PrecisionFDA Crowdsourcing Addresses Veteran Health Needs While Advancing the Science of Real-World Data (RWD)

Anish Prasanna[1], Sarah Prezek[1], Vishal Thovarai[1], Ezekiel J. Maier[1], Elaine Johanson[2], Bocheng Jing[3, 4], Alice Feng [5], Jaleal Sanjak[1], Holly Stephens[6], Amanda Lienau[7] - [1] Booz Allen Hamilton, [2] U.S. Food and Drug Administration, [3]Northern California Institute for Research and Education, [4] San Francisco VA Medical Center, [5] Harvard University, [6] Precigen, [7] VHA Innovation Ecosystem

Abstract

Many of the 18+ million U.S. veterans, experience chronic health conditions including obesity and diabetes. **PrecisionFDA, in collaboration with the Veterans Health Administration (VHA) Innovation Ecosystem (IE), is running public challenges focused on developing AI/ML models of veteran health outcomes**. The VHA IE and FDA COVID-19 Risk Factor Modeling Challenge focused on developing and evaluating models to predict COVID-19-related health outcomes and assessing the value of synthetic RWD. Additionally, the Veterans Cardiac Health and AI Model Predictions (V-CHAMPS) Challenge, cosponsored by VHA IE, DHCoE, precisionFDA, and UK Medicines and Healthcare products Regulatory Agency (MHRA), focuses on developing and evaluating models to predict cardiovascular-related health outcomes.

Twenty-one teams submitted 34 entries to the COVID-19 challenge. Models trained on synthetic data showed similar but inflated model performance metrics and largely overlapping risk factors, to those trained on real data. Over 300 individuals pre-registered for the V-CHAMPS Challenge. These collaborations/participant engagement demonstrate the significant interest in solving veteran health needs/analyzing RWD. **These challenges advance the understanding of how synthetic RWD can be used for AI/ML modeling.** Continued improvements in synthetic data generation will shrink the gap between real and synthetically trained ML models.

Introduction

PrecisionFDA is:

- **a secure, high-performance computing platform for collaborative data analytics** between FDA FTEs and external partners
- a community of experts engaged in the analysis of diverse omics datasets with a focus on supporting FDA's mission and advancing regulatory science
- a framework, methodology, and platform for conducting crowdsourcing challenges

PrecisionFDA Crowdsourcing Activities:

- **Hosted 37 challenges with 816 submissions from participants in 54 countries**
- Cited in 90 Scientific Publications related to these challenges
- Has solved real world problems and results include the development of machine learning tools, expanded sequencing standard adoption, pipeline benchmarks, identification of novel technology, and validation of applied techniques

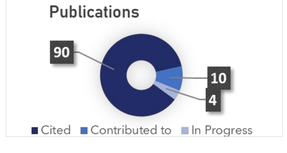

Figure 1. Number of publications in which precisionFDA is cited, has contributed to, or are in progress.

Materials and Methods

COVID-19 Risk Factor Modeling Challenge

- **Problem to be Solved:** Veterans have a higher prevalence of many risk factors for severe COVID-19 illness. Identification of at-risk Veterans can enable preventive measures.
- **Challenge Workflow:**
  - In phase 1, **participants were provided with synthetic health records and were asked to predict target COVID-19 outcomes.**
  - In phase 2, the top three **models were retrained and evaluated on two additional synthetic and real de-identified veteran health record datasets.**

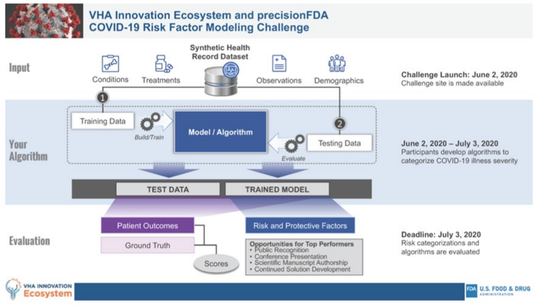

Figure 2. VHA IE and precisionFDA COVID-19 Risk Factor Modeling Challenge workflow graphic.

The V-CHAMPS Challenge

- **Problem to be Solved:** Cardiovascular disease is the leading cause of hospitalization in the VA and a major cause of disability. Identification of risk factors can enable preventive patient care.
- **Challenge Workflow:**
  - In phase 1, **participants are provided with synthetic health records and asked to predict outcome variables for Veterans who have been diagnosed with chronic heart failure.**
  - In phase 2, the best performing **models are validated on the real-world Veterans health data within the VHA.**

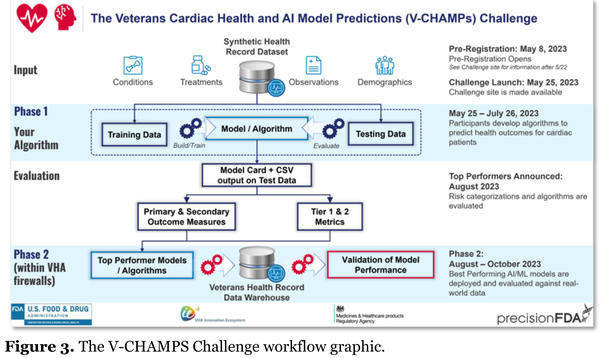

Figure 3. The V-CHAMPS Challenge workflow graphic.

Results and Discussion

COVID-19 Risk Factor Modeling Challenge

- **Within Phase 1, twenty-one teams submitted 34 models to predict COVID-19-related health outcomes including COVID-19 status, length of hospitalization, ICU stay, ventilation status, and mortality that were trained using synthetic Veteran health record data** created by the Synthea synthetic patient generator.
- Phase 1 submissions were evaluated using various metrics such as the AUROC metric, Root-Mean-Square Error (RMSE), and concordance index on a 20% holdout test set.
- The top performers all specifically used gradient-boosted machine learning algorithms

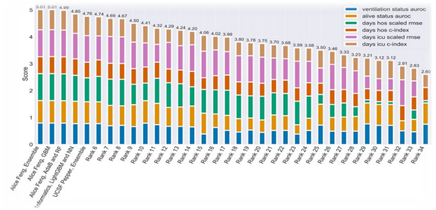

Figure 4. Phase 1 Top Performers: Ranked from left to right. All metrics except for RMSE are between 0 and 1 where 1 is the best possible score. RMSE was scaled to fit this range using the following equation making 6 as the maximum possible score

- Across all COVID-19 health outcome model metric distributions, **severe COVID-19 health outcomes (such as deceased status and days in ICU) were easier to predict than less severe outcomes (e.g., COVID-19 status, ventilation status, days in hospital)**.

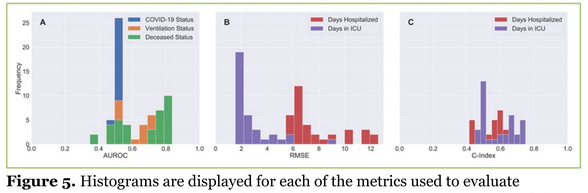

Figure 5. Histograms are displayed for each of the metrics used to evaluate submissions. The average AUROC for COVID-19 status (Figure A) is 0.516 and thus was not considered for identification of top performing models.

Key Takeaways

The COVID-19 Risk Factor Modeling Challenge results provide a public resource for comparing synthetic data generation methodologies and the efficacy of ML algorithms against those respective techniques

The V-CHAMPS Challenge is helping advance the understanding of how synthetic RWD can be used for AI/ML modeling and demonstrate the significant interest in solving veteran health needs and analyzing RWD

- Within Phase 2, top-performing models were adapted/validated on additional CorGAN synthetic and real VA datasets to validate the effectiveness of novel synthetic data generation algorithms.
- **Phase 2 top-performer models significantly outperformed random chance predictions on both the synthetic and real Veteran health datasets and demonstrated minimal variability across outcomes and datasets.**
- All Phase 2 individual models produced significantly superior results on severe COVID-19 health outcomes.
- The models trained on synthetic data showed similar but inflated model performance metrics and largely overlapping risk factors, to those trained on real data.
- The underlying similarity between real and synthetic data generated by packages such as ABSEHRD and Synthea have real practical value for research on Veterans' health.

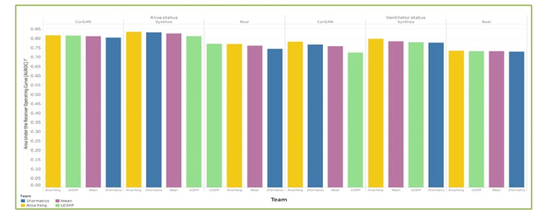

Figure 6. Bar plots show top performer models' AUROC scores on binary health outcomes (e.g., alive and ventilator status) across datasets.

The V-CHAMPS Challenge

- Phase 1 of the Challenge closes on July 26th
- Over 300 individuals have pre-registered for the V-CHAMPS Challenge
- **Participants will submit benchmarked Artificial Intelligence and Machine Learning (AI/ML) models for predicting cardiovascular health related outcomes in Veterans that in Phase 2 will be evaluated against RWD**

Figure 7. The V-CHAMPS Challenge social media communications graphic.

Conclusion

- High participation in the COVID-19 Risk Factor Modeling Challenge and V-CHAMPs Challenge demonstrate significant interest in solving veteran health needs and analyzing RWD
- **The results of these challenges improve our understanding of the strengths and limitations of using synthetic RWD for AI/ML modeling**
- Continued improvements in synthetic data generation will shrink the performance gap between real and synthetically trained ML models
- PrecisionFDA's public community challenges help drive methods development and inform precision medicine