How Public Data Sets Can Address Data Obstacles in Pay for Success

A Demonstration for Veterans With Service-Connected Disabilities

MAY 2019

Burhan Ogut | Garima Siwach
American Institutes for Research
Brian Beachkofski
Third Sector Capital Partners Inc.

How Public Data Sets Can Address Data Obstacles in Pay for Success

A Demonstration for Veterans With Service-Connected Disabilities

MAY 2019

Burhan Ogut | Garima Siwach
American Institutes for Research
Brian Beachkofski
Third Sector Capital Partners Inc.

AMERICAN INSTITUTES FOR RESEARCH®
1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

THIRD SECTOR
225 Franklin Street, Suite 350
Boston, MA 02110
617-912-8957

www.air.org

7116_01/19

Contents

Introduction

Pay for Success (PFS) offers an innovative approach for improving essential social services while doing important social science research. This approach allows policymakers, government agencies, and service providers to drive improved results by paying for achieved outcomes rather than project inputs (i.e., costs) or outputs (e.g., certificates). With a reliance on outcomes, data play an important role in all phases of a PFS project. This brief demonstrates how data obstacles, especially at the early stages of a project, can be overcome by combining information from different sources.

Having access to a range of data sources is essential for PFS and outcomes-oriented projects to describe the needs of the population of interest, determine baseline outcome levels, and assess how such projects might affect those served. However, access to data is not always straightforward. Collecting new data requires significant investment of time and resources, which might not only cause significant delays in the early stages of project planning and development but also prevent some potentially transformative PFS efforts from even launching.

One solution is to use publicly available data, which comes with its own challenges, including the unavailability of the following:

- Variables that align well with the eligibility criteria
- Important outcomes necessary to quantify prior performance and potential improvements
- Identifiers to allow linking across systems and sources

In 2016, the American Institutes for Research (AIR) and Third Sector Capital Partners Inc. (Third Sector) worked with local social service providers in the San Diego, California, area on a project involving veterans. The idea was to connect veterans with service-connected disabilities (SCDs) to resources that would support their employment and re-integration into the community. Third Sector managed the project, and AIR served as the evaluator. In the process of assessing and building this project, AIR developed an important data-related method that could benefit other organizations pursuing outcomes-oriented projects. This method addressed two challenges:

- Detailing beneficiary population eligibility and selection criteria using publicly available data
- Merging publicly available data sets to expand outcome domains available for historical baselining

This brief introduces the PFS approach and summarizes the methodological strategies developed for the project. It also discusses how future work could take advantage of these approaches.

Growing an Outcomes-Oriented Social Sector

More government funding awards are explicitly tied to the achievement of meaningful, long-term outcomes. PFS and Social Impact Bonds (SIB)[1] are the first iteration of this approach, with numerous successful projects launched to date, including Third Sector's Massachusetts Juvenile Justice and Salt Lake County PFS initiatives.[2] In addition, AIR has several PFS projects, including the National Capital Region PFS Demonstration Project on Permanent Supportive Housing and a PFS feasibility study on English Language Acquisition by the third grade.[3] Although the SIB structure and outside financing continue to be available tools, officials increasingly are looking for simpler ways to align incentives and tie the provision of services to achieving meaningful outcomes. Third Sector has helped launch multiple projects using this approach, including a countywide effort in King County, Washington.[4]

To be successful, outcomes-oriented social services projects require a handful of preconditions:

- Available funding streams that can be tied to outcomes
- A service provider that agrees with the concept
- Political will to transform social services procurement processes
- Data that can provide an understanding of potential project beneficiaries and measure both short- and long-term outcomes

Outcomes-oriented projects also require a range of actors:

- A government agency
- Outcome funders
- A service provider or providers
- An organization to measure, validate, or evaluate data (can be internal or external to agencies, funders, or providers)
- A project manager or technical assistance provider, particularly for initial transformation efforts (an intermediary)

---

[1] For a comprehensive introduction to PFS, see https://www.thirdsectorcap.org/wp-content/uploads/2016/06/Third-Sector-Capital-Partners-Intro_PFS-Overview2016.pdf.
[2] For a list of PFS projects by Third Sector, see https://www.thirdsectorcap.org/projects/.
[3] For a list of PFS projects by AIR, see https://www.air.org/resource/pay-success-social-impact-bonds.
[4] The PFS field continues to evolve in the United States; see *The Next Phase of Pay for Success: Driving Public Sector Outcomes* at https://www.thirdsectorcap.org/blog/the-next-phase-of-pay-for-success-driving-public-sector-outcomes/.

Outcomes-oriented projects have three unique phases:

- **Assess,** also referred to as feasibility, is an assessment of current capabilities, desired capabilities, and the gap between these conditions.
- **Build,** also referred to as transaction structuring, is the negotiation and launch of an outcomes contract after the partners agree on the design of the project.
- **Implement** includes the unfolding of the project and the optimization of the operations by engaging stakeholders to continuously improve processes to deliver better outcomes across time.

Data's Role in Outcomes-Oriented Social Services Projects

Data access and measurement remain at the heart of any PFS or outcomes-oriented project, and access to timely and relevant data often is the most critical hurdle faced. The role of data in outcomes-oriented projects goes beyond measuring outcomes in a standard evaluation approach. Outcomes-oriented projects hinge on the ability to use data to define what it means to be successful, what the eligibility is for intended beneficiaries, and how to determine ongoing funding.

During the feasibility phase, access to and the use of both primary and secondary data are essential for assessing the possibility of a PFS project.

Primary data, which is the focus of this brief, are used to detail a specific population's characteristics, an unmet need, and how an outcomes-oriented project might affect population members. For example, a project might aim to address housing needs, and primary data could be used to identify frequent users of public housing options.

Secondary data refer to results from studies designed to inform evidence-based practices. Once the target population is identified, the decision on what services or interventions to deliver relies on the evidence of impact (secondary data) of such interventions.

The subsections that follow discuss the use of primary data for determining success metrics, evaluating success measures, and evaluating the impact of the intervention.

Baselining Existing Levels of Success

Without an understanding of the current performance level of the target population, it is difficult to assess the overall impact of a potential project. In projects where incentives are tied to long-term outcomes, baseline data allow a group of actors to establish whether measurable and meaningful change has occurred. In short, baseline data help justify a specific intervention and inform realistic goals and measurements for success. A benchmark assesses the relative performance of an eligible population, and a baseline allows for the measurement of important performance indicators for future understanding and evaluation based on historical evidence.

Traditional impact evaluations are limited because the counterfactual (what would have happened in the absence of an intervention) cannot be easily observed and quantified. Using relevant baseline data means that an outcomes-oriented project is more likely to address which measures and proposed interventions will have the greatest likelihood for change within the targeted population and, therefore, determines the best return on investment.

Evaluating Success Measures

Consistent and reliable data also are necessary to evaluate the success of a project's outcomes. For example, a labor market-based intervention on a population with employment challenges may decide that the outcome will be the average increase in employment rates for the population. In this case, a benchmark is needed to define whether the outcome was successful. Such a benchmark could be a predetermined constant or one based on observed employment rates of groups other than the population the project targeted. Payment for the project is structured based on these success measures.

Generally, four success payment calculation designs are used in PFS projects:

- **Rate Card Method.** Success is calculated with respect to a predetermined measure.
- **Historical Baseline Method.** Success metrics are computed using the historical records for the intervention group or a similar cohort.
- **Matching Method.** Techniques such as propensity score matching are chosen for a comparison group, and then the outcomes of the intervention group are evaluated against the comparison group.
- **Randomized Controlled Trials (RCTs).** As the most rigorous success calculation method, participants are randomly assigned into treatment or control conditions, and the difference between the two groups is the success metric. RCTs are the most resource-intensive method for evaluating success, partly because of the cost of data collection from a comparison group of participants.

Evaluating an Intervention's Effectiveness

In addition to determining success, a PFS project also evaluates the effectiveness of the intervention used. An evaluation plan focuses on not only existing data but also the means and designs to collect new data as the project is implemented. An evaluation plan also should collect data on a control group to assess the impact of the current program. Whether an RCT or a quasi-experimental design is used, data on both the intervention and comparison groups must be measured and collected consistently. Because the focus of this brief is on the use of data for success measures, evaluation of an intervention's effectiveness will not be discussed in detail.

Potential Challenges—and a Proposed Solution

In PFS projects, multiple challenges may arise when collecting and using data. Collecting data about the population of interest is time and resource intensive. In addition, if data do not exist, then they must be gathered before the PFS project can begin. Data collection is expensive and could drain limited funding during the assess and build phases. In some cases, it may be difficult to access data that contains personally identifiable information because of the extensive protections on it. In addition, when monitoring outcomes across time, the continuity of data may present a challenge if it becomes unavailable when needed later. Although it is possible to clear these hurdles, this requires time and effort, which can delay or impede the start of a PFS project.

One solution is to use publicly available, high-quality data sets. There are advantages to these data sets, such as that access does not require special use agreements or contracts. Further, these data sets, mostly used for research purposes, often include information that is helpful to identify and baseline the population of interest.

At the same time, three caveats should be considered when using public data sets:

- The data may not have the requisite variables that align with the eligibility criteria defined for the target population.
- Data with appropriate variables for the eligibility criteria may be missing important outcomes relevant for the historical baseline and impact analysis.
- Nonadministrative data, or data from disparate sources, are unlikely to have identifiers to allow for linking across systems.

The next section describes how a publicly available data set was used in a PFS study for a population where data typically are difficult to access.

Solving PFS Data Challenges Using Publicly Available Data

Project (re)Launch

The San Diego Veterans Employment PFS Initiative, titled Project (re)Launch, was a partnership that included Third Sector, AIR, and local social service agencies. Funding was provided by a Social Innovation Fund transaction structuring grant from the Nonprofit Finance Fund. The goal of Project (re)Launch was to improve employment outcomes and economic security for veterans with SCDs through intensive case management and wraparound supports while receiving vocational rehabilitation services. To this end, the program targeted veterans with SCDs who were under- or unemployed in the greater San Diego area and were eligible to receive services under the Vocational Rehabilitation and Education (VR&E) program of the U.S. Department of Veterans Affairs (VA).

The evaluation of Project (re)Launch aimed to examine how the program affected economic security and health outcomes for veterans with SCDs. During the feasibility (i.e., assess) phase, an important consideration for evaluation included the availability of data on economic and demographic characteristics of veterans with SCDs in San Diego, as well as baseline measures on their health and economic security. Like any other PFS project, the availability of and access to these data were essential for the impact evaluation and success metrics computations. However, limited information existed on the target population.

The key selection factor for Project (re)Launch was eligibility for VR&E services, but data were not available on the demographic, economic, or social characteristics of the actual participants in the study. In a broader sense, the target population for the project was veterans with SCDs who were eligible for VR&E services. Using this broad definition of the target population allowed the selection of the population of interest from existing publicly available data.

According to the VA, veterans are eligible for VR&E if they

- were discharged from the military within 12 years[5] with a charge other than dishonorable, and
- have at least a 20% SCD rating with an employment handicap or a 10% or more rating with a serious employment related handicap.[6]

In the most current report on VR&E by the VA (Economic Systems Inc., 2018), VR&E recipients—averaged across fiscal years 2010, 2012, and 2014—were more likely to have a serious employment

[5] Or the date that the VA first notified the veteran of an SCD rating.
[6] For more information, see https://www.va.gov/opa/publications/benefits_book/benefits_chap03.asp.

handicap as determined by a vocational rehabilitation counselor and more likely to have served during the Gulf War II era, compared with the general veteran population (Figure 1).

To construct a comparison group for historical baselines for Project (re)Launch, AIR replicated the VR&E eligibility criteria using the information available in that data set. Using these criteria, we created a potential VR&E-eligible sample from existing data sources. Because we did not know whether these potentially VR&E eligible individuals actually received VR&E services, we used information from the VA on actual VR&E recipients to examine whether the potential VR&E-eligible sample that we created looked similar to actual VR&E recipients (Economic Systems Inc., 2018; see Figure 1). In addition, two public data sets were identified to examine employment- and health-related outcomes for this potentially VR&E-eligible population: the American Community Survey (ACS) and the Behavioral Risk Factor Surveillance System (BRFSS).

- ACS is an ongoing survey by the U.S. Census Bureau that provides data every year on the social, economic, demographic, and housing characteristics of the U.S. population.
- BRFSS is sponsored by multiple divisions within the Centers for Disease Control and Prevention and other federal agencies.[7] The BRFSS is a system of nationwide surveys that collects state data about U.S. residents' health-related risk behaviors, chronic health conditions, and use of preventive services.

Baselining Existing Levels of Success

To describe the baseline outcomes for potential participants, data were needed for a similar population in the target area before Project (re)Launch began. The most robust study design would involve access to administrative data from multiple sources, linked together to generate a comprehensive socioeconomic profile of a subsample of veterans in the target geography.

Access to such administrative data is rare, even more so in the feasibility phase of a PFS project. To address the challenge of data accessibility, AIR developed a method that used publicly available data, along with prepublished information on VR&E recipients, to help construct a pseudo-administrative data set to assess the baseline employment and health outcomes of potential project participants. To measure these outcomes, we used ACS data that included information on veterans with SCDs. Using VR&E eligibility criteria from the VA, we determined a potential intervention sample of veterans with SCDs who may have been VR&E eligible.

[7] See https://www.cdc.gov/brfss/about/index.htm for more information.

Although data specific to California were used for Project (re)Launch, for the purposes of this brief and general application of these methods, we analyzed and reported national data. (See the appendix for a detailed description of the method.) We analyzed national-level data from ACS 2011–2015 and restricted the data to 133,985 veterans who were on active duty after 2001. This ensured that we analyzed data only for veterans within 12 years of service discharge, which is a criterion for VR&E eligibility.[8] To construct a narrower sample of veterans who would be potentially representative of the target population, we classified individuals as potentially VR&E eligible if they met both of the following conditions:

- An SCD rating of 10% or higher
- At least one disability (hearing, vision, self-care, independent living, ambulatory, or cognitive difficulty)

The second condition was added to imitate an employment-related handicap, assuming that difficulty in any one of these functions could lead to capability limitation among veterans who also had an SCD. A sample of 11,087 veterans in the ACS data was found to be potentially VR&E eligible. (See Figure 1 for the characteristics of this sample.)

In addition to the ACS, which provides employment-related data, we examined the BRFSS data to identify important health-related outcomes relevant for the historical baseline and impact analysis. However, the BRFSS does not report any information on SCDs or period of service. To address this challenge, we employed a prediction model to create a pseudo-administrative data set using common information that is available in both BRFSS and ACS data. Using the ACS data, we first predicted the likelihood of VR&E eligibility by using variables common between the ACS and BRFSS. We then applied the resulting parameters from this model to predict VR&E eligibility in the BRFSS data. Three minimum conditions had to be met for this prediction model to work:

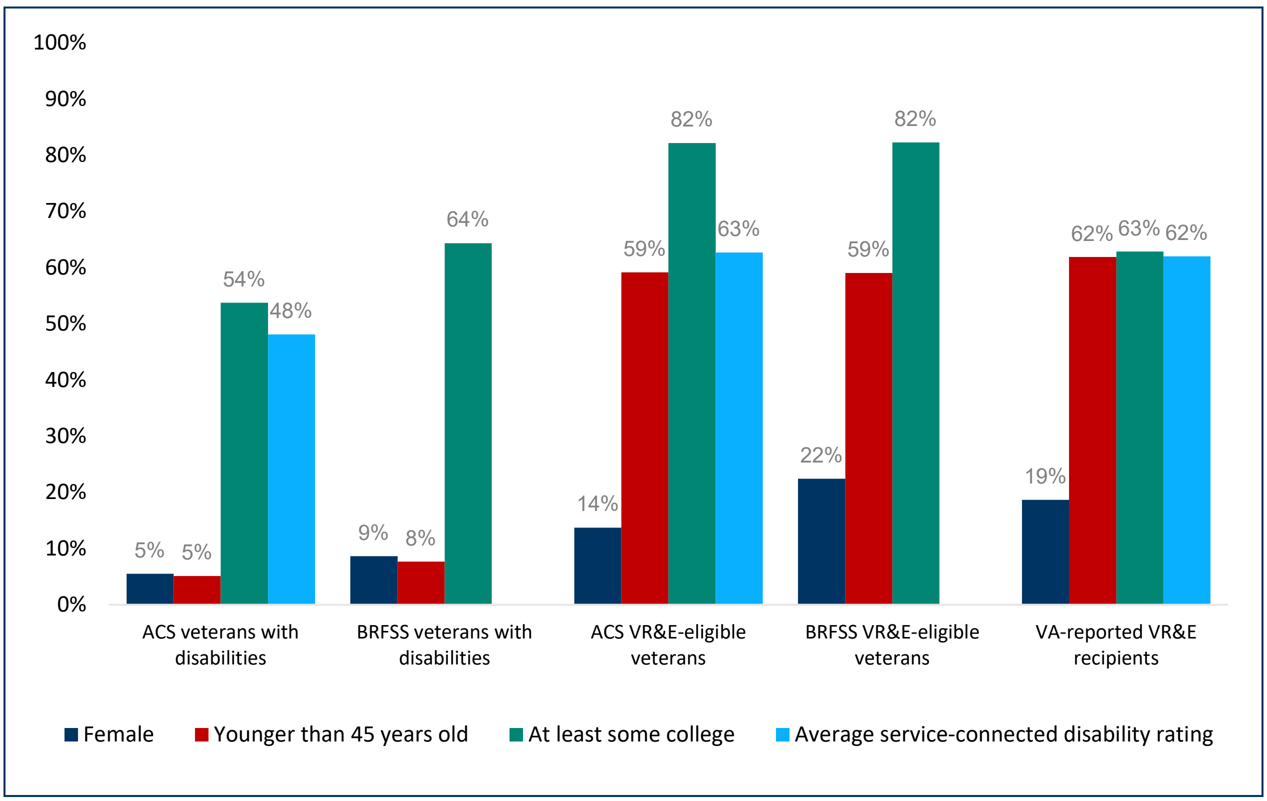
- The starting population on which the model is estimated should be similar across the two data sets.
- The ACS data should include a key set of explanatory variables with high predictive power in estimating VR&E eligibility.
- The BRFSS data should contain the same set of variables used as explanatory variables in estimating VR&E eligibility.

[8] Because we cannot observe the precise period of active duty, the criterion of post-2001 veterans also would include some veterans who have been discharged for more than 12 years (up to 16 years). Although we do not directly address this here, we believe that the differences in characteristics across those who were discharged within 12 years and within 16 years would not be significant.

The details of the modeling process are discussed in the appendix. Because the BRFSS data include veterans with disabilities, we used this as the starting population in both public data sets. We then examined the distribution of these predicted probabilities for VR&E-eligible and non-VR&E-eligible samples to identify the predicted probability that best differentiates between the two samples. To keep the rate of false positives at 5%, we selected a probability threshold of 0.11 to classify cases into a predicted VR&E-eligible sample. This resulted in an accuracy rate of 94% in the ACS data, where 78% of the VR&E-eligible sample was correctly classified as predicted VR&E eligible, and only 5% of the non-VR&E-eligible sample was incorrectly classified as predicted VR&E eligible.

Using the parameters and classification threshold from this prediction model, we predicted VR&E eligibility in the BRFSS 5-year national-level data on veterans with disabilities. This resulted in 9,711 veterans in the potentially VR&E-eligible sample. Figure 1 shows the demographic features for these samples against similar measures on actual VR&E service recipients reported by the VA in the years 2010, 2012, and 2014.

Figure 1. Sample Characteristics: All Veterans With a Disability Versus VR&E-Eligible Veterans



Note. ACS = American Community Survey; BRFSS = Behavioral Risk Factor Surveillance System; VA = Veterans Administration; VR&E = Vocational Rehabilitation and Education.

Comparing columns under "ACS VR&E-eligible veterans" and "BRFSS VR&E-eligible veterans" to "VA-reported VR&E recipients," we found that, in general, the likely VR&E-eligible samples from the ACS and the BRFSS are similar to the VA-reported VR&E recipients in terms of gender, age, and SCD rating. The largest difference was in terms of education: VR&E-eligible samples from both the ACS and the BRFSS are more likely to have at least some college education. However, it is important to note that this measure changes with time. In the VA report, for example, 63% of the veterans had at least some college education at the time of program entry. At the same time, the report said that most of the veterans participating in a plan of services were pursuing a college degree or a graduate degree. Therefore, the education statistics may not be directly comparable. All other measures on explanatory variables used in the model are shown in Table A2 in the appendix.

Baselines for Veterans With Service-Connected Disabilities

Using data for veterans who are likely to be VR&E eligible from the ACS and the BRFSS, we examined the baseline economic and health outcomes. (See Figures 2 and 3; detailed results are shown in the appendix.) It would not be accurate to compare the outcomes of the VR&E-eligible population against all non-VR&E-eligible veterans with any disability because the non-VR&E-eligible veterans are significantly older, out of the labor market, and may not always have an SCD. A better benchmark sample for these outcomes would be post–Gulf War II veterans with an SCD who were not VR&E-eligible. The BRFSS data do not include information on period of service or SCDs; therefore, we show health outcomes only for the first two samples.

Figure 2. Economic Outcomes

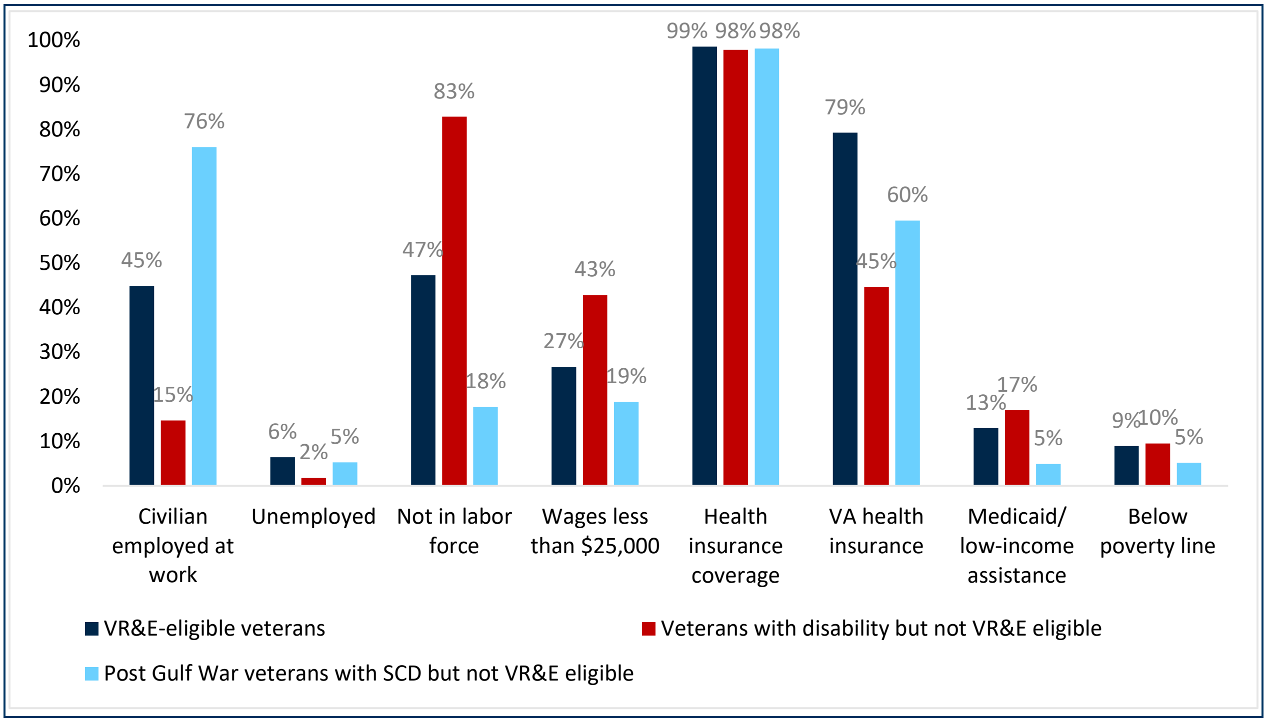
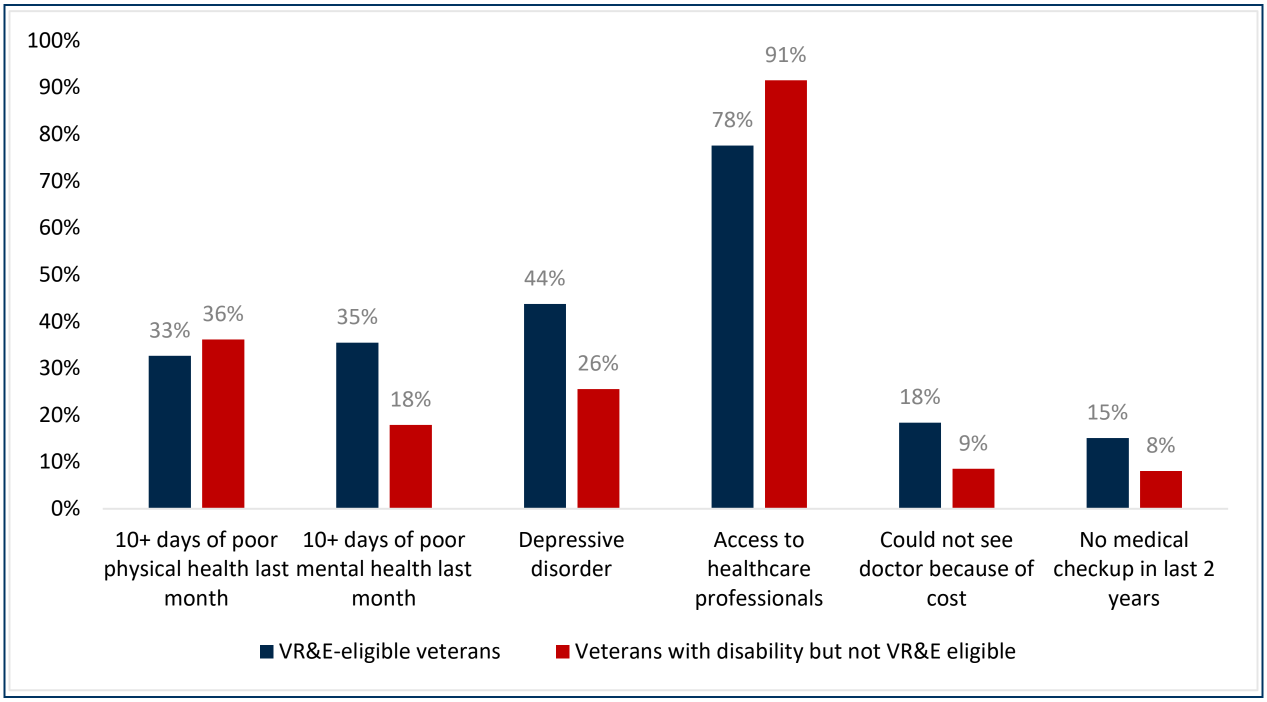


Figure 2 reveals differences among VR&E-eligible veterans and other veterans in terms of economic outcomes:

- The employment rate among VR&E-eligible veterans was 45% compared with 76% of the non-VR&E-eligible post–Gulf War veterans with SCDs.
- Six percent of the VR&E-eligible veterans were unemployed, which is slightly higher than 5% of the non-VR&E-eligible post–Gulf War veterans with SCDs.
- Forty-seven percent of the VR&E-eligible veterans were not in the labor force, whereas only 18% of the non-VR&E-eligible post–Gulf War veterans with SCDs were not in the labor force.
- Only 19% of the non-VR&E-eligible post–Gulf War veterans with SCDs earned less than $25,000 in wages, which is much lower than 27% of the VR&E-eligible veterans.
- Although almost all veterans reported having health insurance coverage, almost 80% of the VR&E-eligible veterans reported having health insurance through the VA, which is much higher than the other subsamples.
- Nine percent of the VR&E-eligible veterans were living below poverty line, which is much higher than other post–Gulf War veterans with SCDs (5%).

Figure 3. Health Outcomes



Looking at the health outcomes in Figure 3, we found the following:

- About 35% of the VR&E-eligible veterans reported having more than 10 days of poor mental health during the previous month compared with 18% of the non-VR&E-eligible veterans with disabilities.
- Forty-four percent of the VR&E-eligible veterans reported being diagnosed with a depressive disorder, whereas only 26% of the non-VR&E-eligible veterans with disabilities were diagnosed with depressive disorders.
- VR&E-eligible veterans had lower access to healthcare professionals (78% versus 91% of the non-VR&E-eligible veterans with disabilities).
- Eighteen percent of the VR&E-eligible veterans reported not being able to see a doctor because of costs, twice the corresponding percentage in non-VR&E-eligible sample.

Discussion

The purpose of this brief is to demonstrate how data obstacles, especially at the earlier stages of a PFS project, can be overcome by combining information from different sources to learn more about the population to be served. As we demonstrated, publicly available data can provide information on potential participants. In addition, we also used statistical prediction models to link two data sources (ACS and BRFSS) because one data set (BRFSS) did not have the required variables that aligned with the eligibility criteria for the target population.

The method we developed for the San Diego veterans can be adapted for other PFS projects, particularly when information on the target population is limited. There are several advantages of the proposed method. First, using publicly available data expedites the initial assess and build stages of a PFS project by eliminating reliance on new data collection. In addition, because access to publicly available data is relatively easy and does not require special use agreements or contracts, such data also reserve resources that would otherwise be spent on gaining access to data with personally identifiable information or a new data collection. Finally, because publicly available data are generally available both before and after an intervention, data for the comparison group selected from the publicly available data can be used for an impact analysis. This lessens reliance on an RCT in favor of a relatively robust quasi-experimental design to estimate the effectiveness of the intervention and also saves resources that would otherwise be used for conducting an RCT.

When demonstrating the use of publicly available data for use in a PFS project, we found some important information on veterans with SCDs. The baseline results from our analyses showed that veterans with SCDs, compared with veterans with disabilities but not eligible for VR&E, have lower job market outcomes (employment and wages) and were more likely to have mental health challenges.

Reference

Economic Systems Inc. (2018). *2016 Vocational Rehabilitation and Employment (VR&E) Longitudinal Study (Pub.L. 110-389 Sec. 334): Annual Report 2018 for FY 2017.* Retrieved from https://www.benefits.va.gov/VOCREHAB/docs/2017LongStdy.pdf

Appendix. Predicting VR&E Eligibility in Data Using Statistical Methods

In the absence of an RCT or a natural experiment, we needed to evaluate the outcomes of project participants against a similar population in the target area before the launch of Project (re)Launch. To measure these baseline outcomes, the most robust study design would involve access to administrative data from multiple sources, linked together to generate a comprehensive socioeconomic profile of a subsample of veterans in the target geography. Access to such administrative data is rare, and even more so in the feasibility phase of a PFS project. To address the challenge of data accessibility, we developed a method that used publicly available data, along with prepublished information on VR&E recipients, to help construct a pseudo-administrative data set to assess the baseline employment and health outcomes. To measure these outcomes, we used the ACS data on veterans with SCDs. Using information on VR&E eligibility criteria, we determined a potential intervention sample of veterans with SCDs who may have been VR&E eligible.

When working with the ACS to define the potential intervention sample, we first restricted the data to 133,985 veterans who were on active duty after 2001. This ensured that we analyzed only those veterans within 12 years of service discharge, which is the first criterion for VR&E eligibility. To construct a narrower sample of veterans who would be potentially representative of the target population, we classified veterans who met both of the following criteria into the potentially VR&E-eligible sample:

- An SCD rating of 10% or higher
- At least one disability (hearing, vision, self-care, independent living, ambulatory, or cognitive difficulty)

A second condition was added to substitute information on employment-related handicaps, assuming that difficulty in any one of these functions could lead to capability limitation among veterans who also had an SCD. This assumption was based on analyzing the correlation between these disabilities and the SCD rating. Specifically, we found that after applying the second condition of having a sensory disability, the average SCD of this potentially VR&E eligible subsample was 62%, which was similar to the rating of actual VR&E participants as reported by the VA. After applying these conditions, we had 11,087 veterans who were classified as potentially VR&E eligible from the ACS data.

Because the ACS data has only economic outcomes, we turned to the BRFSS data to analyze health outcomes. The BRFSS data, however, do not report any information on the SCD rating or period of service. Therefore, we employed a prediction model to create a pseudo-administrative data set using common information available in both the BRFSS and ACS data. Using the ACS data, we first predicted the likelihood of VR&E eligibility using variables common between the ACS and the BRFSS. We then applied the resulting parameters from this model to predict VR&E eligibility in the BRFSS data. This process is conceptually similar to a prediction exercise using cross-validation, where data from a test sample are used to generate a prediction model, and the model is then applied to a validation sample. Here, we used the sample from the ACS data to create a simple prediction model to determine VR&E eligibility and used that model to predict potential VR&E eligibility in the BRFSS sample. Three minimum conditions must be met for this prediction model to work:

- The starting population on which the model is estimated should be similar across the two data sets.
- The ACS data should include a key set of explanatory variables with high predictive power in estimating VR&E eligibility.
- The BRFSS data should contain a similar set of variables that contain similar information.

To predict the VR&E-eligibility status in the BRFSS data, we followed these steps:

1. We limited ACS 5-year national-level data to veterans with at least one ACS-defined disability. ACS-defined disabilities include difficulties with vision, hearing, cognition, ambulation, self-care, and independent living. Because the BRFSS data include information on veteran status and any disability, we used similar information in the ACS data for the starting sample. The starting sample in the ACS included 341,719 veterans with any disability.
2. Within this starting sample, we defined veterans who were VR&E eligible as those who also had an SCD rating of 10% or higher and who were in active duty since 2001. The remaining veterans with disabilities belong to the control group (the non-VR&E-eligible sample). Of the 341,719 veterans with disabilities, 11,087 were classified as VR&E eligible.
3. Using this VR&E eligibility outcome, we analyzed demographic, economic, and disability-related information that would predict VR&E eligibility and were present in the BRFSS sample. We selected the following variables in the prediction model: gender, race, age, income, education, and disability type (each of the disability-related variables listed earlier).

Then using these predictors, we ran a logistic regression model and computed predicted probabilities of VR&E eligibility. See Table A1 for odds ratios from the regression results.

Table A1. Regression Results for Predicting VR&E Eligibility

Variable	VR&E eligibility[a]
Gender (v/s female)	
Male	1.078*
	(0.038)
Race (v/s white alone)	
Black alone	1.181***
	(0.041)
American Indian or Alaskan Native alone	0.998
	(0.100)
Asian only	1.252*
	(0.124)
Native Hawaiian or Pacific Islander only	1.458
	(0.330)
Another race alone	1.167
	(0.104)
Multiracial	1.144*
	(0.067)
Age (in years) (v/s 18—24)	
25–34	1.258**
	(0.093)
35–44	0.312***
	(0.023)
45–54	0.121***
	(0.009)
55–64	0.021***
	(0.002)
65+	0.001***
	(0.000)
Annual Income (USD) (v/s less than 15,000)	
15,000–24,999	1.989***
	(0.085)
25,000–34,999	2.869***
	(0.128)

Variable	VR&E eligibility[a]
35,000–49,999	3.808***
	(0.152)
50,000 or higher	6.283***
	(0.225)
Disability	
Self-care difficulty	0.888**
	(0.035)
Vision difficulty	0.631***
	(0.025)
Independent living difficulty	1.053
	(0.032)
Ambulation difficulty	1.162***
	(0.031)
Cognitive difficulty	1.629***
	(0.042)
Education (v/s High school/GED or less)	
Some college or associate degree	2.169***
	(0.067)
College 4 years or more	3.292***
	(0.120)
Observations	**341,603**

[a]Exponentiated coefficients; standard errors in parentheses.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

After analyzing the distribution of the predicted probabilities (Table A2) for the actual VR&E treatment and control samples, we selected a probability threshold of 0.11. Table A2 indicates that 95% of the non-VR&E-eligible sample has a predicted VR&E eligibility less than 0.11, whereas more than 75% of the VR&E-eligible sample have predicted probability higher than 0.11. With this threshold, we can keep the false negative rate—falsely classifying a non-VR&E-eligible veteran as VR&E eligible—at 5% and compute a predicted VR&E-eligible sample. We achieved an accuracy rate of 94%, where 78% of the actual VR&E-eligible sample was classified as predicted VR&E eligible, and only 5% of actual control sample was classified as predicted VR&E eligible. This test verified the overall performance of the prediction model.

Table A2. Distribution of Predicted Probabilities by VR&E Eligibility

Percentile	VR&E eligible	Non-VR&E eligible
1%	0.0017	0.0001
5%	0.0146	0.0002
10%	0.0425	0.0003
25%	0.1380	0.0006
50%	0.3098	0.0016
75%	0.5364	0.0059
90%	0.7225	0.0443
95%	0.7922	0.1116
99%	0.8549	0.4061

Using the parameters of this prediction model, we estimated the probability of VR&E eligibility in the BRFSS 5-year data on California veterans with disabilities. The BRFSS data compiled from 2012 to 2015 had 305,256 veterans; this sample was further cleaned for disability information and restricted to 83,900 veterans with disabilities. We used the probability threshold of 0.11 to classify those falling over this threshold as VR&E eligible and vice versa. This led to a sample of 9,711 veterans in the potentially VR&E-eligible sample in the BRFSS data.

Table A3 displays the explanatory variables in the treatment and comparison groups, in both the ACS and BRFSS data.

Table A3. VR&E-Eligible Versus Non-VR&E-Eligible Sample Characteristics

	BRFSS		ACS	
Category	**Veterans predicted as VR&E eligible**	**Veterans with a disability but not VR&E eligible**	**Veterans who are VR&E eligible**	**Veterans with a disability but not VR&E eligible**
Number of cases	9,711	72,001	11,087	330,632
Gender				
Female	22.41%	6.77%	13.74%	5.21%
Male	77.59%	93.23%	86.26%	94.79%
Race				
White alone	71.31%	84.65%	75.76%	86.49%
Black or African American alone	10.07%	6.52%	14.21%	8.89%
American Indian or Alaskan Native alone	2.67%	1.90%	1.36%	0.95%
Asian alone	1.32%	0.61%	1.72%	0.91%
Native Hawaiian or Pacific Islander alone	0.70%	0.22%	0.31%	0.12%
Some other race alone	8.83%	3.71%	1.99%	0.77%
Two or more races	5.11%	2.39%	4.65%	1.87%
Age (in years)				
18–24	4.86%	0.01%	3.49%	0.20%
25–34	24.03%	0.13%	32.32%	0.99%
35–44	30.15%	0.63%	23.33%	2.11%
45–54	40.57%	5.27%	26.07%	6.59%
55–64	0.38%	21.94%	11.52%	17.18%
65+	0.00%	72.01%	3.27%	72.93%
Highest educational qualification				
High school/GED or less	17.69%	37.94%	17.81%	47.22%
Some college or associate degree	46.04%	30.54%	55.18%	32.44%

	BRFSS		ACS	
College 4 years or more	36.27%	31.52%	27.00%	20.34%
Income categories				
Less than $25,000	18.82%	33.18%	27.84%	47.11%
$25,000 to $34,999	10.75%	15.10%	11.45%	15.11%
$35,000 to $49,999	17.20%	17.91%	18.27%	15.15%
$50,000 or higher	53.23%	33.80%	42.44%	22.64%
Disability type				
Difficulty in vision	6.66%	13.22%	9.16%	15.71%
Cognitive difficulty	42.09%	22.80%	51.01%	26.89%
Ambulation difficulty	32.29%	51.45%	40.33%	55.04%
Self-care difficulty	9.97%	12.36%	12.69%	20.93%
Independent living difficulty	18.82%	16.37%	27.07%	32.78%

Table A4 shows the baseline outcomes for a target population of veterans with SCDs compared with other veteran groups.

Table A4. Baseline Characteristics for Veterans

	VR&E-eligible veterans	Veterans with disability but not VR&E eligible	Post–Gulf War veterans with SCD but not VR&E eligible
Employment status			
Civilian employed at work	44.88%	14.71%	76.08%
Civilian employed with job but not at work	1.43%	0.66%	0.88%
Unemployed	6.39%	1.78%	5.30%
Not in labor force	47.30%	82.86%	17.73%
N	**11,087**	**330,632**	**35,057**
Income categories (inflation adjusted and conditional on being employed at work)			
Less than or equal to $10,000	11.41%	23.92%	7.74%
$10,001 to $25,000	15.27%	18.92%	11.08%
$25,001 to $50,000	31.17%	25.13%	28.14%
$50,001 to $75,000	21.97%	16.13%	23.85%
Greater than $75,000	20.18%	15.90%	29.19%
N	**11,087**	**330,632**	**35,057**
Health insurance coverage			
Yes	98.62%	97.91%	98.15%
No	1.38%	2.09%	1.85%
N	**11,087**	**330,632**	**35,057**
VA health insurance			
Yes	79.27%	44.69%	59.54%
No	20.73%	55.31%	40.46%
N	**11,087**	**330,632**	**35,057**
Medicaid or other low-income assistance			
Yes	12.97%	16.96%	4.93%
No	87.03%	83.04%	95.07%
N	**11,087**	**330,632**	**35,057**
Below poverty line			
Yes	8.92%	9.50%	5.19%
No	91.08%	90.50%	94.81%
N	**11,087**	**330,632**	**35,057**

	VR&E-eligible veterans	Veterans with disability but not VR&E eligible	Post–Gulf War veterans with SCD but not VR&E eligible
Days with poor physical health last month			
None	40.10%	43.47%	
1–9 days	27.26%	20.41%	
10–19 days	10.82%	9.67%	
20 days or more	21.83%	26.46%	
N	**9,594**	**69,754**	
Days with poor mental health last month			
None	48.11%	73.16%	
1–9 days	16.43%	8.92%	
10–19 days	10.29%	4.09%	
20 days or more	25.17%	13.84%	
N	**9,594**	**69,754**	
Diagnosed with depressive disorder			
Yes	43.76%	25.53%	
No	55.43%	73.93%	
Not sure	0.63%	0.47%	
Missing	0.18%	0.08%	
N	**9,594**	**69,754**	
Heavy alcohol consumption			
No	90.17%	92.73%	
Yes	7.85%	5.06%	
Don't know or refused	1.99%	2.21%	
N	**9,594**	**69,754**	
Access to multiple healthcare professionals			
Yes, only one	65.97%	77.93%	
Yes, more than one	11.58%	13.56%	
No	22.10%	8.16%	
Don't know	0.21%	0.22%	
Refused	0.14%	0.13%	
N	**9,594**	**69,754**	

	VR&E-eligible veterans	Veterans with disability but not VR&E eligible	Post–Gulf War veterans with SCD but not VR&E eligible
Could not see doctor because of cost			
Yes	18.37%	8.52%	
No	81.47%	91.23%	
Missing or not applicable	0.16%	0.25%	
N	**9,594**	**69,754**	
Length of time since last checkup			
Within past year	73.03%	85.26%	
Within past 2 years	11.85%	6.70%	
Within past 5 years	7.03%	3.18%	
5 or more years ago	6.64%	3.43%	
Don’t know/Not sure	0.85%	0.94%	
Never	0.55%	0.39%	
Refused	0.04%	0.10%	
N	**9,594**	**69,754**	

AMERICAN INSTITUTES FOR RESEARCH®

Established in 1946, the American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research on important social issues and delivers technical assistance, both domestically and internationally, in the areas of education, health, and workforce productivity.

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH
1000 Thomas Jefferson Street NW
Washington, DC 20007-3835 | 202.403.5000
www.air.org

Third Sector is transforming the way communities connect people with vital services. We use funding, data, and incentives as levers to impact how governments, service providers, and community stakeholders work together. Founded in 2011, Third Sector is a 501(c)3 nonprofit organization and is supported by its work for government and service providers as well as philanthropic and government grants. To learn more, visit www.thirdsectorcap.org.

LOCATIONS

Domestic: Washington, DC (HQ) | Monterey, Sacramento, and San Mateo, CA | Atlanta, GA | Honolulu, HI | Chicago and Naperville, IL Indianapolis, IN | Metairie, LA | Waltham, MA | Frederick and Rockville, MD | Chapel Hill, NC | New York, NY | Columbus, OH | Cayce, SC Austin, TX | Reston, VA

International: El Salvador | Ethiopia | Haiti | Honduras | Zambia

7116_01/19