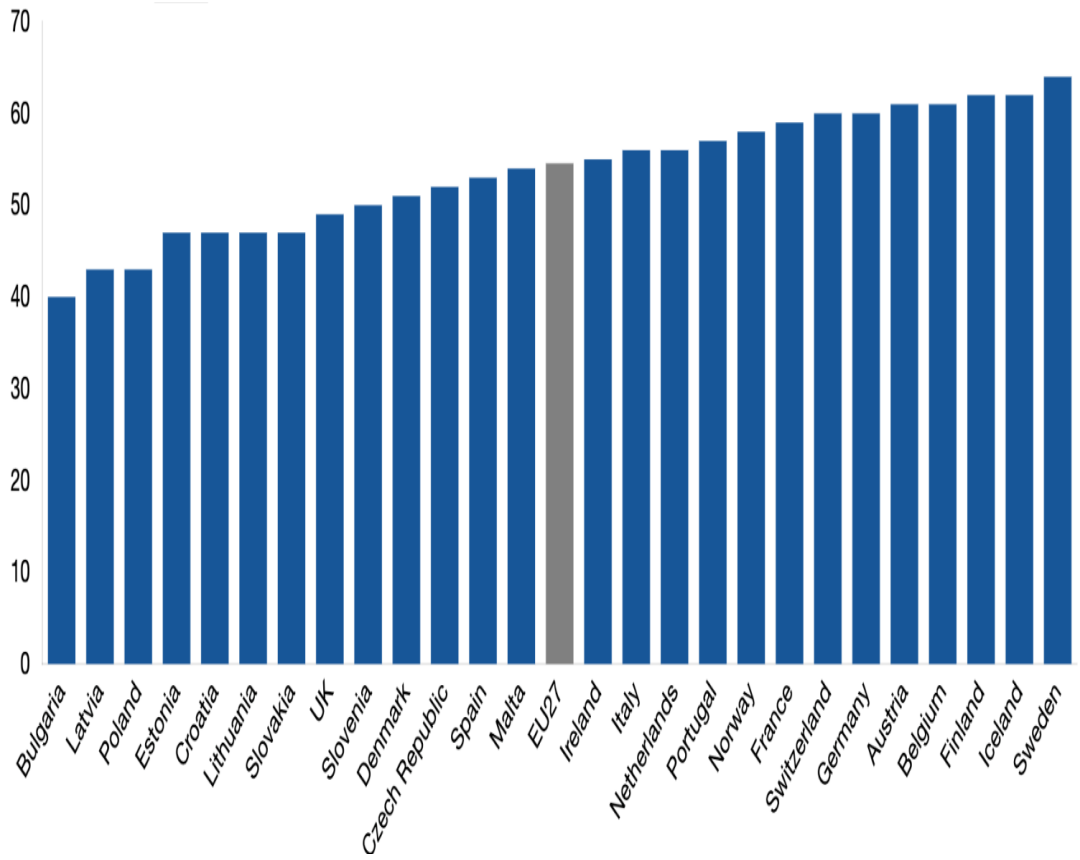
DigiCore

Building a Digital Oncology Network for Europe (DigiONE)

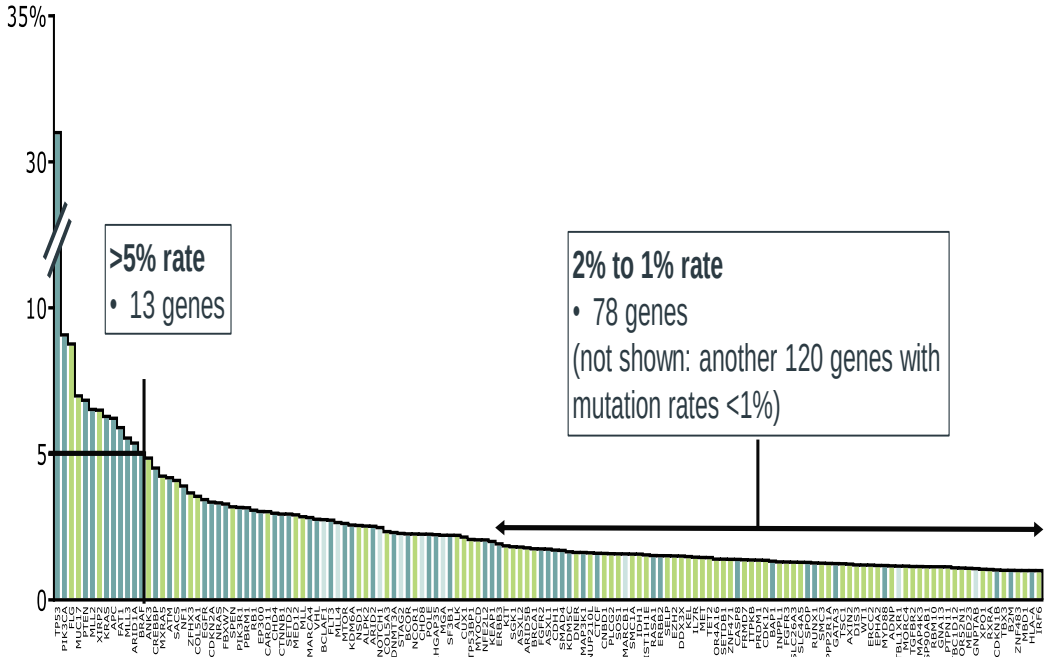
Dr. Piers Mahon

Both care quality improvement and precision oncology research need at-scale international data to improve patient outcomes

5 year age standardised survival (%)



Pan-cancer non-silent mutation frequency (%)



Hofmarcher, T et al. (2019) Comparator Report on Cancer in Europe 2019 - Disease Burden, Costs and Access to Medicines. IHE Report 2019:7
Mahon & Tenenbaum, 2015

DigiONE Pilot: €3M for technology investment in proof of concept to automate and federated cancer outcome research under GDPR

Objectives for DigiONE – Launched in Jan, network meeting in March in Frankfurt

1. Define a **scalable common international minimum dataset for cancer**, building from French OSIRIS
2. **Achieve interoperability and high data quality** on that dataset between 6 centres across Europe under GDPR
3. **Federate those centres** to allow aggregated statistics like counts and to answer simple research questions, with appropriate information governance and contracting
4. **Link routine molecular and clinical data** (despite the format challenges on molecular PDFs)
5. Work out how to **scale up digitally less mature hospitals** with a **variety of technologies and vendors** in DIGICORE's learning – by- doing community

DigiCore

If we can automate disease natural history / outcome studies, we can then automate any observational study (and ultimately pragmatic trials)

THE WORLD TODAY ⟷

DIGIONE PILOT

THE FUTURE…

Traditional eCRF outcomes study	Semi-automated outcomes study	Automated a Cancer outcomes study
• Manual retype post consent into central eCRF	• Protocol specific common data model • Direct extraction from EHR + manual abstraction free text • Meta-analysis	• Network wide CDM, like Cancer OMOP • Data "ready" before protocol in RDR • Key technology PoC • Federated research
• *Unicancer ESME cohorts or traditional clinical registry*	• *ORWIC in ovarian* • *IDEAL4RWE* • *I/o optimise*	• *Federated OMOP disease natural history and outcome studies*

- **Care quality management**
  - \+ *Relevant local analytics*
- **Narrow cohorts like comparative effectiveness**
  - \+ *Bigger network*
- **Biomarker validation**
  - \+ *Routine biomarker linkage*
- **Discovery –omics**
  - \+ *Biobank linkage*
- **Pragmatic trials**
  - \+ *Randomisation, A.E. & prospective consent*

We built international consensus across 16 hospitals in 13 countries to define a minimum data model for cancer outcome research: MEDOC

Step 1: 5 experts, 4 countries

Step 2: 16 Hospitals, 13 countries

1. National cancer datasets (# items)

- UK COSD (~1200)
- German ADT (~300)
- French OSIRIS (65) (and chosen as start point)

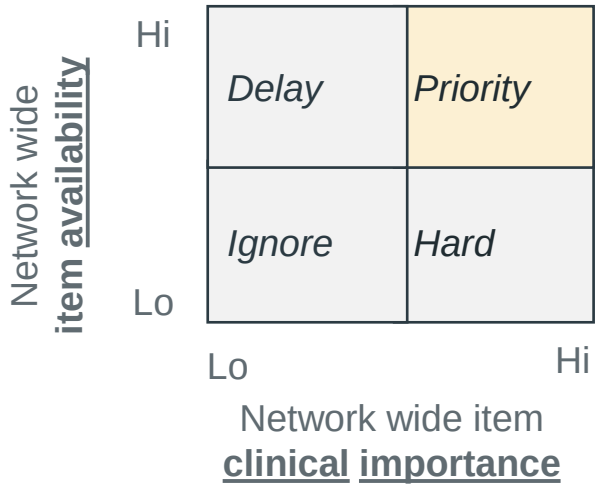
2. International outcomes research experience

- Data availability in EHR
- Importance of item in outcome research

3. Expert hypothesis modified from OSIRIS

- Senior experts in France, Italy, Germany, UK to “slim down” OSIRIS where possible, identify gaps (e.g. weight & height for cachexia or chemotherapy dosing)

4. Clinical priority / feasibility trade-offs by e–survey



5. Traditional consensus process on the “contentious items”

- Item by item discussion on the “Hard” variables to agree pragmatic solutions
- For example, focus on the CCI co-morbidities (not all co-morbidities) to simplify NLP implementation

MEDOC defines a minimum data standard most Cancer Centres can achieve

1. Demographics (=6)

1.1.Date of birth (month)

1.2 Sex

1.3 Weight *& timestamp*

1.4 Height *& timestamp*

1.5 Healthcare ID

1.6 Legal basis for data processing, *e.g.* consent or non-opposition

Notes: legal basis and a healthcare ID are likely to be in national schema, and may be multi-concept in some counties or settings

Key:

Yellow = item must follow local / national rules or norms
Red = Item not in original OSIRIS starting 65 concepts
Italics = implementation notes

2. Clinical Phenotype (=7)

2.1 ICD10 for primary diagnosis and comorbidities (*& timestamp*)

Note: comorbidities often will need NLP and will be optimised for only the 17 CCI co-morbidities

By implication, often not complete

2.2 Charlson comorbidity index (CCI – *timestamped and derived*)

2.3 Date of primary diagnosis

2.4 Method of primary diagnosis

2.5 Performance status (e.g. ECOG, Karnofsky) & *timestamp*

2.6 Disease stage *& timestamp* (e.g. TNM, size, node and metastasis)

2.7 Histological cell type & *timestamp* (e.g. ICD-O-3)

Note: we anticipate multiple cancer specific schema for stage and cell type and will phase implementation

3. Biomarkers (=3)

3.1 Biomarker name *& time stamp*

3.2 Biomarker measure *& time stamp*

3.3 Biological sample identifier & *timestamp*

Notes on biomarkers:

We will aim to get to the same level of detail as in OSIRIS –omics for biomarkers anticipated in the guidelines in 2024 from the drug pipeline (even if from NLP / OCR)

Tests formats will cover:

- *Core routine Blood biochemistry commonly used in cancer*
- *IHC – including HER2+ low*
- *FISH*
- *Somatic mutations, likely as amino acid change or similar*
- *Germline, e.g. BRCA1*

4. Treatment (=14*)

4.1 Line of therapy

4.2 Anti-cancer treatment name

4.3 Molecule generic name

4.4 Start date for drug treatment

4.5 Treatment dose

4.6 End date for drug treatment

4.7 Radiotherapy type (e.g. procedure code of treatment)

4.8 Radiotherapy Start date

4.9 Radiotherapy dose

4.10 Radiotherapy end date

4.11 Surgery type (e.g. procedure)

4.12 Surgery date

4.13 Participation in clinical trial

4.14 Date of trial consent

5. Outcomes (=6)

5.1 Date of death (any location, in-hospital or from national deaths)

5.2 Time to next treatment (*derived*)

5.3 Metastasis presence / absence

5.4 Metastasis location

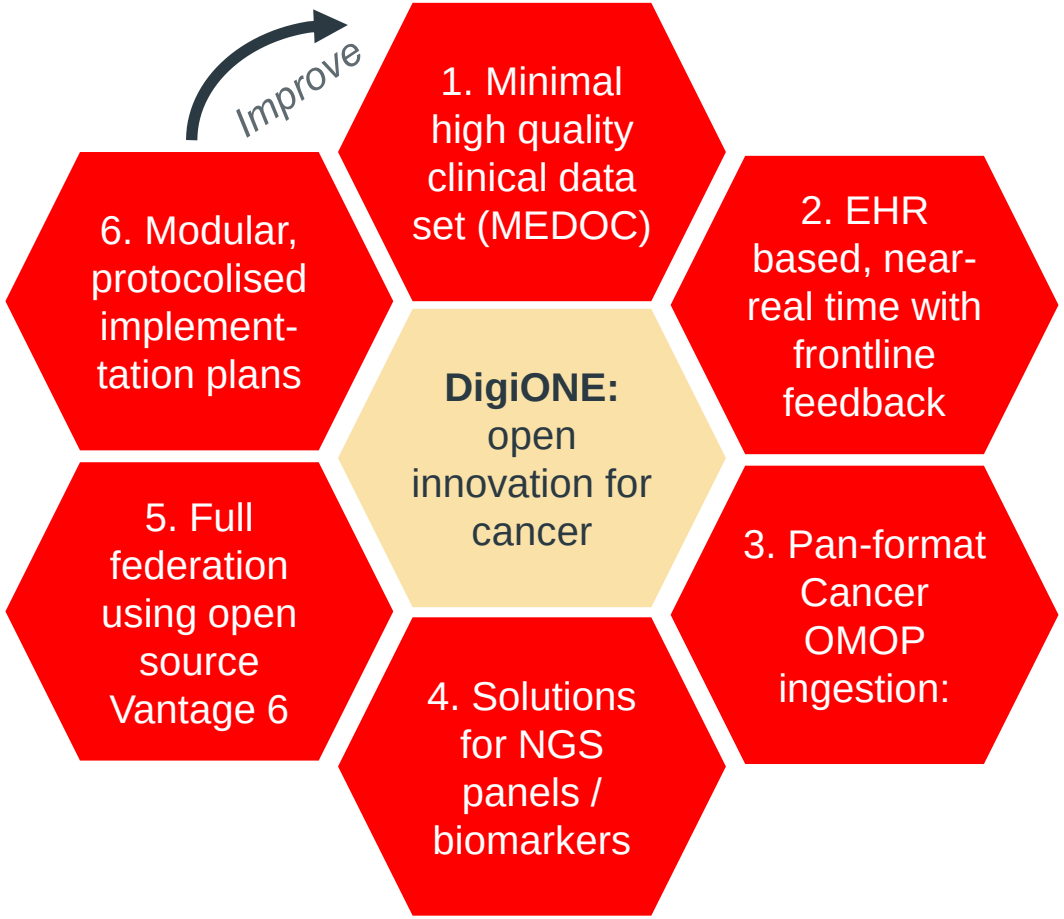
5.5 Date of last visit/follow-up

5.6 Vital status (derived)

Note: routine death registry linkage is not allowed in some European countries, and will require careful design of delivery of 5.1 and 5.6

*** Notes on Treatment**
In some countries we anticipate that claims data is not accessible, only the core EHR which may need NLP routines to extract dates. Where claims data accessible, dates may be derived via timestamps

Beyond MEDOC, DigiONE has many innovative technology features (6 abstracts accepted to OHDSI Europe – annual OMOP conference)



1: **Minimal Essential Description Of Cancer** (MEDOC)

2: **Near-real time frontline feedback loops** to improve data

3: **Pan-format Cancer data ingestion**. Not just ETL also NLP, OCR

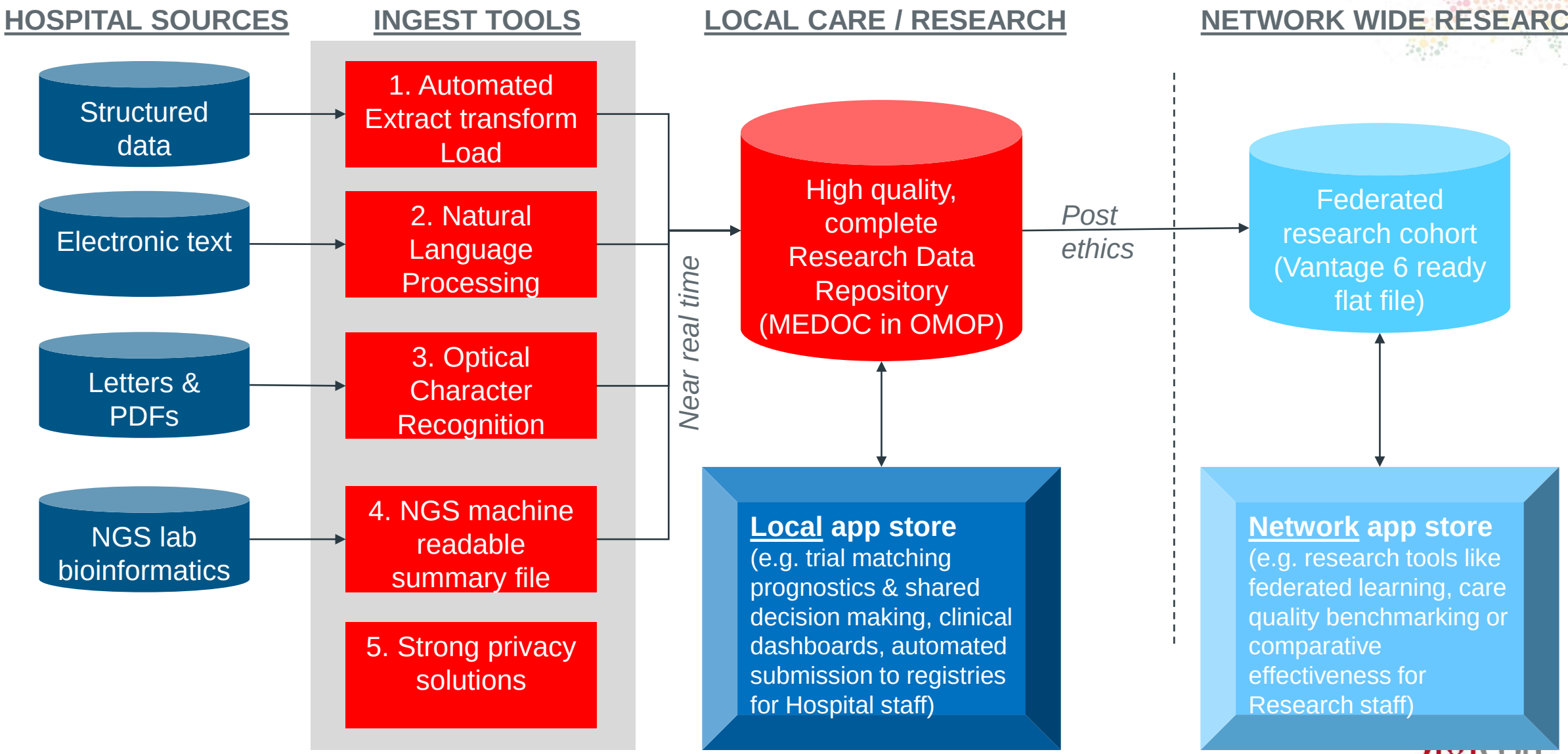
4: **GDPR recital 34 privacy conserving solutions for NGS**

5: **Full federation with open source Vantage6** to allow statistical analysis equivalent to centralised data, but without data pooling

6: **Modular, protocolized implementation plans** to solve for limited data normalisation skills in most hospitals

7. **All in open standards and vendor agnostic**

Feature 3: Pan-format data ingestion



Feature 6: modular, protocolised implementation plans to help hospitals with little OMOP knowledge normalise their data using their research skills

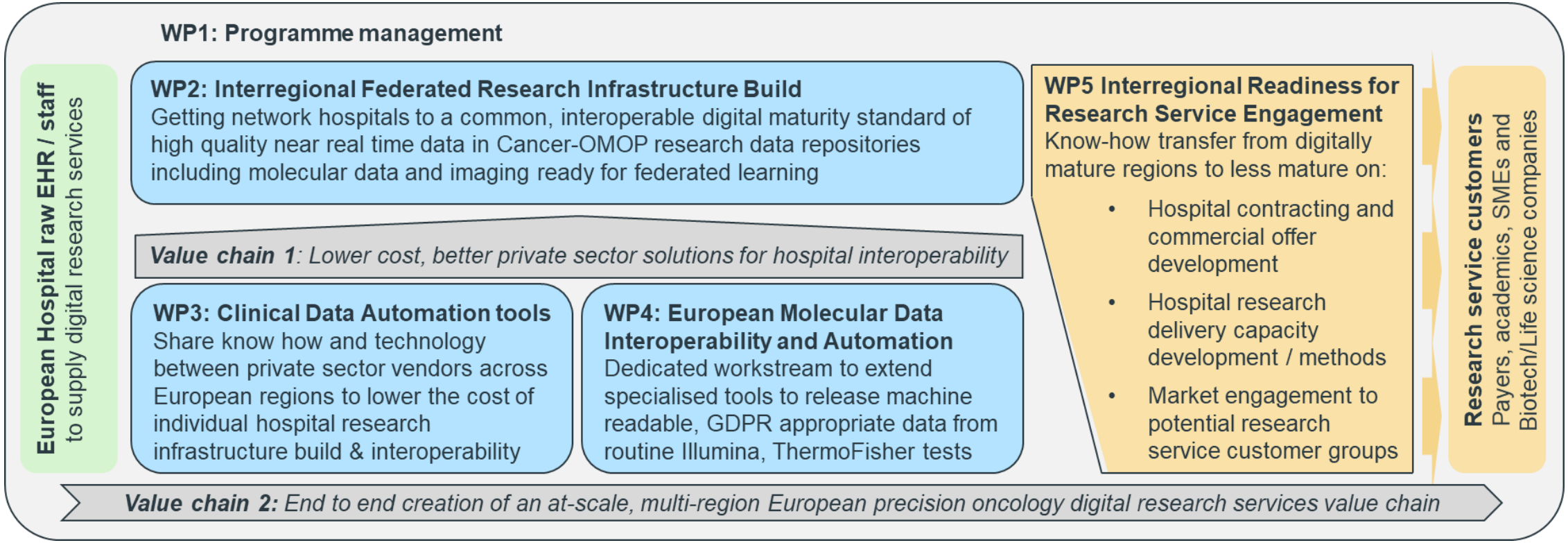
	Dx volumes during Covid19 and recovery	Benchmarking Access To INnovation	Federated Advanced Prognostics	Disease Natural history, outcomes with care quality
Acronym	*C19*	*BATIN*	*FEDAPT*	*DINASTY*
Description	Describe volume of cancer diagnoses, time to Tx initiation and 12-month survival prior to, during and post C19	Examine whether access to innovative drugs, tests or procedures varies by ECOG, sex or age	Predicting 2-year survival; Testing multiple models with increasing # data concepts	Natural history and treatment outcomes studies with care quality assessment
# Data concepts	**7**	**15**	**15**	**36**
Examples (key extension vs C19)	Age, sex, ICD10, Dx date, Tx initiation, date of death, basic staging	+ ECOG, drugs, procedures, trial access	+ full TNM, location metastases, comorbidities, ICD03	All MEDOC: Biomarkers, TTnT, Tx dose
Coverage	Solid cancers (excl. haem and non-melanoma skin)	Wave 1: NSCLC, then ovarian, breast Wave 2: prostate, head & neck, CRC (proposed)		
Complexity	Very simple	Simple	Moderate	Hard

A Disease Natural history study with care quality assessment fully automates outcomes research, and measures guideline compliance

DINASTY care quality features in Ovarian cancer

Pathway element	Care quality feature	Pathway element	Care quality feature
Diagnosis	Diagnosis by formal histology (rather than cytology)	**Monitoring**	Frequency of post primary therapy monitoring
Surgery	Proximity of surgery or chemotherapy to diagnosis		Routine use of full blood count (FBC), urea and electrolytes (U&E) at outpatient appointments
	Surgery by a gynaecological oncologist		Routine use of CA125 at outpatient appointments
	Optimal surgery (no residual or < 1 cm)	**Second line**	Platinum-sensitivity-appropriate chemotherapy used
Testing / initial work-up	Proximity of germline BRCA 1&2 to diagnosis		Use of PARPi in eligible untreated patients (BRCA, HRD or HR +ve)
	Proximity of somatic BRCA 1&2 to diagnosis		
	Proximity of HRD to diagnosis		Platinum resistant relapse treated with appropriate chemotherapy
	Proximity of MSI/MMR to diagnosis		
	Proximity of extended germline R207 panel to diagnosis		Platinum sensitive relapse treated with platinum-based chemotherapy (dose dense, gemcitabine synergy)
Adjuvant / first line	Proximity of platinum-based chemotherapy to surgery		
	Platinum-based chemotherapy on approved list of regimens		
	Use of bevacizumab in eligible patients	**Key**	
	Use of PARPi in germline BRCA patients	normal text	within MEDOC
	Use of PARPi in somatic BRCA patients	*itallics*	requires extension
	Use of PARPi in HRD patients		
	Use of PARPi in HR proficient patients		
	Primary therapy given as part of interventional trial		

We have secured ERDF funding to scale up the network via the €12.5M DigiONE – I3 project due to start by end 2023



18 hospitals have secured technology funding via DIGICORE
Many others have the technology (but most don't know it)

Funding secured via DIGICORE			
#	Country	Hospital	Fund
1	BE	Grand Hopital de Charleroi	I3
2	BE	St Luc	Pilot
3	CZ	Masaryk Memorial	I3
4	DE	Carl Gustav Dresden	I3
5	DE	Charite	I3
6	DE	Frankfurt University	Pilot
7	DE	Greifswald	I3
8	ES	Tartu	I3
9	IR	Trinity	I3
10	IT	Gemelli	other
11	IT	Regina Elena	I3
12	IT	St Raffaele	Pilot & I3
13	LI	Vilnius	I3
14	NL	Groningen	I3
15	NL	Maastricht	Pilot & I3
16	NO	Oslo University	Pilot
17	PL	Marie Curie Warsaw	I3
18	UK	Leeds Teaching	Pilot

Who else has technology or funding?

- **33** of 38 German Academic research hospitals (and another 5 in Austria / German Switzerland)
- **21** ATOMCAT centres (anal cancer)
- **20** of 35 DRUP trial sites in Holland
- ~**14** PIONEER network in prostate
- **11** IDEA4RC clinical partners (Sarcoma, H&N)