The future of Personalized Surgery in Denmark: using OMOP and the OHDSI community tools to predict 90-days mortality after colorectal cancer surgery.

Rosen A, MD[1], Derian N, PhD[1,2], Grafkin P, MSc[3], Allakhverdiiev E[3], MD, Klebanov G, MSc[3], Drakos I, PhD[1,2], Gögenur I, MD[1]
[1]*Center for Surgical Science (CSS), Denmark*
[2]*Department for Personalized Surgery, CSS, Region Zealand, Denmark*
[3]*Odysseus Data Services Inc, USA*

Background

Colorectal cancer is the 3rd most frequent malignant disease worldwide, with 1,400,000 new cases yearly, causing more than 700,000 deaths. The best practice treatment is surgery for localized disease, with additional oncological treatment depending on paraclinical findings. The treatment modalities both cause benefits, but also put the patient at risk for increased morbidity and mortality.
The Danish healthcare data scene offers high granularity data from various registries and electronic health records. All national health data are integrable using a unique identifier, issued to every member of the Danish healthcare system. Hosting a detailed and near complete follow-up of the population's medical history makes the Danish health data scene a fertile field for exploring observational data and applying machine learning methods to predict health outcomes. Utilizing these data scene, the Center of Surgical Science is in continuous development of Decision-Support Tools, helping medical doctors identifying the probability of benefit and complications at an individualized level. The ultimate aim is to define personalized treatment pathways, empower the patient choice and reduce the risk of under- and overtreatment.

Methods

Datasets from Danish national databases for Colorectal Cancer, Anesthesia, Microbiology, Pathology, Cancer, Reimbursed Prescriptions and Causes of Death were integrated in an OMOP CDM. Additionally, data from two population studies were included as well as biochemistry measurements. Mapping from Danish classification (SKS, ICD10, ICD-O3 and ATC) to OMOP was performed on 19,565 concepts. Ongoing quality control is being performed by a multidisciplinary team of data scientists and medical doctors. To predict the 90 days mortality risk for CRC patients after surgery, a model was built and trained using LASSO regression and the OHDSI patient level prediction framework on CDM data from 2,286 patients, including 158 deaths within 90 days.

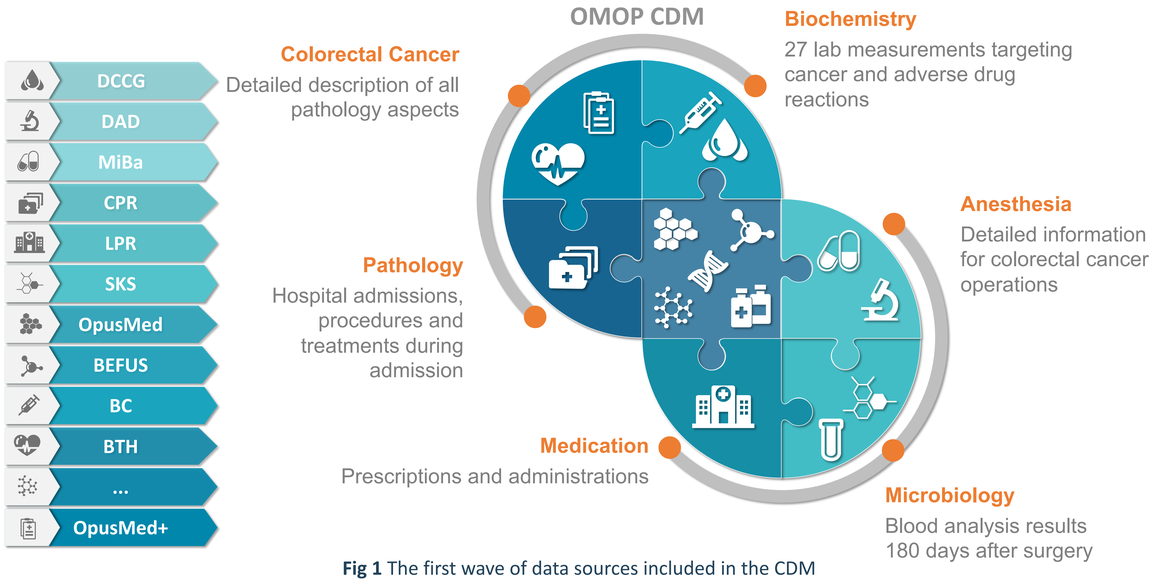

Fig 1 The first wave of data sources included in the CDM

Results

The compiled OMOP CDM managed to integrate heterogeneous data from 11 sources covering around 87% of the involved concepts. The use of OHDSI community tools for cohort creation and patient level predictions was well-received by both clinical and non-clinical researchers, allowing people with different analysis skills to accelerate and advance their work. The developed prediction model for 90 days mortality risk is constantly being refined, with current AUC at 0.78.

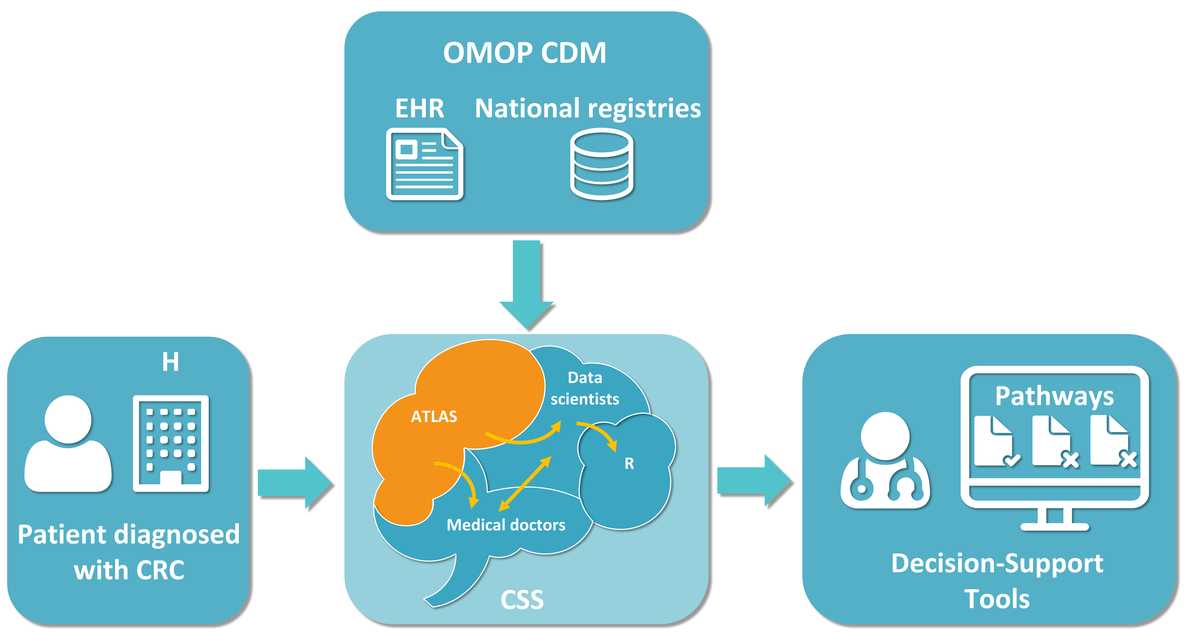

Fig 2 The OHDSI data flow in CSS. Translating health data into improved treatments.

Conclusions

The mapping of various Danish health data sources into the OMOP CDM was successful with satisfactory coverage and preserved granularity. The proof of concept study, regarding mortality risk for patients undergoing surgery for colorectal cancer, is demonstrating encouraging results and highlights the clinical benefits of the OMOP CDM and the OHDSI community tools for personalizing the treatment pathway.
The Danish health data scene enhances the possibilities to create bedside tools for personalized medicine by having a vast number of national registries, spanning for several decades, and being able to integrate these data with the electronic health record. Due to the depth, validity and relevance of such Big Data scene we foresee a vast improvement for the patients and the healthcare actors, through the translation of health data into improved prognosis, diagnosis and treatment.

Contact: anwr@regionsjaelland.dk