Technologies for digital patient finding

(and trial matching too)

Richard Child
Director, Health Research Ecosystem,
Applied Data Science Centre, IQVIA

Problems

Manually intensive

- High burden on Research Coordinators
- Many feasibility requests and repeated searches
- Requires harmony with IT colleagues

EHR systems not designed for searching

- Basic searches by ICD-10 against structured data
- Many fragmented EHR systems
- Data buried in unstructured notes, reports, charts

Patient Finding a major challenge for studies

- Enrollment targets missed at 48% of centers [1]
- Inclusion and Exclusion criteria are more precise
- Studies are delayed or unable to reach primary endpoints – delaying potential new options for Patients

1 Getz, K. Changing Drug Development Landscape and its Anticipated Impact on R&D Operations. Accessed September 21, 20162

Recent improvements

Data interoperability

- HL7 FHIR rapidly adopted
  Interop Santé France, NHS Digital UK
- Data becoming recognised by EHR vendors as Controlled by the Patient and Centre!

Text Mining & Natural Language Processing

- Level 1 - Ontology assisted search, including synonyms
- Level 2 - Parsing : Condition, Labs, Stage, Grade, ECOG
- Level 3 - Contextual: Family vs. Patient history,
  Negations: ”No”, ”Does not”, ”Possible”

Digital Patient Finding tools

- Designed for all users, not just IT data analysts
- Collation of data from multiple EHRs within a centre

USA – Patient Access Rule, Non-blocking rule

21st Century Cures Act

Signed into law in December 2016, defined interoperability and prohibited information blocking. The Cures Act gave HHS further ability to propose and enact regulations.

ONC "Cures Final Rule"

Released March 2020, details FHIR-based Health IT interoperability requirements and exceptions to information blocking

Assessment and Plan of Treatment

Care Team Members

Clinical Notes *NEW
- Consultation Note
- Discharge Summary Note
- History & Physical
- Imaging Narrative
- Laboratory Report Narrative
- Pathology Report Narrative
- Procedure Note
- Progress Note

Goals
- Patient Goals

Health Concerns

Immunizations

Laboratory
- Tests
- Values/Results

Medications
- Medications
- Medication Allergies

Patient Demographics
- First Name
- Last Name
- Previous Name
- Middle Name (including middle initial)
- Suffix
- Birth Sex
- Date of Birth
- Race
- Ethnicity
- Preferred Language
- Address *NEW
- Phone Number *NEW

Problems

Procedures

Provenance *NEW
- Author
- Author Time Stamp
- Author Organization

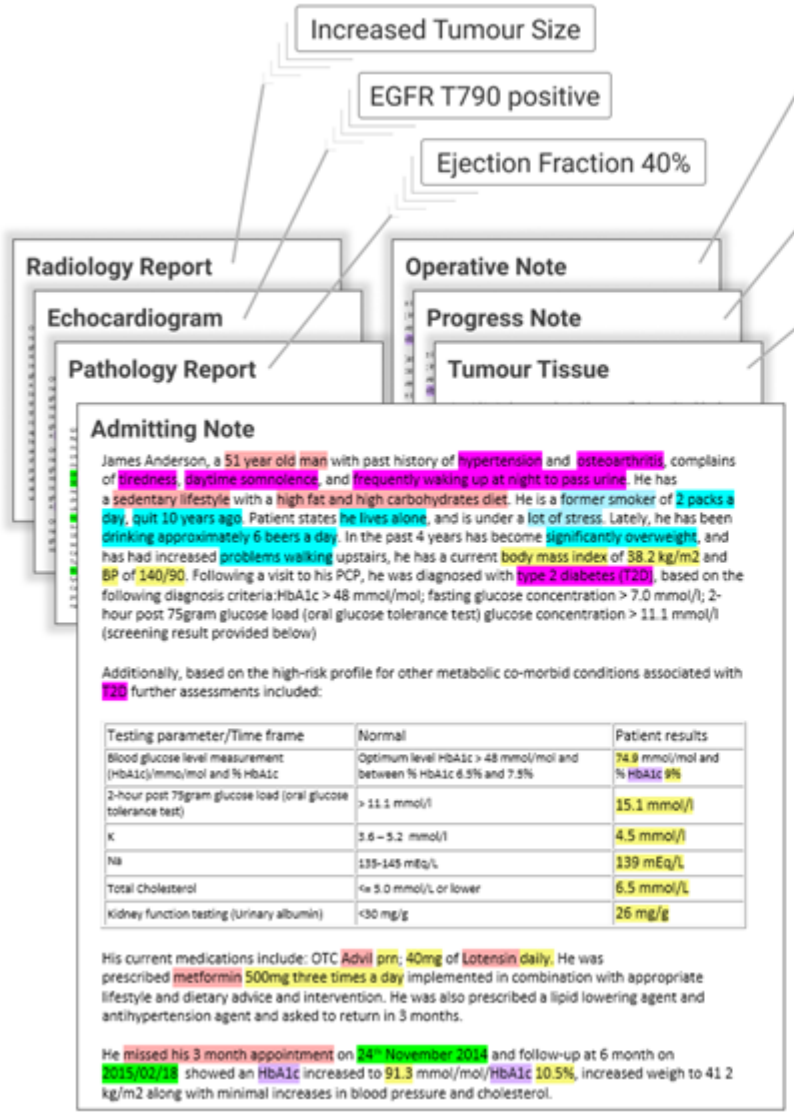
Smoking Status

Unique Device Identifier(s) for a Patient's Implantable Device(s)

Vital Signs
- Diastolic Blood Pressure
- Systolic Blood Pressure
- Body Height
- Body Weight
- Heart Rate
- Respiratory rate
- Body Temperature
- Pulse oximetry
- Inhaled oxygen concentration
- Pediatric Vital Signs *NEW
  - BMI percentile per age and sex for youth 2-20
  - Weight for age per length and sex
  - Occipital-frontal circumference for children < 3 years old

DigiCore

Natural Language Processing (NLP)



Diseases

Text	Normalized Value
Type 2 Diabetes T2D	Diabetes Mellitus, Type 2 SNOMEDid 44054006

Symptoms

Text	Normalized Value
Sleepiness Daytime Somnolence	Fatigue/tiredness
Frequently waking up at night to pass urine	Nocturia Urinary frequency

Social Determinants

Text	Normalized Value
Former smoker Quit 10 years ago	Ex smoker
Problems walking	Ambulatory Status: walking difficulty
Lot of stress	High levels of stress/Stress
He lives alone	Social Isolation
Missed his 3 month appointment	Did not attend

Demographics

Text	Normalized Value
51 year old	51y
man	Male

Dates

Text	Normalized Value
24th November 2014	20141124
2015/02/28	20150228

Measurements

Text	Normalized Value
Body mass index of 38.2 kg/m2	BMI 38.2 kg/m2
15.1 mmol/l	15.1 mmol/l

Medications

Text
On Metformin 0.5g PO three times a day

Normalized value

Medications				
Drug Name	Dose	Unit	Route	Frequency
Metformin	500	mg	Oral	TID

DigiCore

Illustrative examples of modern Patient finding tools

Inclusion

Patients in my population **must** have all of the following terms

Diagnoses
Diagnosis code Maligne neoplasma van prostaat, benig...
Stage IV

AND

PSA
Name PSA
Numeric value ≥ 2

AND

TNM
Name TNM T
Textual value T2

AND

Patient Information (Demographics)
Sex Male
Age ≤ 95 years

AND

Gleason
Name Gleason sum
Numeric value 6 - 9

Patient cohorts

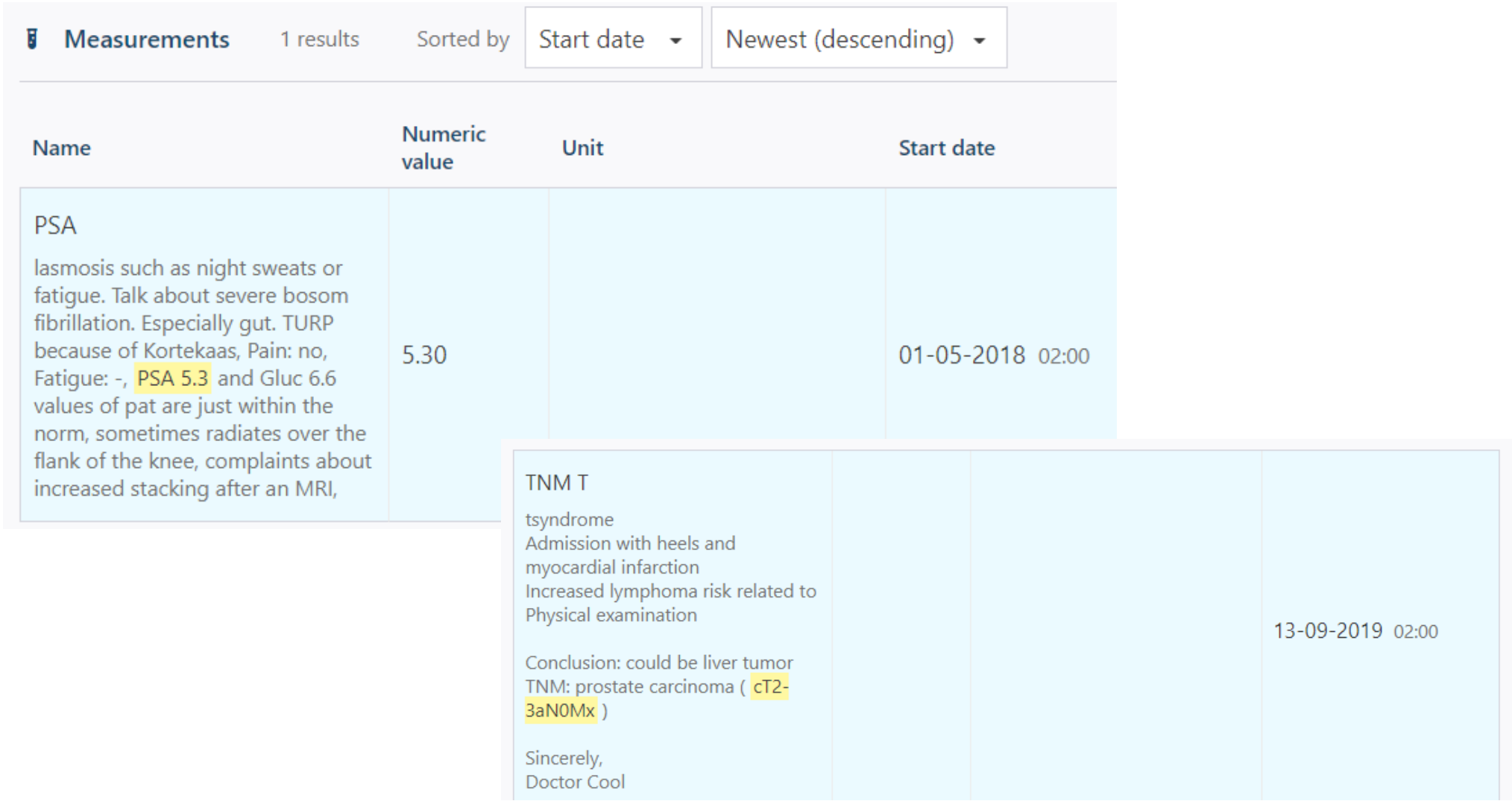
Search results (610) | Included (3) | Excluded (1)

Criteria selection | Manually add patients | Filters

#	Patient Pseudo ID	Sex	Age	Birth year
1	0300_9380-37	Male	44	1976
2	0301_554-1244	Male	61	1960
3	0301_554-13509	Male	72	1949
4	0301_554-14273	Male	74	1947
5	0301_554-1905	Male	64	1954
6	0301_554-9663	Male	53	1968

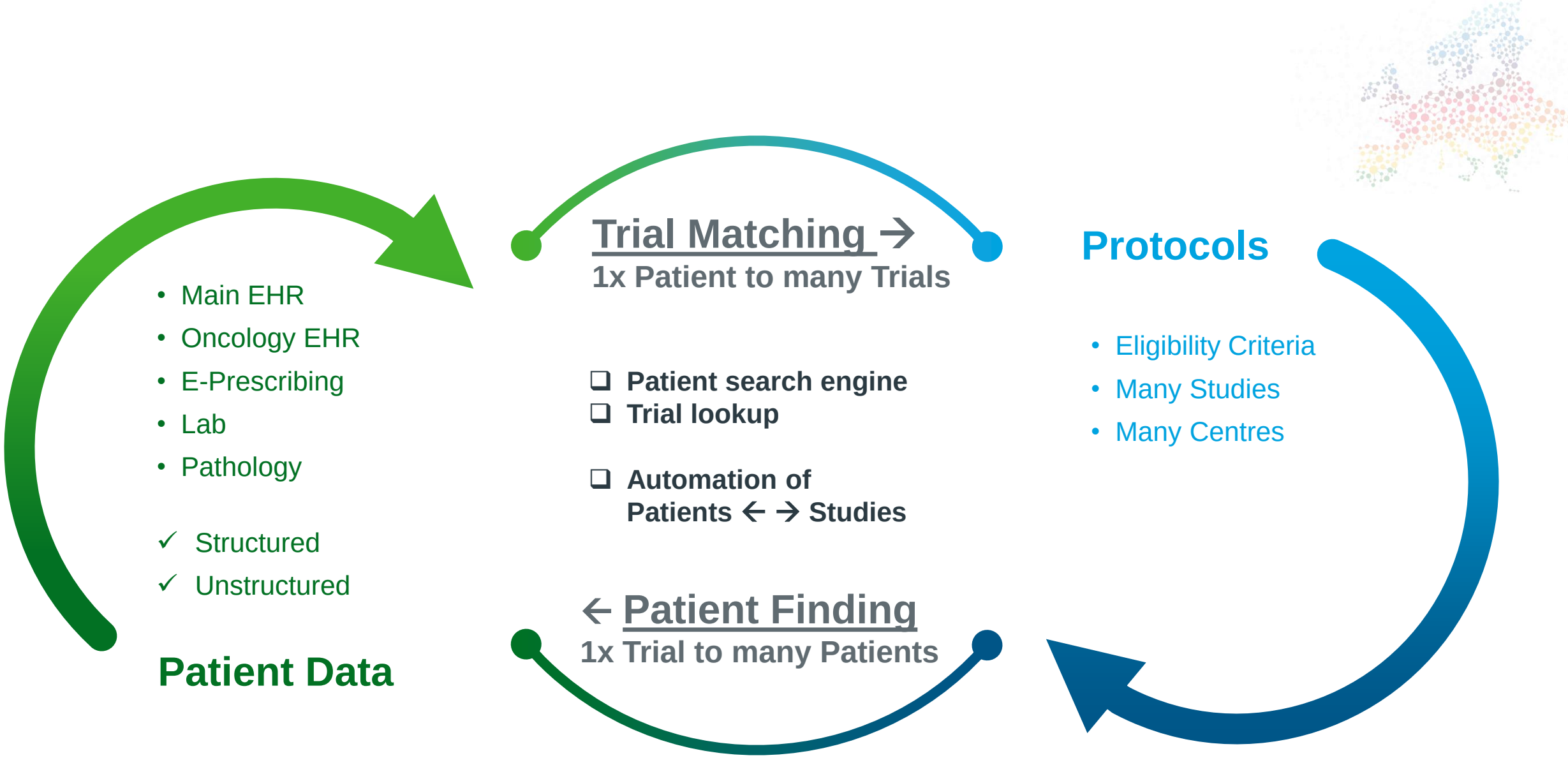
DigiCore

Illustrative examples of modern Patient finding tools



DigiCore

But I know all my Patients and need to find trials!


Trial Matching →
1x Patient to many Trials
Protocols
• Eligibility Criteria
• Many Studies
• Many Centres
❑ Patient search engine
❑ Trial lookup
❑ Automation of
Patients ← → Studies
← Patient Finding
1x Trial to many Patients
• Main EHR
• Oncology EHR
• E-Prescribing
• Lab
• Pathology
✓ Structured
✓ Unstructured
Patient Data

Road from digital Patient Finding to Trial Matching

01

EHR DATA MATURITY

Prerequisite:

Electronic Health Record system(s)

02

PATIENT FINDING

Per-Study protocol digitisation

Share searches (CRO / centres)

03

PROTOCOL DIGITISATION

Digitise ALL protocols

- Tech – NLP
- Standards - PRM
- Community collab.

04

COMPUTABLE PHENOTYPES

Digitised protocols require translation to common data models and terminologies

Carrot of quality

DigiCore

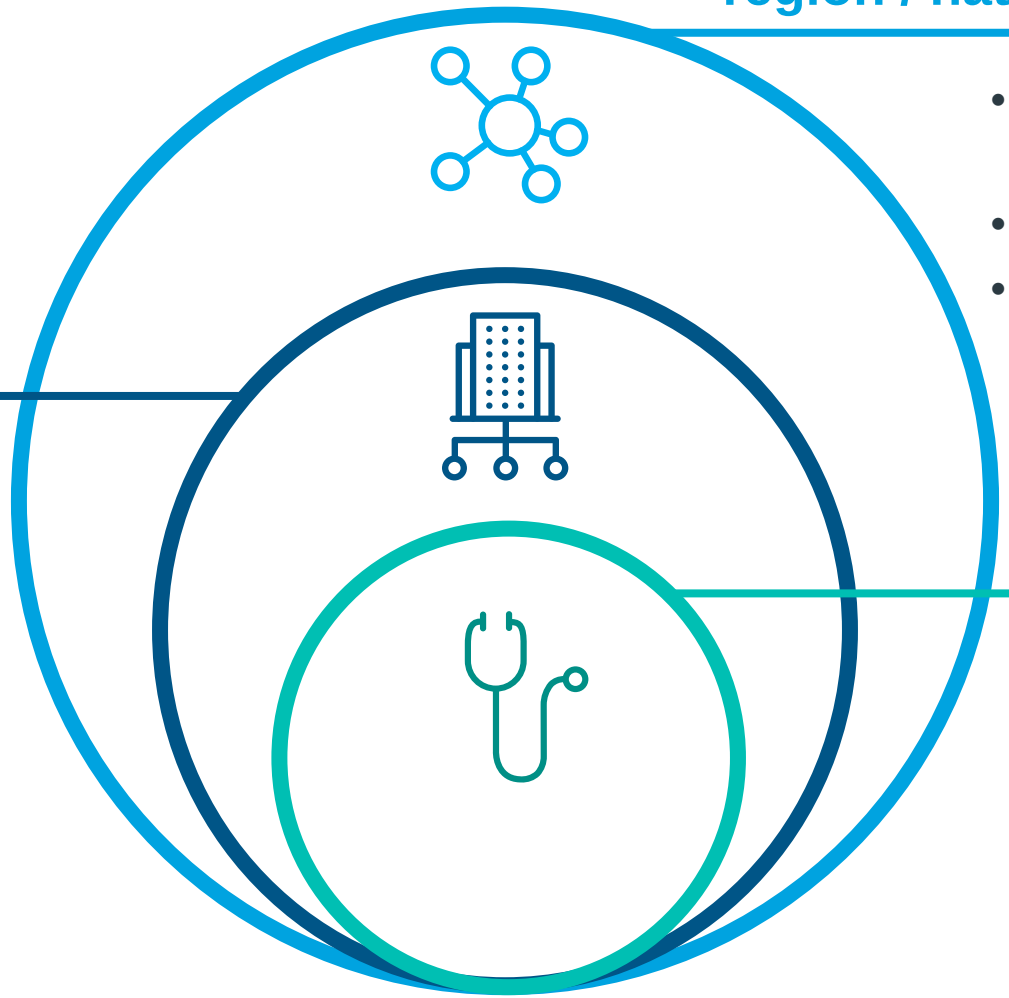
Collaboration and Community

Patients and Studies in a Cancer alliance, or even region / national / Europe

- My Centre may be at capacity, refer Patients to other Centres
- Find Studies with my region
- Share digitised protocol with other Centres and find Patients

Patients and Studies at My Centre

- Visibility of all Studies by colleagues
- Find Patients under care of others



Patients and Studies I know

- Digitise protocols for my studies
- Find Patients for my studies only

Benefits

Benefits of digital Patient Finding

DATA VISIBILITY
- MORE studies & patients surfaced
- Not solely dependent on investigator knowledge

FINANCIAL
- Increased research revenues

OPERATIONAL EFFICIENCY
- Reduce manual searches
- Saves time & effort
- Focus on *eligible* patients

DIVERSITY AND INCLUSION
- Equitable access to studies
- Link with Social Determinants of Health

BEHIND THE FIREWALL
- GDPR privacy by design
- Share digitised protocols and anonymised results, not data

PATIENT
- Avoid unnecessary appointments

FEASIBILITY
- Feedback to sponsor on protocol

USE BEYOND STUDIES
- Cohort / Patient finding for Centre purposes

DigiCore

Digitally-enabled Site ID

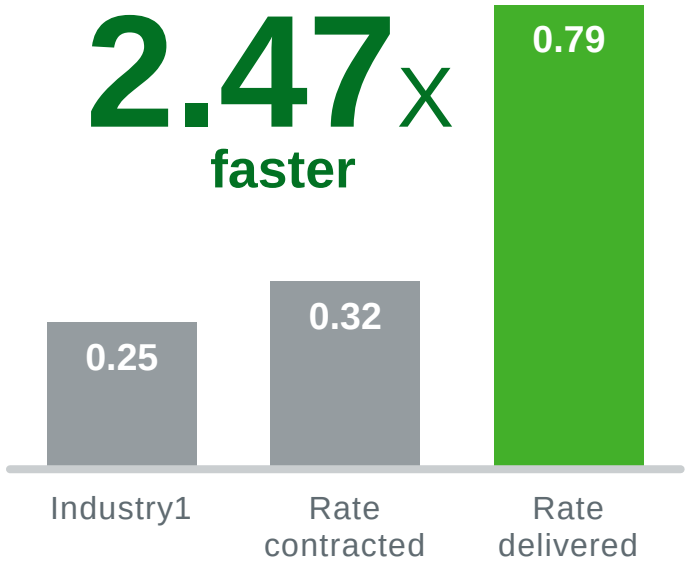
CASE STUDY

Phase II:
cholangiocarcinoma

Higher recruitment rate

Enrollment
(p/s/m = patient / site / month)

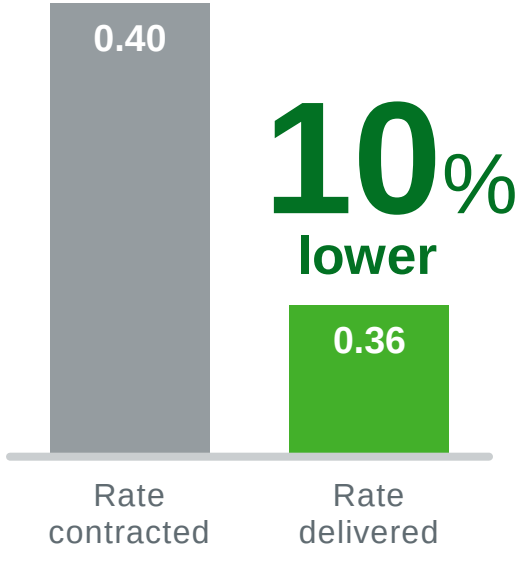
2.47x faster



Lower screen failure rate

Rate of screen failures

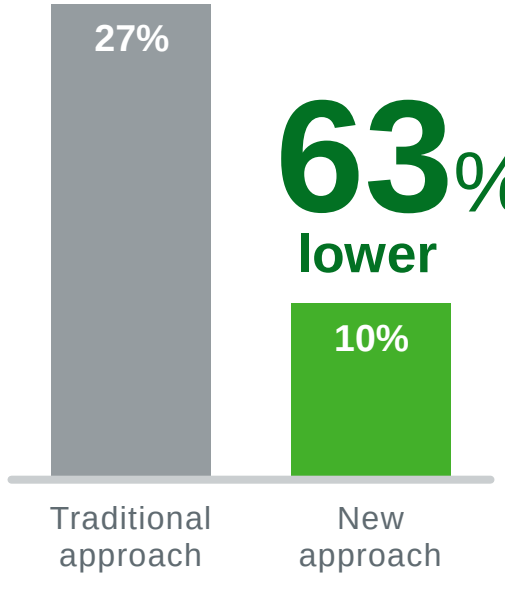
10% lower



Fewer non-enrollers

Non-enrolling sites

63% lower

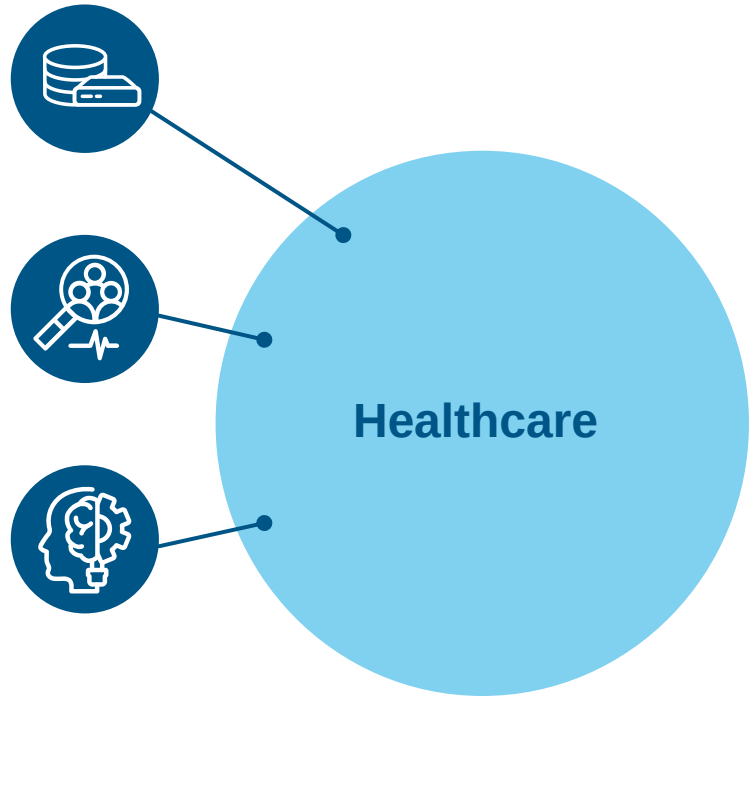


Recruitment completed ~6 months ahead of schedule

DigiCore

Looking forwards

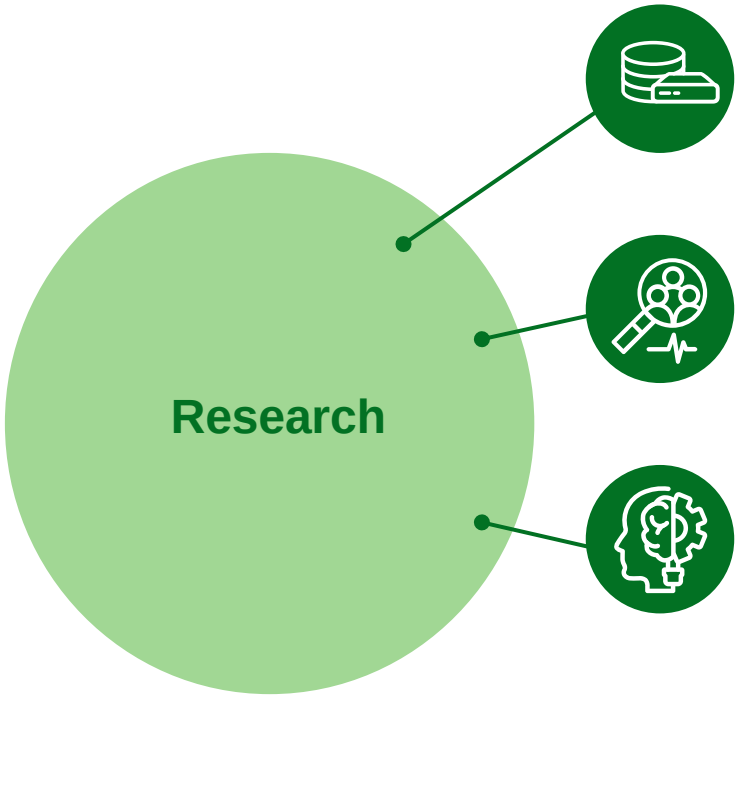
Common problems and needs



Quality and complete data

Find patients
Diagnostic screening
- population health

AI
Enhanced screening
Risk Stratification
Predictive Analytics
Error & Fraud detection



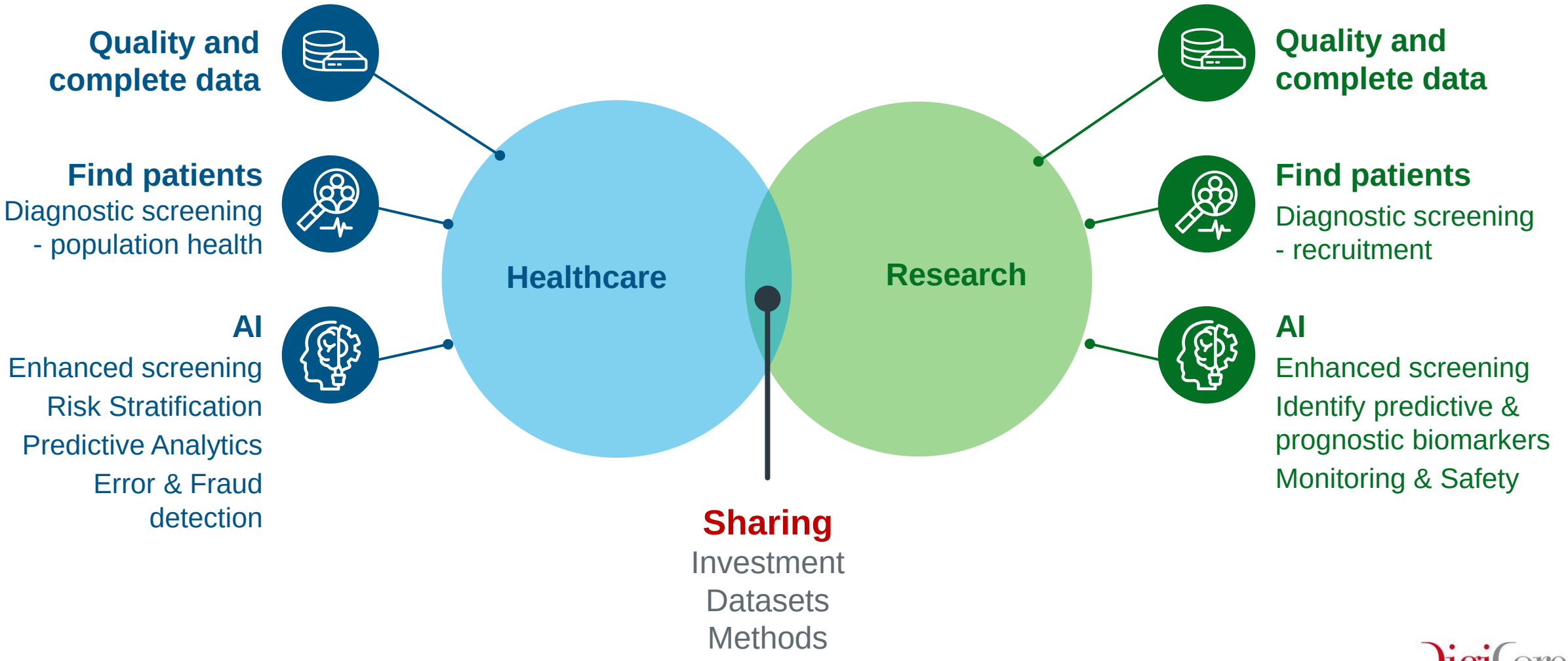
Quality and complete data

Find patients
Diagnostic screening
- recruitment

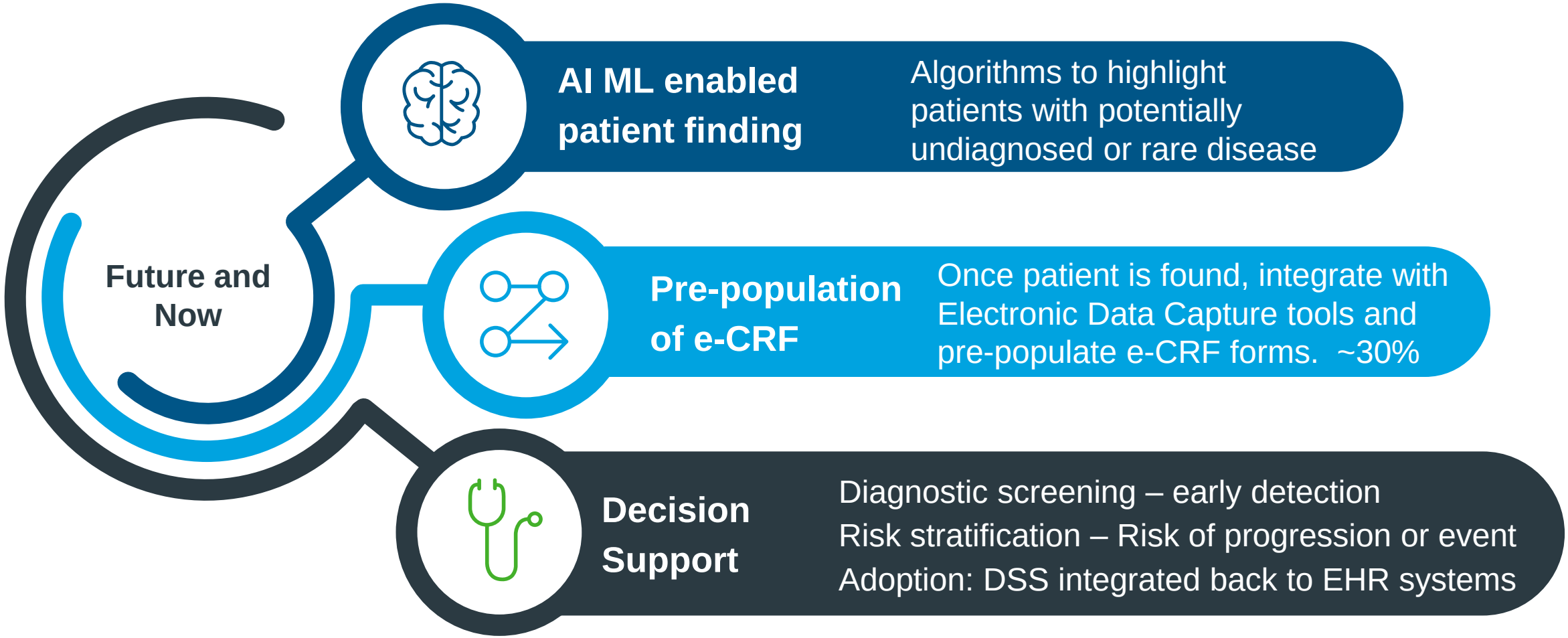
AI
Enhanced screening
Identify predictive & prognostic biomarkers
Monitoring & Safety

Many needs are shared – we need to direct investment
not just to benefit Trials, but also Real World Evidence generation and Patient care

Common problems and needs

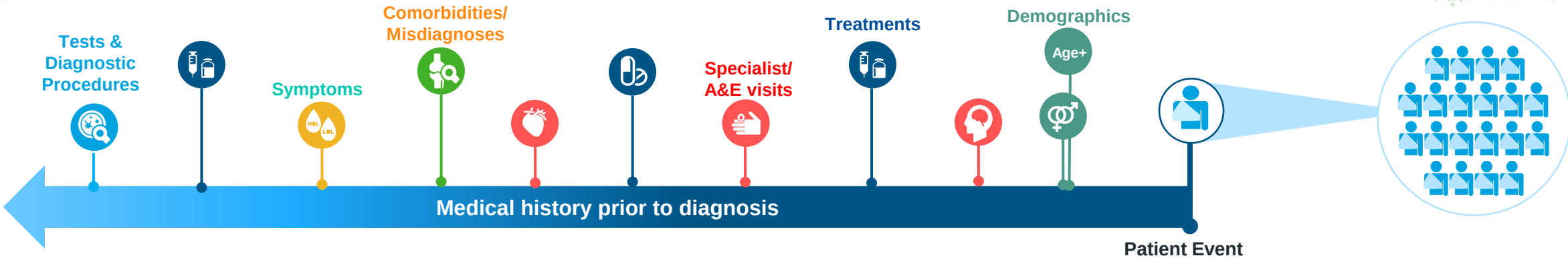


Looking forwards



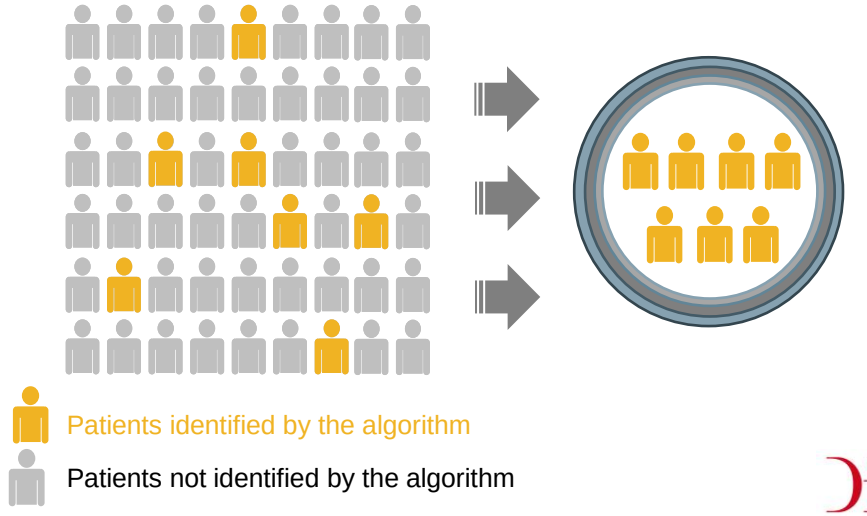
AI Machine Learning enabled Patient Finding

1 Assemble and analyze the often complex medical history of target patients

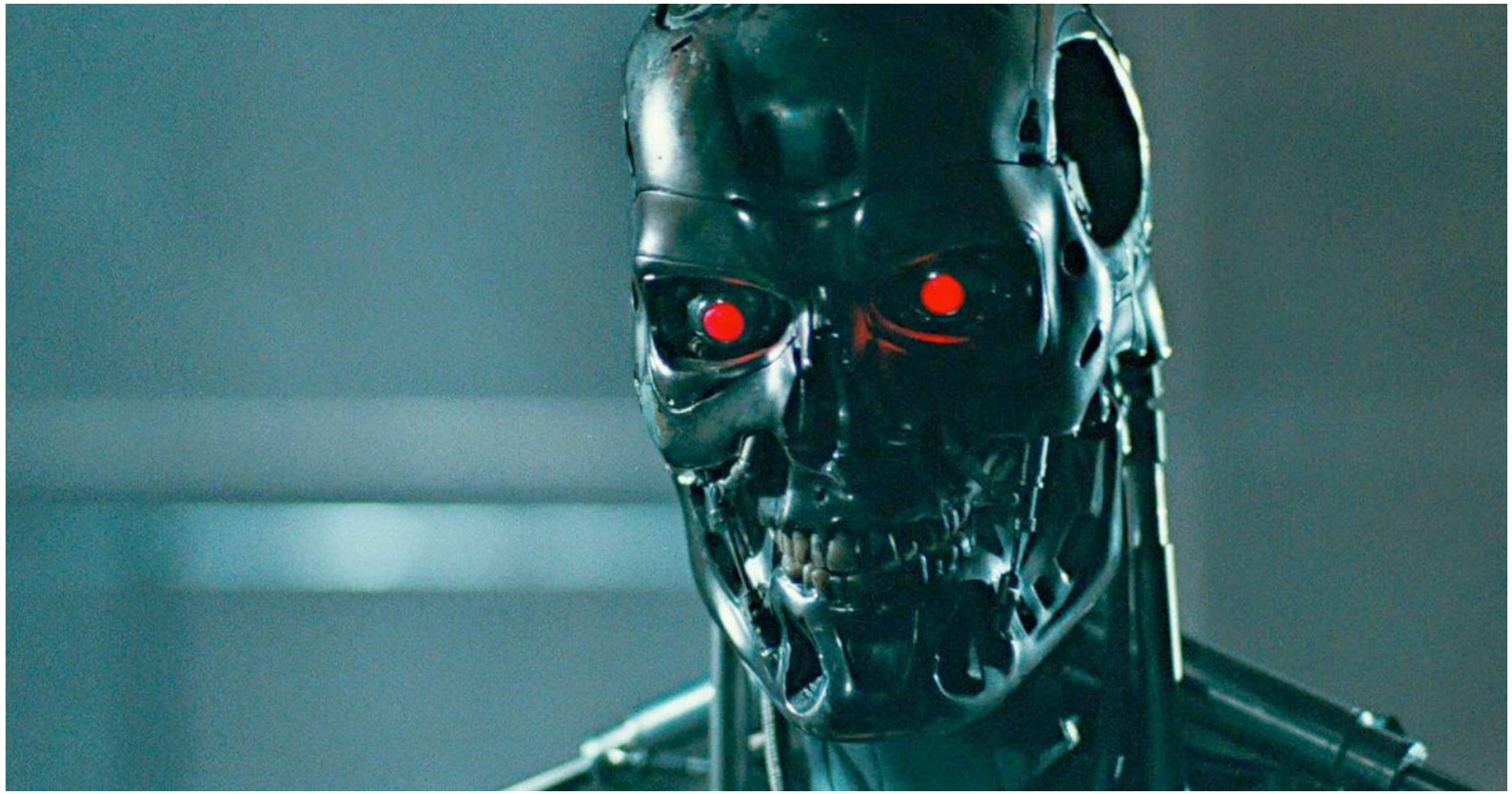


2 Use Machine Learning to identify unique patterns in target patients' medical history PRIOR to Diagnosis or Event

3 Deploy learnings to identify specific target patient populations



Bad AI?



Integrating clinical expertise and machine learning

Case study to identify undiagnosed patients with rare disease showed substantially better results using a hybrid Clinical-AI method in comparison to fully data-driven AI or KOL based knowledge alone

BACKGROUND

- Tuberous sclerosis complex (TSC) is a rare disease with a prevalence of 1 in 15,000
- TSC patients develop Benign multi-system tumours lead to epilepsy, cognitive impairment, kidney failure, etc.
- Manifestations typically start in childhood but correct diagnosis can be delayed into adulthood

CHALLENGE

A client was interested in understanding whether early disease detection algorithms could be used to find undiagnosed patients with TSC

IMPLEMENTATION AND IMPACT

IQVIA conducted a UK study leveraging data from primary (CPRD) and specialist (HES) care. Data covered ~5m patients. **Four algorithms were developed:**

1. KOL-based rules

Rules based on clinical KOL interviews

1 in ~25 patients predicted to have TSC go on to receive a diagnosis of TSC

2. Conventional Statistics

Standard statistical / epidemiology approach based on logistic regression

1 in ~13 patients predicted to have TSC go on to receive a diagnosis of TSC

3. Data-driven AI

Support vector machine based on all available features

1 in ~6 patients predicted to have TSC go on to receive a diagnosis of TSC

4. Hybrid Clinical-AI

Support vector machine incorporating clinically driven feature engineering

1 in ~4 patients predicted to have TSC go on to receive a diagnosis of TSC

AI combined with clinical knowledge performed best

- 2.5 times better than standard epidemiological approach
- 5 times better than rules based on KOL knowledge

Public + Private, Research + Care – DATA-CAN

Our aim is to improve care and outcomes for people with cancer by making high quality health data more accessible for cancer researchers and health professionals

Making high-quality datasets available via an online portal - the Health Data Research Innovation Gateway

Working with patients and the public to make sure that data is used in a transparent and responsible way that benefits the NHS, patients and society

Collecting and using 'real-time' cancer data from the NHS to help health services identify issues and respond to events, such as COVID-19

UK wide partnership of six founding organisations with plans to expand

Supporting more patients into cancer clinical trials through the development of trial matching software

DigiCore

DigiCore

Thank you