

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



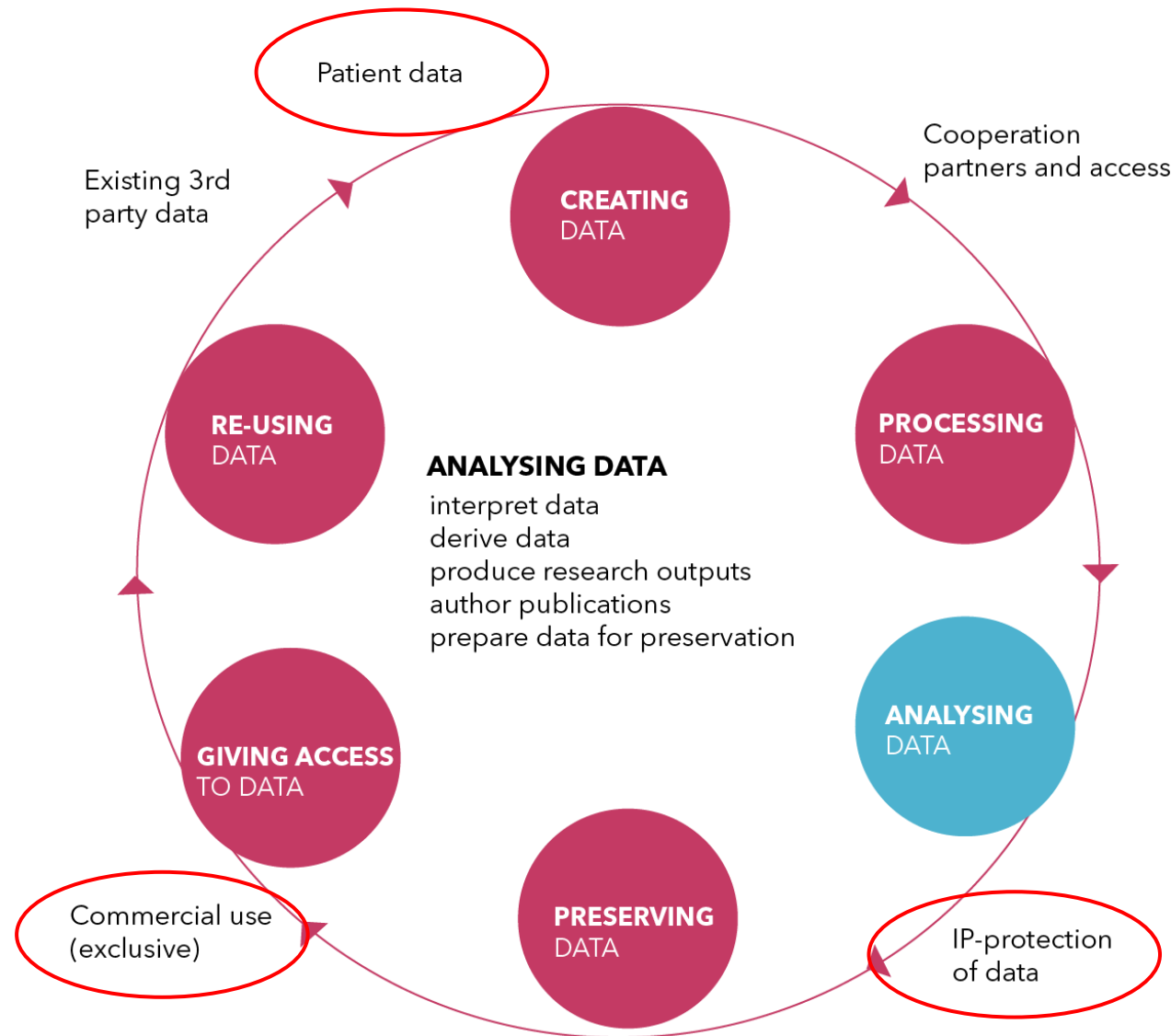
MUNI
ICS

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



eatris-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

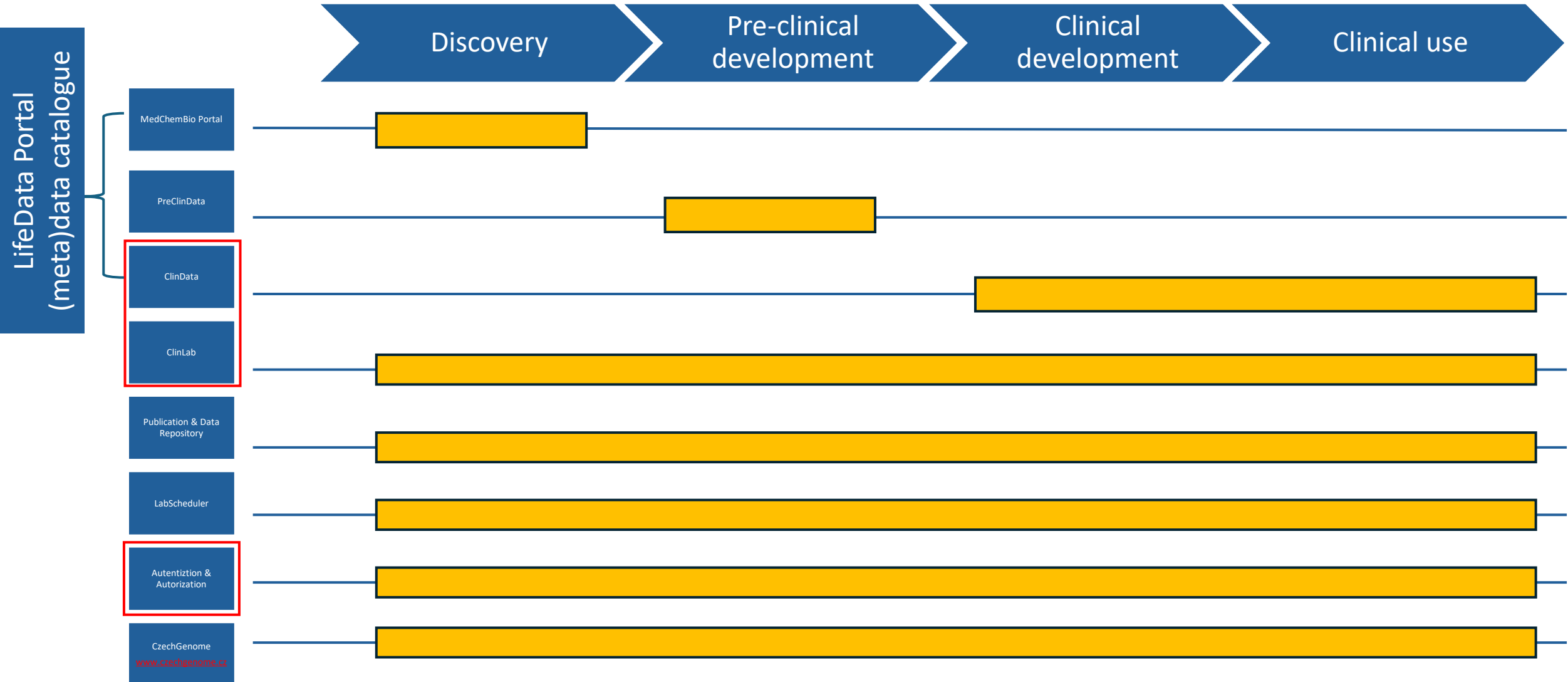


- Storage of laboratory data
 - Information systems ClinLab, PreClinData, MedChemBio Portal, LabScheduler



- 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 shows the 'Privilege list' in the IMTM Admin interface. The main table lists various roles and their descriptions. The role 'ROLE_REGISTRY_PATIENT_ADD_STUDY' is selected, and its details are shown in the 'Privilege details - edit' section at the bottom. The 'Entities with access to the selected Privilege' section on the right lists various entities that have access to this privilege, such as 'Laboratory 65Biocev' and 'Supervisor'.

Name	Description
ROLE_USER	
ROLE_REGISTRY	Registry
ROLE_REGISTRY_DISPLAY_PERSONAL_DATA	Registry display personal data of all
ROLE_REGISTRY_PATIENT	Registry patient
ROLE_REGISTRY_PATIENT_INFORM_CONSENT	Registry show informed consent
ROLE_REGISTRY_PATIENT_INFORM_CONSENT_DELETE	Registry delete informed consent
ROLE_REGISTRY_PATIENT_INFORM_CONSENT_ADD	Registry add informed consent
ROLE_REGISTRY_PATIENT_ADD_STUDY	Registry add patient to study
ROLE_REGISTRY_PATIENT_DEL_STUDY	Registry delete patient from study
ROLE_REGISTRY_PATIENT_FORM	Registry show patient's forms
ROLE_REGISTRY_PATIENT_FORM_EDIT	Registry edit patient's forms
ROLE_REGISTRY_PATIENT_FORM_SHOW_HISTORY	Registry show form history
ROLE_REGISTRY_PATIENT_FORM_CHANGE_STATUS	Registry changes form's status
ROLE_REGISTRY_PATIENT_FORM_CHANGE_STATUS_SIGNED	Registry changes status to signed
ROLE_REGISTRY_PATIENT_FORM_FORCE_CLOSE	Registry force close form
ROLE_REGISTRY_PATIENT_FORM_CHANGE_STATUS_CLOSE	Registry changes status to close
ROLE_REGISTRY_PATIENT_FORM_CHANGE_STATUS_REVIEW	Registry changes status to review
ROLE_REGISTRY_PATIENT_FORM_CHANGE_STATUS_OPEN	Registry changes status to signed
ROLE_REGISTRY_PATIENT_FORM_OPEN	Registry open patient's forms
ROLE_REGISTRY_PATIENT_SHOW_OWNER	Registry show patients owner
ROLE_REGISTRY_PATIENT_STUDY_DETAIL	Registry show patient in study
ROLE_REGISTRY_PATIENT_STUDY_DETAIL_EXPORT	Registry export patient's data
ROLE_REGISTRY_PATIENT_STUDY_DETAIL_EDIT	Registry edit patient in study

Privilege details - edit

Name: ROLE_REGISTRY_PATIENT_ADD_STUDY
Description: Registry add patient to study

Entities with access to the selected Privilege

Name	Description
Laboratory 65Biocev	Role for adding patients
Supervisor	Role with all access rights.
Import data 65BIOCEV manager	Role for importing group of data
CovIT PREVAL sampling site ALL patient user see all patients	
Manager 65BIOCEV laboratory	Role for managing status of patients + SMS
CovIT basic role	
ClinData SARS students	
CovIT PREVAL sampling site NOTall user do NOT see other patients	
Sampling site + Laboratory 5MOTOL	Sampling site and laboratory of 5MOTOL
Manager IMTM laboratory	Role for managing status of patients + SMS
ClinData project admin	Role with unrestricted access to selected project
ClinData project data manager	Role maintains patients in the study, no study editing
ClinData Disulfiram user	
Import data 53IMG manager	Role for importing group of data
Import data IMTM manager	Role for importing group of data
Laboratory 42MUNIBRN	Role for adding patients
Covid ADMIN	
Laboratory IMTM	Role for adding patients
Laboratory 53IMG	Role for adding patients
Sampling site + Laboratory 75UK1LF	Sampling site and laboratory of 75UK1LF
Admin of 42MUNIBRN laboratory	Role for managing status of patients + SMS + import
Sampling site + Laboratory 1SZU	Sampling site and laboratory of 1SZU
ClinData Koproretna gynekolog	Role for gynecologists in Koproretna project
Manager 53IMG laboratory	Role for managing status of patients + SMS

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](#)

Vzorky - odběrové místo [CRF history](#) [Refresh](#) [Pdf](#) [Print](#)

Jméno	<input type="text"/>	Příjmení	<input type="text"/>
Rodné číslo	<input type="text"/>		
Pohlaví	<input type="radio"/> M <input type="radio"/> F	Datum narození	<input type="text" value="DD/MM/RRR"/>
Telefonní číslo	<input type="text" value="telephone"/>	Email	<input type="text"/>
Město	<input type="text"/>	PSČ	<input type="text"/>
Zdravotní pojišťovna	<input type="text"/>		

Přijem vzorku

Datum odběru	<input type="text"/>
Typ materiálu na příjmu *	<input type="text"/>
Externí ID vzorku I.	<input type="text"/>
Externí ID vzorku II.	<input type="text"/>
Externí ID vzorku III.	<input type="text"/>

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				

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

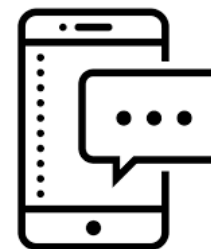
Extraction of
aliquots





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.

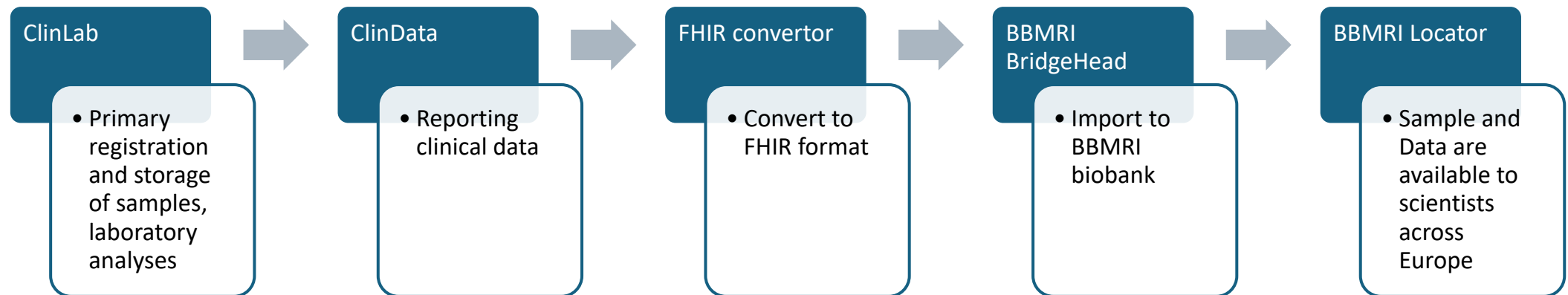
[Vytisknout výsledek](#)

[Stáhnout výsledek](#)





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

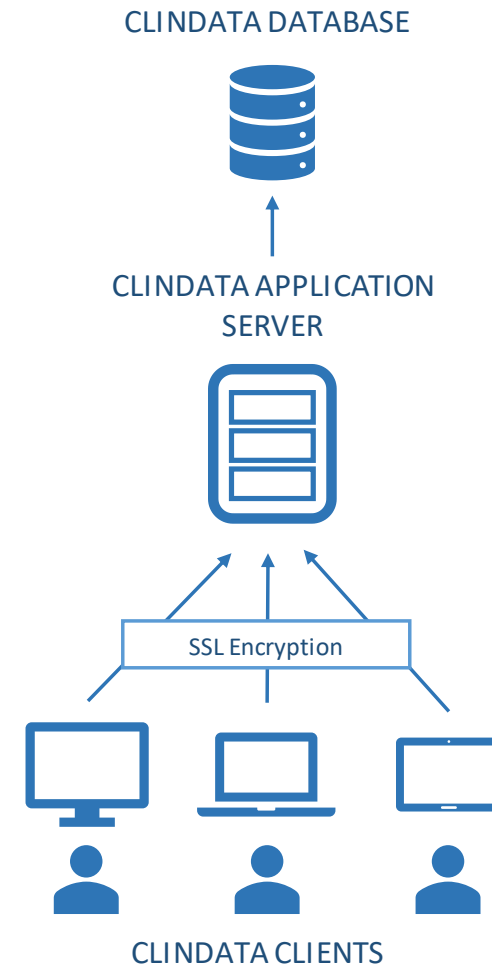


ClinData

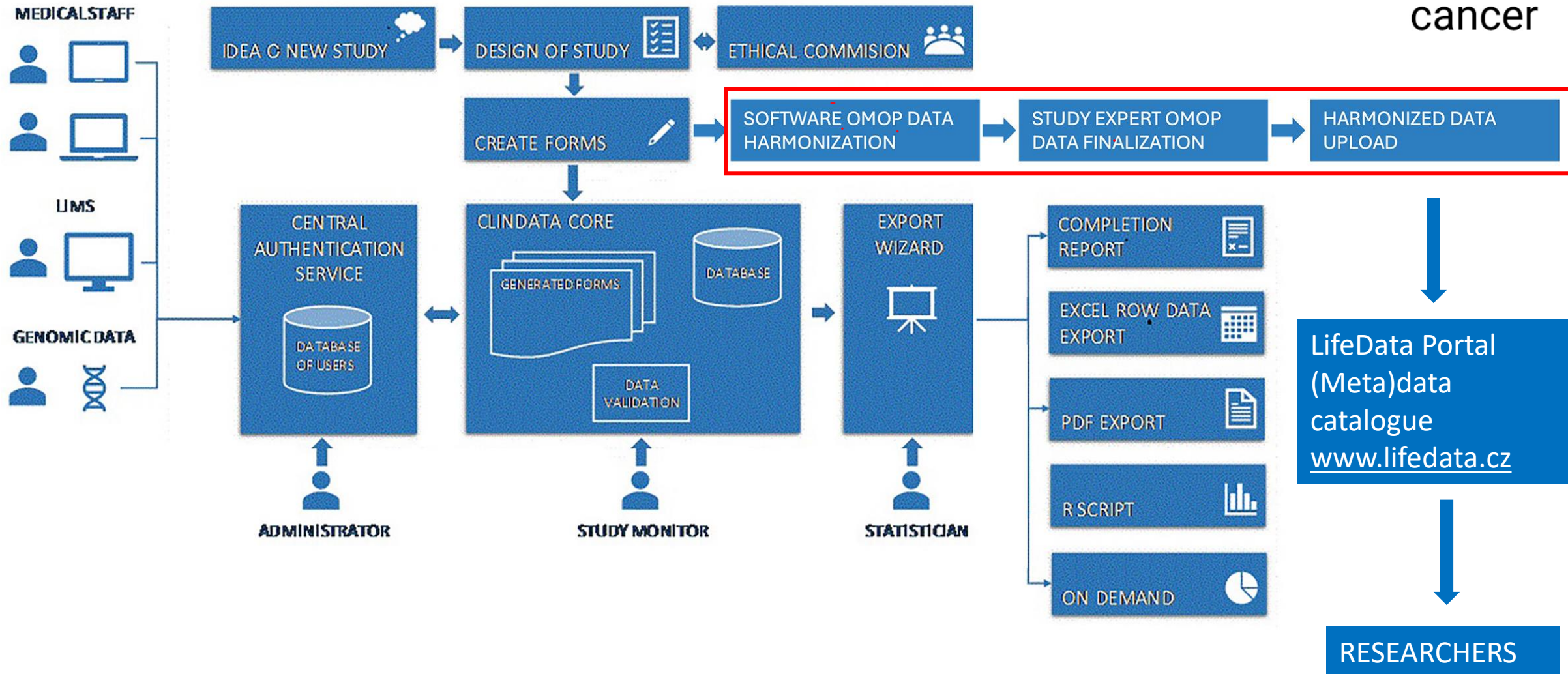
Comprehensive information system for data stewardship in clinical trials, registries and other clinical/scientific databases.

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

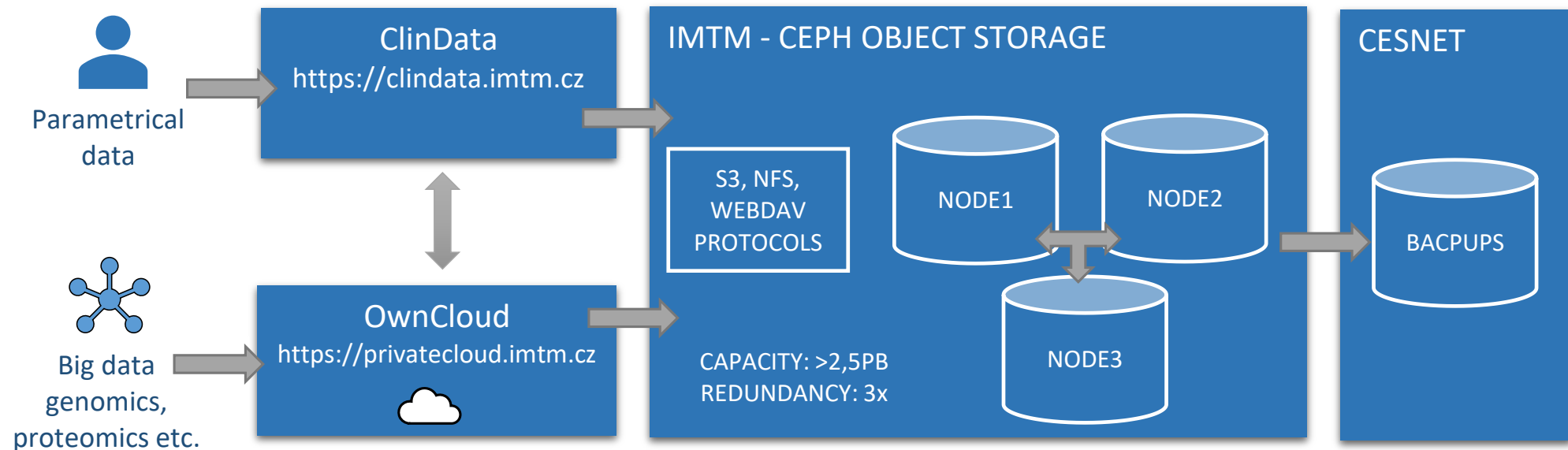


ClinData – basic diagram and data harmonization



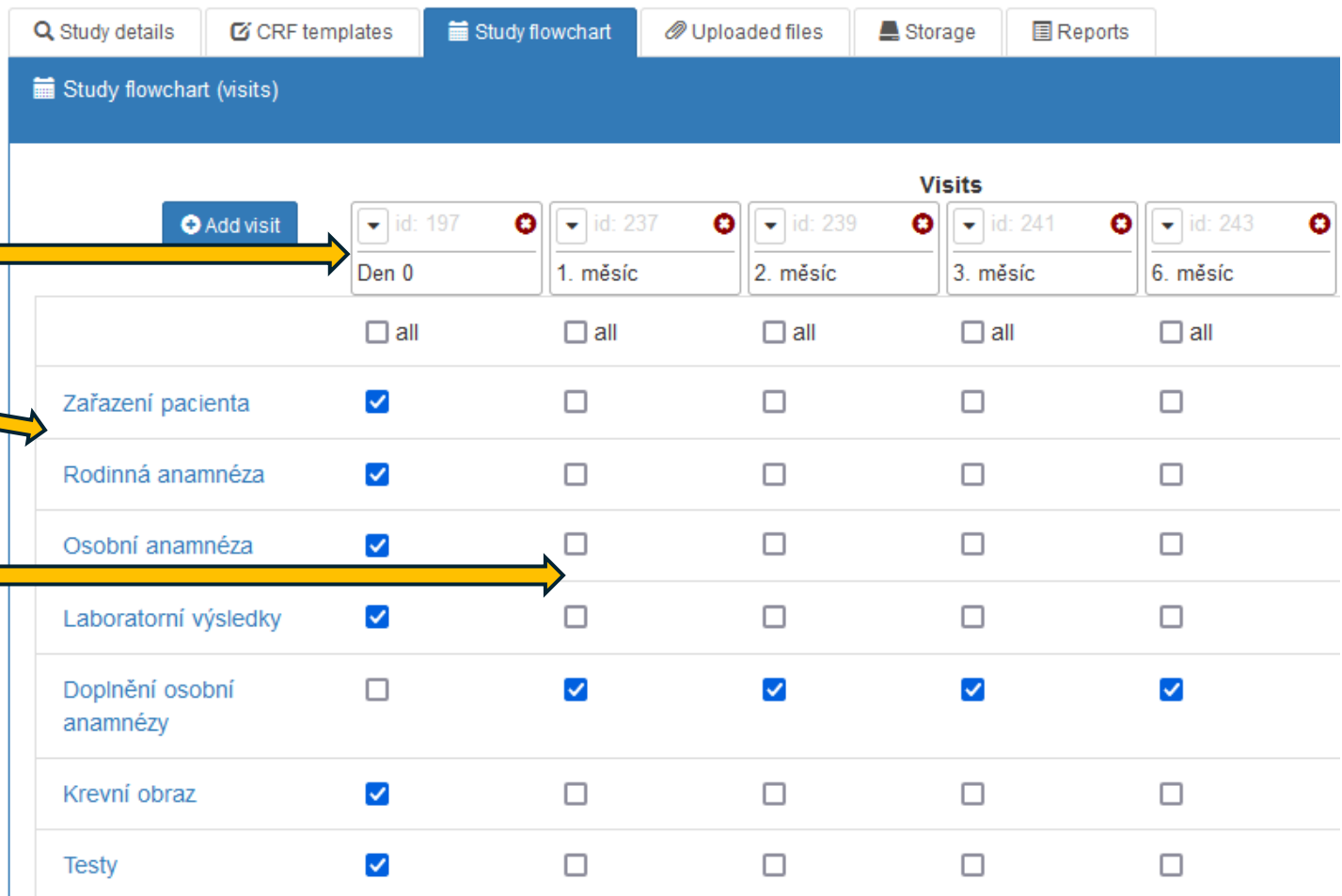
CEPH – Big omics data storage

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

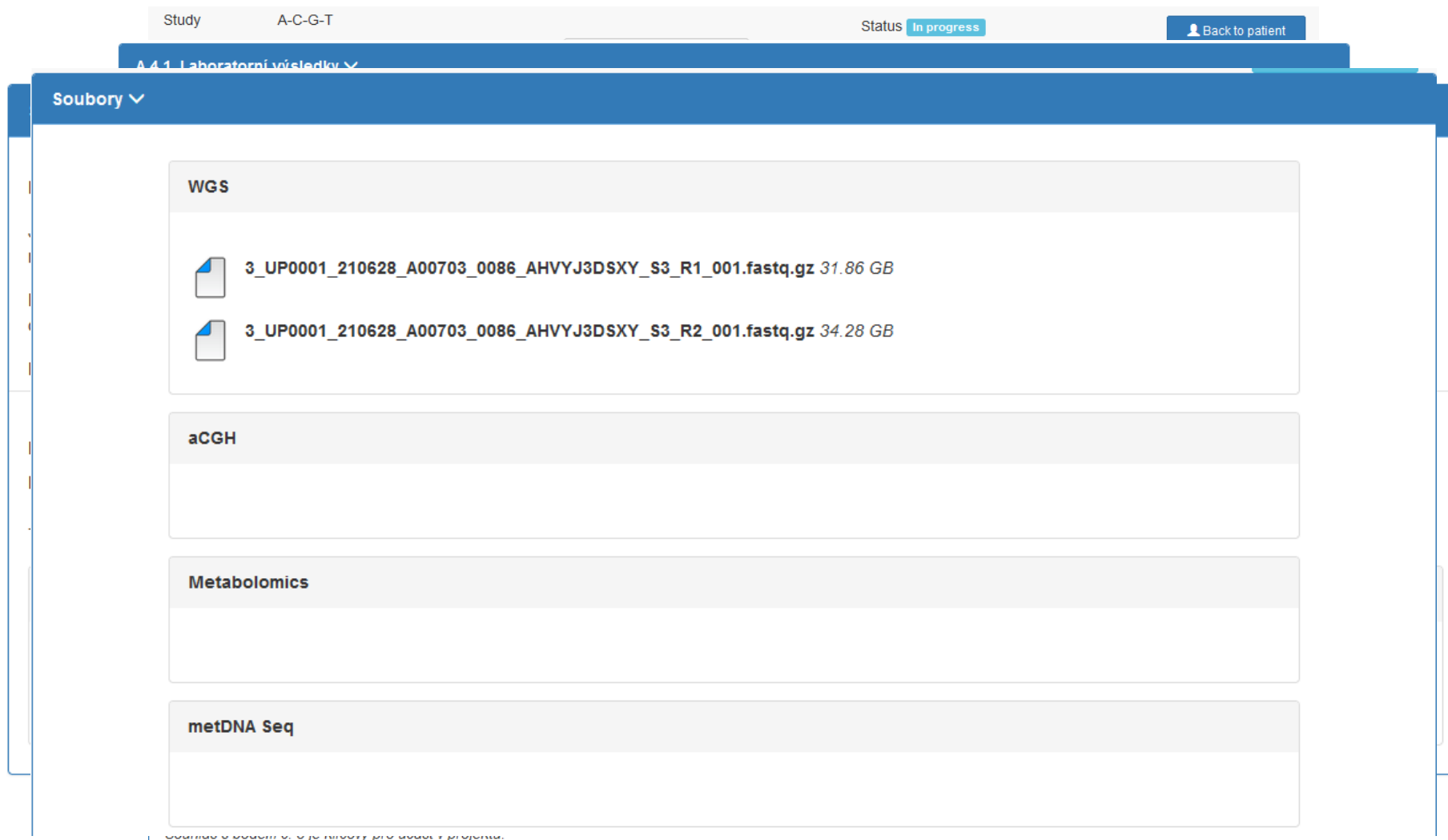
- Basic entity is **study**
- Study defines one or more **events**/visits (schedule)
- Study contains one or more electronic **forms** (eCRF)
- Every event can have **one or more forms**
- Events can be planned, repeated or ad-hoc



The screenshot shows the 'Study flowchart (visits)' interface. At the top, there are navigation tabs: 'Study details', 'CRF templates', 'Study flowchart' (active), 'Uploaded files', 'Storage', and 'Reports'. Below the tabs is a blue header bar with the text 'Study flowchart (visits)'. The main content area is titled 'Visits' and contains a table with columns for different visits and rows for various forms. A yellow arrow points from the 'Add visit' button to the 'Den 0' visit column. Another yellow arrow points from the 'Study contains one or more electronic forms (eCRF)' bullet point to the 'Zařazení pacienta' row. A third yellow arrow points from the 'Every event can have one or more forms' bullet point to the 'Osobní anamnéza' row.

	id: 197 Den 0	id: 237 1. měsíc	id: 239 2. měsíc	id: 241 3. měsíc	id: 243 6. měsíc
	<input type="checkbox"/> all	<input type="checkbox"/> all	<input type="checkbox"/> all	<input type="checkbox"/> all	<input type="checkbox"/> all
Zařazení pacienta	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rodinná anamnéza	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Osobní anamnéza	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Laboratorní výsledky	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Doplnění osobní anamnézy	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Krevní obraz	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Testy	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ClinData – eCRF examples



The screenshot displays a web interface for clinical data management. At the top, the study is identified as 'A-C-G-T' with a status of 'In progress'. A 'Back to patient' button is visible. The main content area is titled 'Soubory' (Files) and is organized into several data type categories: WGS, aCGH, Metabolomics, and metDNA Seq. The WGS section contains two file entries, each with a document icon, a filename, and a size in GB.

Category	File Name	Size
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		

OMOP harmonized data upload to Clindata

The screenshot displays the Clindata Form Designer interface. At the top, the browser address bar shows the URL: <https://clindata.imtm.cz/formDesigner/open?idForm=2828>. The interface includes a top navigation bar with tabs for 'Athena', 'ClinData', and 'Form designer | ClinData'. Below this, there are fields for 'Form name:' (Clinical_Patient), 'Description:', and 'Version:' (0). A toolbar contains buttons for 'Add new section', 'Browse history', 'Preview', and 'Save', along with a 'Developer' dropdown.

The main workspace is divided into three panels:

- Section Panel:** Lists items for the 'Clinical_Patient' section, including 'Patient ID', 'Age at Registration', 'Sex at Birth', 'Study Arm', 'Progression-Free Survival Status mapped)', 'Progression-Free Survival in Months (not mapped)', 'Sidedness of Primary Tumor', 'Time to Liver Metastases Diagnosis', 'Num of Liver Metastases at Baseline', 'Resectability of Liver Metastases at Baseline', and 'Resectability of Liver Metastases at Last Evaluation'.
- Configuration Panel:** A dialog for configuring a 'Text' item. It includes options for 'Label', 'Input', 'List', 'Calculate', and 'Mask'. The 'Display label' checkbox is checked. The 'Text' field contains 'Sex at Birth'. The 'Offset (0-11)' is set to 1, and the 'Width (0-12)' is set to 4. There are two 'Concept ID - Athena link' fields with values '37116947' and 'Biological sex', and one 'Concept Name - Athena link' field.
- Layout Panel:** Shows a vertical stack of 'Layout' controls for each item, with 'Repeatable' and 'Mandatory' checkboxes.

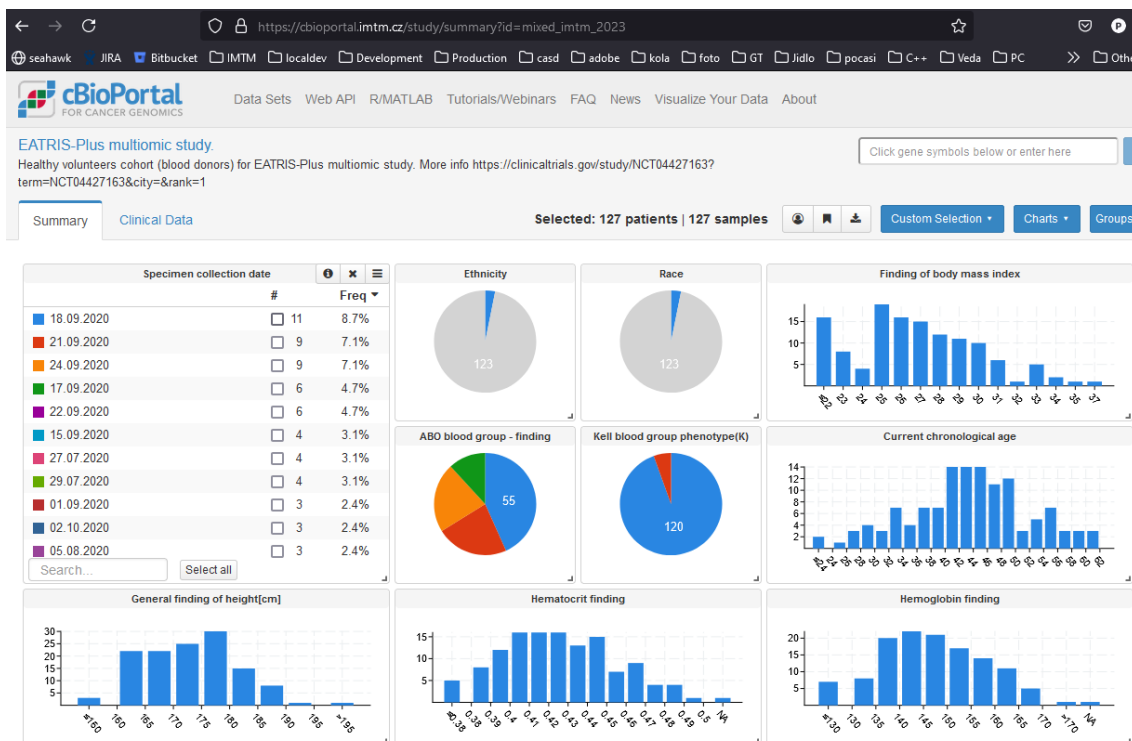
The Windows taskbar at the bottom shows the system tray with the date and time: 14:10, 01.07.2024.

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://cbioportal.imtm.cz>
- Portal to FAIRify data from ClinData – only for cancer genomics
- Development of new public portal for FAIRification of other types studies



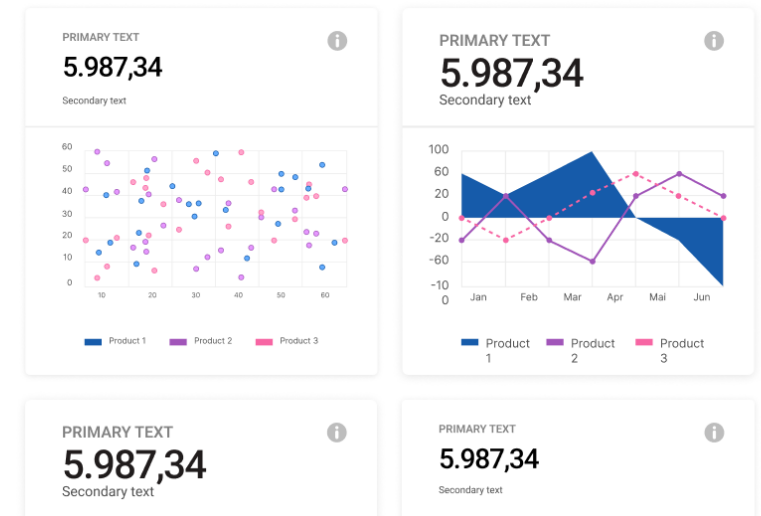
BioPortal

Studies Study Details Data **Graph And Summaries**



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



ClinData – LifeData Portal www.lifedata.cz

LifeData Portal

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

Study Name:

Principal Investigator:

DOI:

Primary Diagnosis:

Disease Population Study Healthy

Other:

[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 infectious 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

Studies Study Details Data Graph And Summaries

Search by name, doi, gender...

Export Result

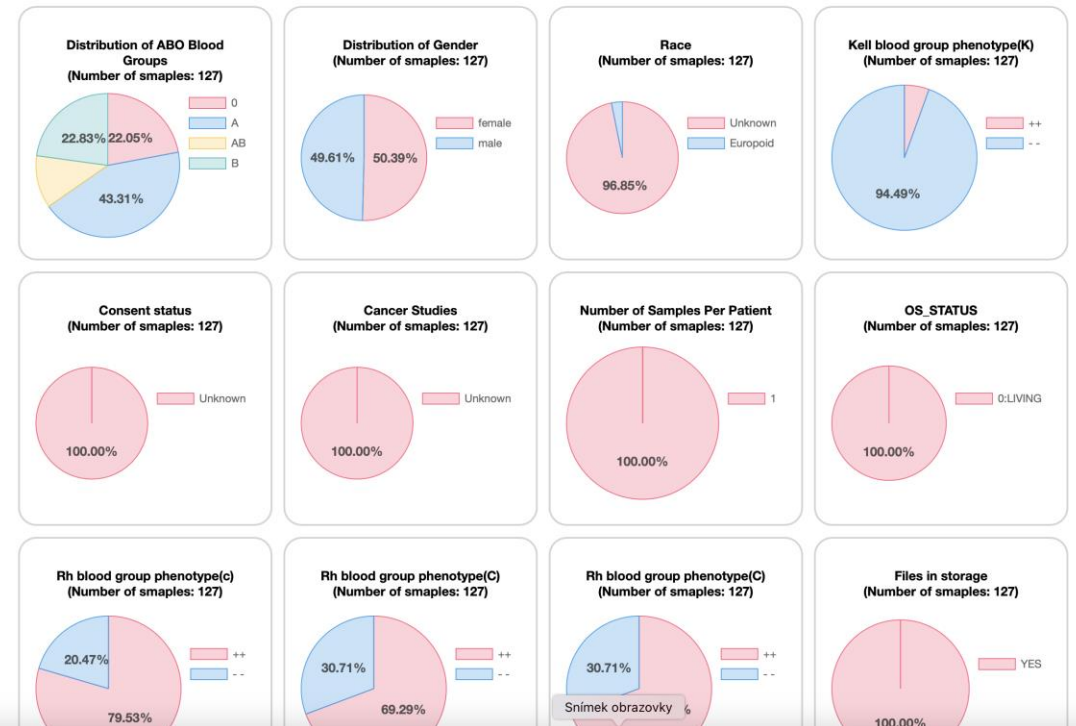
← Prev

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

Studies Study Details Data Graph And Summaries



Thank you for attention!

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



eatris-cz



Co-funded by
the European Union



MUNI
ICS

cesnet
.....

VSB TECHNICAL
UNIVERSITY
OF OSTRAVA

IT4INNOVATIONS
NATIONAL SUPERCOMPUTING
CENTER

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