

Save Our Sisyphus Challenge: Lessons learned from Strategus execution on the OHDSI Evidence Network

Anthony G. Sena^{1,2,3}, Jenna Reps^{1,2}, Chungsoo Kim^{2,4}, Jack Brewster^{2,5}, Adam Black², Linying Zhang², Thomas Falconer^{2,6}, George Hripsak^{2,6}, Cindy X. Cai^{2,7}, Michael Cook^{2,7}, Phan Thanh Phuc^{2,8}, Jason C. Hsu^{2,8}, Phung-Anh Nguyen^{2,8}, Muhammad Solihuddin Muhtar^{2,8}, Brian Toy^{2,9}, Zachary Gilbert^{2,9}, Xiaoyu Lin^{2,10}, Jing Li^{2,10}, Sarah Seager^{2,11}, Yeonjae Han^{2,12}, Seng Chan You^{2,12}, Scott L. DuVall^{2,13,14}, Marc A. Suchard^{2,13,15}

¹Janssen Research & Development, Titusville, NJ, USA, ²Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA, ³ Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands, ⁴Department of Biomedical Sciences, Ajou University Graduate School of Medicine, ⁵Real World Solutions, IQVIA, Brighton, UK, ⁶Department of Biomedical Informatics, Columbia University, New York, USA, ⁷Johns Hopkins University School of Medicine, Baltimore, MD, USA, ⁸Taipei Medical University, Taipei, Taiwan, ⁹Roski Eye Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States, ¹⁰IQVIA Solutions Asia Pte Ltd, ¹¹IQVIA Ltd, UK, ¹²Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Republic of Korea, ¹³VA Informatics and Computing Infrastructure, US Department of Veterans Affairs, Salt Lake City, UT, USA, ¹⁴Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA, ¹⁵Department of Biostatistics, University of California Los Angeles, Los Angeles, CA, USA

Background

The Save Our Sisyphus (SOS) Challenge educated the OHDSI community on the process of leading or participating in an OHDSI network study.¹ A motivation for the SOS challenge was to test the hypothesis that with a well-defined tutorial and the support of OHDSI tools running on a standardized R environment, OHDSI members can execute network studies effectively and efficiently, and present valid results to the study coordinators. A critical step in the OHDSI network study process is the execution of analytical code against patient level data transformed to the OMOP Common Data Model. The SOS challenge made use of the Strategus R package which aims to provide the software infrastructure for running the OHDSI HADES analytical packages.² Here we share lessons learned from the network execution of the SOS Challenge across 11 OHDSI network sites and ideas for improvement of Strategus and study execution moving forward.

Methods

During the SOS Challenge, the OHDSI community learned about the process of running an OHDSI network study through weekly tutorials. These tutorials used two network studies to illustrate the network study process: 1) Intravitreal anti-VEGF and kidney failure risk (Anti-VEGF) and 2) fluoroquinolone and aortic aneurysm risk (FQ). Both studies ran the same sets of analytics for their study population and outcome of interest. Specifically, the studies included characterization analytics for performing cohort diagnostics, computing baseline characteristics, time-to-event analysis, dechallenge/rechallenge and incidence rates. Population-level estimation was performed using two different designs: self-controlled case series and the cohort method. Patient-level prediction was performed to fit a set of models to predict the respective

outcome for several time-at-risk windows.

Two tutorials on network execution of the SOS studies were presented to the OHDSI Community on May 2nd, 2023. These tutorials guided OHDSI community members through the steps of locating the study analysis specification and how to run the code to execute the study, including the Strategus package. OHDSI data partners were then asked to follow the tutorial steps to execute each network study on their data and contribute results back to the OHDSI study coordinators. Office hours were held the following week to support data partners with questions or challenges they faced when attempting to run the network studies using Strategus.

Results

Table 1 summarizes the SOS studies by OHDSI data partner and their contributions to the execution and study results.

OHDSI Data Partner	Study Status (Number of Databases)	
	Anti-VEGF (12)	FQ (17)
Ajou University Medical Center	-	Completed (2)
Columbia University Medical Center	Completed (1)	Completed (1)
IQVIA	-	Completed (5)
Janssen R&D	Completed (6)	Completed (6)
Johns Hopkins University	Completed (1)	-
Northeastern University	Completed (1)	-
Stanford University	Completed (1)	-
Taipei Medical University	-	Completed (1)
University of Southern California	Completed (1)	-
Department of Veterans Affairs	Completed (1)	Completed (1)
Yonsei University College of Medicine	-	Completed (1)

Table 1. Participating OHDSI Network Sites in each SOS Challenge study with number of data sources contributed as of October 1, 2023

The journey from the network execution study tutorial to providing results back to the OHDSI study coordinator had various technical challenges and lessons learned. Strategus makes use of an R package called keyring³ to securely store sensitive information, including database connection information for the data holder's CDM. During the SOS challenge, several users reported challenges in configuring this package on non-Windows operating systems. Solutions to these problems were developed and shared

amongst those data partners that were running on non-Windows operating systems.

Strategus utilizes `renv`⁴ to capture and restore the R packages required to run a HADES analytical package. Strategus automatically invokes the `renv restore` process to set up the execution environment. There were several challenges encountered while attempting to run Strategus related to `renv`. First, when the `renv restore` process encounters any failures in restoring a specific R package it does not automatically raise an error. This then causes the Strategus execution to fail, often with little information on the cause of the failure. This required data partners to dig into the internals of Strategus to properly restore their environment to complete the execution of the study. Additionally, during the `renv restore` process, there were conflicts between the base R version installed on the machine running Strategus and the dependencies required by the HADES analytical package. As a result, the OHDSI HADES working group discussed and adopted a single version of R for use by all HADES analytical packages to mitigate this problem moving forward.

Additional issues and ideas for Strategus were captured via GitHub issues using the “`sos-challenge`” tag⁵.

Conclusion

The SOS Challenge provided an opportunity for the OHDSI community to execute a network study using the Strategus package. The community was able to execute two different studies using the same R infrastructure to perform characterization, estimation and prediction analytical tasks using HADES. This was the first time the community used Strategus for network execution and worked collaboratively to overcome several technical hurdles. We learned that standardization of the R execution environment is a critical task and that there is more work to be done both on the Strategus package and in standardizing the network execution environment.

References

1. SOS Challenge – OHDSI [Internet]. OHDSI (US) 2023 March 28. Available from: <https://ohdsi.org/sos-challenge/>
2. Strategus. GitHub [Internet]. 2023. Available from: <https://github.com/OHDSI/Strategus>
3. keyring. GitHub [Internet]. 2023. Available from: <https://github.com/r-lib/keyring>
4. renv. GitHub [Internet]. 2023. Available from: <https://github.com/rstudio/renv/>
5. Strategus SOS Challenge Issues. GitHub [Internet]. 2023. Available from: <https://github.com/OHDSI/Strategus/labels/sos-challenge>