CORBEL shared services for life-science

PID5815/Integration of EU-wide cardiovascular datasets

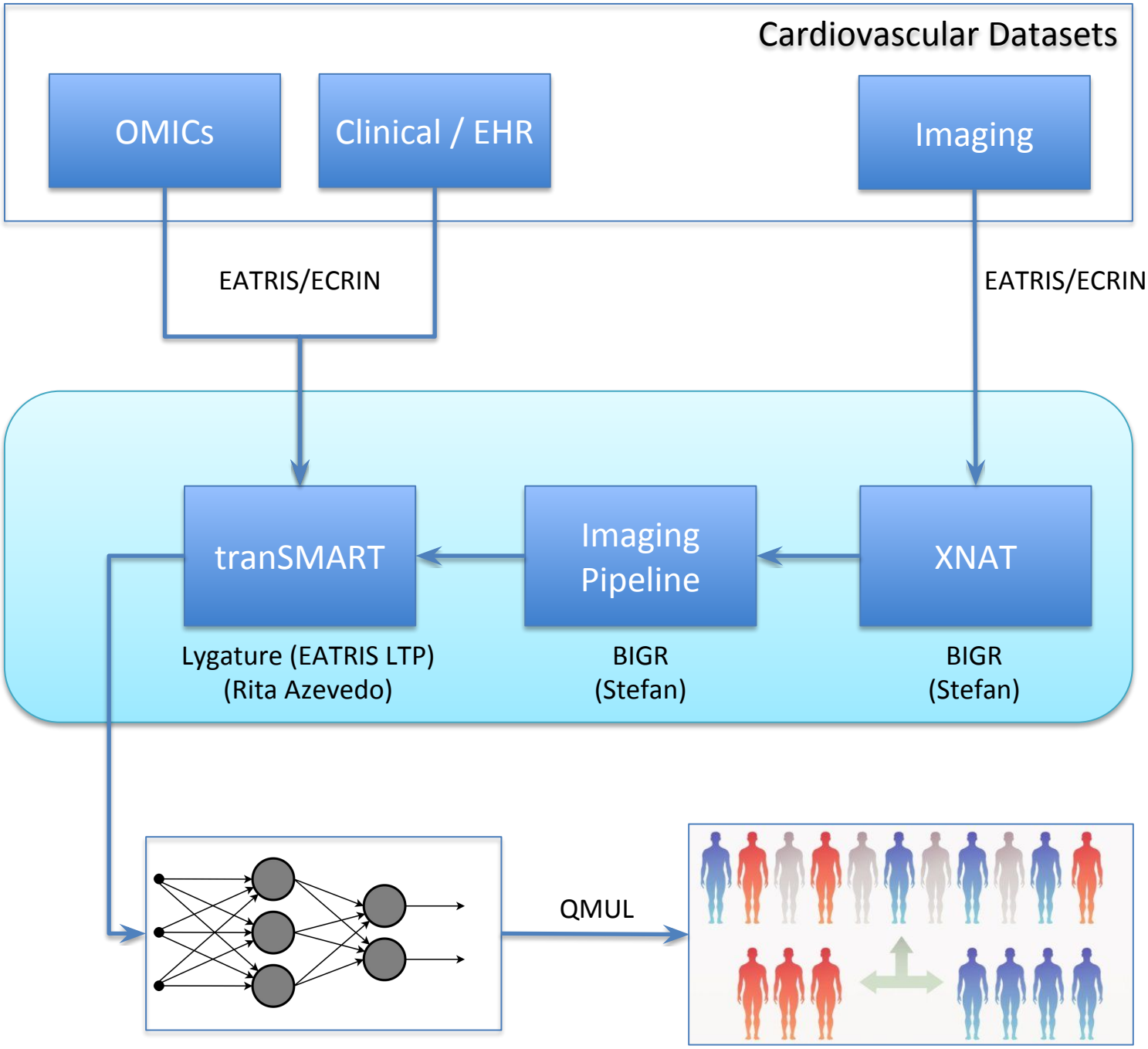
Adriano Barbosa da Silva, QMUL

RI 1: BBMRI-ERIC, Petr Holub
RI 2: EATRIS, Florence Bietrix
RI 3: EATRIS LTP, Rita Azevedo
RI 4: EurobioImaging, Stefan Klein (EPI2)

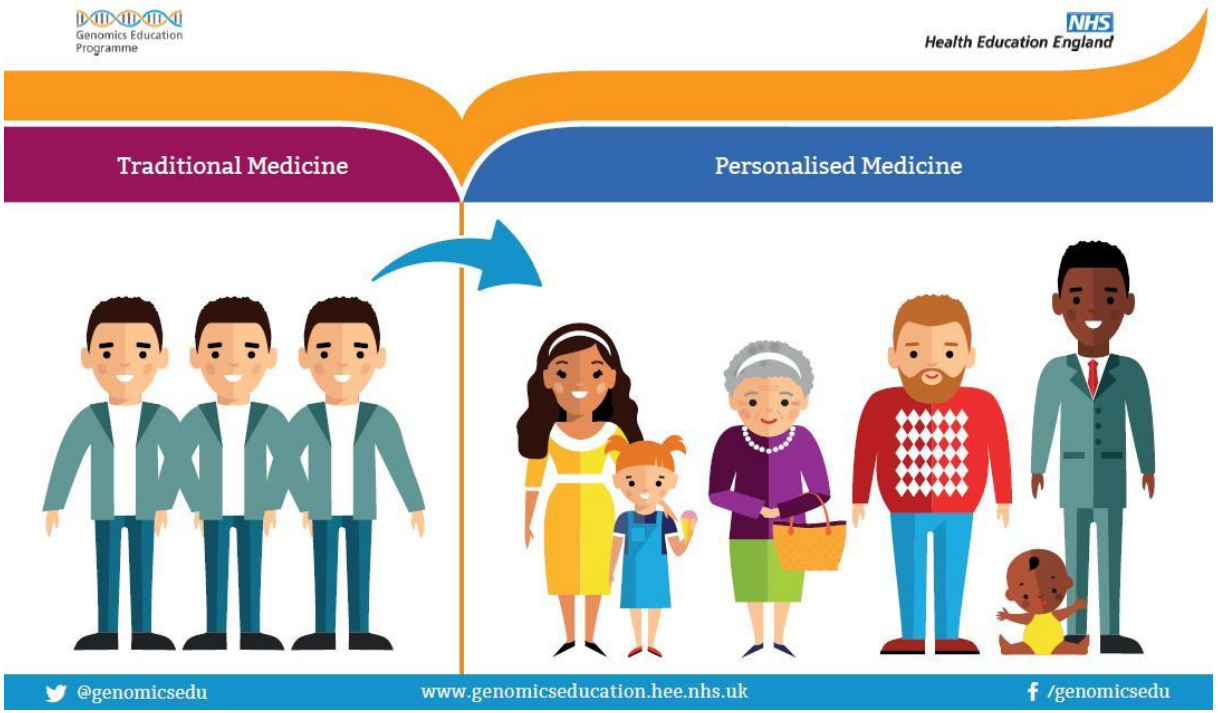
Queen Mary University of London

BBMRI-ERIC
(Petr Holub)
EATRIS (Florence Bietrix)

Queen Mary University of London



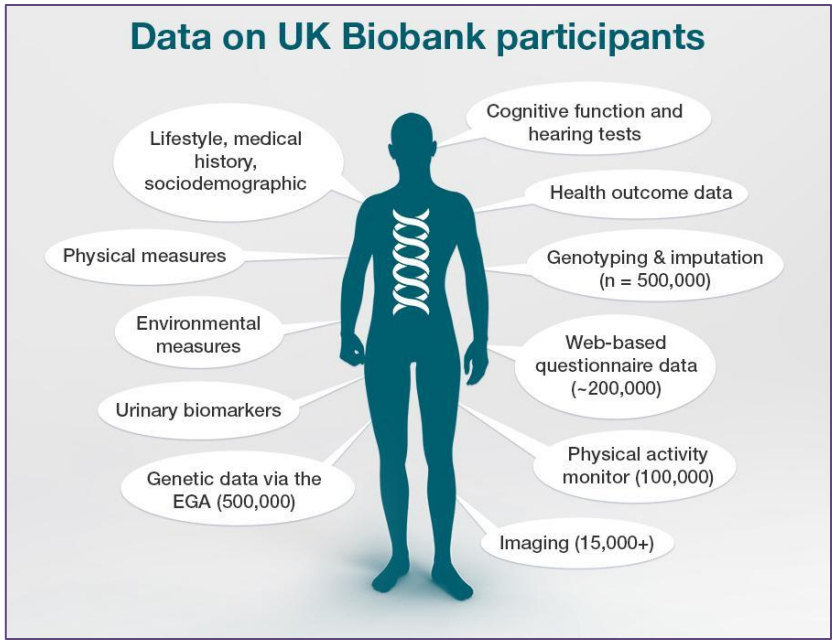
In the future the NHS should treat the Patients and not the Diseases, the main postulate of **Precision Medicine**

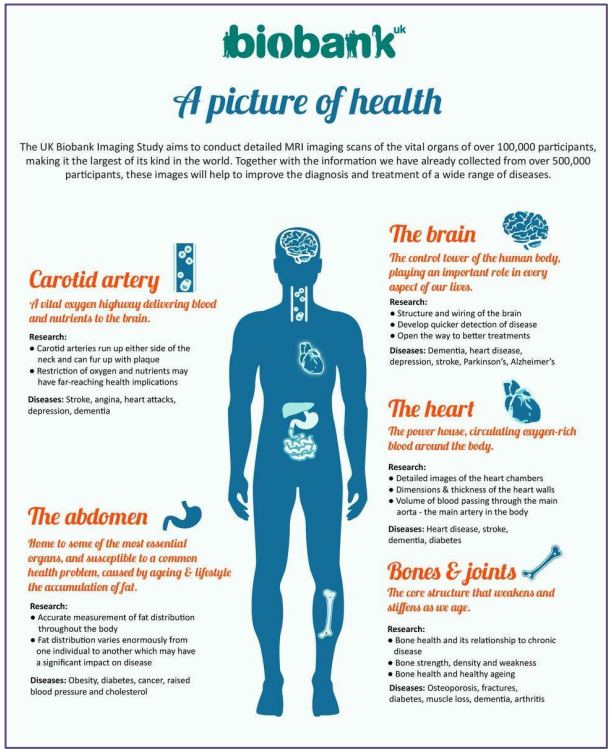


Images credits: NHS Education England

Need for data collection

- Projects Like the UK Biobank aims to collect a wide range of patient's data





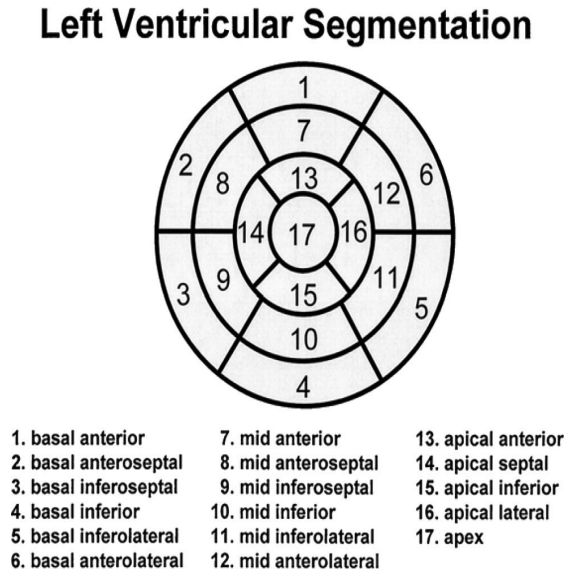
Cardiac Biomarkers

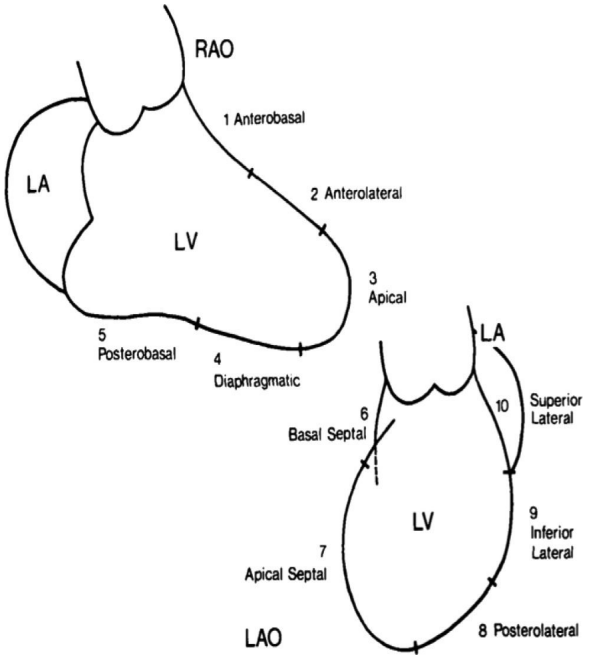
Table 6
The list of the cardiac landmarks being included in the model

Landmark name	Mean model to landmark distance (mm)
Ostium left coronary artery	5.2
Ostium right coronary artery	6.3
Apex	7.1
Aortic valve center	4.2
Mitral valve center	5.3
Tricuspid valve center	5.5
Pulmonary valve center	4.2
Ostium coronary sinus	7.7
Ostium vena cava superior	5.0
Ostium vena cava inferior	8.1
Ostium left inferior pulmonary vein	7.6
Ostium left superior pulmonary vein	7.9
Ostium right inferior pulmonary vein	7.8
Ostium right superior pulmonary vein	6.7
Center left ventricle	4.1
Center right ventricle	6.2
Center left atrium	4.9
Center right atrium	6.9
Heart center	6.9
Bifurcation point LAD-LCX	4.7
Center left atrial appendage	6.9
Center right atrial appendage	7.9
Center LV anterior papillary muscle	5.1
Center LV posterior papillary muscle	6.8
Center RV anterior papillary muscle	7.4

Using the transformations T_i resulting from the procrustes analysis of the surface model, mean landmark positions have been calculated (see Section 5.3). The right column shows the residual mean distances between mean model and landmark samples (see Eq. (14)).

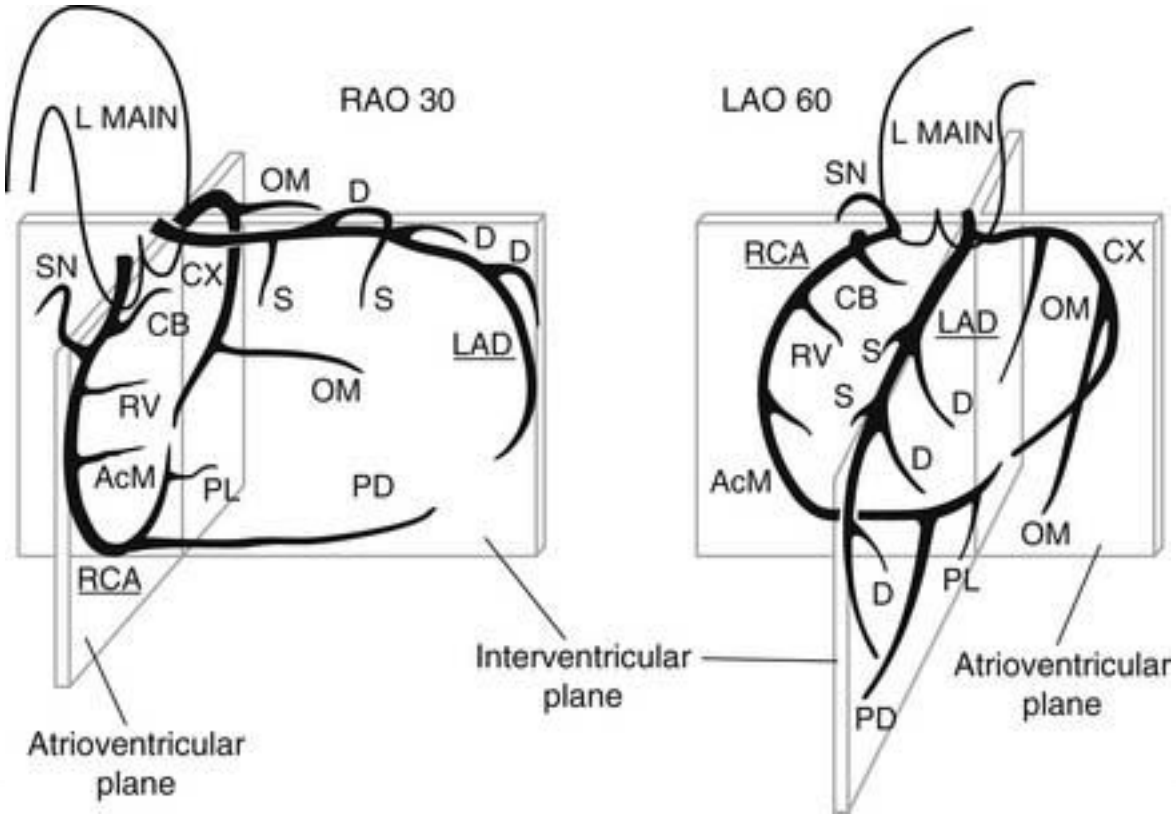
Cardiac Biomarkers





Cardiac Biomarkers

Coronary Angiography: Valve and Hemodynamic Assessment



Data Tsunami

Sanger Institute @sangerinstitute · 20h

As of 15:00 BST, all of our 17 @illumina NovaSeq sequencing machines are up and running, and we're ramping up to sequence 10,000 whole genomes a month for the @uk_biobank Whole Genome Sequencing Project.

3 66 205

Show this thread

Data
Tsunami

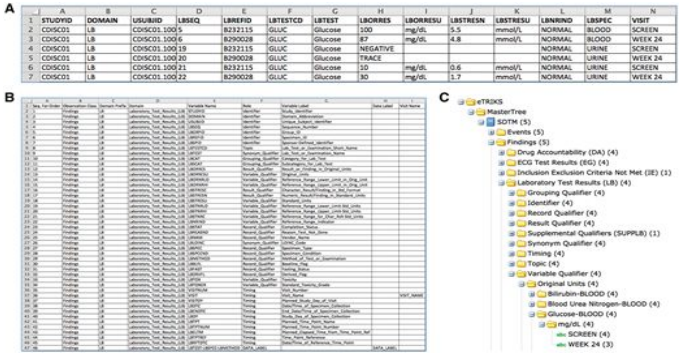
Solutions based on open source software

CORRECTED PROOF

Presenting and sharing clinical data using the eTRIKS Standards Master Tree for tranSMART

Adriano Barbosa-Silva ✉, Dorina Bratfalean, Wei Gu, Venkata Satagopam, Paul Houston, Lauren B Becnel, Serge Eifes, Fabien Richard, Andreas Tielmann, Sascha Herzinger, ...
Show more

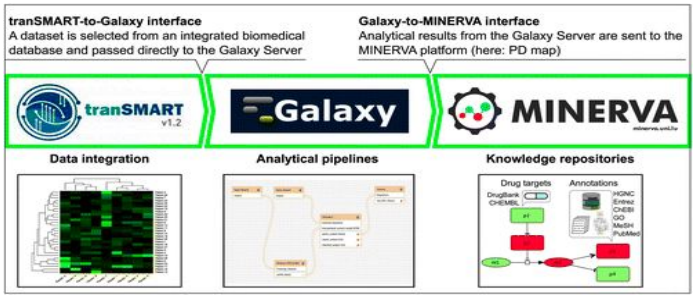
Bioinformatics, bty809, https://doi.org/10.1093/bioinformatics/bty809
Published: 26 September 2018 **Article history** ▾



Big Data, Vol. 4, No. 2 | Original Articles Open Access | CC

Integration and Visualization of Translational Medicine Data for Better Understanding of Human Diseases

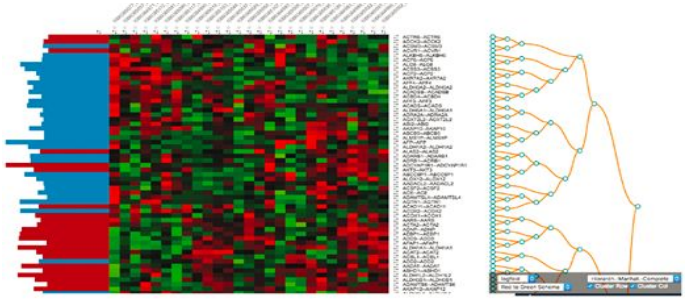
Venkata Satagopam ✉, Wei Gu, Serge Eifes, Piotr Gawron, Marek Ostaszewski, Stephan Gebel, Adriano Barbosa-Silva, Rudi Balling, and Reinhard Schneider

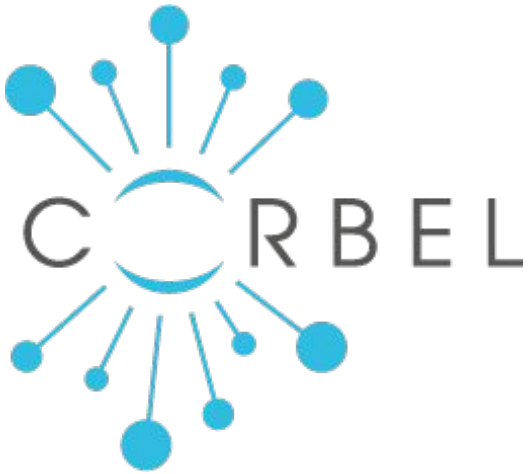


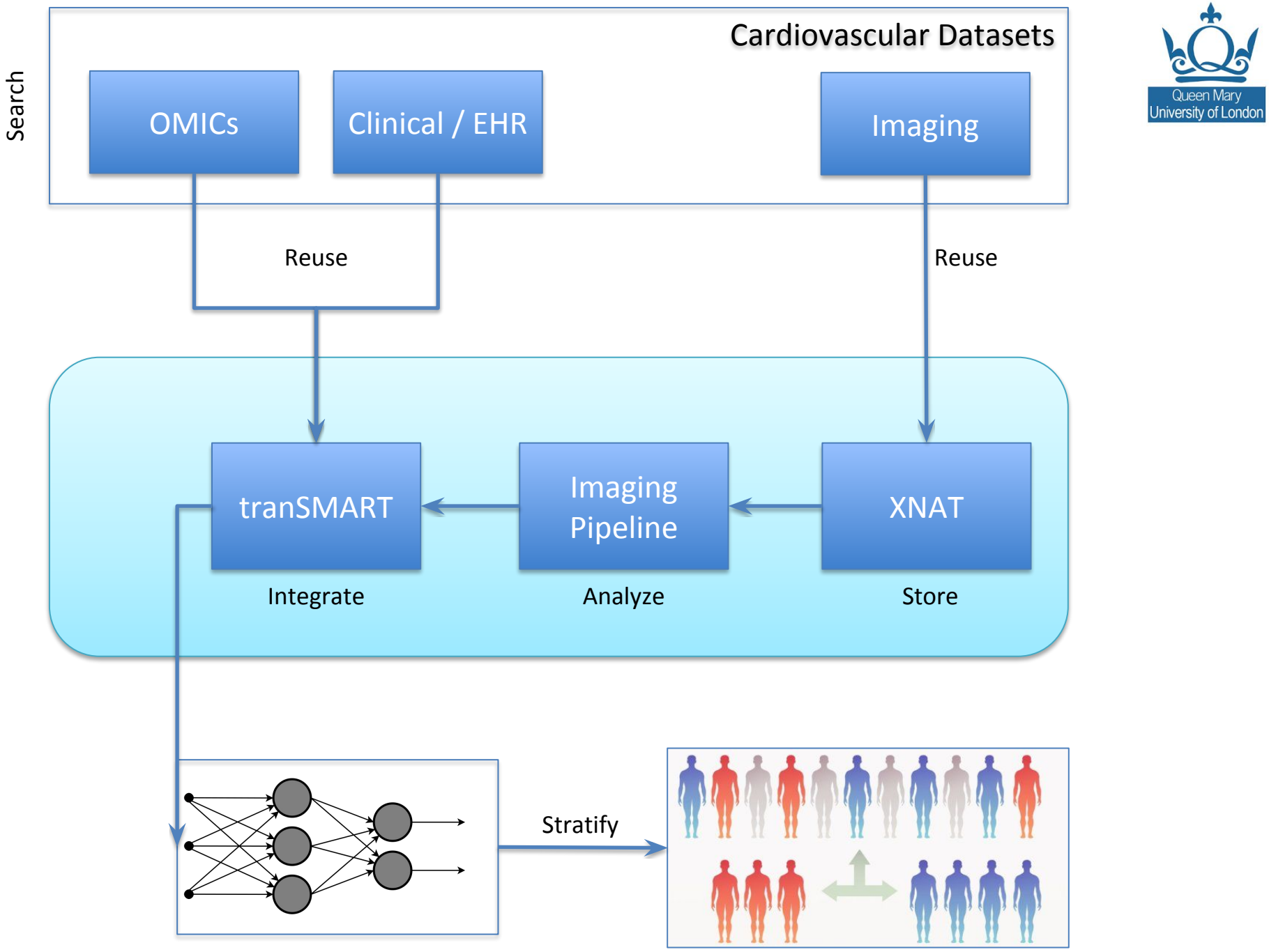
SmartR: an open-source platform for interactive visual analytics for translational research data

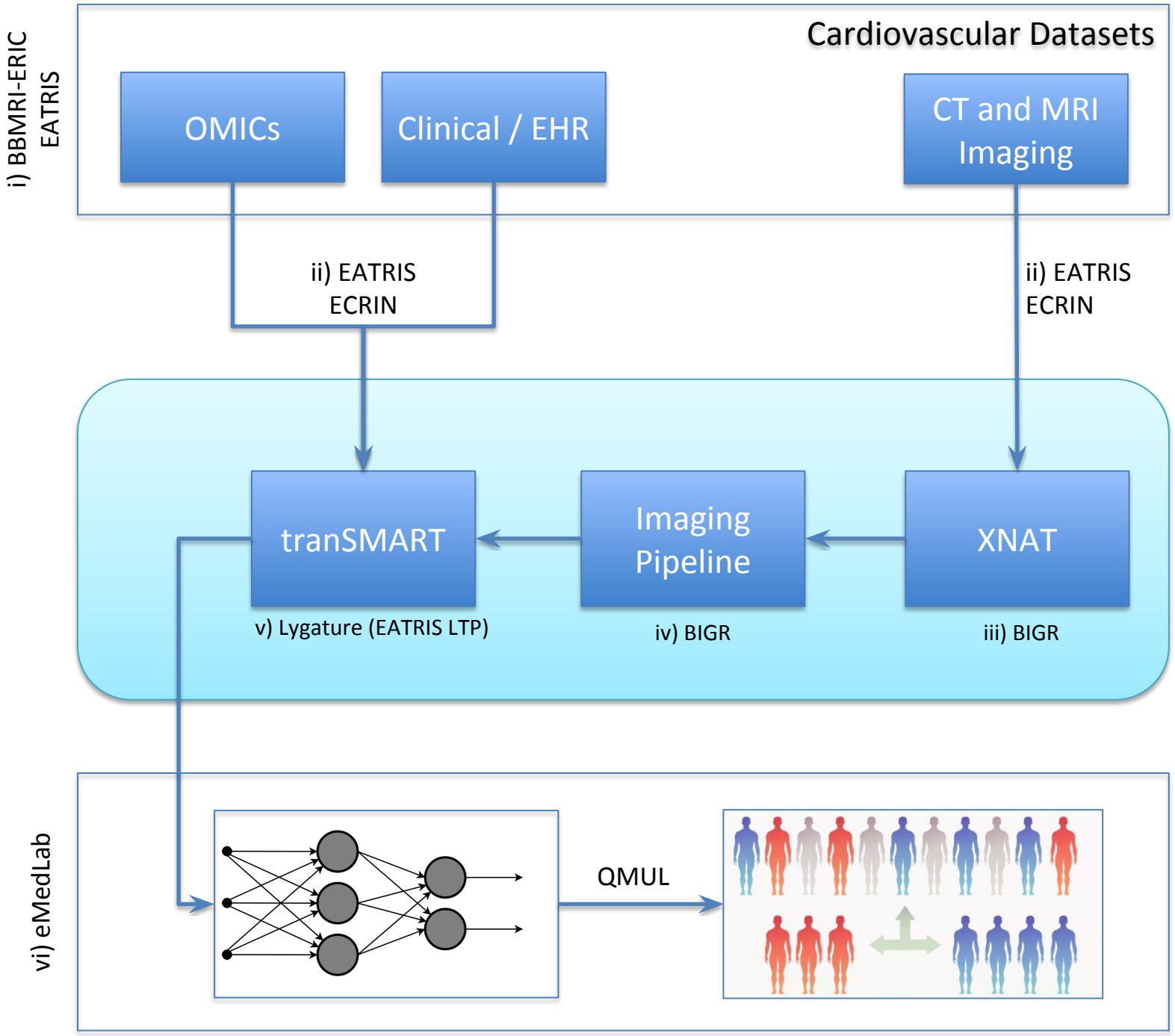
Sascha Herzinger, Wei Gu, Venkata Satagopam, Serge Eifes, Kavita Rege, Adriano Barbosa-Silva, Reinhard Schneider ✉, eTRIKS Consortium

Bioinformatics, Volume 33, Issue 14, 15 July 2017, Pages 2229–2231, https://doi.org/10.1093/bioinformatics/btx137
Published: 09 March 2017 **Article history** ▾



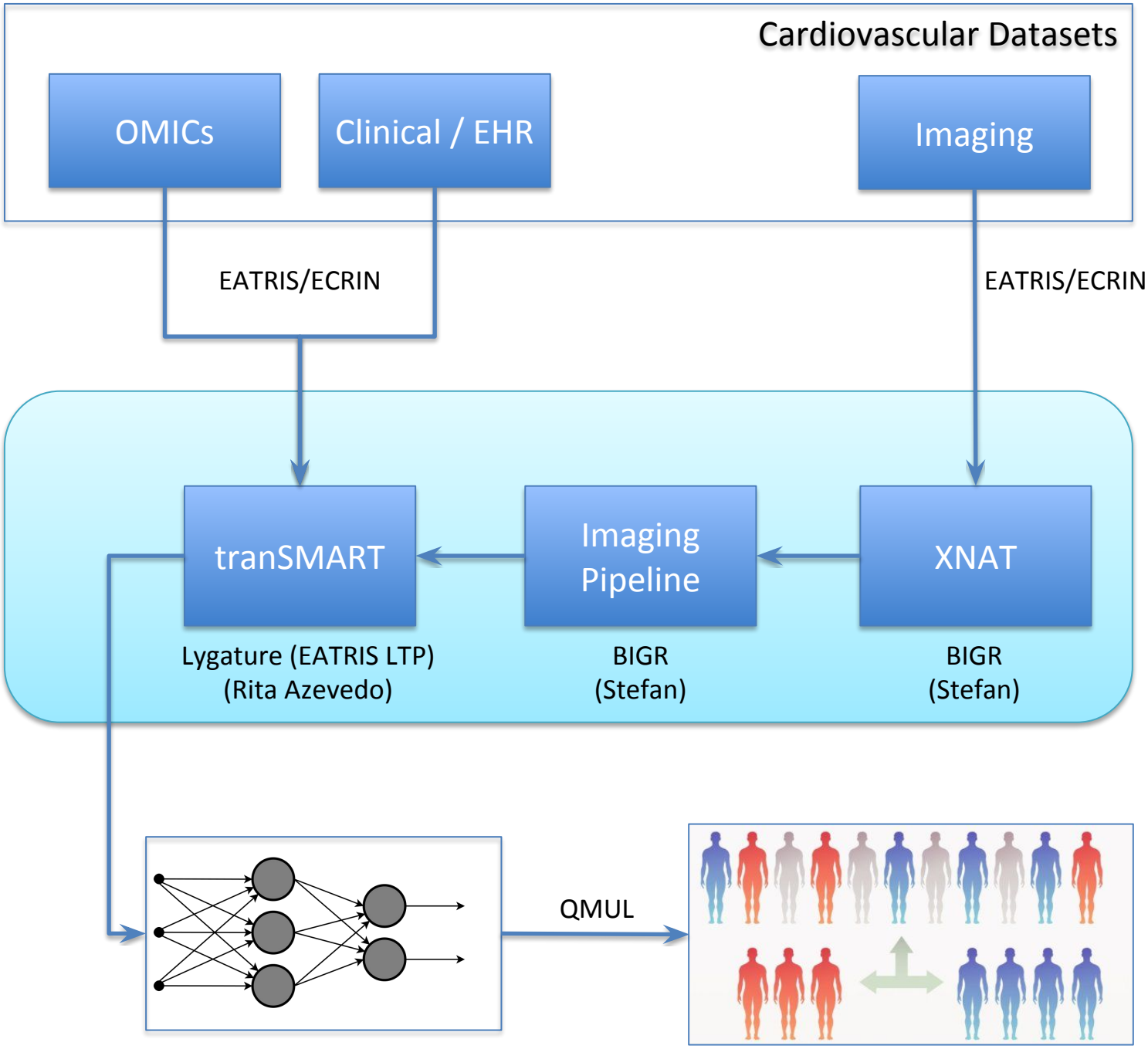

CORBEL


Search
Cardiovascular Datasets
OMICs
Clinical / EHR
Imaging
Reuse
Reuse
tranSMART
Imaging Pipeline
XNAT
Integrate
Analyze
Store
Stratify

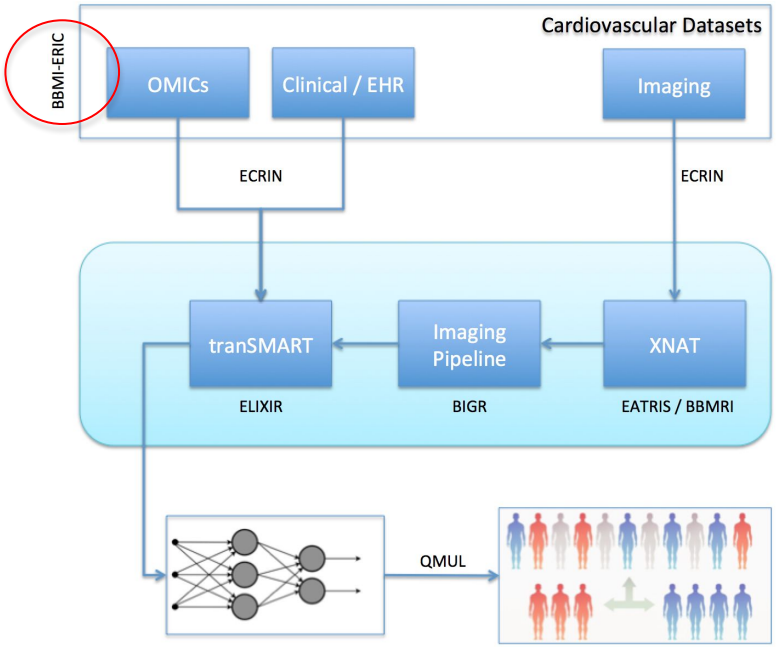

i) BBMRI-ERIC
EATRIS
Cardiovascular Datasets
OMICs
Clinical / EHR
CT and MRI Imaging
ii) EATRIS ECRIN
ii) EATRIS ECRIN
tranSMART
Imaging Pipeline
XNAT
v) Lygature (EATRIS LTP)
iv) BIGR
iii) BIGR
vi) eMedLab
QMUL

BBMRI-ERIC
(Petr Holub)
EATRIS (Florence Bietrix)

Queen Mary University of London



BBMRI-ERIC: Petr Holub
EATRIS: Florence Bietrix



Aim: identify EU datasets for cardiovascular research that could be used as input of the *data capture system.*

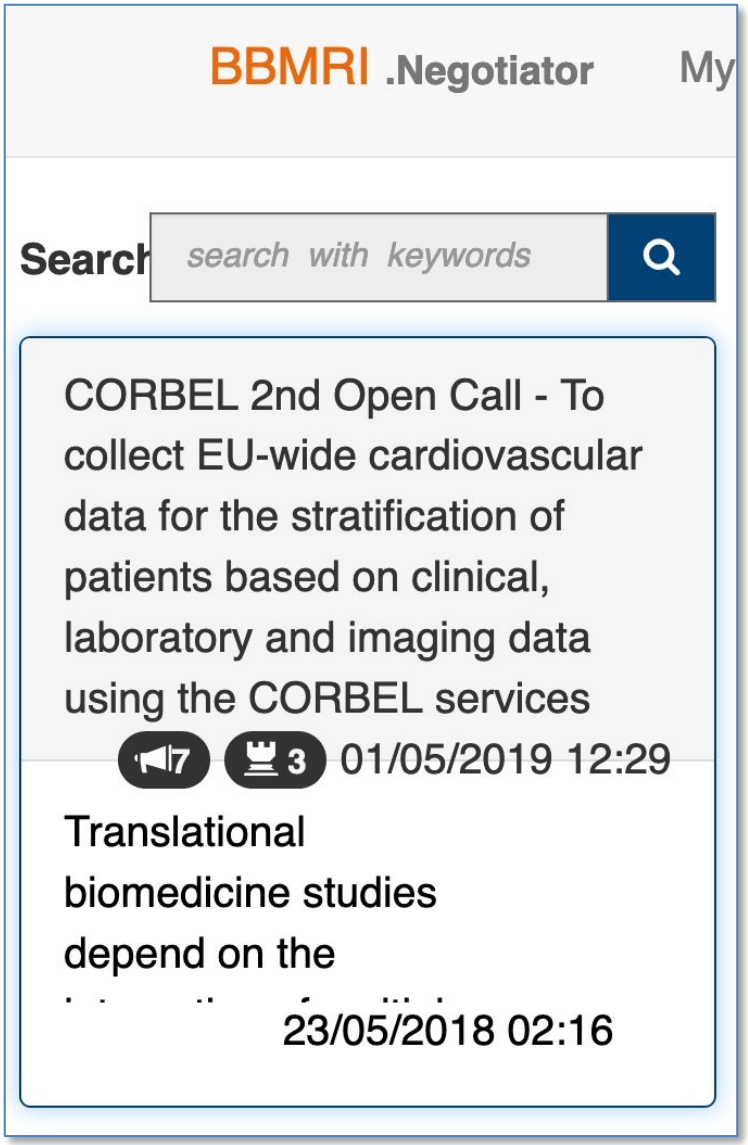
CORBEL Services:

- To identify biobanks that have relevant biological material, associated **data** and services
- To access biological samples and **associated data**

BBMRI-ERIC: Petr Holub

Actions

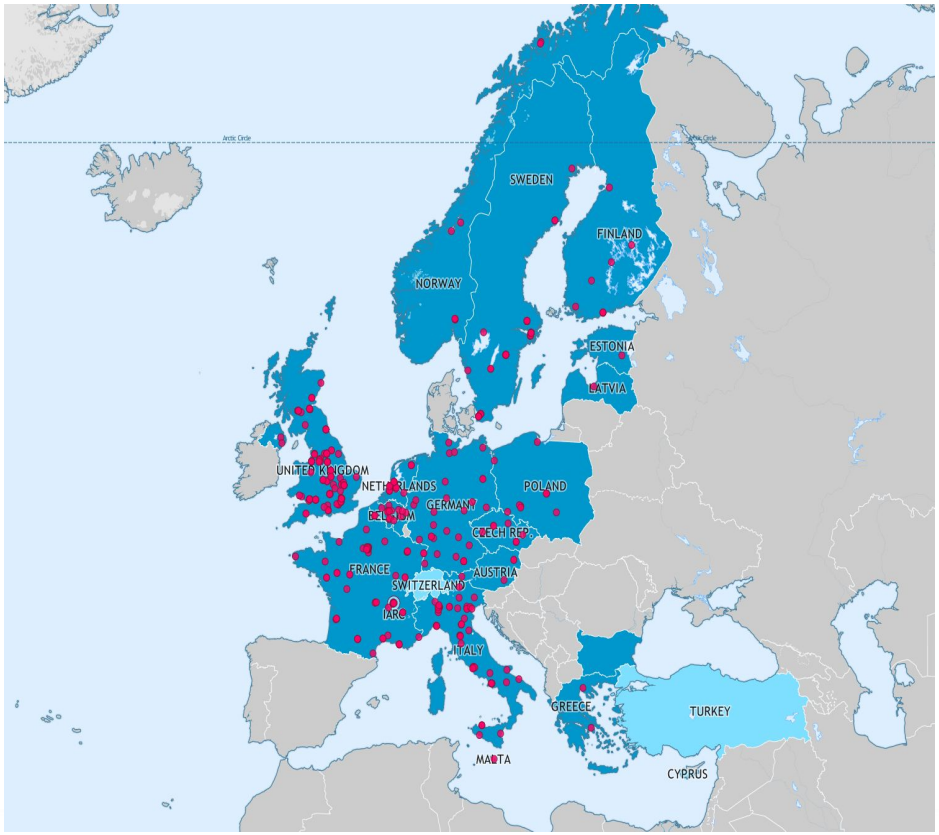
- Added QMUL as a registered institution;
- Helped to create a descriptive query for my request;
- Increased query limits from 100 to 500+ biobanks at once;



BBMRI-ERIC: Petr Holub

- Directory 5.0
- 702 biobanks
- 603 biobanks with collections
  - 1487 collections
- Estimated

  >100,000,000 samples and associated data



BBMRI-ERIC: Petr Holub

Results

- Query sent to the system

Request:

Identify EU-wide cardiovascular research/medical datasets including

- DICOM data (CT/MRI)
- Clinical data (Electronic Health Record)
- Omics data

BBMRI-ERIC: Petr Holub

Results

- 465 matches

BBMRI
Biobanking and Biomolecular Resources Research Infrastructure

Matching Biobanks: 465 (330 Collections reachable, 1150 Collections unreachable) **Show matching Biobanks**

- Medical University of Lodz Clinical Biobank
- Biobank Graz
- Biobank Innsbruck
- MedUni Wien Biobank
- VetBioBank
- Biobank @ UZ Brussel
- University Biobank Limburg
- Biobank@UZA
- Bimetra Biobank @ UZ Gent

EATRIS-ERIC

Search across EATRIS member institutions (90+)

BBMRI Biobanking and Biomolecular Resources Research Infrastructure

- 5 institutions identified:

Matchmaking Report

Client: Corbel Open Call - Adriano Barbosa da Silva
Project Title: PID5815 - Integrating EU-wide cardiovascular research datasets
EATRIS Project ID: 20181017-1400

EATRIS:

Results

- 5 Negotiations started:
  - Instituto Ramón y Cajal de Investigación Sanitaria (IRYCIS, Spain)
  - University Hospital Bellvitge (IDIBELL, Spain)
  - Clinical Research Centre of Tartu University and Tartu University Clinics (Estonia)
  - Fondazione Ri.MED (Italy)
  - St. Anne´s University Hospital in Brno (Czech Republic)

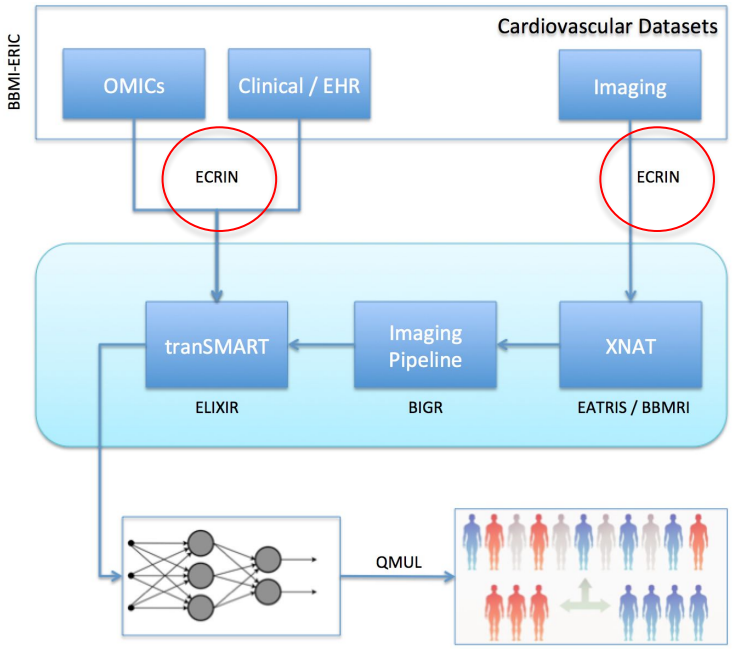
eatris

European infrastructure for translational medicine

EATRIS:

CORBEL Services:

- Regulatory and ethical context
- Data sharing/reuse (prospective and **retrospective data**)



Actions:

Contacted the relevant partners at the hospitals with compatible data;

eatris
European infrastructure for translational medicine

EATRIS: Florence Bietrix

Results

- 2 Negotiations followed:

  - **Fondazione Ri.MED (Italy);**

    - **Collaboration possible** via **ISMETT** (Mediterranean Institute for Transplantation and Advanced Specialized Therapies);
    - Needs a **Data Processing Agreement**.

eatris
European infrastructure for translational medicine

EATRIS: Florence Bietrix

Results

- 2 Negotiations followed:

  - **Fondazione Ri.MED (Italy);**

DPO, JMRO

Data Processing Agreement (DPA)

by and between

ISTITUTO MEDITERRANEO PER I TRAPIANTI E TERAPIE AD ALTA SPECIALIZZAZIONE S.r.l., with registered office in Palermo (Italy), at Via Discesa dei Giudici, 4 (hereinafter referred to as "**ISMETT**")

and

UPMC ITALY S.r.l., with registered office in Palermo (Italy), at Via Discesa dei Giudici, 4 (hereinafter referred to as "**UPMCI**")

- Jointly referred to as the "**Joint-Controllers**"

and

.....

Queen Mary University of London of Mile End Road, London, E1 4NS, United Kingdom

.....

- hereinafter referred to as "**QMUL**"

EATRIS: Florence Bietrix

Results

ISMETT Istituto di Ricovero e Cura a Carattere Scientifico

37 e-mails,

over

1 year

EATRIS: Florence Bietrix

Results

- 2 Negotiations followed:

FNUSA ICRC
ST. ANNE'S UNIVERSITY HOSPITAL BRNO
INTERNATIONAL CLINICAL RESEARCH CENTER

- **St. Anne´s University Hospital in Brno (Czech Republic)**
  - **Collaboration possible**
  - We want to use UKBIOBANK mapping as a way to aggregate value to the hospital's data;
  - Needs a **MoU** for **Collaboration Agreement** between the universities prior to deliver data;

EATRIS: Florence Bietrix

Results

- 2 Negotiations followed:

  - **St. Anne´s University Hospital in Brno (Czech Republic)**

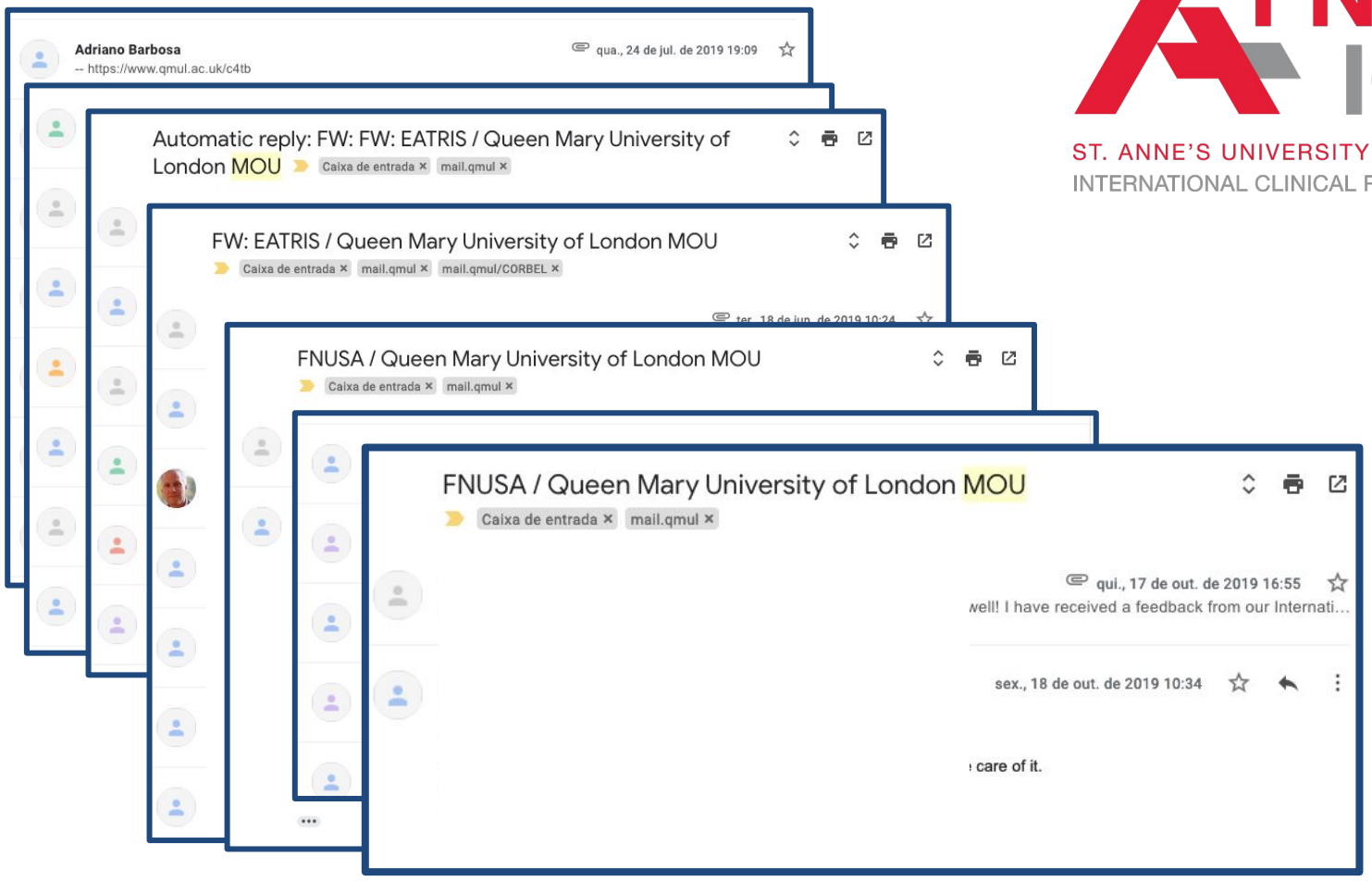
insert the logo of the partner institution here xxxxxxxxx

Číslo dohody FNUSA: Číslo dohody Partnera:	FNUSA´s agreement number: Partner´s agreement number:
DOHODA O SPOLUPRÁCI uzavřená dle § 1746, zákona č. 89/2012 Sb., občanský zákoník, ve znění pozdějších předpisů (dále jen „občanský zákoník“)	**COLLABORATION AGREEMENT** concluded under Article 1746, of Act No. 89/2012 Coll., the Civil Code as amended (hereinafter referred to as “Civil Code”)

International Cooperation Office

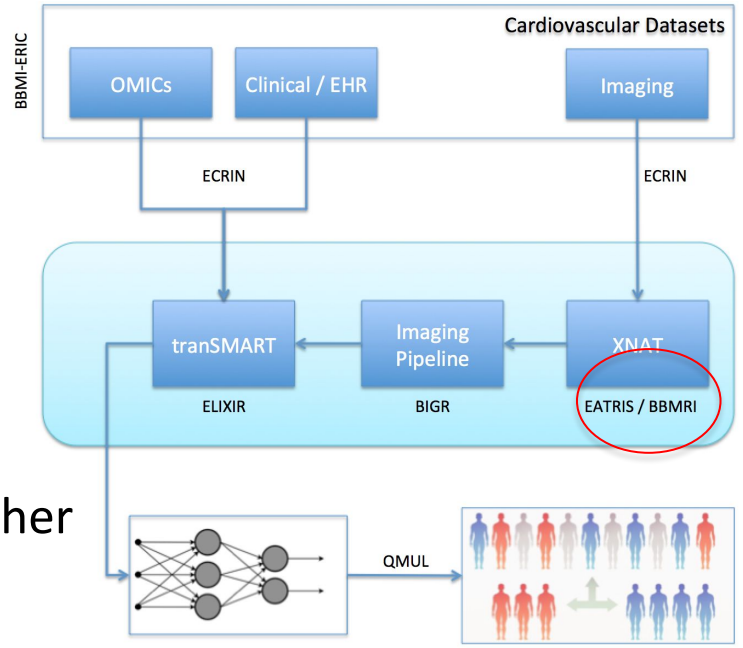
EATRIS: Florence Bietrix

Results



FNUSA ICRC
ST. ANNE'S UNIVERSITY HOSPITAL BRNO
INTERNATIONAL CLINICAL RESEARCH CENTER

EurobioImaging (EPI2)



Stefan Klein

Aim: Promote the integration of the imaging datasets within the data capture system

CORBEL Service:

- Deploy XNAT.bmia.nl imaging archive (together with EATRIS/BBMRI)

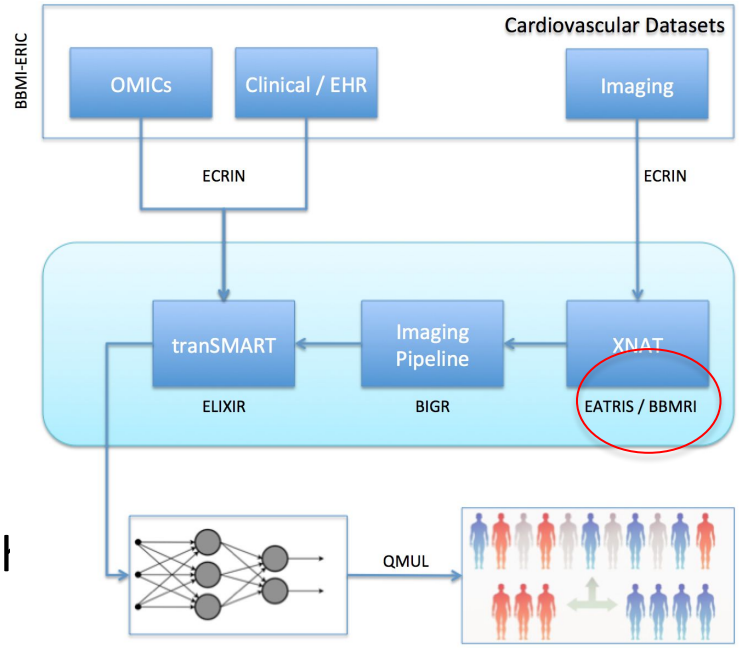
EurobioImaging (EPI2)

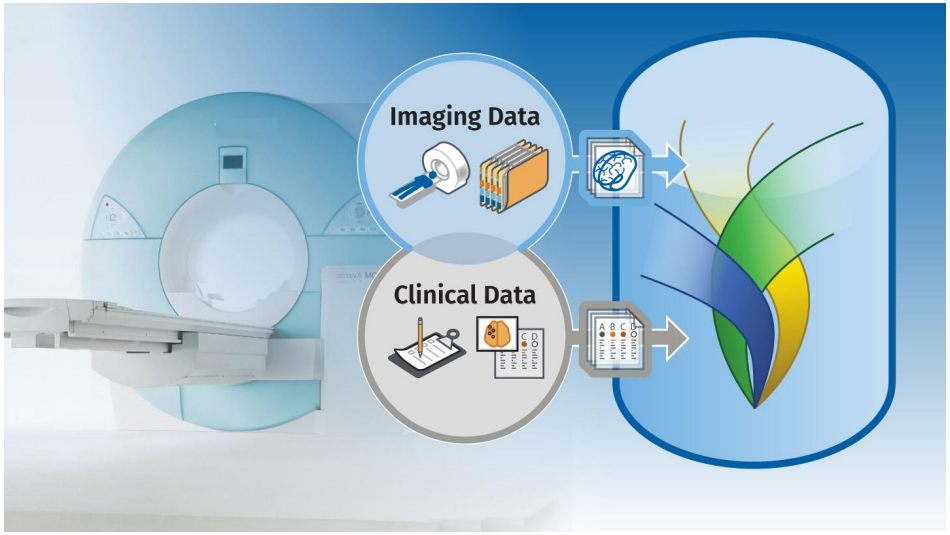
Stefan Klein

Aim: Promote the integration of the imaging datasets within the data capture system

CORBEL Service:

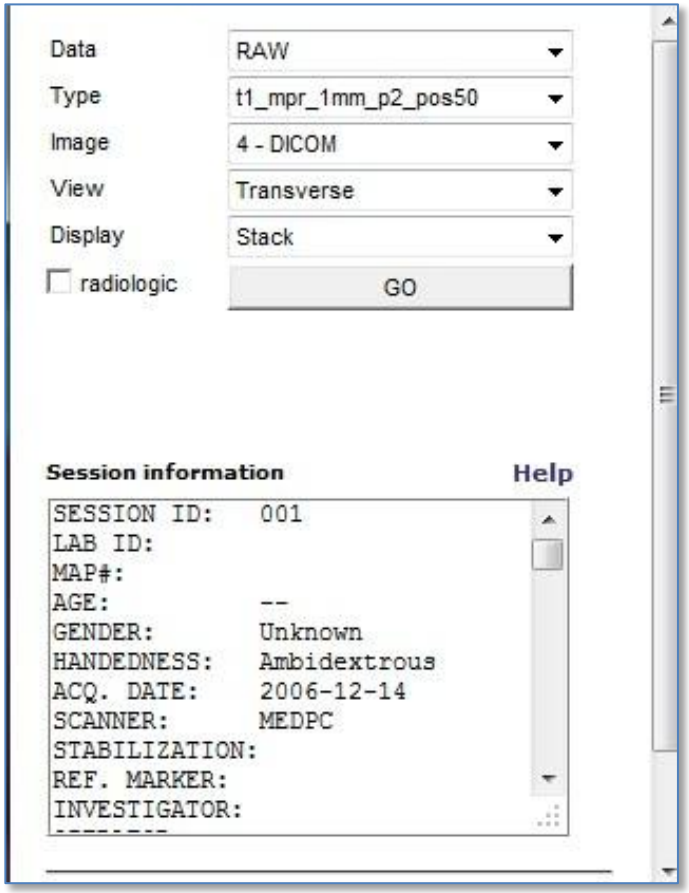
- Deploy XNAT.bmia.nl imaging archive (togetl EATRIS/BBMRI)

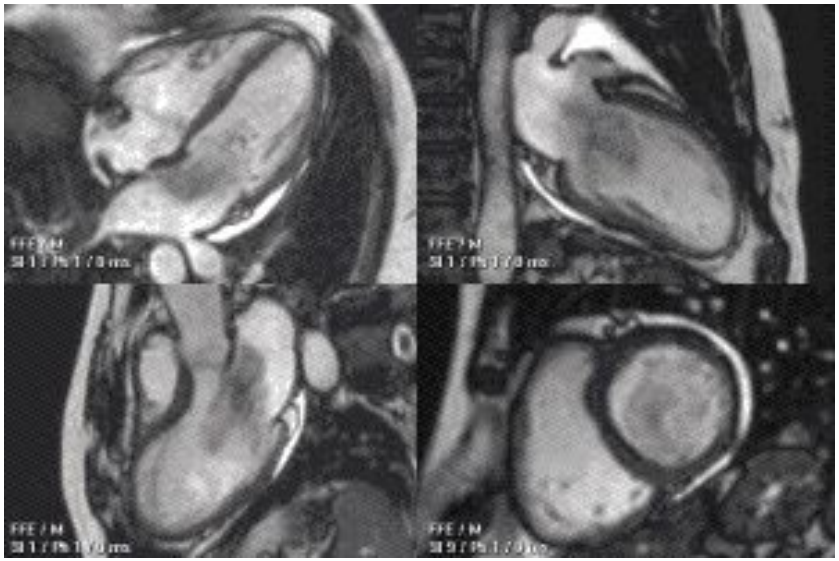




EurobioImaging (EPI2)

Stefan Klein

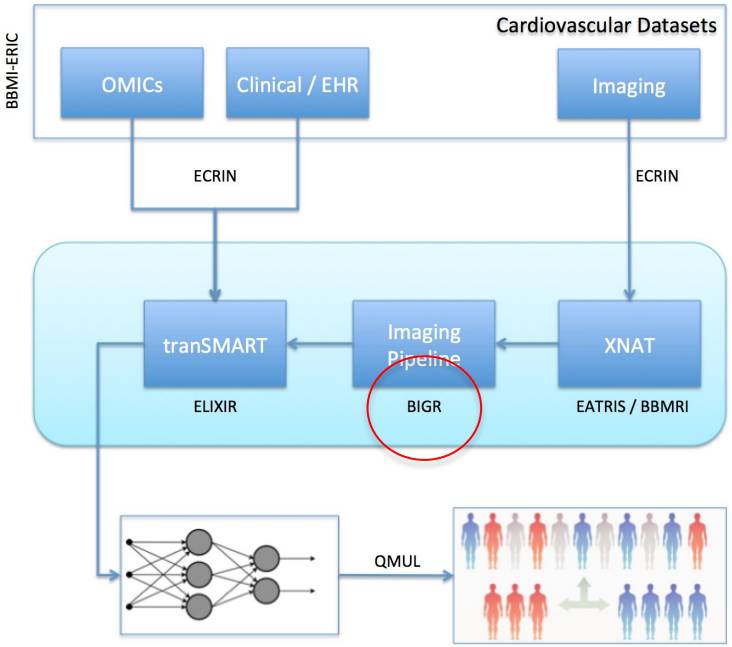




XNAT

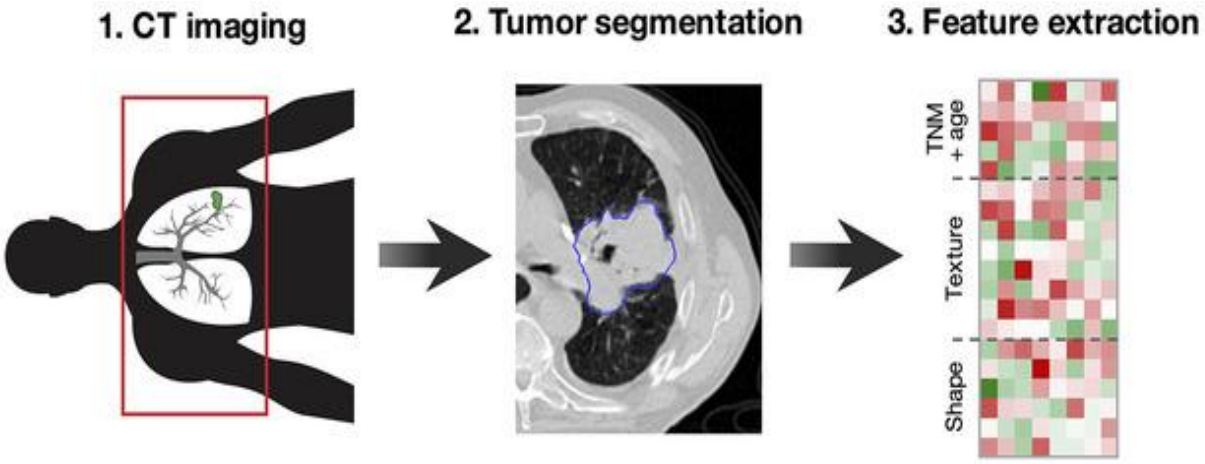
EurobioImaging (EPI2)

Stefan Klein



Future actions

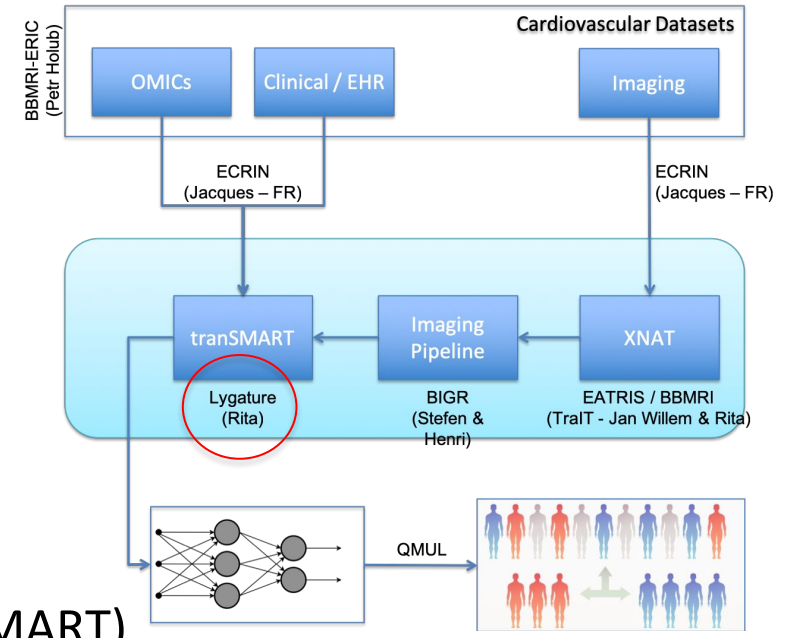
Collaborate with the Radiomics pipeline to extract imaging features from the raw data



DOI: 10.18632/oncotarget.22251

Lygature (EATRIS):

Rita Azevedo



Aim: Extract biomarkers from the acquired images

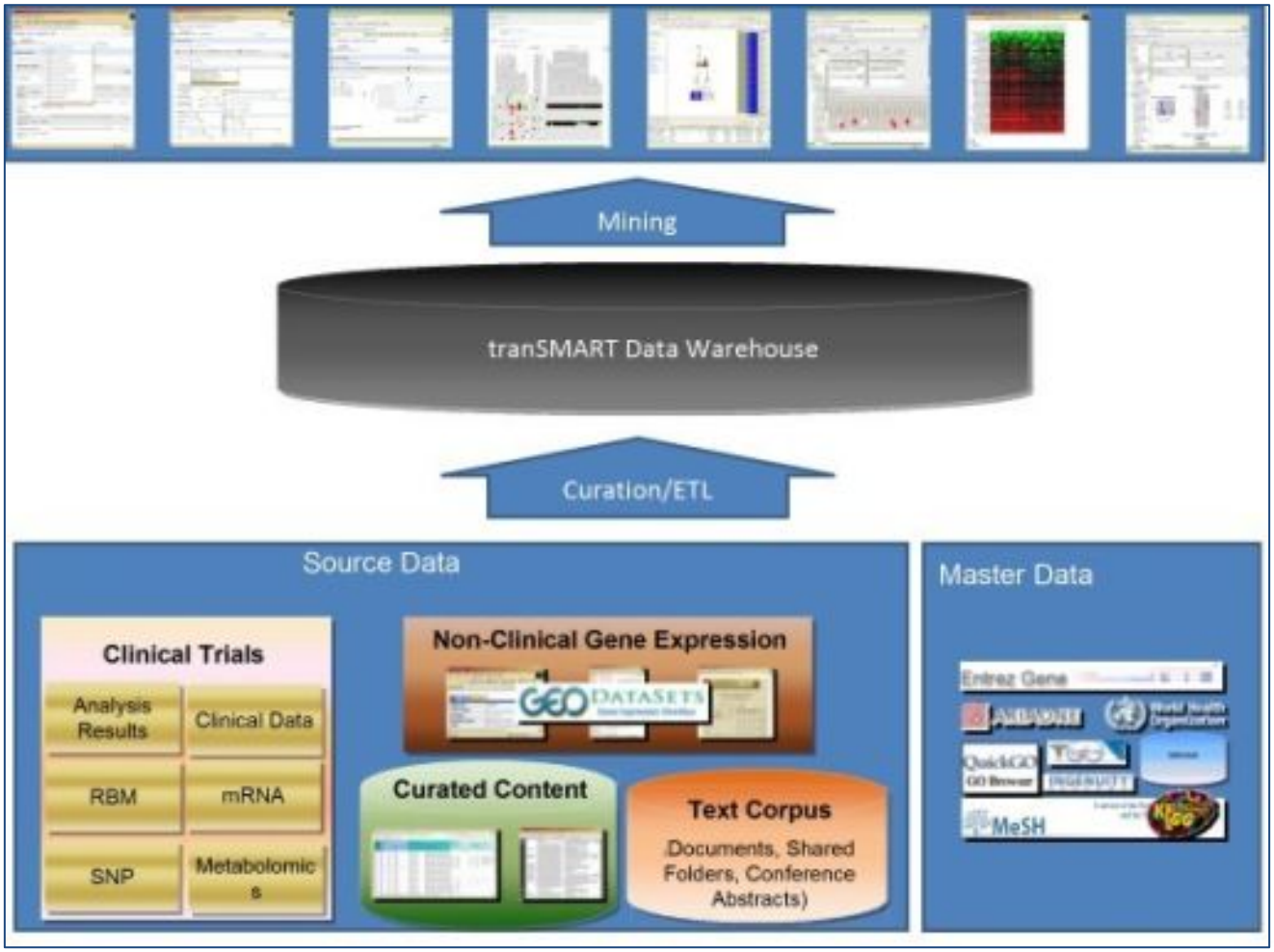
CORBEL Service:

a) Provide services and tools to deposit store and access their data (tranSMART)

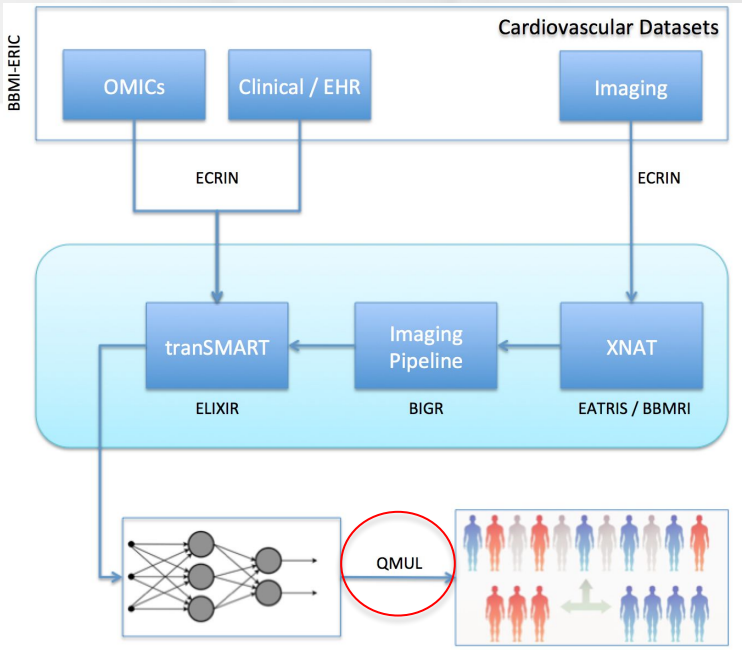
tranSMART

Lygature (EATRIS):

Rita Azevedo



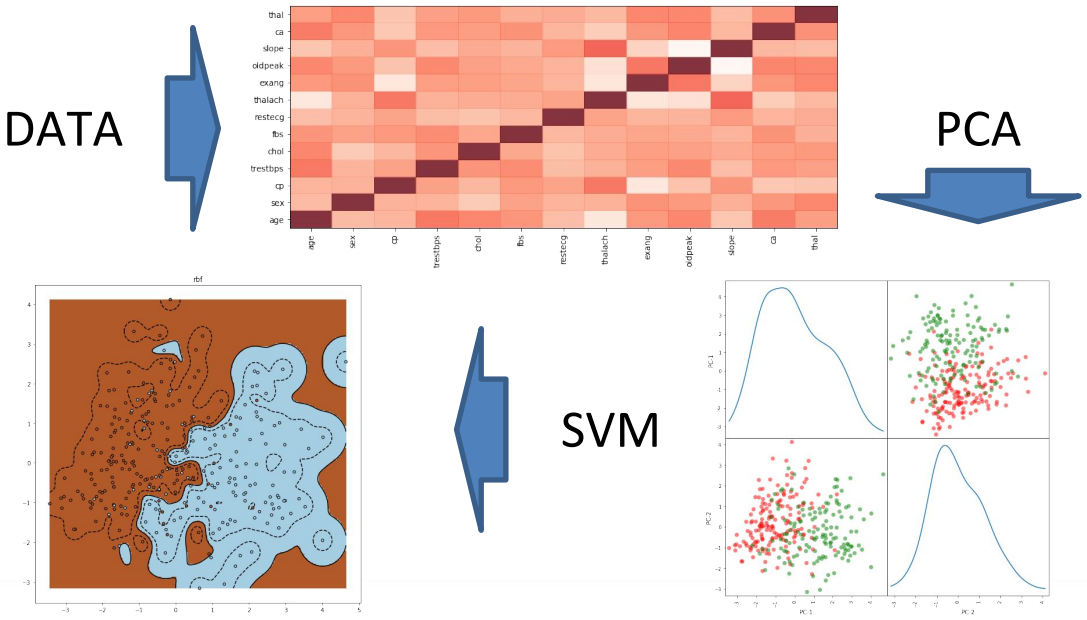
QMUL



Aim: Apply machine learning methods to stratify cardiovascular patients based on machine learning

Methods:

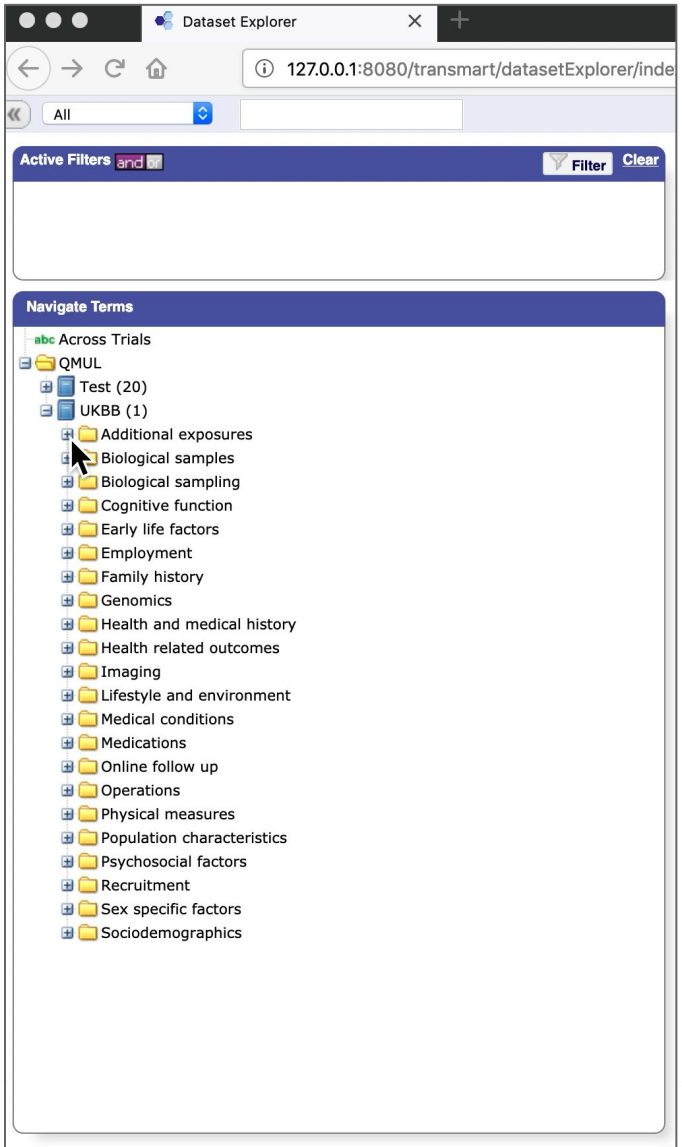
- Dimensionality reduction: PCA
- Machine learning classifiers:
  a) Supervised
  b) Un-supervised
- Tree based-methods:
  a) Random forests
  b) Gradient boost
- Deep learning



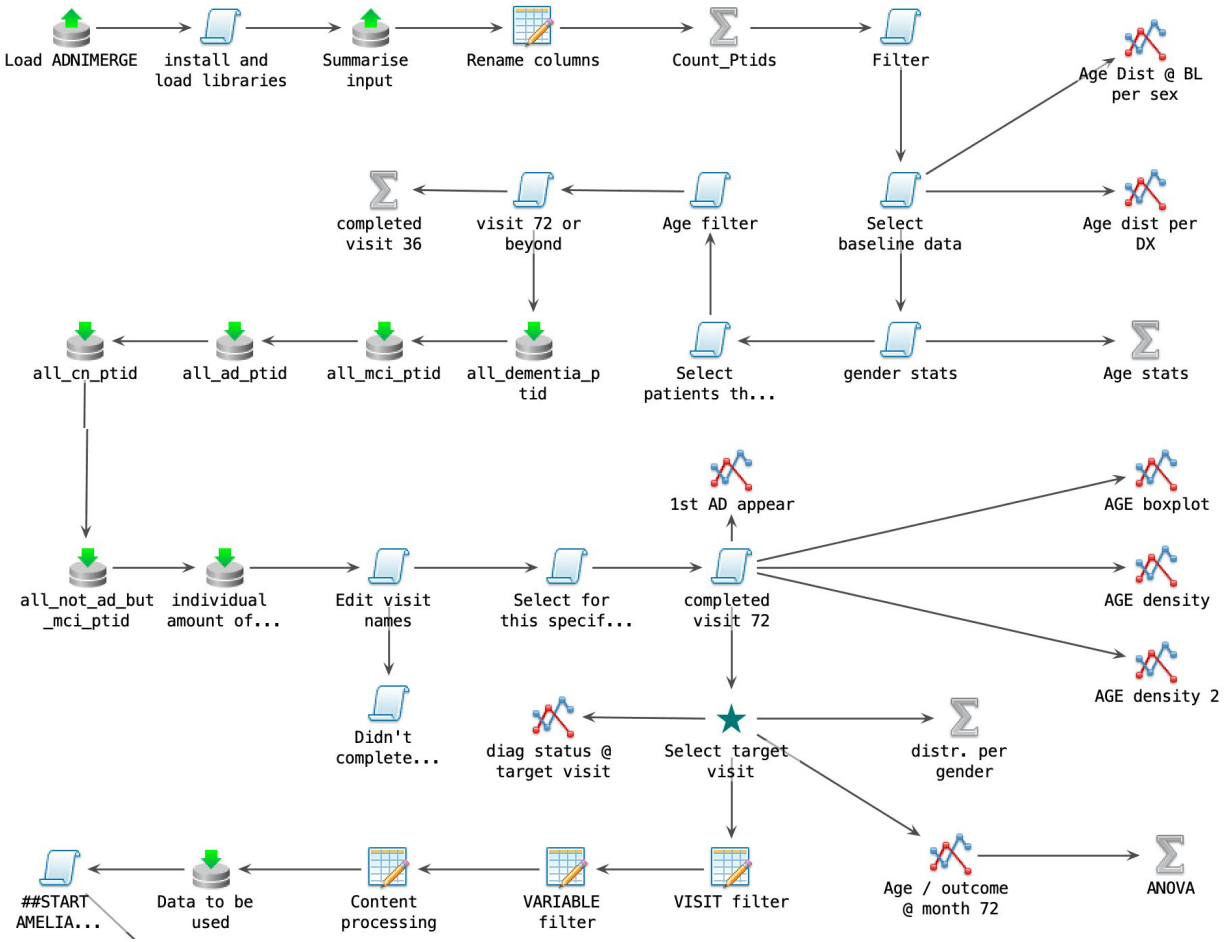
- i2b2 categorization of UKBB data
  - 11,870 variables

A	B	C	D	E
file_name	category_cd	col_n	data_lab	data_label_source
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Body_size_measures+Weight_pre_imaging	1706		bio.weight..pre.imaging..0.imaging
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Body_size_measures+Weight_method	3475		bio.weight.method.0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Body_size_measures+Weight_method	3476		bio.weight.method.0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Body_size_measures+Weight_method	3477		bio.weight.method.0.imaging
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Whole_body_water_mass	4656		bio.whole.body.water.mass.0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Whole_body_water_mass	4657		bio.whole.body.water.mass.0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_of_arm	4671		bio.impedance.of.arm..right..0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_of_arm	4672		bio.impedance.of.arm..right..0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_of_arm	4673		bio.impedance.of.arm..right..0.imaging
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_of_arm	4674		bio.impedance.of.arm..left..0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_of_arm	4675		bio.impedance.of.arm..left..0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_of_arm	4676		bio.impedance.of.arm..left..0.imaging
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Arm_fat_percentage	4693		bio.arm.fat.percentage..right..0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Arm_fat_percentage	4694		bio.arm.fat.percentage..right..0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Arm_fat_percentage	4701		bio.arm.fat.percentage..left..0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Arm_fat_percentage	4702		bio.arm.fat.percentage..left..0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Spirometry+Smoked_cigarette_or_pipe_within_last_hour	6255		bio.smoked.cigarette.or.pipe.within.last.hour.0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Spirometry+Smoked_cigarette_or_pipe_within_last_hour	6256		bio.smoked.cigarette.or.pipe.within.last.hour.0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Spirometry+Smoked_cigarette_or_pipe_within_last_hour	6257		bio.smoked.cigarette.or.pipe.within.last.hour.0.imaging
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Arterial_stiffness+Arterial_pulse_wave_stiffness_device_ID	7209		bio.arterial.pulse.wave.stiffness.device.id.0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Arterial_stiffness+Arterial_pulse_wave_stiffness_device_ID	7210		bio.arterial.pulse.wave.stiffness.device.id.0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Arterial_stiffness+Arterial_pulse_wave_stiffness_device_ID	7211		bio.arterial.pulse.wave.stiffness.device.id.0.imaging
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_device_ID	8390		bio.impedance.device.id.0.baseline
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_device_ID	8391		bio.impedance.device.id.0.calibration
biobank_decoded_decoded_r8a.list.csv	Physical_measures+Anthropometry+Impedance_measures+Impedance_device_ID	8392		bio.impedance.device.id.0.imaging
biobank_decoded_decoded_r8a.list.csv	Psychosocial_factors+Mental_health+Ever_highly_irritable_argumentative_for_2_days	8456		bio.ever.highly.irritable.argumentative.for.2.days.0.baseline
biobank_decoded_decoded_r8a.list.csv	Psychosocial_factors+Mental_health+Ever_highly_irritable_argumentative_for_2_days	8457		bio.ever.highly.irritable.argumentative.for.2.days.0.calibration
biobank_decoded_decoded_r8a.list.csv	Psychosocial_factors+Mental_health+Ever_highly_irritable_argumentative_for_2_days	8458		bio.ever.highly.irritable.argumentative.for.2.days.0.imaging

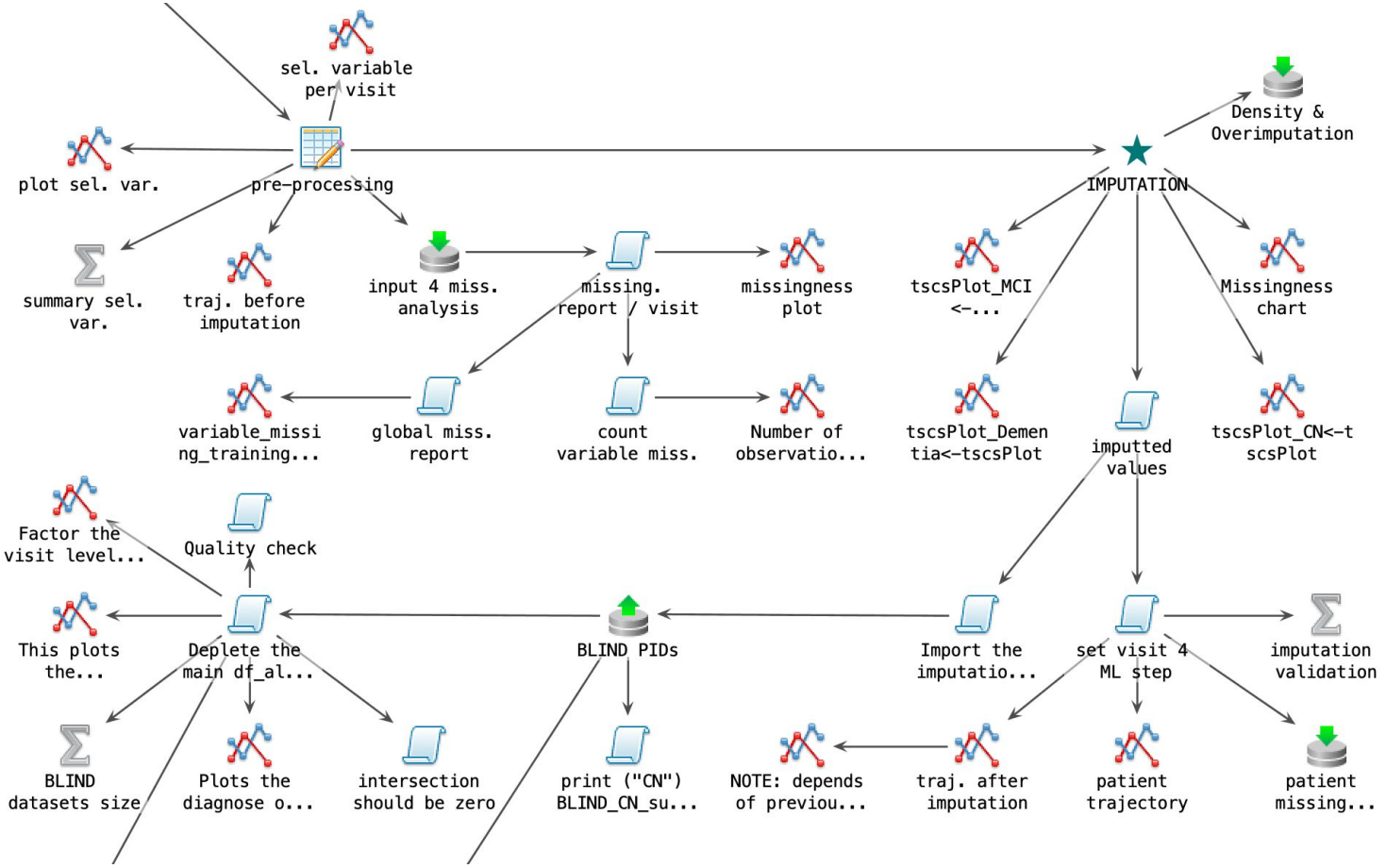
- i2b2 categorization of UKBB data
  - 11,870 variables



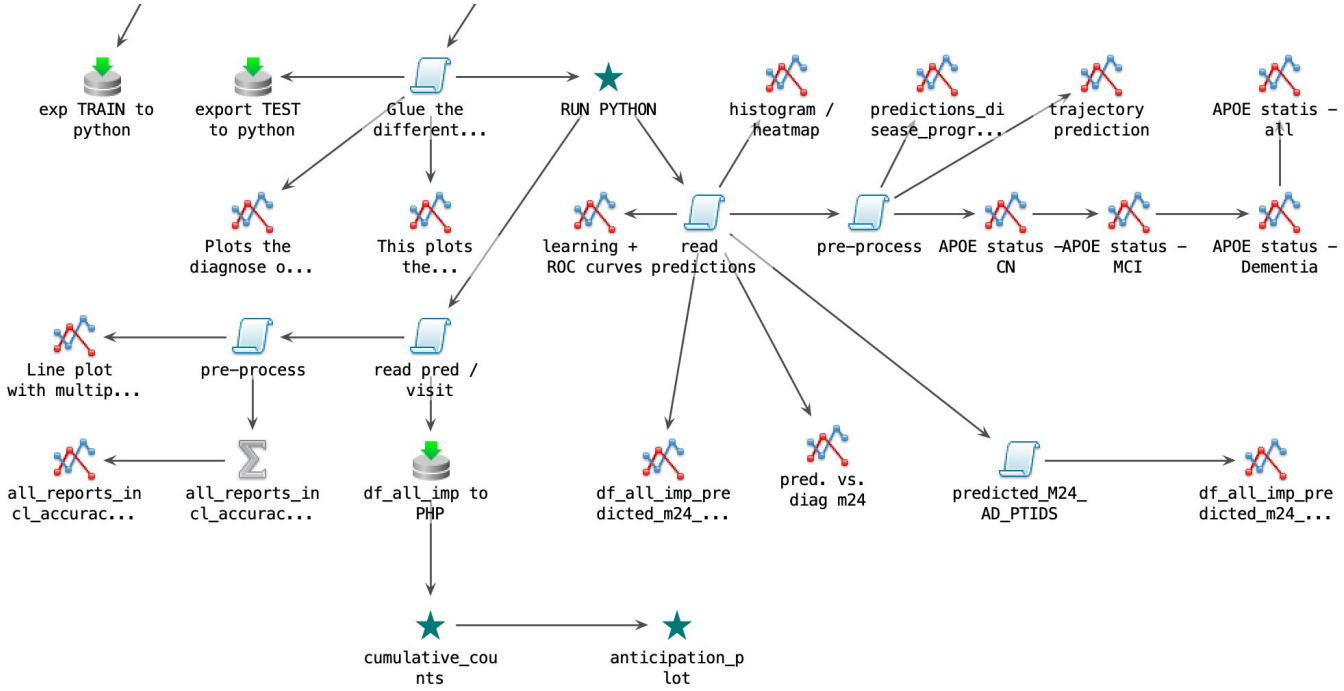
- Modelling the analytic environment for Machine Learning while waiting.



- Modelling the analytic environment for Machine Learning while waiting.

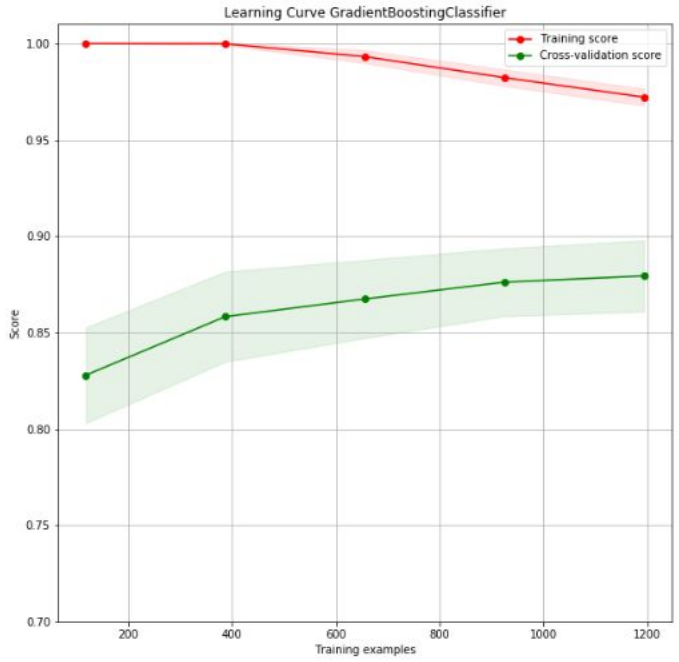


- Modelling the analytic environment for Machine Learning while waiting.

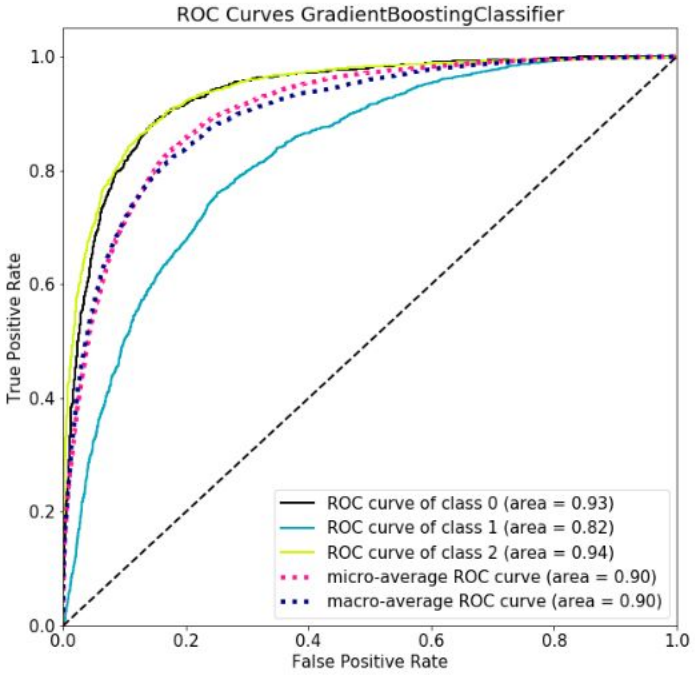


- Modelling the analytic environment for Machine Learning while waiting.

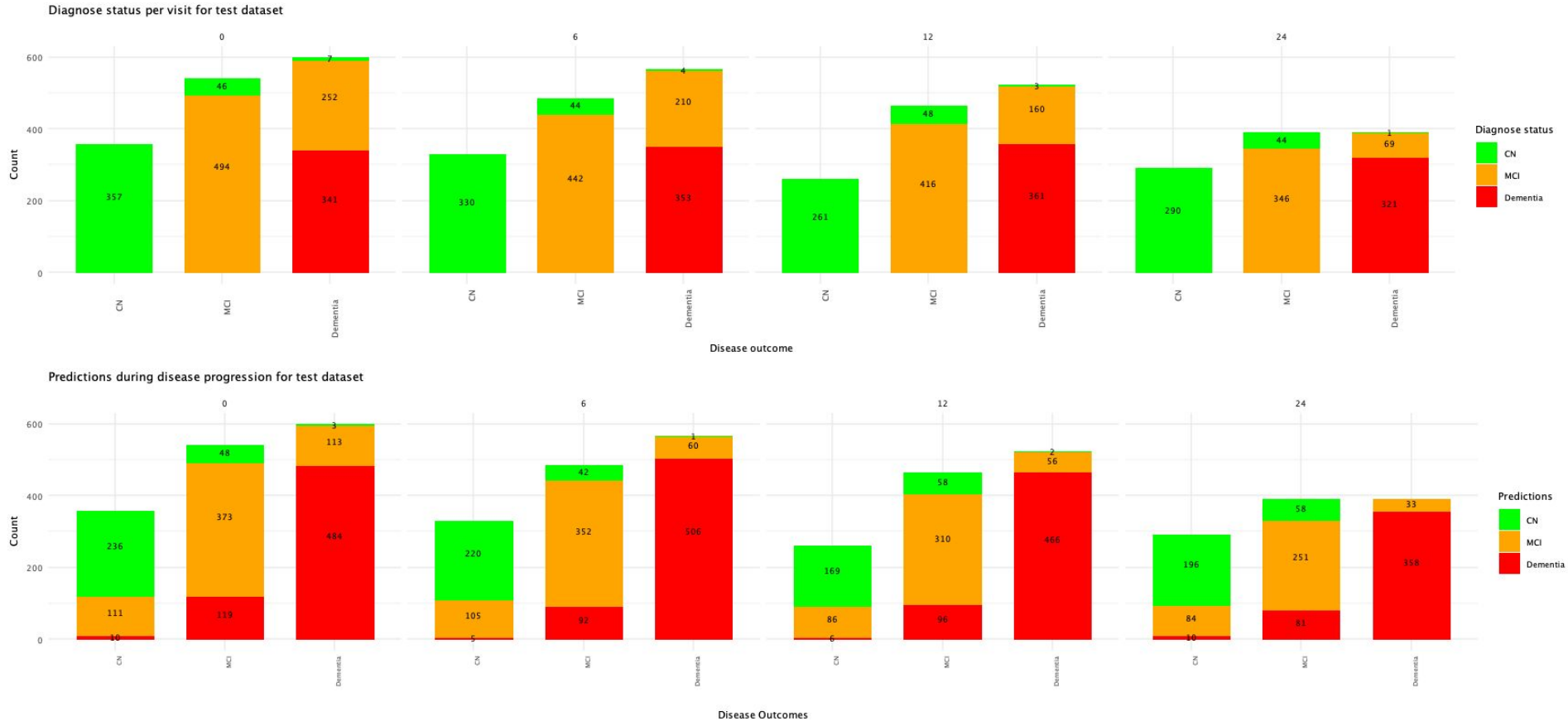
a



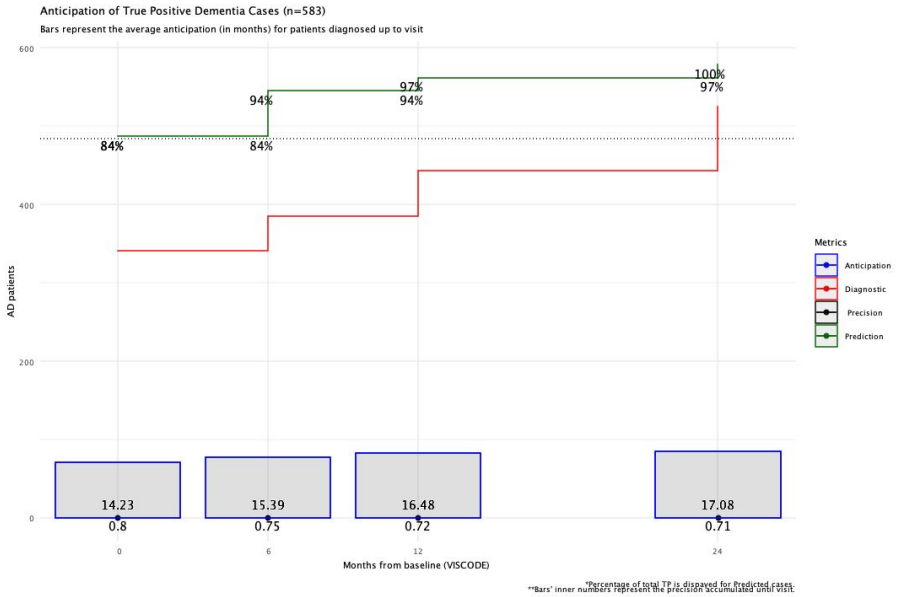
b

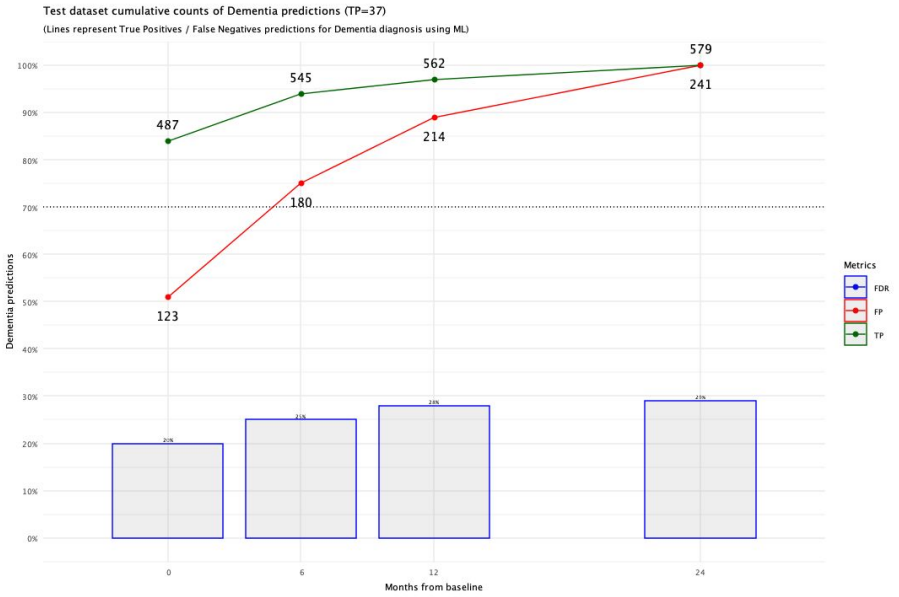


- Modelling the analytic environment for Machine Learning while waiting.



- Modelling the analytic environment for Machine Learning while waiting.





Overall evaluation and feedback

What did we appreciate?

Quick V.I.P. treatment and support.

What went wrong? Was there anything missing?

Need to fund a position to extract data from the hospitals.

How could one improve in future projects?

Be very specific on what you need.

Beyond CORBEL: what is needed to maintain common access provision across RIs?

Continued funding.

Summary

- We used R.I. offered by CORBEL to identify 2 potentially used Cardiovascular Datasets (**Petr Holub**);
- We're setting up the legal framework to re-use this data (**Florence Bietrix**);
- We received training at the EPI2-EMC to use XNAT as platform for imaging integration (**Stefan Klein**);
- We received consulting from Lygature to compare our modelling to their i2b2-tranSMART mapping (**Rita Azevedo**) ;
- We are developing data structures using UKB for tranSMART, and ML modules with ADNI data while waiting.

CORBEL

Queen Mary University of London

European Union

EUandME

Language EN change

What matters most to you? From beachcombing to backpacking, surfing the internet to saving the world, living in the European Union provides endless opportunities to do what you love – wherever that may lead. As a young citizen, the EU can give you all the tools you need to unlock your potential and ignite your inspiration. Whether you're an energetic entrepreneur, a digital devotee or a protector of the planet, it all comes down to #EUandME.

Get inspired by our #EUandME short films directed by renowned European filmmakers and learn more about the link between your pastime and the EU.

Acknowledgements

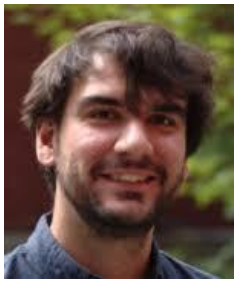
Mike Barnes

Katriona Goldmann

Chris John

Sandra Ng

David Watson



Claudia Cabrera &
Evan Tzanis

Barts and The London School of Medicine and Dentistry
William Harvey Research Institute

MRC Medical Research Council eMedLab

HDRUK Health Data Research UK

NHS Barts Health NHS Trust

UK Research and Innovation

Horizon 2020 European Union Funding for Research & Innovation

biobank uk
Improving the health of future generations

THANK YOU!

Obrigado!
Merci!
Danke je wel!
Danke schön!
Děkuju!

- BBMRI-ERIC — biobanks
- ELIXIR — curated databases
- EMBRC — marine model organisms
- ISBE — systems biology
- EMPHASIS — Plant phenotyping centres
- eatris — translational research
- INFRAFRONTIER — mouse disease models
- EU-OPENSCREEN — Chemical Keys for Life's Locks — screening & medicinal chemistry
- MIRRI — microorganisms
- ECRIN — clinical trials
- Instruct Integrating Biology — structural biology
- EURO-BIOIMAGING — biological/medical imaging
- erinha — European Research Infrastructure on Highly Pathogenic Agents — Highly Pathogenic Agents

Follow us: @qmul_c4tb