

Universal Screening for Lynch Syndrome in Switzerland: Implementation of a Precision Medicine Protocol for Population Health.

Background: Lynch Syndrome (LS) is a genetic condition predisposing to colorectal and endometrial cancer, as well as a battery of other malignancies. It is caused by mutations in the DNA mismatch repair genes (*MSH1*, *MSH2*, *MLH6*, *PMS2*) or *EPCAM* deletions that are inherited in an autosomal dominant way. This suggests that first-, second-, and third-degree relatives of index cases (proband) have 50%, 25%, and 12.5% probability of carrying the same pathogenic variant. This offers a unique opportunity for effective cancer control through cascade screening. An implementation framework has been developed to promote this concept.

Notation:

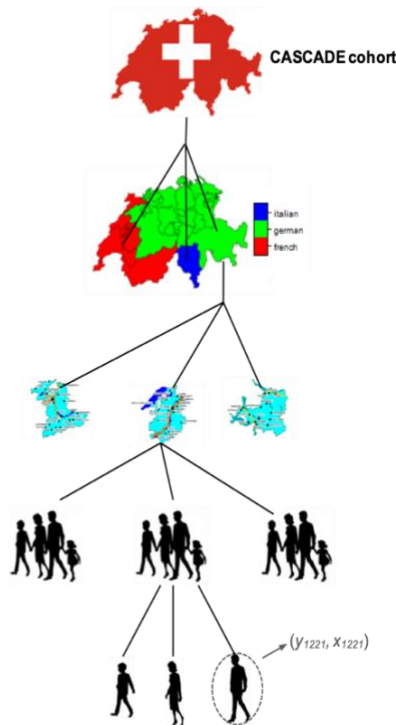
Person: $sijk$
Outcome: Y_{sijk}
Predictors: X_{sijk}

Language region: $s=1, \dots, S$

Community: $i=1, \dots, I_s$

Family: $j=1, \dots, J_{si}$

Person: $k=1, \dots, K_{sij}$



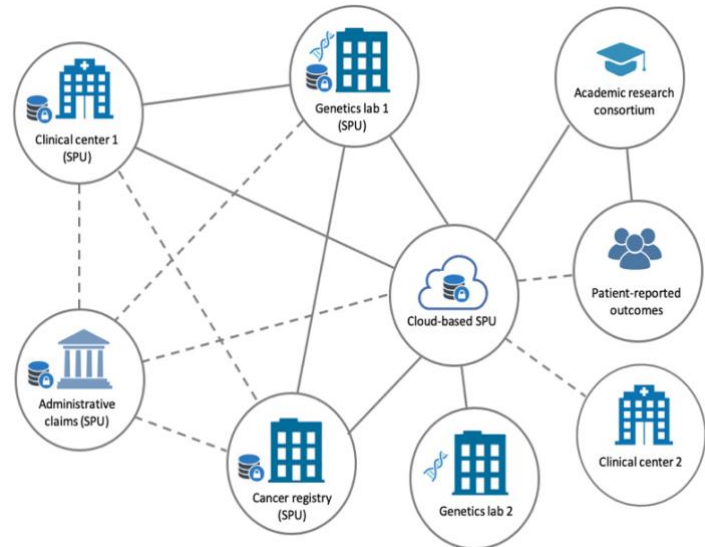
Purpose: Assessment of real-world health outcomes is central to fulfillment of personalized medicine. This largely depends on the possibility of collecting, linking, and reusing large amounts of health-related data. To achieve this goal there is a need for greater integration, and increased interoperability between different data sources (e.g., clinical, epidemiological, insurance claims, and patient-reported outcomes). Access to this sensible information is typically hindered by privacy and security risks, and further complicated by governance-related, and other management constrains. Identifying new ways for sharing these data, while controlling for associated risks is necessary to accelerate future medical research. Here we propose a **data cooperative model** for clinical/genetic, epidemiological, and registry data integration for a more efficient LS patient diagnosis, management, and care in Switzerland.

Proposed research model: The designed approach sets the foundations for the creation of a **learning healthcare system** around LS management and care in Switzerland, which will actively, and prospectively integrate insights from multiple levels (i.e., patient, provider, institution, and system level). The proposed approach has the potential to improve statistical power and facilitate economies of scale, while at the same time mitigate the risks and develop shared solutions (e.g., streamlined ethical approvals for LS research). This infrastructure can support **pragmatic trials**, comparative effectiveness studies, as well as accelerate research efforts to improve population health, including the translation of genomics into clinical practice (i.e., a T3->T4 transition in the translational medicine continuum).

Existing infrastructure: The research plan will build on the developments of the **Swiss Personalized Health Network (SPHN)**, and in particular the **BioMedIT** infrastructure. Each participating research center will be affiliated to one of three existing **High Performance Computing (HPC)** facilities, namely the BioMedIT nodes. Coordinated by the **Swiss Institute of Bioinformatics (SIB)**, BioMedIT provides a secure infrastructure for computing and storage resources operating under the scope of the SPHN Information Security Policy. **sciCORE**, the scientific computing center of the University of Basel, acts as the main BioMedIT node in Basel. **OpenStack** will provide the security framework, as well as computing and storage resources to all affiliated researchers.

Advanced cryptography for genetic data: The **Data Protection in Personalized Health (DPPH)** consortium has created a platform that allows sharing of data between Swiss institutions (e.g. medical centers, laboratories, registries etc.) in a scalable, secure, and privacy-conserving way. The operating system developed called **MedCo** uses lattice homomorphic encryption, distributed ledger technologies (blockchains), multiparty computation, and control systems for advanced security, privacy, and interoperability, while at the same time allowing use of widespread tools to conduct data analysis (e.g., i2b2, TranSMART, and SHRINE).

A conceptual design. The data cooperative approach is achieved by distributing trust in a set of different storage and processing units (SPUs), where participating sites can securely store their **encrypted data** using a shared key that is collectively generated. This creates a **secure, federated and interoperable data network** where investigators can set up research queries as if it was a single unified database. Only encrypted data from different STUs can be accessed and used, and this can be achieved without the need of decrypting them. Thus preserving safety and security of sensitive, personal, identifiable information.



Design of pilot study: Proof-of-concept will be achieved by using both **simulated** and **real data** from institutions that will agree to participate in this initiative. Genetic testing centers will be contacted, and invited to participate in the study. Upon acceptance and approval of Ethics Committee, hardware installations (in- house, or cloud-based) will be envisaged. Only simulated data will be used at first. The MedCo platform will be used to set up queries on those simulated dataset to validate interoperability between participating nodes. Example query: “*how many male patients, age <50 identified with a PMS2 pathogenic variant*”, which will return an N value, stratified to N_1 and N_2 from each participating center.

Impact: This will allow independent data collection and sharing in the form of a unified **virtual registry**. If this pilot phase is successful, an attempt will be made to also link other nodes willing to participate, prioritizing Universitätsspital Basel to integrate clinical information of LS patients, and potentially a cancer registry (i.e., Krebsregister beider Basel) to integrate cancer incidence and mortality. It is estimated that Universal Screening for LS in Switzerland can prevent 131-278 new colorectal cancer cases, and dozens other related malignancies on an annual basis, which translates into **thousands of cancer cases in a decadal period**, and a substantial long-term public health benefit.

Relevance: The **Swiss Personalized Health Network (SPHN)** has initiated a strategic driver project, the **Swiss Personalized Oncology (SPO)** that aims to standardize the molecular profiling of tumor biopsies in Switzerland, and facilitate research to identify new predictive biomarkers. Screening for LS involves: i) Microsatellite instability (MSI) analysis and/or immunohistochemistry to determine the expression of MMR proteins in tumor tissues; ii) additional MLH1 promoter methylation testing to eliminate the possibility of loss of MLH1 expression due to epigenetic effects, and/or identification of somatic **BRAF V600E** pathogenic variants; followed by iii) full germline sequence and large rearrangement testing of MMR, and EPCAM genes for final diagnostic confirmation. Immune checkpoint inhibitors, anti-PD-1 and anti-PDL-1 antibodies (*pembrolizumab*, and *nivolumab*) have recently gained FDA-approval to be used against metastatic solid tumors that shows evidence of MSI. This creates also an opportunity for collaboration with the pharmaceutical industry in the prospect of medicines development and/or pharmacovigilance (real-world monitoring of adverse events).