Put VA Data to Work for Veterans, ORD-22-D4V

Director: Scott L. DuVall, PhD
Precision Medicine Director: Julie A. Lynch, Phd, RN, MBA, FAAN
Associate Directors: Michael E. Matheny, MD, MS, MPH; B. Kevin Malohi, BS;
Jeffrey P. Ferraro, PhD, MSc; Steven M. Johnson, DBA, MSM, MS, MA
Administrative Officer: Della A. Richter, BS

Mission, Vision, and Strategic Goals

VINCI Mission: Reduce the time and effort required to appropriately access, properly understand, and effectively use Veteran data for research.

VINCI Vision: Become VA researchers' preferred data source and data-processing environment.

	1. Put VA Data to work (D4V) to improve the lives of Veterans by **curating and linking the world's most comprehensive healthcare data to maximize its potential** for clinical decision-making and scientific discovery.	✅
	2. Develop infrastructure and tools, to include but not limited, to the infrastructure and tool initiatives (e.g. VINCI, VA Data Commons, AWS-VAEC-Prospect.)	🚧
	3. Participate in cooperative research to include but not limited to, set an example in accelerating scientific discovery by contributing data, analysis, clinical, informatics, and other subject matter expertise in research projects that provide impact to Veterans and help improve the infrastructure, methods, data, and tools. Both leading and working collaboratively in supporting projects done by other VA, affiliate university, other federal agency, non-profit, and industry colleagues.	🚧
	4. Champion VINCI as a community of excellence made up of skilled, diverse members by developing and coordinating core functions that support a professional and productive working environment (e.g. invest in individuals, optimize processes, maintain a strategic mindset, adequate resource functions, etc)	✅

Currently in the process of developing strategic goals #2 & #3 that will become VINCI's long-term ambitions (≥5 years), which consolidates sets of activities in a coordinated line of effort (LOE), to improve the situation through forward-thinking, transformational outcomes.

Choose VA

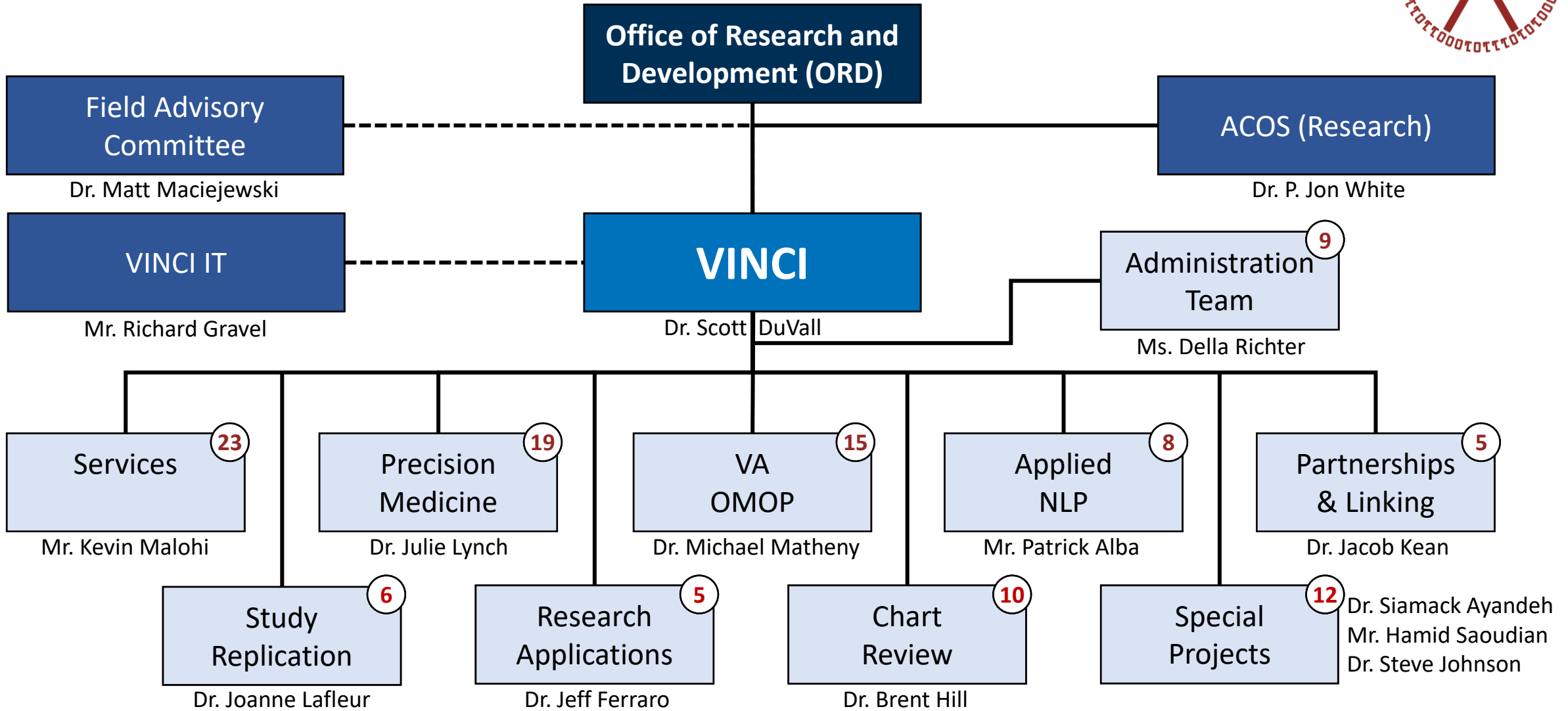VA | **U.S. Department of Veterans Affairs**

2

Partronerships and Customer Base

- To support this mission, VINCI partners with the VA Office of Information and Technology (OI&T) to deliver a full-service analytic environment (e.g. infrastructure) with access to data, tools, and support services.

- These services are available to any VAMC where thousands of researchers and projects depend on VINCI for their success to accelerate scientific discovery.

- VINCI provides additional services to its local VA medical center, which includes guidance to the local Research Review Committee (RRC), R&D Committee, University of Utah IRB, and partnership with CSP NODES-SLC.

- VINCI also partners with our customers to provide outcomes together that cannot otherwise be achieved individually:

- ORD	- Point-of-Care	- OEHRM	- Oncology
- HSRD	- NDS	- ONI	- Office of Mental Health
- OIT	- RAPID	- LGBTQ+	- PERC
- Field Investigators	- CSDE	- GRECC	- Connected Care
- CSP	- CTO	- CMS	- Cerner
- NODES	- CDO	- DoD	- VSSC
- MVP	- Lab and Pathology	- Palantir IPO	- VIReC*

Organizational Structure (by Team)

- - - - - - Coordination Line
————— Operational Line
(#) # of Team Members

Office of Research and Development (ORD)

Field Advisory Committee
Dr. Matt Maciejewski

ACOS (Research)
Dr. P. Jon White

VINCI IT
Mr. Richard Gravel

VINCI
Dr. Scott DuVall

Administration Team (9)
Ms. Della Richter

Services (23)
Mr. Kevin Malohi

Precision Medicine (19)
Dr. Julie Lynch

VA OMOP (15)
Dr. Michael Matheny

Applied NLP (8)
Mr. Patrick Alba

Partnerships & Linking (5)
Dr. Jacob Kean

Study Replication (6)
Dr. Joanne Lafleur

Research Applications (5)
Dr. Jeff Ferraro

Chart Review (10)
Dr. Brent Hill

Special Projects (12)
Dr. Siamack Ayandeh
Mr. Hamid Saoudian
Dr. Steve Johnson

The personnel who support these product and service lines are mostly funded by VA and are either VA employees or WOCs administered by academic affiliates or non-profit corporations. See demographics slide for breakdowns.

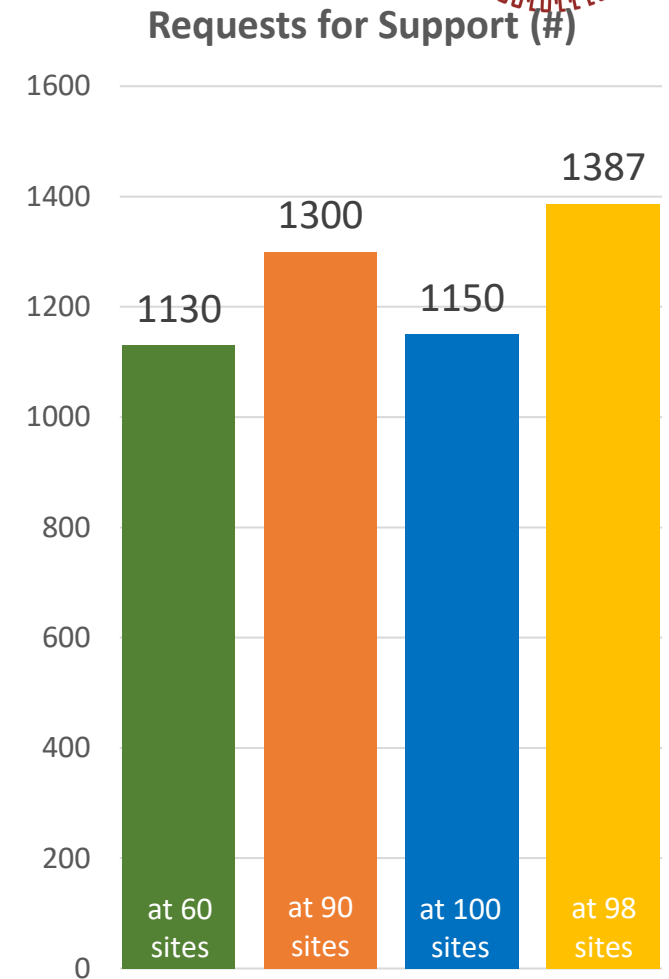U.S. Department of Veterans Affairs

4

VINCI Services

- Primary focus is comprised of education and outreach, feasibility studies, data needs assessments, clinical trial recruiting support, data provisioning, cohort development, and other data support.

- Works in accordance with published CDW and VINCI best practices to obtain, clean, map, provision, analyze, and quality check VA data (e.g. code and quality review) to ensure high quality and customer satisfaction.

- Supports all aspects of the research project lifecycle by facilitating access to and providing expertise in VINCI hosted data that increases the efficient and effective use of VA data by investigators.

- While not engaging in research, VINCI Services becomes an extension of the research team supporting all aspects of their project lifecycle to include preparatory to research activities, provisioning data, data transformations, and transfers that require elevated access privileges and that use best practices and optimizations to facilitate more effective use of the data.

Data Services

- In addition to its core services, continued to support VA programs such as MVP, CSP, NPOP, MAVERIC as well as the COVID-19 Shared Data Resource (CSDR).

- Provided data support including a variety of services particular to the creation of analytical files, data curation, data provisioning, data extractions, data transfers, data mappings, and documentation.

- These collaborations have led to the continued growth of the VA Phenomics Library as we contribute all our definitions to the Centralized Interactive Phenomics Resource (CIPHER).

Requests for Support (#)

	Value	Sites
FY20	1130	at 60 sites
FY21	1300	at 90 sites
FY22	1150	at 100 sites
FY23	1387	at 98 sites

Data Services: Data Needs Assessments

- Supports researchers' data questions regarding CDW and other VA data sources.

- Deliverables: Data domains, tables, columns, regulatory consults, & source documents
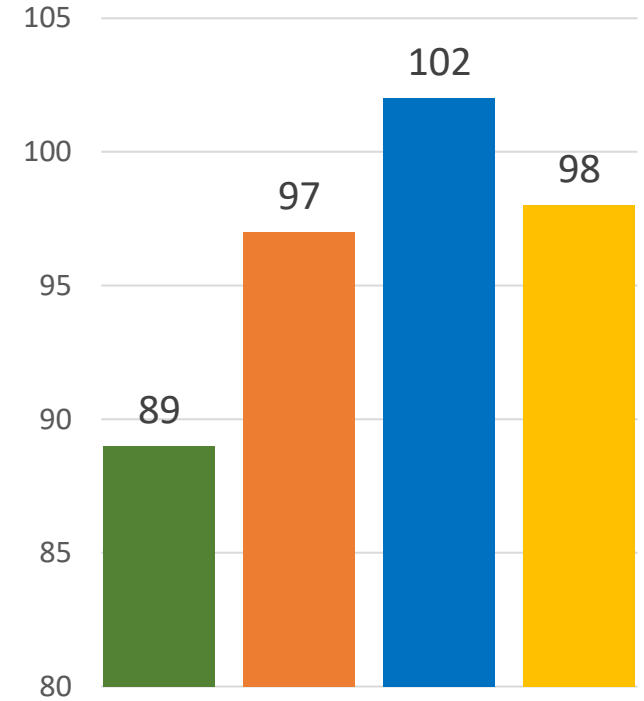
Data Requests (#)

FY	Value
FY20	341
FY21	452
FY22	662
FY23	716

Mean Turn-Around (days)

FY	Value
FY20	2
FY21	2
FY22	2.5
FY23	3.2

VA Stations Supported (#)

FY	Value
FY20	89
FY21	97
FY22	102
FY23	98

FY20 FY21 FY22 FY23
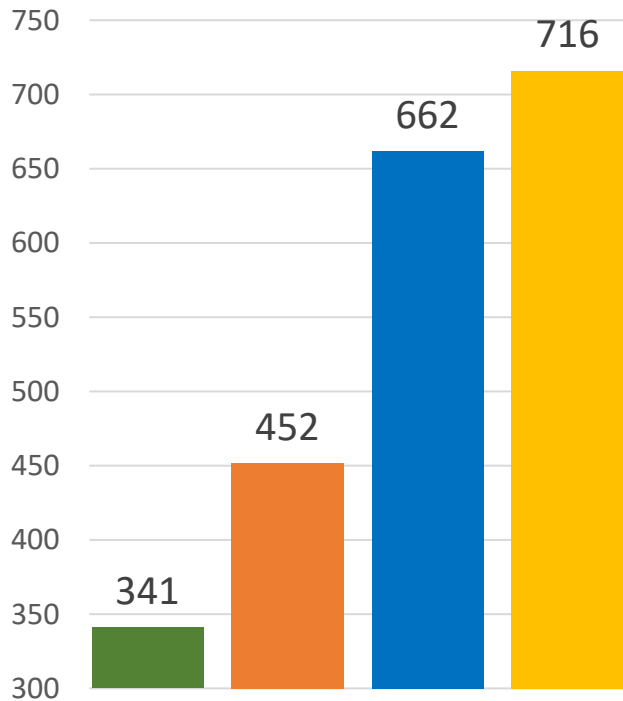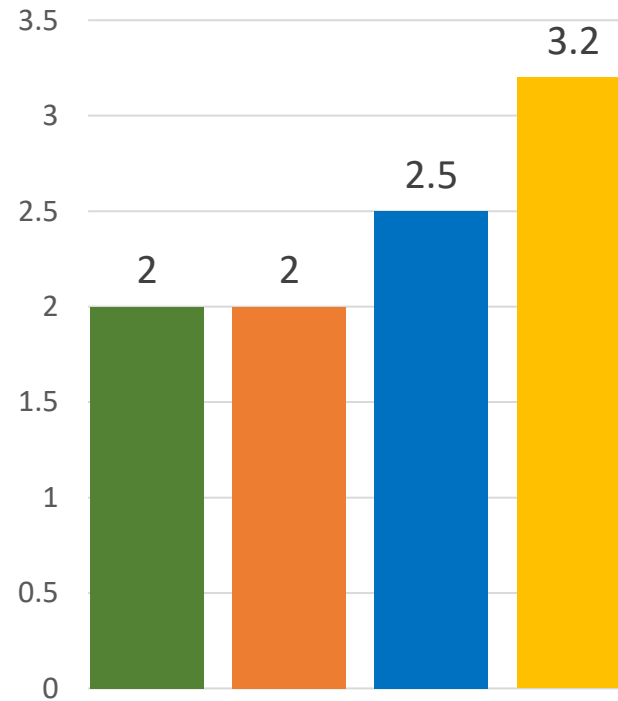
Data Services: Feasibility Requests

- Provides aggregate counts for researchers during the preparatory-to-research phase.

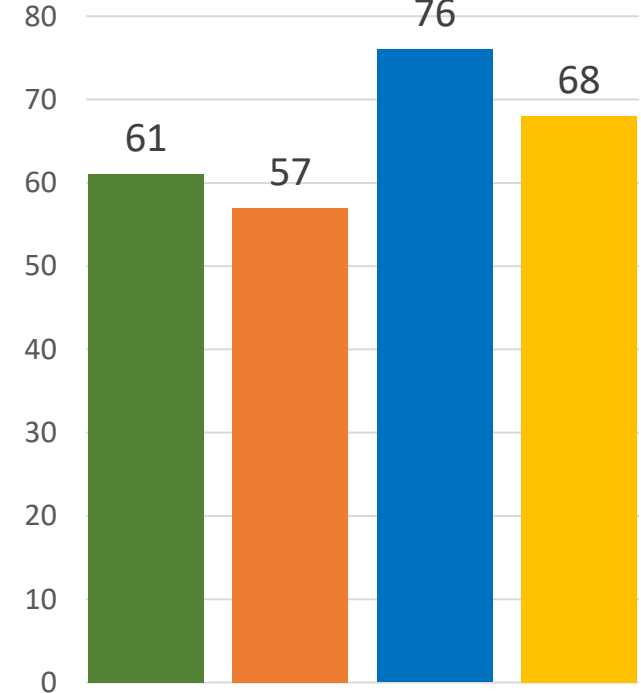- Deliverables: Funnel analysis, query definitions and demographics

Feasibility Studies (#)

FY20	FY21	FY22	FY23
364	288	274	254

Mean Turn-Around (days)

FY20	FY21	FY22	FY23
7	4	3.5	5.1

VA Stations Supported (#)

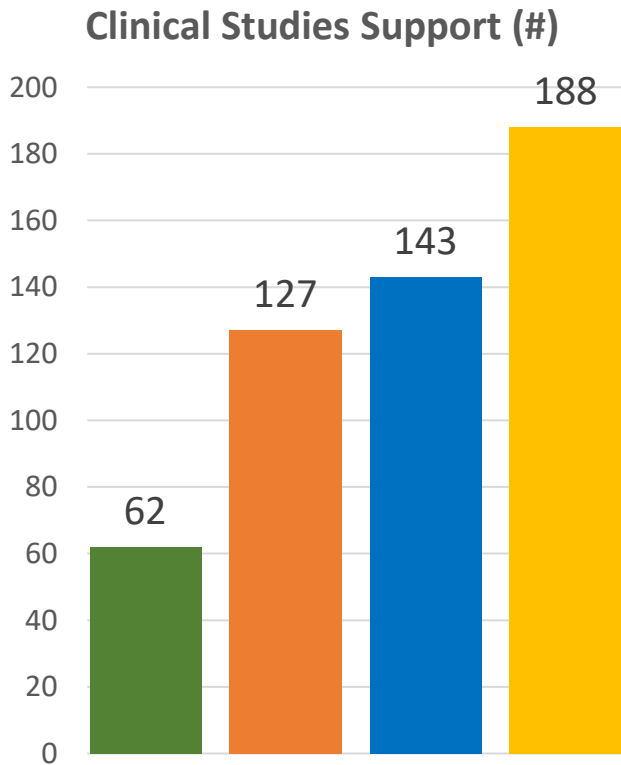| FY20	FY21	FY22	FY23
61	57	76	68
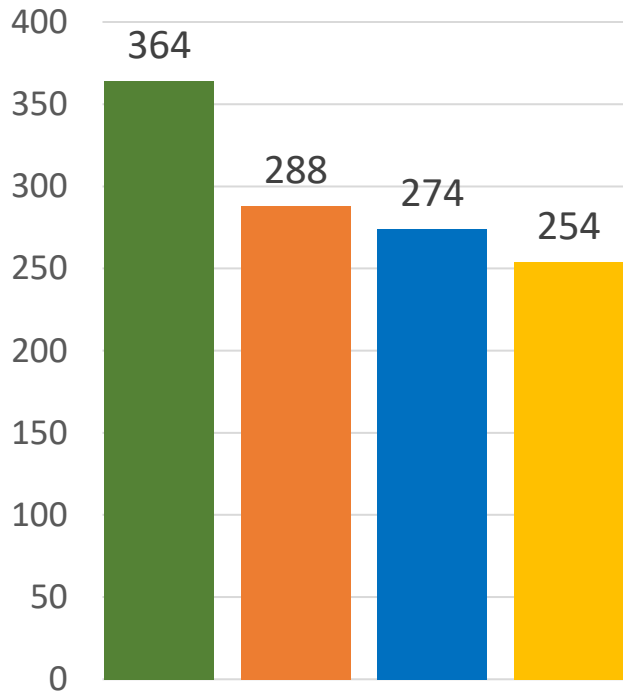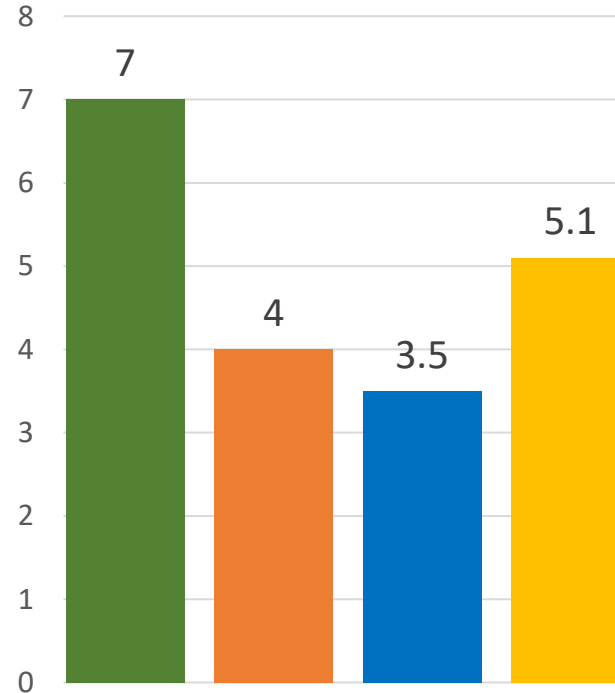
Data Services: Clinical Trial Support

- Translates study protocols into searchable items regarding how many patients correspond to inclusion and exclusion criteria.

- Deliverables: Patient lists with demographics, appointments, addresses, clinical conditions, & contact info

Clinical Studies Support (#)

Mean Turn-Around (days)

Patient Lists Delivered (#)
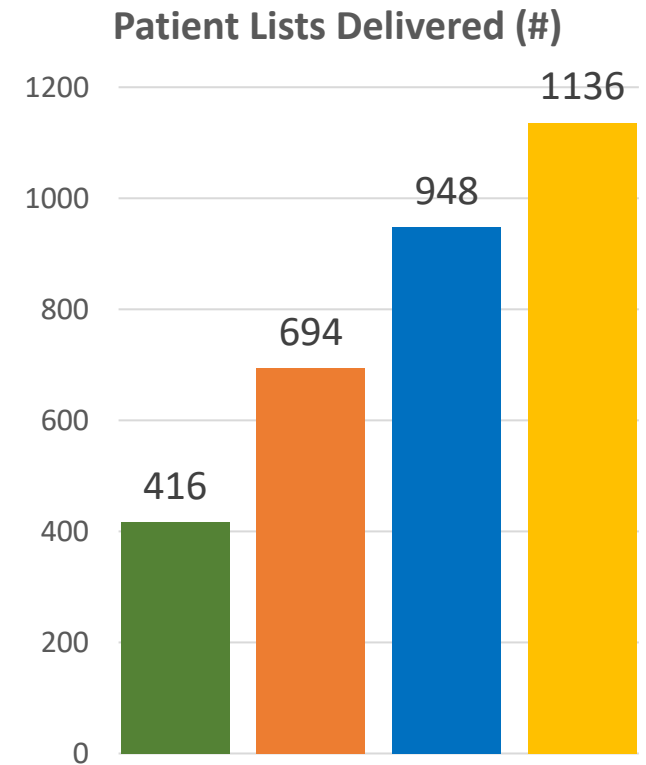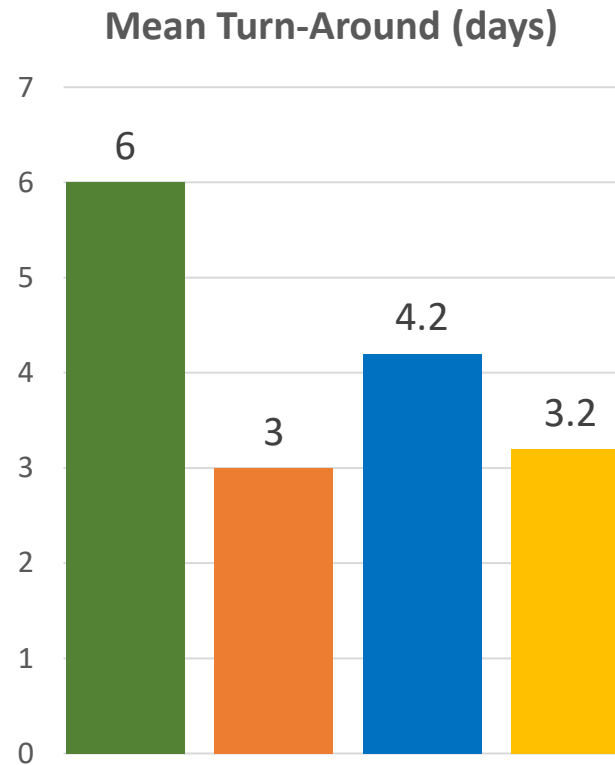
Data Services: Cohort Development

- Translates study protocols into searchable items regarding how many patients correspond to inclusion and exclusion criteria.

- Deliverables: cohort for the approved research project data base, attrition table, SQL script, cohort description
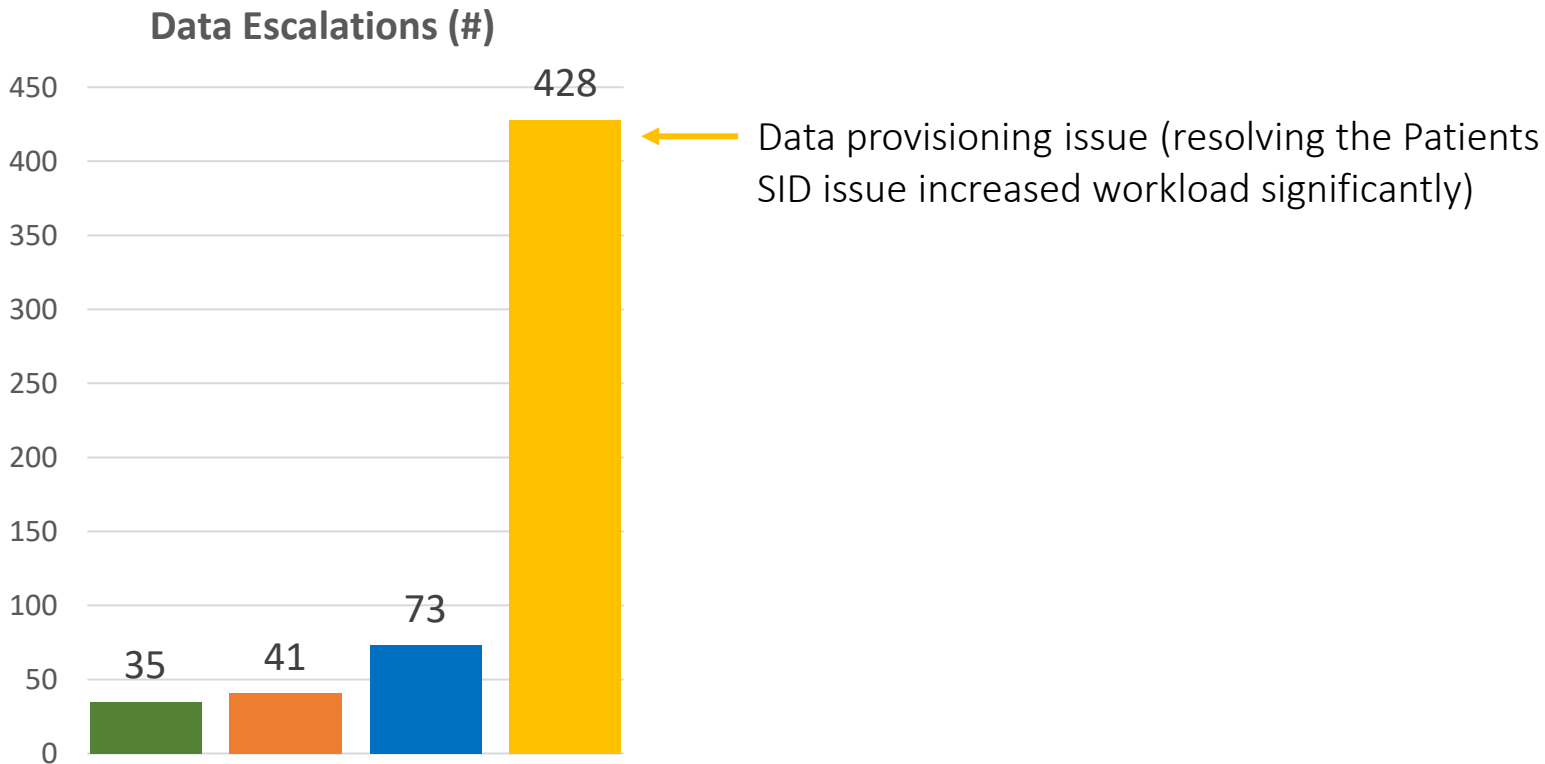
Cohort Development (#)
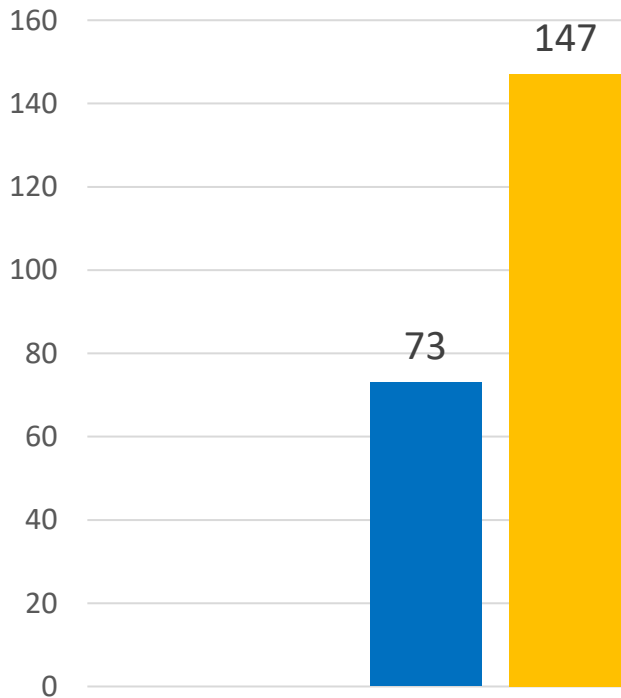
Mean Turn-Around (days)
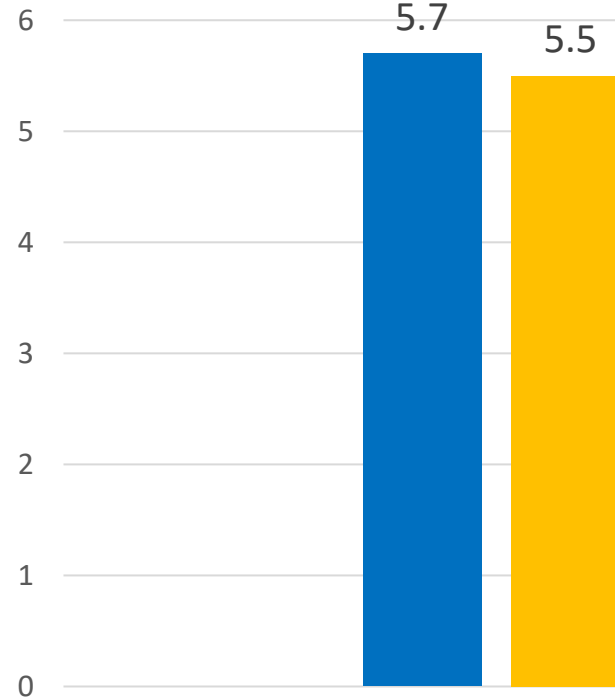
Data Services: Data Escalations

- Research projects requiring further data refinement beyond the scope of VINCI IT Data Managers.

- Deliverables: Analytical files, refined cohorts, SQL scripts, and data dictionaries; ETL data, data transferring, provisioning Prostate Cancer data core files; Millennium data

Data Escalations (#)

Data provisioning issue (resolving the Patients SID issue increased workload significantly)

428

73

35   41

Data Services: Quality Assurance Measures

- Developed and maintained SQL Standard Operating Procedures (SOPs)
- Conducted daily and weekly Code Reviews
- Conducted Code Reviews for all cohort development project including OIT data managers
- Managed version control code repository
- Implemented CDW Best Practices

SQL SOPs Developed and Maintained (#)

Reviewed Projects

Feasibilities #

Patient List #

Cohort Development #
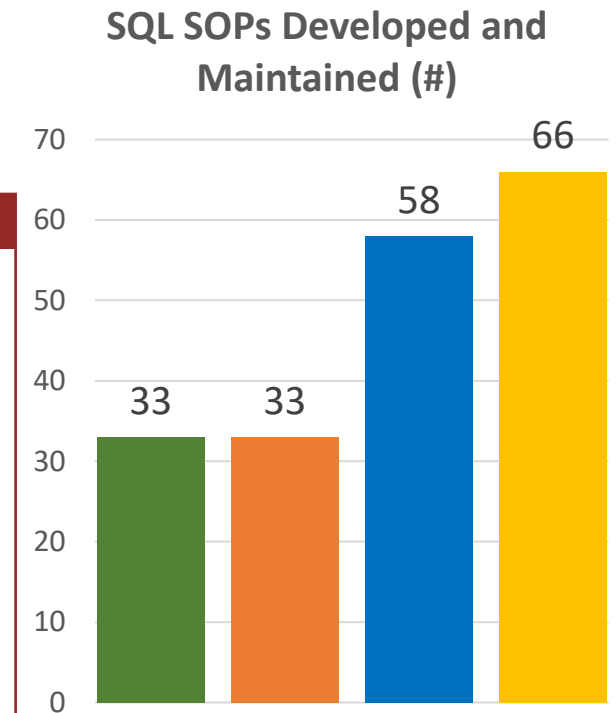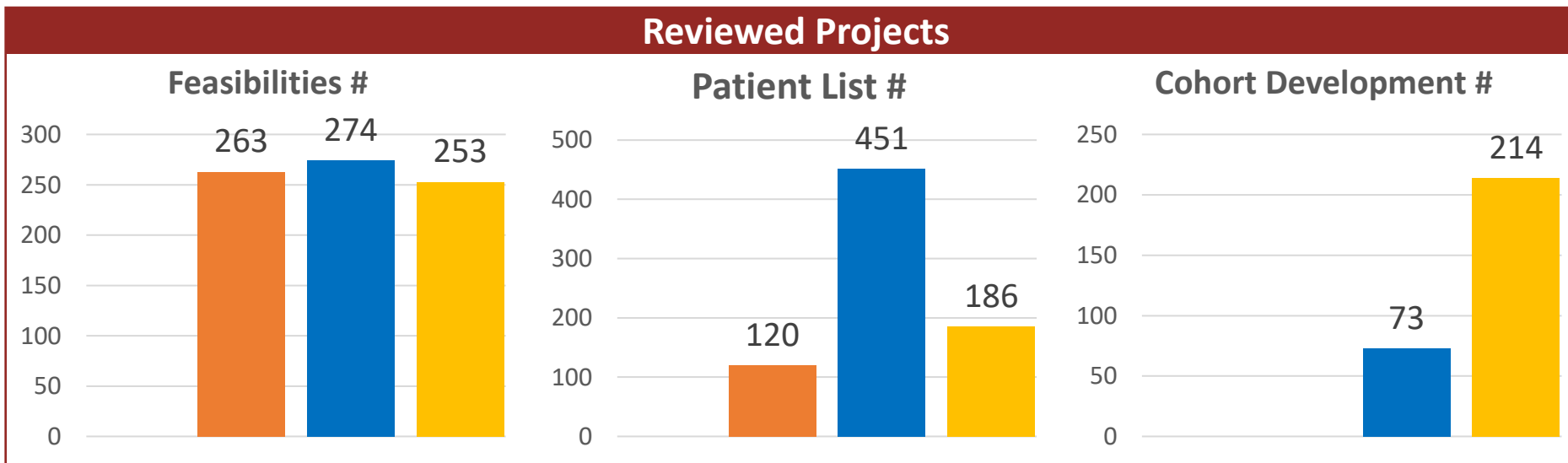
VINCI Services: Support to Data Provisioning Issue

- Supported the ORD/OIT joint task force that guided responses to a problem of OIT distributing incomplete datasets to several hundred studies, as first noted on August 5, 2022.

- **364 of 400 help tickets** had valid Patient SID Issue:
  - 312 Research Studies
  - 52 Prep to Research projects
  - **Reprovisioned Data to 188 studies** (were given an updated cohort based on PatiendSID issue).

- Post Surveys received:
  - **81 of the 188 Studies that were reprovisioned data replied to the survey**
  - **52 of those 81 Studies have rerun analyses**:
    - 26 No meaningful change
    - 4 Changes to at least one primary outcome or conclusions
    - 22 Some change to specific estimates, but no change in primary outcomes or conclusions
  - **29 of those 81 studies have not completed the rerun of their analyses**

- **VINCI Services effort to resolve**:
  - **700 hrs for meetings** with ORD/OIT leadership, data managers planning, and ORO support.
  - **750 hrs to determine scope** of impacted projects; **reviewed 1,138 SQL scripts**.
  - **250 hrs developing and maintaining PBI Dashboard**; categories include reported and impacted projects, survey data, Help ticket summary, PatientSiD analysis, and SQL script review.
  - **750 hrs in outreach & coordination** involving **1000+ email correspondence, consultations, and tracking**.
  - **200 hrs revising SQL scripts for 173 impacted projects**.
  - **20 hrs reviewing SQL scripts** for quality assurance during the data provisioning pilot.

- **Cost to resolve $187,228**

VINCI Services: Field Impact of Data Provisioning

- The VINCI Field Advisory volunteered to lead an analysis to create an estimate of how much **the field put effort in to resolving** the data provisioning issue first noted on August 5, 2022.

- **Total Cost for the Field:**
  - Lower bound: $2,085,530 + $493,587 = **$2,525,116 ($6,787.95 per project)**
  - Upper bound: $4,171,060 + $493,587 = **$4,610,646 ($12,394.21 per project)**

- Data analyst time:
  - Lower bound: $5606 * 372 projects = $2,085,530
    - 10% of GS-12/1 analyst for 6 months: $112,125 ($86,250 salary + $25,875 fringe) * 10% * 0.5 = $5606 where $71,099 is GS 12/1 GS base from opm.gov
  - Upper bound: $11,213 * 372 projects = $4,171,060
    - 20% of GS-12 analyst for 6 months: $112,125 * 20% * 0.5 = $11,213

- 1% of PI time (GS-14/1) for 9 months: $1,182 * 372 projects = $493,587
  - $157,558 ($121,198 salary + $36,360 fringe) * 1% * 0.75 = $1,182

  * Assuming an average locality adjustment across the country of 21.31% and 30% fringe benefits.

VINCI Services: Data Provisioning Process Pilot

- ORD and OIT collaborated in a series of meetings to update their current joint standard operating procedures (SOP) to improve accuracy and completeness of datasets for research.

- Resulted in testing a new pilot provisioning process (**processed 91 projects**)
  - The current pilot utilizes SQL SOPs developed by both ORD and OIT Data Managers.

  - Incorporates an attrition table in the process detailing the inclusion and exclusion criteria that allows researchers to see how each criteria impacts their cohort.

  - Includes Quality assurance (QA) in which a second ORD VINCI DM peer reviews each new project to accomplish the following:
    1. Manually audit SQL script for coding and logic errors.
    2. Determine the relevant SQL SOPs and data domains were applied appropriately.
    3. Determine SQL best practices were applied.
    4. Clarify definitions and recommended data sources if necessary.
    5. Check alignment between cohort criteria and script.
    6. Generate caveats and limitations for researchers.

Concierge and Help Desk

- First contact with almost all VINCI users.

- Primary focus is to simplify and standardize access to and use of VINCI resources by actively reaching out to VINCI users to educate and help set expectations.

- Answers all questions asked from research studies or finds the experts who can from the studies' conception to completion, which includes questions about data availability, data attributes, and the data request process.

- Provides support, education, and outreach related to VINCI, the VINCI Workspace, DART, data needs, project needs, study setup, training, and data access.
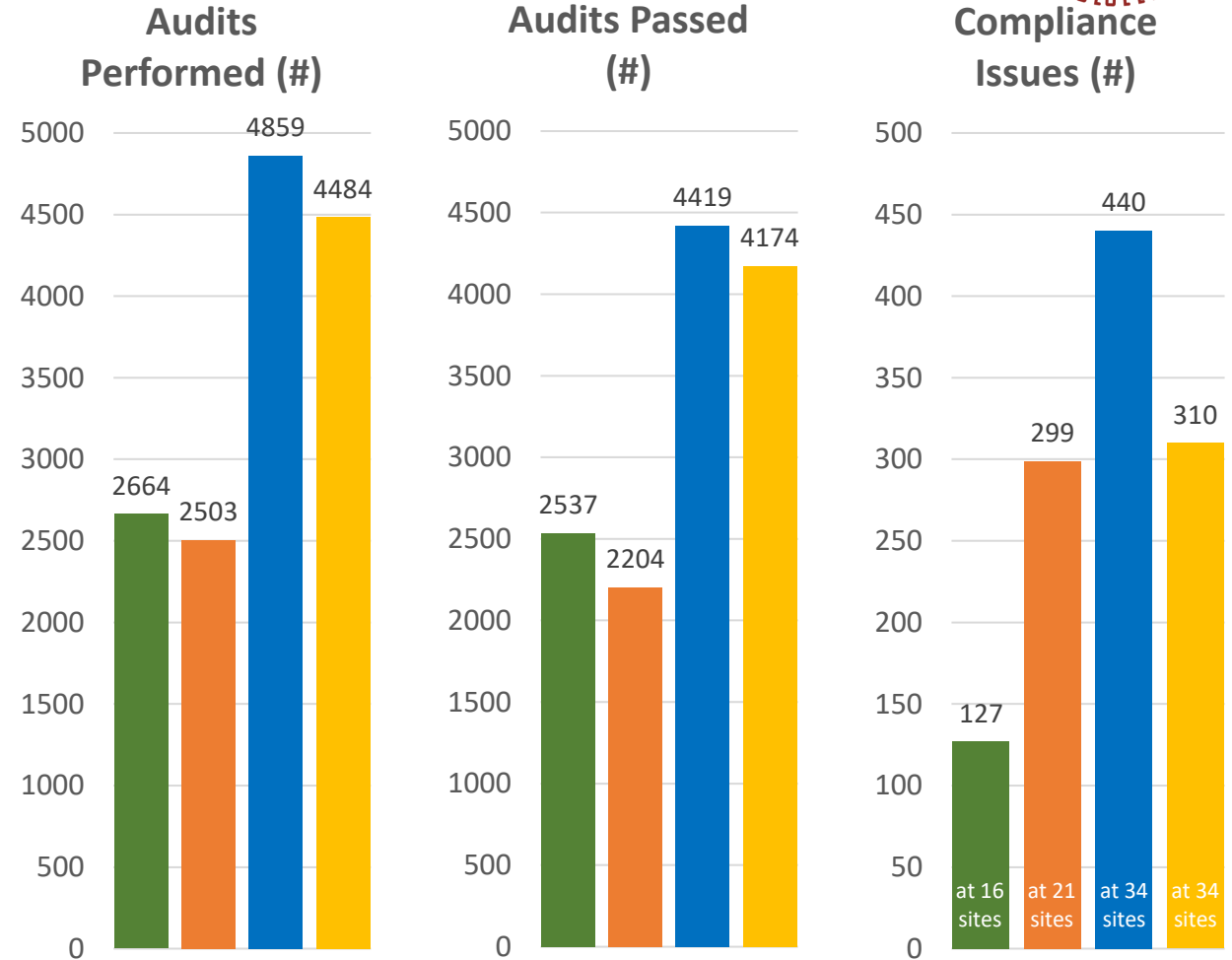
Tickets Resolved (#)

Value
5534
4603
2756
2972
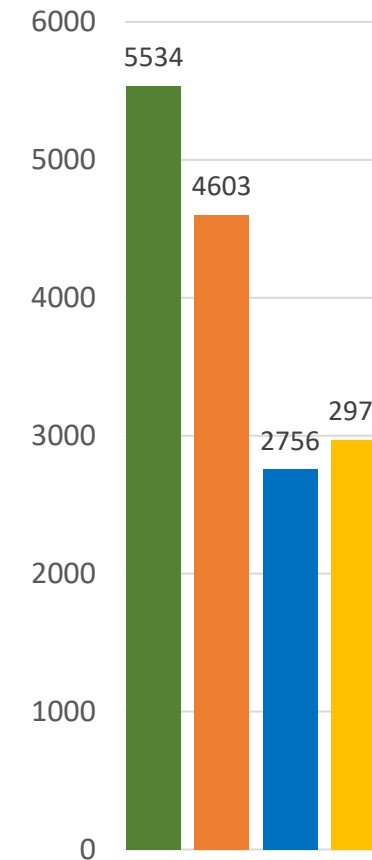
New Research Projects (#)

Value
627
480
448
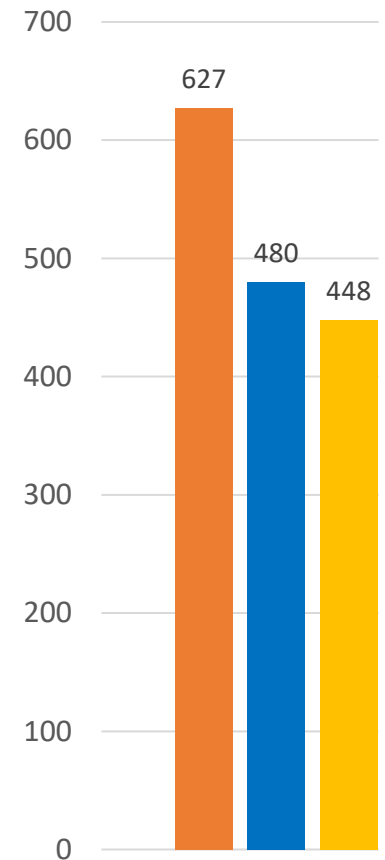
Audits

- The team conducted 4,484 audits with a compliance rate of 93% during the year.

- To prioritize, the team audits a sampling of all downloads using an algorithm that weighs particular variables heavier than others (e.g. new users, users with red flags, etc).

Audits Performed (#)

- FY20: 2664
- FY21: 2503
- FY22: 4859
- FY23: 4484

Audits Passed (#)

- FY20: 2537
- FY21: 2204
- FY22: 4419
- FY23: 4174

Compliance Issues (#)

- FY20: 127 at 16 sites
- FY21: 299 at 21 sites
- FY22: 440 at 34 sites
- FY23: 310 at 34 sites

FY20 FY21 FY22 FY23

Education and Outreach

- Offers numerous education and outreach activities dedicated to meeting all of researchers' VINCI training and educational needs.

- Provides videos on VINCI data, the VINCI Workspace, administrative processes, and technical procedures on our VINCI video site in addition to our extensive collection of user guides.

- Additionally, provides VA-wide training and information provided through Meet VINCI Initiatives, Listening Sessions, Cyberseminars, VINCI Training Hour, VINCI SQL and CDW Bootcamp, Q&A sessions, VINCI Office Hours, Workshops, Education and Outreach/Concierge Campaigns, and Educational Products made available through the VINCI Central Website.

- Also provides VINCI Updates on VHA Data User Calls

- Research team consultations (**332 in FY23**), Millennium consultations (**82 in FY23**)

Education and Outreach: VHA Data Users Call

- VINCI is a long-time participant in the VHA Data Users Call even though it focuses on Operations

- VINCI provides a monthly report of VINCI training, what VINCI has to offer, and matters that may impact operations users of VINCI

- This reporting takes place along side reporting from other VHA entities of: VIReC, ARC, Performance Measurement, MCAO, OPES, NDS, CSAR, Chief Strategy Office-Strategic Analysis Service, CDW, and VSSC

- This informs our counterparts as to VINCI offerings, ongoings, and services

VINCI Update to Participants (#)

| FY	Participants
FY20	2399
FY21	2658
FY22	2107
FY23	1001
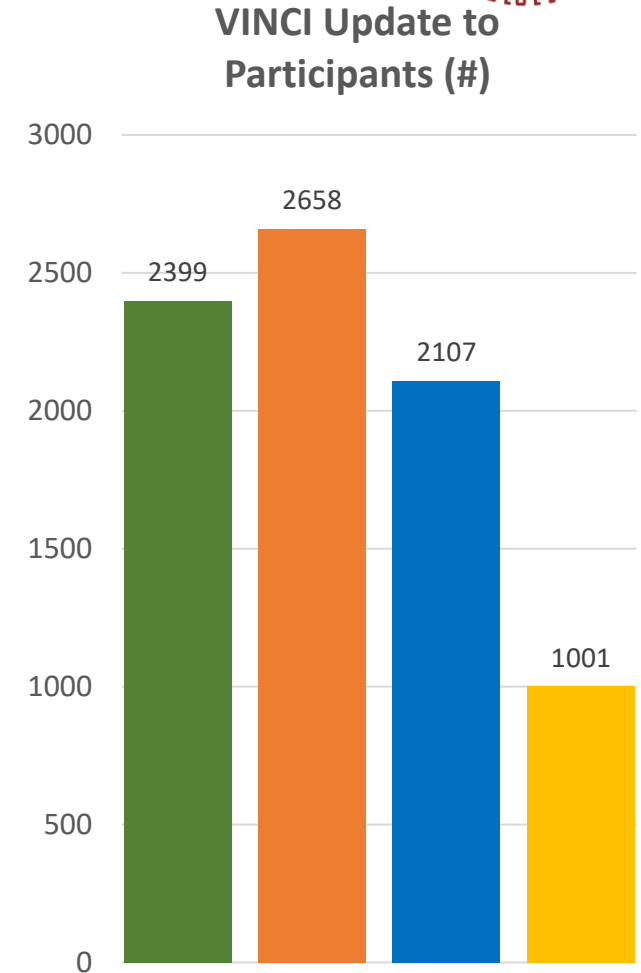
Education and Outreach: VINCI Cyberseminars

- In 2023, provided **8 sessions** on topics including:
  - A Practical Approach to Working With VA Community Care Data
  - VINCI Data Services: An Overview
  - Navigating CDW Data
  - Cerner Millennium Data for Researchers
  - What is VINCI and How Does It Help Researchers
  - Basic Research and Preparation with the SAS Grid
  - How to Use the VINCI Workspace
  - Managing Research Data

- *And collaborated with partners at VIReC on:
  - Database and Methods Seminar (2 sessions)
  - Good Data Practices Miniseries (3 sessions)

Total Participants (#)

FY20	FY21	FY22	FY23
1564	1497	1182	1090*
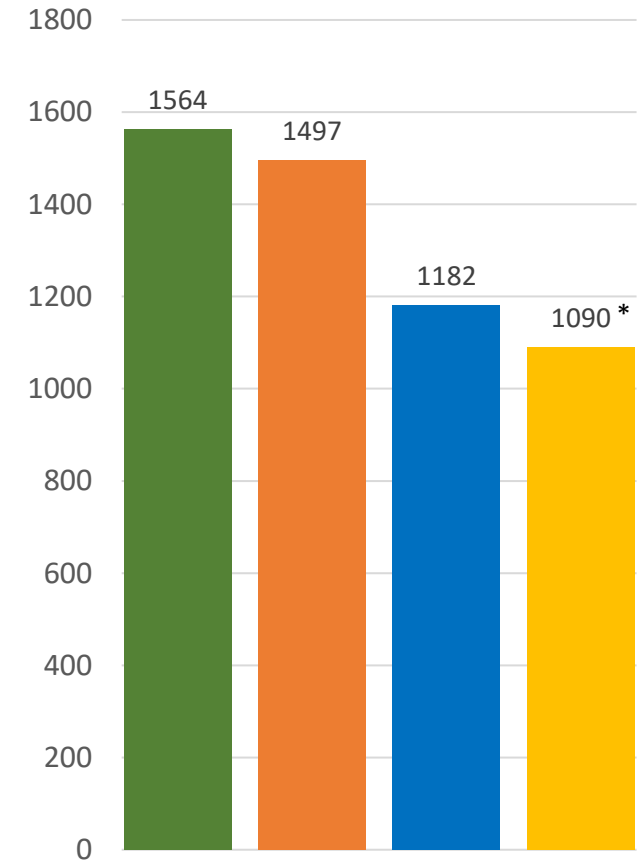
*Not counting collaboration

FY20 FY21 FY22 FY23

Education and Outreach: VINCI Training Hours

- In 2023, provided **11 sessions** on topics including:

  - An Overview of OMOP in the VA

  - LOINC Searching/Lab SOP Demos

  - Inpatient Pharmacy Domain (IV and Unit Dose orders)

  - Outpatient Pharmacy Domain (Prescriptions)

  - CPRS Orders and Pharmacy Orders and Data, including CPRSOrder, BCMA, RxIV, and RxUD

  - VINCI SAS – Efficient SAS Grid Utilization

  - Accessing Cerner-Millennium Data as a Researcher

  - New Features of the Centralized Interactive Phenomics Resource (CIPHER)

  - A Brief Introduction to a New Community Care Data Source: IVC_CDS

  - Batch Job Techniques with the SAS Grid
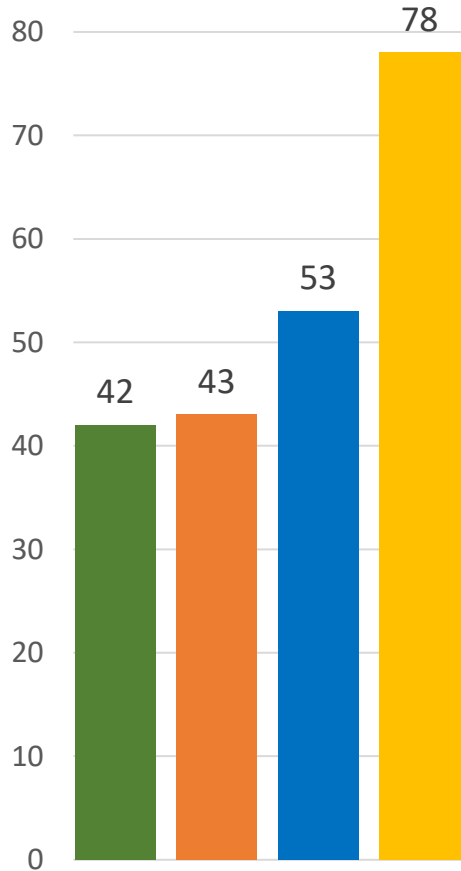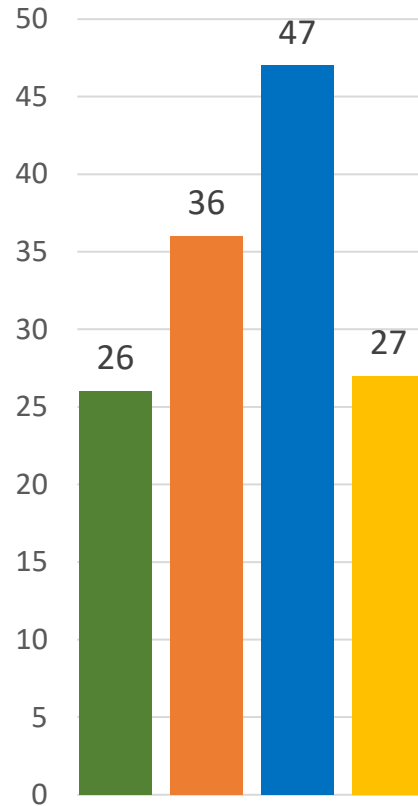
  - VINCI Compliance

Total Participants (#)

	Total Participants (#)
FY20	1454
FY21	1301
FY22	1468
FY23	1608
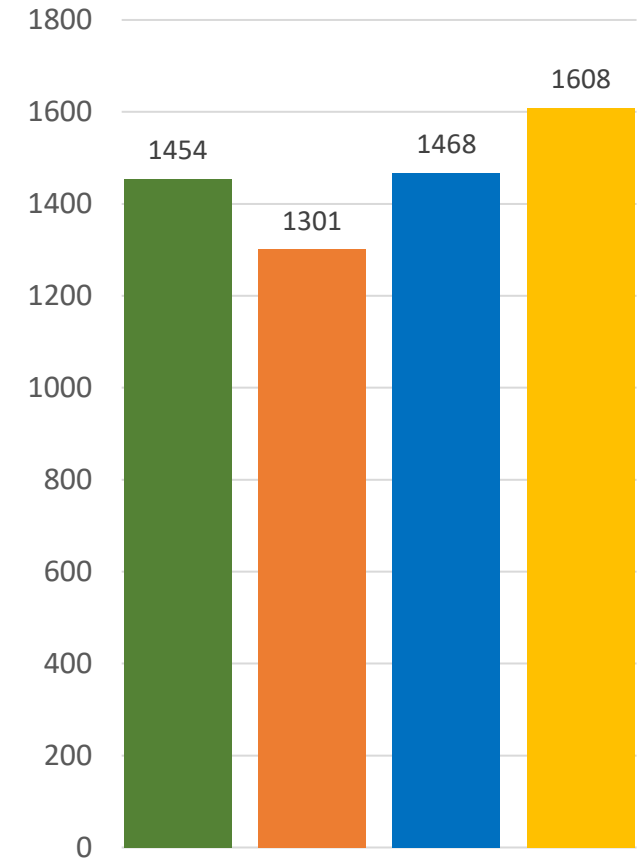
Education and Outreach: VINCI Materials Available
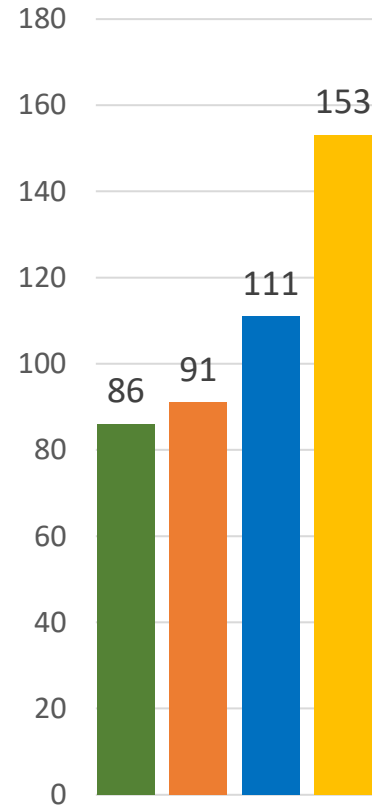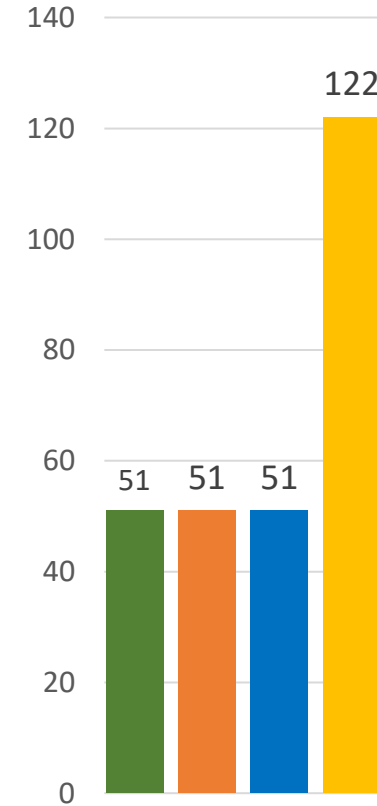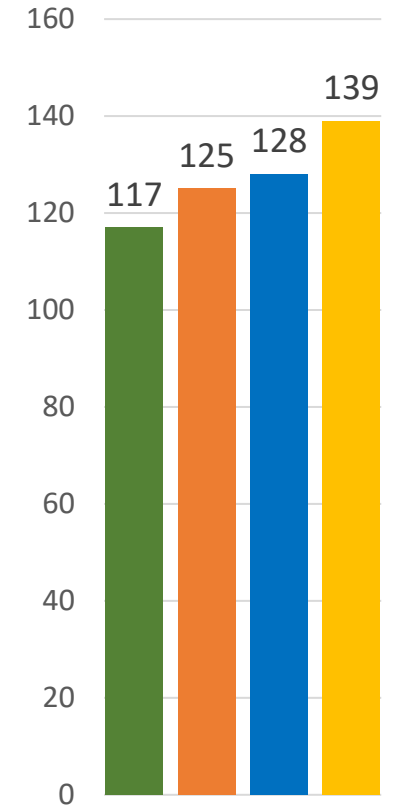
Education and Outreach: VINCI Academies

VA OMOP: Teaches users about the VA OMOP Common Data Model, the CDW data that it contains, and how to use it effectively.

29 Videos
w/ slides & 10 SQL Examples

DaVINCI: Familiarizes users with the Military Health System Data Repository data it contains and DaVINCI OMOP data.

33 Videos
w/ slides

NLP Academy: Introduces Natural Language Processing course and a Clinical NLP for Analysts and Developers course.
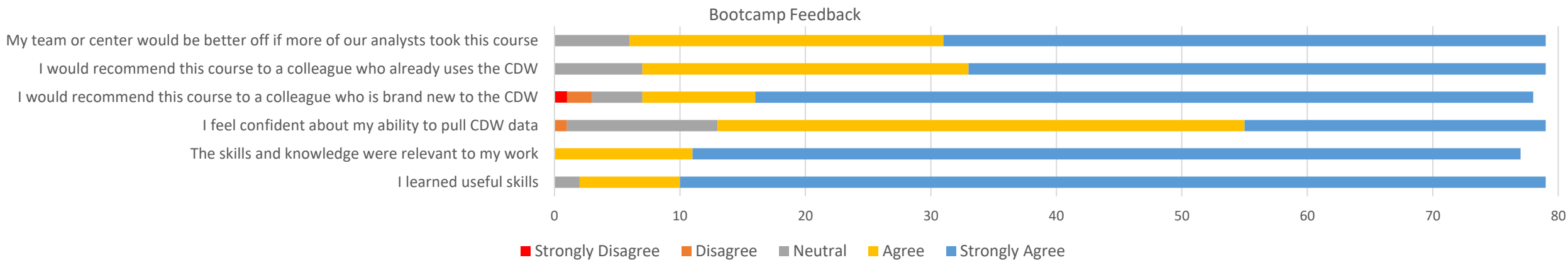
13 Videos
w/ slides

VINCI R: Familiarizes user with the R programming language and how to use it in the VINCI Workspace environment including its use with SAS.

5 Videos
w/ slides & 16 Training Modules
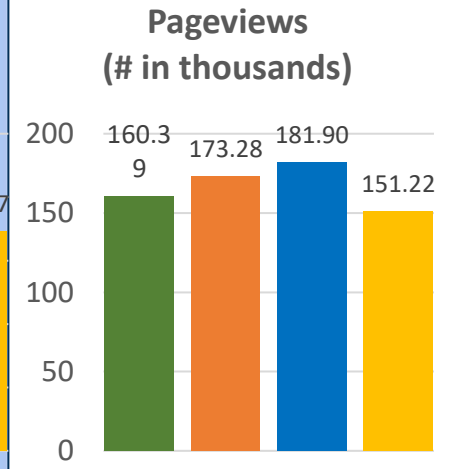
Choose VA

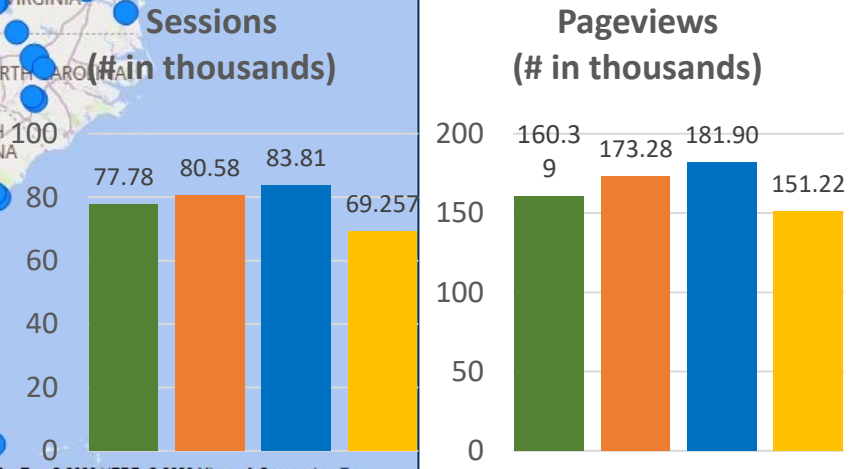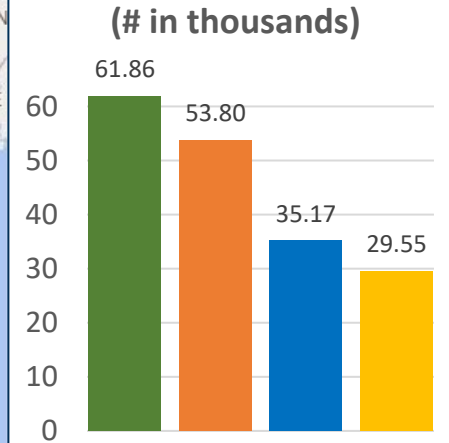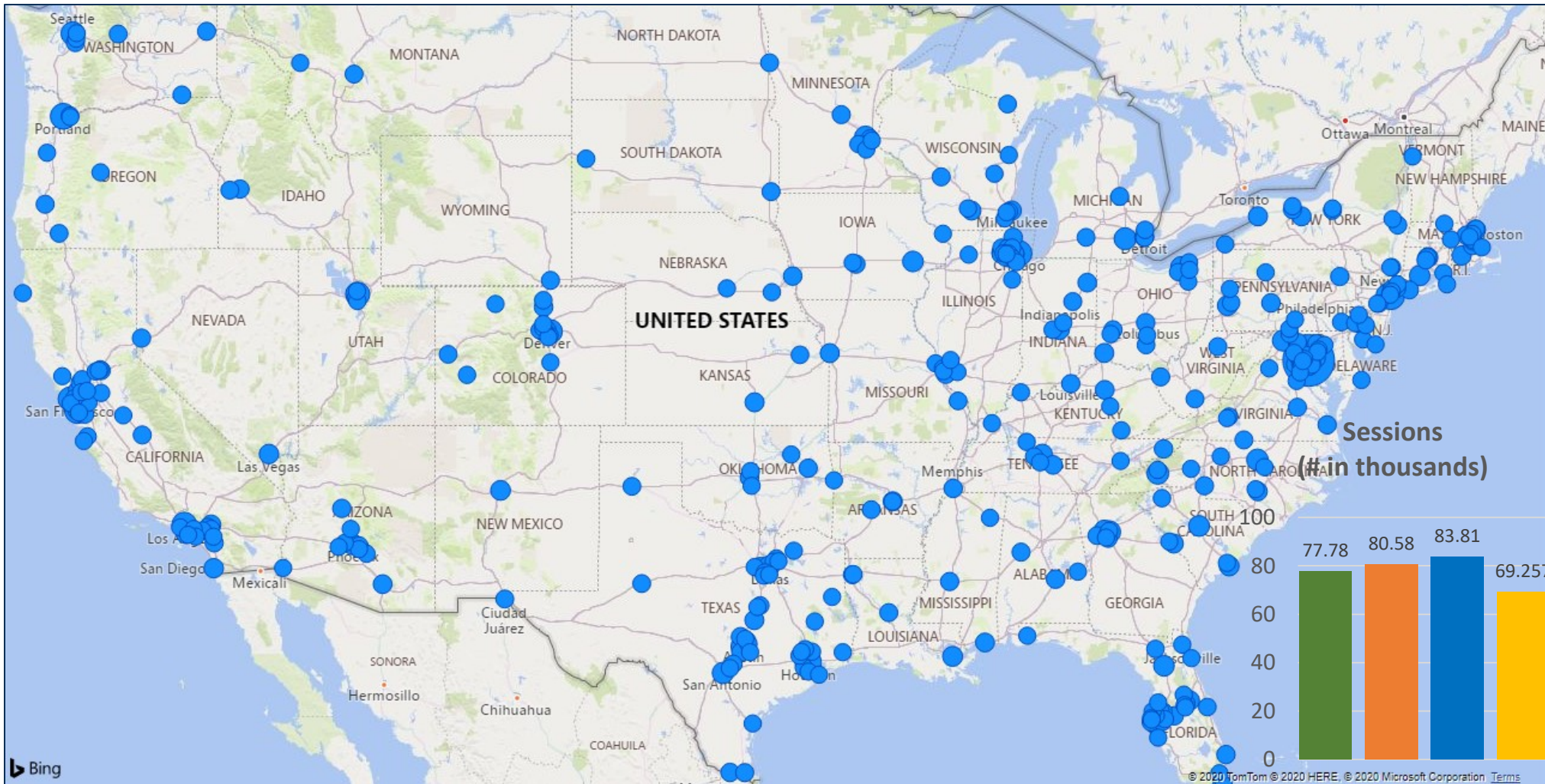VA | U.S. Department of Veterans Affairs

Education and Outreach: VINCI SQL & CDW Bootcamp

- Supports VINCI's goal of driving VA research by training research analysts on VA data basics. For new analysts, the training will rapidly accelerate their learning curve to productively contribute to research projects sooner and avoid crippling beginner data mistakes, especially the ones that unnecessarily burden shared data resources.

- Offered quarterly to 30-45 students (mostly COIN analysts).

- **204 students trained**, not counting the pilot.

- Course includes 8 hours of lecture, 10 hours of group code review, 8-10 hours of office hours, exercise sets and cohort definitions with access to a research training database, CDW references and documentation, and tutorials on using VINCI resources (the feasibility and cohort creation processes, the workspace, standard scripts, etc.).

Bootcamp Feedback

My team or center would be better off if more of our analysts took this course
I would recommend this course to a colleague who already uses the CDW
I would recommend this course to a colleague who is brand new to the CDW
I feel confident about my ability to pull CDW data
The skills and knowledge were relevant to my work
I learned useful skills

0   10   20   30   40   50   60   70   80

■ Strongly Disagree   ■ Disagree   ■ Neutral   ■ Agree   ■ Strongly Agree

Education and Outreach: VHA Data Portal

Users
(# in thousands)

61.86 | 53.80 | 35.17 | 29.55

Pageviews
(# in thousands)

160.39 | 173.28 | 181.90 | 151.22

Sessions
(# in thousands)

77.78 | 80.58 | 83.81 | 69.257

FY20 FY21 FY22 FY23

Precision Medicine

- The VINCI Precision Medicine team collects, links, and analyzes clinical and genomic data to support equity in the use and benefit of providing the right treatment to the right patient at the right time. They lead and support national initiatives and research related to genetics.

- **GDX:** Data product containing detailed patient level information about genetic services including tests results.

- **VINCI Prostate:** Data product containing detailed information about all-things prostate for patients with this type of cancer.
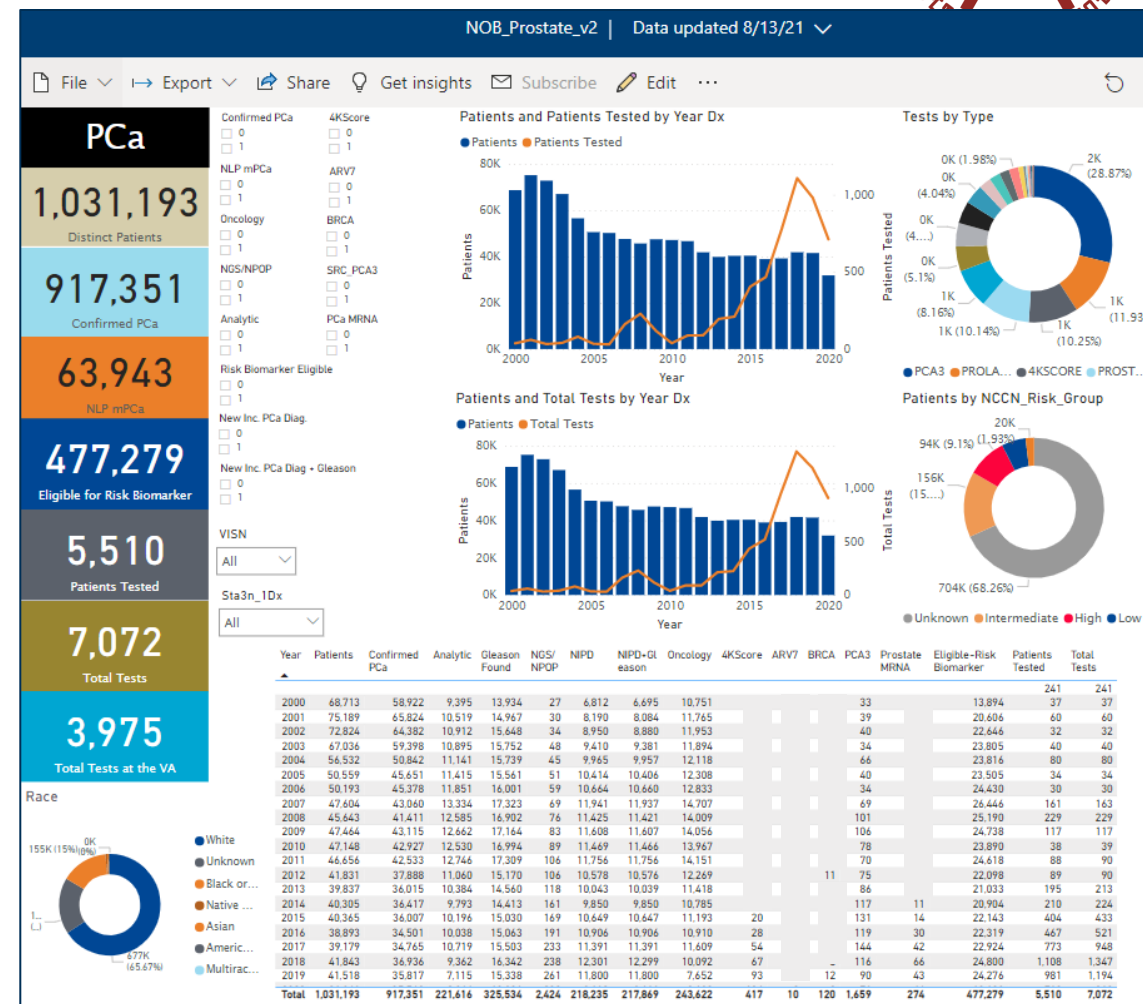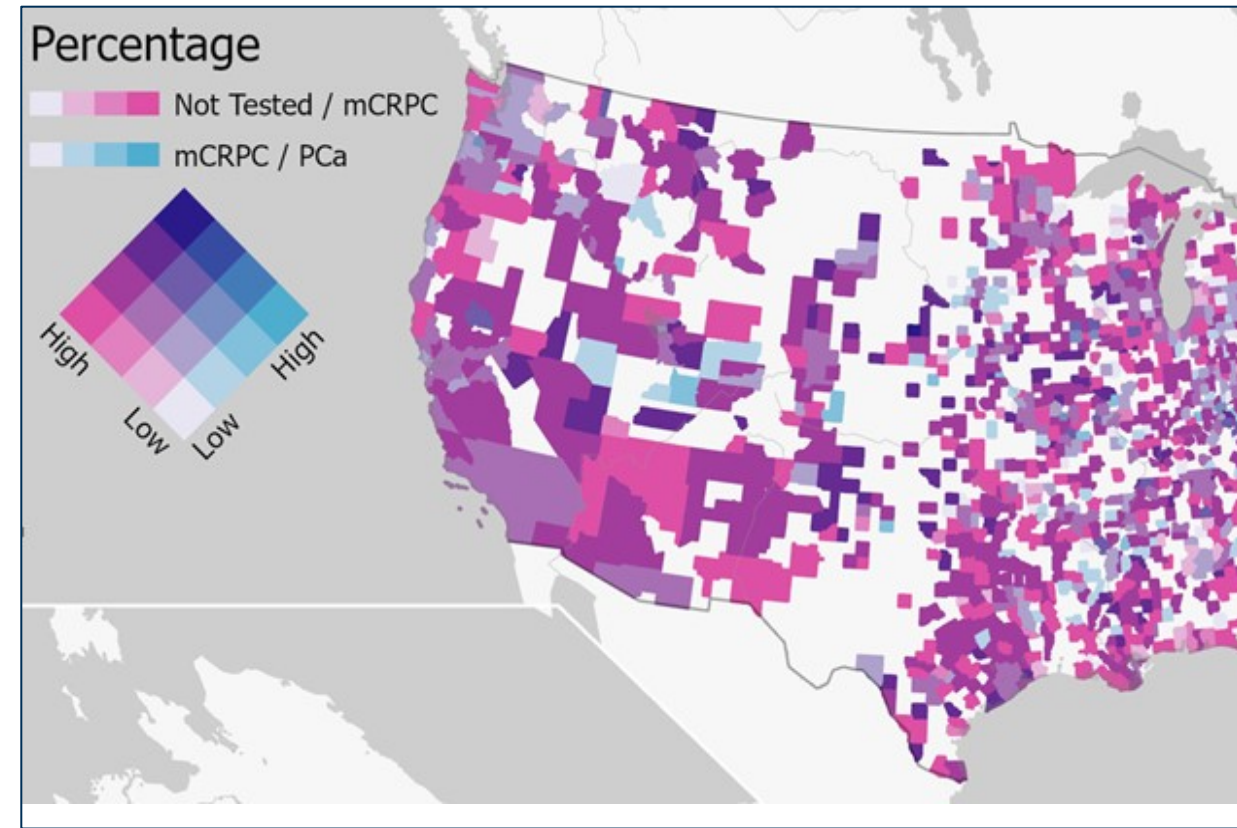
Precision Medicine

- **Genetic Services:** Dashboard to access notes related to genetic services.

- **VATCGA/VINCI Cancer GDx:** The Cancer Genome Atlas for the VA.

- **Lung Nodule:** Dashboard for physicians to review radiology reports and assess patient's access to XL2 tests to rule out lung cancer, and NLP tool to determine dimensions of lung nodules from radiology reports.

- **Hereditary Transthyretin Amyloidosis (hATTR) Dashboard:** Analytic file and dashboard that allows physicians to prioritize genetic services recommendations for patients at high risk of hATTR.

Precision Medicine

- **VINCI Enviro:** Database repository for geospatial and patient level data related to environmental and socioeconomic factors. Datasets include Area Deprivation Index and Social Determinants of Health.

- **VINCI Cardiac:** NLP algorithm for Carotid Stenosis and Abdominal Aortic Aneurysm.

- **VINCI RadOnc:** Nation-wide database of radiation oncology services.

Precision Medicine: National Radiation Oncology

- Developed, validated, and deployed code to identify Veterans who underwent radiation treatment.

- Linked these data to the RADHinge data frame that has been built to extract more complete dosage data.
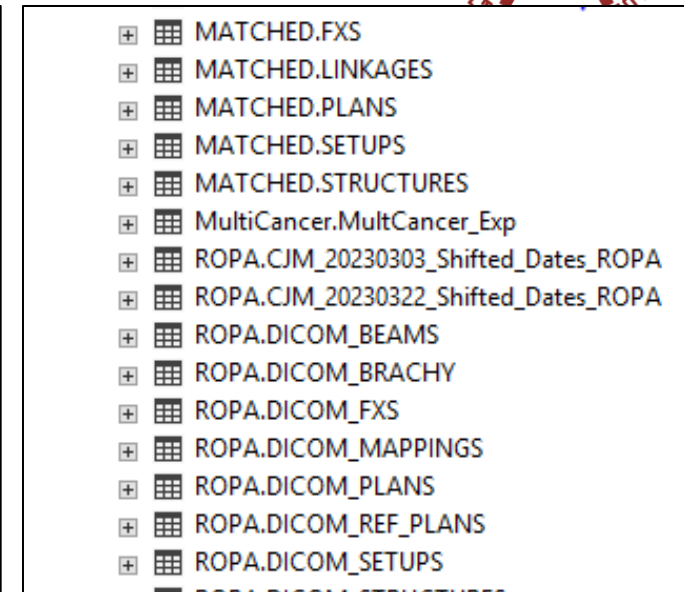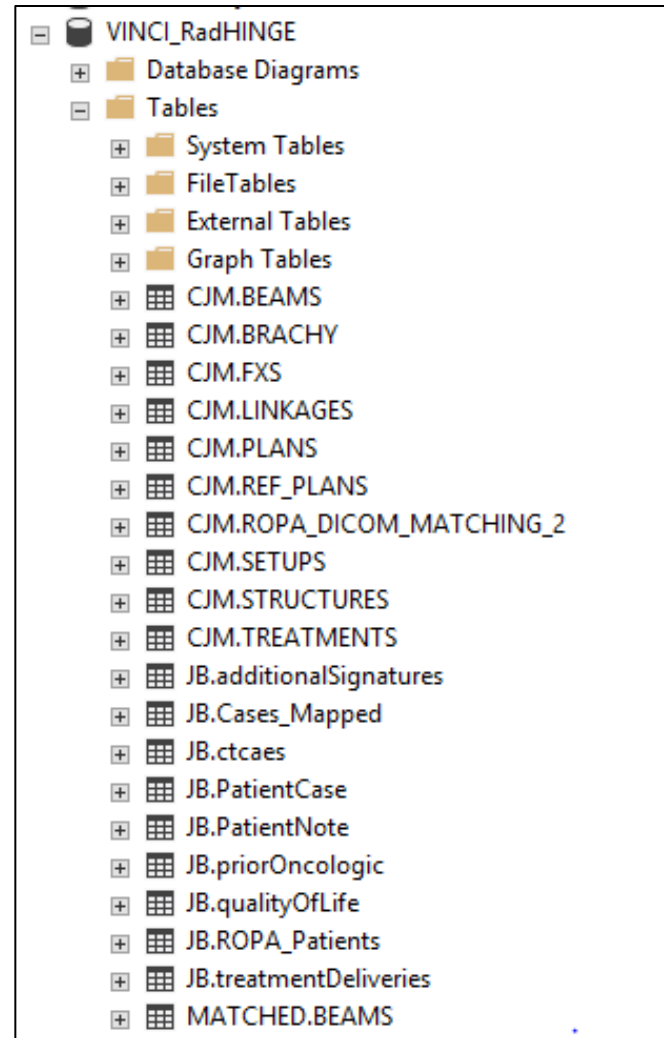
Precision Medicine: Genetic Epidemiology

- Describing Genetic Testing Rates (Germline and Somatic) in mCRPC Population

  - MVP Analysis:
    - Incorporating genetic marker data into risk models:
    - Prostate cancer and Alzheimer's disease
    - Pan-cancer polygenic risk scores
    - Autoimmune genetic markers and melanoma risk
    - Prostate biopsy/PSA modeling

  - GWAS:
    - Left ventricular wall thickness
    - Bilateral carpal tunnel syndrome
    - hidradenitis suppurativa
    - FIB4 slope

  - PheWAS:
    - TTR genotypes
    - XXY/XYY/XXX karyotypes
    - HSD3B1 genotypes

Precision Medicine: MVP Research

- Performed as a research site on several MVP Research studies serving as experts in conducting deep phenotyping, developing, validating, and deploying algorithms that rely on more extensive clinical data to define patient phenotypes.

- MVP003 Genetics of cardiometabolic disease has generated more than 50 high impact publications.

- Activities include:

  - Developed numerous standardized phenotype data frames that can be used across studies:
    - Disaggregated PheWAS data frame
    - MedWAS data frame
    - ProcedureWAS data frame
    - LabWAS data frame

  - Curated and structured phenotype data elements from clinical notes using NLP:
    - Ankle brachial index
    - Aortic root diameter
    - Amyloid deposits

Precision Medicine: NLP Work

- Collaborated extensively with other VINCI teams to achieve more and better results together than could otherwise be potentially done individually, such as leveraging NLP for:

    - Prostate DataCore (NLP variables)

    - Metastatic Prostate Cancer Flag

    - Gleason Score Variables

    - Gene Extraction Tool

    - Race_Eth NLP- Cleaner variable for Race/Ethnicity using NLP for those previously marked as 'unknown'

    - Castrate Resistant Prostate Cancer (CRPC)/Metastatic Castration Resistant Prostate Cancer (mCRPC) Tool

    - Used a combination of structured data and NLP tool to develop CRPC and mCRPC algorithm to identify underreported patients and further will be implementing into Pca Datacube and used in current and future epidemiological studies.

    - Using NLP to create a tool to extract genetic result and variant information from clinical notes for genetic epidemiology studies from lab data. (Lab Data Feeds and CDW Lab Chem)

NLP

Applications Development

Precision Medicine

Industry Team

Data Services/MVP Core

Annotation Team
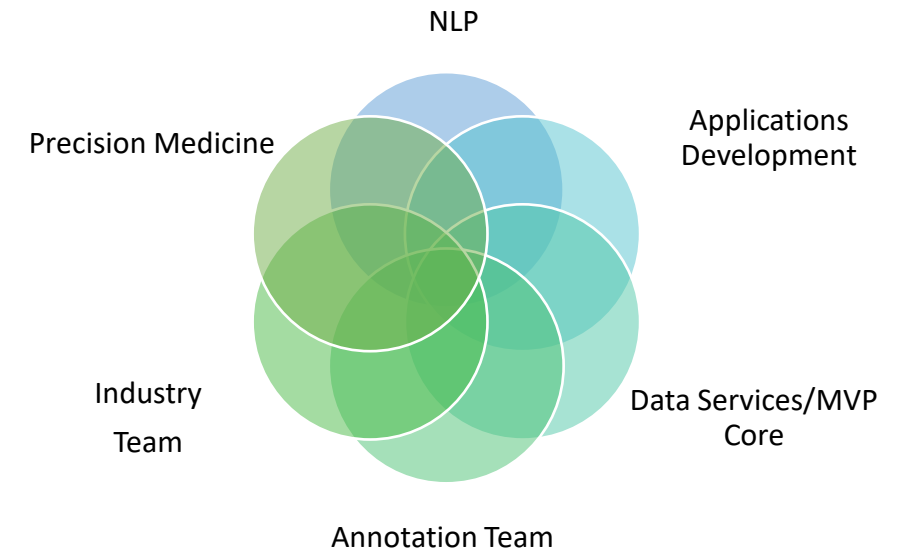
Pharmacogenomics

- **VINCI PHASER:** Data core supporting the implementation and evaluation of the national clinical program PHASER (Pharmacogenomic Testing for Veterans) to improve medication safety and optimize treatment based on patient genetic profile.

- **VINCI PGx:** Data initiative to facilitate the curation, interpretation and dissemination of pharmacogenomics results in practice (collaboration with Pharmacy Benefit Management – MedSafe and Office of Informatics; PERC STORM Dashboard).

VINCI PHASER To Support Workflow and Operations

Dashboards with Aggregated data
Wide access to site champions, site coordinators, leadership, department heads and providers

Operations Workflow

Dashboards with Patient-level data
Access restricted to site coordinators and central team

Performance metrics
- Volume of ordering
- Provider participation by services
- Patient demographics

Patient visit
- Consent to testing
- Test order
- Sample Collected

Increase patient participation
- List upcoming appointments for educational mailings and outreach

Test performed by lab

Increase workflow efficiency
- Facilitate order follow-up
- Prevent duplicate ordering
- Refresh lapsed orders

Results return to VA
- Result entry in EHR
- Lab record created
- Provider and patient informed of results

Patient outcome tracking
- Follow provider action after testing using medication prescription data and CROC alerts

Support quality and reliability
- Detect delay in result entry
- Audit entry of test results
- Facilitate result mailings to patient

Medication order
- Test result alert

Data Workflow for VINCI PHASER Data Core
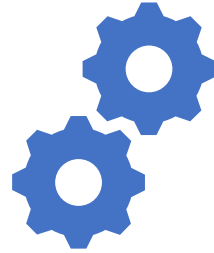
Data Collect

CDW Data   Lab Data feed

CDW data (daily)
+
Structured data
from reference lab
(monthly)

Production

Automated procedures to
extract and transform CDW
data
+
Dashboard Development

Publishing

Dashboard content refreshed daily

- Report cards in PowerBI
  - National
  - Site

- Operations reports on
 BISL Sharepoint
  - Patient level

Iterative feedback from team and users
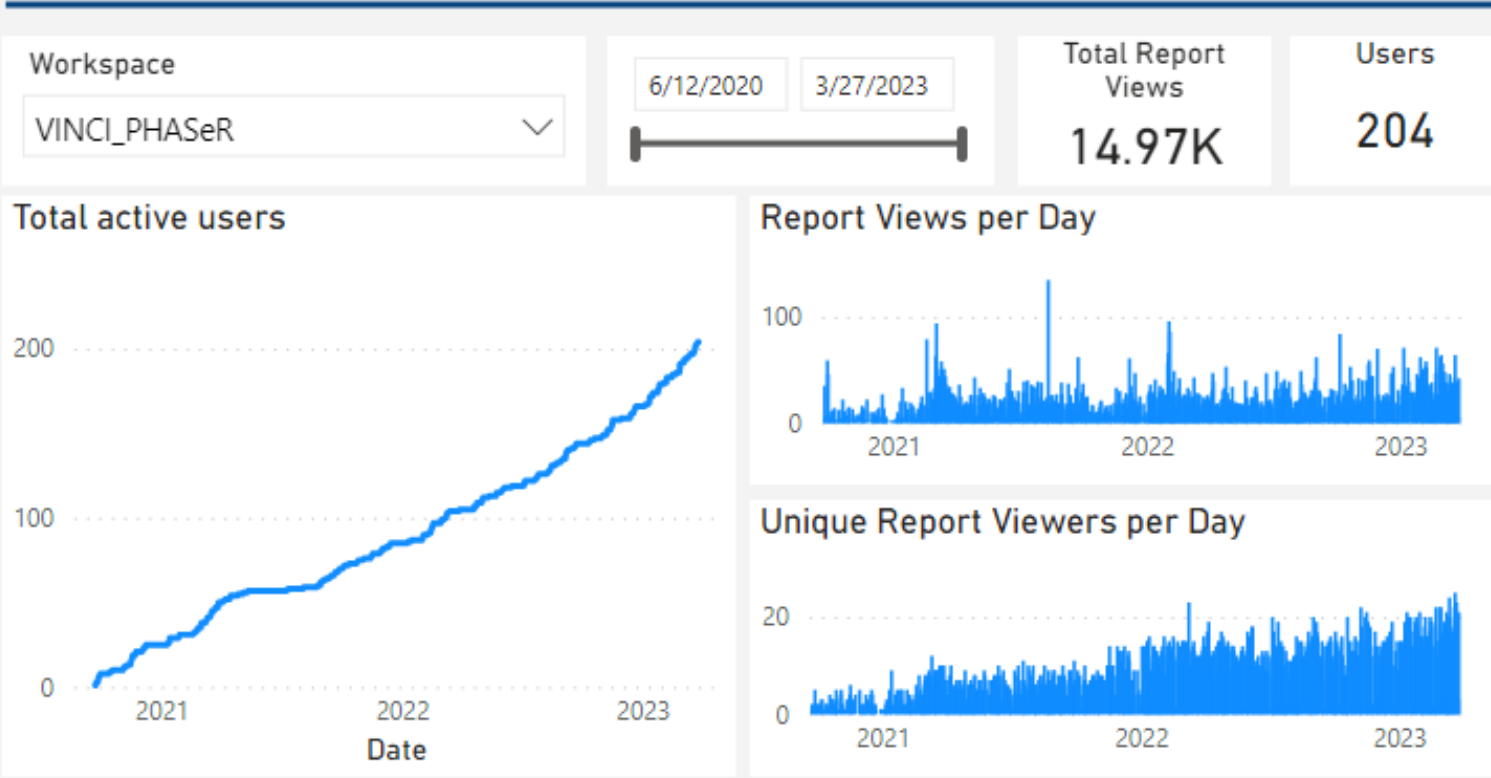
Quality Dashboards and Performance Metrics

- Audience: Site champions, site coordinators, leadership, department heads and providers

- Dynamic reports in **Power BI** with aggregated data pulled daily from CDW using an automated process

- User-friendly access on a web browser after joining a Team

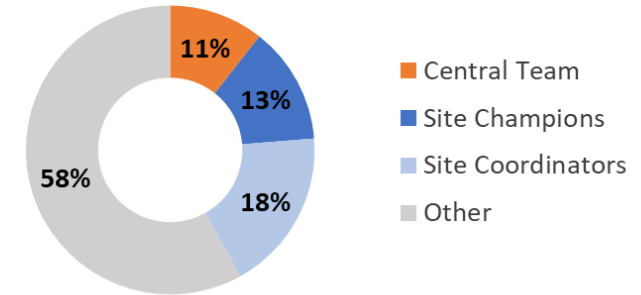- View can be customized and saved by user
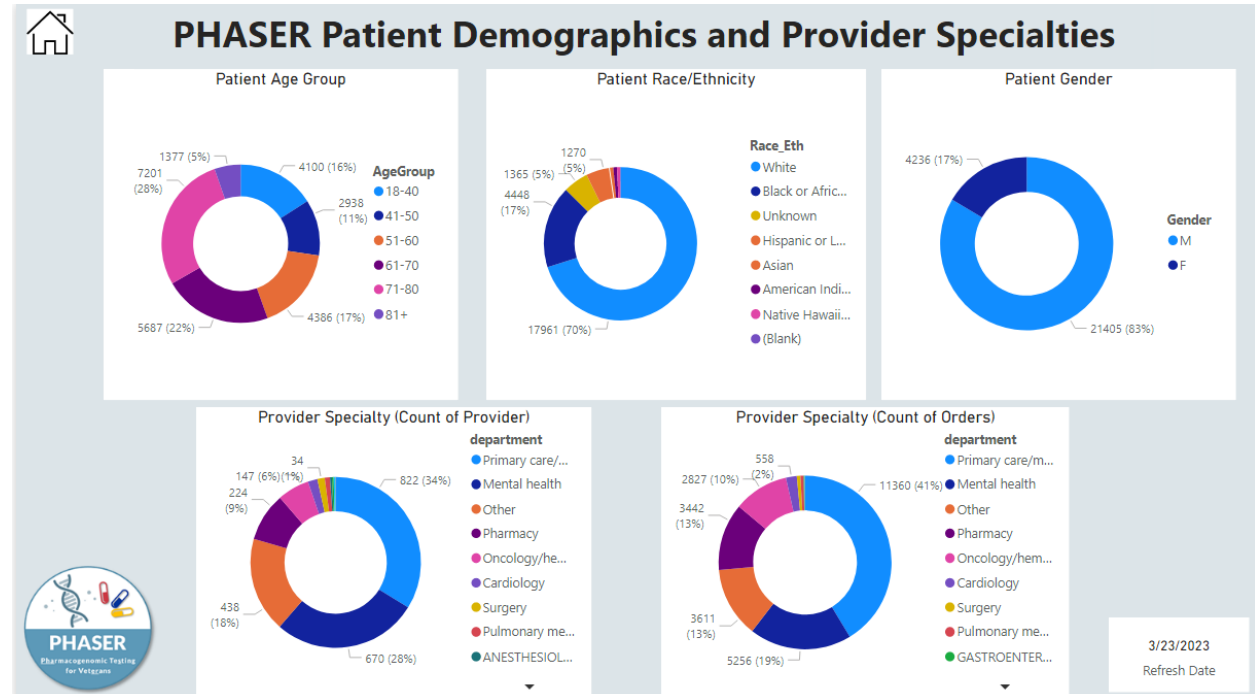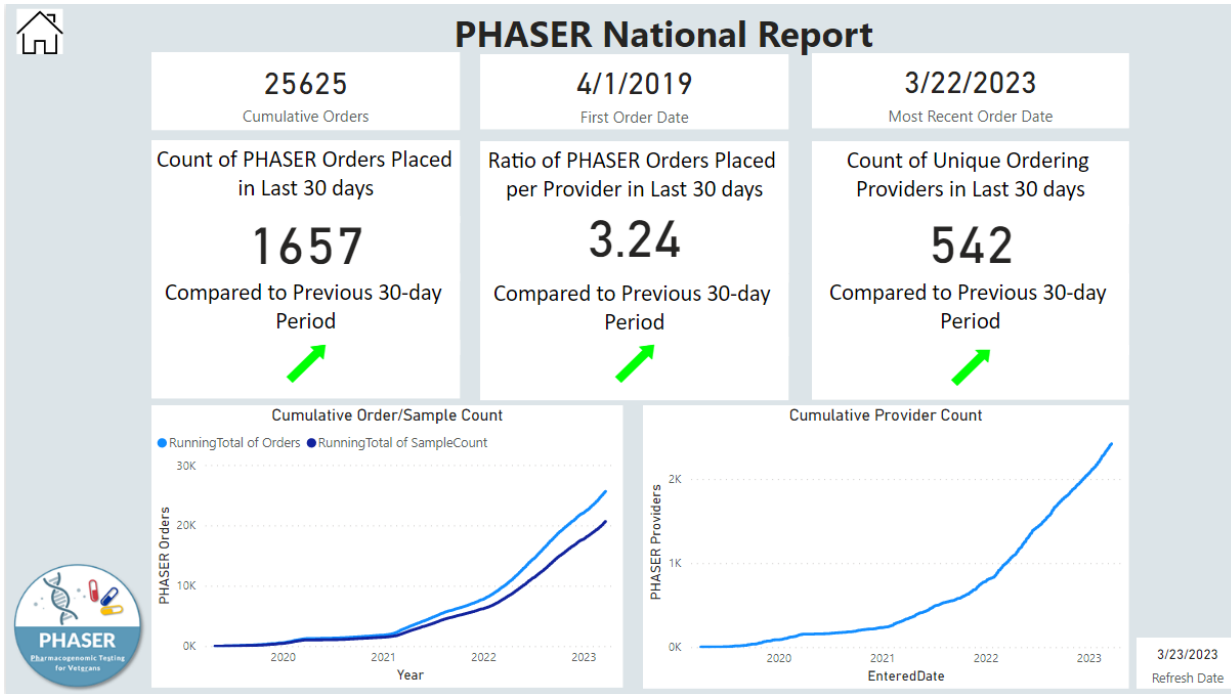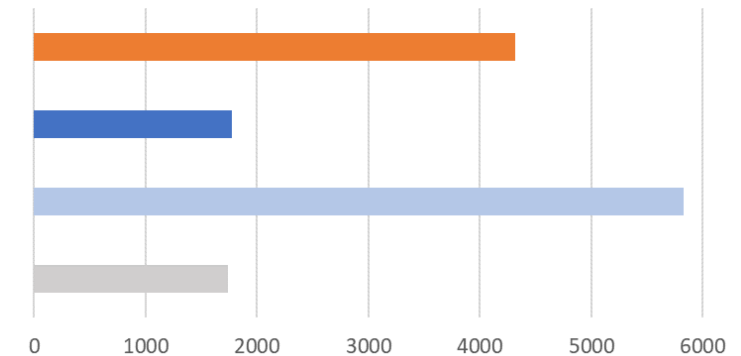
Quality Dashboard Usage

Clinical Decision Support in CPRS

Example of alerts in response to an order for codeine if a patient has a test result indicating a PGX interaction

CYP2D6 ULTRARAPID METABOLIZER – Codeine – Severity HIGH

```
Pharmacogenomics Warning - Consider Alternative Agent
=========================================================
ACTION: Strongly consider removing codeine order.
        Patient at increased risk of toxicity from codeine
        based on pharmacogenomics (PGx) results.

ALTERNATIVES:
        * Non-pharmacological treatments
        * Non-opioid analgesics (e.g., NSAIDS)
        * If an Opioid is indicated, do NOT use:
            - Tramadol
            - Hydrocodone
            - Oxycodone

RATIONALE: Patient is an ultrarapid metabolizer of CYP2D6.
        Higher levels of codeine active metabolite occur than
        expected. Possible toxicity includes:
          - Life-threatening respiratory depression
          - Increased fall risk
          - Increased sedation

RESOURCES:
        * Clinical assistance: order local or interfacility
        pharmacogenomics consult
        * Information on testing: go to https://bit.ly/PHASERhome
        * Feedback on this alert: email PHASER-CROC@va.gov
        * Non-clinical support: IM or email PHASERtechsupport@va.gov
```

CYP2D6 POOR METABOLIZER – Codeine – Severity MEDIUM

```
Pharmacogenomics Warning - Consider Alternative Agent
=========================================================

ACTION: Consider alternative to codeine (and meds metabolized by
        CYP2D6). Patient may experience insufficient pain relief
        based on pharmacogenomics (PGx) results.

ALTERNATIVES:
        * Non-pharmacological treatments
        * Non-opioid analgesics (e.g., NSAIDS)
        * If an Opioid is indicated, do NOT use:
            - Tramadol
            - Hydrocodone
            - Oxycodone

RATIONALE: Patient is a poor metabolizer of CYP2D6.
        Far fewer active metabolites form than expected
        resulting in reduced pain relief.

RESOURCES:
        * Clinical assistance: order local or interfacility
        pharmacogenomics consult
        * Information on testing: go to https://bit.ly/PHASERhome
        * Feedback on this alert: email PHASER-CROC@va.gov
        * Non-clinical support: IM or email PHASERtechsupport@va.gov
```
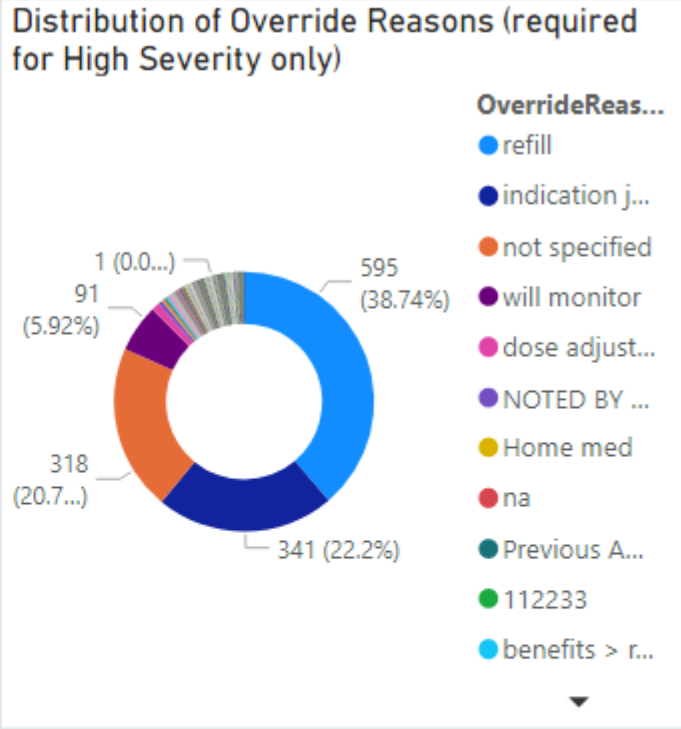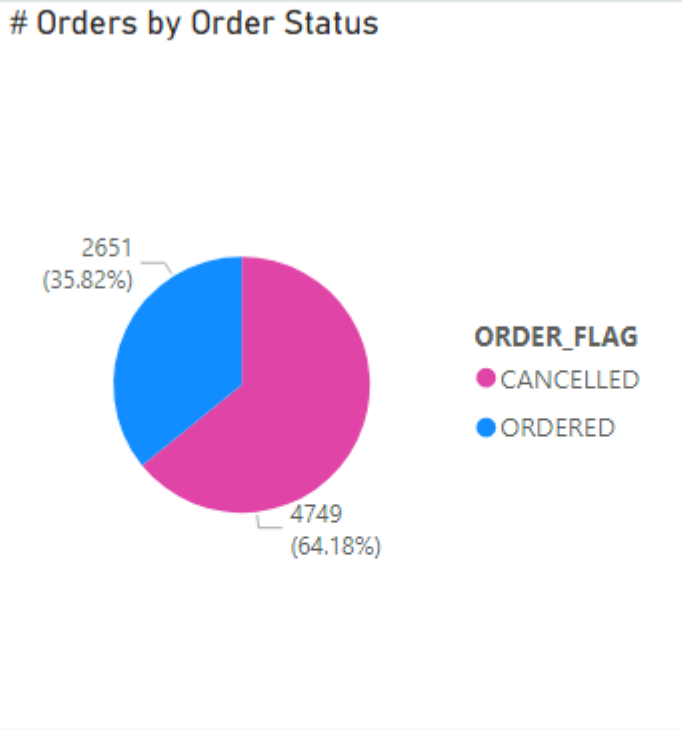
Provider Actions Following CROC Alerts During Prescription

Total Alerts
7400

Unique Patients
2661

% Of Patient Tested
13.44%

Orders by Order Status

2651 (35.82%)

4749 (64.18%)

ORDER_FLAG
- CANCELLED
- ORDERED

Distribution of Override Reasons (required for High Severity only)

OverrideReas...
- refill
- indication j...
- not specified
- will monitor
- dose adjust...
- NOTED BY ...
- Home med
- na
- Previous A...
- 112233
- benefits > r...

1 (0.0...)
91 (5.92%)
595 (38.74%)
318 (20.7...)
341 (22.2%)

Total Providers
2037

Total Stations
30

First Occurrence Date
7/23/2021

Last Occurrence Date
3/28/2023

Applied Natural Language Processing (NLP)

- Offers a customized sets of natural language processing (NLP) tools, services and data that allow investigators to utilize accurate, trustworthy, and large-scale NLP systems in the electronic health record.

- This led to focusing on the following major initiatives:

  - Investigated new methods:
    - ✓ Implemented language model architectures for NLP systems across clinical domains (oncology; pathology; cardiology; mental health; infectious disease; etc.)
    - ✓ Developed large language models trained on VA clinic notes, allowing for more control, and improved accuracy and trustworthiness of models used in the VA -- Completed on the Health Data Analytics Platform (aka summit).
    - ✓ Investigated new ways to conduct chart review & annotation working with open-source software (John Snow Labs).
    - ✓ Investigated 5 open-source note deidentification systems performance on VA notes.

  - Contributed to open-source software and collaborative global science initiatives
    - ✓ Delivered stable Version 1.01 Release of MedSpaCy software, which has experienced significant adoption in the open-source community.
    - ✓ Continued support of the open-source NLP software Leo, which has been used to develop over 100 NLP pipelines at the VA.
    - ✓ Released methods and code for multiple concept-specific systems and publications.

FY23 Medspacy Community Contributions
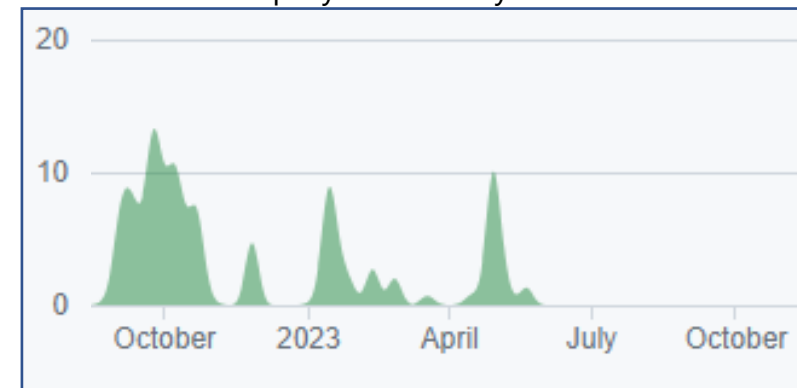
Applied Natural Language Processing (NLP)

- Training, Education, and Publication Efforts

  ✓ Delivered 4 posters, 1 tutorials, 5 published papers as first authors, and 1 talk at events, which does include efforts supporting other leads in their efforts as submitters and first authors. Supported participation in conferences hosted by the National Institute of Health, National Cancer Institute, American Medical Informatics Association, and International Medical Informatics Association.

  ✓ This consultation and education of 30+ requests for NLP consultation and provided education for new/existing studies

  ✓ Developed a standard set of frequently asked questions related to NLP at the VA for VA researchers.

- Continued support for operational projects and partners

  ✓ Developed new complex phenotypes for VA COVID -19 Shared Data Resource, Prostate Data Core, and VINCI Precision Medicine team.

  ✓ Successfully completed regular updates and quality checks for provisionable data sources including the CSDR, Prostate Data Core, and LVEF NLP data; processing over hundreds of millions notes for various systems to date.

  ✓ Provided technical support for other operational partners including MVP, CSP and BASIC.

NLP System	Notes Processed	Patients with Output Data
COVID-19 Screening	29,611,151	2,108,938
Mechanical Ventilation	607,005,349	103,679
Menopausal Status	148,884,005	175,123
Reason for Stay	7,815,991	1,509,853
Respiration and Oxygen Therapy	64,664,427	397,772
Vital Signs	584,379,994	2,362,487
All Systems	1,449,360,947 (total) 721,010,729 (distinct)	2,406,293 (distinct)

Table 1: The scale of the amount of data processed for the CSDR, including total number of notes processed and total number of patients with NLP output from each system (as of March 1, 2023).

Table (Above): Table quantifies the amount of data processed, and resulting patients identified and available in the CSDR. This data would be unavailable to researchers it were not for the NLP pipelines produced.
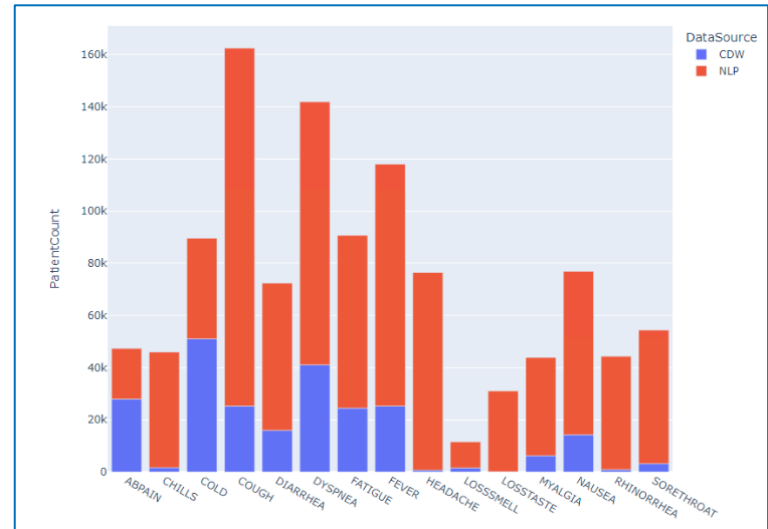

Table (Above): Symptoms available in the CSDR for all researchers that identified that were identified by NLP versus those that were not.

Chart Review

- Offers Chart Review as a service to researchers to provide education, training, and support of chart review, medical record abstraction, and chart annotation of clinical records for studies.

- This can include validating NLP programs and reviewing medical records to collect data for research projects.

- Completed chart review or annotation projects for:
  - ✓ Precision Medicine
  - ✓ MVP
  - ✓ VINCI NLP Team
  - ✓ National Surveillance Team
  - ✓ VA Investigators
  - ✓ VINCI Investigators
  - ✓ Sponsored research projects

Chart Review

- Completed **17 chart review or annotation projects**:

✓ Adam - ADRD

✓ COVID Death Review

✓ COVID State Reporting

✓ COVID19 Operational Support

✓ COVID Death Note Titles

✓ Damrauer – AAA

✓ Damrauer – Bicuspid Aortic Valve

✓ Frey - Social Determinants of Health

✓ Heredity Cancer Genetic Testing

✓ Liver Biopsy Annotations

✓ Long Covid Symptoms

✓ OPS & MVP – XL2 Lung Nodule

✓ Schroeck - Qual Path Cancer Synoptic Reporting

✓ Solid Tumor NLP

✓ VINCI Data – Deidentification

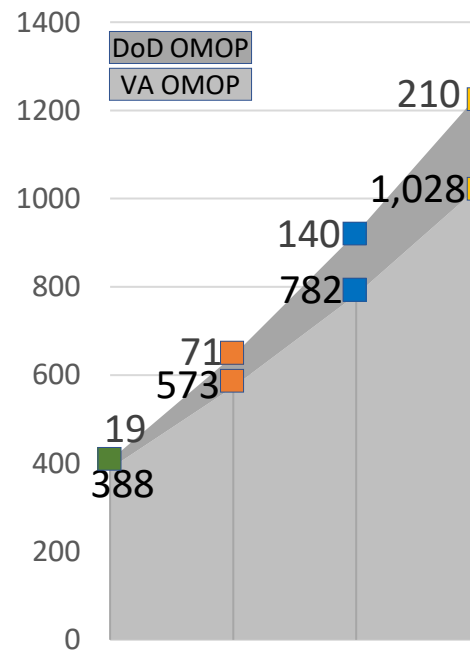✓ Janssen Vaccine Safety – Special Project

✓ J/P NDMM

VA OMOP

- Sustainably transforms VA data into the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) to optimally organize it with standardized content and vocabularies that allows investigators to conduct more efficient analysis by reducing the level of deep and idiosyncratic understanding required of the source data.

- The resulting VA OMOP datasets provide VA researchers and operations users a significantly more concise and simpler to use form of the VA clinical data, resulting in substantial time savings when exploring, searching and utilizing data.
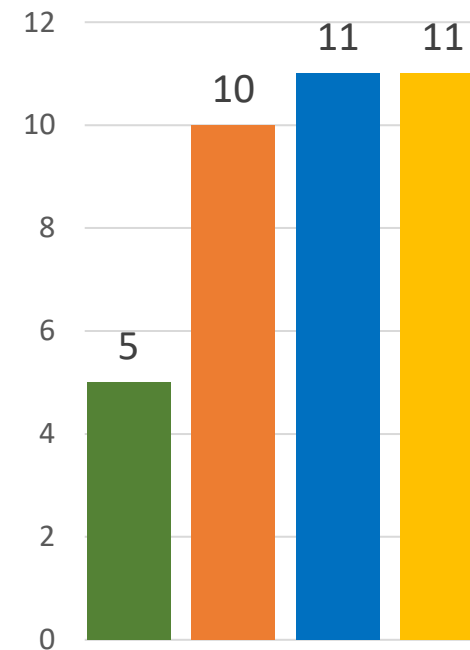
VA OMOP

- 1,238 active projects using OMOP and supported
  - 1,208 using VA OMOP
  - 210 using DoD OMOP

- Average of 18 new projects supported monthly

- 166 Help Desk Tickets resolved
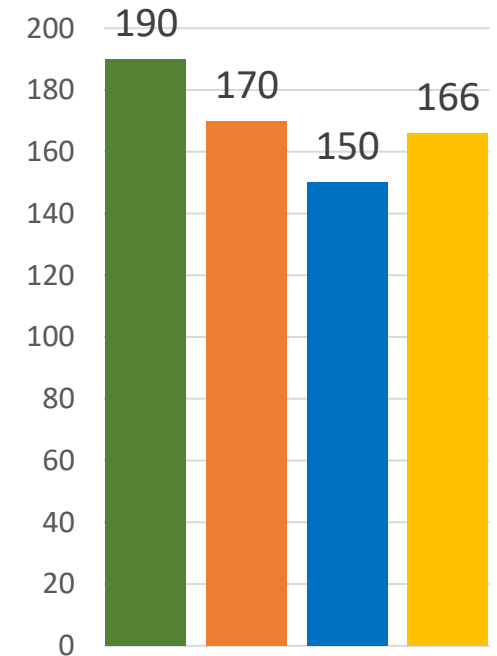
- 11 refreshes completed in FY23

Active Projects using OMOP

DoD OMOP
VA OMOP

- 210
- 1,028
- 140
- 782
- 71
- 573
- 19
- 388

Refreshes Completed

- 5
- 10
- 11
- 11

Help Desk Tickets Resolved

- 190
- 170
- 150
- 166

VA OMOP

- Continued support of *All of Us* Research Program with targeted table deduplications and concept mappings.

- Improvements, new data sources, and mappings:
  - ✓ Created a second OMOP pipeline and production instance coming from Cerner
  - ✓ Created a full Dev pipeline of both OMOP instances, allowing us to test complex code on a limited subset of patients with highly sufficient data to evaluate performance
  - ✓ Significantly reduced duplicates, developed over 20 new QA scripts, reports and in-process checks to ensure quick evaluation and remediation of data quality for both transforms
  - ✓ Added Rad_RadiologyNuclearMedicineOrder source data
  - ✓ Added Bilirubin, Carbamazepine, Complement C3, Leukocyte Esterase, Ammonia and over a dozen other mappings

VA OMOP

- Created an historical snapshot synch process that significantly reduces storage requirements while enabling evaluation of changes between each refresh in addition to critical reproducibility.

- Began developing an all-OMOP (CDW, Cerner, CMS, DoD) architecture merger to enable better population synching between sources and reducing cost and deduplication for each researcher.

- Developed a semi-automated mapping, tool stack, and mapping 'age' reporting process.

- Developed critical vocabulary mapping evaluation checks for new OHDSI releases.

- Updated source control methods and tools for Change Management.

- Various document and training material updates.
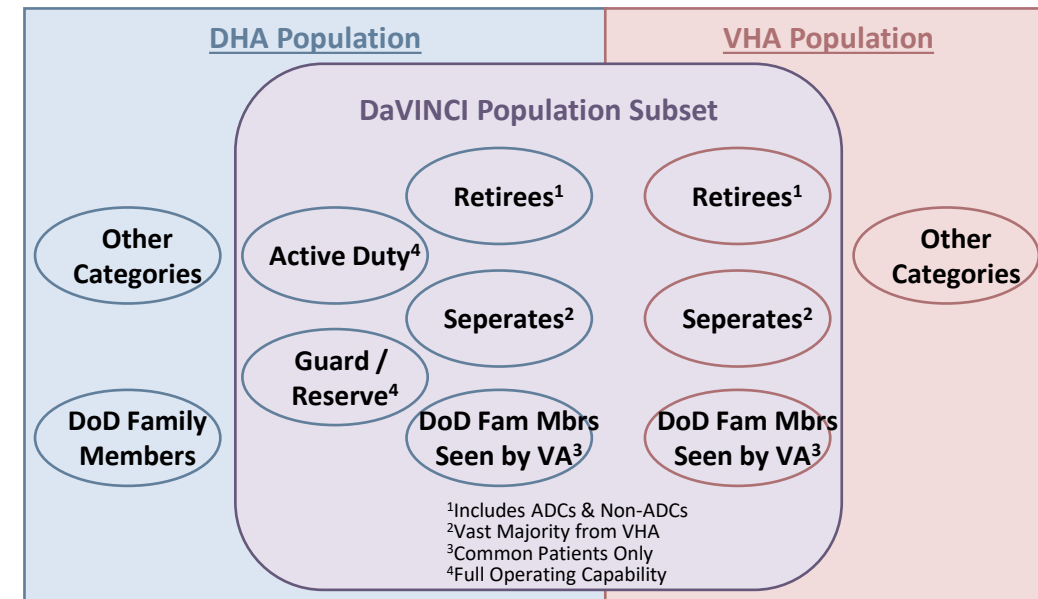
Partnerships & Linking

- Explores and builds capacity to support operations and research use cases that seek to link data sources to mitigate fragmentation of health and other data that cannot be addressed by existing interoperability efforts (e.g., Health Information Exchange), a problem that is growing since the passage of the Mission Act.

- The team is:

  - Partnering with ORD, National Center for Ethics in Health Care, CDC, GSA and NIH on initiative to define ethics of privacy-preserving records linkage across the federal government

  - Collaborating with VA/DoD JDA Executive Committee to leverage DaVINCI governance and infrastructure to realize a single, bidirectional, interagency exchange medium for large data and to enhance the exchange infrastructure to increase efficiency

  - Working with National Data Systems and ORPPE to effectively implement VA Data Ethics/HIPAA Minimum Necessary Criterion in the review of large VA data requests

Partnerships & Linking: DaVINCI

- The Department of Defense (DoD) and Department of Veterans Affairs (VA) Infrastructure for Clinical Intelligence (DaVINCI) provides governance related to data stewardship and oversees transfer of data between VHA and DHA. In this way, DaVINCI provides a consolidated view of individuals from accession to internment.

- Creates a consolidated view of healthcare from accession to interment for service women and men, Veterans, and other eligible patients receiving care from DoD to VA.

- Individual who have records in each system: 38.0M
  - DoD Service Women and Men: 11.1M
  - VA Veterans: 26.9M

- Individuals who have records in both systems: 6.1M

Partnerships & Linking: DaVINCI

- The EIDS Program office in PEO DHMS and the VINCI service in VHA's ORD have partnered with DoD and VA IT departments to successfully implement the Joint Virtual Private Network (JVPN) data exchange medium to regularly share large data files between DHA and VHA.

- The JVPN conforms to both Departments' privacy and security standards and is capable of bidirectional delivery of large files containing confidential information and PHI between environments of the MHS Information Platform (MIP) and the Corporate Data Warehouse (CDW).

- To date, this capability has been used to share health record data between Departments to serve each Departments' analytics use cases as part of the DaVINCI program.

- All DoD and VA stakeholders with a need to share large data files between the Departments are encouraged to coordinate with the DaVINCI program sponsors/designees, who manage the data exchange.

- Guidance is currently being proposed that any new request for a similar capability seeking JEC approval will need to include a statement explaining why the JVPN (or latest enterprise-endorsed technology deployed by the DaVINCI program) cannot meet its needs.
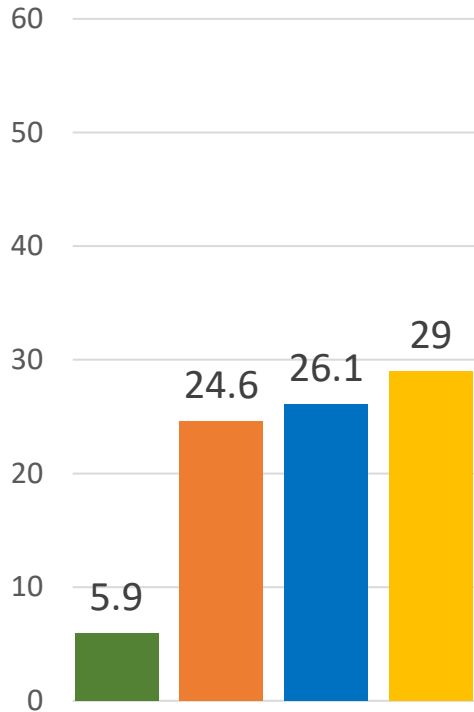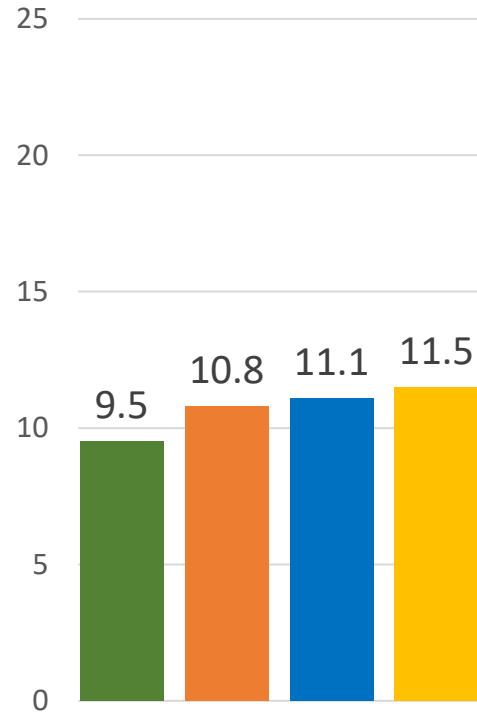
Partnerships & Linking: DaVINCI

Patient Data transferred from DHA to VHA

Patient Data transferred from VHA to DHA

Rows Transferred (# in billions) — DHA to VHA: 5.9, 24.6, 26.1, 29

Patients (# in millions) — DHA to VHA: 9.5, 10.8, 11.1, 11.5

Rows Transferred (# in billions) — VHA to DHA: 20, 52, 54, 58

Patients (# in millions) — VHA to DHA: 23, 23, 23, 23

Legend: FY20, FY21, FY22, FY23
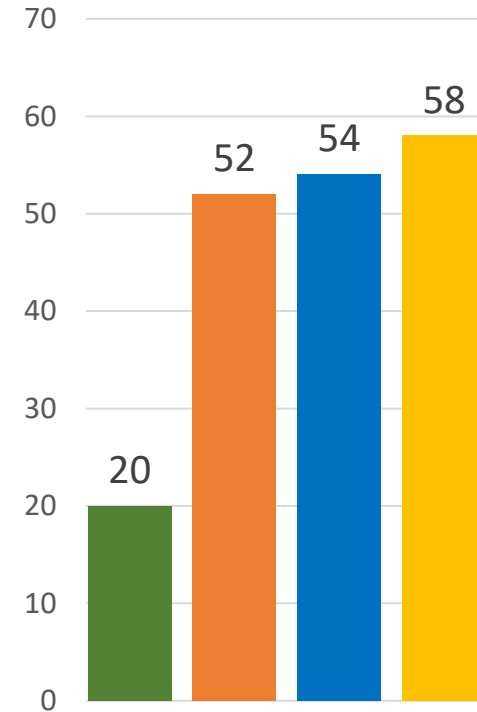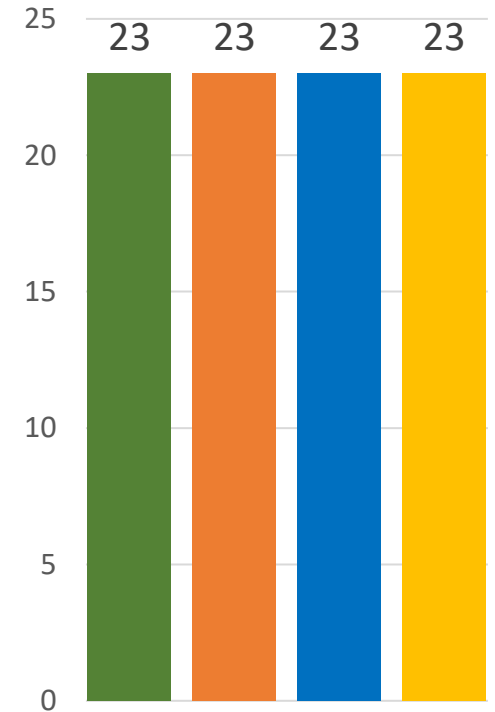
Partnerships & Linking: DaVINCI

Research Projects Supported (#)
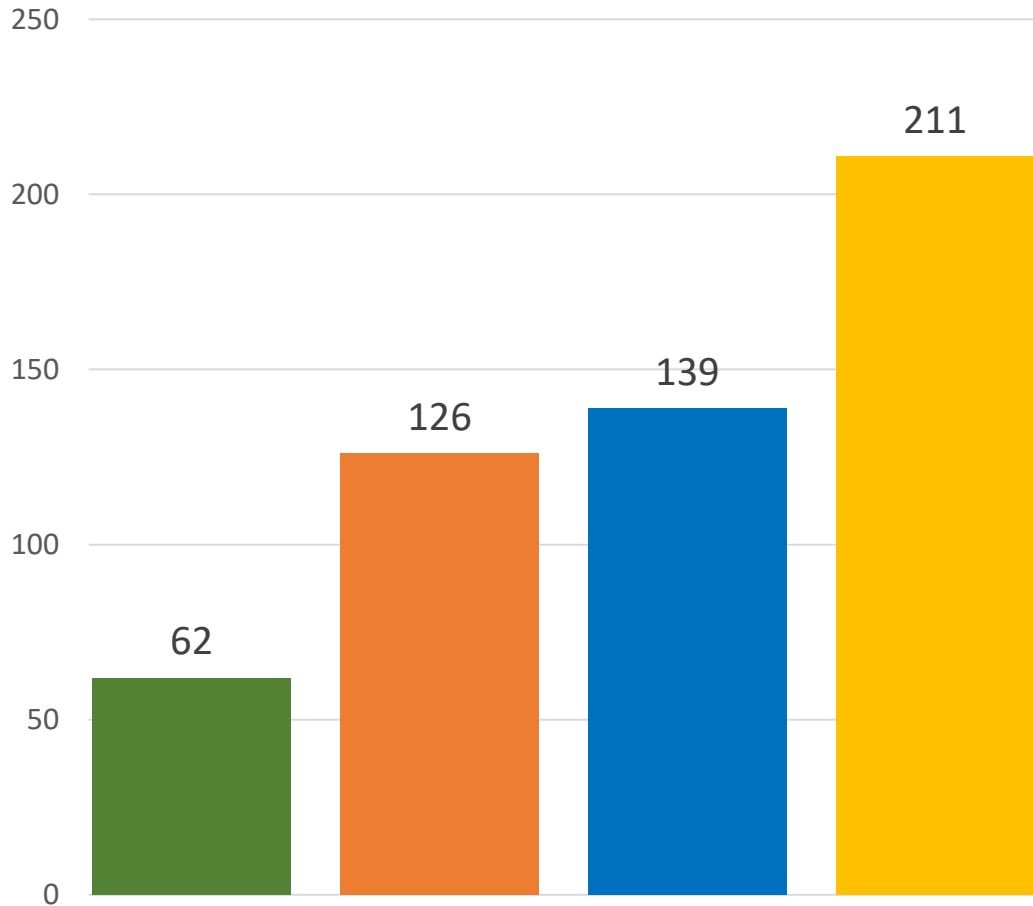
Operational Groups Supported (#)

FY20 FY21 FY22 FY23
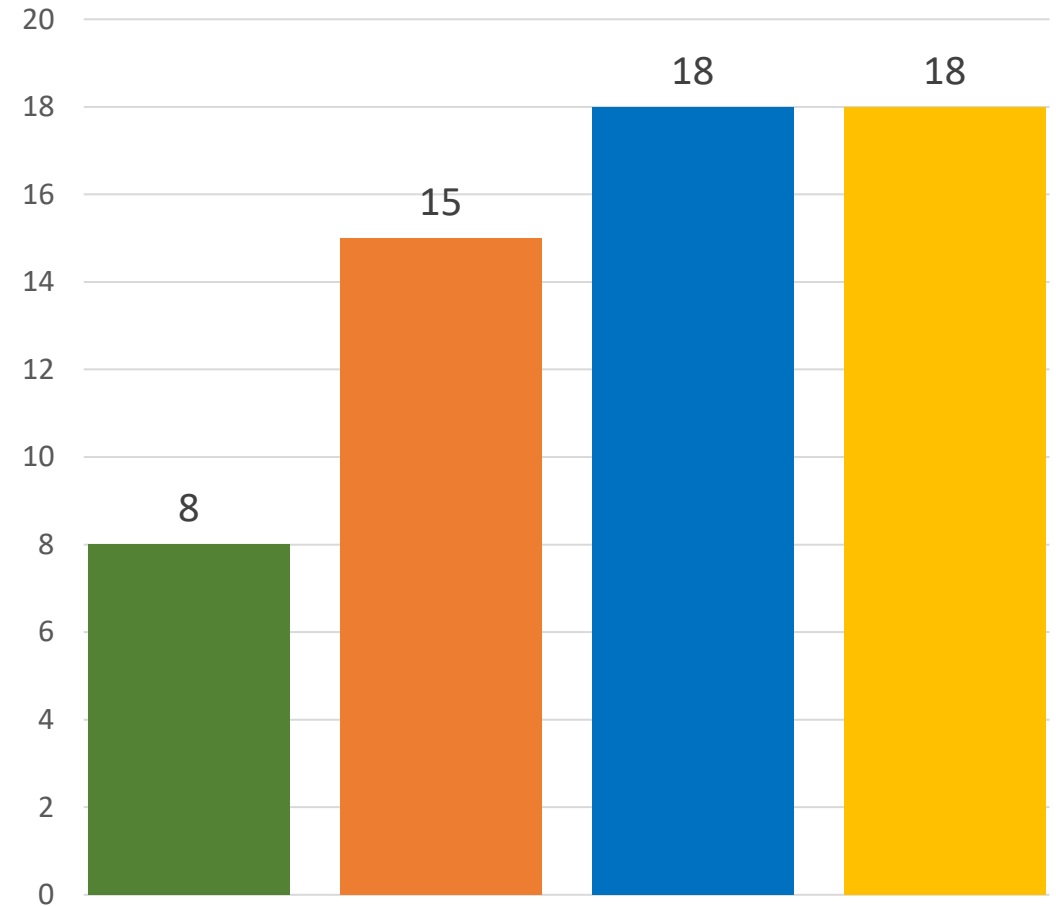
Research Applications

- Develops, pilots, deploys, and/or host algorithms and tools (e.g. applications) that result in new data products or supports common research tasks to facilitate data-driven research and secondary analysis of medical record data.

- In addition, VINCI has developed applications to support back-office research operations.

- Much of data-driven research and secondary analysis of medical record data requires development of these algorithms and applications.

- Resource center staff create and pilot algorithms and applications that result in new data products or support common research tasks.

- In addition to developing, resource center staff, where appropriate, prepares a path to deploy or provide access to these tools for other VINCI Resource Center users.

Research Applications

- **National strategic AI initiatives.** VINCI has been instrumental in contributing to national strategic AI initiatives within the VA. We continually provide thought leadership to the National Artificial Intelligence Institute (NAII) and subject matter expertise in the application of Artificial Intelligence within the VA.

- **ORD Application Hosting Cloud Initiative.** VINCI is tasked with overseeing the strategy and development of this initiative that:
  - Plays a pivotal role in ensuring the seamless continuity of services while establishing standardized practices for the operational facets of IT operations specifically associated with hosting applications for ORD.
  - Creates an environment that serves as a dedicated platform for hosting applications, embodying a comprehensive approach to streamline and enhance the efficiency of IT operations within ORD. By taking charge of the ORD Application Hosting Cloud Initiative, VINCI aims to implement a cohesive framework that not only sustains the delivery of services but also fosters consistency in the operational procedures.

Research Applications

Current projects include:

- **VINCI Abstract:** A chart abstraction web enabled tool that supports several operational and research COVID-19 related initiatives across VHA, specifically helping with manual text annotation and abstraction.

- **VAEDA:** The VA Electronic Determination Aid system that provides VA Internal Review Board (IRB) determinations to VA researchers before they initiate any research project within VA. This is a national initiative project developed for VA Central Office.

- **ProjectFlow:** A clinical trials management system that the VINCI team has supported for several ongoing studies conducted by the MAVERIC team. VINCI ensures the continuity of service and maintains the stability of this application.

Research Applications

- **COSMOS Advanced Analytic Environment:** VINCI has played a pivotal and transformative role in shaping the trajectory of research initiatives that delve into the cutting-edge realm of Artificial Intelligence (AI). At the forefront of this paradigm shift is VINCI's instrumental contribution to providing strategic direction and actively participating in the formulation of the overarching strategy for Advanced Analytics Computing Infrastructure.

- **Advice and Consultation:** In the ever-evolving landscape of AI research, where innovation and breakthroughs are paramount, VINCI's influence has been nothing short of profound. The strategic guidance offered by VINCI goes beyond mere consultation; it represents a strategic compass that navigates the complex terrain of Advanced Analytics Computing Infrastructure. This involvement underscores the strategic acumen and forward-thinking approach that VINCI brings to the table, positioning it as a linchpin in steering research endeavors towards unprecedented heights.

AWS-VAEC Prospect

- Consists of an enclave in the Amazon Web Services (AWS) VA Enterprise Cloud (VAEC) that supports research and offers elastic resources such as storage, compute, software applications, and services which are hard to find elsewhere in VA at scale.

- This project started after ORD recognized the VAEC's potential use to expand research capability and it provides a bridging strategy for a cloud analytic environment until a strategy is implemented for VAEC.

- With latest in infrastructure, platforms, and cloud services, Prospect fills this gap to support various types of research workloads beyond the capacity of on-prem systems.

- The project has continued with ORD controlling its rapid on-boarding process that as of this year, includes over 40 studies in a variety of research topics that have engaged and are in various stages. Many have resulted in publications in major journals or results used to support further grant applications.

VA Data Commons

- Currently there are over 1,000,000 Veterans enrolled in the VA's Million Veteran Program (MVP), which became the largest genomic database in 2016 and is larger than the next three (Veteran or non-Veteran) combined.

- The VA Data Commons is secure, cloud infrastructure that can scale to support thousands of simultaneous genetic studies of this data overcoming current limits of MVP's ability to have concurrent access to only a few dozen projects.

- Its cloud-based infrastructure overcomes current limitations of on-premise storage space and computing power as well as providing limited tools for exploring, visualizing, and analyzing to support conducting approved scientific activities.

- Its permission structure overcomes access barriers to allow investigators to partner with non-VA credential teams who can yield scientific value from the data in accordance with patient data sharing desires.

COVID-19 Shared Data Resource

- Data domain for studying the pandemic comprised of data from various sources and uses case definitions, concept definitions, VA data mappings, and other information that are collected, validated, and updated collaboratively across VA.

- Data contain a wide range of information on all patients who have received a COVID-19 laboratory test within VA or who tested positive outside VA and information pertaining to the positive test was recorded in VA clinical notes.

- Includes timing and nature of test results, and key data elements such as pharmacological and non-pharmacological interventions, patient outcomes, and pre-existing conditions.

- Data elements within the tables are defined using data from the CDW and leverage experiences and data knowledge from across the VA research community.

COVID-19 Shared Data Resource

- Tables Generated:
  - ✓ 28 COVID-19 analytical tables (Patient, symptoms, labs, meds, etc)
  - ✓ 7 Dim/mapping tables (meds, conditions, labs, etc)
  - ✓ 11 Community tables

- Highlights:
  - ✓ Accessible to operational users on VINCI research servers and A01
  - ✓ Supported 70 new research study groups in FY23
  - ✓ Completed 18 COVID DNA requests and 9 COVID feasibility projects and 2 COVID clinical trial recruitment projects
  - ✓ Contributed 299+ phenotype definitions to VA Phenomics Library

Research Studies Supported

Bar chart values: 94, 204, 286, 356
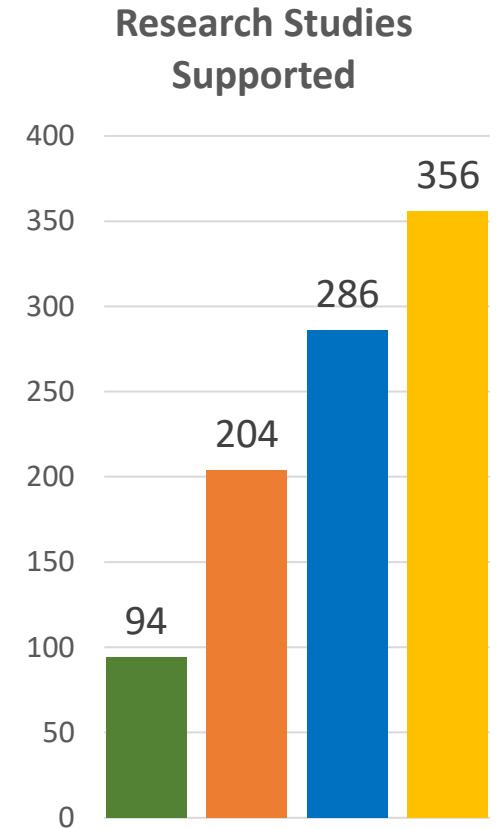
Central Office Interactions

- ORD
- HSRD
- OIT
- CSP
- MVP
- Point-of-Care
- NDS
- RAPID

- CSDE
- CTO
- CDO
- Lab & Pathology
- OEHRM
- ONI
- LGBT
- GRECC

- CMS
- DoD
- Palantir IPO
- Oncology
- Office of Mental Health
- PERC
- Connected Care
- Cerner

Recommendations and Assessing Utility

- VINCI has 2 initiatives to determine customer needs:
    1. A customer survey gathers information on participant's impressions of VINCI and its usage with the intent to improve services and offerings as well as understanding VINCI customers and their needs.
    2. The Field Advisory Committee creates a directed channel for regular feedback and prioritization.

- These new initiatives add to existing channels for interacting with customers such as Meet VINCI visits, listening sessions, satisfaction surveys associated with consulting requests, and other Concierge outreach.

- These activities provide extremely valuable insight of what our customers really want from VINCI. The main themes from the feedback we have received are threefold:
    1. Assistance with and more efficient access and a better understanding of data needed for their projects;
    2. Increased availability and reuse of high-quality, value-added data; and
    3. Increased access to the tools for analysis in a reliable and efficient environment.

Regulatory Initiatives

- Supporting Ethics Principles for Access to and Use of Veteran Data in VINCI data requests

- Developing standard operating procedures for privacy-preserving records linkage requests

- Representing VA on White House Office of Science & Technology Policy Subcommittees

- Open Science

- Equitable Data

- Developing regulatory and technical infrastructures for research reuse repositories

- Serving on ORD IT & Data Governance committees

VINCI Research and Capabilities

- VINCI not only supports others' research with data, applications, services, and infrastructure, but we use those same elements to conduct & publish our own research to support Veterans.

- Research capabilities includes:
  - Population Characterization
  - Health Outcomes Research
  - Expanding available variables
  - Comparative Effectiveness
  - Pharmacovigilance / Safety Studies
  - Health Economics
  - Phenotype Exploration and Validation
  - Advanced Analytics
  - EMR-driven Clinical Trial Recruitment
  - Registry Studies
  - Virtual Registries
  - Point-of-Care Clinical Trials
  - Predictive Models
  - Other Prospective Studies

**# of Publications
(VINCI 1st Author)**

**# of Publications
(w/ VINCI Members)**

FY20 FY21 FY22 FY23
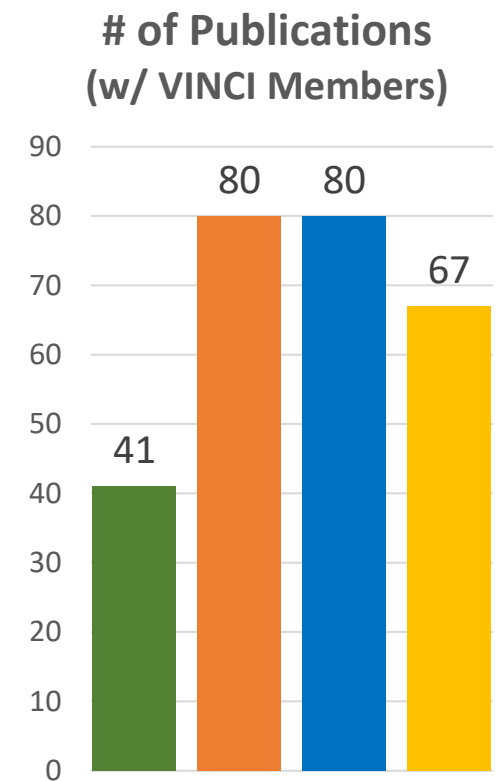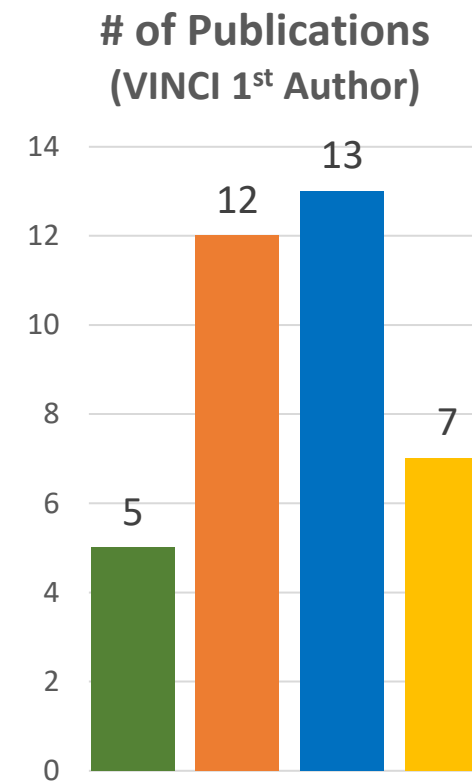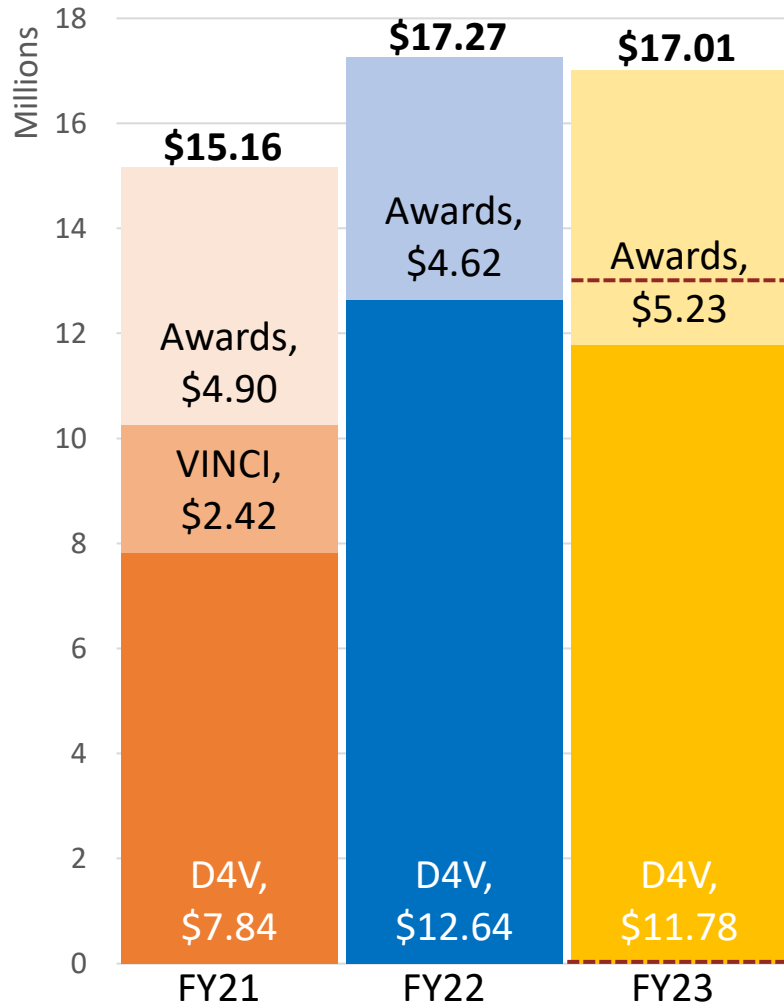
Administration

- Stewardship of financial resources
  - Develops budgets, manages funds, and tracks spending for VINCI and related projects.
  - Advises leadership on financial matters, and is responsible for generating financial reports at project, program, and appropriation level for annual reports.

- Manages human resources (HR) functions
  - Hiring, position classifications, promotions, evaluations, off-boarding and space requirements.

- Assists leadership in management
  - Planning, organizing, and implementing policies and functions across the resource center.

- Supervises staff
  - Includes providing program and project management to teams and for overall portfolio.
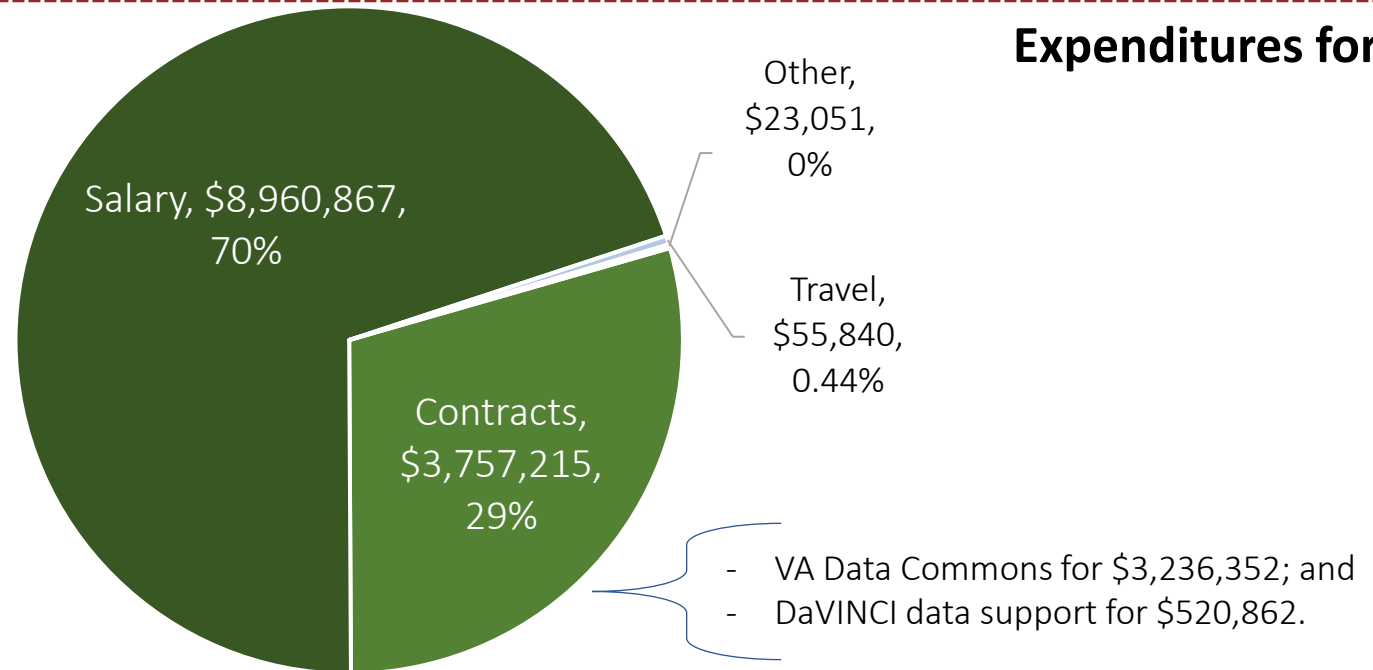
Administration: Funding

Total Budget for FY21 - FY23

Bar chart values:
- FY21: $15.16 (D4V, $7.84; VINCI, $2.42; Awards, $4.90)
- FY22: $17.27 (D4V, $12.64; Awards, $4.62)
- FY23: $17.01 (D4V, $11.78; Awards, $5.23)

- VINCI supports its mission with a combination of VA and Non-VA awards (e.g grants, contracts) and operational funding (D4V) that are mostly spent on effort in the form of salaries and contracts.
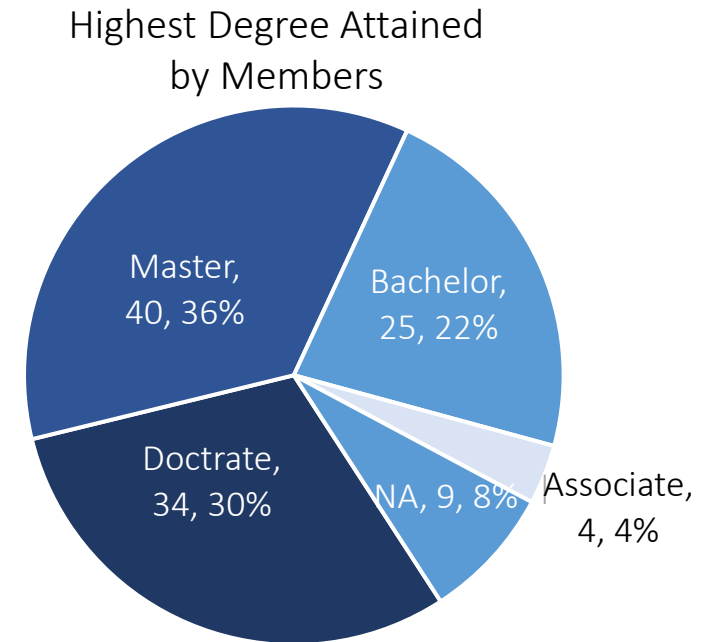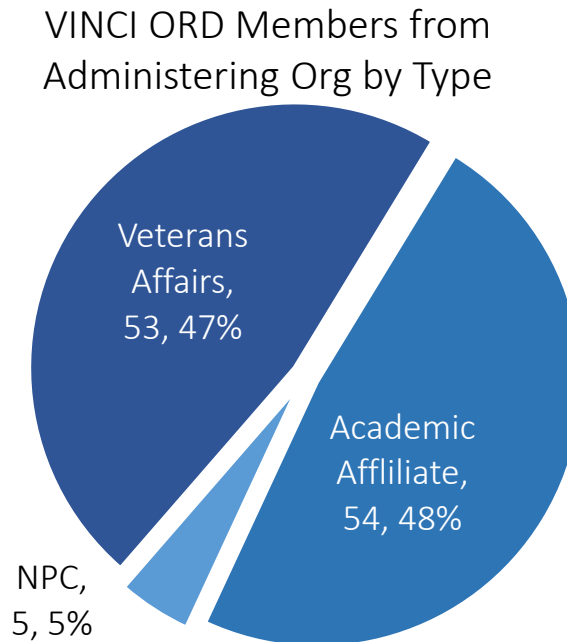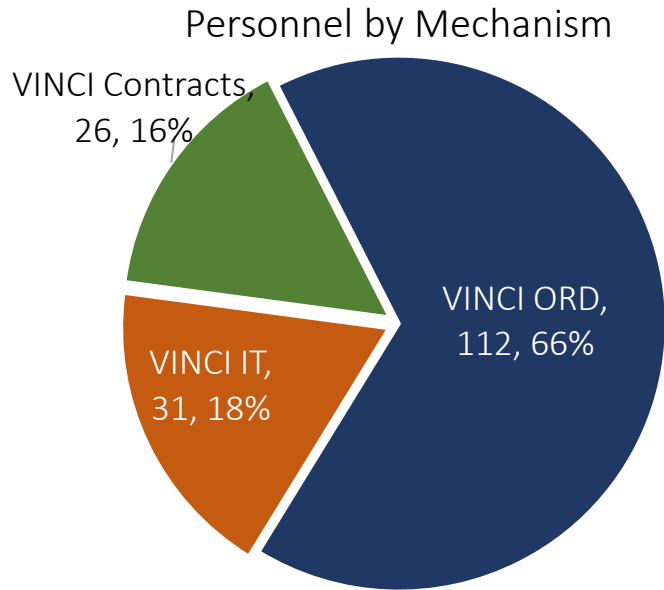
Expenditures for VA Only

Pie chart:
- Salary, $8,960,867, 70%
- Contracts, $3,757,215, 29%
- Other, $23,051, 0%
- Travel, $55,840, 0.44%

- VA Data Commons for $3,236,352; and
- DaVINCI data support for $520,862.

Administration: Demographics

- To fulfill its mission, VINCI organizes these operations with support from highly educated members who are administered by several organizations that provide expertise and skills ranging from information technology, clinical, research, management, and administration.

- This does not include the additional personnel who provide critical support provide through contracts and from VA OI&T.

Personnel by Mechanism

VINCI Contracts, 26, 16%
VINCI ORD, 112, 66%
VINCI IT, 31, 18%

VINCI ORD Members from Administering Org by Type

Veterans Affairs, 53, 47%
Academic Affliliate, 54, 48%
NPC, 5, 5%

Highest Degree Attained by Members

Master, 40, 36%
Bachelor, 25, 22%
Doctrate, 34, 30%
NA, 9, 8%
Associate, 4, 4%

Administration: FY23 Key Personnel Changes

- **Hired (29)**

  - Barna, Samuel R. (Sam)
  - Barter, Rebecca L.
  - Bell, Ruth
  - Boyce, Richard D.
  - Cox, Amy
  - Crook, Jacob L.
  - Danilov, Peter N.
  - Dorn, Chad A.
  - Halls, Corinne (Cori)
  - Heath, Lauren J.
  - Henkels, Rebecca
  - High, Craig A.
  - Jacobsen, Cassandra
  - Jansson, Eva C.
  - Katsoulakis, Evangelia (Eva)

  - King, Jamie N.
  - Knippenberg, Kristin A.
  - LaFleur, Joanne F.
  - Lenoue-Netwon, Michele L.
  - Liang, Xi
  - Madison, Cecilia J. (Celia)
  - Martino, Matthew R. (Matt)
  - Nelson, Tyler
  - Pridgen, Kathryn M. (Katie)
  - Roddy, Lisa M.
  - Schwartz, Samuel (Sam)
  - Smuin, Daniel J.
  - Williams, Roger
  - Yu, Bin (Ben)

- **Resigned (6)**

  - Elize Gatsby
  - Valerie Gonzales
  - Kushan Hewa
  - Sharidan Parr
  - Cristina (Cris) Perez
  - Anee Uphadyaya

Choose VA

U.S. Department of Veterans Affairs