Portfolio Review

Interim Summary

Background

- In 2010, VA sponsored a State of the Art (SOTA) conference on access to care.
- Since that time, VA has invested in a wide array of research and operational projects related to access.
- But we only have a limited understanding of this body of recent and ongoing work.
- Better understanding this body of work could inform future funding priorities and identify opportunities for partnered research in this domain.

Background

Purpose

- To understand the body of ongoing access-related work in VHA

Objectives

- To review and synthesize recent and ongoing VA-funded research projects and operational initiatives focused on access to care (the "**access portfolio**")
- To identify gaps in the existing research portfolio to inform priorities for future research
- To inform partners about access-related interventions that may be ready for implementation

Methods

- We conducted an environmental scan of current/recent VA research and operational projects focused on access to care.

- We collected data in two ways:
  1) text analysis of VA and National Library of Medicine websites
  2) structured interviews with operational partners

- We developed and refined a rubric to categorize these projects.

- The rubric incorporated descriptive components (e.g., clinical care setting, general research methodology) as well as elements specific to access (e.g., Fortney model domain, operational priority area).

Rubric

Underlying questions

- Did the project have an access impact?
- On what clinical domain was the project focused?
- Did the project directly measure access? If so, was the measure administrative or self-reported?
- What barriers to access did the project address (using Fortney's model)?
- Was the project observational, interventional, or program evaluation?
- Did the project align with any of the research priority areas of the Office of Veterans Access to Care (OVAC)?
- Did the project have a specific focus on the MISSION Act or Choice Act?
- Did the project involve the use or evaluation of virtual care?
- Did the project use a non-VA care dataset?
- Are the deliverables ready for implementation?

Rubric

Code 1: Care Setting	
Type	Definition
Primary Care	The project is focused on primary care access
Specialty Care	The project is focused on specialty care access
Mental Health	The project is focused on mental health access
Inpatient Care	The project is focused on inpatient care
Long-term care	The project is focused on long-term care

Code 2: Access specific versus Access relevant	
Type	Definition
Access Specific	A project that has an access impact and measures actual or perceived access
Access Relevant	A project that has an access impact, but does not incorporate specific measurements of access

Code 3: Access Distinction	
Access Distinction	Definition
Actual Access	The direct measurement of access, usually through administrative data (i.e., EHR, GIS, etc.)
Perceived Access	The measurement of perceived access, typically through self-report (e.g., SHEP, V-Signals)

Code 4: Fortney Model Classification Tags	
Tag(s)	Definition
Geographical	The primary purpose of the project is to address the difficulties of traveling to healthcare provider locations.
Temporal	Wait time due to appointment availability.
Cultural	The project evaluates the acceptability of health services in terms of comfort with, trust in, or preference for their medical provider.
Digital	The connectivity that enables synchronous or asynchronous digital communications with formal providers, informal caregivers, peers, and computerized health applications
Financial	The healthcare system eligibility issues and the cost of utilizing healthcare services

Code 5: Project Type	
Type	Overall Grant Product
Observational	Secondary data analysis, mixed methods, qualitative methods, modeling
Program Evaluation	Evaluation of a programmatic initiative designed to improve access
Interventional	Prospective evaluation of an intervention designed to improve access

Code 6: OVAC priorities (select primary and secondary themes)
Priority
System redesign (PACT integration, clinical delegation, MISSION)
Overuse/low-value care/appropriateness
Prioritization/urgency/wait list management
Virtual care/technology
Burnout
Workforce satisfaction/retention/expansion
Clinical Operations
Access measurement
Improving patient satisfaction/experience (must have product to address satisfaction)

Code 7: Community Care and Virtual Care Focus (select one, both, or note neither)	
Type	Definition
MISSION Act	Is there a MISSION Act focus to the project?
Veterans Choice Act/Program	Is there a Veterans Choice Act/Program focus to the project?
Non-VA Care/Data	Is there a non-VA care dataset affiliated with the project unrelated to MISSION or Choice?
Virtual Care	Does the study involve the use or evaluation of virtual care programs such as secure messaging, mobile apps, telehealth, virtual video connect, etc.)?

Code 8: Implementation Status (select one)	
Implementation not part of objectives	Does not address implementation in objectives
Pre-Implementation	The stage prior to the main implementation - this can include small implementation pilots or demonstration/feasibility projects.
Implementation	Stage at which implementation has occurred

Code 9: Non-VA data
What non-VA data were considered and/or utilized? Distinguish between (a) non-VA data available through the VA such as community care, CMS data linked to Veterans, fee-basis, and (b) data obtained by an external entity (e.g., Blue Cross, Kaiser, University hospital system, etc.)

Methods

- Each project was independently reviewed and coded by two study team members.
- Discrepancies were resolved by consensus.
- 20% of projects will be reviewed by investigators, with a specific focus on those with discrepancies (review was ongoing at the time this document was developed).

Methods

We **included** projects that:

- ✓ Were funded or supported by a VA office
- ✓ Were active from January 2015 – July 2020
- ✓ Directly or indirectly impacted access to care
- ✓ Provided some level of detail for scope of work greater than project title (i.e., abstract, specific aims, etc.)

Methods: Web-based Search

Data Collection

- We used “web-scraping” to rapidly and reproducibly extract large numbers of grants/abstracts from publicly available websites.
- A text analytic toolkit was locally developed, validated, and then used to identify potentially relevant abstracts for further, detailed review.

Data Sources

- VA: HSR&D, CSR&D, RR&R, BLR&D, and QUERI websites
- NIH ExPORTER: Provides detailed information for federally-funded research for all federal agencies
- Clinicaltrials.gov

Methods: Operational Interviews

- We developed a brief interview guide to elicit information about recent and ongoing access work from operational offices.
- The guide was refined with input from the Ann Arbor COIN Qualitative Core and the Access CORE team.
- Interviewers were trained to use the guide through mock interviews prior to application.

Methods: Operational Interviews

- We then developed a list of all operational offices that may have engaged in access-related work, refining this list with input from Access CORE team members.
- Email invitations were sent to operations office leads to request a brief interview, with the goal of identifying operationally-funded or supported work that was unlikely to be publicly reported.
- Source documents were also obtained when possible.
- Interviews were audio recorded, but not transcribed.
- Projects were coded according to the rubric, similar to research projects.

What is web scraping?

- **Definition**
  - Web scraping, web harvesting, or web data extraction is the process of extracting data from websites into a structured form
- **Common Example**
  - Search engines will **scrape webpages** with the intention of creating visibility to content

What can web scraping do?

- It can create an automated approach for extracting unstructured web data and processing it into a more structured form
- It allows the ability to systematically open URLs and sub-URLs to extract information
- For example, a web scraping tool could scan an online forum and create a dataset by extracting all subject names, messages, individual user IDs, and dates of messages
    - Normally, this would be too difficult and time consuming for staff to compile all of this information into a dataset

How does web scraping work?

- The algorithm parses webpage code based on specific parameters and extracts data

- For example, if we wanted to obtain a dataset of all abstracts on the **Current HSR&D studies and Projects** webpage, we would find that all of the links to abstract pages have the same substring: "*https://www.hsrd.research.va.gov/research/abstracts.cfm?Project_ID=*"

- We would then generate a list of all URLs that contain this substring, export them into a list, loop a web scraper through each URL, and extract the abstract/project information into a dataset

What application is needed?

- We created our web scraper in Python using the Beautiful Soup Module.
- However, other applications have packages for web scraping such as R, Matlab, Java, and SAS
- There are also apps, browser add-ons, and private services that can conduct web scraping for a fee

Methods: Web Scraping and Text Analysis

- Our approach involved 3 steps:
    1) Web scraping algorithm development
    2) Algorithm deployment and data cleaning
    3) Data analysis
- Our goal was to develop an algorithm that was sensitive and reasonably specific for identifying access-related projects

Step 1: Algorithm Development

- We first identified keywords and phrases associated with access to care in PubMed
  - Our team searched PubMed for projects with the MeSH term “Healthcare Accessibility” to identify consistently occurring keywords and phrases
- However, just using these terms was quite non-specific
- Therefore, we added additional terms related to types of providers, care type, and healthcare settings
- Two investigators manually reviewed 50 abstracts that were identified with the final algorithm to determine if they were access-related, serving as the gold standard

Step 1: Algorithm Development

- Following the manual review, we extracted simple terms or phrases that were highly specific for whether a project was access-related
- To do this, we compared terms/phrases in the gold standard access abstract dataset to a group of abstracts that were clearly not access-related
- We called these newly extracted terms/phrases “safe access terms”
- These were subsequently used to enhance the sensitivity of our overall algorithm

Step 2: Data Cleaning

Sentence Tokenization

- The first thing we did to clean our dataset is separate each abstract into individual sentences, also called sentence tokenization or separation
- To do this, we used Regular Expressions to separate abstract sentences
  - A Regular Expression is a sequence of characters that define a search pattern
- Sentence tokenization is important because the words that occur within a sentence are part of a phrase and thus within the same context
- This is key for our scoring tool that counts the close proximity of access-related terms in sentences (which we will discuss later)

Step 2: Data Cleaning

Sentence Tokenization Example

- Example: “Wait time metrics are a valuable tool to **improve** appointment **access**. **Care** and precision should be at the forefront of metric development.”
- The context of the 2nd sentence is completely different than the 1st sentence
- If we looked at the close proximity of “access” and “care” from the lens of the entire abstract, a keyword searching algorithm would likely score this as a positive

Step 2: Data Cleaning

Lemmatization

- Each word from our 3 lists as well as from each abstract were lemmatized or modified to remove inflectional endings only and to return the base or dictionary form of a word
  - Example: “Accessibility,” “Accessible,” “Accessing”, “Accesses” = Access
  - This ensured that our text analytic tool did not miss potentially access-related work because of different word tenses

Step 2: Data Cleaning

Stop Words

- We removed stop words, or common words which would appear to be of little value for identifying access projects, from abstracts
- Stop word examples include “to,” “of”, “a”, “in”, “this”, etc.
- This would treat “access to care,” “accessibility of care,” and “accessible care” as the same

Step 3: Data Analysis

Scoring

- We used four methods to score potentially access-related projects
  - Proximity Score: We looped through our list of providers, care types, and healthcare settings to see if multiple terms within 5 words of each other in a sentence.
    - Example: **Subspecialty care** burnout may have serious **access** implications.
  - Title Search: We used our safe list of access terms to scan project titles
  - Topic Generation Model: We used our safe list of access terms to detect matches of terms derived from the TextRank topic generating model
  - Safe List Counts: we ran a frequency count for the number of safe list of access terms that occurred in an abstract

Step 3: Data Analysis

Classification

- We found that scored abstracts typically fell into three categories
  - Reliably access-related – no manual review of abstract required (total score of ≥ 10)
  - Possibly access-related – manual review of abstract required (total score of 1-9)
  - Unlikely access-related – no manual review of abstract required (total score of 0)

Summary of Portfolio Review: N=266 Projects

Web-based review

211
total projects identified

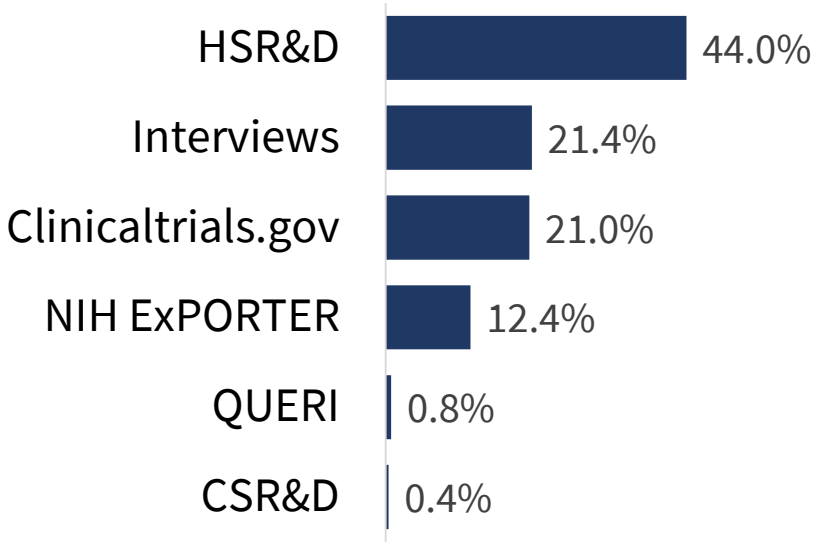
Operational Interviews

11 Interviews completed

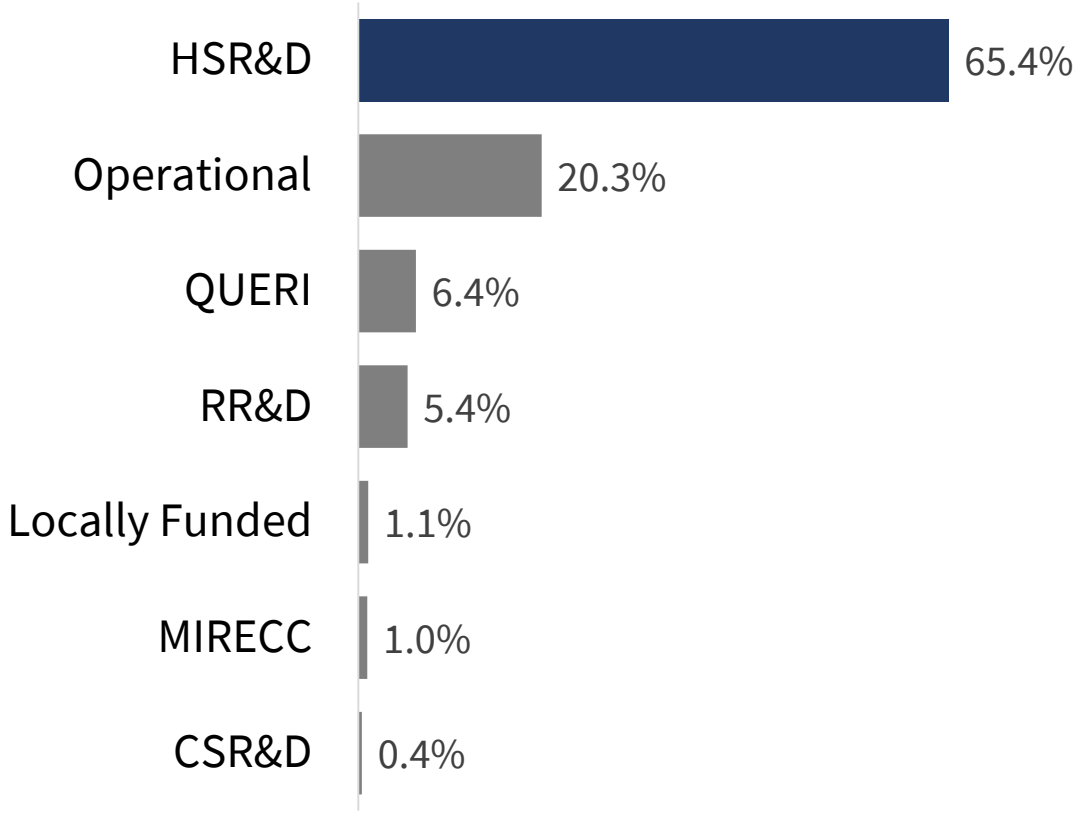
55 Operational projects identified

Summary of Portfolio Review

A variety of **data sources** were used to identify projects.

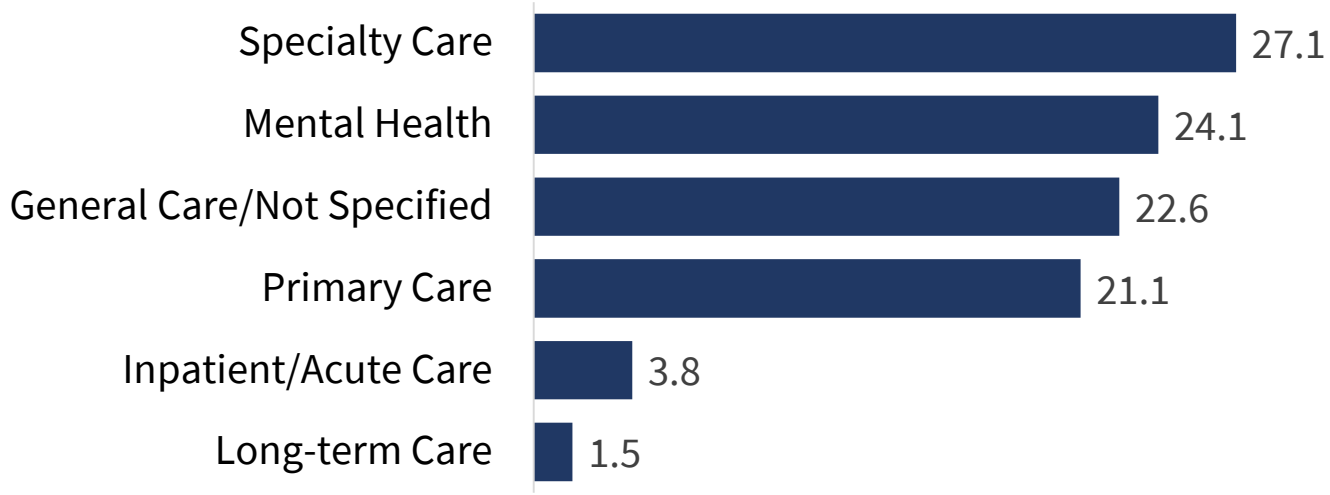


HSR&D was the most common **source of funding**.

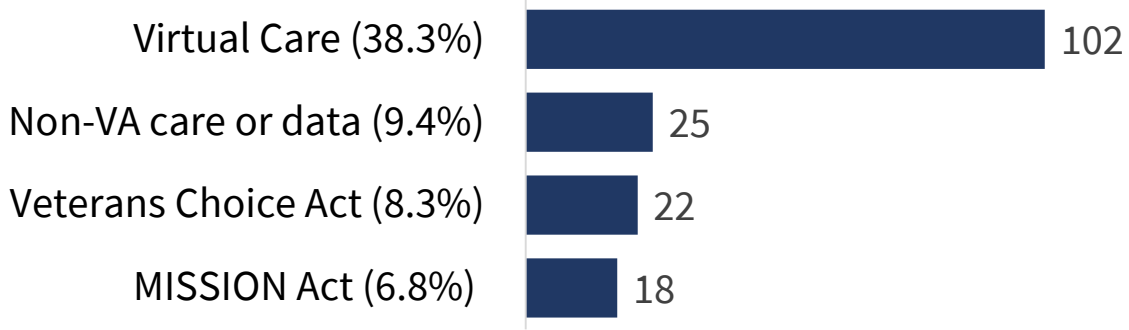


Results – Healthcare Environment & Implementation

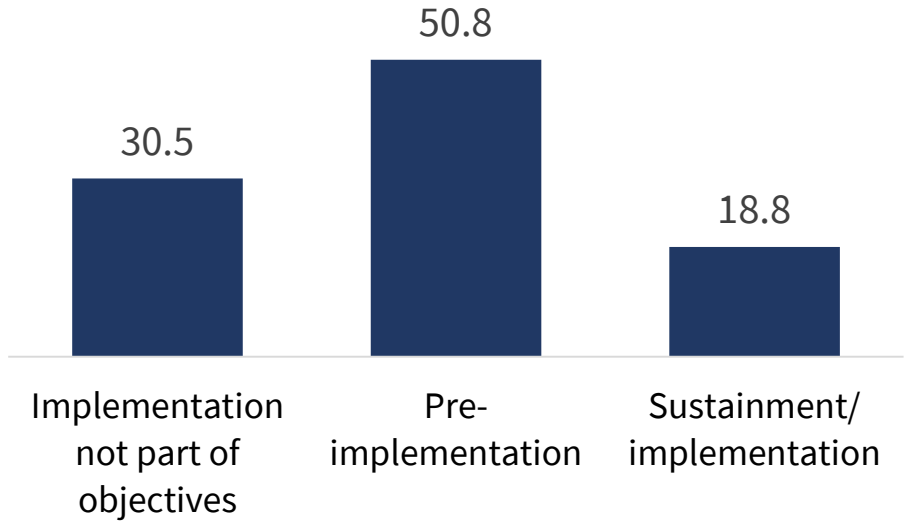
Clinical settings of projects



Projects related to virtual care or non-VA care



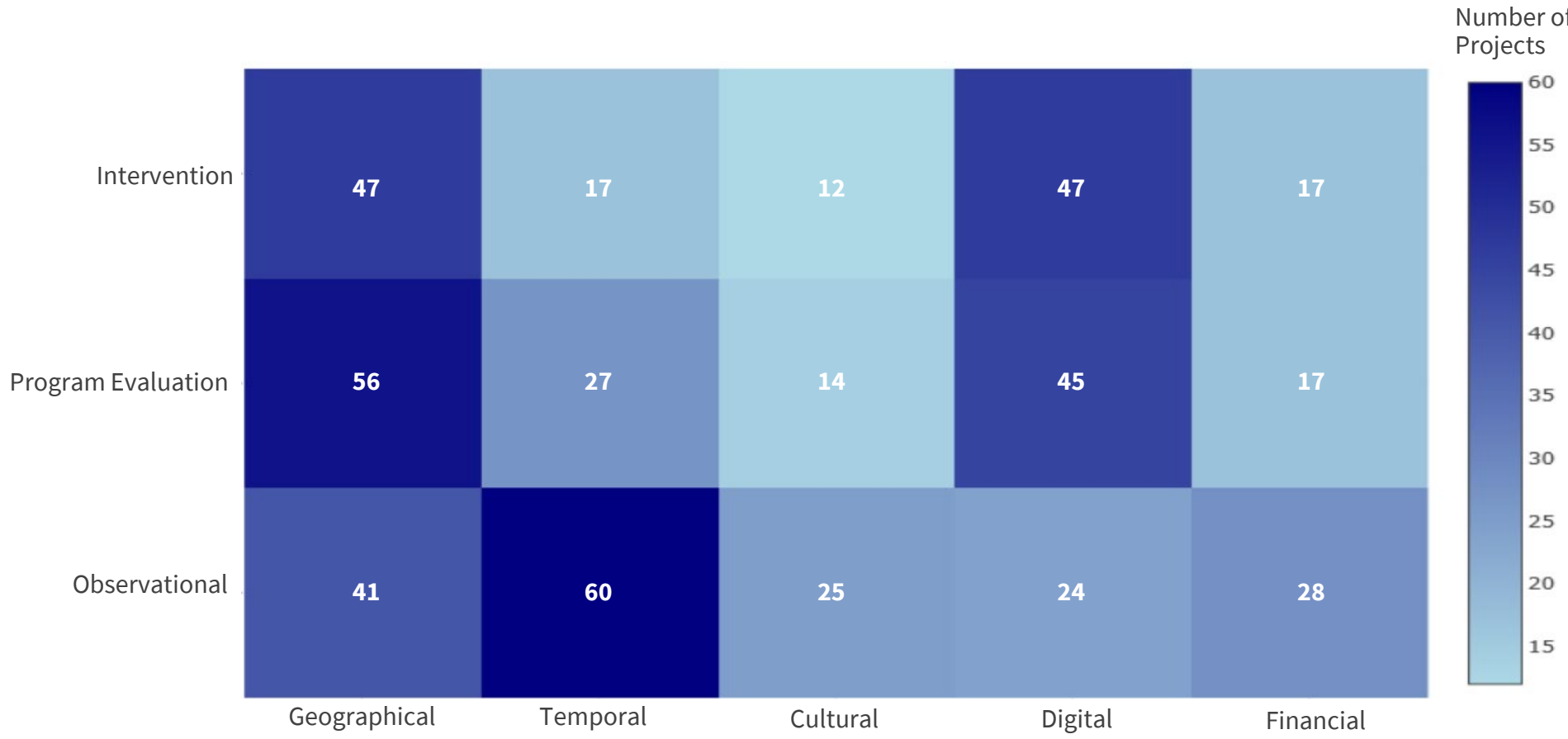
Implementation stage at end of projects



Results – Fortney Model Dimensions

Relatively few projects were **interventions**, but those that were interventions focused on **geographical** and **digital** barriers to access.

Few projects focused on cultural or financial barriers to access.



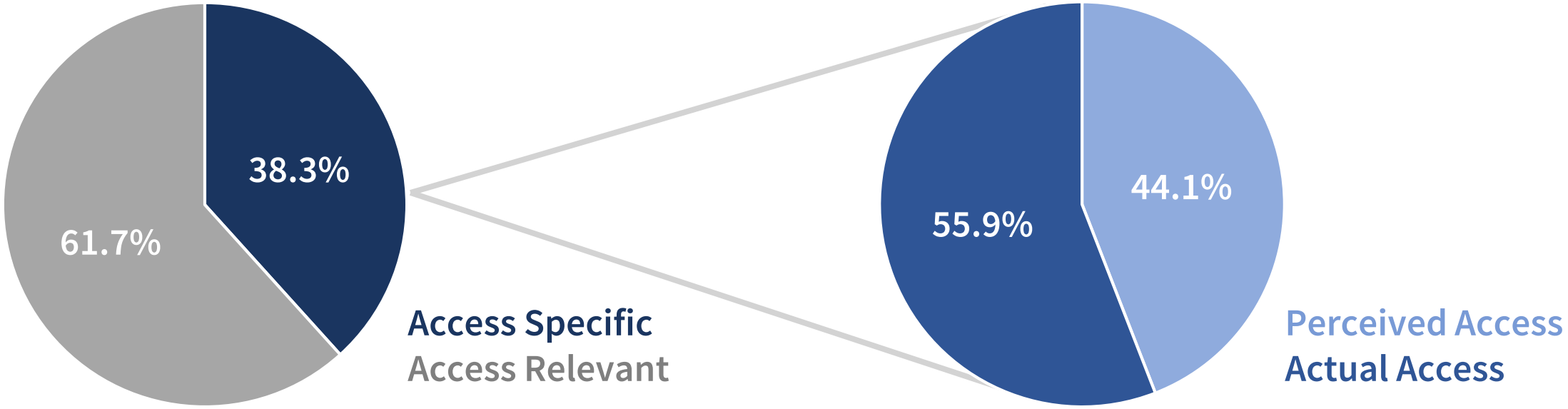
Results – Access Focus

About half of access-related projects **directly measured access** (were access specific). N=266

Access specific projects measure actual or perceived access, whereas *access relevant* projects do not incorporate specific measurements of access.

More projects tended to directly measure **actual patient access** as opposed to **patient perceptions** of access. N=102

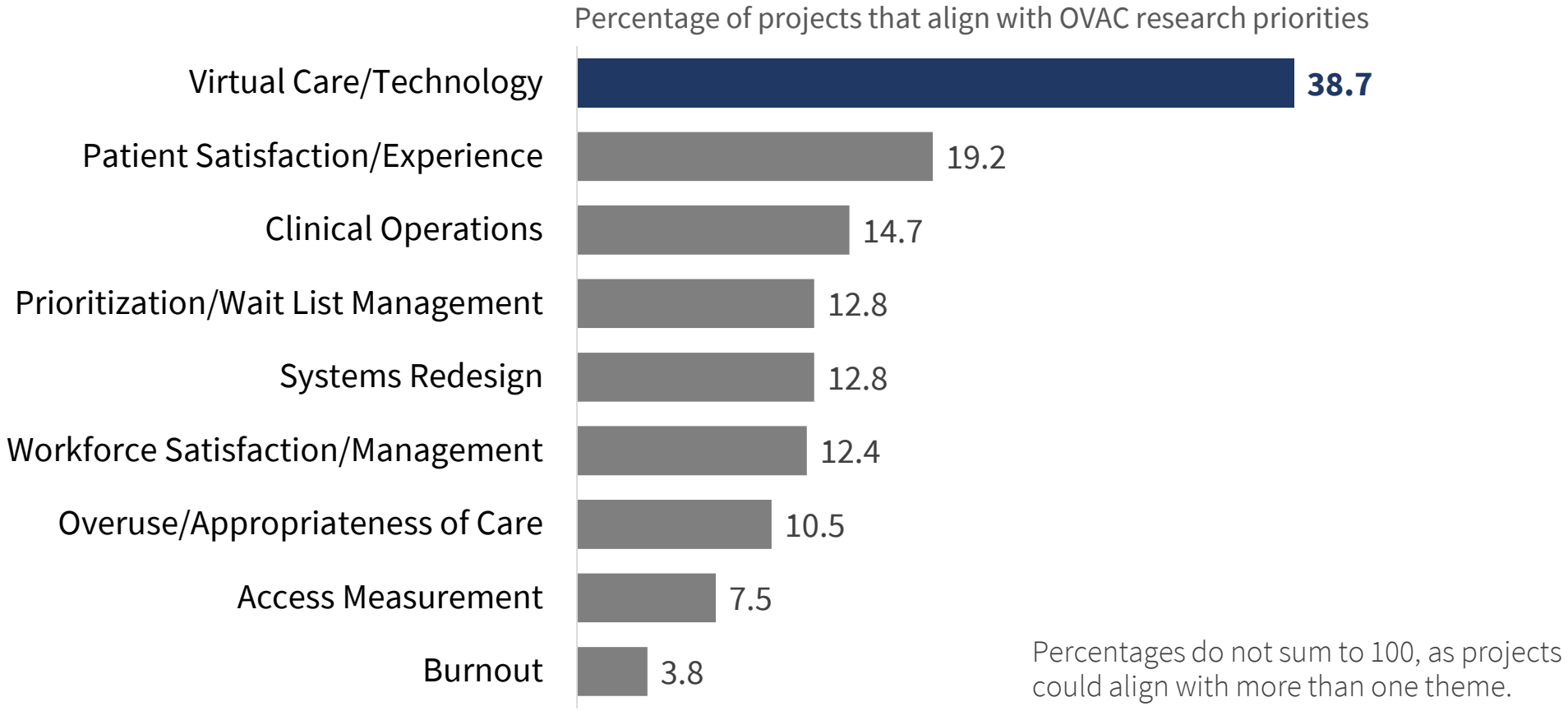
Access type (i.e., actual or perceived access) are subsets of *access specific*.



Results – Research Priority Areas

Of projects aligned with **OVAC priorities** (N=237), **virtual care and technology** was the most common area of alignment.

Patient experience and satisfaction and clinical operations were also among the more common aligning themes.



Percentages do not sum to 100, as projects could align with more than one theme.

Summary

- Over the last 5 years, VA has developed a robust access portfolio with research and operational work across clinical domains.
- Much of the intervention / evaluation work has focused on digital and geographical barriers.
- A substantial proportion of interventions show promise and engage in pre-implementation work, but fail to be translated / operationalized.
- Opportunities for translation of research to impact in virtual care space.
- Need for more research related to overuse, access measurement, burnout / workforce satisfaction.

Acknowledgements

- Special thanks to:
  - Our VARC colleagues in Bedford, Denver, and Iowa City
  - The Portfolio Review Co-Investigators
    - Megan Adams, MD, JD
    - Tanner Caverly, MD, MPH
    - Erika Sears, MD, MS
  - The Portfolio Review Project Staff
    - Jennifer Caldararo, MSW
    - Joan Gargaro, MPH, MBA
    - Pia Roman, MS