Breakout session - WG9 Cancer

1+MG / B1MG Stakeholder Forum

Chair: Giovanni Tonon

Co-chair: Astrid Vicente

04

Slido: **Cancer**

‘1+ Million Genomes’
1+MG initiative + B1MG project

Giovanni Tonon
ACC/UniSR
WG9 Cancer

https://b1mg-project.eu

Outcome

1. How can stakeholders from each use case help scale up and make the work of 1+MG/B1MG in data governance, data standards and technical infrastructural aspects sustainable?

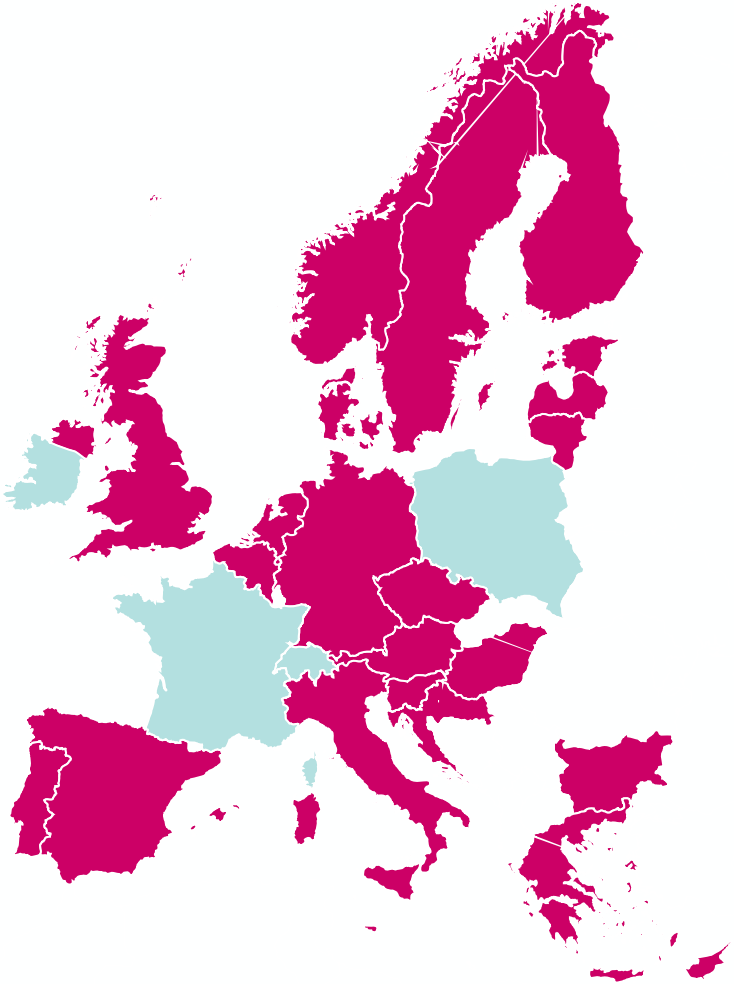
2. How can industry facilitate this process/ gaps industry can fill?

Slido: **Cancer**

Learning network of 1+MG National Mirror Groups

- Ministries of Health and Science
- National Stakeholders
- Guiding national implementation along 1+MG recommendations

1+MG Working Groups and National Mirror Groups



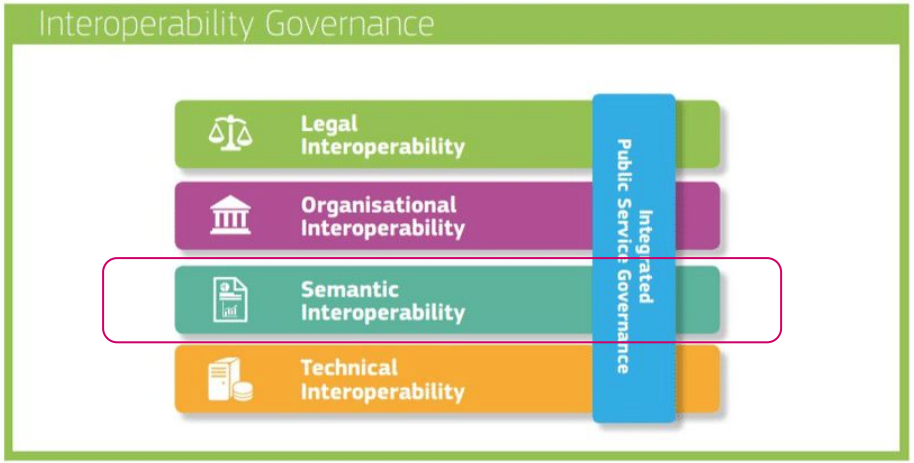
***) ELIXIR nodes involved in NMG's**

Slido: **Cancer**

Phenotypic and clinical metadata framework

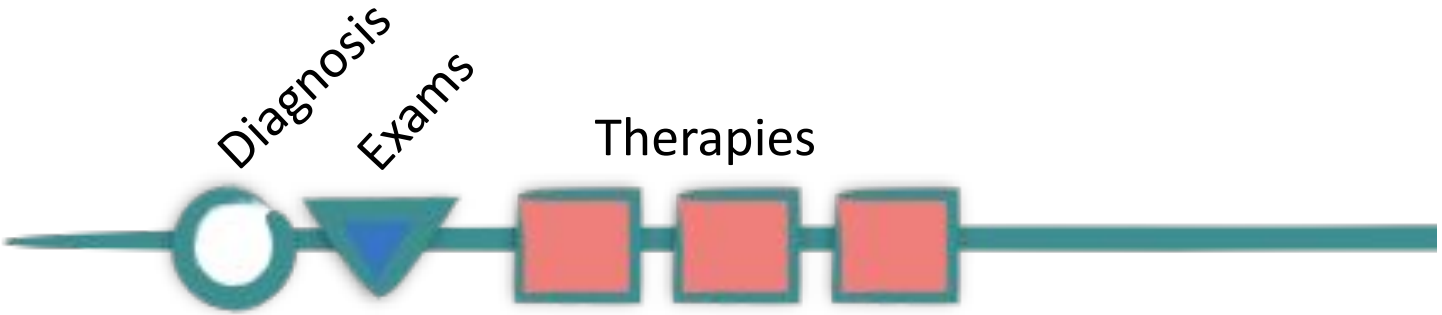
Principles, models, standards and recommendations for sharing and linking phenotypic and genetic metadata between the member states

- Guidance on which standards, terminologies and tools to use
- Best practices, describing which ontologies are currently implemented in each member state
- Enabling FAIR data



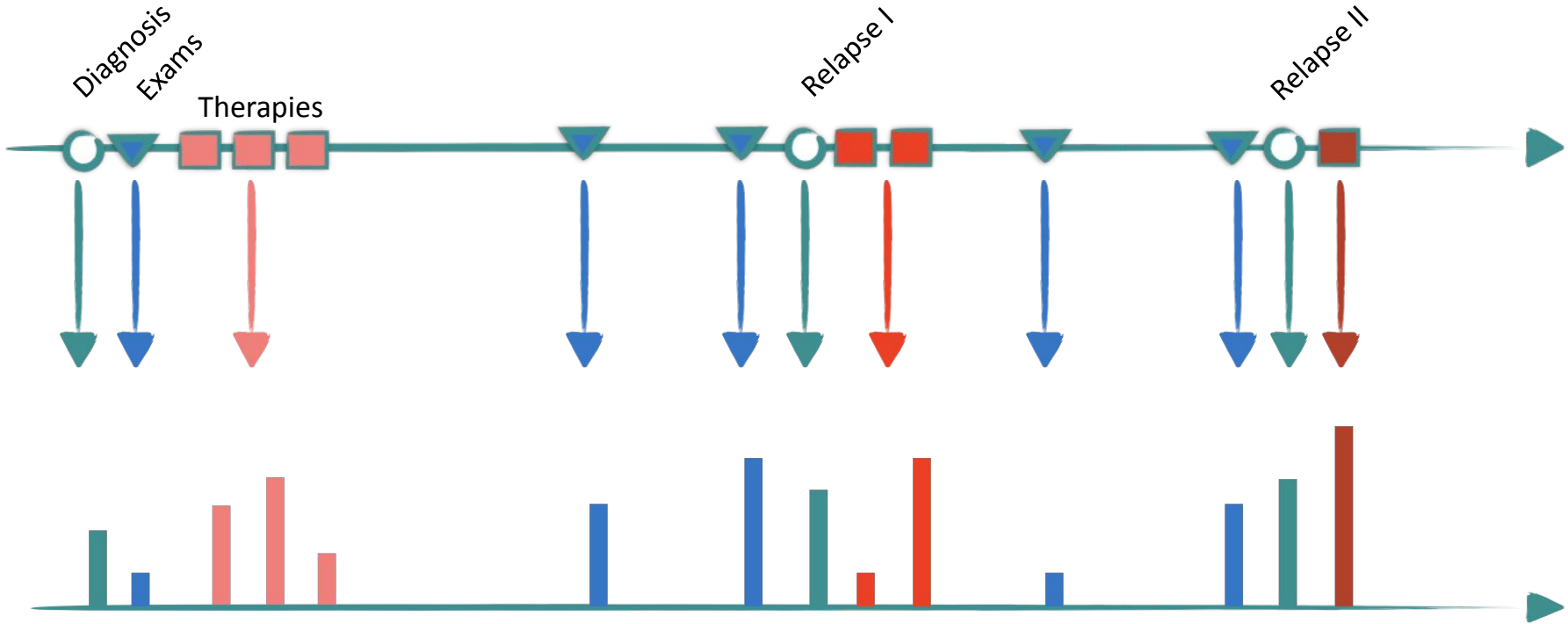
Slido: **Cancer**

The oncology problem



Slido: **Cancer**

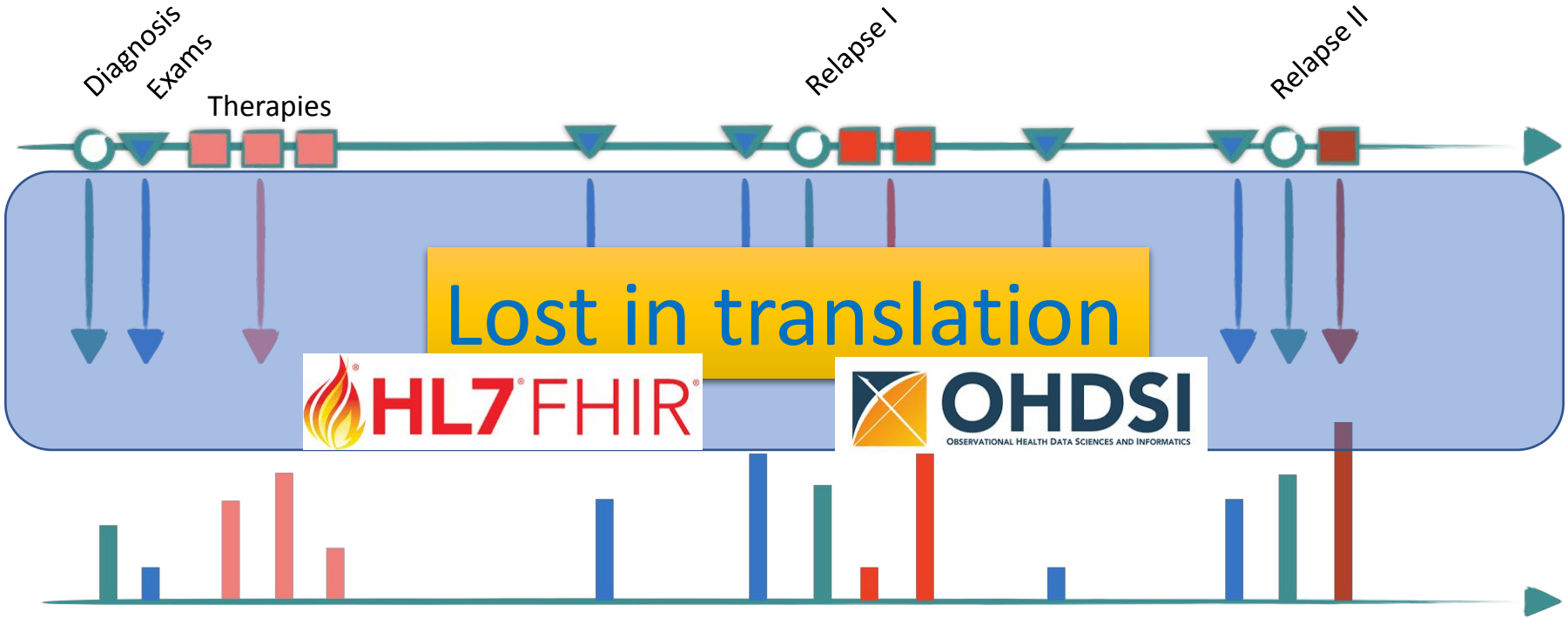
The oncology problem



Credits: X. Fernandez, M. Curie

Slido: **Cancer**

The oncology problem



Minimal Data Set for Cancer

Slido: **Cancer**

Motivation

Taking into account the existing published data models, we would like to propose a data model fitted for cancer genomics in Europe with the following features:

- *Minimal* for easiness of use but comprehensive for both clinicians and researchers
- Capturing the *longitudinal* and complex aspects of cancer
  - recording data coming from different "points" of the disease evolution *timeline*,
  - accepting *various samples* (not only primary tumour) from the same patient
- Considering data on *treatments* and their *outcomes* to benefit also genomics research's needs
- Maximizing the use of *standards* for needs for *interoperability*

Slido: **Cancer**

Data Models for Oncology +

mCODE

Osterman, T. J., et al., (2020). Improving Cancer Data Interoperability: The Promise of the Minimal Common Oncology Data Elements (mCODE) Initiative. JCO Clinical Cancer Informatics, https://doi.org/10.1200/cci.20.00059

OSIRIS

Guérin, J., et al.,. (2021). OSIRIS: A Minimum Data Set for Data Sharing and Interoperability in Oncology. JCO Clinical Cancer Informatics, https://doi.org/10.1200/cci.20.00094

ICGC-ARGO

https://cancerres.aacrjournals.org/content/79/13_Supplement/1698#

ODHSI-OMOP

Belenkaya, R., et al., (2021). Extending the OMOP Common Data Model and Standardized Vocabularies to Support Observational Cancer Research. JCO Clinical Cancer Informatics, https://doi.org/10.1200/cci.20.00079

…

Slido: **Cancer**

Data Models for Cancer Arise from Different Assumptions and Focus on Different Aspects

Slido: **Cancer**

	mCODE data model	ICGC-ARGO data model	OSIRIS data model
Clinical data model	YES	YES	YES
Genomics data model	Small-scale		Small-scale
Longitudinal data		NO	YES
Pharmacogenomics		NO	YES
Focused on		mainly primary tumors	solid tumors
Raw data re-analysed	NO	YES	NO
Scalability	YES		YES
Modularity		YES	YES
Use of standard terminologies	YES	YES (vocabularies or lists)	YES (vocabularies or lists)
Interoperability	YES, native on FHIR		YES, not native on FHIR

The Process towards a Minimal Set for Cancer

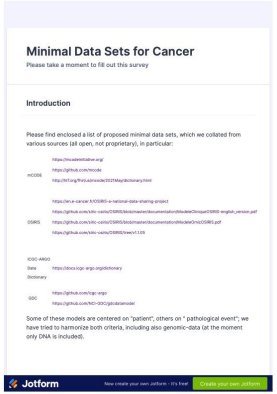
Dialogue with clinicians and researchers

Definition of a set of tables, sub-tables and items

Revision by the WG9 NMG community

Refinements

Minimal Set



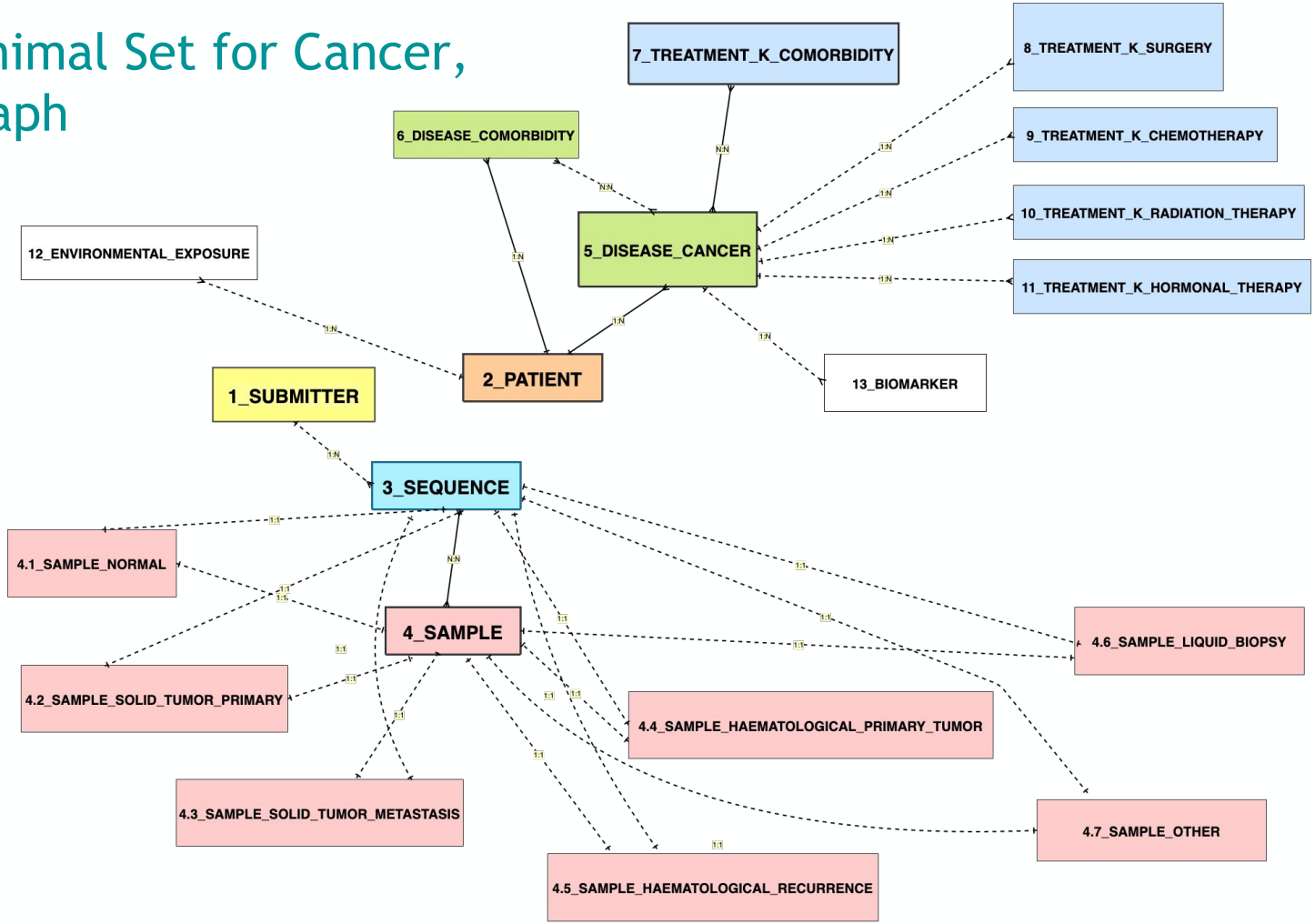
WG9 survey

March 1, 2022

1+MG WG3/B1MG WP3

Slido: **Cancer**

Minimal Set for Cancer, Graph



Slido: Cancer

Minimal Set for Cancer

- A set of 13 excel tables

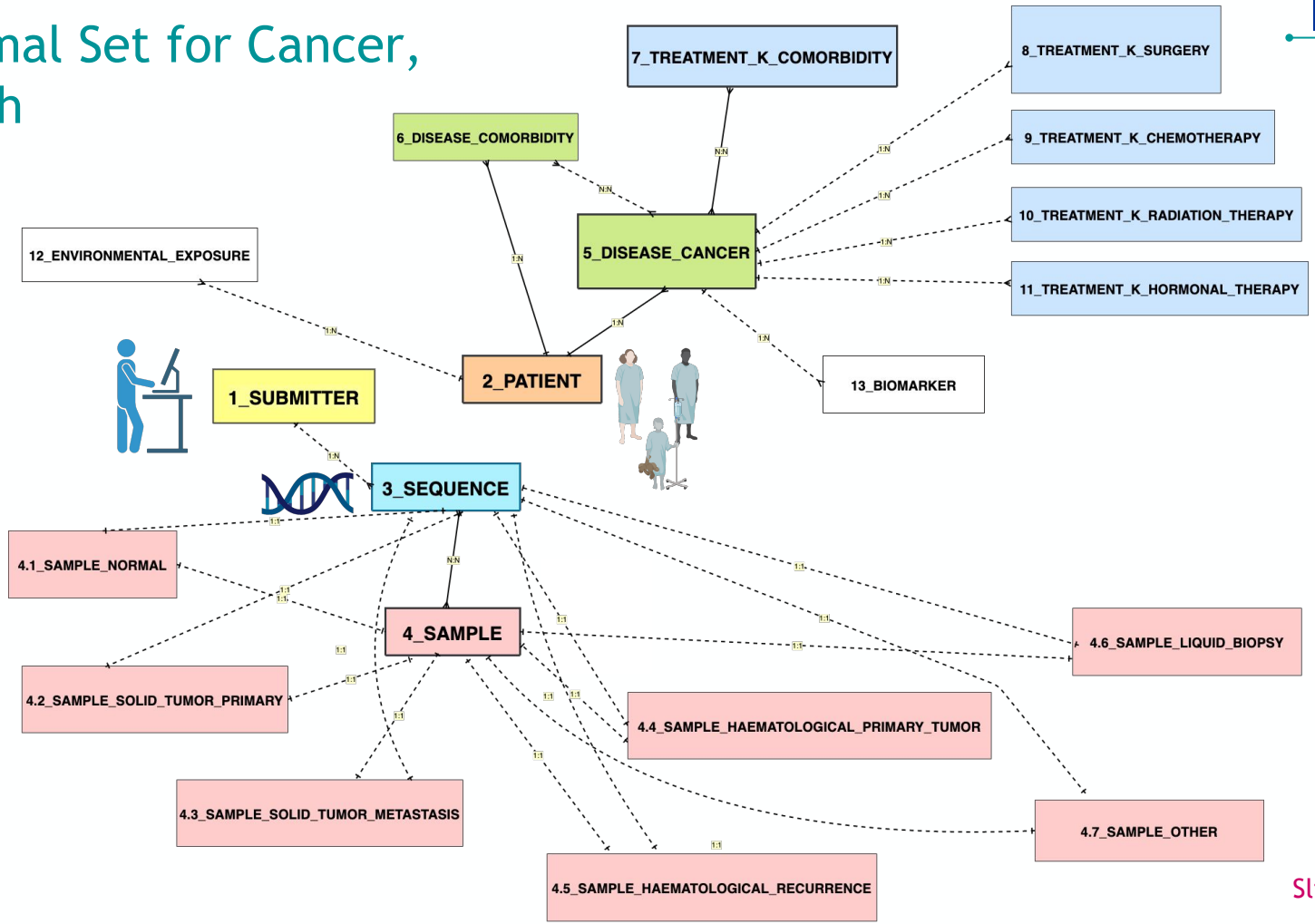
collection status

Table and sub-tables	Item	Item definition	Examples	Source	Mandatory recommended optional
	Name of the item	Description	Possibible standard values	Reference to existing data models	

1+MG, NMG WG9 March 1, 2022

Slido: **Cancer**

Minimal Set for Cancer, Graph



Slido: **Cancer**

Value Sets to Fill in the Items in the Minimal Set

The use of free text would be minimised when dictionaries or pre-defined lists of possible choices are identified
This survey aims to identify the most commonly used dictionaries or pre-defined lists for each item of the Minimal Set for Cancer

WG9 survey

Dictionaries for the Minimal Data Set for Cancer

This survey aims to identify the most commonly used dictionaries or pre-defined lists for each item of the Minimal Set for Cancer.
The use of free text would be minimised when dictionaries or pre-defined lists of possible choices are identified.

Lists have been mostly derived from:
ICGC-ARGO Data Dictionary
https://docs.icgc-argo.org/dictionary

OSIRIS data model
https://github.com/siric-osiris/OSIRIS/tree/v1.1.05

mCODE data model for cancer https://mcodeinitiative.org/

Please select the proposed dictionary, or pre-defined list (first choice) in case of approval; otherwise, when provided, please select the items of the list;

the "Other" option could be used for suggestions and comments.

- 138 items over 13 tables
- towards the definition of filling value sets

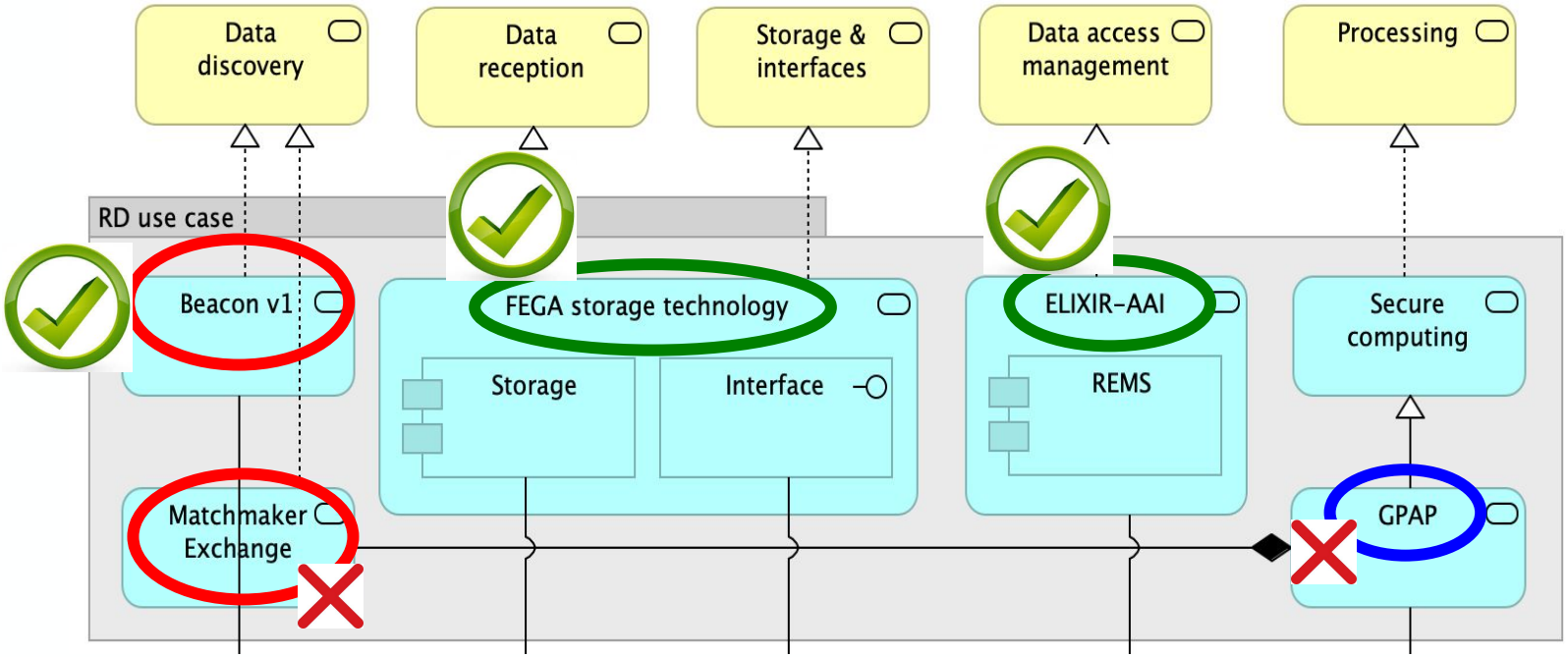
Slido: **Cancer**

Rare Disease PoC

Generic

RD-specific

Slido: Cancer



Data model?

Re-analysis?

Collaboration with Dylan Spalding, WG4

1. **lung cancer**

(formerly B1MG PoC's use case)
Recreate a fully synthetic dataset (collab. EUCANCAN team in Barcelona, David Torrents), using simulated and real data from publicly available mutational profiles

2. **melanoma**

We'll use publicly available metastatic cancer and related lymphoblastoid cell lines data (COLO829, EGA project PRJEB27698, Edwin Cuppen)

Slido: **Cancer**

Acknowledgements to 1+MG/B1MG team


M. Riba


G. Bucci


M. Morelli

Slido

Go to www.slido.com from your phone or computer and enter the code below OR scan the QR Code to start answering the breakout questions

Cancer

Feedback

Breakout session - WG9 Cancer

Chair: Giovanni Tonon

Co-chair: Astrid Vicente


Popular ideas
7
ELSI & Data Governance: Which challenges to the scale up (within this Use Case) could be addressed by industry?
Anonymous
5
Harmonisation
Anonymous
4
after data harmonization we will need data quality control - an expensive chapter
Anonymous
3
Data ownership
Anonymous
1
Confidentialy
Anonymous
1
ELSI terminology/contracts only open to 1 interpretation

ELSI & Data Governance: Which challenges to the scale up (within this Use Case) could be addressed by other stakeholders (please specify stakeholder)?

Anonymous — 8

Patient organisations

Anonymous — 7

Policy-makers

Anonymous — 3

Citizens

Anonymous — 3

Creating trust ↔ all relevant stakeholders

Anonymous — 0

Data & Quality Standards: Which challenges to the scale up (within this Use Case) could be addressed by industry?

Anonymous — 10

incorporate standards into their products (e.g. EHR, equipment) and provide easy means of use

Anonymous — 7

Transparent adoption of FAIR data standards

Anonymous — 4

Limited return on engagement for effort required

Anonymous — 1

Capacity

Data & Quality Standards: Which challenges to the scale up (within this Use Case) could be addressed by other stakeholders (please specify stakeholder)?

Anonymous — 10

Data curators - working in semantic interoperability

Anonymous — 9

Hospital IT personnel

Anonymous — 6

Patients having access to their own data is a good leverage to data stndards

Anonymous — 4

make text mining / natural language processing a reality (not all is structured data)

Technical Infrastructure: Which challenges to the scale up (within this Use Case) could be addressed by industry?

Anonymous — 11

Develop Doctor-friendly user interfaces for tools in federated infrastructure

Anonymous — 8

pseudonimisation/key management/persistent identifiers

Anonymous — 8

trusted working/research environments

Anonymous — 3

Linguistic technologies

Technical Infrastructure: Which challenges to the scale up (within this Use Case) could be addressed by other stakeholders (please specify stakeholder)?

Anonymous — 10

connecting hospitals with the European HPC environment - connectivity and security issues.

Anonymous — 5

IT storage

Anonymous — 3

Hospitals

Anonymous — 3

Academia + research

General recommendations

Anonymous — 3

Engage citizens especially Patients

G Giovanni — 3

Engage Doctors

Anonymous — 2

Agree on a date that all stakeholders stick to/execute made agreements/standards/etc

Anonymous — 1

Return on engagement for clinicians who enter data

Anonymous — 0

Thank you!

company/b1mg-project/

@B1MG_project

//b1mg-project.eu/