



#HL7

April 2016 Taha Kass-Hout

Joint work with many colleagues at FDA, White House, NIH, and ONC

Any views or opinions expressed here do not necessarily represent the views of the FDA, HHS, or any other entity of the United States government. Furthermore, the use of any product names, trade names, images, or commercial sources is for identification purposes only, and does not imply endorsement or government sanction by the U.S. Department of Health and Human Services.



*FDA advancing precision medicine with precisionFDA:
a collaborative informatics community
to explore regulatory science*





... AGCATCGATGCAGAAAGATTACAAGACGATCCGCTC ...

THE PRECISION MEDICINE INITIATIVE

... AGCATCGATGCAGAAGATTACAACATAAAAAGATTACACGCGATCCGCTC ...

reference

... AGCATCGATGCATAAGATTACAAGATAAAAAGATTACACGCGATCCGCTC ...

george

... AGCATCGATGCA AAGATTACAACATAAAAAGATTACATGCGATCCGCTC ...

elaine

... AGCATCGCTGCAGAAGATTACAA ATAAAAGATTACATGCGATCCGCTC ...

omar



Global Genes™

Allies in Rare Disease

RARE DISEASE TYPES



7,000+

distinct types of rare diseases exist, with more being discovered every day.

THE CAUSE

80%

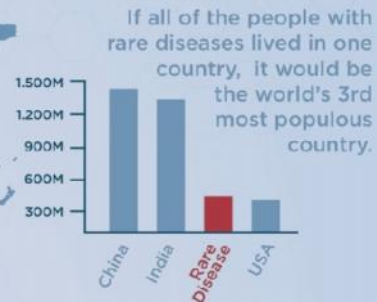
of rare diseases are caused by faulty genes.



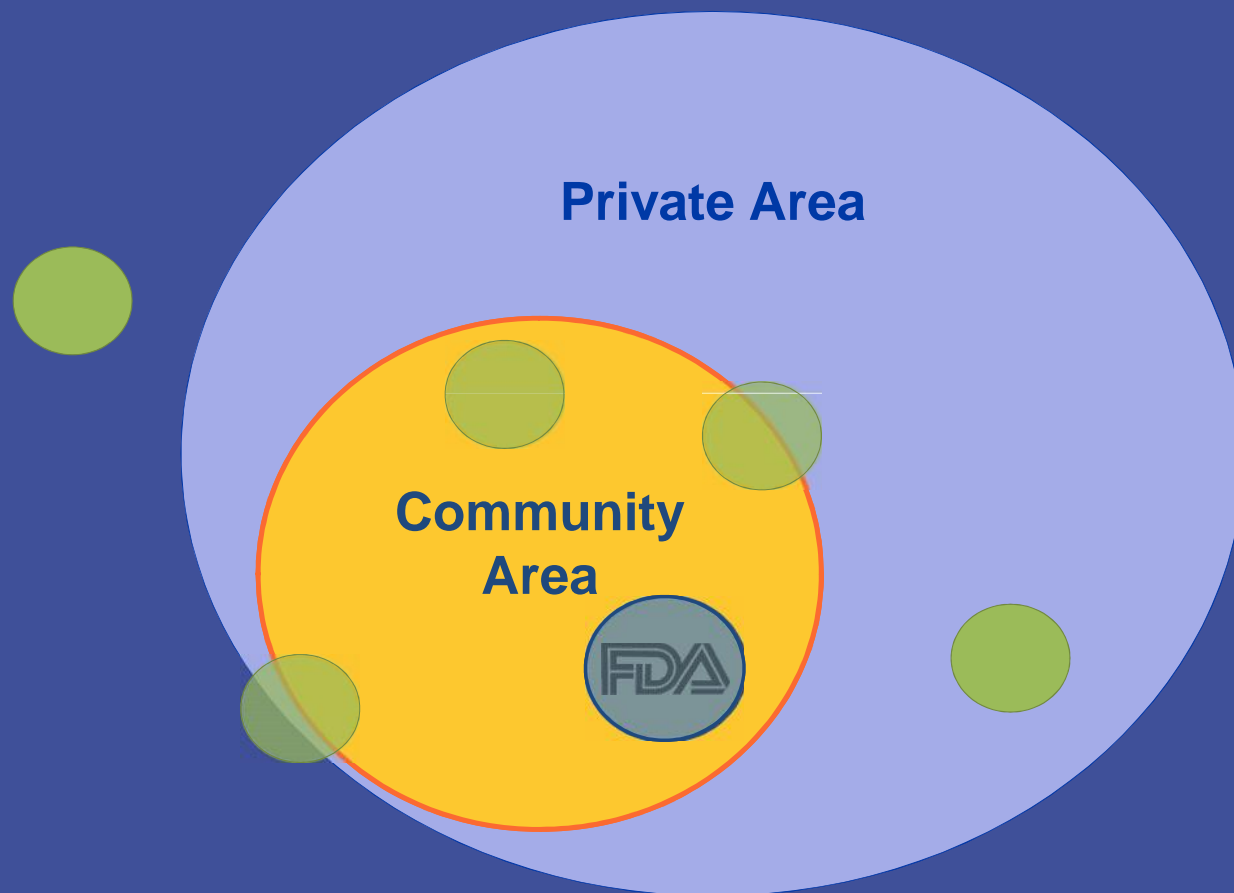
RARE DISEASE EFFECT

30 MILLION AMERICANS

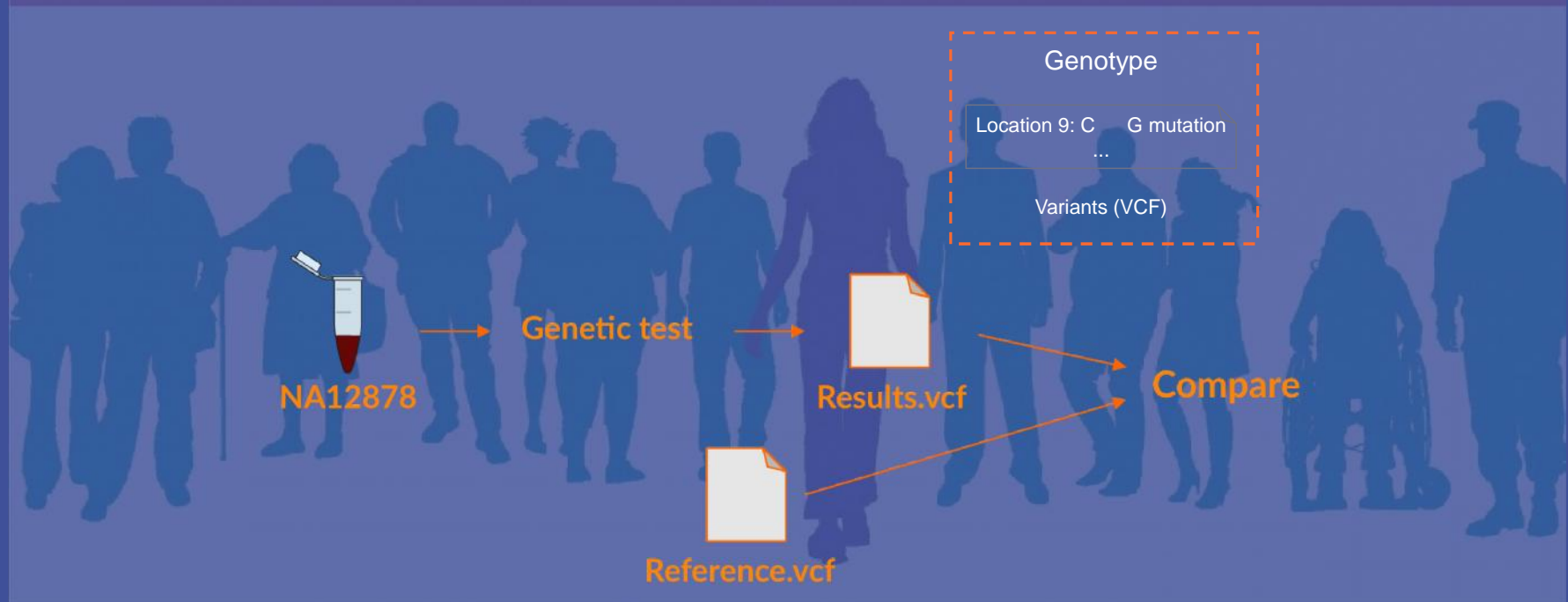
350 MILLION WORLDWIDE



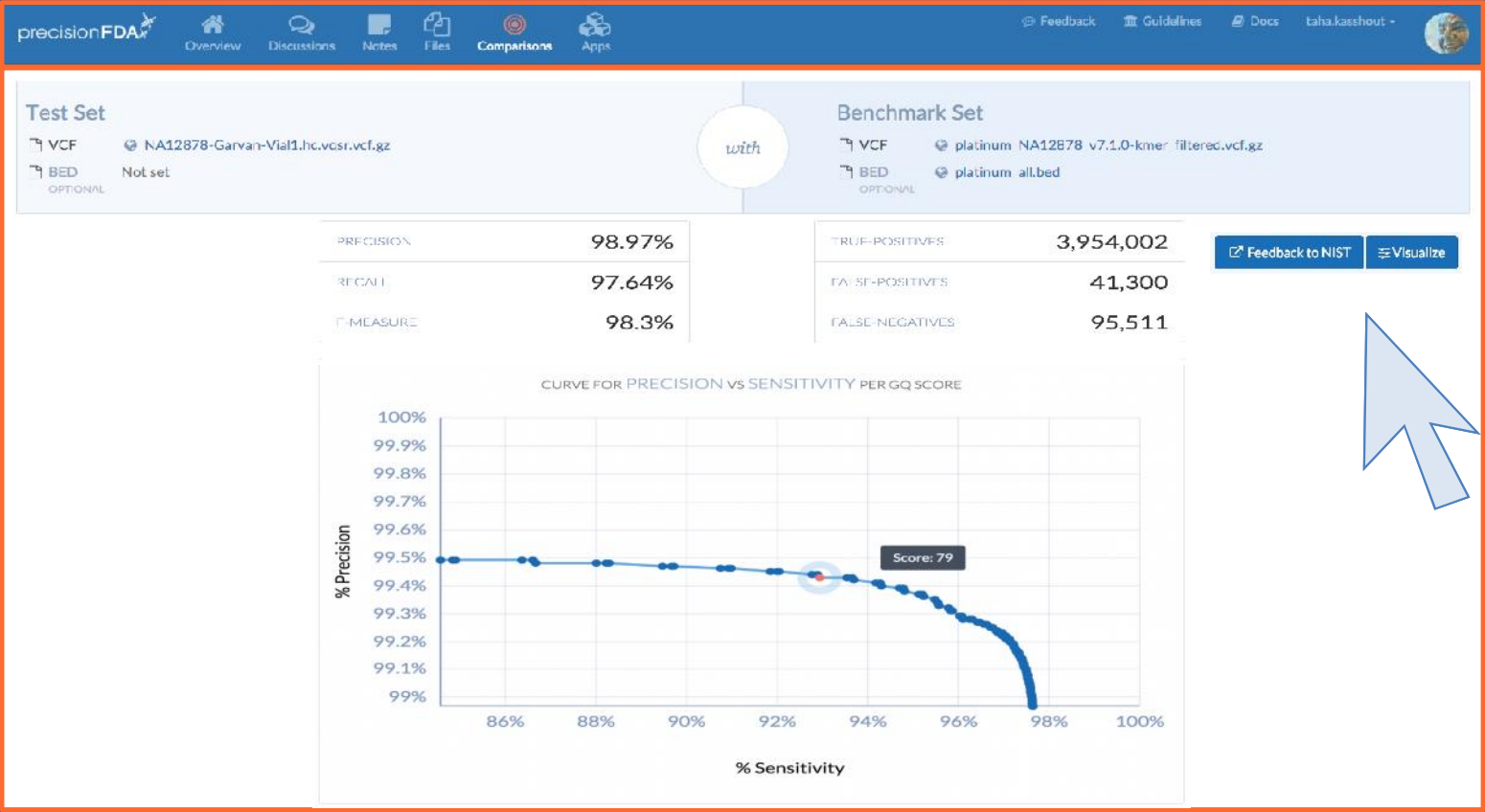
precisionFDA Eco-System



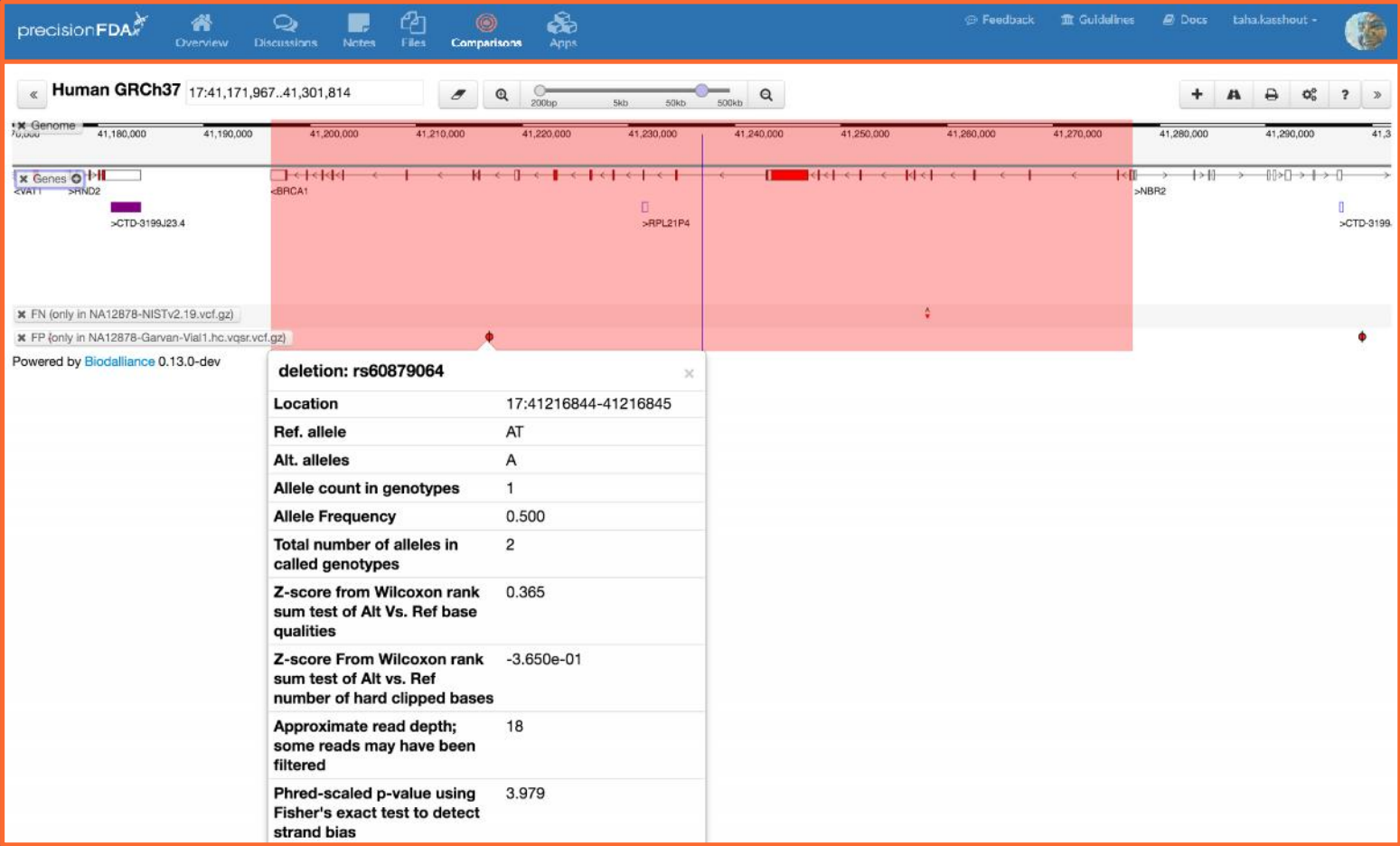
THE PRECISION MEDICINE INITIATIVE



Comparisons



Comparisons



FHIR Use Case



Quality Resource Extension

of tp
of fp
of fn
Recall
Precision
f-score

Community Challenges



Join the precisionFDA Consistency Challenge!

Join the FDA in engaging and improving DNA test results with the first community challenge. Learn more: precision.fda.gov.



President Obama's Precision Medicine Initiative envisions a day when an individual's medical care will be tailored in part based on their unique characteristics and genetic make-up.



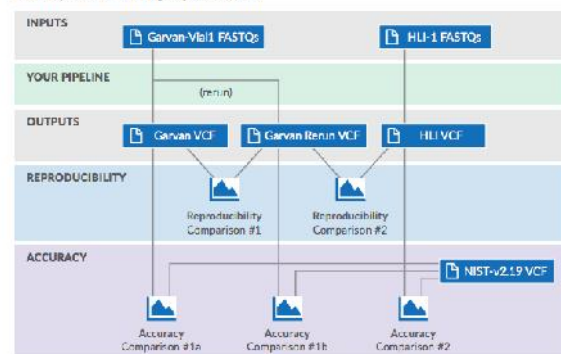
The goal of the FDA's first precisionFDA challenge is to engage the genomics community in order to achieve more consistent results in the context of genetic tests (relative to whole human genome sequencing), advancing the goal of better personalized care.



PrecisionFDA invites all innovators to take the challenge and assess their software on the supplied reference human datasets. Participation is voluntary, but instrumental in helping the community prepare for the coming genomic data revolution.

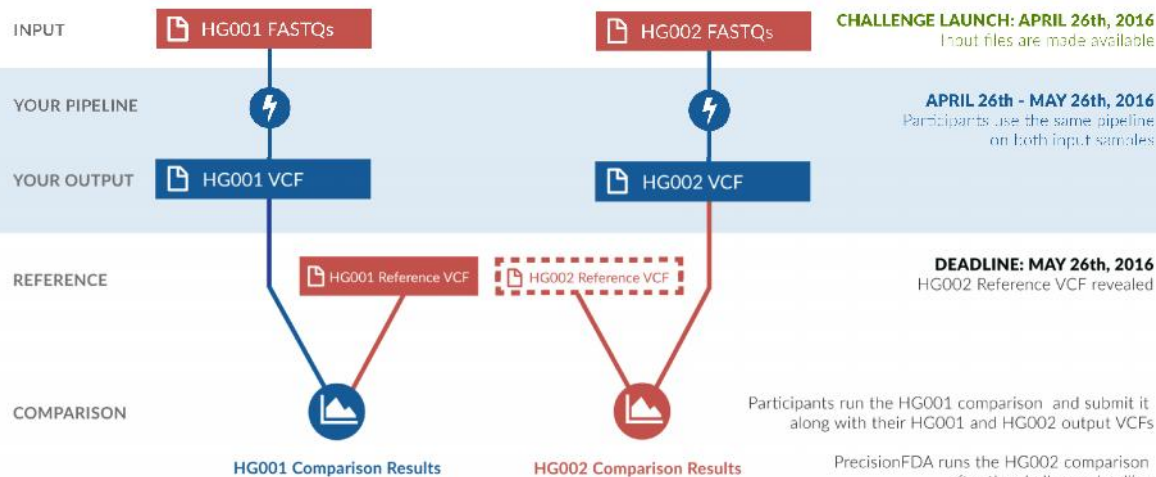
AT A GLANCE

PrecisionFDA Consistency Challenge
February 25, 2016 through April 25, 2016



PrecisionFDA Truth Challenge

April 26, 2016 through May 26, 2016



precision.fda.gov | precisionFDA@fda.hhs.gov | [@precisionFDA](https://twitter.com/precisionFDA) | github.com/fda



Discussions



+ Start a discussion

Elaine Johanson (FDA) posted
PrecisionFDA Truth Challenge (NEW!)
0 comments

Deepak Grover (Sanofi-Genzyme) posted
Comparing multiple VCF files
0 comments

Bahram Kerani (Crystal Genetics, Inc.) posted
Call for Participation in Data Sharing
0 comments

Elaine Johanson (FDA) posted
PrecisionFDA Consistency Challenge
0 comments

precision.fda.gov | precisionFDA@fda.hhs.gov | @precisionFDA | github.com/fda

Apps

precisionFDA									
Overview Discussions Notes Files Comparisons Apps Feedback Guidelines Docs taha.kasshout									
RELEVANT APPS FEATURED EXPLORE									
+ Create App Manage Assets									
NAME	TITLE	REVISION	EXPLORERS	ORG	ADDED BY	CREATED ↑	ACCESS	RUN BY YOU?	
sentieon-hc--grch37	Run Sentieon DNaseq 201603.01 using GRCh37 bundle files	14	1	sentieon	rafael.aldana	2016-04-06 20:52:33	Public	Try	
bai	Bam Index Maker	3	1	crystalgenetics	bahram.kermani	2016-04-03 09:15:59	Public	Try	
bamsubsample	BAM Subsample	1	1	crystalgenetics	bahram.kermani	2016-04-03 05:13:30	Public	Try	
varsim-art-simulation	VarSim: Simulation	15	4	roche	marghoob.mohiyuddin	2016-02-29 20:55:51	Public	Try	
vcf-comparison	VCF Comparison	3	6	precisionfda	george.asimenos	2016-02-29 19:17:39	Public	Try	
bwa_mem_bamsormadup	BWA-MEM Map Illumina reads	2	8	precisionfda	george.asimenos	2016-01-13 18:38:43	Public	Try	
hive-insilico	HIVE Insilico Simulate GRCh37 reads from BED and VCF	3	6	precisionfda	george.asimenos	2015-12-23 20:25:48	Public	Try	
key_gene_coverage	Coverage of Key Genes	44	1	stanford	rachel.goldfeder	2015-12-18 19:23:21	Public	Try	
elprep	elPrep	19	1	intel	pascal.costanza	2015-12-18 10:54:09	Public	Try	
pharmgkb_pipeline_v03	PharmGKB NGS pipeline - Phase 1	9	1	stanford	lester.carter	2015-12-16 07:24:18	Public	Try	
rfmix-v170-r0-fda	Local Ancestry Analysis with RFMIX	33	1	stanford	mark.wright	2015-12-12 21:50:45	Public	Try	
varsim-vcfcompare	VarSim: Validation	11	3	roche	marghoob.mohiyuddin	2015-12-11 01:11:51	Public	Try	
freebayes	FreeBayes	14	3	roche	marghoob.mohiyuddin	2015-12-10 01:13:24	Public	Try	
bwa_mem_bamsormadup	BWA-MEM Accepts interleaved paired read fastq	3	1	stanford	mark.wright	2015-12-07 23:30:41	Public	Try	
varsim-art-clinvar-simulation	VarSim+ART Simulate FASTQ files with ClinVar variants	1	6	precisionfda	george.asimenos	2015-12-01 06:18:43	Public	Try	

<https://precision.fda.gov/apps>

precision.fda.gov | precisionFDA@fda.hhs.gov | [@precisionFDA](https://twitter.com/precisionFDA) | github.com/fda

Notes

Welcome to precisionFDA!

Welcome to precisionFDA, the community platform for NGS assay evaluation and regulatory science exploration. You have been given [early access to this closed beta](#) so that you

NA12878

The NA12878 Coriell DNA sample is a widely studied DNA biospecimen; it has been sequenced with various technologies and has been the reference point of many studies. The precisionFDA team has...

1

NA12878 on HiSeq X Ten

How many variants of NA12878 can be recovered using a single lane of X Ten sequencing?

We set to answer this by comparing [NA12878-Garvan-Vial1.hc.vqsr.vcf.gz](#) (which is a single lane of H GATK pipeline) to the reference datasets [NA12878-NISTv2.19.vcf.gz](#) and [NA12878-Illumina-Plati](#)

In [NA12878-Garvan vs NA12878-PlatinumGenome](#), sensitivity (recall) was 96%.

But when comparing against NIST, given its more narrow [NA12878-NISTv2.19.bed](#) footprint, results v

NA12878-Garvan vs NA12878-NISTv2.19		
99.31%	99.14%	99.22%
PRECISION	RECALL	F-MEASURE

Do you think there is room for improvement? Upload your own NA12878 variants files and publish your c

precisionFDA

OverviewDiscussionsNotesFilesComparisonsAppsFeedbackGuidelinesDocsHelp

Coverage of ACMG Coding Bases for NA12878 Garvan Vial 1

ACCESS	ID	ORG	ADDED BY	CREATED
Public	na12878	Standard	richard.hill@fda.hhs.gov	2015-12-18 19:47:49

Track

Here, I walk through an example application of the [Coverage of Key Genes](#) app.

Preparation

Input data

The following FASTQ files are used as input:

- NA12878-Garvan-Vial1_K1_L1.fastq.gz
- NA12878-Garvan-Vial1_K2_L1.fastq.gz

As mentioned in a previous note, these files are the result of sequencing the NA12878 sample on one lane of an Illumina HiSeq X Ten instrument.

Alignment

I aligned these files to the [hg19](#) reference genome using the [BWA-MEM](#) app.

Note that it is important to provide a Read Group ID (any text is fine) for the downstream GATK coverage calculations.

The alignment job created a BAM and BAI file that we will use for the next step:

- NA12878-Garvan-Vial1.bam
- NA12878-Garvan-Vial1.bai

Coverage Calculation

Input data and parameters

Using the newly generated BAM and BAI files, I calculated the coverage in the ACMG Gene coding bases with the [Coverage of Key Genes](#) app.

The ACMG Gene coding bases are found in the following file:

- acmg.coding.bed

I used the following parameters:

- Minimum Base Quality: 30
- Minimum Mapping Quality: 2
- Minimum Depth of Coverage: 20

To run a similar analysis yourself, you'll need to upload your copy of the [coverageKeyGenes.txt](#) file. This can be easily obtained, here: <https://www.ncbi.nlm.nih.gov/ftp/gen/na12878/coverageKeyGenes.txt>

Results

The app creates plots and a table showing the number and percentage of bases in each gene with < 20x coverage of >=30 bases in reads with mapQ >= 1:

- coverageKeyGenes.txt
- NA12878-Garvan-Vial1_L100seq.L1na20.coverage.metrics.txt

FILES

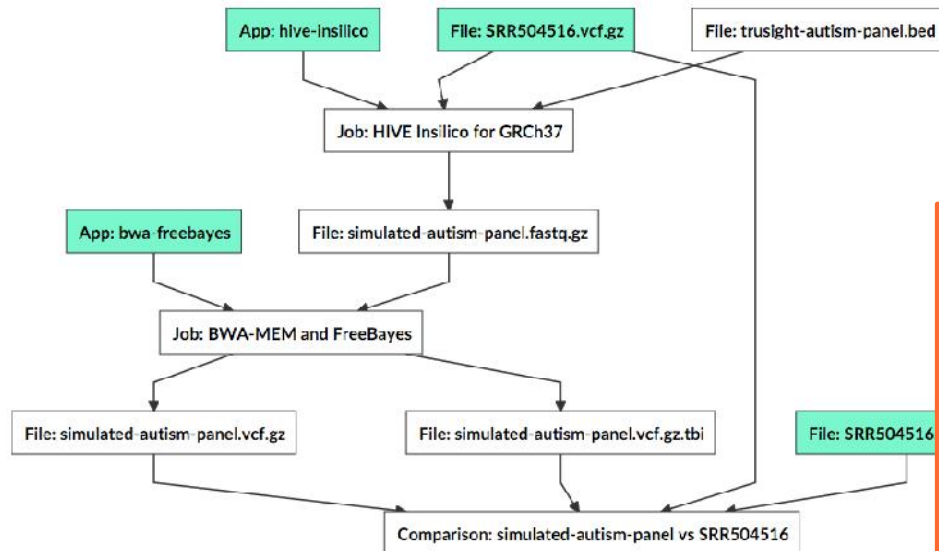
- NA12878-Garvan-Vial1_L100seq.L1na20.coverage.metrics.txt
- coverageOfKeyGenes.txt
- coverageOfKeyGenes.txt
- NA12878-Garvan-Vial1_L1na20.coverage.metrics.txt
- NA12878-Garvan-Vial1_L1na20.coverage.metrics.txt
- NA12878-Garvan-Vial1_L1na20.coverage.metrics.txt

JOB

- Coverage of Key Genes
- BWA-MEM | Map Illumina reads

Sharing

Welcome to the publishing wizard. You have chosen to publish the item shown at the bottom of the graph. This item was generated through other items (files, apps, jobs). By also publishing these related items, you allow people to trace back your steps and even repeat them. The graph below shows how the item got generated, and allows you to choose what else you would like to publish. (Items shown in green are already published).



Taha Kass-Hout ✓
@DrTaha_FDA

A community member shared their @BaseSpace data on @precisionfda using the Fetch from URL feature #Interoperability precision.fda.gov/about/how#file...

RETWEETS
5

LIKES
3



4:25 PM - 22 Apr 2016



Documentation

About precisionFDA

WHY?

Background and motivation

WHAT?

A genomics community and platform

WHO?

The team behind the initiative

Introduction

Welcome to precisionFDA, the community platform for NGS assay evaluation and regulatory science exploration.

We are excited to have you on board, and have prepared this guide to help you make the most out of the precisionFDA platform. We hope that this guide will answer many of your questions and provide you with additional insights to further empower you to use the system. If your favorite topic is not covered, please don't hesitate to contact us, and we'll make sure to expand the guide accordingly.

If you've already logged onto the system, you will have noticed that the precisionFDA website is divided into four sections: Notes, Files, Comparisons and Apps. These are further discussed in the sections below.

TIP: The precisionFDA platform is currently in beta, and is expected to evolve and change over the next few weeks. If you encounter an error page during your interaction with the site, we encourage you to report feedback and tell us what you were trying to do at the time and how you got to the error page. We appreciate your help!

Introduction

Files

Comparisons

Apps

Creating Apps

Notes

As with all systems, precisionFDA relies on files to store data. Files can be uploaded from your computer or generated by running apps, and can be shared with the precisionFDA community.

Listing files

Clicking on "Files" at the top navigation bar takes you to the page that lists all the files accessible by your account. Your private files are marked as "Private" in the Access column. Members' contributions (including yours) are shown as "Public". You can further filter files (for example, show only your private files, or search for a given filename keyword) by clicking on the "eye" icon, which will present you with the filter bar. Choose any filters and apply them by clicking on the "filter" icon.

Uploading files

You can upload files from your computer or generate them from apps. To begin the upload process, click the "Upload" button in the top right corner of the Files page. The system will automatically calculate a checksum for the file and upload it to the cloud side.

Once the upload is complete, the file is usually taken a few seconds, during which the file becomes "closed" and can no longer be accessed. Its content is provided as input to apps. Its content is provided as input to apps. Its content is provided as input to apps.

If the upload gets interrupted for any reason, the file is marked with an "open" state. If the upload gets interrupted for any reason, the file is marked with an "open" state. If the upload gets interrupted for any reason, the file is marked with an "open" state.

TIP: Do you need to upload large files? The precisionFDA team allows you to upload large files into the system via a dedicated interface.

Introduction

Files

Listing files
Uploading files
Tracking file origin
File details

Comparisons

Apps

Creating Apps

Notes

The concept of comparing two sets of variants (VCF files) is central to the execution of regulatory science, and to the evaluation of NGS assays. It is therefore represented as a first-class entity on precisionFDA.

The process of comparing VCF files involves several steps. The precisionFDA platform provides a set of tools to facilitate this process. The platform provides a set of tools to facilitate this process. The platform provides a set of tools to facilitate this process.

This initial framework compares two variants using a heuristic and a benchmark set. The underlying comparison methodology is mostly symmetric with respect to each assignment; however, reporting of the results is based on the assumption that the benchmark set represents the truth, and that the test set represents a prediction. Therefore, the results of a comparison will be an implicit evaluation of the performance of whatever method was used to generate the predictions.

Introduction

Files

Comparisons

Creating a new comparison
BED files
Understanding comparison results
Visualizing a comparison
Tools for editing comparisons

Apps

Creating Apps

Notes

Creating a new comparison

To create a new comparison, click the "Comparisons" button in the top navigation bar. The system will prompt you to select a test set and a benchmark set. You can select a test set and a benchmark set. You can select a test set and a benchmark set.

To create a new app, click "Create App" in the Apps page. The following section walks you through important concepts of app development.

TIP: Want to learn by example? Simply choose any of the public apps in precisionFDA and click "Fork". This will load up the app editor, where you can take a look at the internals of the app and see what it is comprised of. You can then hit the Back button in your web browser -- unless of course you truly want to fork the app and make a private copy with which you can experiment, in which case click the new Fork button from inside the app editor to complete the operation.

App naming conventions

Apps have a machine-readable name that cannot contain spaces (such as "bwa-mem-and-freebayes") and a human-readable title (such as "BWA-MEM and FreeBayes"). Among apps that you create, names need to be unique (you cannot author two distinct apps with the same name). This restriction is only per-user, meaning that you can still create an app with the same name as someone else's app. In fact, the system encourages you to use someone else's app as a starting point and make further tweaks and save it as your own app (a process called "forking" an app). This model was inspired from the model of GitHub repositories.

Input and Output spec

Apps require an input/output specification, which mandates what inputs they need from the user, and what outputs they are expected to generate. Note that an "input" is anything that needs to be received from the user and which can potentially vary between executions. These can be not only input files but also numerical or boolean values, and strings. In that sense, the "input" can be used both for receiving data to operate on as well as receiving configuration parameters. Each input field has the following properties:

Property	Explanation
Class	The kind of input. There are exactly five classes supported: file, string, integer, float, and boolean.

Introduction

Files

Comparisons

Apps

Creating Apps

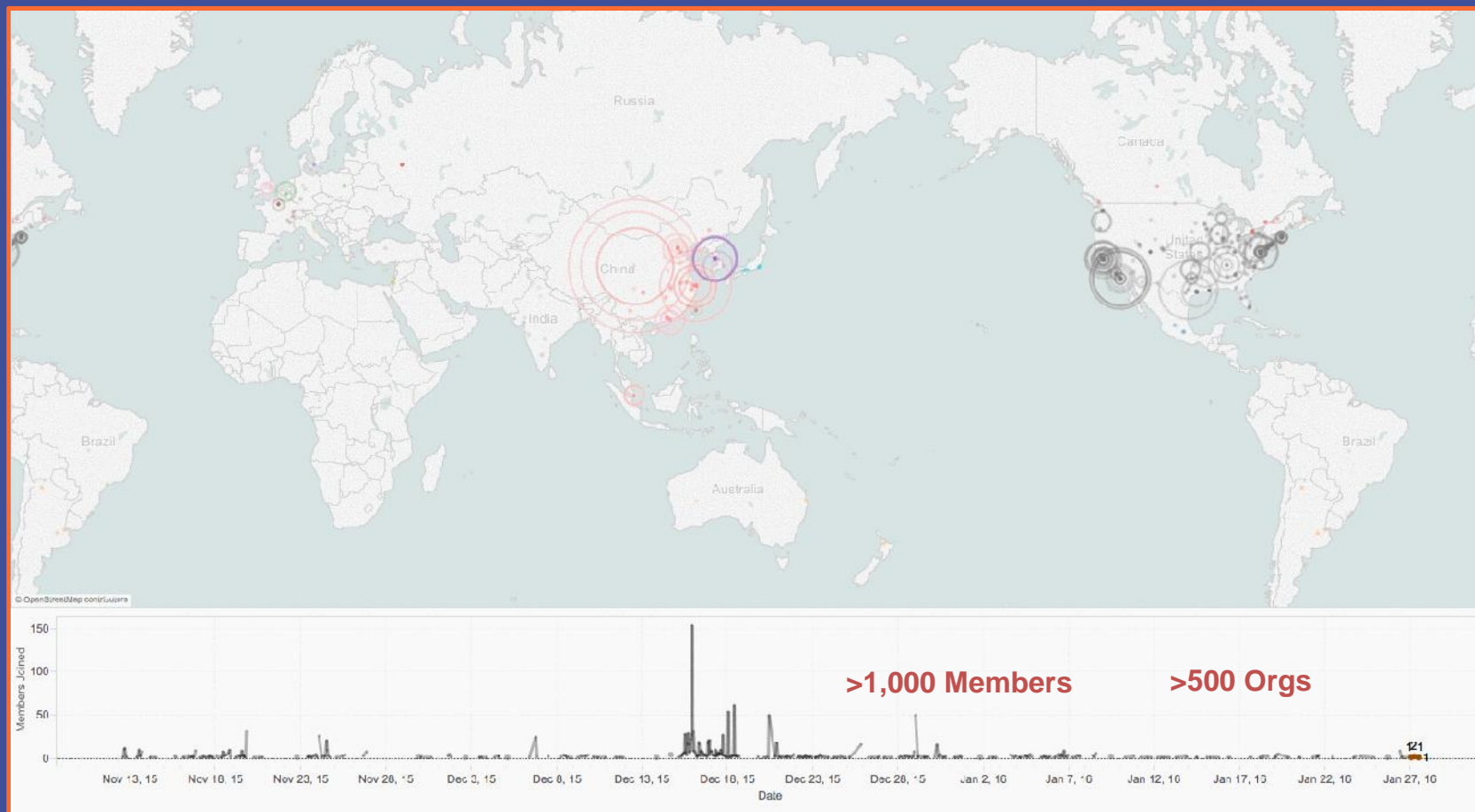
App naming conventions
Input and Output spec
VM Environment
Available instance types
App assets
Your own assets
App script
Forking an app

Notes

FierceBiotechIT

FDA hits tight timeline for start of precisionFDA open beta

December 21, 2015 | By Nick Paul Taylor



precision.fda.gov | precisionFDA@fda.hhs.gov | [@precisionFDA](https://twitter.com/precisionFDA) | github.com/fda

A research roadmap for next-generation sequencing informatics

Russ B. Altman,^{1*} Snehit Prabhu,² Arend Sidow,³ Justin M. Zook,^{4,11} Rachel Goldfeder,⁵ David Litwack,⁶ Euan Ashley,⁷ George Asimenos,⁸ Carlos D. Bustamante,⁹ Katherine Donigan,⁹ Kathleen M. Giacomini,⁹ Elaine Johansen,⁶ Natalia Khuri,¹⁰ Eunice Lee,⁶ Xueying Sharon Liang,⁶ Marc Salit,^{4,10,11} Omar Serang,⁸ Zivana Tezak,⁶ Dennis P. Wall,¹² Elizabeth Mansfield,⁶ Taha Kass-Hout⁶

Next-generation sequencing technologies are fueling a wave of new diagnostic tests. Progress on a key set of nine research challenge areas will help generate the knowledge required to advance effectively these diagnostics to the clinic.

The Precision Medicine Initiative (PMI) is a U.S. national effort “to enable a new era of medicine through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized care” (1). One goal is to bring about the routine use of next-generation precision diagnostics to benefit individuals and public health. Central to the introduction of safe and effective new precision diagnostic technologies is an adequate understanding of how well they perform. Through the PMI, the U.S. Food and Drug Administration (FDA) is seeking to address this issue by providing dynamic, flexible, and well-balanced regulation of precision diagnostics. Because these complex technologies pose new challenges in understanding their likely benefits and their limits in terms of accuracy, precision, and clinical validity, FDA is advancing a robust research agenda in regulatory science. New knowledge gained from this agenda will inform the next generation of regulation for precision medicine.

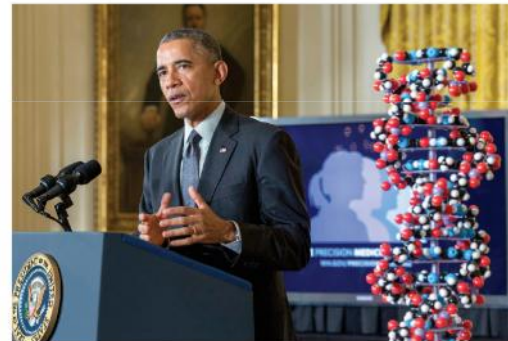
The UCSF-Stanford Center for Excellence in Regulatory Science & Innovation (CERSI) hosted a series of meetings in September 2015

that included a public workshop and discussions on identifying key activities needed to evaluate the clinical implications of next-generation nucleic acid sequencing (NGS). Here we summarize the ideas and directions that were proposed and put forth a working “roadmap”

Eventually, these developments are likely to culminate in the routine sequencing of patients’ genomes. In the meantime, there will be several years during which the process of DNA sequence determination remains challenging and in which cost-, quality-, and goal-driven trade-offs result in a large diversity of testing strategies. In this Perspective, we lay out the technological challenges that are slowing the routine clinical use of a new generation of genetic tests and propose questions that regulatory science should address to arrive at a flexible yet robust regulatory framework that results in maximum benefit for patients.

As part of its PMI effort, FDA seeks to undertake and support regulatory science research that will enhance our understanding of NGS test products and their development and validation, as well as how the results of such tests are best communicated in an evolving health care environment.

A centerpiece of this effort is precisionFDA, a research and development portal that



Sharing the stage with a DNA double helix, U.S. President Barack Obama discusses the Precision Medicine Initiative.

for NGS evaluation, as a possible exemplar of how many other new next-generation diagnostics may be understood.

DEFINING THE TASK

Technological breakthroughs have recently led to DNA sequencing methods that can generate the raw data necessary for determining nearly the entire genome sequence of any individual.

will allow community members to better understand, develop, and improve existing and new bioinformatics approaches for processing the vast amount of genomic data that is collected using NGS technology. precisionFDA is a public, cloud-based platform developed by FDA and its contractor DNAnexus that hosts shared tools, crowdsourced testing, and community challenges, to improve and share

¹Bioengineering, Genetics, and Medicine, Stanford University, Stanford, CA 94305, USA. ²Biomedical Data Science and Genetics, Stanford University, Stanford, CA 94305, USA. ³Pathology and Genetics, Stanford University, Stanford, CA 94305, USA. ⁴Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. ⁵Biomedical Informatics, Stanford University, Stanford, CA 94305, USA. ⁶Food and Drug Administration, Silver Spring, MD 20993, USA. ⁷Medicine, Genetics, and Pathology, Stanford University, Stanford, CA 94305, USA. ⁸DNAnexus, Mountain View, CA 94040, USA. ⁹Bioengineering and Therapeutic Sciences, University of California at San Francisco, San Francisco, CA 94143, USA. ¹⁰Bioengineering, Stanford University, Stanford, CA 94305, USA. ¹¹Material Measurement Laboratory, National Institute of Standards and Technology, Stanford University, Stanford, CA 94305, USA. ¹²Systems Medicine and Psychiatry, Stanford University, Stanford, CA 94305, USA. *Corresponding author. E-mail: russaltman@stanford.edu

Got Side Effects? We have an (~~App~~) API for that!



**Application Programming Interface (API) provides an easy way for a software application to access, query, or update source data*

A unified set of open-source APIs for FDA drug, device, and food data

Available APIs

Product	Data	Timeframe	# of records
Drug	Labeling	Current	>74K
	Adverse event reports	Since 2003	>6M
	Recalls	Since 2012	>5K
Medical device	Classification	Current	>6K
	Registration & Listing	Current	>24K establishments >100K devices
	510(k)s (including de novos)	Since 1976	>143.5K
	PMAs (including supplements)	Since 1977	>31.6K
	Adverse event reports	Since 1991	>4.7M
	Recalls	Since 2002	>9.5K
	Recalls	Since 2012	>9.6K
Food	Recalls	Since 2012	>9.6K

Query: URL

[https://api.fda.gov/drug/event.json?search=receivedate:\[20040101+TO+20150101\]+AND+_missing_:companynumb&count=receivedate](https://api.fda.gov/drug/event.json?search=receivedate:[20040101+TO+20150101]+AND+_missing_:companynumb&count=receivedate)

Results: Disclaimer

```
{
  meta: {
    disclaimer: "openFDA is a beta research project and not for clinical use. While we make every effort to ensure that data is accurate, you should assume all results are unvalidated.",
    license: "http://open.fda.gov/license",
    last_updated: "2014-08-06"
  },
```

Results: JSON Format

```
  results: [
    {
      time: "20040102",
      count: 83
    },
    {
      time: "20040104",
      count: 53
    },
  ]
}
```

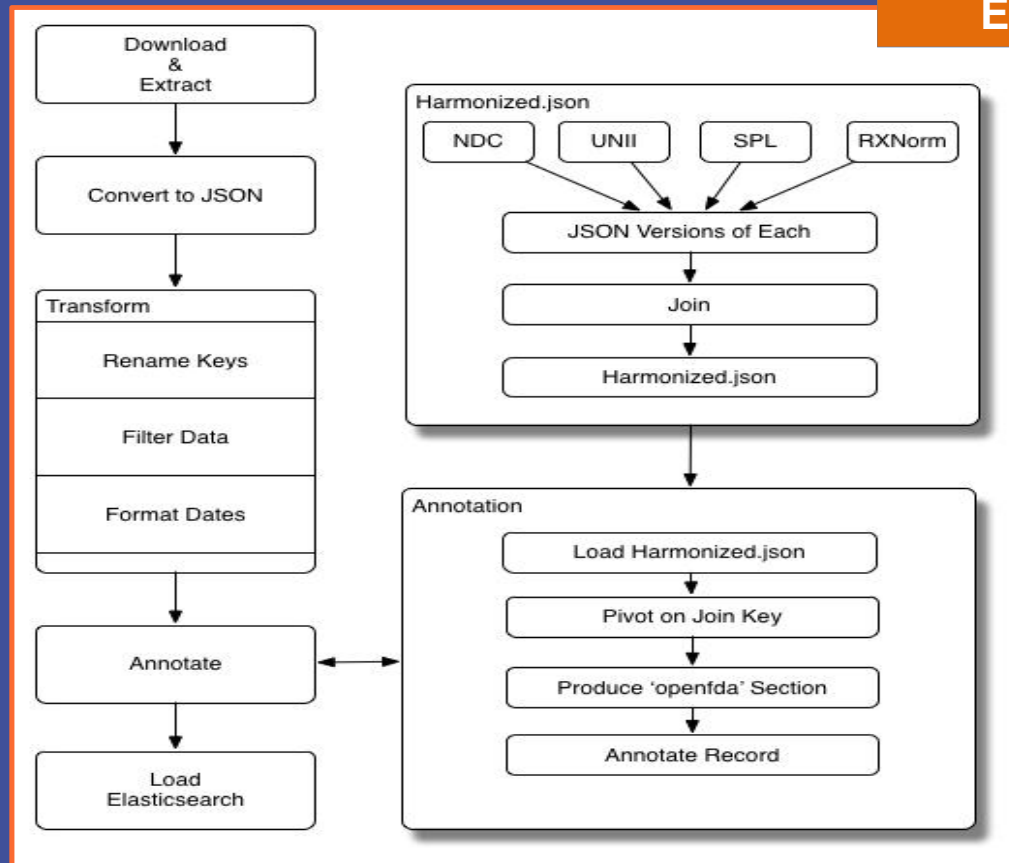
Harmonization

Joins | Identifiers

- National Drug Code
- Unique Ingredient Identification
- Structured Product Labelling
- RxNorm links drug terminologies.

got
- EHR?

ETL



Harmonization

NDC and SPL datasets contribute the openFDA fields that we add to adverse events and recalls records that mention ≥ 1 drug

Contains identifiers such as brand name, generic name, SPL SetID, SPL ID, NDC, NDA, UNII, Drug Class and many more

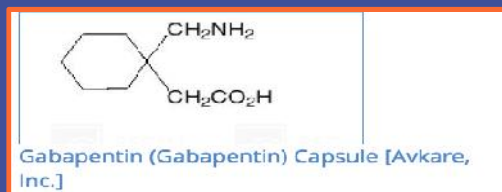
```
openfda: {  
  unii: [  
    "75J73V1629"  
  ],  
  spl_id: [  
    "86e3103c-9d8b-4693-b5db-3fd62330c754"  
  ],  
  product_ndc: [  
    "0004-1963",  
    "0004-1964"  
  ],  
  substance_name: [  
    "CEFTRIAXONE SODIUM"  
  ],  
}
```


Harmonization – Adverse Events

```
{
  drugcharacterization: "2",
  medicinalproduct: "ROCEPHIN",
  openfda: {
    unii: [
      "75J73V1629"
    ],
    spl_id: [
      "86e3103c-9d8b-4693-b5db-3fd62330c754"
    ],
    product_ndc: [
      "0004-1963",
      "0004-1964"
    ],
    substance_name: [
      "CEFTRIAXONE SODIUM"
    ],
    rxcui: [
      "204871",
      "105212"
    ],
    spl_set_id: [
      "9467f6c9-3e59-45c6-a1be-77200f2d4554"
    ],
    product_type: [
      "HUMAN PRESCRIPTION DRUG"
    ],
    pharm_class_cs: [
      "Cephalosporins [Chemical/Ingredient]"
    ],
  },
}
```

Join key is NDA/ANDA (application number) and medicinal products

Harmonization – Enforcement Reports



PAXIL CR (Paroxetine HCL) Controlled-Release Tablets 25 mg, 30-count bottle, Rx only, Manufactured by GlaxoSmithKline, RTP, NC 27709, Manufactured for Apotex Corp., Weston, FL 33328, NDC 60505-3669-3

NDC Code

NDC 0049-2710-30

Free-text fields requiring structure

UPC code #'s FA02-90GM, FA02-180GM, and FA02-270GM

UPC Code(s)

The PREBIOTIC FORMULA was distributed in three sizes (90 gram, 180 gram, and 270 grams). UPC code #'s FA02-90GM, FA02-180GM, and FA02-270GM respectively. It is a berry flavored powder in a white plastic screw off container and contains a scoop. It is distributed by Eco-Health, Inc. and contains the following address on the label. 25876 The Old Road #158, Newhall, CA 91321. The dates of distribution are between June 2011 - Aug 2012 in the quantities mentioned above.

<http://open.fda.gov> | open@fda.hhs.gov | [@openfda](https://twitter.com/openfda) | github.com/fda

openFDA
open.fda.gov

For Drugs

NDC Code

UPC Code

FDA
Application

Brand Name

Dosage Form

Generic Name

Manufacturer
Name

Drug Type

Route Of
Administration

Active
Ingredient

Pharmacologic
Class

Mechanism
of Action

Chemical
Structure

Functional
Activity

UNII Code

RxNorm
Identifier

```
{"logo": "API Evangelist"}
```

Another Strong API Implementation In Federal Government With OpenFDA

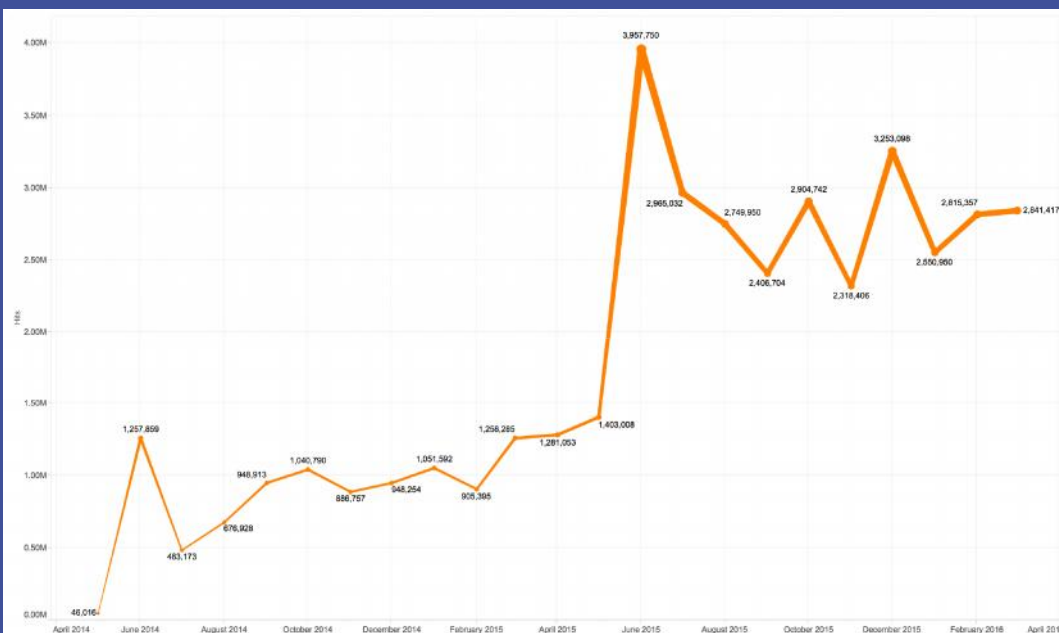
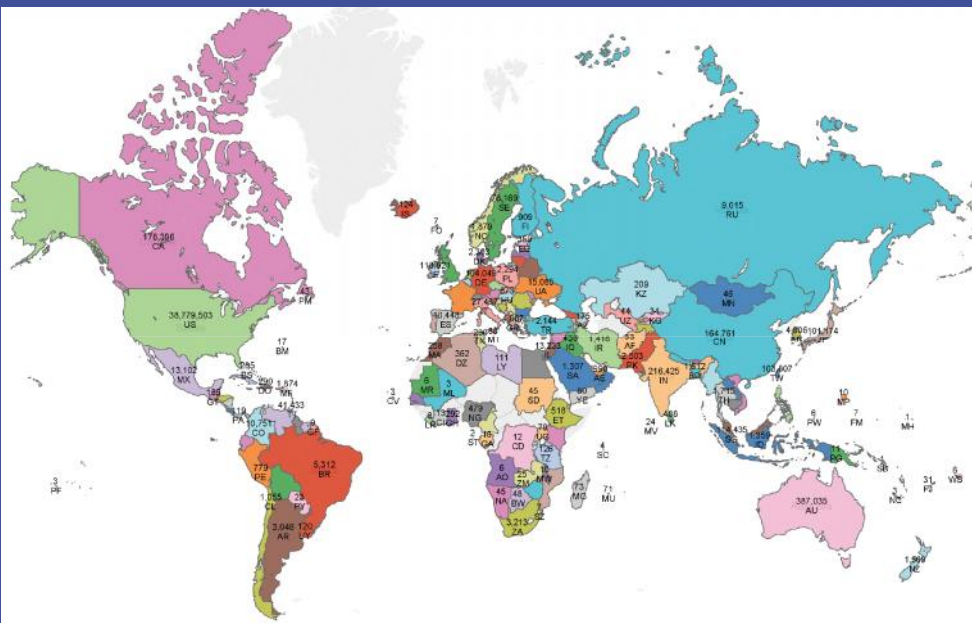
04 Jun 2014

I am really impressed with the quality of API deployments coming out of the federal government recently. I wrote about the **FBOpen API** from 18F a couple months ago, and the latest is the **OpenFDA API from the Food & Drug Administration**. I've been watching the rollout of the API from behind the scenes for a while now, but with all my travel and speaking I haven't had time to write about or participate, but now that they've officially launch publicly, I wanted to help showcase what they've been up to at the FDA.



Community

>7,600 registered API users



41,924,767

hits

1,140

unique users

236,355

unique ip addresses

105ms

average response time

<http://open.fda.gov> | open@fda.hhs.gov | [@openfda](https://github.com/fda) | github.com/fda

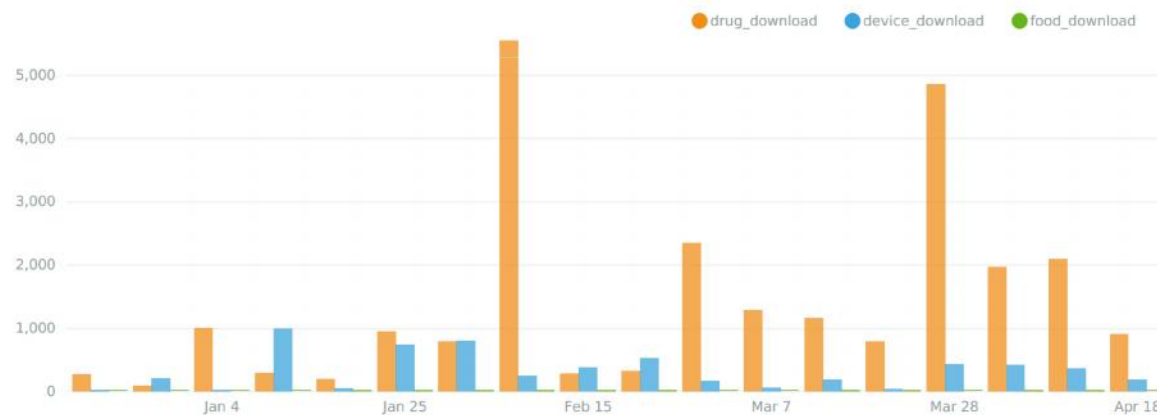
openFDA
open.fda.gov

Total File Downloads By Center Results

center	total
drug	24711
device	5848
food	241
data	159

>30K Downloads in 3.5 Months

Downloads by Center and Week



Sample App

EhrScape @EhrScape · 13h

Preview - @EhrScape #OpenEHR
ePrescribing demo - using @NLM_news
#rxTerm, @medlineplus & @openFDA API.

Form: Oral Tablet Strength: 500 mg Route: Oral Pill

Frequency: 3 /d

Therapy duration: 08-10-2014 until canceled days Stop Date ☒ PRN

Comment:

Medications

Medicine: Aspirin (Oral Pill)

+ Info

Dose: Aspirin 500 MG Oral Tablet

2

Form: Oral Tablet Strength: 500 mg Route: Oral Pill

Frequency: 3 /d

Therapy duration: 08-10-2014 until canceled days Stop Date ☒ PRN

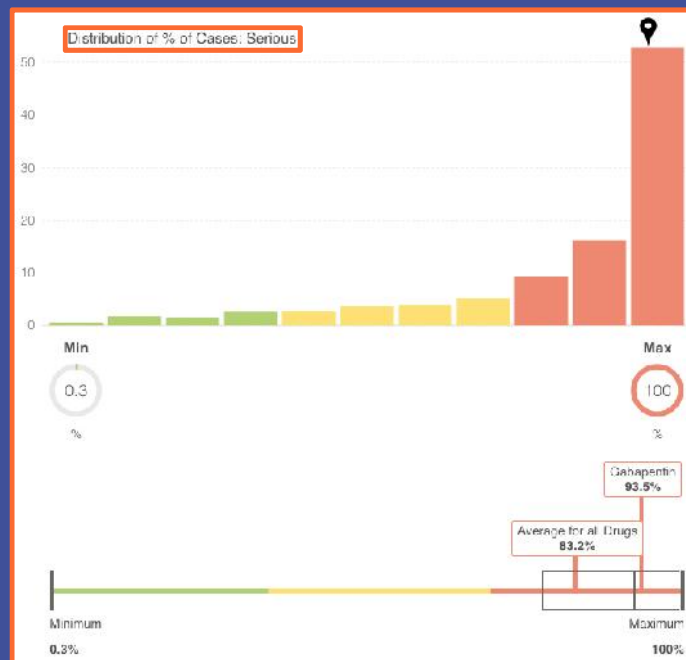
Comment:


+ ADD MEDICATION

History

Medicine	Dose	Frequency	Start Date	Stop Date	Route
Akineton 2 mg tbl.	3 tbl	1 /wk	26-09-2014	03-10-2014	PO - Oral
LEKADOL 120 mg/5 ml si:up	3 tbl	1 /d	26-09-2014	29-09-2014	PO - Oral
ABILIFY 15 mg tbl.	2 tbl	1 /d	26-09-2014	28-09-2014	PO - Oral

Sample App





Gabapentin

★★★★★ Be the first to review

Generic Drug: Gabapentin (- gab-a-PEN-tin)

Brands: Neurontin, Gralise, Gabarone, FusePaq Fanatrex,

Treats: Postherpetic Neuralgia

5 Most Common Adverse Effects for Gabapentin:

Drug Ineffective, Suicide Attempt, Suicidal Ideation, Completed Suicide, Pain

# of Cases	% of Cases: Primary Drug	% of Cases: Serious
27,747	37.1%	93.5%

Gabapentin Fair Price

The fair price for this drug is:

\$14

Fair Price provides a single price for the most-prescribed form, dosage and quantity of a specific drug. (Powered by GoodRx)

- Form: capsule
- Dosage: 300mg
- Quantity: 90

PER PRESCRIPTION

Research/Clinical Trials

The following research reports and clinical studies cover the latest findings on Gabapentin.

Title	Journal	Date
Efficacy of gabapentin in migraine prophylaxis	Headache	2001 Feb
Pharmacologic management part 1: better-studied neurologic pain disorders	Pain Med	2004 Mar
Survey of management of acquired nystagmus in the United Kingdom	Evo	2007 Sep
Gabapentin in bipolar disorder: a placebo-controlled trial of adjunctive therapy	Gabapentin Bipolar Disorder Study Group	Bipolar Disord

Data Sources

Data from Truven, Drugbank, DailyMED, openFDA, Medicare Part D (2011), Open Payments (2013) and other government/private data sources.

Connecting with Data Scientists and Researchers

```

data = data[patient_count>0 && patient_count<100000 && drug != '']
data.columns = ['patient', 'patient_count', 'drug']

data_count = {'patient': patient_count, 'drug': drug}

data_count['patient_count'] = 1
data_count['drug'] = 1

if (is.null(df)) { return(data.frame()) }

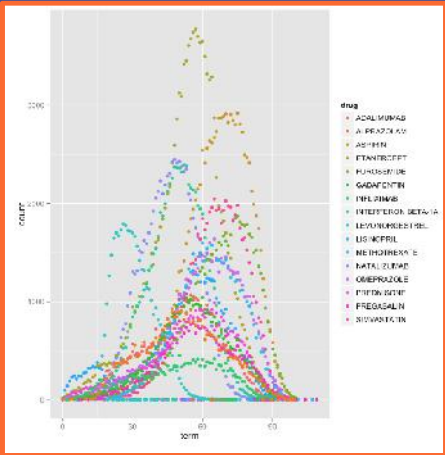
df[df$drug == drug]

##

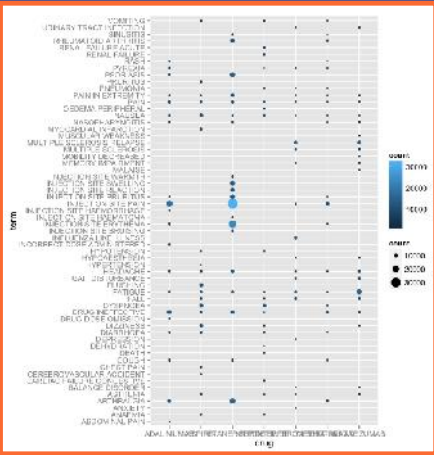
}

plot(data$ages, xmain, ymain, color=drug)

```



Compute cross correlations of drug reports vs meddra categories

[illegible]

Download food reports as a CSV

```

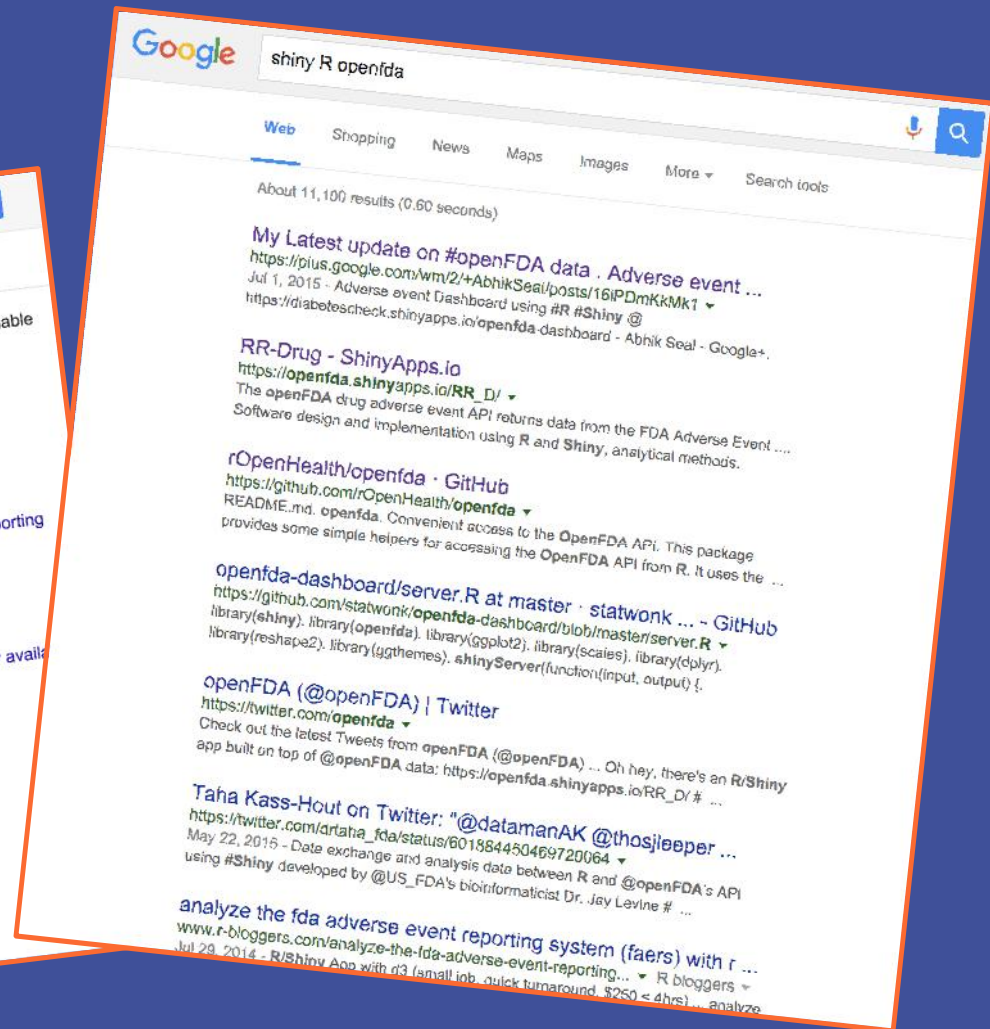
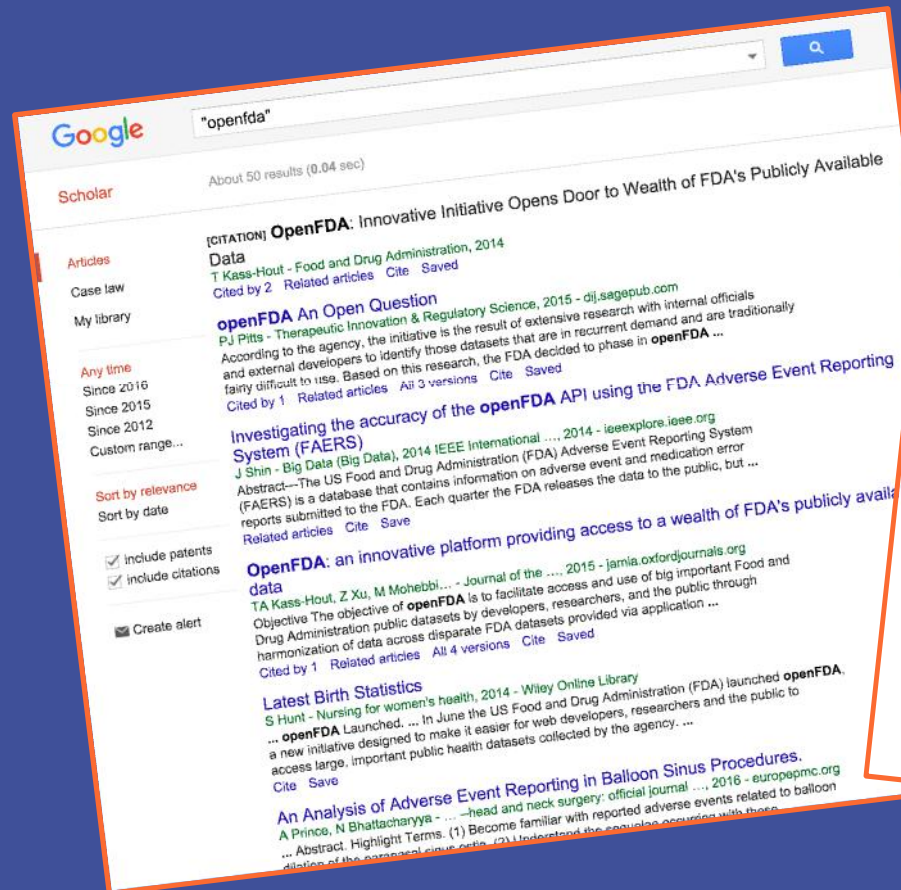
trainData = function() {
  load("wsg0, 8000, 100", function(wslap) {
    df = jsoemlites::fromJSON(paste0(
      "https://api.rda.gov/training/combined_trainData", 100,
      "-slap-", wslap, ".json"))
    df = df[,1:10]
    obj$trainData <- df
  })
}

```

An open source R library
for interfacing directly with
openFDA APIs has been
adopted by the
rOpenHealth project

github.com/ropenhealth/openfda

Sample Analytics & Peer-Reviewed Publications



<http://open.fda.gov> | open@fda.hhs.gov | [@openfda](https://twitter.com/openfda) | github.com/fda



Thank you!

**LETTUCE**

Canada, Chile, Dominican Republic, Mexico, Peru, USA

**CUCUMBERS**

Canada, Honduras, India, Mexico, Spain, USA

**FETA CHEESE**

Canada, Denmark, Egypt, Germany, Greece, Israel, Italy, Turkey, UK, USA

**VINAIGRETTE**

Argentina, Brazil, Canada, Chile, China, France, Germany, Greece, India, Indonesia, Italy, Mexico, Morocco, Peru, Portugal, Spain, Thailand, Tunisia, Turkey, USA, Vietnam

**OLIVES**

Greece, Israel, Mexico, Spain, USA

**SPROUTS**

Argentina, Australia, Bangladesh, Canada, China, Egypt, France, India, Morocco, Nepal, Pakistan, South Africa, Spain, Turkey, USA

**CROUTONS**

Argentina, Australia, Brazil, Canada, China, France, India, Mexico, Netherlands, Poland, Russia, Switzerland, Uruguay, USA, Vietnam

**TOMATOES**

Canada, Dominican Republic, Holland, Israel, Italy, Mexico, USA

**ONIONS**

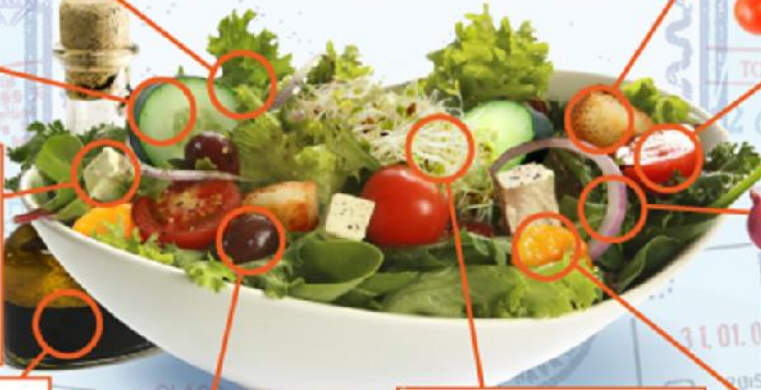
Canada, China, Germany, India, USA

**MANDARIN ORANGES**

Israel, Mexico, Morocco, South Africa, Spain

The Well-Traveled Salad.**Do You Know Where Your Food Has Been?**

As consumers, many of us fail to recognize that even our domestic and local food supplies are part of a global network. The daily activity of consuming food directly links our health as humans to the health of crops and produce, food animals, and the environments in which they are produced.



A "One Health" approach to food safety—bringing together expertise and resources from the clinical, veterinary, wildlife health, and ecology communities—has the potential to reveal the sources, pathways, and factors driving the outbreaks of foodborne illness and possibly prevent them from occurring in the first place.

NOTE: Countries are listed in alphabetical order and not by volume of export.



INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

www.iom.edu



Food Safety News

Breaking news for everyone's consumption

Home Foodborne Illness Outbreaks Food Recalls Food Politics Events Subscribe

Jury Verdicts: Guilty, Guilty, and Guilty in PCA Criminal Trial

BY DAN FLYNN | SEPTEMBER 19, 2014

Former Peanut Corporation of America owner Stewart Parnell, his brother and one-time peanut broker, Michael Parnell, and Mary Wilkerson, former quality control manager at the company's Blakely, GA, plant, were all found guilty today by a federal jury in Albany, GA.

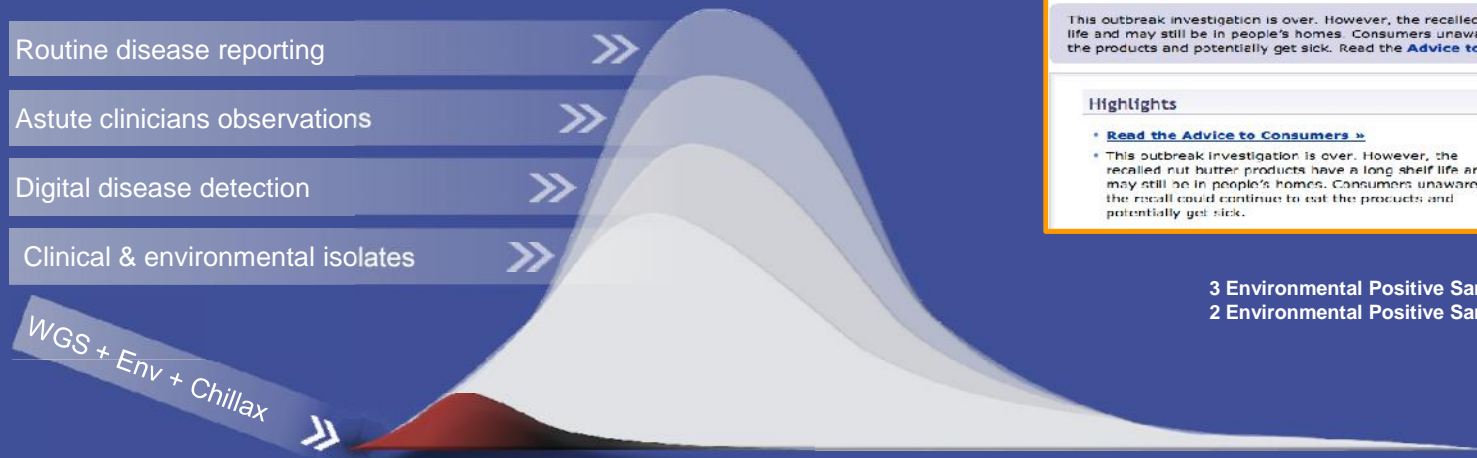
Sentencing will come later. Announcement of the jury verdicts brought an emotional outburst from the two Parnell families, while Stewart Parnell, age 60, simply put his head down.



Parnell, former chief executive of the now-defunct company with plants in three states producing peanut and peanut paste used for its own products and as an ingredient in almost 4,000 others, was convicted by a 12-member jury for his role in a deadly Salmonella outbreak that began almost six years ago.

The government accused Parnell, his brother, and the quality control manager of a mammoth conspiracy involving fraud, wire fraud, obstruction of justice, and knowingly introducing both adulterated and misbranded products.





Food enters commerce

Source of contamination
identified from few
cases



**6 Clinical isolates
(5 States)**

Multistate Outbreak of *Salmonella* Braenderup Infections Linked to Nut Butter Manufactured by nSpired Natural Foods, Inc. (Final Update)

Posted October 16, 2014 4:30 PM ET

This outbreak investigation is over. However, the recalled nut butter products have a long shelf life and may still be in people's homes. Consumers unaware of the recall could continue to eat the products and potentially get sick. Read the [Advice to Consumers](#).

Highlights

- [Read the Advice to Consumers »](#)
- This outbreak investigation is over. However, the recalled nut butter products have a long shelf life and may still be in people's homes. Consumers unaware of the recall could continue to eat the products and potentially get sick.

At a Glance:

- Case Count: [5](#)
- States: [5](#)
- Deaths: [0](#)
- Hospitalizations: [1](#)
- Recall: [Yes](#)

3 Environmental Positive Samples by FDA in 2/2014
2 Environmental Positive Samples by FDA in 7/2014

