PHEMS: validating novel federated ecosystems for analytics and synthetic data generation methods through real-world data investigations, particularly in the context of pediatric healthcare

New strategies in Health Data Sharing – Clinical Use Cases

Pediatric Hospitals as European drivers for multi-party computation and synthetic data generation capabilities across clinical specialties and data types (PHEMS, an EU-funded project) aims to establish an open and decentralized health data ecosystem for accessing health data across multiple European hospitals. Data federation (federated analytics and learning) will be applied to overcome significant obstacles in cross-border collaboration while complying with the European Union’s General Data Protection Regulations (GDPR) and AI Act. Additionally, PHEMS will develop an innovative data synthesis and anonymization pipeline for use in the federated ecosystem for rare disease research.

Variable categories
The variables of interest for each use case are further grouped into categories.

Variable overlap
An overview of the number of variables to be collected for each use case and their overlap.

1. Cardiology operations benchmarking
   Supporting creation of benchmarking standard and promote a culture of benchmarking across pediatric cardiac institutions, enabling the adoption of ‘best-practice’ across institutions
   Led by Great Ormond Street Hospital for Children (GOSH)

2. Pediatric Intensive Care Unit Sepsis
   Investigating the benefits of the federated ecosystem to develop, train and test algorithms to predict sepsis on a large scale between pediatric intensive care units in four large European children’s hospitals
   Led by Sant Joan de Déu Barcelona Children’s Hospital (HSJD)

3. Hematology - hemophilia
   Developing and testing a machine learning-based prediction algorithm to improve treatment for pediatric patients with hemophilia A or B
   Led by Erasmus University Medical Centre Rotterdam (Erasmus)

At the current state of the project, the achievement worth mentioning is the curation of the variables essential to the use case studies. This preliminary work, in the first months of the consortium, lays the foundation for the OMOP data standardization at the hospitals and the synthetic data generation.