FinOMOP Swarm Learning - Deep learning for patient-specific modelling of Acute Myeloid Leukemia

Salma Rachidi[1], Hartmut Schultze[2], Vytis Vadoklis[3], Perre Gustafsson[4], Johansson Markus[5], Kauko Tommi[4], Anna Hammais[6], Kukkurainen Sampo[7], Niemelä Sami[7], Tuomas Hakala[3], Alexey Ryzhenkov[1], Valtteri Nieminen[8], Tomi Mäkelä[8], Oscar Brück[1], Joachim Schultze[9], Tarja Laitinen[2], Arho Virkki[3], Kimmo Porkka[1,8], Eric Fey[1,8]

1 HUS Helsinki University Hospital, 2 Hewlett Packard Enterprise, 3 Tietoevry Oy, 4 Hospital District of Southwest Finland, VARHA, Auria Clinical Informatics, 5 Istekki Oy, 6 THL Finnish Institute for Health and Welfare , 7 Tampere University Hospital, PIRHA, 8 iCAN Digital Precision Cancer Medicine Flagship, 9 German Center for Neurodegenerative Diseases

Background

As common data model for structured, real world data from hospitals, OMOP facilitates the construction and use of personalized models for precision medicine. The aim of FinOMOP Swarm Learning is to establish a comprehensive, privacy-regulation compliant health data network that will leverage the high-quality real-world medical data available in Finnish hospitals and other health data partners to cooperatively build and test personalized predictive models for precision medicine and optimizing patient care.

While building predictive models using machine learning and artificial intelligence (AI) is data hungry, data sharing is challenged by privacy (GDPR, local regulations) and logistic concerns (infeasibility of copying large volumes of patient-level data). Swarm learning has emerged as a powerful tool to cooperatively build predictive models without the need of sharing the raw patient-level data between the registry owners.[1]

Secondary healthcare in Finland has been centralized to five university hospitals. Hospital-level raw data residing in vendor-independent, secure data lakes provides a robust and granular source of clinical data for research and development. Of those, the three largest university hospitals have already adopted the OMOP common data model, thus providing harmonized data sources for federated analyses and machine learning, including deep learning.

Our aims are threefold. Firstly, to establish a swarm learning network between the key public health data owners in Finland. Secondly, to evaluate the technical and legal framework for performing swarm learning analytics in a secure and privacy-preserving manner within the FinOMOP network. And lastly, to perform technical and clinical proof-of-concept studies.

The envisioned outcomes are a unique population-based resource for national and international collaborative work, the international recognition of Finland as a highly-valuable, ready-to-go collaboration partner, attraction of EU and global funding, improved patient care in the participating hospitals through the use of the robust and clinically actionable AI models.

Methods

Overall approach:

1. Start with the three largest university hospitals in Finland, HUS in Helsinki, VARHA in Turku, PIRHA in Tampere, who together cover about 70% of the Finnish population.

2. Use a predictive model for acute myeloid leukemia (AML) as the first use case.[2]
3. Federated feasibility analysis leveraging OMOP to define the scope of the first AML model.
4. Develop the mathematical and computational framework to integrate time-to-event (survival) modelling into swarm learning, using OMOP as the data backend, and that can handle longitudinal data and missing values.
5. Implement and iteratively refine the swarm-learning technology in all three university hospitals, and seek certification from the national data authority (Findata).
6. Build a first model locally at HUS, then train and validate in the swarm using OMOP data from all three hospitals.

To build the first AML predictive model locally at the HUS, the following setup was used: We included all patients who were diagnosed with AML between 2000 and 2022 and had at least three blast measurements within 21 days after diagnosis. We extracted the following features from the OMOP CDM: LDH, hemoglobin, and blood count values; leukocytes, and platelets (number of cells/volume), blast, lymphocytes, monocytes, neutrophils, eosinophils, basophils (per 100 leukocytes in blood). We imputed missing values with -1, allowing the model to learn that -1 means missing. Data were split into training, test, and validation sets. We used a two-layer fully connected artificial neural network (ANN) based on PyTorch as the base model, connected to a survival model called DeepHit as the output layer.[3] DeepHit is a deep learning approach that can handle competing risks and provide individualized survival predictions. Dropout layer with a dropout rate of 0.4 were used to avoid the risk of overfitting. We trained the model using cross validation on the training split, and validated it on the independent test set. Kaplan Meier plots were created by splitting the cohort into three strata (upper, intermediate, and lower quantile) based on the model predicted patient-specific risk.

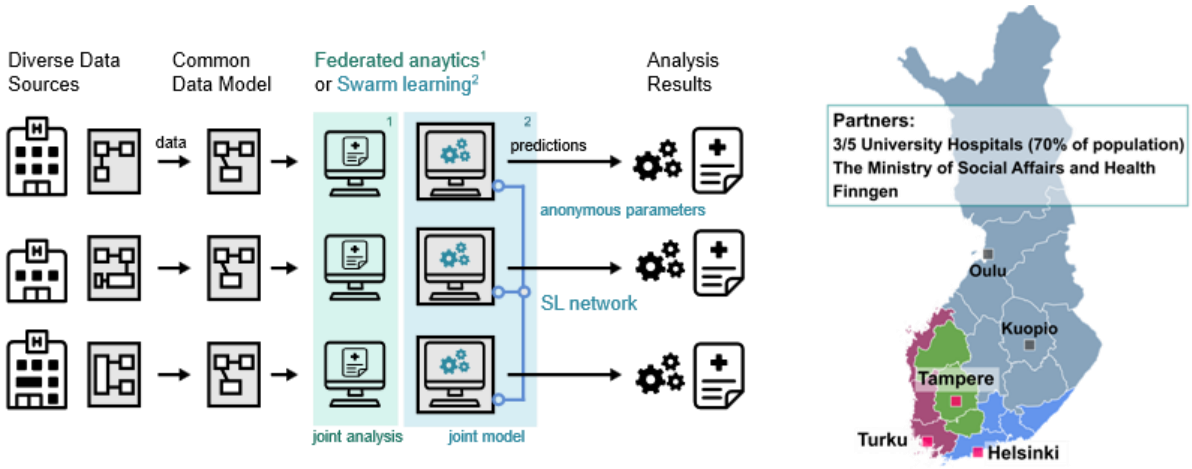

Figure 1. *Study concept.* ***Left:*** *EHR data from the three largest university hospitals in Finland were harmonized to OMOP enabling federates analytics (joint analysis) and swarm learning (joint model), a form of federated machine learning (ML). In contrast to federated ML, swarm learning does not require a central orchestrator. In contrast to [1]federated analytics where scripts are run locally and results can be shared asynchronously afterwards, [2]training the joint model (federated ML, swarm learning) requires the coordinated exchange of model parameter in real-time during training. Both, training data and personalized model predictions stay local, but can be shared afterwards if so desired. In contrast to federated learning, swarm learning logs all transactions, such as parameter updates during training, in a blockchain, thereby providing a secure, immutable audit log.* ***Right:*** *Map of Finland illustrating the catchment areas of the three participating university hospitals, covering about 70% of the Finnish*

population.

Results

1. **SL has been implemented in all three participating hospitals.** For this implementation, we designed a solution that meets the requirements of our secure processing environments in the hospitals as mandated by Findata. The architecture of this solution separates the communication with the external swarm network from the internal traffic within the secure processing environment (machine learning nodes and data-access) using a reverse proxy setup.
2. **Federated analyses using OMOP scripts scoping the available data** across the three hospitals, identified a set of measurements that are routinely taken for AML patients in all three hospitals.
3. **A deep learning framework was developed** that integrates deep learning and survival modelling into SL. This framework uses pytorch to implement artificial neuronal networks for deep learning with freely configurable hidden layers, followed by an output layer for survival modelling, that can either consist of DeepHit (time-discretized survival model) or pyCox (continuous model based on the cox proportional hazard assumption). Longitudinal OMOP data are preprocessed (time-discretized and flattened) to feed into the input layer, and missing values are imputed out-of-distribution (e.g. with -1 for data that is normally non-negative). The rationale is that such out-of-distribution imputation (raw feature values were strictly non-negative) allows the deep model to learn that -1 means missing (= not useful information).
4. **Results of the first predictive model using the HUS data only:** After applying the inclusion criterion of having at least three blast measurements within 21 days after diagnosis, 614 patients remained for the analysis. The data were split into training, validation, and test sets, of 393, 98, and 123 patients, accordingly. The model was trained on the training set using cross validation, and the best performing model was selected based on the validation set. The model converged quickly, requiring less than 100 epochs to reach the optimal performance. We experimented with different architectures of the ANN, varying the number of hidden layers and nodes. We found that the model performance was not sensitive to the architecture, and even a simple two-layer network with 16 nodes in each layer achieved good results. The final model was able to stratify the patients into three risk groups with markedly different survival probabilities (Fig. 1). 2-year survival for the model-predicted low versus high-risk patient groups was 71% vs 10% in the training and 52% vs 11 % in the independent test cohort. Further, we investigated how much longitudinal data (days post diagnosis) were needed to achieve good results. A expected, predictive power deteriorated with decreased availability of follow-up, with a marked loss of performance when less than 15 days of data were used. Note that the model can also provide individualized survival probability curves for each patient and the uncertainty of the predictions over time.
5. **A test run of the 3-node SL network has been performed** with the developed AML model architecture but using artificially generated “fake” data in the first instance. “Real” training of the AML across all three participating hospitals using real-world data is currently ongoing.
6. **The implemented swarm learning system has been audited** by the Finnish National Data Authority (Findata, https://findata.fi/en/) providing a first-of-its-kind certification for such federated system for the use of real-world hospital data.

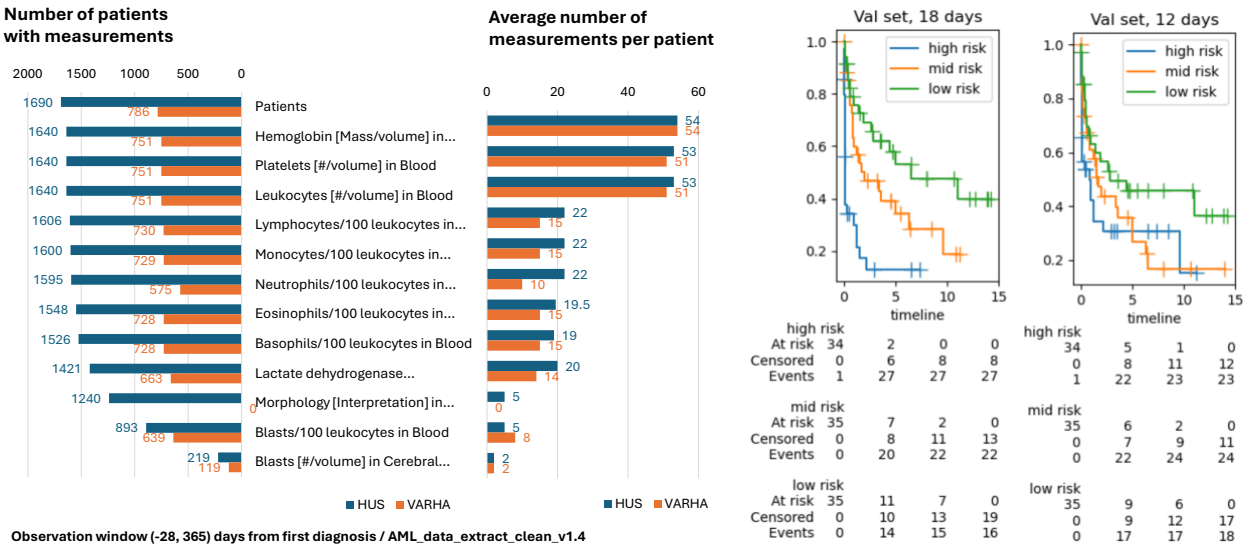

Figure 2. *Results AML model.* ***Left:*** *OMOP data availability for the selected model features at HUS (Helsinki) and VARHA (Turku).* ***Right:*** *Model performance in the validation dataset of HUS when 18 days and 12 days of feature data were used as indicated. Days measured relative to the date of first diagnosis.*

Conclusion

Building joint predictive models from distributed longitudinal, real-word data spread across multiple institutions is feasible and has great potential. OMOP and SL facilitate this model building. The framework is highly secure, as certified by Finland National Data Authority, and generally applicable to predictive problems, including predicting treatment responses, patient survival, adverse events, and optimal treatments. Further, our first use-case demonstrates that using dynamic, short term follow up data only - blood count measurements up to 21 days post diagnosis - holds valuable information for predicting long-term prognosis, including overall survival. These models will be further refined using more traditional baseline features (cytogenetics, mutation profiles, clinical characteristics). OMOP Genomics mapping is ongoing (see poster #9). In conclusion, our OMOP-based swarm-learning framework facilitates constructing, training, and iteratively maturing predictive models for precision medicine in multicenter network settings.

References

1. Warnat-Herresthal, S. *et al.* Swarm Learning for decentralized and confidential clinical machine learning. *Nature* **594**, (2021).
2. Malani, D. *et al.* Implementing a Functional Precision Medicine Tumor Board for Acute Myeloid Leukemia. *Cancer Discov* **12**, (2022).
3. Lee, C., Zame, W. R., Yoon, J. & Van Der Schaar, M. *DeepHit: A Deep Learning Approach to Survival Analysis with Competing Risks*. www.aaai.org.