

Developing Efficient, Reusable SAS Programs in PCORnet: PCORnet Modular Program 1 (PMP1)

Jensen N. Hu, MPH¹, Julia A. Fearington, MPH¹, Jessica L. Sturtevant, MS¹, Casie E. Horgan, MPH¹, Érick Moyneur, MA², Dave Martin, MA², Darcy M. Louzao, PhD³, Jeffrey S. Brown, PhD¹

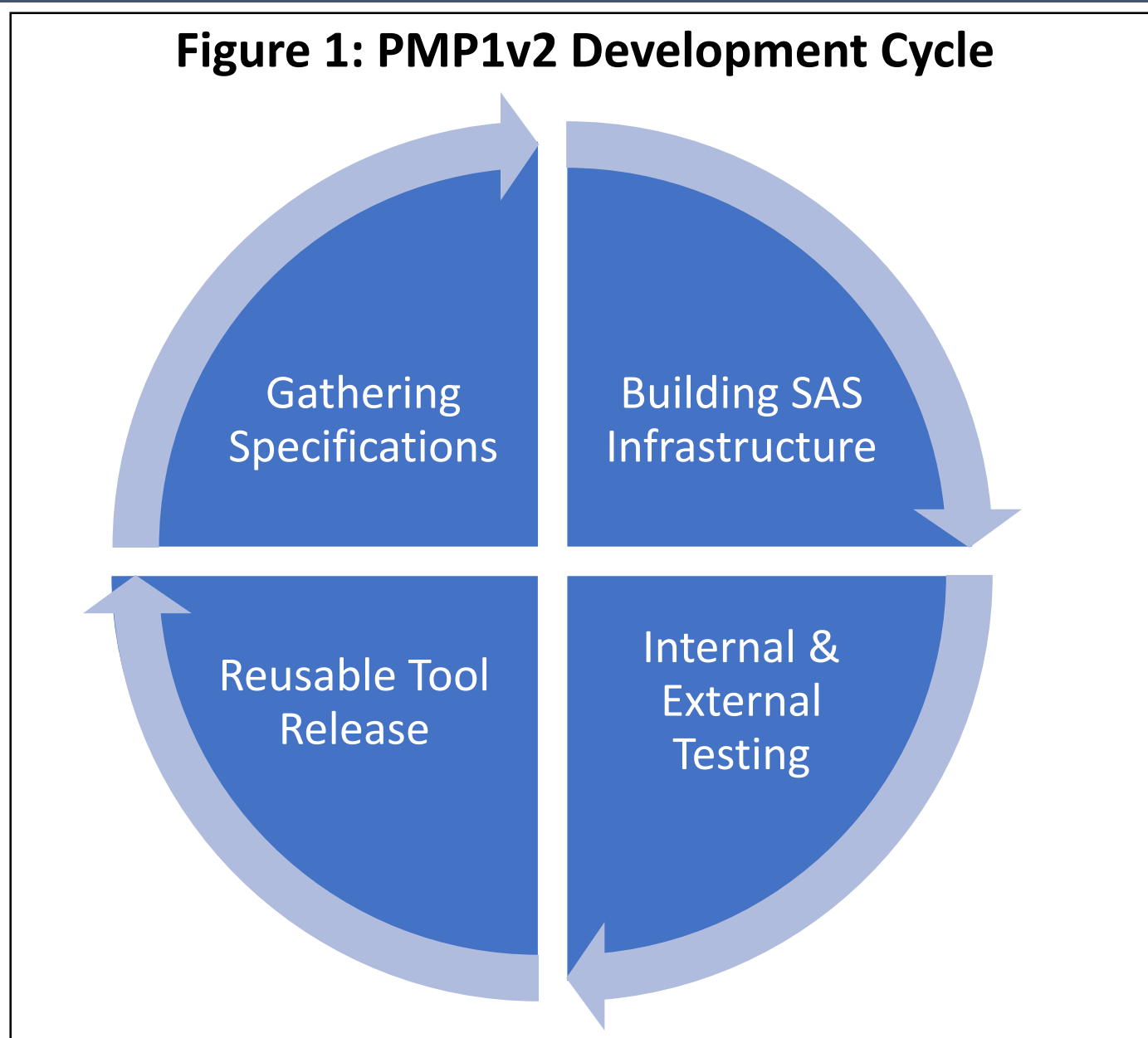
¹Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA; ²Statlog Consulting Inc., Ottawa, ON, Canada; ³Duke Clinical Research Institute, Durham, NC;

Background

- The National Patient-Centered Clinical Research Network (PCORnet) developed a Common Data Model (CDM) to enable distributed querying to 80+ data contributing Network Partners (NPs).
- Distributed querying can be accomplished using ad hoc code, however this process is time intensive and costly.
- Efficient, reusable SAS-based tools can generate answers in weeks instead of months.
- Reusable tool development incorporates key stakeholder input and thorough testing to enable execution across 80+ diverse technical computing ecosystems.
- Development and enhancement of the reusable SAS program, PCORnet Modular Program 1 (PMP1), served to improve efficiency, usability and facilitate easier output review and subsequent partner governance and approval process.

Methods

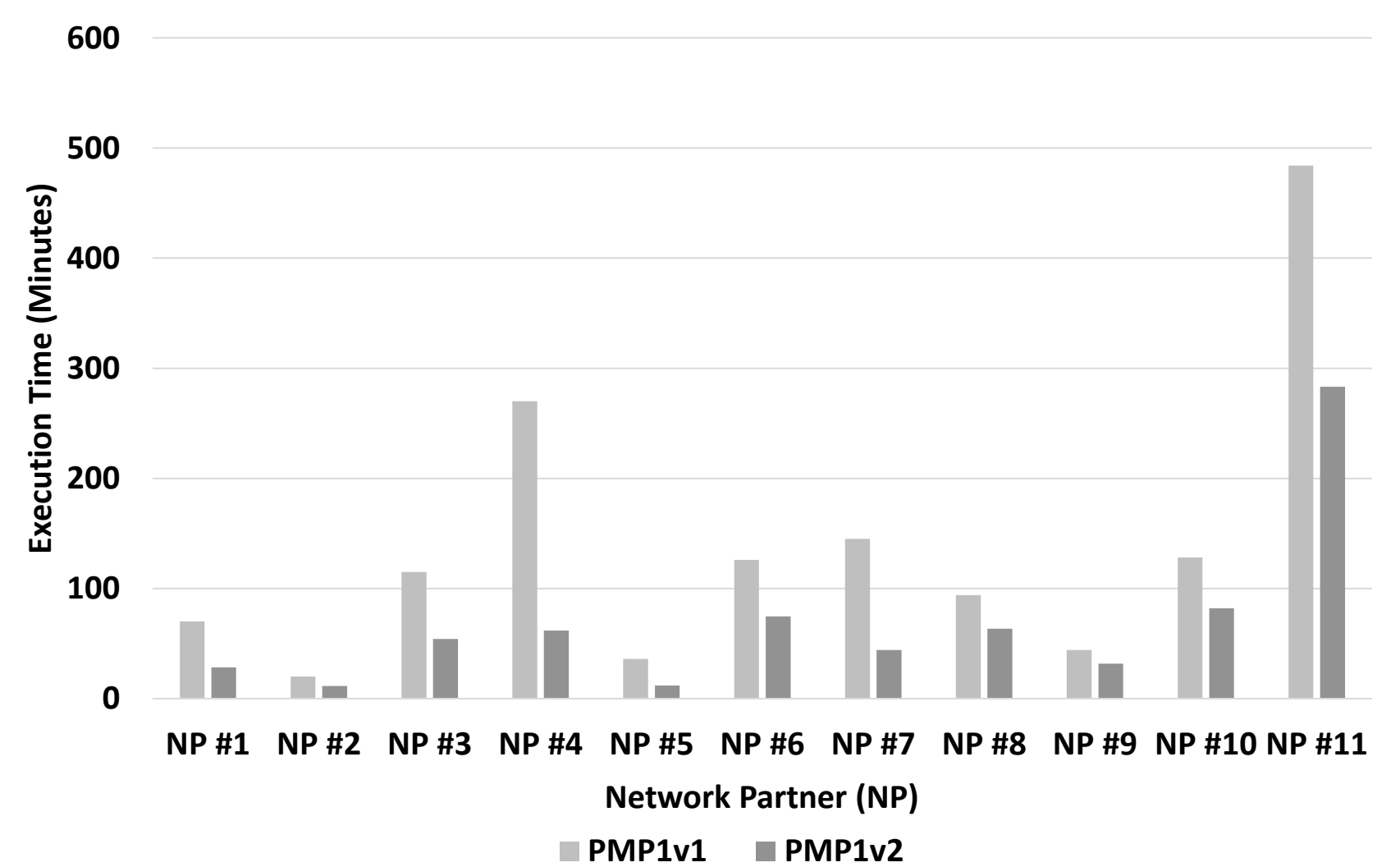
Figure 1: PMP1v2 Development Cycle



- **Gathering Specifications** involved a working group of 3 Clinical Data Research Networks (CDRNs)—PaTH, Mid-South, and REACHnet—that provided feedback on content and format of a summary report to improve output review.
- **Building SAS Infrastructure** involved adding and modifying SAS code to create the report and to improve efficiency and usability.
- **Internal & External Testing** involved execution of a PMP1v1-documented and tested diagnosis-based phenotype query on 13 NPs selected to represent PCORnet computing and infrastructure heterogeneity.
- **Reusable Tool Release** involved a release communication plan and a distribution of a survey to gauge network agreement on interpretability of output, facilitation of internal review and governance process, and usability of PMP1v2.

Results

Figure 2: PMP1v2 and PMP1v1 Execution Time by NP



Network Partner(s)	Population Size (Number of Patients in Diagnosis Table)	Operating System
NP #1, NP #2	<1 million	WIN
NP #3	<1 million	LIN X64
NP #4	1-5 million	AIX 64
NP #5, NP #6	1-5 million	WIN
NP #7	1-5 million	LIN X64
NP #8, NP #9	1-5 million	WIN
NP #10	1-5 million	LIN X64
NP #11	>5 million	WIN

Table 1: Survey of Network Partner Sentiment toward PMP1v2 Enhancements (Based on response from 8 CDRNs)

PMP1v2 Improves:	Agreement	Neutral	Disagreement
Usability & Execution	59%	41%	0%
Site Interpretation & Review of Output	86%	14%	0%
Site Governance Process & Expedited Return of Results	60%	30%	10%

Conclusions

- Collaboration with NPs is beneficial in architecting and implementing enhancements in distributed, reusable SAS programs.
- PMP1v2 improved interpretability of results by creating a summary report and program execution efficiency by dramatically decreasing program runtime.
- Utilizing reusable tools in the PCORnet DRN enable questions to be asked of 80+ NPs.
- The PCORnet DRN benefits from NP, stakeholder and development team input in reusable tool architecture and development.

Contact Information

Jensen Hu, MPH
Harvard Medical School, Harvard Pilgrim Health Care Institute;
Jensen_Nicolas_Hu@harvardpilgrim.org

Acknowledgements

This work was supported through a Patient-Centered Outcomes Research Institute (PCORI) Program Award (CC2-Duke-2016). All statements in this poster, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of PCORI, its Board of Governors or Methodology Committee.

AMIA 2018
Informatics Summit
March 12th-15th

