

# DEVELOPMENT OF AN ELECTRONIC HEALTH RECORD DERIVED MULTICENTER INPATIENT PEDIATRIC DATA REPOSITORY

---

**Christoph P Hornik, MD MPH**

Department of Pediatrics, Duke University  
Duke Clinical Research Institute

Andrew Atz, Catherine Bendel, Francis Chan, Kevin Downes, Robert Grundmeier, Ben Fogel, Debbie Gipson, Matthew Laughon, Michael Miller, Michael Smith, Chad Livingston, Cindy Kluchar, Anne Heath, Robert McDonald, Chanda Jarrett, Brian McKerlie, Hetalkumar Patel, Christina Hunter



**PEDIATRIC  
TRIALS NETWORK**

Making drugs safer & more effective  
for use in the youngest patients



*Eunice Kennedy Shriver* National Institute  
of Child Health and Human Development

**A project of the Best Pharmaceuticals for Children Act**

# Disclosures

- Nothing to disclose

# Electronic Health Record Data in Pediatric Clinical Research

- Most pediatric diseases are rare, limiting the sample size of clinical studies
- Electronic health records (EHR) contain large volumes of data collected during routine medical care
  - Applications in health services research, trial design/planning, other
- Barriers to widespread and inter-institutional use remain
  - Data structure and codes
  - Data privacy

# NICHD Pediatric Trials Network (PTN)

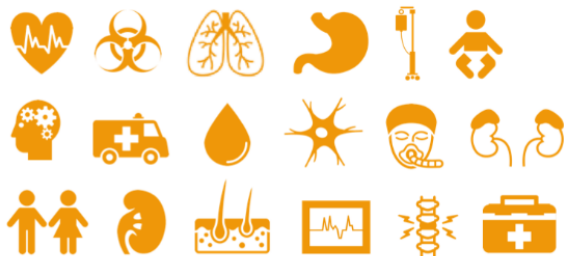
## Scope

26 clinical trials  
12 other studies  
38 total studies



Phase 1–4 studies

18 therapeutic areas\*



>7000 children



## Needs

- Epidemiology
  - Natural history
  - Efficacy endpoints
  - Safety endpoints
  
- Feasibility and simulation
  
- Alternate data source for efficacy and safety

# Objective

- Develop an EHR based pediatric inpatient repository to support the design and conduct of drug trials in children

# Participants

- 9 PTN sites
- Inclusion criteria:
  - Age at admission <17 years
  - Admitted as inpatient between 06/01/2013 and 06/31/2017
  - Admission source records available from EHR or Data Warehouse
- Exclusion criteria:
  - Lack of EHR record for hospital encounter
- Oracle based repository at Duke Clinical Research Institute in FISMA moderate environment

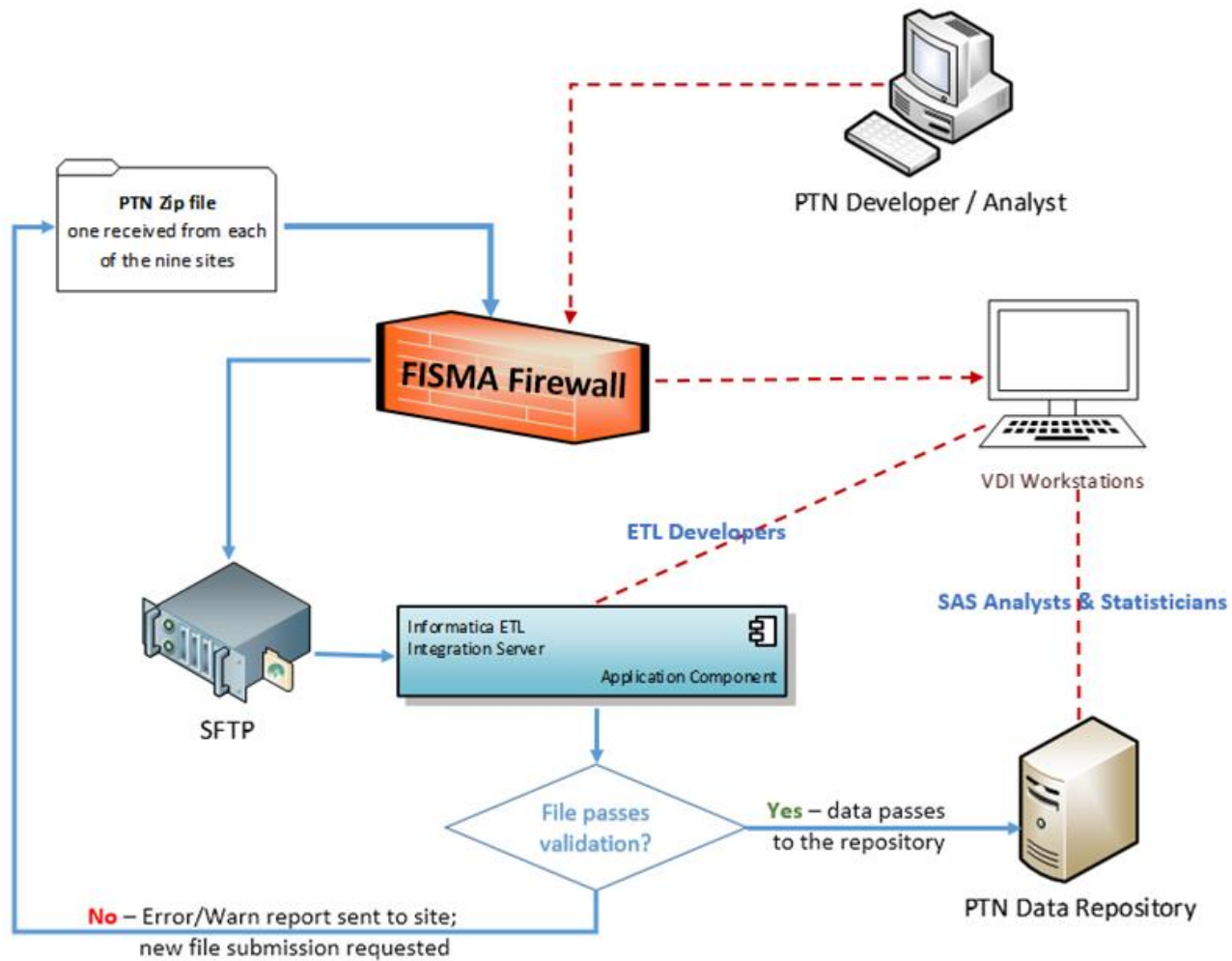
# Data Model

- Guidance from PCORnet model v2.0
- 147 mandatory and 99 optional data elements
- 14 domains

<b>Demographic</b>	<b>Encounter</b>	<b>Conditions/ Problems</b>	<b>Diagnoses</b>	<b>Procedures</b>	<b>Allergies</b>	<b>Immunization</b>
<b>Labs</b>	<b>Medications administered</b>	<b>Medications prescribed</b>	<b><i>Vitals Signs</i></b>	<b><i>Fluids</i></b>	<b><i>Flowsheet assessments</i></b>	<b><i>Respiratory Support</i></b>

- Site based data mapping using implementation guide
- Data master list:
  - Validation rules
  - Coded value sets, data formats
  - Updated version after first submissions

# Infrastructure





# Testing

- Iterative submission process:
  - Training sessions and follow up teleconferences with site staff
  - FAQ document for site use
  - Multiple test submissions – 177 total submissions
- 520 rules and warnings:
  - 0% errors
  - <2% warnings
- Database and logic tested in development and validation environments prior to moving to production

# 386,455 encounters - 264,709 children

Variables	Median (range) or %
Age	2.3 (0, 17)
Female	53%
Race	
American Indian or Alaska Native	<1%
Asian	4%
Black or African American	19%
Native Hawaiian or Other Pacific Islander	<1%
White	58%
Multiple Races	2%
Other	9%
Hispanic Ethnicity	18%
Discharge Year	
2013	11%
2014	25%
2015	26%
2016	26%
2017	12%

# Commonly administered drugs

31 million drug administration records

	Albuterol	Acetaminophen
Total number of doses	1,113,838	1,083,344
Unique patients	29,699	110,967
Average doses per patient	37.5	9.7
Dose, median (25 <sup>th</sup> , 75 <sup>th</sup> %ile)	2.5mg (2.5, 12.5)	13mg/kg (10, 15)

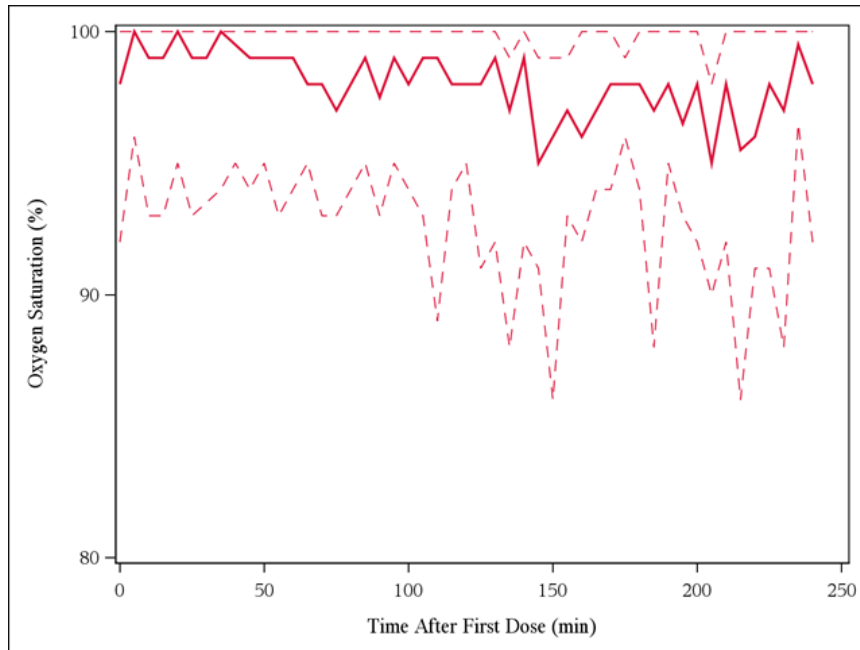
# Commonly diagnoses

4.7 million diagnosis records

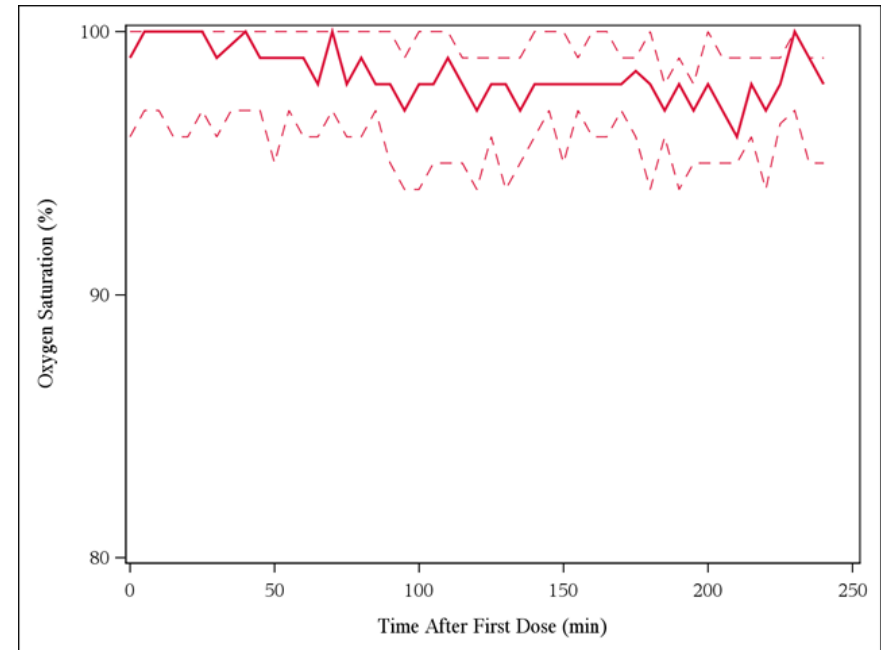
	<b>Asthma</b>	<b>Fever</b>
Unique encounters	35,647	29,173
Unique patients	23,883	19,059

# Oxygen saturation following ketamine administration

0 to <2 years



2 to < 6 years



# Limitations and lessons learned

## Limitations

- Number of sites
- Inpatient EHR data
- Data model partially based on PCORnet
- Scalability

## Lessons learned

- Site support
- Early validation
- Flexibility
  - Rules
  - Formats

# Conclusions

- Data repository of pediatric inpatient EHR data from 9 sites
- Data model based on PCORnet but enriched with key variables for pediatric drug development
- Future applications of the repository include:
  - Pharmacoepidemiology
  - Feasibility and clinical trial simulation
  - Efficacy and safety assessments

# Acknowledgement

- We thank the principal investigators and research teams.

Site PIs	Study Team	City	State
Michael Miller	George Lales Larry Blackburn Herman Wagner Tracie Smith	Chicago	IL
Janice Sullivan Mike Smith	Steve Heilman Kamakhya Singh Sarah Penny	Louisville	KY
Andrew Atz	Katie Kirchoff Tomoko Goddard	Charleston	SC
Kevin Downes Bob Grundmeier	Mark Ramos Shawn O'Connor	Philadelphia	PA
Francis Chan	John Loor Stephanie Fan	Loma Linda	CA
Deb Gipson	Don Liamini Amitava Shee Samara Attalla	Ann Arbor	MI
Catherine Bendel	Brian Harvey Sonya Grillo Duy Dong	Minneapolis	MN
Ben Fogel	Matt Bolton Nमित Govil	Hershey	PA
Matthew Laughon	Adam Lee	Chapel Hill	NC



**PEDIATRIC  
TRIALS NETWORK**

Making drugs safer & more effective  
for use in the youngest patients



*Eunice Kennedy Shriver* National Institute  
of Child Health and Human Development

**A project of the Best  
Pharmaceuticals for Children  
Act**

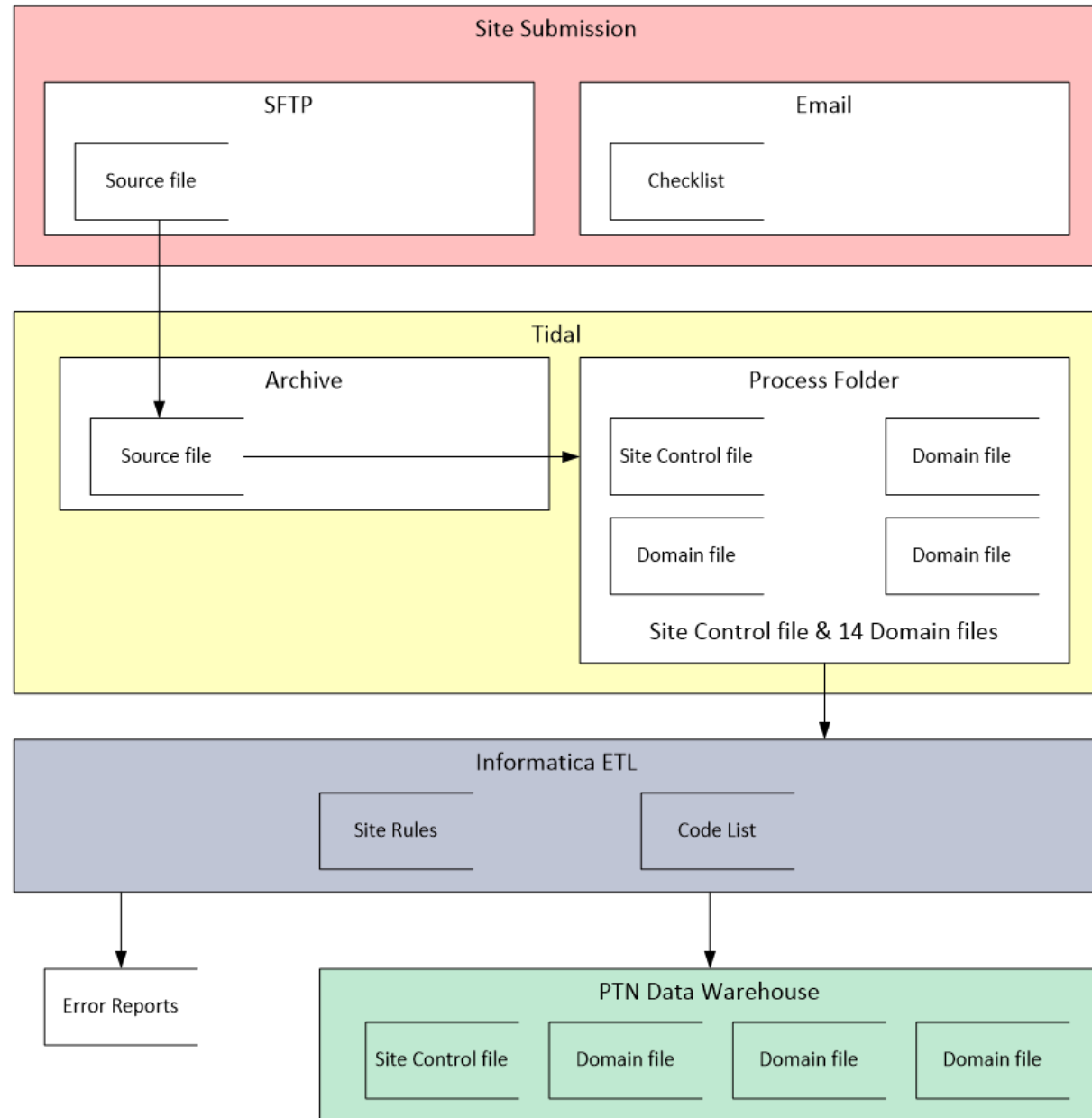


# Additional slides

# Site Selection

- 30 high enrolling sites from the NICHD PTN (~80 sites)
- Site survey:
  - EHR availability
  - Site resources to analyze data and map
  - Availability of IT infrastructure to extract files
  - Interest
- Site selection calls:
  - 21 questions sent in advance of call
- Selection criteria:
  - PCORNET experience
  - Experienced informatics/technical teams
  - Prior experience using EHR data in-house for clinical research

# Process Flow



# Other workflow steps

- Data point selection:
  - No domains with 100% compatibility with PCORnet
  - New data points were developed
- Inventory assessments: goal >60% of data points
- File transmission and processing:
  - Implementation guide
  - FAQ document
  - sFTP delivery method to drop site
  - 15 informatica ETL tools workflow processes
  - Error reports generated and returned to site
  - Average of 20 submission needed per site

# Data Domains

DEMOGRAPHICS	CONDITIONS/PROBLEM LIST	IMMUNIZATION	ALLERGY	MEDS ADMINISTERED	MEDS PRESCRIBED	FLWSHEETS
SITE ID PATIENT ID PATIENT DATE OF BIRTH GESTATIONAL AGE (weeks) DEATH DATE DEATH DATE IMPUTE DEATH CAUSE PATIENT RACE PATIENT RACE OTHER MATERNAL RACE MATERNAL RACE OTHER PATIENT GENDER PATIENT ETHNICITY HISPANIC MATERNAL ETHNICITY HISPANIC SMOKING STATUS SMOKING STATUS OTHER PATIENT ZIP CODE	SITE ID PATIENT ID ADMIT DATE/TIME COND REPORT DATE COND DX CODE COND COND CODE SYSTEM COND CODE OTHER COND ONSET DATE COND RESOLVE DATE COND STATUS COND SOURCE CHRONIC COND INDICATOR PRINCIPAL COND INDICATOR	SITE ID PATIENT ID ADMIT DATE/TIME VACC CODE VACC ADMIN DATE VACC NAME VACC ROUTE VACC DOSE QUANTITY VACC DOSE UOM VACC REACTION DATE ADVERSE REACTION DETAIL VACC STATUS VACC REPORT SOURCE	SITE ID PATIENT ID ADMIT DATE/TIME ALLERGEN NAME TYPE OF ALLERGEN ONSET DATE ADVERSE REACTION RECORDED DATE CRITICALITY CRITICALITY OTHER INTOLERANCE STATUS INTOLERANCE STATUS OTHER	SITE ID PATIENT ID ADMIT DATE/TIME MED ADMIN START DATE/TIME MED ADMIN CODE CODE SYSTEM CODE SYSTEM OTHER MED ADMIN STOP DATE/TIME MED ADMIN BRAND NAME MED ADMIN GENERIC NAME MED ADMIN STATUS MED ADMIN STATUS OTHER MED ADMIN DOSE FORM MED ADMIN DOSE MED ADMIN DOSE UNIT MED ADMIN DOSE FREQUENCY MED ADMIN DOSE FREQ OTHER MED DOSING WEIGHT MED DOSING WEIGHT UNIT MED ADMIN ROUTE MED ADMIN ROUTE OTHER	SITE ID PATIENT ID ADMIT DATE/TIME RX ORDER DATE/TIME MED PRESCRIBED CODE CODE SYSTEM CODE SYSTEM OTHER RX MED NAME RX SOURCE RX START DATE RX END DATE RX QUANTITY RX DOSE RX DOSE UNIT RX FREQUENCY RX FREQUENCY OTHER RX REFILLS RX PACKAGE DESCRIPTION RX DAYS SUPPLY RX INSTRUCTIONS RXNORM CUI	SITE ID PATIENT ID ADMIT DATE/TIME  <b>VITALS</b> VITAL SIGN REPORTED MEASURE DATE/TIME VITAL SIGN REPORTED OTHER VITALS SOURCE RESULT NUM UNIT OF MEASURE RESULT QUAL  <b>RESPIRATORY SUPPORT</b> RESP SUPPORT TYPE RESP SUPPORT TYPE OTHER RESP SUPPORT START DATE RESP SUPPORT STOP DATE  <b>FLUIDS</b> INTAKE/OUTPUT INDICATOR MEASURED DATE/TIME TYPE/ROUTE TYPE/ROUTE OTHER FLUID VOLUME MEASURE FLUID VOLUME UNITS INTAKE FLUID MEASURED  <b>ASSESSMENTS</b> ASSESSMENT TOOL ASSESSMENT AREA ASSESSMENT DATE/TIME ASSESSMENT CATEGORY/TYPE ASSESSMENT SCORE/VALUE NUM ASSESSMENT SCORE/VALUE QUAL SCALE MIN VALUE SCALE MAX VALUE ASSESSMENT ANALYSIS
ENCOUNTER	DIAGNOSES	PROCEDURES	LAB RESULTS		CONTROL FILE	
SITE ID PATIENT ID ADMIT DATE/TIME ENC TYPE ADMITTING SOURCE FACILITY LOCATION DISCHARGE DATE/TIME DISCHARGE STATUS DISCHARGE STATUS OTHER DISCHARGE DISPOSITION DRG CODE DRG CODE TYPE ENROLLED IN STUDY	SITE ID PATIENT ID ADMIT DATE/TIME DX DATE TIME DX CODE DX CODE SYSTEM DX CODE SYSTEM OTHER DX SOURCE DX SOURCE OTHER POX POX OTHER DX RESOLVE DATE	SITE ID PATIENT ID ADMIT DATE/TIME PX DATE/TIME PX CODE PX CODE SYSTEM PX CODE SYSTEM OTHER PX SHORT DESCRIPTION PX LONG DESCRIPTION PX REASON PX SOURCE PX STATUS PX ADVERSE EVENT DESCRIPTION	SITE ID PATIENT ID ADMIT DATE/TIME LAB ID LAB TEST RESULT STATUS RESULT DATE/TIME LAB TEST COMPONENT ID LAB ORDER DATE LAB TEST SHORT NAME LAB TEST LONG NAME LAB TEST LOINC CODE LAB TEST PX CODE LAB TEST PX CODE SYSTEM LAB TEST PX CODE SYSTEM OTHER LAB SUB CATEGORY CODE LAB SUB CATEGORY NAME LAB CULTURE ORGANIZISM CODE LAB CULTURE ORGANIZISM NAME LAB PANEL LAB TEST SPECIMEN SOURCE	LAB TEST SPECIMEN SOURCE OTHER LAB TEST NORM RANGE LOW LAB TEST NORM RANGE HIGH SPECIMEN COLLECTED DATE/TIME LAB ORDER PRIORITY LAB ORDER PRIORITY OTHER LAB RESULT DX CODE ORIG LAB RESULT UNIT RESULT QUAL RESULT QUAL OTHER RESULT NUM RESULT NUM_UNIT RESULT MODIFIER RESULT MODIFIER OTHER NORM MODIFIER LOW NORM MODIFIER LOW OTHER NORM MODIFIER HIGH NORM MODIFIER HIGH OTHER ABN IND ABN IND OTHER	SITE ID DEMOGRAPHICS RECORD COUNT ENCOUNTER RECORD COUNT COND/PROB LIST RECORD COUNT DIAGNOSES RECORD COUNT PROCEDURE RECORD COUNT LAB RECORD COUNT MEDS ADMIN RECORD COUNT MEDS PRESC RECORD COUNT IMMUNIZATION RECORD COUNT ALLERGIES RECORD COUNT VITALS RECORD COUNT FLUIDS RECORD COUNT ASSESSMENT RECORD COUNT RESP SUPP RECORD COUNT NOTES	

# Commonly administered drugs

Drug name	N=
Albuterol Sulfate	974,551
Acetaminophen	656,159
Lorazepam	494,086
Furosemide	482,635
Morphine	385,013
Ranitidine	379,373
Fentanyl	357,729
Dexmedetomidine	348,053
Midazolam	326,321
Oxycodone HCl	270,333

# Common Diagnoses

Diagnosis	N=
Feeding problems/feeding difficulties	40,767
Gastroesophageal reflux disease	38,178
Fever, unspecified	25,388
Dehydration	18,143
Gastrostomy tube	18,107
Other dyspnea and respiratory anomaly	18,039
Asthma, unspecified	17,498
Acute respiratory failure	16,778
Neonatal and fetal jaundice, unspecified	13,709
Atrial septal defect	12,589