket in Relapsed/Refractory cancers across multiple diagnoses		NCI-COG Pediatric MATCH - COG/NCI AcSé-eSMART - Gustave Roussy
Disease-specific umbrella in patients with progressive disease	Ruxolitinib or Dasatinib with Chemotherapy in Ph-Like ALL - MD Anderson NEPENTHE (Neuroblastoma) - CHOP	
Single-agent targeted therapy in advanced cancers	Larotrectinib in NTRK Fusion Positive Tumors - LOXO Oncology EZH2 Inhibitor Tazemetostat in INI-1 Negative tumors - Epizyme Crizotinib for Tumors with an ALK, MET or ROS1 alteration - UNICANCER LDK378 (Ceritinib) in ALK-activated Pediatric Tumors - Novartis Dabrafenib with Trametinib for BRAF V600 Positive Tumors - Novartis Afatinib in Pediatric Tumors with ErbB Pathway Deregulation - Boehringer Ingelheim	
Disease-specific trials in newly-diagnosed patients	Total Therapy XVII JAK/STAT Mutations in ALL and Lymphoma - St. Jude Addition of Dasatinib for ALL with TKI-targetable Fusions - DFCI Combination Therapy Plus Dasatinib for Ph-Like B-ALL - COG/NCI Clinical and Molecular Risk-Directed Therapy (Medulloblastoma) - St. Jude BIOMEDE (DIPG) - Gustave Roussy	

# Precision trials in pediatric oncology

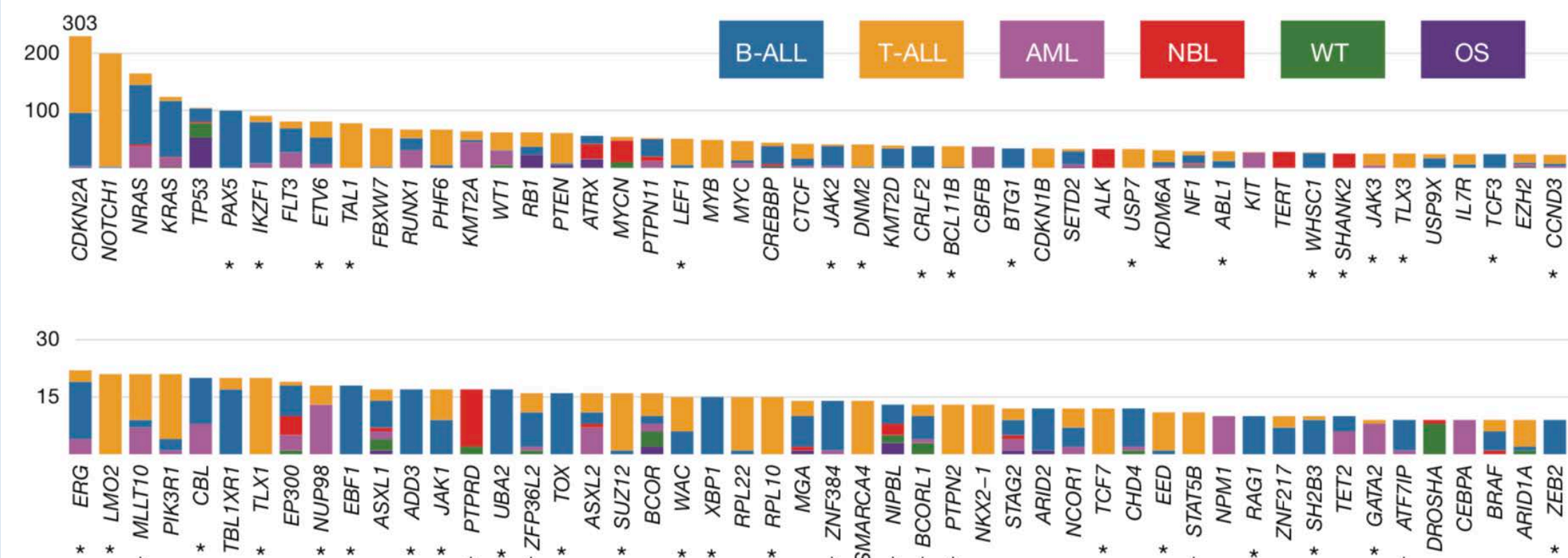
Trial type	Examples	Sponsor
Basket in Relapsed/Refractory cancers across multiple diagnoses		NCI-COG Pediatric MATCH - COG/NCI AcSé-eSMART - Gustave Roussy
Disease-specific umbrella in patients with progressive disease	Ruxolitinib or Dasatinib with Chemotherapy in Ph-Like ALL - MD Anderson NEPENTHE (Neuroblastoma) - CHOP	
Single-agent targeted therapy in advanced cancers	Larotrectinib in NTRK Fusion Positive Tumors - LOXO Oncology EZH2 Inhibitor Tazemetostat in INI-1 Negative tumors - Epizyme Crizotinib for Tumors with an ALK, MET or ROS1 alteration - UNICANCER LDK378 (Ceritinib) in ALK-activated Pediatric Tumors - Novartis Dabrafenib with Trametinib for BRAF V600 Positive Tumors - Novartis Afatinib in Pediatric Tumors with ErbB Pathway Deregulation - Boehringer Ingelheim	
Disease-specific trials in newly-diagnosed patients	Total Therapy XVII JAK/STAT Mutations in ALL and Lymphoma - St. Jude Addition of Dasatinib for ALL with TKI-targetable Fusions - DFCI Combination Therapy Plus Dasatinib for Ph-Like B-ALL - COG/NCI Clinical and Molecular Risk-Directed Therapy (Medulloblastoma) - St. Jude BIOMEDE (DIPG) - Gustave Roussy	

# Landmark work in paediatric oncology

## Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours

Xiaotu Ma<sup>1\*</sup>, Yu Liu<sup>1\*</sup>, Yanling Liu<sup>1</sup>, Ludmil B. Alexandrov<sup>2</sup>, Michael N. Edmonson<sup>1</sup>, Charles Gawad<sup>1</sup>, Xin Zhou<sup>1</sup>, Yongjin Li<sup>1</sup>, Michael C. Rusch<sup>1</sup>, John Easton<sup>1</sup>, Robert Huether<sup>3†</sup>, Veronica Gonzalez-Pena<sup>4</sup>, Mark R. Wilkinson<sup>1</sup>, Leandro C. Hermida<sup>5</sup>, Sean Davis<sup>6</sup>, Edgar Sioson<sup>1</sup>, Stanley Pounds<sup>7</sup>, Xueyuan Cao<sup>7</sup>, Rhonda E. Ries<sup>8</sup>, Zhaoming Wang<sup>1</sup>, Xiang Chen<sup>1</sup>, Li Dong<sup>1</sup>, Sharon J. Diskin<sup>9</sup>, Malcolm A. Smith<sup>10</sup>, Jaime M. Guidry Auvil<sup>5</sup>, Paul S. Meltzer<sup>6</sup>, Ching C. Lau<sup>11,12</sup>, Elizabeth J. Perlman<sup>13</sup>, John M. Maris<sup>9</sup>, Soheil Meshinchi<sup>8</sup>, Stephen P. Hunger<sup>9</sup>, Daniela S. Gerhard<sup>5</sup> & Jinghui Zhang<sup>1</sup>

doi:10.1038/nature25795

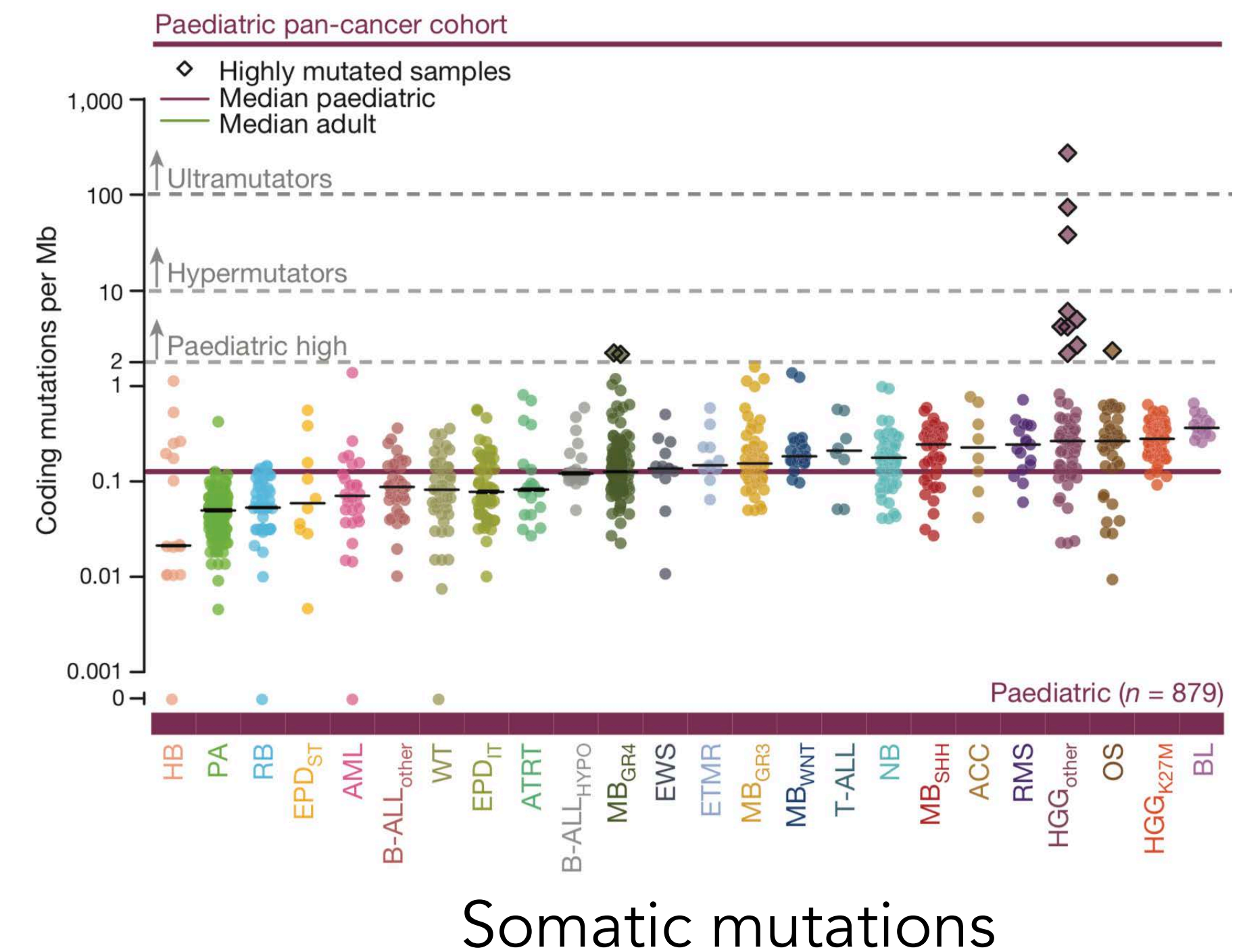


Top 100 recurrently mutated genes

## The landscape of genomic alterations across childhood cancers

A list of authors and affiliations appears at the end of the paper.

doi:10.1038/nature25480



Somatic mutations





# Ongoing data aggregation studies in pediatric oncology

Repository	Notes	Cases
Genomic data commons	TARGET Basic data standardization	3,236
St. Jude Cloud	Less clinical annotation	3,054
GENIE	Hospital-based sequencing Potential for greater clinical annotation	>1000
Foundation Medicine	Limited clinical information	1,215

Courtesy: K. Janeway



# Infrastructure innovations



# UChicago clinical genomics lab



Illumina HiSeq 2500



MiSeq Desktop



Ion Torrent



# Division of Genomic and Molecular Pathology

## Clinical Genomics and Molecular Diagnostics Laboratories

### Faculty

Jeremy Segal, MD, PhD (Director)

Lauren Ritterhouse, MD, PhD (Co-Dir, Molecular Laboratories)

Lynn Wang, MD, PhD

Megan McNerney, MD, PhD

Carrie Fitzpatrick, PhD

Nifang Niu, PhD

### Bioinformatics

Sabah Kadri, PhD (Director of Bioinformatics)

Sushant Patil, PhD

Pankhuri Wanjari, PhD

### Laboratory Staff

Sonia Benhamed (Manager)

Chaojie Zhen

Rutika Puranik, MS

Neda Joudeh

Filippo Galbo

David Montes

Bekim Ameti

Anastasiya Mendybaeva

Candace Henderson

Vidya Balagopal, PhD

### Oversight

Daniel Arber, MD (Chairman)

Candis Kinkus (VP, Laboratories)



# Division of Genomic and Molecular Pathology

## Clinical Genomics and Molecular Diagnostics Laboratories

### Faculty

Jeremy Segal, MD, PhD (Director)

Lauren Ritterhouse, MD, PhD (Co-Dir, Molecular Laboratories)

Lynn Wang, MD, PhD

Megan McNerney, MD, PhD

Carrie Fitzpatrick, PhD

Nifang Niu, PhD

### Bioinformatics

Sabah Kadri, PhD (Director of Bioinformatics)

Sushant Patil, PhD

Pankhuri Wanjari, PhD

### Laboratory Staff

Sonia Benhamed (Manager)

Chaojie Zhen

Rutika Puranik, MS

Neda Joudeh

Filippo Galbo

David Montes

Bekim Ameti

Anastasiya Mendybaeva

Candace Henderson

Vidya Balagopal, PhD

### Oversight

Daniel Arber, MD (Chairman)

Candis Kinkus (VP, Laboratories)

# Clinical Genomics Laboratory

## NGS Test Menu

- **OncoPlus v3.0**

- 147 (out of 1213) gene panel for comprehensive profiling
- Mutations, indels, structural variations, CNVs, fusions, MSI
- Specimens: FFPE tissue, blood, bone marrow, FNAs
- Turnaround time: 10-14 business days

- **OncoScreen v2.0**

- 50 gene panel for hot-spot mutation detection
- Specimens: FFPE tissue, blood, bone marrow, FNAs
- Turnaround time: 4-10 business days
- Minimal DNA input

- ***NPM1* MRD detection**

- Molecular barcode amplicon sequencing for low% *NPM1* mutation detection.
- Allelic discrimination at 1 mutant molecule per 10,000 normal molecules (0.01%)
- Specimens: blood, bone marrow
- Turnaround time: 4-10 business days

- ***APOL1* risk variant detection**

- Detection of common renal disease risk variants in the *APOL1* gene.
- Turnaround time: 4-10 business days



# Clinical Genomics Laboratory

## NGS Test Menu

- **OncoPlus v3.0**

- 147 (out of 1213) gene panel for comprehensive profiling
- Mutations, indels, structural variations, CNVs, fusions, MSI
- Specimens: FFPE tissue, blood, bone marrow, FNAs
- Turnaround time: 10-14 business days

- **OncoScreen v2.0**

- 50 gene panel for hot-spot mutation detection
- Specimens: FFPE tissue, blood, bone marrow, FNAs
- Turnaround time: 4-10 business days
- Minimal DNA input

- ***NPM1* MRD detection**

- Molecular barcode amplicon sequencing for low% *NPM1* mutation detection.
- Allelic discrimination at 1 mutant molecule per 10,000 normal molecules (0.01%)
- Specimens: blood, bone marrow
- Turnaround time: 4-10 business days

- ***APOL1* risk variant detection**

- Detection of common renal disease risk variants in the *APOL1* gene.
- Turnaround time: 4-10 business days



# Clinical Genomics Laboratory

## NGS Test Menu

- **OncoPlus v3.0**

- 147 (out of 1213) gene panel for comprehensive profiling
- Mutations, indels, structural variations, CNVs, fusions, MSI
- Specimens: FFPE tissue, blood, bone marrow, FNAs
- Turnaround time: 10-14 business days

- **OncoScreen v2.0**

- 50 gene panel for hot-spot mutation detection
- Specimens: FFPE tissue, blood, bone marrow, FNAs
- Turnaround time: 4-10 business days
- Minimal DNA input

- ***NPM1* MRD detection**

- Molecular barcode amplicon sequencing for low% *NPM1* mutation detection.
- Allelic discrimination at 1 mutant molecule per 10,000 normal molecules (0.01%)
- Specimens: blood, bone marrow
- Turnaround time: 4-10 business days

- ***APOL1* risk variant detection**

- Detection of common renal disease risk variants in the *APOL1* gene.
- Turnaround time: 4-10 business days

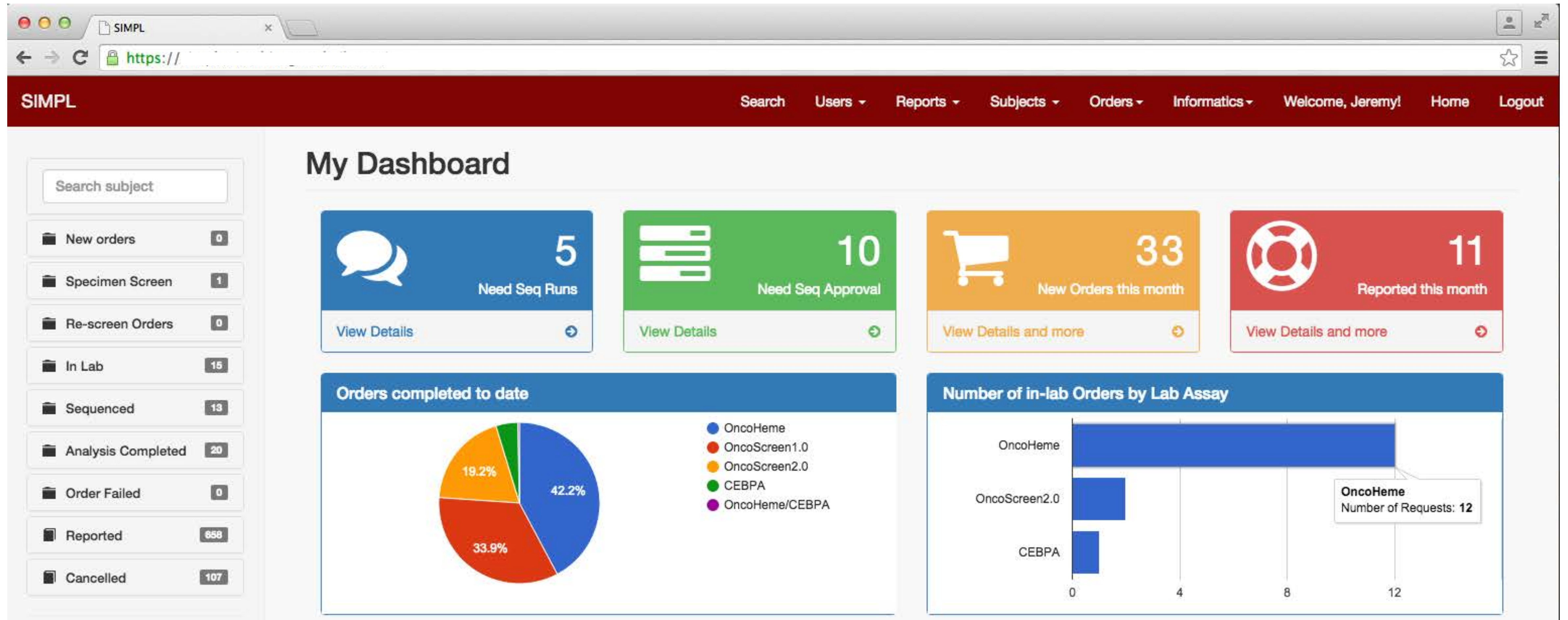








# System for informatics in the molecular pathology laboratory (SIMPL)



# SIMPL sample tracking

The screenshot shows the SIMPL Test Order Detail page. The interface includes a navigation bar with 'SIMPL' and various menu items like 'Search', 'Users', 'Reports', 'Subjects', 'Orders', 'Informatics', 'Welcome, Jeremy!', 'Home', and 'Logout'. A left sidebar contains filters for 'New orders', 'Specimen Screen', 'Re-screen Orders', 'In Lab', 'Sequenced', 'Analysis Completed', 'Order Failed', 'Reported', and 'Cancelled'. The main content area is divided into several sections:

- Subject Information:** Fields for Name, MRN, Gender (M), DOB, and IDs (CGL659).
- Order Information:** Fields for Order ID, Order Date (09/16/15), Test Requested (OncoScreen FFPE), Physician, Case ID, Copath PO, and Note.
- Order Status:** A table showing the order's history.
 

Date	User	Status	Comment
09/29/15	imujacic	Analysis Completed	None
09/28/15	czhen	Sequenced	None
09/21/15	dlopez2	In Lab	None
09/16/15	surban	Specimen Screen	None
- Specimen Process:** A table showing specimen processing details.
 

Date	Type	Adquacy	Status	ID	Detail
09/16/15	FFPE	Yes	Delivered to lab	CGL659.S1	<a href="#">View</a>
- Lab Assay:** Fields for Sample (CGL659.T1.S1 (OncoScreen2.0)), Assay Requested (OncoScreen2.0), and Results Uploaded on (09/28/15). It also includes a sub-table for Sequence Info.
 

Date	Run ID	Pipeline
09/24/15	AJ9U7	OncoST2.0 v2.2.0
- Variants Interpretation:** A table showing variant information.
 

Variant	P Level	Interpretation	DBData	Is Final	References
chr7:116411923, C>T	3	This sequence variant in exon 14 of the MET gene results in a arginine to cysteine substitution at amino acid of 988 of MET protein. MET is a receptor tyrosine kinase that transduces signals from the extracellular matrix into the cytoplasm by binding to hepatocyte growth factor(HGF) ligand, and regulates several physiological processes including proliferation, morphogenesis and survival. Many tumors show high level of MET expression, and in colon cancer, high MET expression has been associated with poor prognosis (World J Gastroenterol. 2015 Mar 28; 21(12): 3706-3710). However, missense mutations in MET have only been reported in approximately 3% of colorectal cancers whereas MET R988C mutation was also identified in the same series (BMC Cancer 2010, 10:101). The	1/2/3	true	

# SIMPL sample data annotation

**Specimen Process - CGL659.T1.S1 (OncoScreen2.0)**

Specimen type: FFPE

Collection date:

UC specimen Copath ID:

UC Block/Slide: D1

Specimen source (organ): Liver mass, segment 3

Disease category: Cancer

Disease site (original):  Rectum +

Disease:  Rectal adenocarcinoma, metastatic +

Disease notes: Metastatic Rectal cancer

Disease subtype:  +

Note:

Hospital:  UCM +

Outside case id:

Outside Block/Slide:

Recut request date: 09/17/2015

Recut received date: 09/21/2015

Percent tumor: 50.00

Background dashboard elements:

- Search subject:
- New orders: 0
- Specimen Screen: 1
- Re-screen Orders: 0
- In Lab: 15
- Sequenced: 13
- Analysis Completed: 29
- Order Failed: 0
- Reported: 668
- Cancelled: 107
- Lab tools:

Navigation: Search, Users, Reports, Subjects, Orders, Informatics, Welcome, Jeremy!, Home, Logout

IDs: CGL659

Buttons: View, Update, Add Variants

Run ID	Pipeline
AJ9U7	OncoST2.0 v2.2.0

DBData	Is Final	References
f988	1/2/3	true

# SIMPL variant report

SIMPL Search ▾ Users ▾ Reports ▾ Subjects ▾ Orders ▾ Informatics ▾ Welcome, Jeremy! Home Logout

Alamut **AnnoVar** Variants for **.T1.S1 (OncoPlus (Large Panel)) with Thrombocytosis for OncoPlus Periph Blood** Generate Report

Search:  Reset all filters

a/u	chr	Pos	Gene	Seen	Path	Max1000	MaxESP	DP30	AF	Transcript	varType	varLocation	codingEffect	cNomen	alt pNomen	exon
update	2	25464544	DNMT3A	1	3			1074	0.556	NM_022552.4	substitution	exon	missense	NM_022552.4:c.1969G>T	p.Val657Leu	17
add	2	198267373	SF3B1	1				1095	0.332	NM_012433.3	substitution	exon	missense	NM_012433.3:c.1984C>G	p.His662Asp	14
add	3	142281428	ATR	3		0.0053	0.00748979	1440	0.513	NM_001184.3	substitution	exon	synonymous	NM_001184.3:c.816A>G	p.Ser272Ser	4
update	4	55564576	KIT	1	3		0.000226963	1354	0.496	NM_000222.2	substitution	exon	missense	NM_000222.2:c.464C>T	p.Pro155Leu	3
add	7	101845115	CUX1	2		0.0038	0.00522252	439	0.528	NM_001202543.1	substitution	exon	synonymous	NM_001202543.1:c.2571G>A	p.Thr857Thr	18
update	9	5073770	JAK2	38	1		0.000232558	1131	0.126	NM_004972.3	substitution	exon	missense	NM_004972.3:c.1849G>T	p.Val617Phe	14
update	11	118362473	KMT2A	1	3			912	0.216	NM_001197104.1	substitution	exon	missense	NM_001197104.1:c.4834A>G	p.Ile1612Val	15
update	14	95562387	DICER1	1	3		0.00068089	713	0.46	NM_030621.4	substitution	exon	missense	NM_030621.4:c.4870G>C	p.Ala1624Pro	25
update	15	90630406	IDH2	1	3			1193	0.499	NM_002168.3	substitution	exon	missense	NM_002168.3:c.905G>C	p.Gly302Ala	7
update	20	31017743	ASXL1	1	3			893	0.493	NM_015338.5	substitution	exon	missense	NM_015338.5:c.605C>T	p.Pro202Leu	8
add	X	76855257	ATRX	1				1143	0.497	NM_000489.4	substitution	exon	synonymous	NM_000489.4:c.5730T>C	p.Phe1910Phe	24

Showing 1 to 11 of 11 entries (filtered from 187 total entries)

Tip: Hold SHIFT key and click the arrows to sort by multiple columns.

# SIMPL variant report

**Variant Interpretation: 194. AHMLW5BCXX- (OncoPlus (Large Panel))-chr9:5073770, G>T \***

**Nomenclature:** JAK2 c.1849G>T, p.V617F (NM\_004972.3)

**Interpretation:** This valine to phenylalanine substitution in codon 617 within exon 14 of the Janus kinase 2 (JAK2) gene is a well characterized somatic mutation resulting in constitutive activation of the JAK2 protein and upregulation of the downstream signaling pathways including JAK-STAT, PI3K, and MAPK (Nature 2005;434:1144-8). The V617F substitution is found in 95% of patients with polycythemia vera, in 50-60% of patients with essential thrombocythemia (ET), and 50% of patients with primary myelofibrosis (PMF) (N Engl J Med 2006; 355:2452-66). This V617F mutation has also been reported in approximately 50% of refractory anemia with ring sideroblasts and thrombocytosis (RARS-T) (Am J Hematol. 2016, 91(5):492-8).

**Pathogenic level:** 1

**References:**

Is final  
 Delete

**Previous Interpretations:** 191. BHM3NSBCXX-CGL961.T1.S1 (OncoPlus (Large Panel))-chr9:5073770, G>T interp. This variant has been seen 33 times with max path level of None.

**Test order:** Disease: Primary Myelofibrosis Path Level: 1 [Copy to top](#)

**Interpretation:** This valine to phenylalanine substitution in codon 617 within exon 14 of the Janus kinase 2 (JAK2) gene is a well characterized somatic mutation resulting in constitutive activation of the JAK2 protein and upregulation of the downstream signaling pathways including JAK-STAT, PI3K, and MAPK (Nature 2005;434:1144-8). The V617F substitution is found in 95% of patients with polycythemia vera, in 50-60% of patients with essential thrombocythemia (ET), and 50% of patients with primary myelofibrosis (PMF) (N Engl J Med 2006; 355:2452-66).

**References:**

a/u	chr	Pos	Gene	See
update	2	25464544	DNMT3A	1
add	2	198267373	SF3B1	1
add	3	142281428	ATR	3
update	4	55564576	KIT	1
add	7	101845115	CUX1	2
update	9	5073770	JAK2	38
update	11	118362473	KMT2A	1
update	14	95562387	DICER1	1
update	15	90630406	IDH2	1
update	20	31017743	ASXL1	1
add	X	76855257	ATRX	1



# SIMPL pathology report

The screenshot displays the SIMPL web interface. A central modal window titled "Order Report -" is open, showing the following details:

- Report text**
- Clinical Indication:** Thrombocytosis
- Specimen Information:** Specimen Type: Peripheral Blood, Collection Date: 04/26/16, Accession Number:
- Pathology Assessment:** Concurrent peripheral blood CBC showed elevated platelet (627k/uL) and reduced RBC ( $2.87 \times 10^6/uL$ ).
- Pathogenic Variants Detected:** JAK2 c.1849G>T, p.V617F (NM\_004972.3)
- Variants of Uncertain Clinical Significance:** DNMT3A c.1969G>T, p.V657L (NM\_022552.4), KIT c.464C>T, p.P155L (NM\_000222.2), KMT2A c.4834A>G, p.I1612V (NM\_001197104.1), DICER1 c.4870G>C, p.A1624P (NM\_030621.4), IDH2 c.905G>C, p.G302A (NM\_002168.3), ASXL1 c.605C>T, p.P202L (NM\_015338.5)
- Interpretation:** JAK2 c.1849G>T, p.V617F (NM\_004972.3). This valine to phenylalanine substitution in codon 617 within exon 14 of the Janus kinase 2 (JAK2) gene is a well characterized somatic mutation resulting in constitutive activation of the JAK2 protein and upregulation of the downstream signaling pathways including JAK-STAT, PI3K, and MAPK (Nature 2005;434:1144-8). The V617F substitution is found in 95% of patients with polycythemia vera, in 50-60% of patients with essential thrombocythemia (ET), and 50% of patients with primary myelofibrosis (PMF) (N Engl J Med 2006; 355:2452-66). This V617F mutation has also been reported in approximately 50% of refractory anemia with ring sideroblasts and thrombocytosis (RARS-T) (Am J Hematol. 2016, 91(5):492-8).

Below the interpretation, there are fields for "Upload to copath date" and "Is final" (checkbox). A "Version message" field is also present. At the bottom of the modal are buttons for "View Job Status", "Close", and "Save Changes".

The background interface shows a list of orders on the left, with "CGL964.T1" selected. The main content area displays a table with columns "DBData", "Is Final", and "References".

DBData	Is Final	References
1/1/3	true	
1/1/3	true	



# Clinical trial innovations

# GAIN Consortium



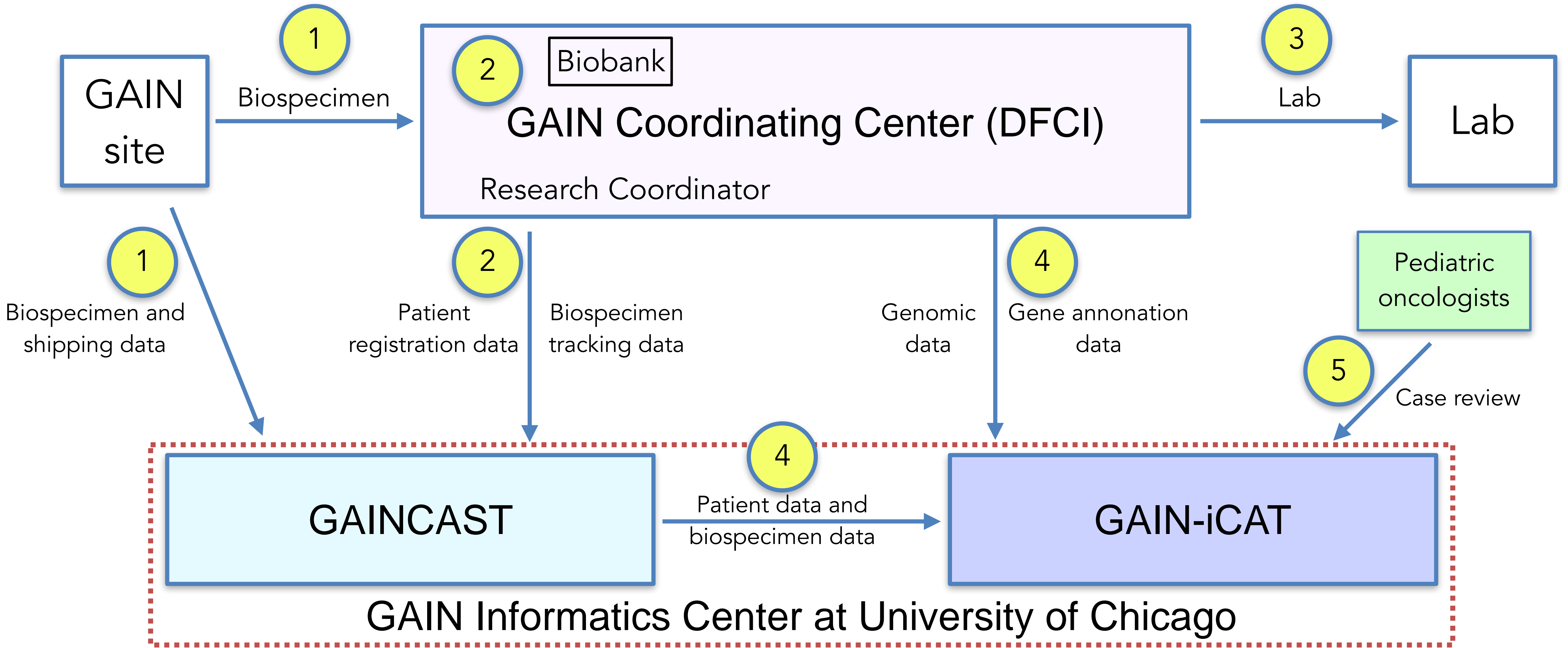
- Boston Children's Hospital
- Children's Hospital at Montefiore
- Children's Hospital of Philadelphia
- Children's National Medical Center
- Columbia University Medical Center
- Dana-Farber Cancer Institute
- Huntsman Cancer Institute, University of Utah
- Nationwide Children's Hospital
- Seattle Children's Hospital
- UCSF Benioff Children's Hospital
- University of Chicago Comer Children's Hospital
- Children's Hospital Colorado
- UT Southwestern Medical Center



Multi-institutional collaboration in precision oncology for children with difficult-to-cure cancers



# GAIN Consortium - infrastructure



# GAIN iCAT - patient tracking

The screenshot displays the GAIN iCAT web application interface. The top navigation bar includes the GAIN logo, a menu icon, and links for Users, Subjects, Curations, Admin, Home, and Logout. A left sidebar shows navigation options: Pending seq upload (1), Pending curation (1), Pending Clinical Interp (1), and Report sent (1). The main content area is titled "Existing Subjects" and features a search bar and a "Show 10 entries" dropdown. Below this is a table with columns for Name, MRN, DOB, Gender, ID, and Institution. The table contains six entries, with the first one being "Doe, John" and the last one being "Test, test". At the bottom, it indicates "Showing 1 to 6 of 6 entries" and includes "Previous", "1", and "Next" navigation buttons.

Name	MRN	DOB	Gender	ID	Institution
Doe, John		01/02/1999	M	1	DFCI
Fast, Mary	23423423	05/02/2011	F	564	BCH
John, Nancy	234324	03/02/1993	F	234	U. Chicago
Nice, Benjamin	234234	05/02/1999	M	566	U. Chicago
Test, Gain	2342342	03/02/1992	F	45	DFCI
Test, test		03/02/2012	F	24	U. Chicago

# GAIN iCAT - analysis tracking

The screenshot displays the GAIN iCAT web application interface. The top navigation bar includes the GAIN logo, a menu icon, and links for Users, Subjects, Curations, Admin, Home, and Logout. A left sidebar shows a search bar and a list of subjects with status indicators (e.g., Pending seq upload, Pending curation). The main content area is divided into several panels:

- Subject Information:** Displays Name (Mary Fast), MRN (23423423), Gender (F), DOB (05/02/2011), ID (564), and Site (BCH).
- Test Information:** Displays Test ID (5), Test Date (03/17/16), Test (OncoPanel Germline), Specimen Type (FFPE), Assession ID (S213234), Diagnosis (Cancer B), and Note.
- Interpretation status:** A table showing the history of interpretations.
- Step 1: Test Results Upload/Input:** Displays Report Date (03/17/16), File (/bootstrap\_tutorial\_dx6RLDB.pdf), Results (View Variants), and DLM (03/17/16 by admin).
- Step 3: Assign Curation and Clinic Review Teams:** Displays Curator (Ted Counselor1), Molecular pathologist (Mary MP1), Pedi Onc Fellow (John Pofellow1), Pedi oncologist (PD Oncologisit1), and Start Date (03/17/2016).
- Step 4: Curation:** A section for Gene/Variant Curation with a Start button.

Date	User	Status	Comment
03/17/16	admin	Pending curation	None
03/17/16	admin	Seq uploaded	None
03/17/16	admin	Pending seq upload	None

# GAIN iCAT - Facilitated gene curation

The screenshot displays the GAIN iCAT web interface for gene curation, specifically for the BRAF gene. The interface is organized into several sections:

- Gene curation:** Shows subject and test information, including ID (CA10001Diagnosis), Neuroendocrine Tumor, Report Date (05/27/17), Gender (M), Age (14), Test (OncoPanel), and BRCA (BRCA, Farber, Cancer Institute).
- Gene Name:** Lists BRAF with HGNC, UniProt, NCBI Gene, GeneReviews, PubMed, and Google links. It also shows the Official Gene Name (BRAF), Common use name (BRAF), and Entered date (07/28/17 by Lori.gstrand).
- Gene Function:** Provides a function summary: "BRAF is a well-characterized oncogene that encodes a serine/threonine kinase in the RAS/RAF/MEK/ERK signaling pathway. Alterations leading to constitutive signaling in BRAF leads to evasion of apoptosis, greater cell proliferation, angiogenesis, increased mitotic potential, and lowered susceptibility to immune responses (PMID: 21577205)." It includes a table with references and entered dates.
- Germline Gene Associations:** A section for germline gene associations, currently showing "No previous knowledge".
- Gene and Cancer:** A section for gene and cancer associations, currently showing "No previous knowledge".
- Gene & Diagnosis:** A section for gene and diagnosis associations, currently showing "No previous knowledge".
- Gene & Therapy:** A section for gene and therapy associations, showing potential drugs and clinical trials. It includes a table with columns for Clinical trials, References, Entered, and Action.
- Gene Curation Summary:** A section for a curation summary, currently showing "No previous knowledge".



# GAIN iCAT - variant curation summary

## variant curation summary

CCNE1 amplification (~6 copies) was detected in this tumor. CCNE1 amplification in ovarian, uterine, and endometrial cancers has been associated with poor prognoses (PMID: 20336784, 26647729). Preclinical studies have shown that CCNE1 amplification sensitized ovarian or uterine carcinoma cells to CDK2 knockdown or CDK2 inhibitors (PMID: 24004674, 27351214). Numerous companies are developing cyclin dependent kinase inhibitors (PMID: 28127048). It should be noted however that these are not selective CDK2 inhibitors and efficacy specific for CCNE1-amplified tumors has not been conclusively demonstrated. Below are several CDK inhibitor clinical trials that are actively recruiting patients at least 18 years of age for which your patient may be eligible.

(1) NCT01434316 is a phase I clinical trial of dinaciclib (CDK2/5/1/9 inhibitor) in combination with veliparib (PARP inhibitor) in advanced solid tumors.

(2) NCT00999401 is a phase I clinical trial of seliciclib (CDK2/7/1 & ERK2 inhibitor) in combination with sapacitabine (a DNA damaging agent) in advanced solid tumors.

(3) NCT02503709 is a phase I clinical trial of AT7519 (CDK9/5/2/4/6/1 & GSK3B inhibitor) in combination with onalespib (HSP90 inhibitor) in metastatic solid tumors.

Number of iCat recommendations (for this variant): 2

- iCat recommendation therapeutic class: CDK2 Inhibitors
  - iCat recommendation tier: 4B
    - iCat recommendation specific drugs: dinaciclib
      - Treatment availability: Clinical trial active and currently open for enrollment
    - iCat recommendation specific drugs: seliciclib
      - Treatment availability: Clinical trial active and currently open for enrollment
    - iCat recommendation specific drugs: AT7519
      - Treatment availability: Clinical trial active and currently open for enrollment
- iCat recommendation therapeutic class: DNA Damage Response Inhibitor
  - iCat recommendation tier: 5A
    - iCat recommendation specific drugs: AZD1775
      - Treatment availability: Clinical trial active and currently open for enrollment
    - iCat recommendation specific drugs: Prexasertib
      - Treatment availability: Clinical trial active and currently open for enrollment
    - iCat recommendation specific drugs: VX-970
      - Treatment availability: Clinical trial active and currently open for enrollment

## Variant curation summary



# GAIN iCAT - Final curation report

A TP53 p.R290H variant was identified in this tumor with an allelic frequency of 51%. This represents a modest change from arginine to histidine (Grantham distance of 29) at a moderately conserved residue at the edge of the DNA binding domain. This missense variant, p.R290H, is frequently reported in multiple cancer databases (COSMIC, cBioPortal, TCGA). However, it has also been reported in population databases at a frequency of up to 0.046%. It has been submitted to ClinVar multiple times with conflicting interpretations. The p.R290H variant was identified in 6 families with inherited cancer predisposition, 2 with classic clinical criteria for LFS (PMID: 28861920). However, the variant scored as having wild type activity in a transactivation assay (PMID: 12826609). Coupled with the low penetrance with which carriers develop cancer, this variant likely has reduced pathogenicity. This test does not distinguish between germline and somatic alterations. If clinically indicated, genetic counseling and germline testing may be helpful.

Number of iCat recommendations (for this variant): 1

iCat recommendation therapeutic class: WEE1 Inhibitors

iCat recommendation tier: 2A

iCat recommendation specific drugs: AZD1775

Treatment availability: Drug in clinical development but no appropriate trial for this patient

## Curation Report:

### **POLB p.I260M 56.7% in 171 reads**

POLB is a DNA polymerase involved in base excision and repair, also called gap-filling DNA synthesis. POLB variants have been identified in a wide range of cancers, including colon, prostate, gastric, cervical, breast, and esophageal cancer.

A POLB p.I260M variant was identified in the tumor with an allelic frequency of 56%. This represents a seemingly modest change from an isoleucine to a methionine (Grantham distance of 10) at a highly conserved residue in the nucleotidyltransferase (NT) domain. It has been identified in population databases at frequencies as high as 0.001%. p.I260 is a key residue of the hydrophobic hinge important for polymerase closing, which is critical for polymerase fidelity (PMID: 22914675). The p.I260M variant alters protein function and is a sequence-context-dependent mutator in preclinical models (PMID: 16313169, 28862868, 22914675). This variant is included in the Curation Section as it likely contributes to an increased mutation rate in this cancer.

### **RAD51 p.G305Efs\*68 57.5% in 153 reads**

RAD51 functions in homologous recombination (HR)-mediated repair of double-stranded DNA breaks. RAD51 germline mutations may increase breast cancer risk in certain populations (PMID: 10807537, 26108708). Heterozygous truncating mutations have also been associated with mirror movements-2 (MRMV2) while dominant-negative missense mutations have been linked with an atypical form of Fanconi anemia. RAD51 is infrequently mutated in human cancer. However, RAD51 overexpression has been linked to increased oncogenic potential in several tumor





# Innovations in data collection

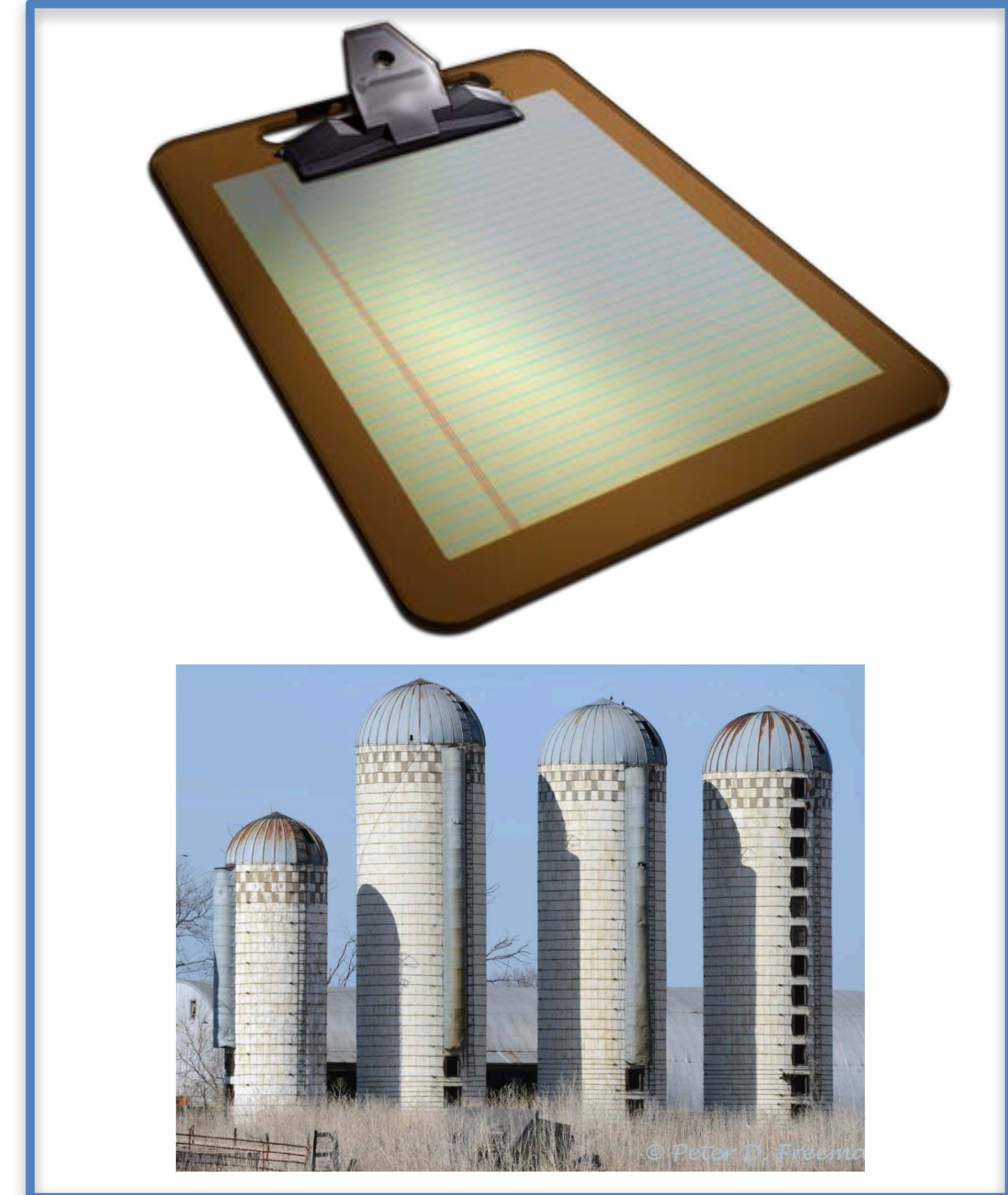
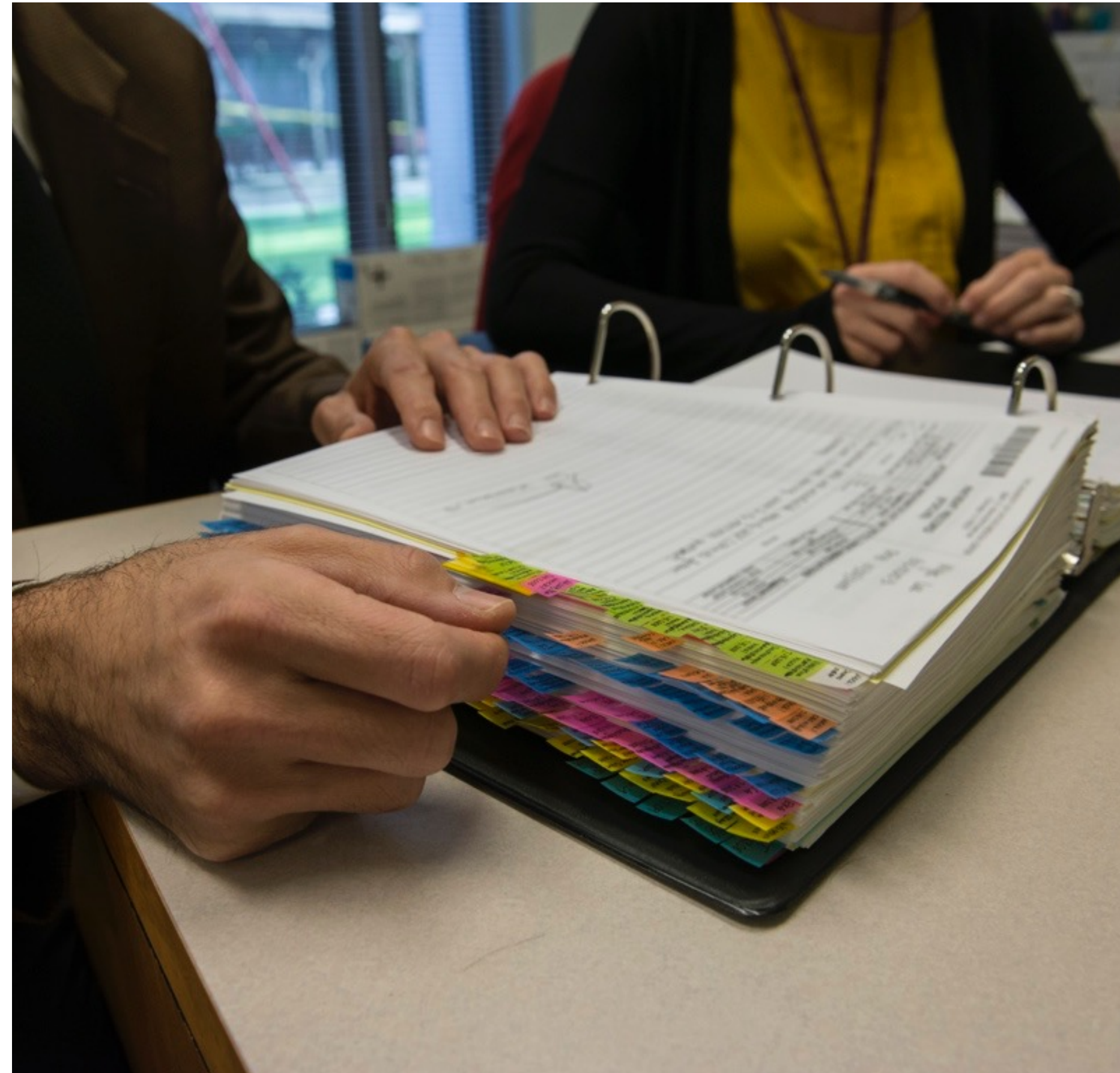


# Lack of rich phenotype data hinders progress

- Many samples for genomics lack sufficient clinical information
- Without deep phenotype data, analyses are limited
- Deep phenotype data should be collected at the time of sample acquisition, directly from electronic health record systems



# Manual processes and lack of data standards for clinical trials



# Manual processes and lack of data standards for clinical trials

CTCAE CODE (per protocol)	CTCAE SHORT NAME (per protocol)	Current Grade	Maximum Grade This Course* (for occurrence)	Maximum Grade This Occurrence**	Attribution	Date Onset	Resolved Yes / No	Date Resolved	Comments
10020943	Hypoalbuminemia	0	1	1	3	05/23	Yes	06/02	
10020943	Hypoalbuminemia	0	1	1	3	06/05	Yes	06/09	
10021038	Hyponatremia	0	1	1	3	05/26	Yes	05/27	
10021038	Hyponatremia	0	1	1	3	05/29			
10021059	Hypophosphatemia	0	1	1	3	05/10			
10021059	Hypophosphatemia	0	1	1	3	05/23			
10016256	Fatigue	0	1	1	3	05/23			

Activation Date: 8/8/01      Version Date: 9-4-10

NANT 99-02

TITLE: MODULATION OF INTENSIVE MELPHALAN (L-PAM) BY BUTHIONINE SULFOXIMINE (BSO) WITH AUTOLOGOUS STEM CELL SUPPORT FOR RESISTANT/RECURRENT HIGH-RISK NEUROBLASTOMA (IND # 68,112)

Phase I Study

COORDINATING CENTER  
Children's Hospital Los Angeles / NANT Operations Center

PARTICIPATING INSTITUTIONS AND CONTACT PHYSICIANS

Children's Hospital Los Angeles, 4650 Sunset Blvd., MS-04, Los Angeles, CA 90027.  
Judy Villablanca, M.D., Phone: 323-361-5654; Email: jvillablanca@chla.usc.edu

USCF School of Medicine, 505 Parnassus Avenue, M547, San Francisco, CA 94143.  
Katherine K. Mathay, M.D., Phone: 415-476-3831; Email: kmathay@ucsf.edu

Ludie Salkin, Pediatric Children's Hospital, 300 Pasteur Drive, Newark, CA 94505.  
Clare Teitel, M.D., Phone: 650-723-5535; Email: clare.teitel@stanford.edu

Cincinnati Children's Hospital, 3333 Burnet Avenue, MLC 7015, Cincinnati, OH 45229.  
Brian Weiss, M.D., Phone: 513-636-8663; Email: brian.weiss@ccmc.com

Children's Hospital of Philadelphia, 3615 Civic Center Boulevard, ARC 907C, Philadelphia, PA 19104.  
Yael Meese, M.D., Phone: (215) 980-0965; Email: ymeese@chop.edu

C.S. Mott Children's Hospital, 1500 East Medical Center Drive, B1-208, Ann Arbor, MI 48109.  
Gregory Yank, M.D., Phone: 734-558-8795; Email: gyan@umich.edu

Texas Children's Hospital, Baylor College of Medicine, 6701 Fannin, 14<sup>th</sup> Floor, MC CC1410, Houston, TX 77030.  
Susan Blaney, M.D., Phone: (832) 822-4588; Email: sblaney@bcm.edu

Children's Hospital and Regional Medical Center, 4800 Sand Point Way NE, MS 401, Seattle, WA 98105.  
Julie Park, M.D., Phone: (206) 987-1447; Email: julie.park@seattlechildrens.org

Children's Hospital Boston, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.  
Suzanne Shusterman, M.D., Phone: (617) 632-3725; Email: Suzanne.Shusterman@dfci.harvard.edu

University of Chicago, Comer Children's Hospital, 5841 South Maryland Avenue, Chicago, IL 60637.  
Susan Cohen, M.D., Phone: (773) 703-9271; Email: scohen@bsd.uchicago.edu

Children's Healthcare of Atlanta, 2015 Uppergate Drive, Atlanta, GA 30322.  
MD, Phone: (404) 795-0863; Email: Howard.halperin@choc.org

Stoll for Sick Children, 685 University Avenue, Toronto, Ontario, Canada M5C 1X8.  
B. Baruchel, M.D., Phone: (416) 813-7792; Email: bbaruchel@chcscs.on.ca

Ben's Hospital of New York-Presbyterian, Herbert Irving Division of Child & Adolescent Onc, 161 Ave. P, New York, NY 10032, Julia Grade-Bender, MD, Phone: (212) 305-3379; Email: jgrade@nyp.org

St. Granger, M.D., Phone: (862) 885-2560; Email: mgranger@goodchildrens.org

Protocol Chairperson  
Justin G. Villablanca, M.D.  
Children's Hospital Los Angeles  
4650 Sunset Blvd., MS-04, Los Angeles, CA 90027  
Phone (323) 361-5654 or (323) 361-5687  
Email: jvillablanca@chla.usc.edu

1.1.7. Laboratory Studies  
Patients must have a performance status of 0 or 1 (Appendix I) and a life expectancy of  $\geq 2$  months.

1.1.8. Organ Function  
1.1.8.1 Kidney Function  
Patients must have adequate renal function defined as a glomerular filtration rate (GFR) of 12 hour urine collection for creatinine clearance  $\geq 100$  mL/min/1.73m<sup>2</sup>. ASD a serum creatinine  $< 1.5 \times$  normal for age.

1.1.8.2 Serum Creatinine  
Age Range      Serum Creatinine  
 $\leq 5$  years       $< 0.8$  mg/dL

N9902 And #7 9-4-2010

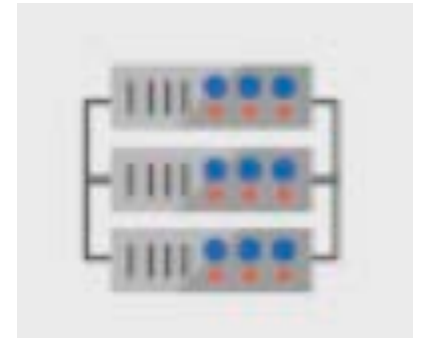
17

October		William				
Sun	Mon	Tue	Wed	Thu	Fri	Sat
	<b>1</b> Bactrim Cytoxin 20mg	<b>2</b> Bactrim Cytoxin 20mg	<b>3</b> cycle 2 DTR 9:30am Avastin IV Cytoxin 20mg	<b>4</b> DTR 9:30am Cytoxin IV Zometa IV	<b>5</b> Cytoxin 20mg	<b>6</b> Cytoxin 20mg
<b>7</b> Cytoxin 20mg	<b>8</b> Bactrim Cytoxin 20mg Labs and blood pressure at home	<b>9</b> Bactrim Cytoxin 20mg	<b>10</b> Cytoxin 20mg	<b>11</b> Cytoxin 20mg	<b>12</b> Cytoxin 20mg	<b>13</b> Cytoxin 20mg
<b>14</b> Cytoxin 20mg	<b>15</b> Bactrim Cytoxin 20mg	<b>16</b> Bactrim Cytoxin 20mg	<b>17</b> DTR 9:30am Avastin IV Cytoxin 20mg	<b>18</b> Cytoxin 20mg	<b>19</b> Cytoxin 20mg	<b>20</b> Cytoxin 20mg
<b>21</b> Cytoxin 20mg	<b>22</b> Bactrim Cytoxin 20mg SSKI	<b>23</b> Bactrim Cytoxin 20mg MIBG, x-ray knees Bone marrow biopsies SSKI	<b>24</b> Cytoxin 20mg MIBG CT scan SSKI	<b>25</b> Cytoxin 20mg SSKI	<b>26</b> Cytoxin 20mg SSKI	<b>27</b> Cytoxin 20mg
<b>28</b> Cytoxin 20mg	<b>29</b> Bactrim Cytoxin 20mg	<b>30</b> Bactrim Cytoxin 20mg	<b>31</b> Cycle 3 DTR 9:30am Avastin IV Cytoxin 20mg	<b>1</b> DTR 9:30am Cytoxin IV Zometa IV		

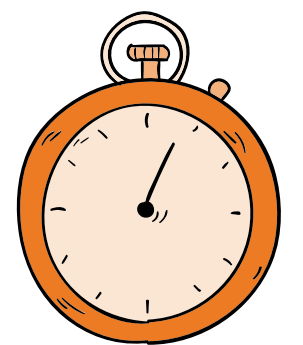
# Why do we need data commons?



Too much data to store



Takes too long to transfer



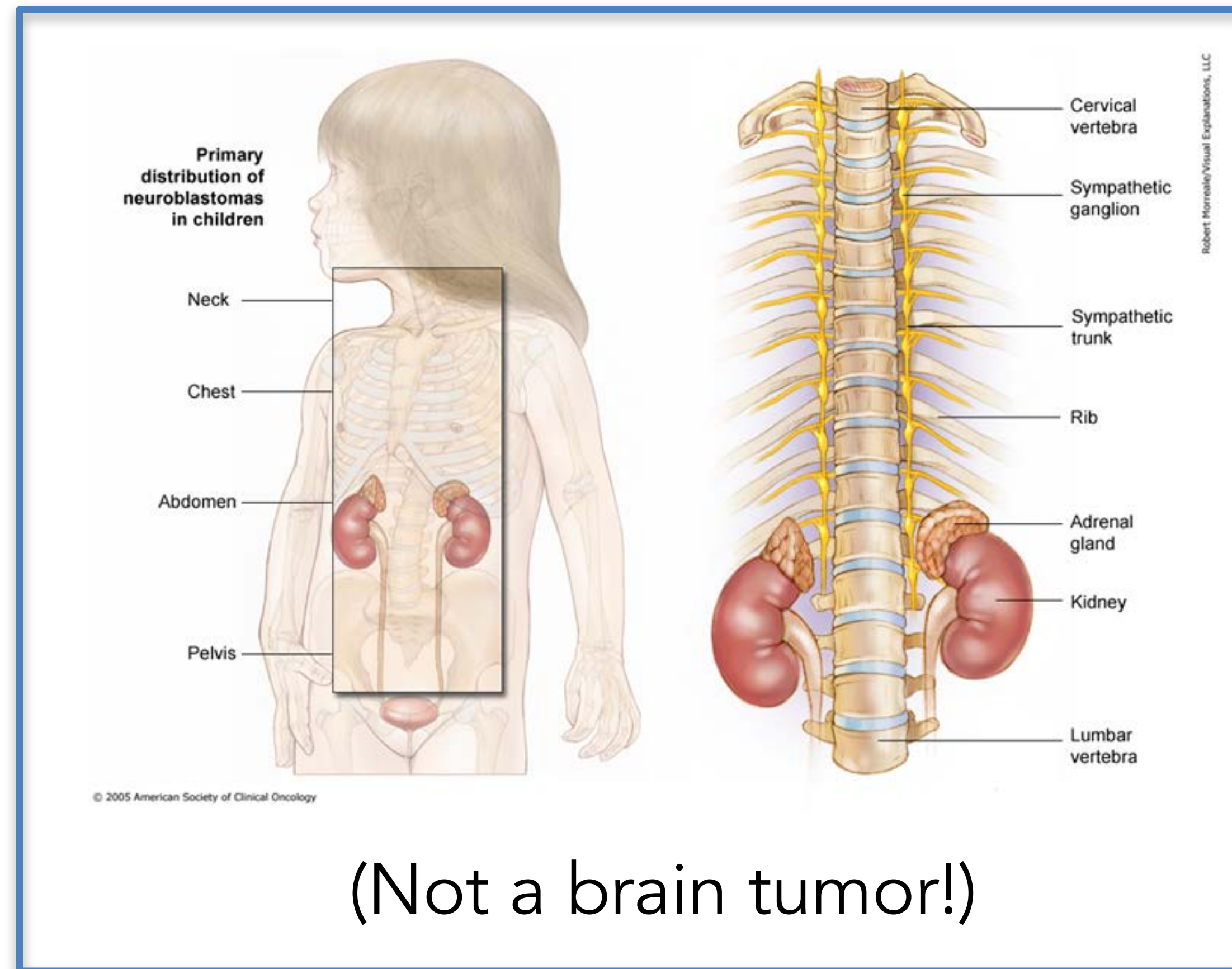
Too expensive to analyze



Lack of data standardization



# Neuroblastoma



# International Neuroblastoma Research Group (2004)



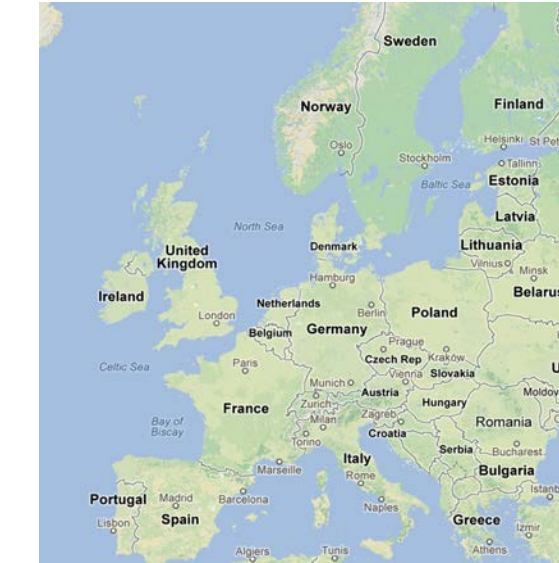
Children's Oncology Group (COG)



Germany



Japan



SIOPEN



# International Neuroblastoma Research Group (2004)



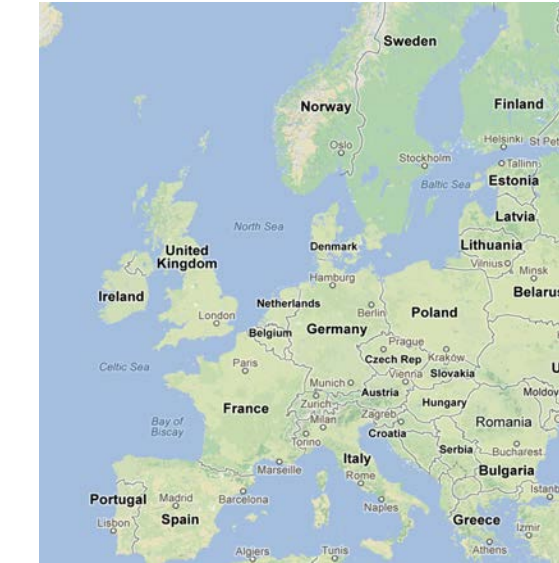
Children's Oncology Group (COG)



Germany



Japan



SIOPEN

COG	SIOPEN	Japan	Germany
Study 1	Study 4	Study 7	Study 10
Study 2	Study 5	Study 8	Study 11
Study 3	Study 6	Study 9	Study 12
COG	SIOPEN	Japan	Germany
Consensus	Consensus	Consensus	Consensus



# International Neuroblastoma Research Group (2004)



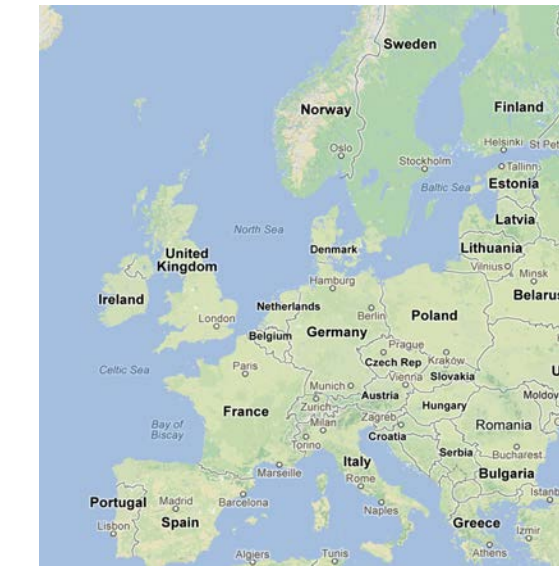
Children's Oncology Group (COG)



Germany



Japan



SIOPEN

COG	SIOPEN	Japan	Germany
Study 1	Study 4	Study 7	Study 10
Study 2	Study 5	Study 8	Study 11
Study 3	Study 6	Study 9	Study 12
COG	SIOPEN	Japan	Germany
Consensus	Consensus	Consensus	Consensus



Consensus standard

# International Neuroblastoma Research Group (2004)



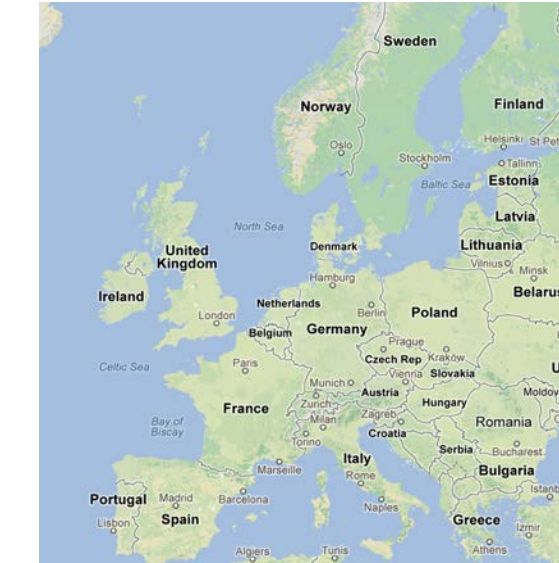
Children's Oncology Group (COG)



Germany



Japan



SIOPEN



# International Neuroblastoma Research Group (2004)



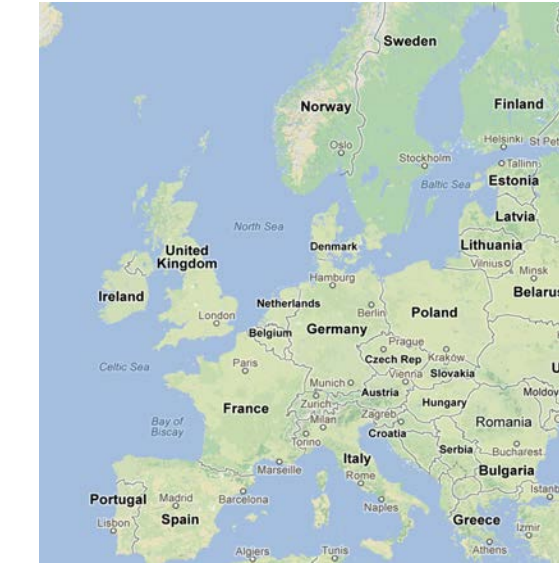
Children's Oncology Group (COG)



Germany



Japan



SIOPEN

Group	Number
COG	4235
Germany	1938
Japan	470
SIOPEN	936
Total	8800

>15 high-impact publications that changed clinical practice



# International Neuroblastoma Research Group (2004)



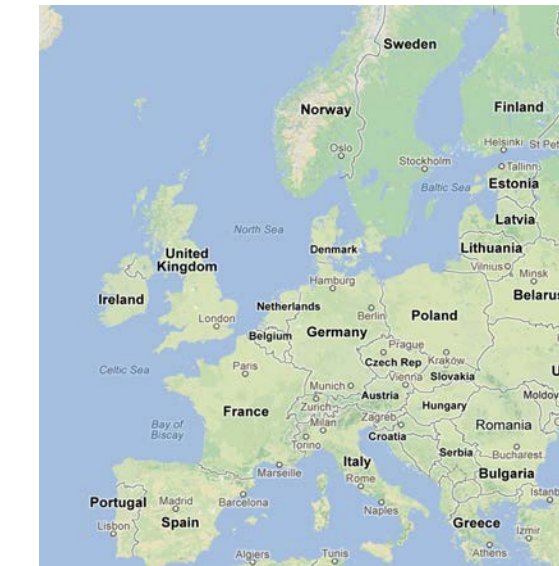
Children's Oncology Group (COG)



Germany



Japan



SIOPEN

Group	Number
COG	4235
Germany	1938
Japan	470
SIOPEN	936
Total	8800

The good news: 8800 patients  
The bad news:

>15 high-impact publications  
that changed clinical practice





# Neuroblastoma Commons Cohort Discovery

Favorable tumor biology, tissue available

## Cohort Search

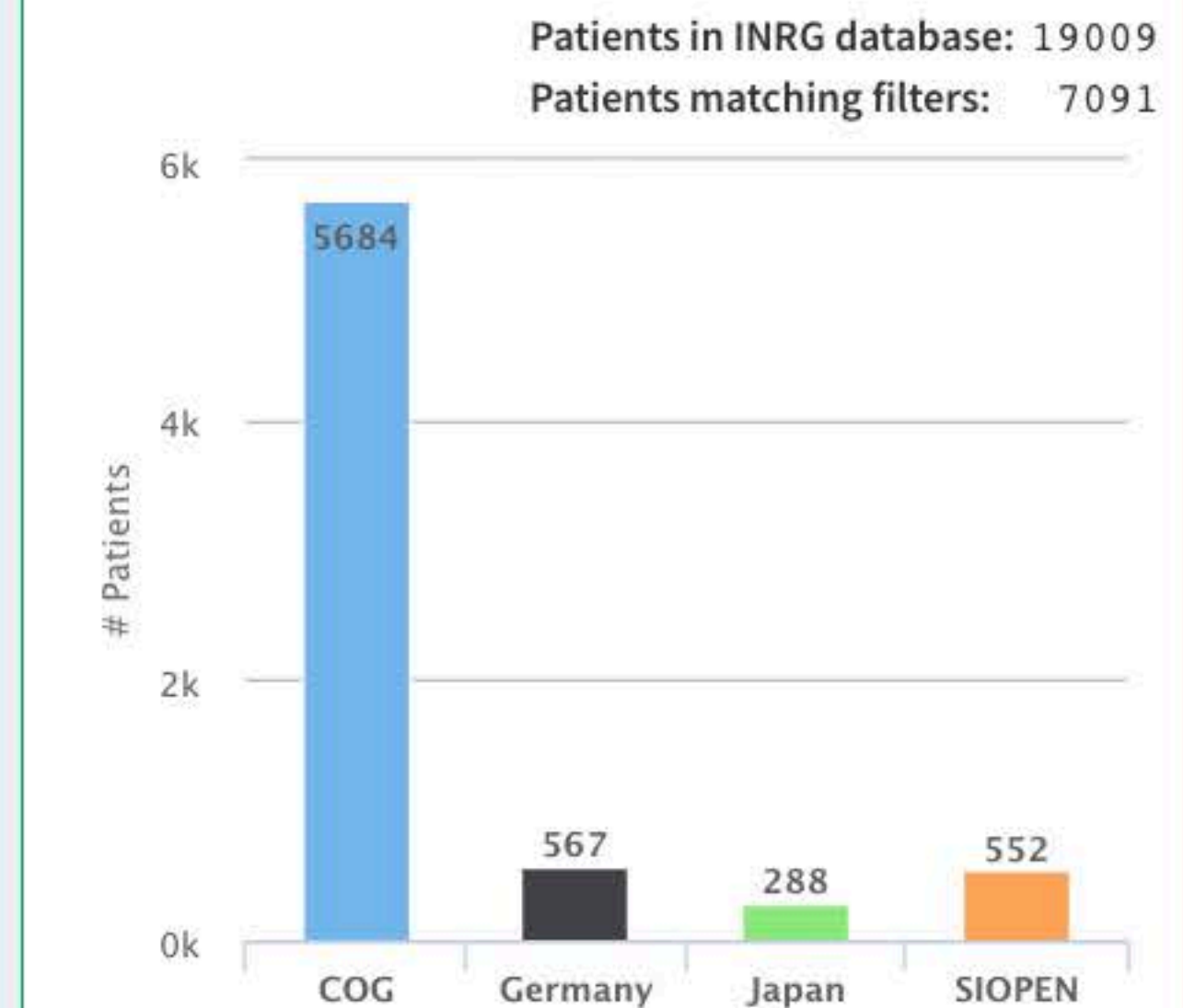
**Search: New** Save Search

Add a filter: - Please select - Clear Filters

Revised INPC Prognostic Group/ Shimada Diagnostic Category ×

- Favorable
- Unfavorable
- Unknown/Slides not received/Inadequate

## Results



# Neuroblastoma Commons Cohort Discovery

Cohort Search

Search: New Save Search

Add a filter: - Please select - Clear Filters

- Please select -
- Primary Tumor-Adrenail
- Primary Tumor-Neck
- Primary Tumor-Other
- Primary Tumor-Pelvis
- Primary Tumor-Thorax
- Race
- Revised INPC Prognostic Group/ Shimada Diagnostic Category
- Site of Relapse
- Time from Dx to Death or Last Contact
- Time from Dx to Event or Last Contact
- Year of Diagnosis
- GEO Data - **Note! External Data**
- GWAS Data - **Note! External Data**
- Nationwide Tissue Bank - **Note! External Data**
- Nucleic Acid Data - **Note! External Data**
- TARGET Data - **Note! External Data**

Revised INP

Favorable  
 Unfavorable  
 Unknown

Results

Patients in INRG database: 19009  
Patients matching filters: 7091

Group	# Patients
COG	5684
Germany	567
Japan	288
SIOPEN	552



# Neuroblastoma Commons Cohort Discovery

Cohort Search

Search: New Save Search

Add a filter: - Please select - Clear Filters

- Please select -
- Primary Tumor-Adrenail
- Primary Tumor-Neck
- Primary Tumor-Other
- Primary Tumor-Pelvis
- Primary Tumor-Thorax
- Race
- Revised INPC Prognostic Group/ Shimada Diagnostic Category
- Site of Relapse
- Time from Dx to Death or Last Contact
- Time from Dx to Event or Last Contact
- Year of Diagnosis
- GEO Data - **Note! External Data**
- GWAS Data - **Note! External Data**
- Nationwide Tissue Bank - **Note! External Data**
- Nucleic Acid Data - **Note! External Data**
- TARGET Data - **Note! External Data**

Results

Patients in INRG database: 19009  
Patients matching filters: 7091

Group	# Patients
COG	5684
Germany	567
Japan	288
SIOPEN	552

Immediately see cohort counts





# Neuroblastoma Commons Cohort Discovery

Cohort Search

Search: New Save Search

Add a filter: - Please select - Clear Filters

- Please select -
- Primary Tumor-Adrenail
- Primary Tumor-Neck
- Primary Tumor-Other
- Primary Tumor-Pelvis
- Primary Tumor-Thorax
- Race
- Revised INPC Prognostic Group/ Shimada Diagnostic Category
- Site of Relapse
- Time from Dx to Death or Last Contact
- Time from Dx to Event or Last Contact
- Year of Diagnosis
- GEO Data - **Note! External Data**
- GWAS Data - **Note! External Data**
- Nationwide Tissue Bank - **Note! External Data**
- Nucleic Acid Data - **Note! External Data**
- TARGET Data - **Note! External Data**

Results

Patients in INRG database: 19009  
Patients matching filters: 7091

Group	# Patients
COG	5684
Germany	567
Japan	288
SIOPEN	552

Immediately see cohort counts

Links to external data sets

# Neuroblastoma Commons Cohort Discovery

## Cohort Search

**Search: New** Save Search

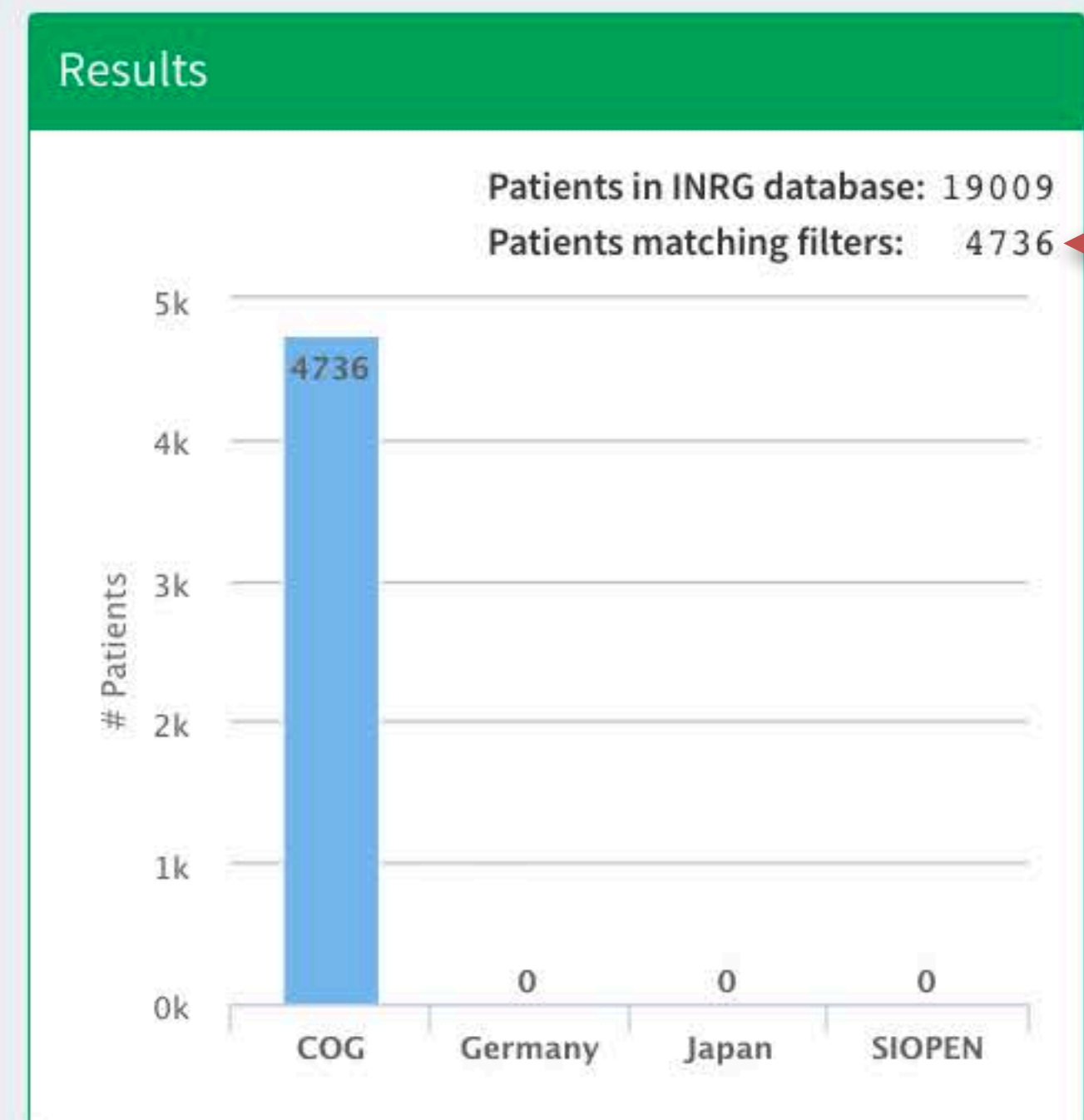
Add a filter: Nationwide Tissue Bank Clear Filters

Revised INPC Prognostic Group/ Shimada Diagnostic Category ×

- Favorable
- Unfavorable
- Unknown/Slides not received/Inadequate

Nationwide Tissue Bank ×

- No data exists
- Data exists



Favorable biology, tissue available

# Neuroblastoma data commons growth

Year	COG	SIOPEN	GPOH	Japan	Total
2004	4235	2157	1938	470	<b>8800</b>
2012	6127	2504	1938	470	<b>11039</b>
2013	11642	2504	1938	470	<b>16554</b>
2015	13060	2504	1938	470	<b>17972</b>
2016	13937	2664	2154	470	<b>19225</b>

Data upload can be automated using a standardized data dictionary with error and consistency checking.

# Paradigm for building a pediatric cancer commons

1. Engage cooperative group(s)
2. Define scope
3. Identify funding source
4. Identify infrastructure
5. Engage project team
6. Identify data sources
7. Establish governance, create policies and procedures
8. Create contributor / use agreements
9. Create standards working group to create data dictionary, map elements
10. Create database
11. Build front-end query engine
12. Create and execute communication and education plans
13. Create sustainability model

Volchenbom SL, Cox SM, Heath A, Resnick A, Cohn SL, Grossman R  
"Data Commons to Support Pediatric Cancer Research"



# Governance / Regulation / Compliance

## Data contributor agreement

### NEUROBLASTOMA DATA CONTRIBUTOR AGREEMENT

This Neuroblastoma Data Cloud Agreement (this “**Agreement**”) is made as of [DATE] (the “**Effective Date**”), by and between The University of Chicago (the “**University**”), and [PARTNER], a [JURISDICTION OF INCORPORATION] [ENTITY TYPE], [ADDRESS] (“**Partner**”), and, together with the University, the “**Parties**”).

#### RECITALS

WHEREAS, the University has created a technology platform (the “**Platform**”) software, hardware, and other technologies, for storing and harmonizing massive data sets of genomic, electronic medical record, and other information related to neuroblastoma (“**Neuroblastoma Data**”);

WHEREAS, as part of the Platform, the University owns and operates a data service that provides authorized researchers and other users with access to Neuroblastoma Data provided by data contributors;

WHEREAS, Partner has assembled large data sets of Neuroblastoma Data from anonymous individuals and associated clinical data (“**Clinical Data**”);

WHEREAS, Partner desires to: (i) contribute certain of its Neuroblastoma Data (“**Contributed Data**”), as further described on one or more Contributed Data Agreements (the “**Contributed Data Agreements**”), below), to the Platform and (ii) permit the University to provide researchers and other users with access to such Contributed Data, subject to the restrictions set forth in this Agreement; and

WHEREAS, the University is willing to accept such Contributed Data.

## Data sharing agreement

### INTERNATIONAL NEUROBLASTOMA RISK GROUP MASTER DATA USE AGREEMENT

This International Neuroblastoma Risk Group Data Use Agreement (this “**Agreement**”) is made as of [DATE] (the “**Effective Date**”), by and between The University of Chicago (the “**University**”), and [PARTNER], a [JURISDICTION OF INCORPORATION] [ENTITY TYPE], [ADDRESS] (“**Partner**”), and, together with the University, the “**Parties**”).

#### RECITALS

WHEREAS, the University has created a technology platform (the “**Platform**”), including software, hardware, and other technologies, for storing and harmonizing massive data sets of genomic, electronic medical record, and other information related to neuroblastoma;

WHEREAS, as part of the Platform and in collaboration with the International Neuroblastoma Risk Group (“**INRG**”) the University owns and operates a data service that provides authorized researchers and other users with access to such genomic, electronic medical record and other information (“**Contributed Data**”) provided by various data contributors (each a “**Data Contributor**”);

WHEREAS, Partner desires to permit its researchers to access the Contributed Data, subject to the restrictions set forth in this Agreement; and

WHEREAS, the University is willing to provide such access subject to the terms and conditions set forth in this Agreement.



# COG RMS data dictionary

Field Name	Data Type	Description	Value Constraints	Notes
PATIENT_ID	INTEGER	Unique Patient Identification number, assigned after data submission		
USI	TEXT	Universal specimen index (COG patients)		
STUDY	TEXT	Protocol Number		
AGE	INTEGER	Age (in years) at date of diagnosis	99 = Unknown	
GENDER	INTEGER		1 = Male 2 = Female	
HISTOLOGY	INTEGER	Type of Rhabdomyosarcoma	1 = ARMS 2 = ERMS 3 = BRMS 4 = NOS 5 = Undifferentiated Sarcoma 6 = Sarcoma, not classifiable 7 = Spindle cell 8 = Ectomesenchymoma 9 = Other 10 = Mixed RMS 99 = Unknown	
STAGE	INTEGER	IRS TMN classification for pretreatment clinical assessment of disease	1 = Favorable Site, M0 2 = Other Site, any T, a, N0, M0 3 = Other Site, any T, a, N1, M0, any T, b, N0/N1, M0 4 = Metastases, M1	
GROUP	INTEGER	IRS Surgical-Pathologic Grouping System	1 = Completely removed 2 = IIA Microscopic residual, Margin + Nodes - 3 = IIB Microscopic residual, Margin - Nodes +	
PRI_ORBIT	INTEGER	Pr		
PRI_HN	INTEGER	Pr		

Field Name	Data Type	Description	Value Constraints	Notes
PRI_PM	INTEGER	Primary site of tumor is parameningeal	0 = No 1 = Yes 9 = Unknown	12 = Other Head & Neck 13 = Infratemporal Fossa 14 = Middle Ear 15 = Nasal Cavity & Sinus 16 = Nasopharynx 17 = Paranasal Sinus 18 = Parapharyngeal Area 19 = Pterygopalatine 20 = Cheek & PM Extension 21 = Larynx & PM Extension 22 = Orbit & PM Extension 23 = Oropharynx & PM Extension 24 = Other H&N & PM Extension 25 = Parotid & PM Extension 26 = Scalp & PM Extension
PRI_GU	INTEGER	Primary site of tumor is GU, non-te	0 = No 1 = Yes 9 = Unknown	27 = Cervix 28 = Epididymis 29 = Kidney 30 = Ovary 31 = Penis 32 = Spermatic Cord 33 = Testis-Paratestis 34 = 35 = 36 = 37 = 38 = 39 =
tumor is bladder/prostate	INTEGER		0 = No 1 = Yes 9 = Unknown	40 = 41 = 401 =
tumor is extremity	INTEGER		0 = No 1 = Yes 9 = Unknown	42 = 43 = 44 = 45 = 46 = 47 = 48 = 49 = 50 = 51 = 62 =
tumor is intrathoracic	INTEGER		0 = No	

			1 = Yes 9 = Unknown	63 = Diaphragm 64 = Heart 65 = Hilum 66 = Lung & Local Sites 67 = Lung & Other Sites 68 = Lung 69 = Mediastinum 70 = Pericardium 71 = Pleura 72 = Pleural Effusion 73 = Thymus 74 = Trachea
PRI_PERI_ANUS	INTEGER	Primary site of tumor is perineum/anus	0 = No 1 = Yes 9 = Unknown	75 = Anus 76 = Perineum
PRI_RETROPERI	INTEGER	Primary site of tumor is retroperineum	0 = No 1 = Yes 9 = Unknown	77 = Pelvis, Site Indeterminate 78 = Retroperitoneum
PRI_TRUNK	INTEGER	Primary site of tumor is trunk	0 = No 1 = Yes 9 = Unknown	79 = Abdominal Wall 80 = Breast 81 = Chest Wall 82 = Paraspinal
PRI_OTHER	INTEGER	Primary site of tumor is other	0 = No 1 = Yes 9 = Unknown	52 = Esophagus 53 = Gall Bladder & Biliary Tree 54 = Intestine, large 55 = Intestine, small & duodenum 56 = Liver 57 = Omentum 58 = Pancreas 59 = Peritoneal Nodules 60 = Peritoneum 61 = Stomach 83 = Adrenal Glands 84 = Ascites 85 = Bone

METASTATIC	INTEGER	Metastases is distant	0 = No 1 = Yes 9 = Unknown	
MET_LUNG	INTEGER	Metastatic site of tumor at the time of diagnosis is lung	0 = No 1 = Yes 9 = Unknown	
MET_BONE	INTEGER	Metastatic site of tumor at the time of diagnosis is bone or bone marrow	0 = No 1 = Yes 9 = Unknown	
MET_NODE	INTEGER	Metastatic site of tumor at the time of diagnosis is distant lymph nodes	0 = No 1 = Yes 9 = Unknown	
MET_STS	INTEGER	Metastatic site of tumor at the time of diagnosis is soft tissue sites	0 = No 1 = Yes 9 = Unknown	
MET_PE	INTEGER	Metastatic site of tumor at the time of diagnosis is pleural effusion	0 = No 1 = Yes 9 = Unknown	
MET_OTHER	INTEGER	Metastatic site of tumor at the time of diagnosis is other	0 = No 1 = Yes 9 = Unknown	
MEASUREMENT	INTEGER	Maximum diameter of tumor (cm)	99 = Unknown	
INVASIVE	INTEGER	Tumor invasiveness	1 = Yes [confined to organ of origin], T1 2 = No [not confined to organ of origin], T2 9 = Unknown	
NODAL_CLINICAL	INTEGER	Clinical or imaging evidence for nodal involvement	0 = No, N0 1 = Yes, N1 9 = Not Evaluated/Unknown	

NODAL_PATH	INTEGER	Pathological evidence for Nodal Involvement	0 = No, N0 1 = Yes, N1 9 = Not Evaluated/Unknown	
FUSION	INTEGER	Fusion Status	0 = FOX01- 1 = FOX01+ 9 = Unknown	
PAX	INTEGER	PAX fusion partner	3 = PAX3 7 = PAX7 9 = Unknown	
ANAPLA	INTEGER	Anaplasia status	0 = None 1 = Focal 2 = Diffuse 9 = Unknown	
EFS_EVENT	INTEGER	Censoring flag for event	0 = Patient is Censored (No Event) 1 = Patient had an event	Event = first occurrence of relapse, progression, secondary malignancy, or death from any cause
EVENT_TYPE	INTEGER	Type of first event	0 = No Event 1 = Relapse/Progression 2 = Second Malignancy 3 = Death 9 = Cannot determine if first event	
EFS_TIME	INTEGER	Time (in days) from diagnosis/enrollment to first event, or to last contact if no event occurred		
OS_EVENT	INTEGER	Censoring flag for death	0 = Patient is Censored (No Death) 1 = Patient died	Death = death from any cause
OS_TIME	INTEGER	Time (in days) from diagnosis/enrollment to death, or to last contact if patient is alive		



# International RMS meeting - October 2017



## INSTRuCT

(INTERNATIONAL SOFT TISSUE SARCOMA CONSORTIUM)

Chicago Meeting

Tuesday, October 17 8:00am – Wednesday, October 18 4:00pm

Polsky Exchange North  
1452 East 53rd Street, 2nd floor

Room: Promontory Point

Chicago, IL 60615

773-795-0209



THE UNIVERSITY OF CHICAGO MEDICINE & BIOLOGICAL SCIENCES



CENTER FOR RESEARCH INFORMATICS

@samvolchenboum

# Consensus harmonization of primary site

CWS

COG

EpSSG

Major Primary Site	CWS	COG	EpSSG/MMT Name
ORBIT	Eyelid	1=Eye	Eyelid
	Orbit	2=Orbit	Orbit
HEAD & NECK (non PM)	Scalp	10=Scalp	Soft tissue of scalp External auricular canal Ear soft tissue, external ear Temporal muscle
	Parotid	9=Parotid	Parotid, soft tissue
	Oral Cavity	7=Oral cavity	Gum Base of tongue Lip Lower lip Upper lip Tongue
	Larynx	5=Larynx	Larynx
	Oropharynx	8=Orophaynx	Oropharynx Lingual tonsil Mandible soft tissue Bone of face (Maxillar) Masseter Oral cavity Cheek
	Cheek	3=Cheek	Hypopharynx
	Hypopharynx	4=Hypopharynx	Thyroid
	Thyroid & Parathyroid	11=Thyroid & Parathyroid	Neck Neck Supra-clavicular soft tissues Neck, nodes Nos Chin
	Neck	6=Neck	Soft tissue face (non specified region) Face specified region Nasolabial fold (skin) Nostril
			12=Other Head & Neck





# Consensus harmonization of primary site

CWS COG EpSSG

Major Primary Site	CWS	COG	EpSSG/MMT Name
ORBIT	Eyelid	1=Eye	Eyelid
	Orbit	2=Orbit	Orbit
HEAD & NECK (non PM)	Scalp	10=Scalp	Soft tissue of scalp External auricular canal Ear soft tissue, external ear Temporal muscle
	Parotid	9=Paratoid	Parotid, soft tissue
	Oral Cavity	7=Oral cavity	Gum Base of tongue Lip Lower lip Upper lip Tongue
	Larynx	5=Larynx	Larynx
	Oropharynx	8=Orophaynx	Oropharynx Lingual tonsil Mandible soft tissue Bone of face (Maxillar) Masseter Oral cavity
	Cheek	3=Cheek	Cheek
	Hypopharynx	4=Hypopharynx	Hypopharynx
	Thyroid & Parathyroid	11=Thyroid & Parathyroid	Thyroid
	Neck	6=Neck	Neck Neck Supra-clavicular soft tissues Neck, nodes Nos Chin
		12=Other Head & Neck	Soft tissue face (non specified region) Face specified region Nasolabial fold (skin) Nostril



Orbit		CUI
	Eyelid	<a href="#">C0015426</a>
	Orbit	<a href="#">C0029180</a>
	Other orbit	<a href="#">C0700042</a>
Head and Neck		
	Scalp	<a href="#">C0036270</a>
	Parotid	<a href="#">C0030580</a>
	Oral cavity	<a href="#">C0226896</a>
	Larynx	<a href="#">C0023078</a>
	Oropharyngeal	<a href="#">C0521367</a>
	Cheek	<a href="#">C0007966</a>
	Hypopharynx	<a href="#">C0020629</a>
	Thyroid and/or parathyroid	<a href="#">C0574117</a>
	Neck	<a href="#">C0027530</a>
	Other head and neck	<a href="#">C0460004</a>



# Consensus example: Maximum tumor diameter

Old

- Maximum diameter
- **or**
- X,Y,Z
- **or**
- >5 cm vs. <5 cm



# Consensus example: Maximum tumor diameter

## Old

- Maximum diameter
- **or**
- X,Y,Z
- **or**
- >5 cm vs. <5 cm

## New

- Discrete measurement (in cm)
  - X (or max diameter if single)
  - Y
  - Z
- Category (if no discrete meas.)
  - $\leq 5$  cm
  - $> 5$  cm
  - Unknown



# Building pediatric cancer data commons

April 10 COG meeting to engage all disease groups

- Acute myeloid leukemia
- Acute lymphoblastic leukemia
- Bone tumors
- Central nervous system tumors
- Germ cell tumors
- Hodgkin Disease
- Neuroblastoma
- Non-Hodgkin lymphoma
- Renal tumors
- Soft tissue sarcoma



# Building pediatric cancer data commons

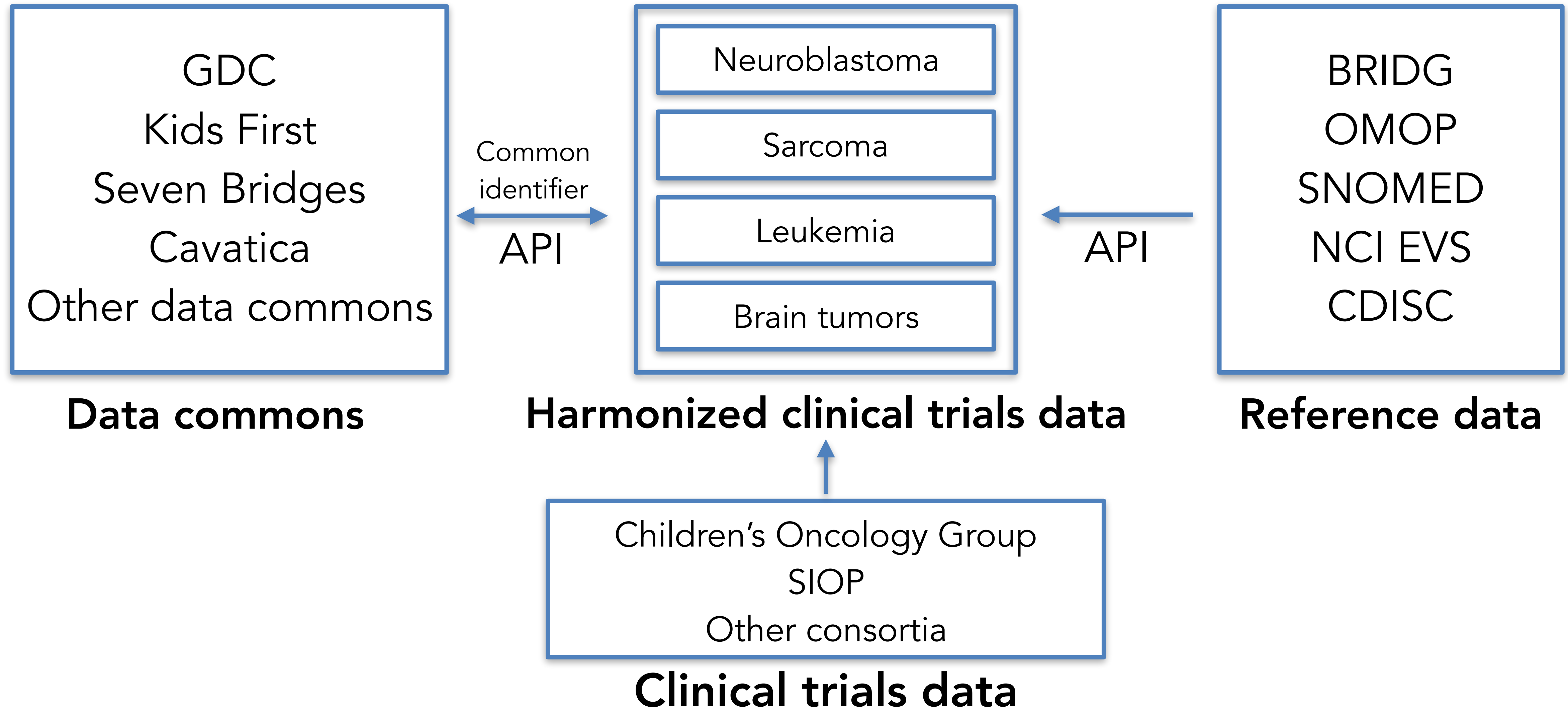
**April 10 COG meeting to engage all disease groups**

- Acute myeloid leukemia
- Acute lymphoblastic leukemia
- Bone tumors
- Central nervous system tumors
- Germ cell tumors
- Hodgkin Disease
- Neuroblastoma
- Non-Hodgkin lymphoma
- Renal tumors
- Soft tissue sarcoma

**Goal to discuss data and sample collection for all pediatric oncology.**



# Paradigm for pediatric cancer clinical trials data commons



# Summary / Call to action

- Pediatric cancer requires novel, genotype-based therapies
- Discovery requires innovations in sample collection, processing, and annotation
- Harmonized data leads to shared phenotype/clinical data
- Data and samples must have universal identifiers
- We must envision data collection and sharing at all stages of care
- The goal is all data from all patients at all times





# Center for Research Informatics

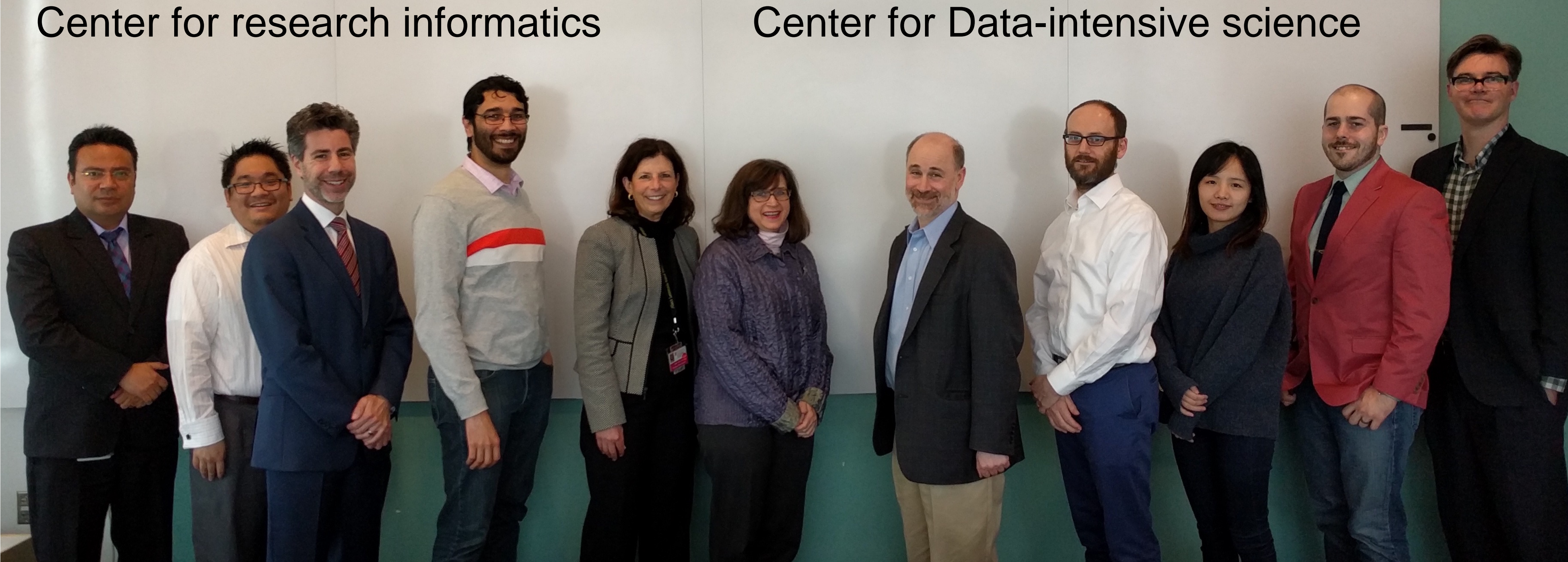
Applications - Systems - Bioinformatics - Data warehousing - Clinical trials



# Acknowledgements

Center for research informatics

Center for Data-intensive science





THE UNIVERSITY OF  
CHICAGO MEDICINE &  
BIOLOGICAL SCIENCES



CENTER FOR  
**RESEARCH**  
INFORMATICS

@samvolchenboum