

Examples of good practice in sensitive data management

Marián Hajdúch

Institute of Molecular and Translational Medicine, Faculty of Medicine and
Dentistry, Palacky University

European Infrastructure for Translational Medicine (EATRIS-ERIC, EATRIS-CZ)



Co-funded by
the European Union



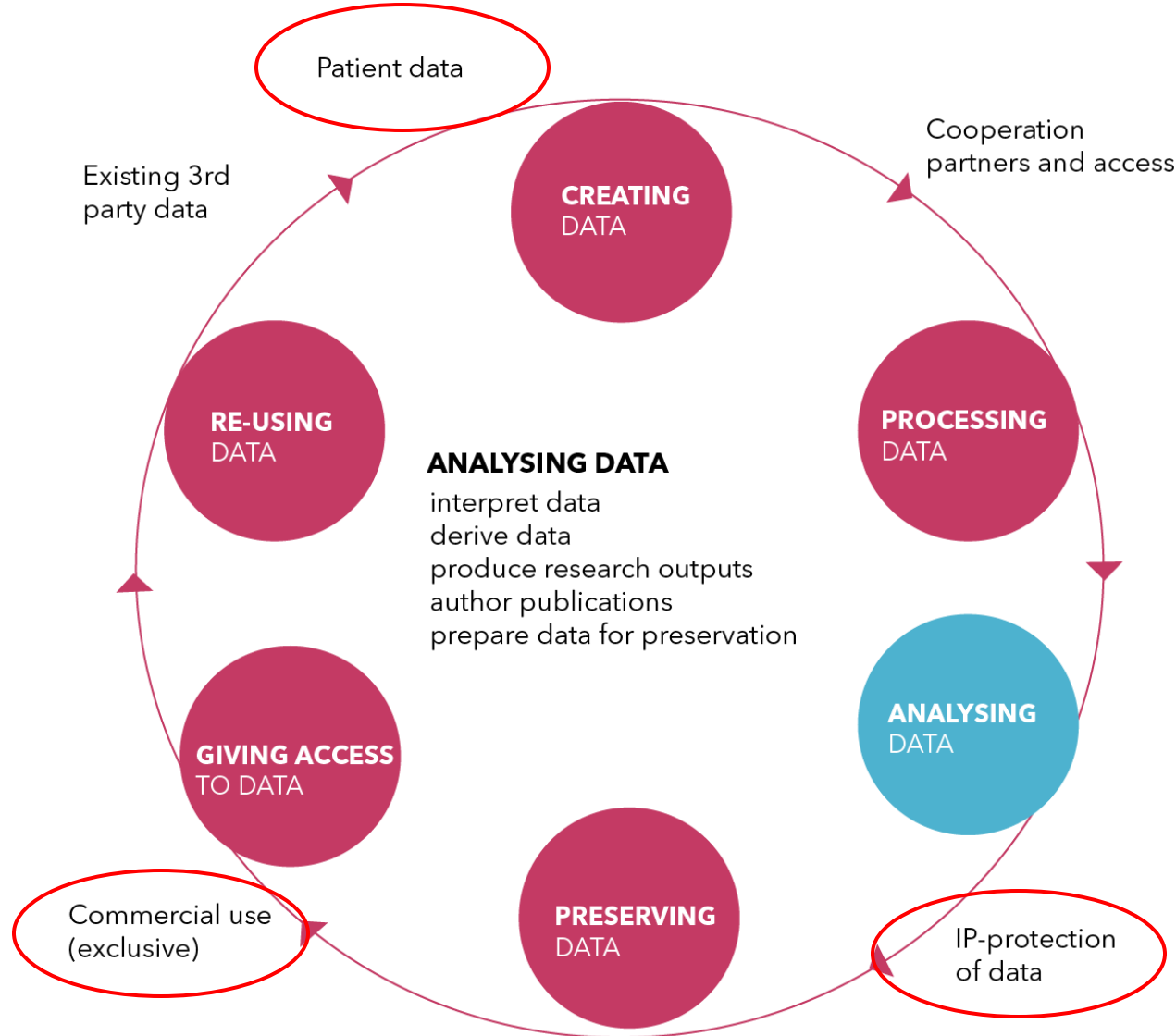
M U N I
I C S

IPs EOSC-CZ registration number
CZ.02.01.01/00/22_004/0007682



eattris-cz

Data stewardship in health (human) based research: How to be FAIR and OPEN?



Sensitive data management at IMTM/EATRIS-CZ



- Users and their rights
 - Information system Admin



- Storage of clinical/parametric data
 - Information system ClinData

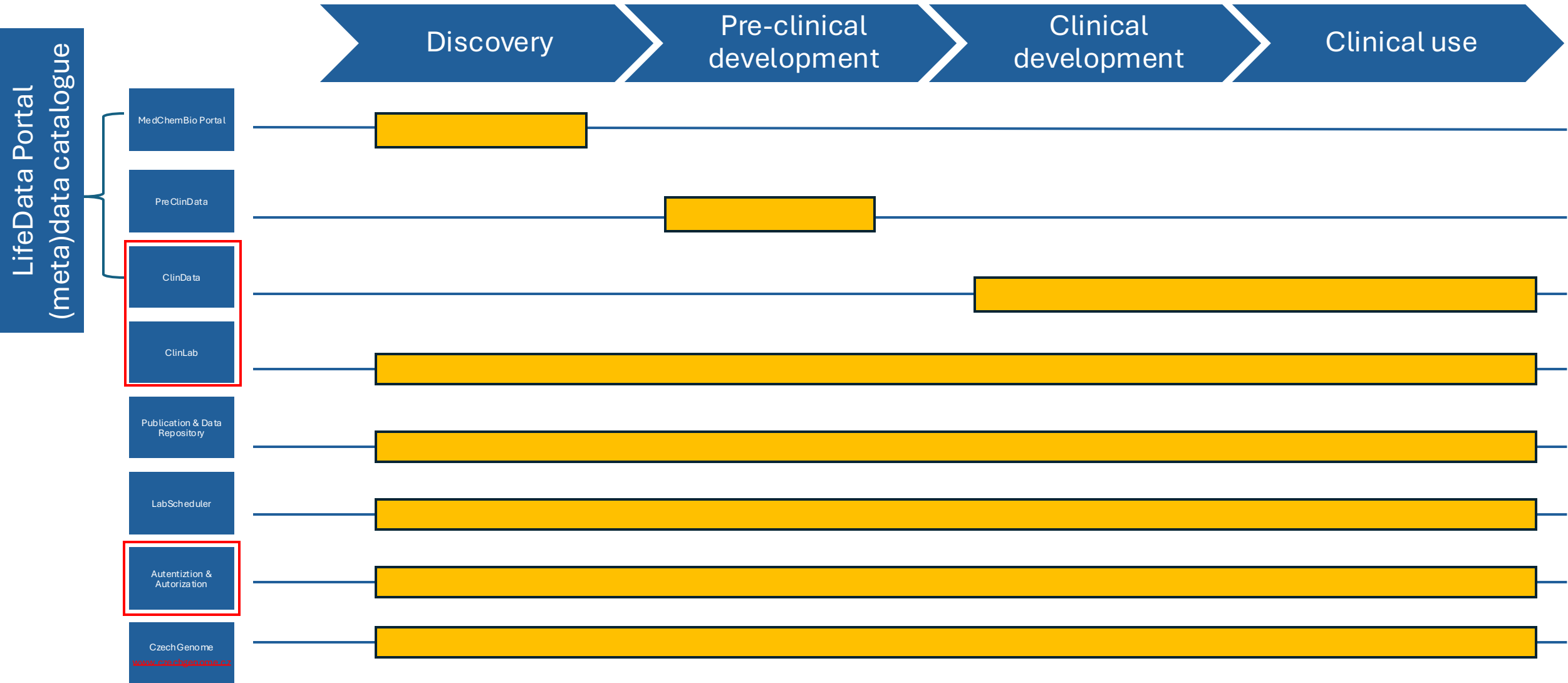


- Storage of laboratory data
 - Information systems ClinLab, PreClinData, MedChemBio Portal, LabScheduler



- Storage of big omics data
 - Object storage CEPH, Own Cloud
- (Meta)data direktory
 - LifeData Portal

IMTM/EATRIS-CZ Digitalization Hub



Admin

User management system

IMTM Admin

User administration



AUTHENTICATION (WHO IS WHO)

User definition:

Name, mail, phone, organization, team



AUTHORIZATION (WHAT IS ALLOWED)

Two types of restrictions:

- Access to functionality
- Access to data

PRIVILEGE

Every feature in system has its own item in admin

ROLE

Virtual entity connecting more privileges together:

- ClinData project admin
- ClinData principal investigator
- ClinData project data manager

The screenshot displays the IMTM Admin interface. At the top, there's a navigation bar with 'Administration' and 'Site configuration' tabs. Below this, the 'Privilege list' is shown, featuring a table with columns for 'Name' and 'Description'. The table lists various roles like 'ROLE_USER', 'ROLE_REGISTRY', and 'ROLE_REGISTRY_PATIENT'. A 'new' button is visible at the top left of the table. To the right, the 'Entities with access to the selected Privilege' section is active, showing a table with columns for 'Name' and 'Description'. This table lists specific roles and their descriptions, such as 'Laboratory 65Biocev' and 'Supervisor'. At the bottom, the 'Privilege details - edit' form is visible, showing fields for 'Name' and 'Description' with the values 'ROLE_REGISTRY_PATIENT_ADD_STUDY' and 'Registry add patient to study' respectively.

ClinLab

Laboratory information system for management of samples from various projects.



ClinLab

Laboratory information system

Based on CovIT

Specialized system for management of Sars-Cov2 samples

Functionality of LIMS for single laboratory procedure



ClinLab is generic system

Evidence and management of samples in IMTM/BBMRI biobank

Other LIMS functionality planned and being developed

ClinLab

Sampling site



Sample registration



Informed consent signed



Biobanking - odběrové místo

Status **In progress**

Save and close

Close

CRF history



Pdf



Vzorky - odběrové místo

Jméno

Příjmení

Rodné číslo

Pohlaví

☐ M ☐ F

Datum narození

Telefonní číslo

Email

Město

PSČ

Zdravotní pojišťovna

Přijem vzorku

Datum odběru

Typ materiálu na příjmu *

Externí ID vzorku I.

Externí ID vzorku II.

Externí ID vzorku III.

Informovaný souhlas

Informovaný souhlas je k
dispozici☐ Ano ☐ Ne

ClinLab Laboratory



Action	ID	First name	Last name	Address	Sex	Birth date	Enrollment date
Open	1	Jan	Novák		M	1972/01/01	2024/06/14 14:47:25

SMS

Open

Visit: default

Visit date: 14/06/2024

Form: Vzorky

25/29

In progress

SMS

Open

Visit: default

Vi

SMS

Open

Visit: default

Vi

New sample

Visit: default

Extraction of
aliquots

Sample storage

Date of receiving sample

14/06/2024

Reason

Indication by a doctor

Study

ID sample IMTM

IMTM-null-P1-S00000124

Number of aliquots

1

0

ID aliquot

IMTM-null-P1-S00000124-A01

Tube number

4565654

Material to store

RC RNA celková

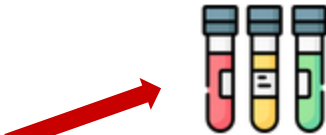
Amount

12

ml

Storage temperature [C]

-80





Direct communication channel with patients or medical professionals

- Gathering patient's data via predefined registration forms
- Patients can receive results or certificates
- Authentication by information sent via SMS or emails
- Login by personal id and unique hash value sent in SMS



COVIDLAB

Pick up the test result

Your code from SMS:

Your personal id:

I am a foreigner or I do not have Czech birth identification number.

Show test results



Online výsledky Kontakt

Datum odběru vzorku: 01/04/2020
Datum doručení vzorku do laboratoře: 23/03/2020
Datum provedení vyšetření: 10/03/2020

Jméno: **Pacient**
Příjmení: **Dva**
Pohlaví: **M**
Datum narození: **24/02/1994**
Rodné číslo: **940224/9602**
Zdravotní pojišťovna: **201 Vojenská zdravotní pojišťovna**
Adresa: **Brnenska 5**
Telefonní číslo: **731119683**

Výsledek vyšetření přítomnosti SARS-CoV-2 pomocí rRT-PCR:
Výsledek testu SARS-CoV-2 nelze stanovit kvůli špatnému odběru. Doporučujeme odběr opakovat.

Metoda vyšetření přítomnosti SARS-CoV-2:
Vzorek byl vyšetřen metodou rRT-PCR za použití detekce virálního genomu pomocí LightMix® Modular Sarbecovirus E-gene a LightMix® Modular Wuhan CoV RdRP-gene kitu s kontrolou EAV RNA extraction control (Roche). Negativní výsledek metody nevylučuje přítomnost viru SARS-CoV-2 pod detekčním limitem metody.

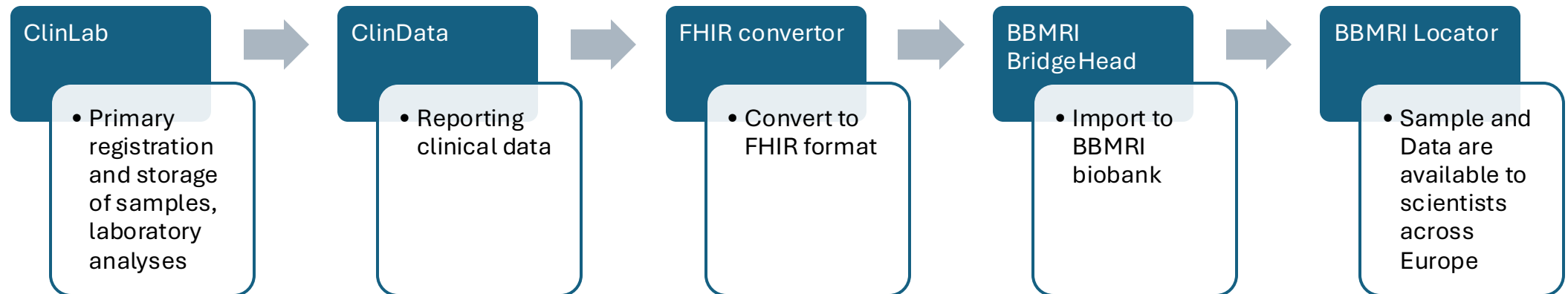
Vytisknout výsledek

Stáhnout výsledek



Connection to European bio bank BBMRI

Samples registered in ClinLab can be reported to BBMRI
European bio bank

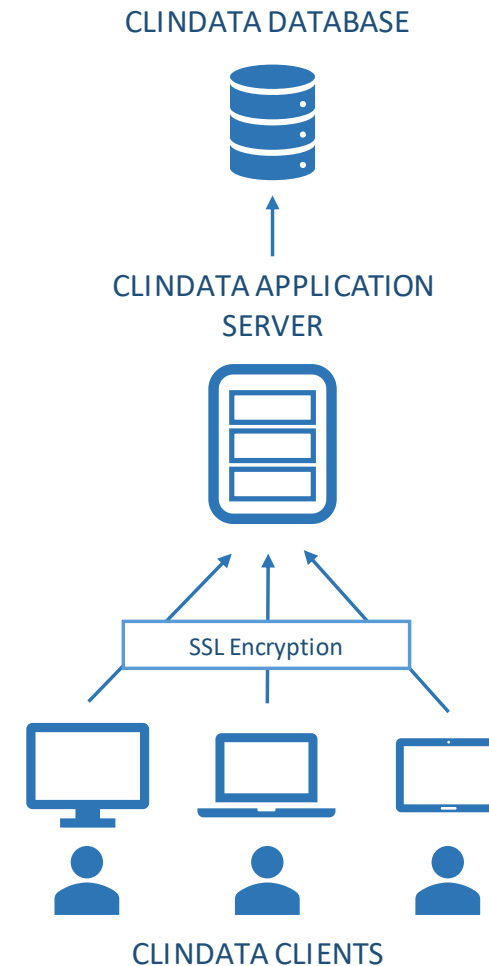


ClinData

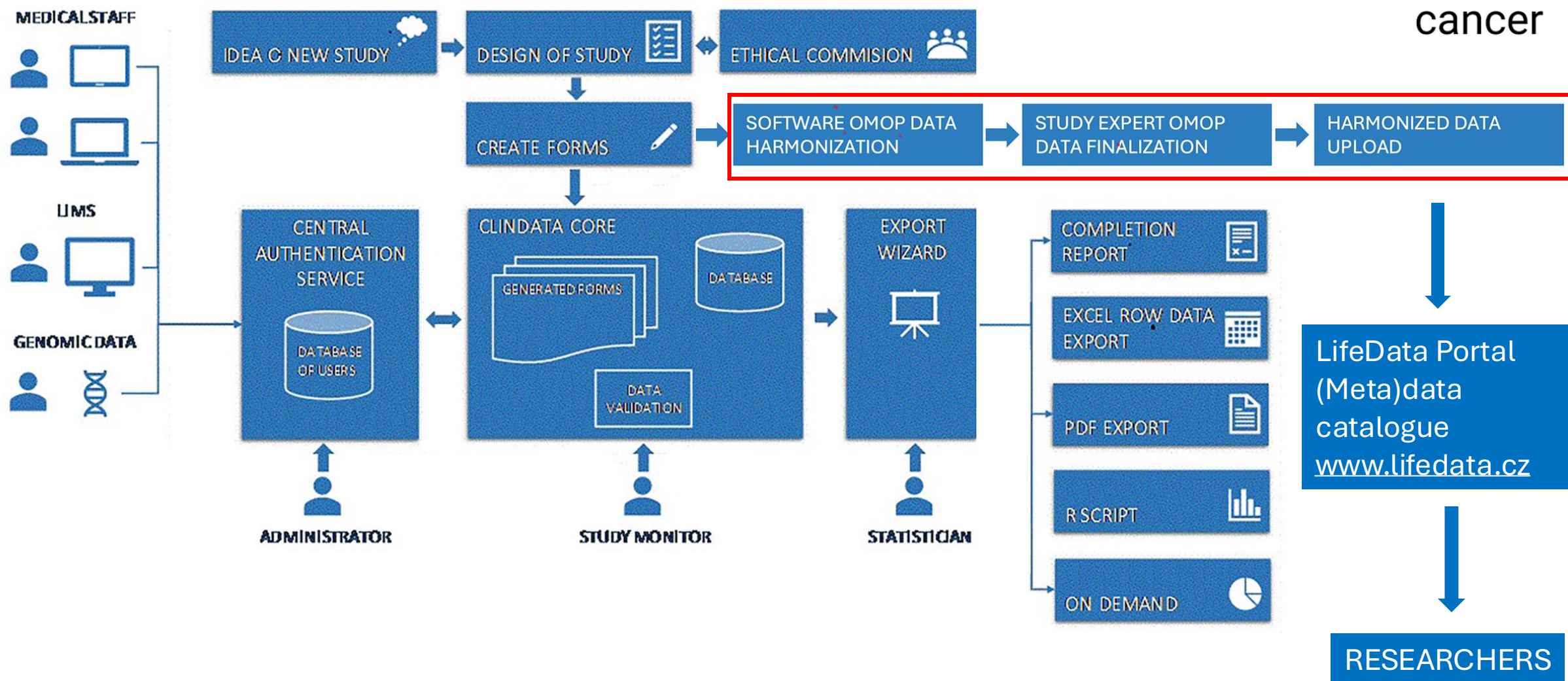
Comprehensive information system for data stewardship in clinical trials, registries and other clinical/scientific databases.

ClinData – Parametric data storage

- ClinData software has web-based client/server architecture, the application runs on the server, the user is connected through a web browser from computer, tablet or smart phone.
- All communication between the client and the server is secured by SSL encryption, which is currently industry standard.
- Available at <https://clindata.imtm.cz/>
- Manual: <https://confluence.imtm.cz/x/NAFp>
- Leaflet: <https://confluence.imtm.cz/x/VwCzBQ>
- Service desk

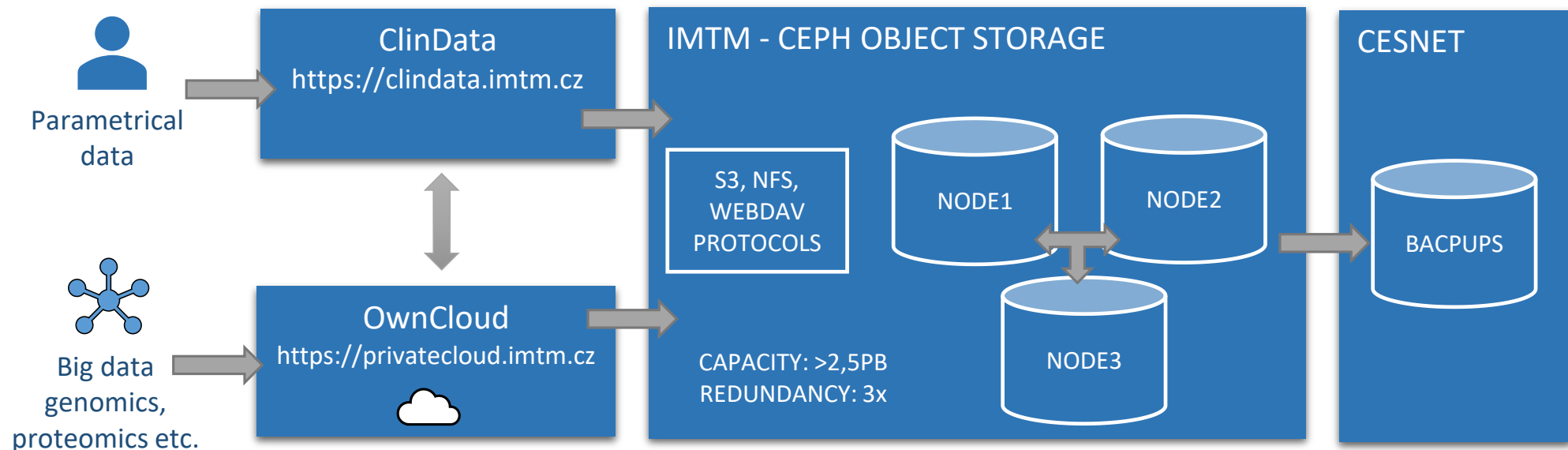


ClinData – basic diagram and data harmonization



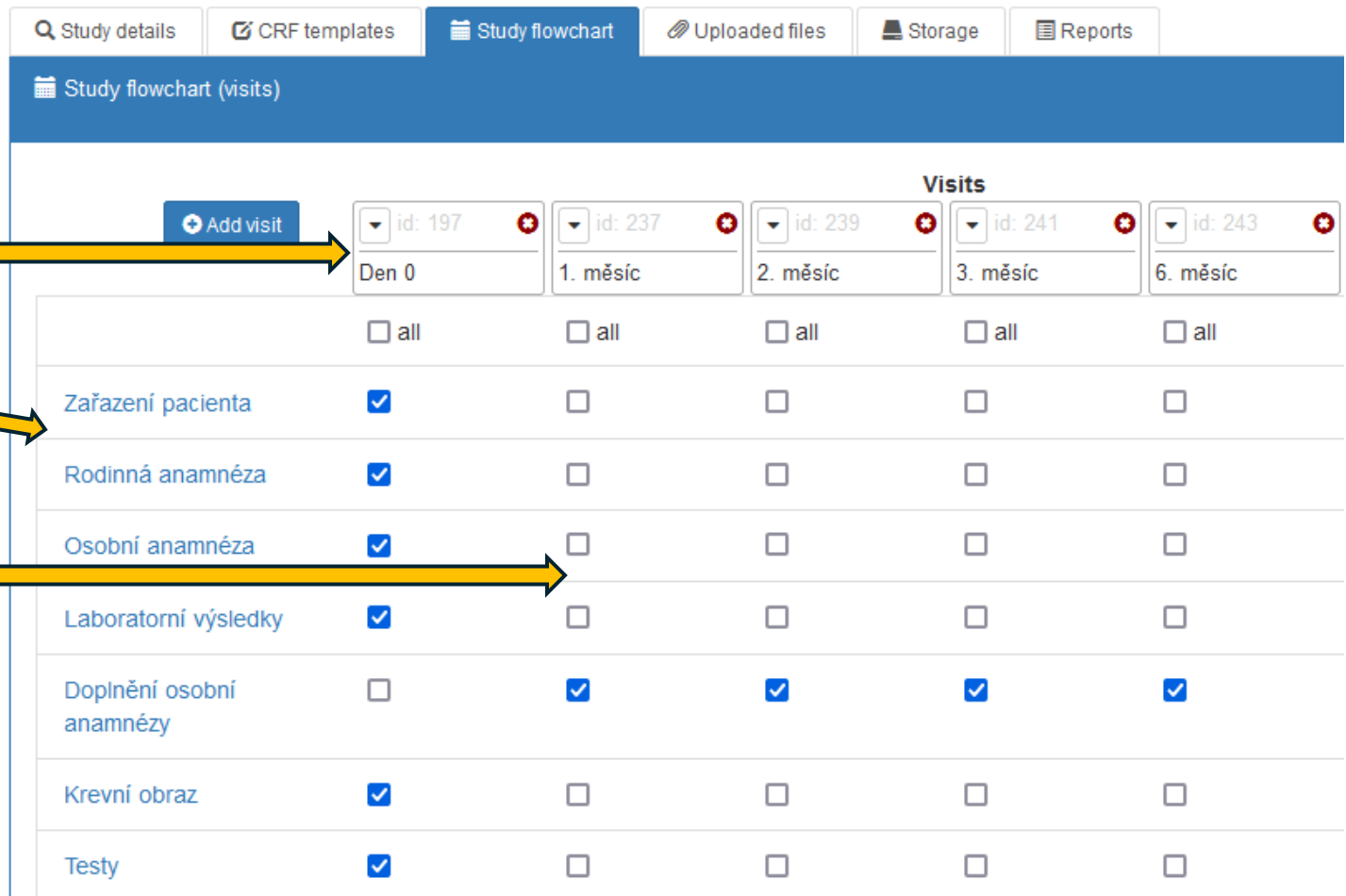
CEPH – Big omics data storage

ClinData software is connected to IMTM Object storage which allows storing of big scientific data (imaging, genomic, proteomics) in safe way. Capacity of the object storage is 2.5PB. Objects can be directly linked to ClinData forms to allow accessing all data from one place.



ClinData - study model

- Basic entity is **study**
- Study defines one or more **events**/visits (schedule)
- Study contains one or more electronic **forms** (eCRF)
- Every event can have **one or more forms**
- Events can be planned, repeated or ad-hoc



	id: 197	id: 237	id: 239	id: 241	id: 243
	Den 0	1. měsíc	2. měsíc	3. měsíc	6. měsíc
	☐ all	☐ all	☐ all	☐ all	☐ all
Zařazení pacienta	☒	☐	☐	☐	☐
Rodinná anamnéza	☒	☐	☐	☐	☐
Osobní anamnéza	☒	☐	☐	☐	☐
Laboratorní výsledky	☒	☐	☐	☐	☐
Doplnění osobní anamnézy	☐	☒	☒	☒	☒
Krevní obraz	☒	☐	☐	☐	☐
Testy	☒	☐	☐	☐	☐

ClinData – eCRF examples

Study

A-C-G-T

Status

In progress

Back to patient

A 4.1. Laboratorní výsledky

Soubory

WGS

 3_UP0001_210628_A00703_0086_AHVVJ3DSXY_S3_R1_001.fastq.gz 31.86 GB

 3_UP0001_210628_A00703_0086_AHVVJ3DSXY_S3_R2_001.fastq.gz 34.28 GB

aCGH

Metabolomics

metDNA Seq

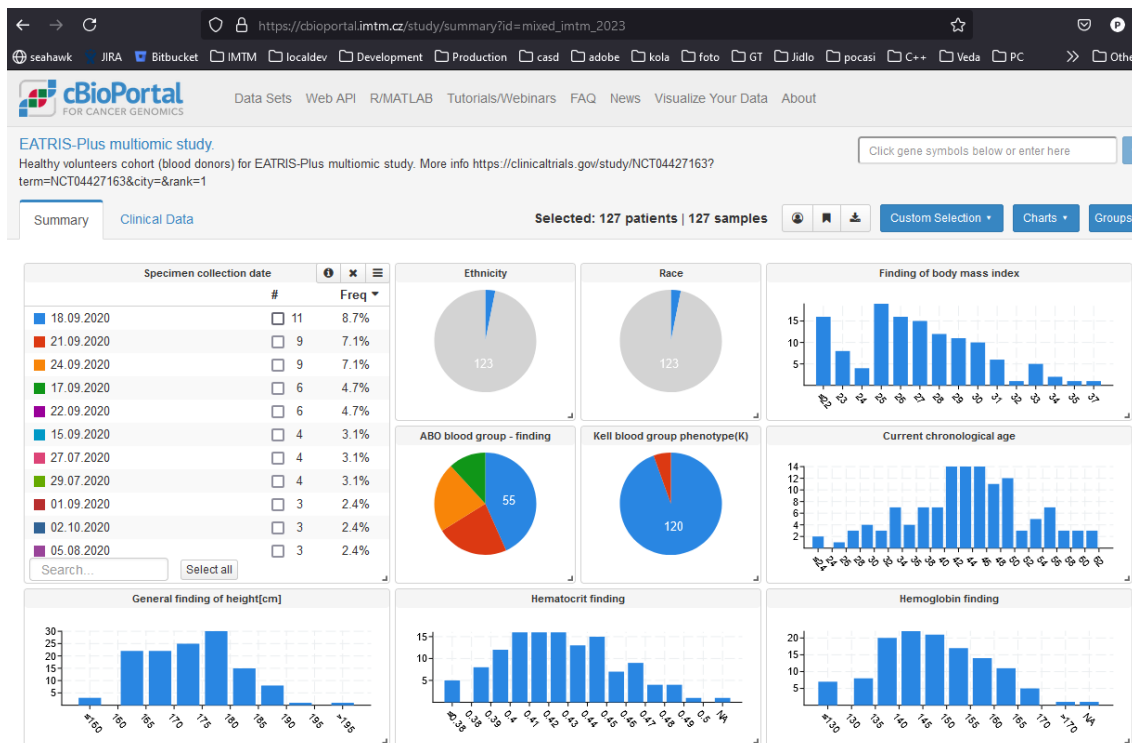
[illegible]

Main obstacles:

1. Some terms are not standardized yet - e.g. Progression free survival, some treatment regimens
 - The list of the missing parameters was discussed within OMOP Oncology workgroup meeting during OHDSI European symposium 2024.
 - If there is a specific parameter to be mapped to the standard OMOP vocabulary, this can be done by us during the regular OMOP Oncology workgroup meetings.
 - The new OMOP vocabulary update release is planned for October/November 2024.

ClinData – cBioPortal (cancer data primarily)

- Local IMTM instance of cBioPortal - <https://cbiportal.imtm.cz>
- Portal to FAIRify data from ClinData – only for cancer genomics
- Development of new public portal for FAIRification of other types studies



BioPortal

[Studies](#)[Study Details](#)[Data](#)[Graph And Summaries](#)[mullemar](#)

Details

Sample size: 300 participants

Age range: 20-70 years

Median age: 45 years

Gender ratio (M:F): 1:1.2

Baseline systolic BP: 130 mmHg

Baseline diastolic BP: 85 mmHg

Participants with prior conditions: 30%

Average BMI: 27 kg/m²

Smokers in sample: 20%

Median follow-up period: 5 years

Treatment adherence rate: 85%

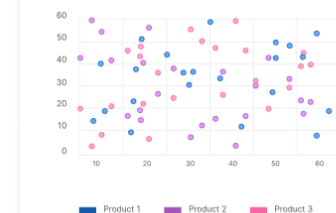
Placebo group size: 150 participants

Treatment group size: 150 participants

PRIMARY TEXT

5.987,34

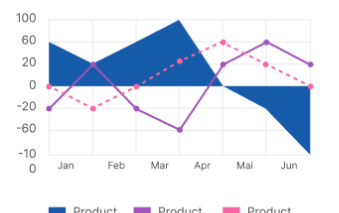
Secondary text



PRIMARY TEXT

5.987,34

Secondary text



PRIMARY TEXT

5.987,34

Secondary text

PRIMARY TEXT

5.987,34

Secondary text

ClinData – LifeData Portal www.lifedata.cz

LifeData Portal

Studies Study Details Data Graph And Summaries

Study Name:

Enter study name

Principal Investigator:

Enter principal investigator

DOI:

Enter DOI

Primary Diagnosis:

☐ Disease

☐ Population Study

☐ Healthy

Other:

Enter any string

Q Search

Study Name	Principal Investigator	DOI	Details
Assesment of Multiomics Profiles in Health and Disease – Correlation With the Disease Phenotype.	Marian Hajduch, MD, PhD.		Select →
Study of resistance to cytostatic treatment in oncological patients	Ass. Prof. Marian Hajduch, MD, Ph.D.		Select →

LifeData Portal

Studies **Study Details** Data Graph And Summaries

Basic Information

Name: Assesment of Multiomics Profiles in Health and Disease – Correlation With the Disease Phenotype.

Type: STUDY

DOI:

Status: IN_PROGRESS

Start Date: 2018–09–01

Finish Date: 2029–03–30

Primary diagnosis: Genetic Predisposition to Disease

Enrollment Criteria: Subjects with age 18 – 68 years. (for the first 1100 subjects): Healthy volunteers without genetically dependent disease and without such a disease in a family. Healthy volunteer without the preliminary evidence of civilizational diseases such as hypertension, diabetes, autoimmune and tumor diseases or acute infectional diseases; clinically manifesting cardiovascular or pulmonary disability. Subject without permanent of long-term medication in the time of biological sampling.

Description

This study will determine reference genomic, transcriptomic, proteomic and metabolomic profiles in Czech population. Initially, there will be 1000 healthy volunteers, with the planned expansion to 10.000 participants (healthy volunteers and patients with different types of disease). Formation of the reference database of healthy volunteers and their parameters will allow a correct interpretation of the potential pathological findings in patients. It is very important to obtain healthy controls from the region of the Czech Republic, Central Europe respectively; since it is not possible to reliably compare ethnically and geographically diverse populations, which have generated in a different context and where the diseases manifest with other etiology and phenotype. Although, in the limited measure, the similar molecular data exist in foreign databases, these are not compiled from the inhabitants of the Czech Republic, Central Europe not even from Slavic population. Study participants may volunteer for archiving of remaining biological materials for future studies.

Abstract

This study will determine reference genomic, transcriptomic, proteomic and metabolomic profiles in Czech population. Initially, there will be 1000 healthy volunteers, with the planned expansion to 10.000 participants (healthy volunteers and patients with different types of disease). Formation of the reference database of healthy volunteers and their parameters will allow a correct interpretation of the potential pathological findings in patients. It is very important to obtain healthy controls from the region of the Czech Republic, Central Europe respectively; since it is not possible to reliably compare ethnically and geographically diverse populations, which have generated in a different context and where the diseases manifest with other etiology and phenotype. Although, in the limited measure, the similar molecular data exist in foreign databases, these are not compiled from the inhabitants of the Czech Republic, Central Europe not even from Slavic population. Study participants may volunteer for archiving of remaining biological materials for future studies.

ClinData – LifeData Portal www.lifedata.cz

LifeData Portal

StudiesStudy DetailsDataGraph And Summaries

Search by name, doi, gender...

Export Result

Prev

Next

Study ID	Patient ID	Sample ID	ABO blood group - finding	Ethnicity	Finding of body mass index
mixed_imtm_2023	2759	20-0408-DG-11583	A		28
mixed_imtm_2023	2898	20-0409-DG-11584	B		26
mixed_imtm_2023	2706	20-0414-DG-11590	O		26
mixed_imtm_2023	3774	20-0415-DG-11591	O		26
mixed_imtm_2023	2675	20-0418-DG-11596	A		25
mixed_imtm_2023	2740	20-0419-DG-11597	B	Czech	27
mixed_imtm_2023	2552	20-0425-DG-11607	AB		37
mixed_imtm_2023	2710	20-0426-DG-11610	B		31
mixed_imtm_2023	2708	20-0427-DG-11611	O		26
mixed_imtm_2023	2547	20-0428-DG-11612	B		30
mixed_imtm_2023	2836	20-0432-DG-11617	A		34
mixed_imtm_2023	2803	20-0433-DG-11618	A		31
mixed_imtm_2023	2830	20-0434-DG-11619	B		27
mixed_imtm_2023	2843	20-0435-DG-11620	A		29
mixed_imtm_2023	2558	20-0436-DG-11622	A		28
mixed_imtm_2023	4156	20-0437-DG-11623	A	Czech	22
mixed_imtm_2023	2550	20-0441-DG-11628	B		28
mixed_imtm_2023	2715	20-0442-DG-11631	B		23

LifeData Portal

StudiesStudy DetailsDataGraph And Summaries

Distribution of ABO Blood Groups
(Number of smaples: 127)

22.83%

22.05%

43.31%

Distribution of Gender
(Number of smaples: 127)

49.61%

50.39%

Race
(Number of smaples: 127)

96.85%

Kell blood group phenotype(K)
(Number of smaples: 127)

94.49%

Consent status
(Number of smaples: 127)

100.00%

Cancer Studies
(Number of smaples: 127)

100.00%

Number of Samples Per Patient
(Number of smaples: 127)

100.00%

OS_STATUS
(Number of smaples: 127)

100.00%

Rh blood group phenotype(c)
(Number of smaples: 127)

20.47%

79.53%

Rh blood group phenotype(C)
(Number of smaples: 127)

30.71%

69.29%

Rh blood group phenotype(C)
(Number of smaples: 127)

30.71%

Files in storage
(Number of smaples: 127)

100.00%

Thank you for attention!

E: marian.hajduch@upol.cz; www.imtm.cz; www.portal.imtm.cz



Co-funded by
the European Union



IPs EOSC-CZ registration number
CZ.02.01.01/00/22_004/0007682