Workshop #1

Harmonization of Clinical and Biological Data

Session 3

Eva Steliarova-Foucher, PhD

Serban Negoita, MD, DrPH

Workshop General Goals and Structure

Goal: Discuss strategies and plan an international partnership to advance the harmonization of biological and clinical data in support of childhood cancer research

Session 1: Core data elements: patient, tumor, prognostic factors

Session 2: Treatment and outcome + standards mapping

Session 3: Genetic/molecular data + challenges to harmonization

Session 4: Summary of discussions + pilot project

Intro to session 3

- Pediatric tumors differ from tumors in adults in cellular origin, genetic complexity, driver mutations, and potential mutational processes, therapy decision and response to treatment
- Genomic sequencing, molecular profiling, and liquid biopsy characterization at the time of diagnosis
- Genetic/Molecular characterization helps survivors and medical providers select the best and most appropriate treatment

Intro to session 3 (continued)

- Childhood Cancer Data Initiative, an NCI program that ensures data are shared by survivors, clinicians and researchers in the USA; similar projects in Europe
- Increased concerns about protection of personal health information
- NCI platforms with controlled access- dbGAP, Seven Bridges

Session-Specific Questions and issues for consideration

- Identify genomic variables and data standards relevant for childhood cancer research
- Discuss solutions to challenges in harmonization of genetic and molecular data
- Plan activities to integrate data and standards for phenotypic and genetic/molecular data into a childhood cancer data resource/platform

Session 3 topics and Discussants

- **Dr. Arnaud Petit**

Challenges in harmonization and sharing of pediatric AML data

- **Dr. Frank Westermann**

Combining genetic and molecular data with phenotypic data for childhood tumors

Discussants

Arnaud PETIT
France
Armand Trousseau Hospital
French Society for Childhood and Adolescent Cancer and Leukemia
Pediatric Oncologist

Frank WESTERMANN
Germany
German Cancer Research Cancer (DFKZ)
Head of the Department
of Neuroblastoma Genomics

Paris Conference for an International Childhood Cancer Data Partnership
Newcap Event Center (Paris, France), November 7-8, 2023

DOREMy

Base de **DO**nnées cliniques et biologiques harmonisées pour une **RE**cherche intégrée à la prise en charge des leucémies aiguës **MY**éloïdes pédiatriques

Harmonized Clinical and Biological Data Bases for an Integrated research dedicated to Pediatric Acute Myeloid leukemia

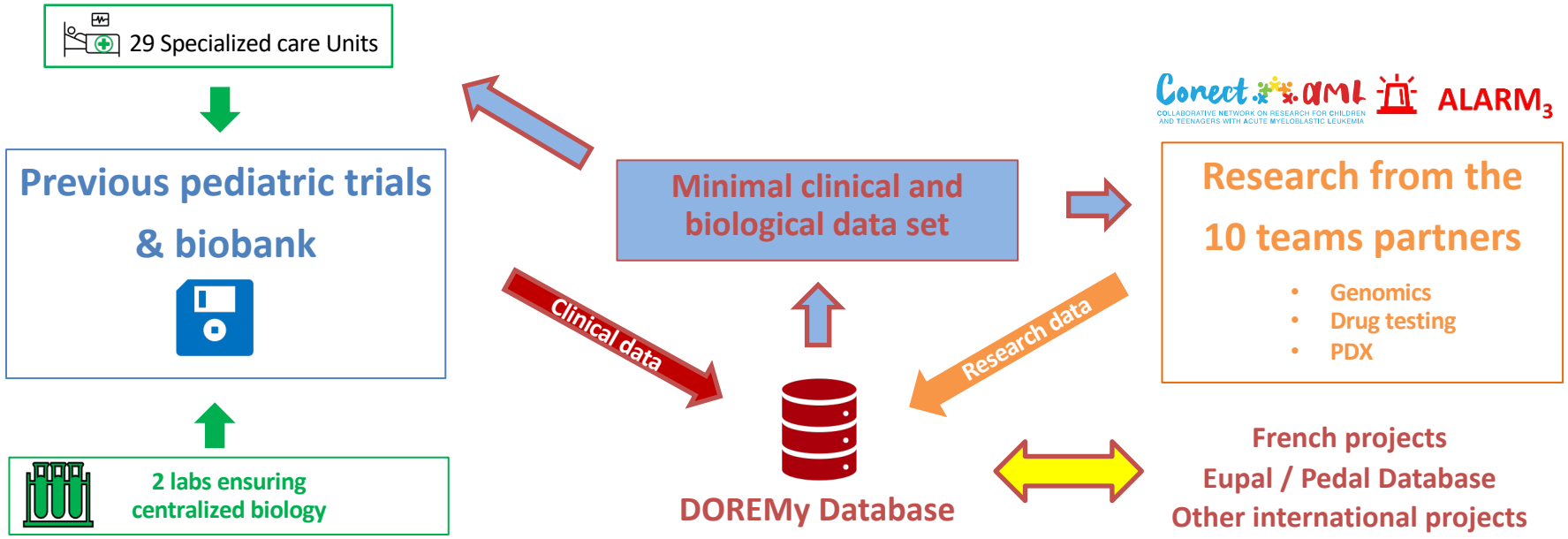
Arnaud Petit, MD, PhD, Paris, France

Conect aML
COLLABORATIVE NETWORK ON RESEARCH FOR CHILDREN AND TEENAGERS WITH ACUTE MYELOBLASTIC LEUKEMIA

DOREMy is supported by a grand from InCa (parpedia19-008)
CONECT-AML is supported by a grant from InCA, Fondation ARC, Ligue nationale contre le cancer (InCa-ARC-LIGUE_11905), and the following partners

Objectives of DOREMy

- Better characterize rare high-risks subgroups of AML
- Identify new prognosis or therapeutic markers
- To identify new therapeutic targets
- The ultimate goals are to reduce relapse and improve survival



Funding : 2019, Start of the project : September 2020

REGULATORY PROCESS

Main blocking difficulty

1. *Authorization by the CNIL*

- Prerequisite for any further steps
- 1st submission : summer 2021
- Withdrawal for inadequate security of health data

2. *Health data host*

- Completed mid-2023

CASD

3. *Next step*

- 2nd submission to CNIL, Q4 2023

CLINICAL DATABASE

1. *Data collection from previous trials*

- Selection of clinical and biological data

2. *Data harmonization (clinical data)*

- According to PDCD INTERACT* AML Dictionnary

LAME89/91 | ELAM02 | Myechild01

Post-ELAM02 registry

3. *Integration into a closed database*

- REDCAP, completed Q3 2023

1989-2022
n= 1326

4. *Next steps*

- Update survival data (link with LEA project)
- Prospective data : ALARM[3+] and Co-DoReMy registry

GENOMICS DATABASE

1. *Data from CONECT-AML*

- Generated in 2021 and 2022
- n≈90 paired sample diag/CR
- WGS/RNAseq
- SNVs, Indels, SVs, CNVs, fusion transripts

2. *Integration*

cBioPortal FOR CANCER GENOMICS

- Started mid-2023
- Still ongoing

ARTbio IBPS

*International Pediatric Acute Myeloid Leukemia Consortium

[+] Granted by Imagine For Margo (FKC), Enfants Cancer Santé, AFER

DATA ACCESS WITHIN DOREMy

Scientific committee (SC)

P-DOREMy
actor who is a partner DOREMy

NP-DOREMy
actor who is not a partner DOREMy

TYPE 1 REQUEST
Made by a P-DOREMy limited to the only data generated by the center or team

TYPE 2 REQUEST
Made by a P-DOREMy associating other P-DOREMy limited to the only data generated by all of these actors

TYPE 3 REQUEST
Nor type 1 or 2

Information to scientific committee

Need for approval by SC

BENEFITS

Clear need for international cooperation to make progress on childhood myeloid leukemia

STRENGTHS

IBFM AML SG : International cooperative group to discuss and promote collaborative projects

Common standardized language for clinical data

KEY CHALLENGES

1. Raise the good question to built the good database
2. Link the data to ensure it relates the same patient
3. Selection and harmonization for Omics data (Raw vs processed data ?)
4. Obtain regulatory authorizations for data reuse and data sharing
5. Terms and conditions of data sharing

Discussants

Arnaud PETIT
France
Armand Trousseau Hospital
French Society for Childhood and Adolescent Cancer and Leukemia
Pediatric Oncologist

Frank WESTERMANN
Germany
German Cancer Research Cancer (DFKZ)
Head of the Department
of Neuroblastoma Genomics

Paris Conference for an International Childhood Cancer Data Partnership
Newcap Event Center (Paris, France), November 7-8, 2023

ITCC Data Integration Portal

Frank Westermann

David TW Jones, Patrick Kemmeren, Stefan M. Pfister

The ITCC-data integration platform

Germany (+12)

MAPPYACTS
France (+x)

SM-PAEDS
UK

>5000 retrospective datasets
>1000 new patients per year

iTHER
The Netherlands

ZERO
Australia

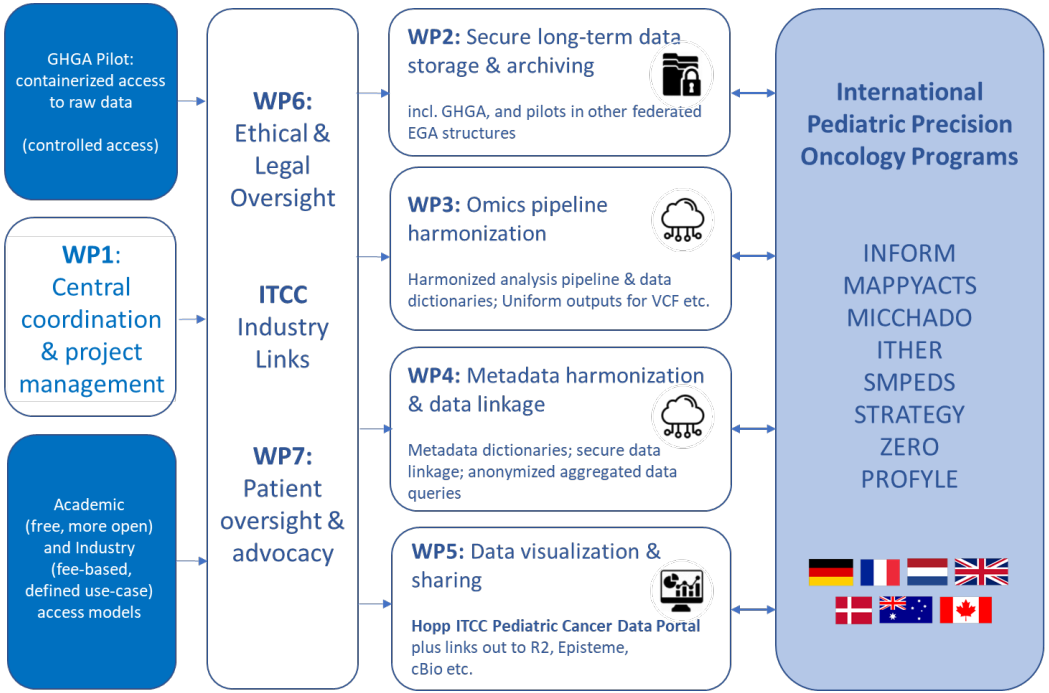
PROFYLE
Canada

iCOPE/STRATEGY
Denmark

The ITCC-data integration platform: Aims

- Real-time archiving of raw data in national EGA instances
- Harmonization of data dictionaries (molecular & clinical) as well as ICFs (e.g., also allowing data sharing with industry)
- Harmonization of omics pipeline and central collection of processed omics data for overarching analyses
- Data vizualization and sharing through www.pedcanportal.eu (internal and external) utilizing existing tools (R2, cbio, PedCan portal at St. Jude, etc.)
- Coordinating overaching analyses, similar to PCAWG (ICGC)
- Involving patient representatives as well as ethical and legal experts as key project partners

The ITCC-data integration platform: WP structure



David Jones (KiTZ)
Patrick Kemmeren (PMC)
Stefan Pfister (KiTZ)
Frank Westermann (KiTZ)

+ fantastic global team!

10.5M € funding obtained from the Dietmar Hopp Foundation 1/2023

Key decisions in the process of integrating and exchanging genetic/molecular and clinical/phenotypic data

- **Establish a federated and harmonized bioinformatics pipeline**

=> assessment of existing bioinformatic pipelines

=> develop a harmonized consensus pipeline

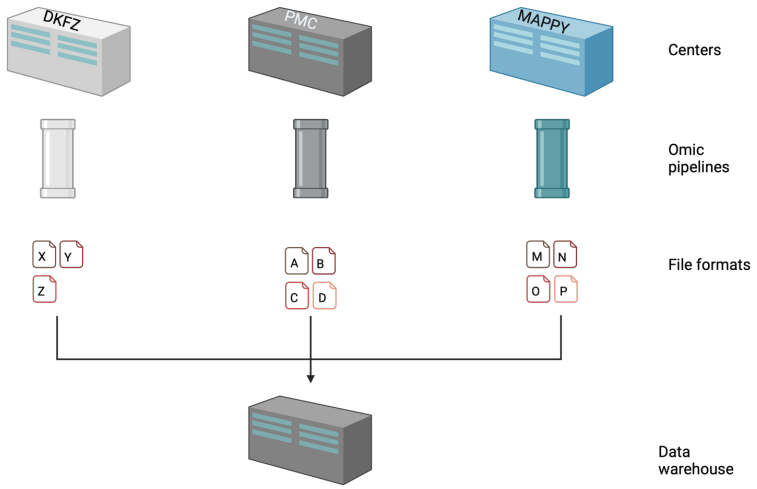
=> Establishing a validation procedure/quality management for the pipeline

- **Complexities from data formats from multi omics layers**

=> flag restricted/permissive data outputs,

=> define standard data formats (e.g., annotated variants in MAF, chromosome copy number as CBS segments)

- **Establishing a data warehouse for secure access**



Metadata harmonization – action items

- **Buid a Common Data Model (CDM)**

=> Benchmarking of initiatives for clinical data harmonization (semantic interoperability)

=> Development of a CDM encompassing the relevant Common Data elements (CDE)

=> Prospective use of CDE

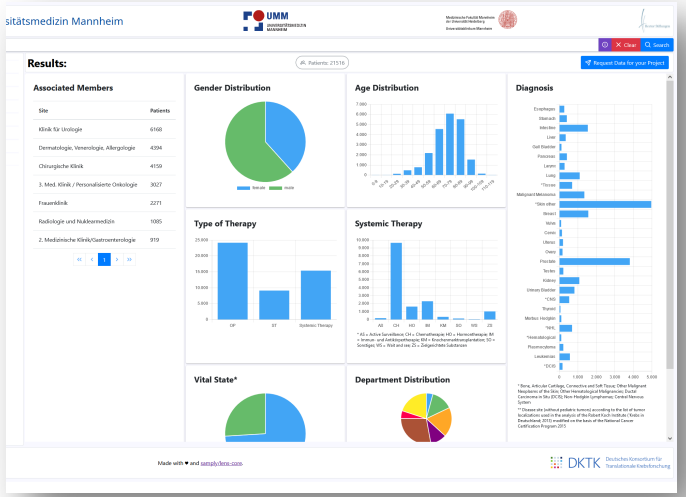
- **Develop Secure Data Exchange**

=> Design and implement pseudonymization processes

=> Design application programming interface (API) for data submission

=> Provide local node package

=> Operation and sustainabliity



Looking forward to the discussion!

KiTZ Hopp-Kindertumorzentrum Heidelberg

Princess máxima center pediatric oncology

dkfz. DEUTSCHES KREBSFORSCHUNGSZENTRUM IN DER HELMHOLTZ-GEMEINSCHAFT