

OHDSI Japanイブニングカンファレンス

ウェビナー聴講報告

2024年6月28日 浅尾 啓子



**OHDSI**

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

*Applying Machine Learning in Distributed Networks to Support  
Activities for Post-Market Surveillance of Medical Products:  
Opportunities, Challenges, and Considerations*

*CBER BEST Initiative Seminar Series*

June 26, 2024

Presenter: Dr. Jenna Wong, PhD, Assistant Professor, Department of Population  
Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute

# The Biologics Effectiveness and Safety (BEST) Initiative



Search ... 🔍

- BEST VACCINES & ALLERGENICS BLOOD & BLOOD PRODUCTS TISSUES & ADVANCED THERAPEUTICS OUTCOMES
- OTHER PROGRAMS

## Center for Biologics Evaluation & Research (CBER), FDA

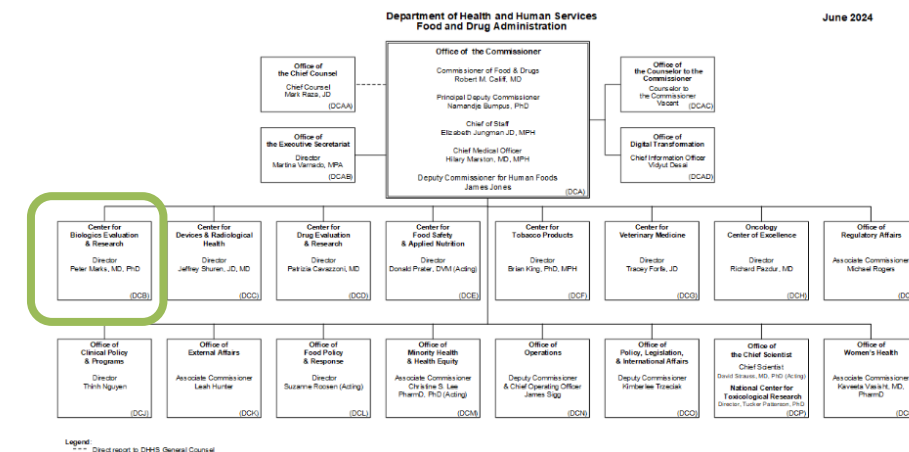
- Vaccines
- Allergenic
- Blood and blood products
- Cells
- Tissues
- Gene therapies

### The Biologics Effectiveness and Safety (BEST) Initiative

- A part of the Center for Biologics Evaluation and Research (CBER) surveillance program
- A part of the Sentinel Initiative housed within the Office of Biostatistics and Pharmacovigilance (OBPV)

### Data & Surveillance Activities

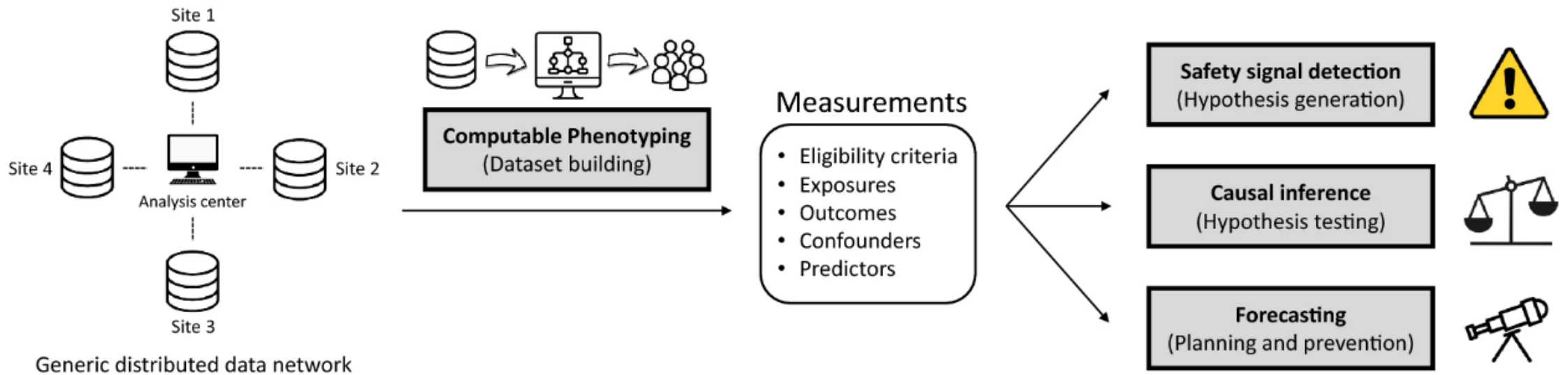
- Distributed Network and Common Data Model
- Artificial Intelligence and Natural Language Processing
- BEST Innovative Methods Exchange Platform
- FDA's Survey of Health and Patient Experience SHAPE Platform





# 分散データネットワーク(DDN)の中で 機械学習(ML)が応用される場面

## From: Applying Machine Learning in Distributed Data Networks for Pharmacoepidemiologic and Pharmacovigilance Studies: Opportunities, Challenges, and Considerations



# 分散データネットワーク(DDN)での4つのシナリオ

シナリオ	1. 基本	2. データの標準化↓	3. データ構造の複雑さ↑	4. データ粒度↓
データ種別	構造化データのみ	構造化データのみ	構造化・非構造化データ	構造化データのみ
標準化	全項目がCDM	一部CDMではない入力	全項目がCDM	全項目がCDM
粒度	全サイトで個人レベル	全サイトで個人レベル	全サイトで個人レベル	全サイトで集団レベル
MLを適用する際の考慮	<ul style="list-style-type: none"> <li>科学的な観点から、全サイトのデータをプールしてMLを適用してよいのか？(DB間の不均一性を考慮しないでよいのか)</li> </ul>	<ul style="list-style-type: none"> <li>アプローチ1. 標準化されていないデータを標準化してからMLを適用</li> <li>アプローチ2. 各サイトでサイト別に標準化されていないデータのモデリングを実施し、これらの追加変数で結果が変わるかどうかを確認</li> </ul>	<ul style="list-style-type: none"> <li>アプローチ1. 各サイトでサイト別に標準化されていないデータのモデリングを実施し、これらの追加変数で結果が変わるかどうかを確認</li> <li>アプローチ2. 全現データテキストを直接CDMの1フィールドとして投入(ストレージ容量とプライバシーへの懸念)</li> </ul>	<ul style="list-style-type: none"> <li>アプローチ1. 各サイトでモデルフィッティングし、各サイトのモデルを比較、あるいは、メタ解析として各モデルを統合</li> <li>アプローチ2. 1つのサイトでモデルを構築し、他のサイトでモデルをバリデーション(1つのサイトでモデルを構築するためDDNの強みを生かし切れない)</li> </ul>

連合学習(federated learning)により、複数サイトが協力してMLモデルをプライバシーの懸念なしにトレーニングできる(か)



# References

- “Learning Patient-Level Prediction Models across Multiple Healthcare Databases: Evaluation of Ensembles for Increasing Model Transportability | BMC Medical Informatics and Decision Making | Full Text.” Accessed June 28, 2024. <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-022-01879-6>.
- Lee, Dong Yun, Byungjin Choi, Chungsoo Kim, Egill Fridgeirsson, Jenna Reps, Myoungsuk Kim, Jihyeong Kim, et al. “Privacy-Preserving Federated Model Predicting Bipolar Transition in Patients With Depression: Prediction Model Development Study.” *Journal of Medical Internet Research* 25 (July 20, 2023): e46165. <https://doi.org/10.2196/46165>.
- Maro, Judith C., Michael D. Nguyen, Joy Kolonoski, Ryan Schoepfle, Ting-Ying Huang, Sarah K. Dutcher, Gerald J. Dal Pan, and Robert Ball. “Six Years of the US Food and Drug Administration’s Postmarket Active Risk Identification and Analysis System in the Sentinel Initiative: Implications for Real World Evidence Generation.” *Clinical Pharmacology and Therapeutics* 114, no. 4 (October 2023): 815–24. <https://doi.org/10.1002/cpt.2979>.
- Reps, Jenna M., Ross D. Williams, Seng Chan You, Thomas Falconer, Evan Minty, Alison Callahan, Patrick B. Ryan, Rae Woong Park, Hong-Seok Lim, and Peter Rijnbeek. “Feasibility and Evaluation of a Large-Scale External Validation Approach for Patient-Level Prediction in an International Data Network: Validation of Models Predicting Stroke in Female Patients Newly Diagnosed with Atrial Fibrillation.” *BMC Medical Research Methodology* 20, no. 1 (May 6, 2020): 102. <https://doi.org/10.1186/s12874-020-00991-3>.
- Rieke, Nicola, Jonny Hancox, Wenqi Li, Fausto Milletari, Holger R. Roth, Shadi Albarqouni, Spyridon Bakas, et al. “The Future of Digital Health with Federated Learning.” *NPJ Digital Medicine* 3 (2020): 119. <https://doi.org/10.1038/s41746-020-00323-1>.
- Wong, Jenna, Daniel Prieto-Alhambra, Peter R. Rijnbeek, Rishi J. Desai, Jenna M. Reps, and Sengwee Toh. “Applying Machine Learning in Distributed Data Networks for Pharmacoepidemiologic and Pharmacovigilance Studies: Opportunities, Challenges, and Considerations.” *Drug Safety* 45, no. 5 (May 2022): 493–510. <https://doi.org/10.1007/s40264-022-01158-3>.