

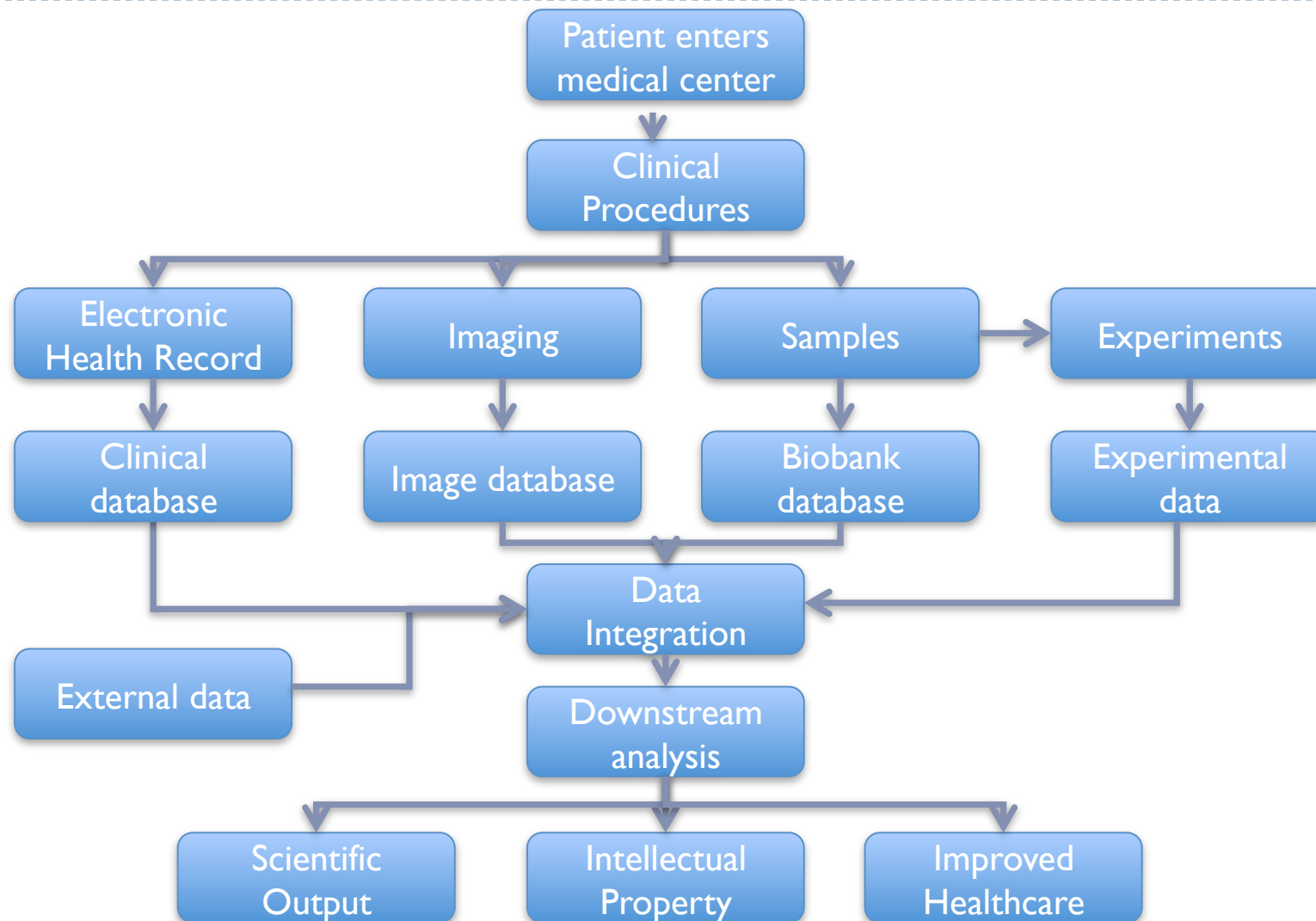


Imperial College
London Data Science Institute

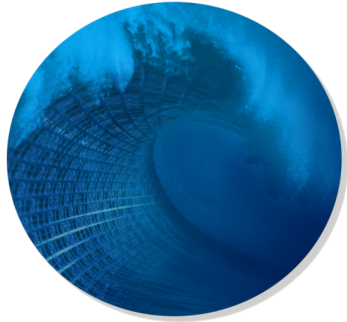
The tranSMART open source platform and the IMI-eTRIKS project

Ioannis Pandis, PhD \ Life Science Research Manager, Data Science Institute
- Imperial College London

Translational Research Project Process



Data Challenges...: Not Just Analysis



Tsunami of genomic and clinical data



Consortia underestimate Data Management Effort



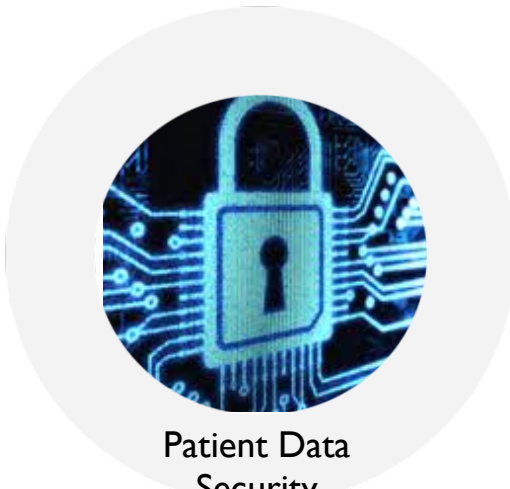
Multiple Partners; Fragmented Platforms



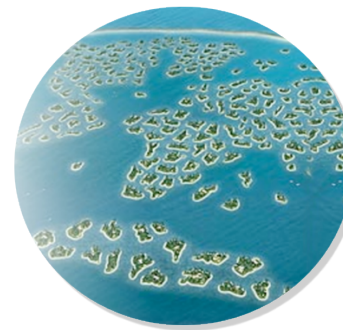
IT Research £ >>
IT Service £



Weak standards
Poor interoperability



Patient Data
Security

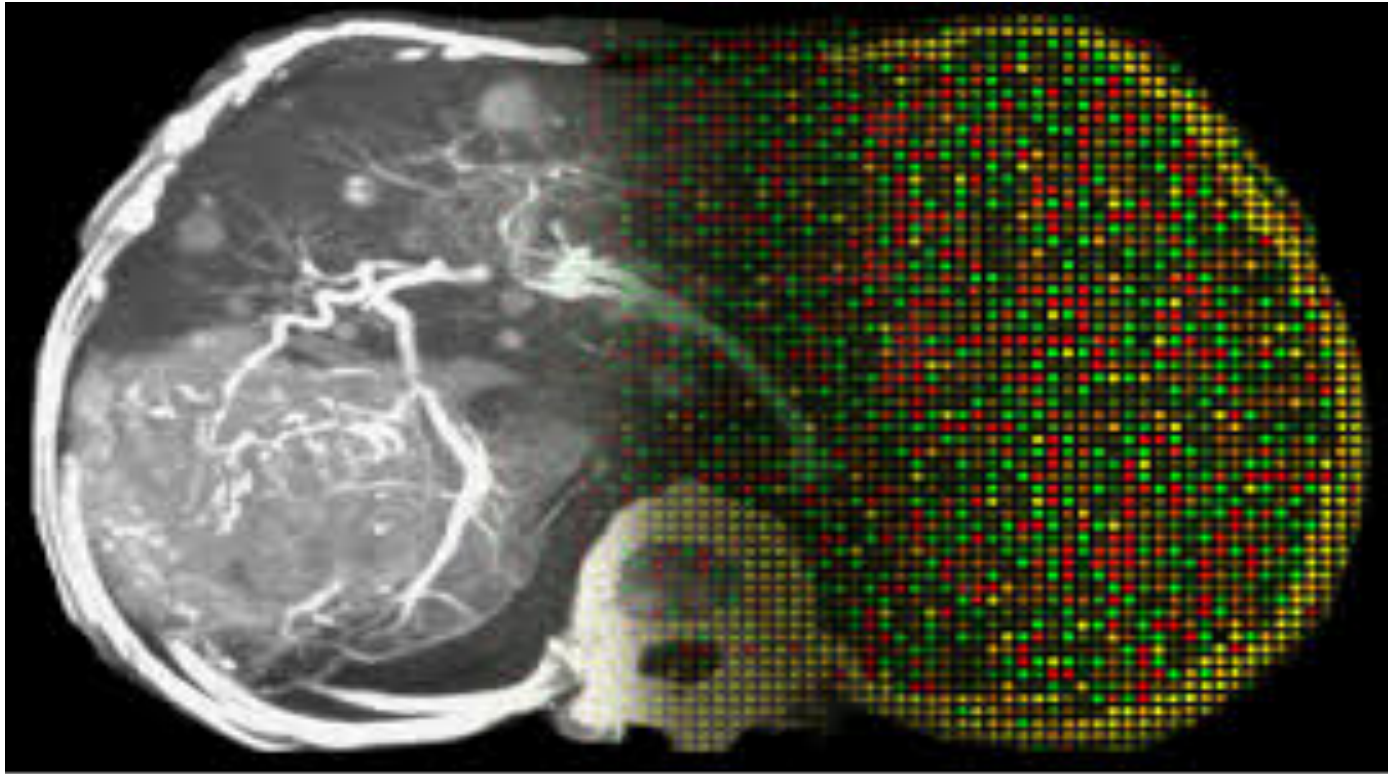


Islands of
information



Sustainability

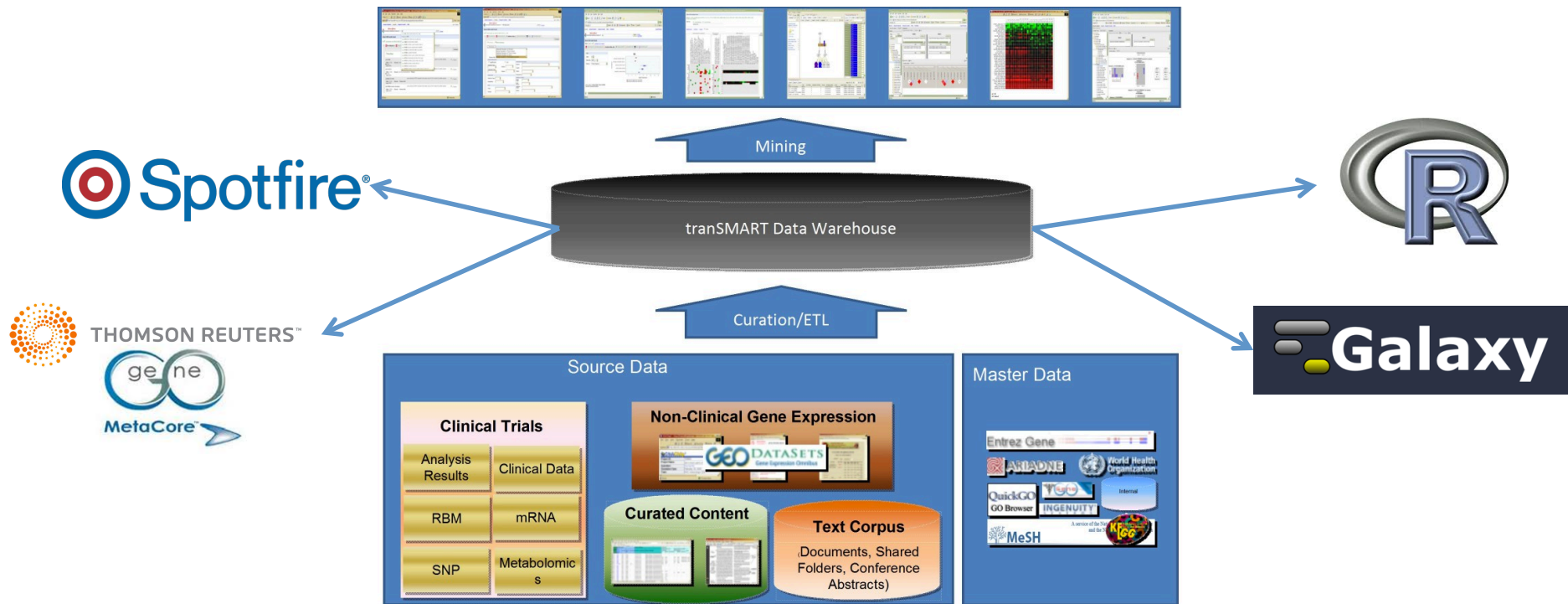
tranSMART : a fast growing platform



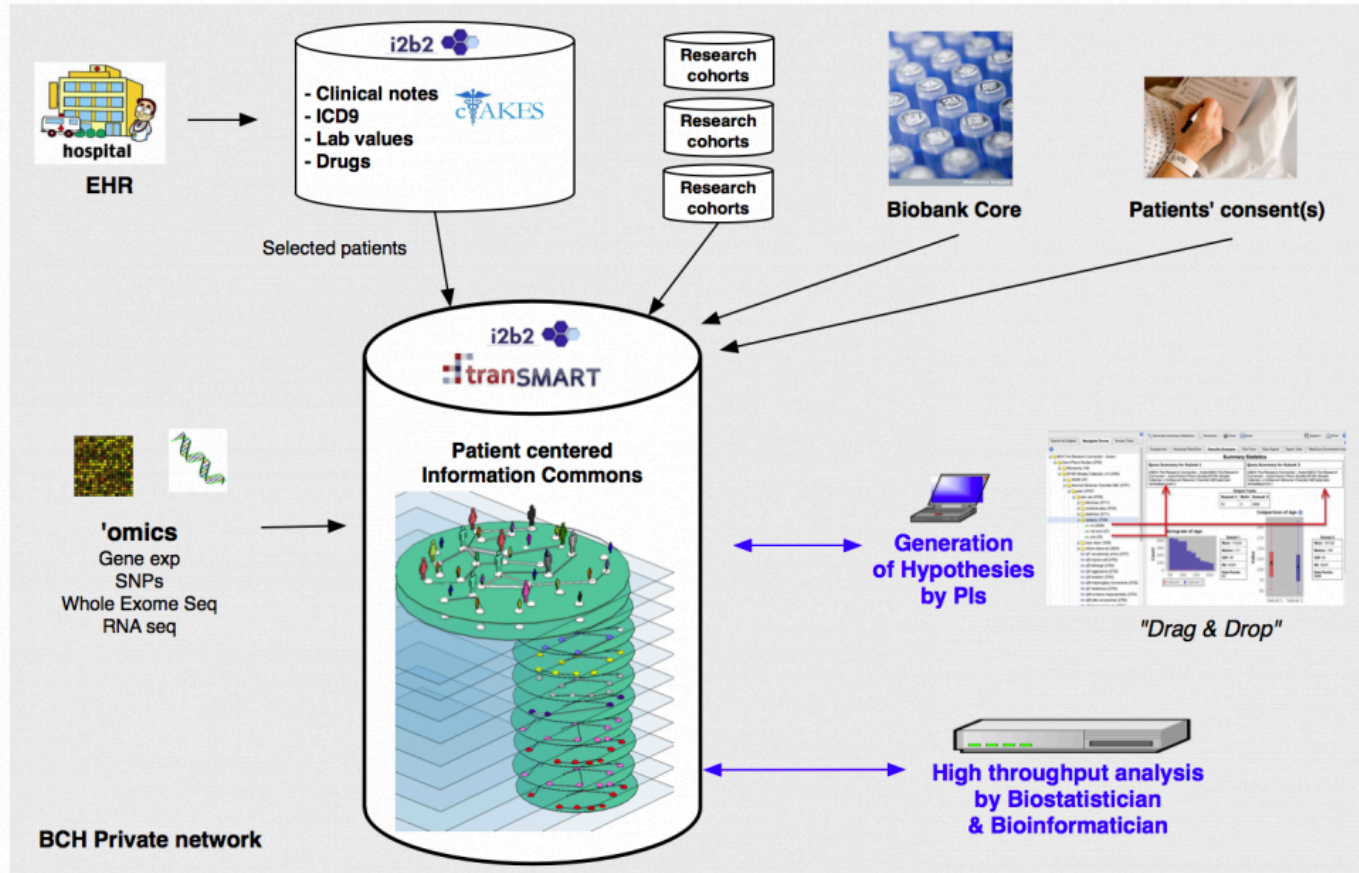
Objectives :

1. **Integration of clinical, biological and 'omics data in one place – hypothesis free –**
2. **Generation of hypothesis by Clinicians / Researchers**

tranSMART

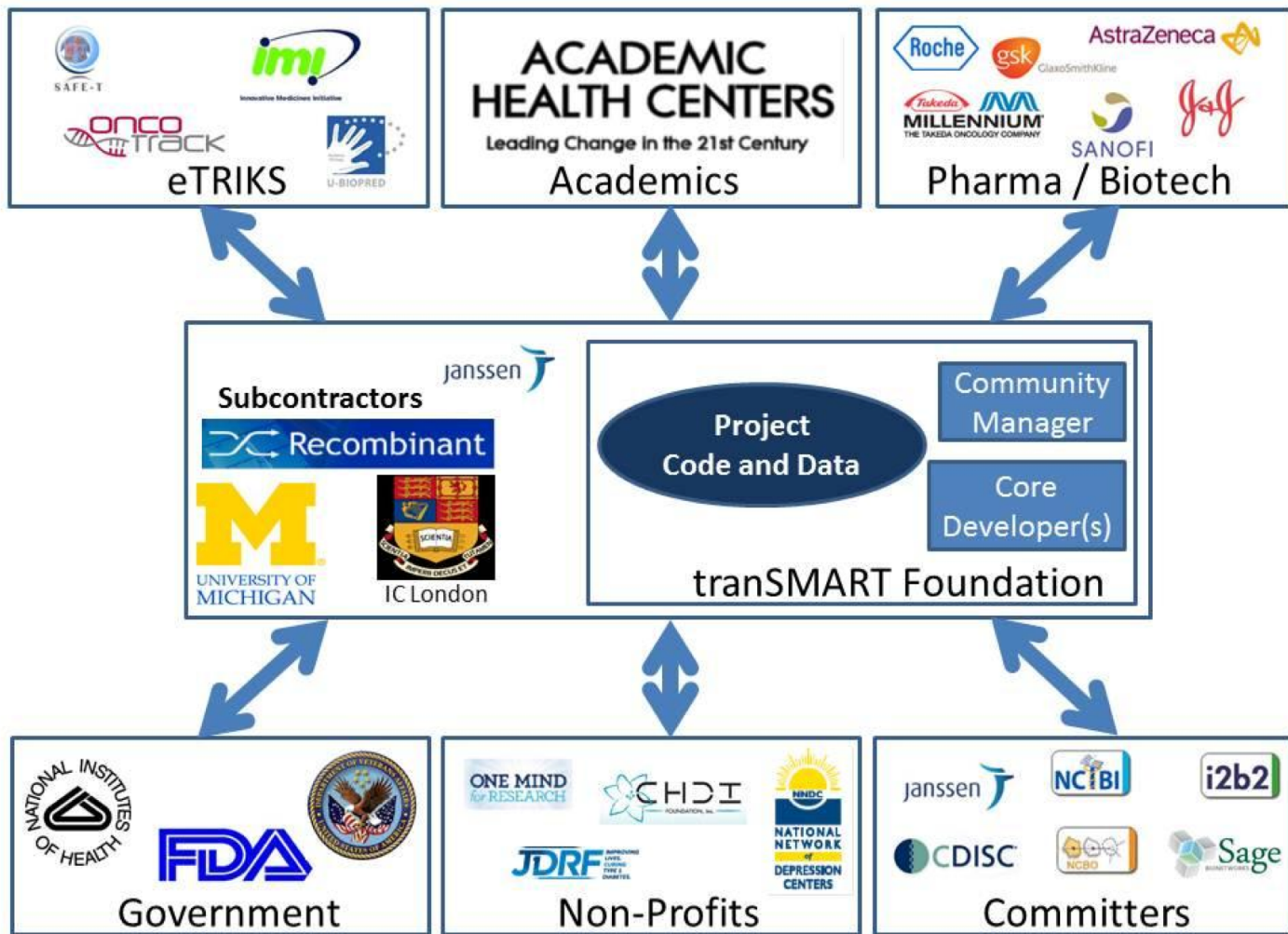


tranSMART @ Boston Children's Hospital



→ integration
 ↔ analysis

tranSMART Foundation





European Translational Research Information and Knowledge Management Services



1 Billion € **Public**

1 Billion € **Private**

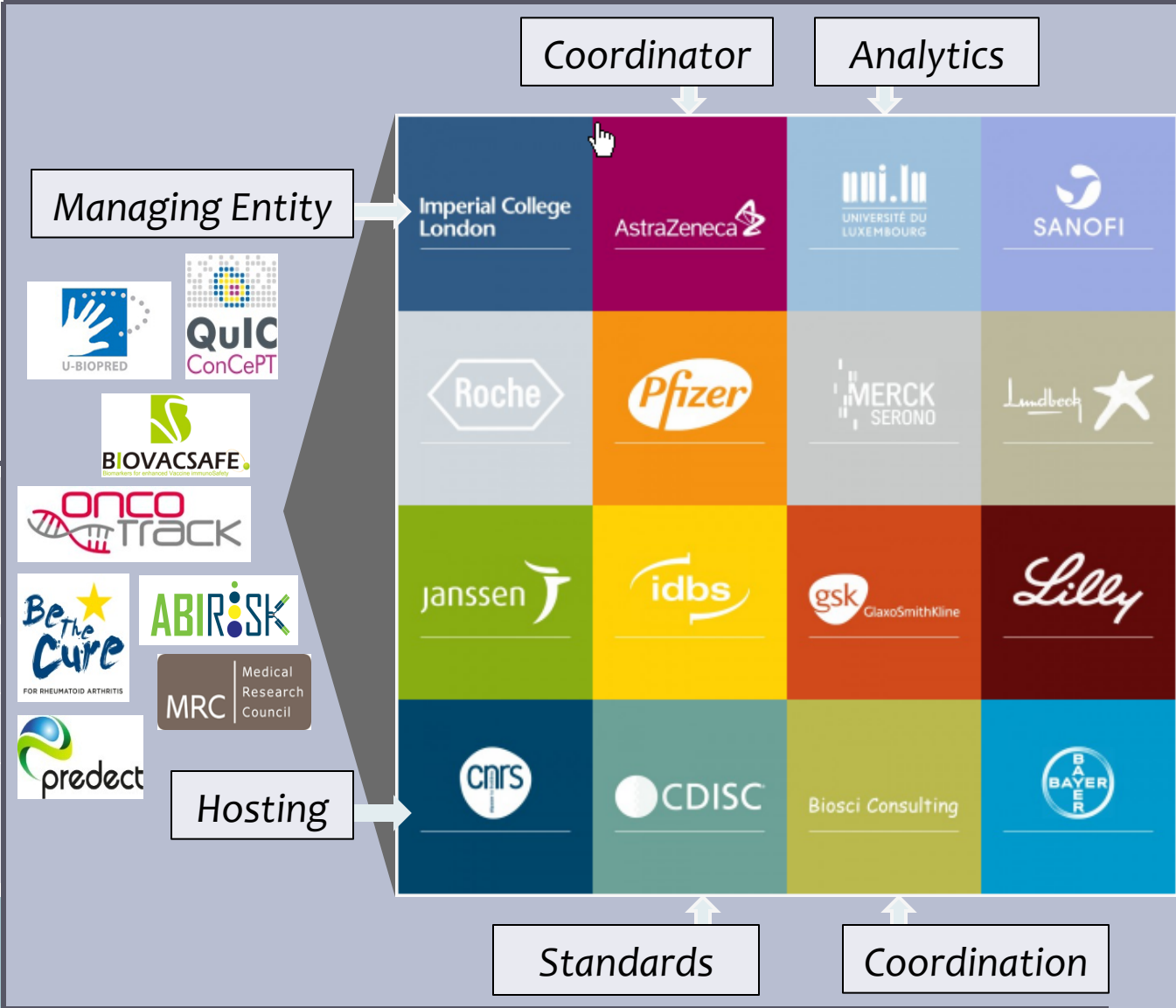
Partnership



~ \$23M Euro

Oct-2012 – Sept-2017

Sustainable Open Platform



Goals from the *Full Project Proposal*

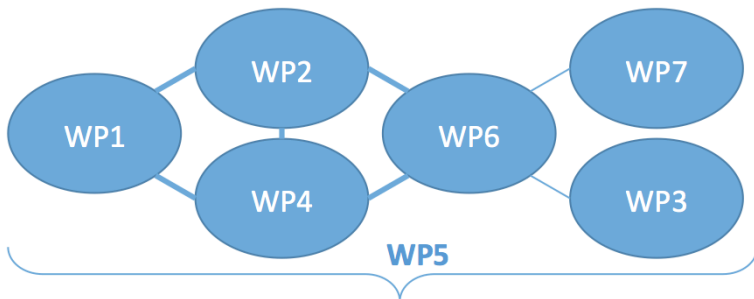
- A cloud-based biomarker discovery infrastructure**
 - Support for OMICs, clinical data, cohorts
 - A rich set of analytical methods
- A collaborative research management system**
 - Support large scale IMI programs
- An integrative knowledge and management environment**
 - QC, ontology and terminology management, data provenance
- A scientific knowledge service delivery platform**
 - Knowledge sharing, “knowledge as a service”
- A federated knowledge sharing environment**
 - Support local instances of the eTRIKS platform (data privacy)
 - Share public data and data across projects where appropriate

Constraints/Priorities

- Must use *tranSMART***
 - Infrastructure must be a fully *open technology stack*
- Priority: Support client projects**

Work Packages

	WP Number	WP Name	WP Leads
<i>Biosci (Program Mngmt)</i>	WP1	Platform Deployment	CNRS
	WP2	Platform Development	Imperial/Sanofi/Pfizer
	WP3	Data Standards	Roche/IDBS/Lilly/CDISC
	WP4	Curation and Analysis	Luxembourg/Sanofi/Merck Serono
	WP5	Management/Sustainability	AstraZeneca
	WP6	Community and Outreach	Janssen/Biosci
	WP7	Ethics	GSK/CNRS/Bayer



Horizontal Cooperation

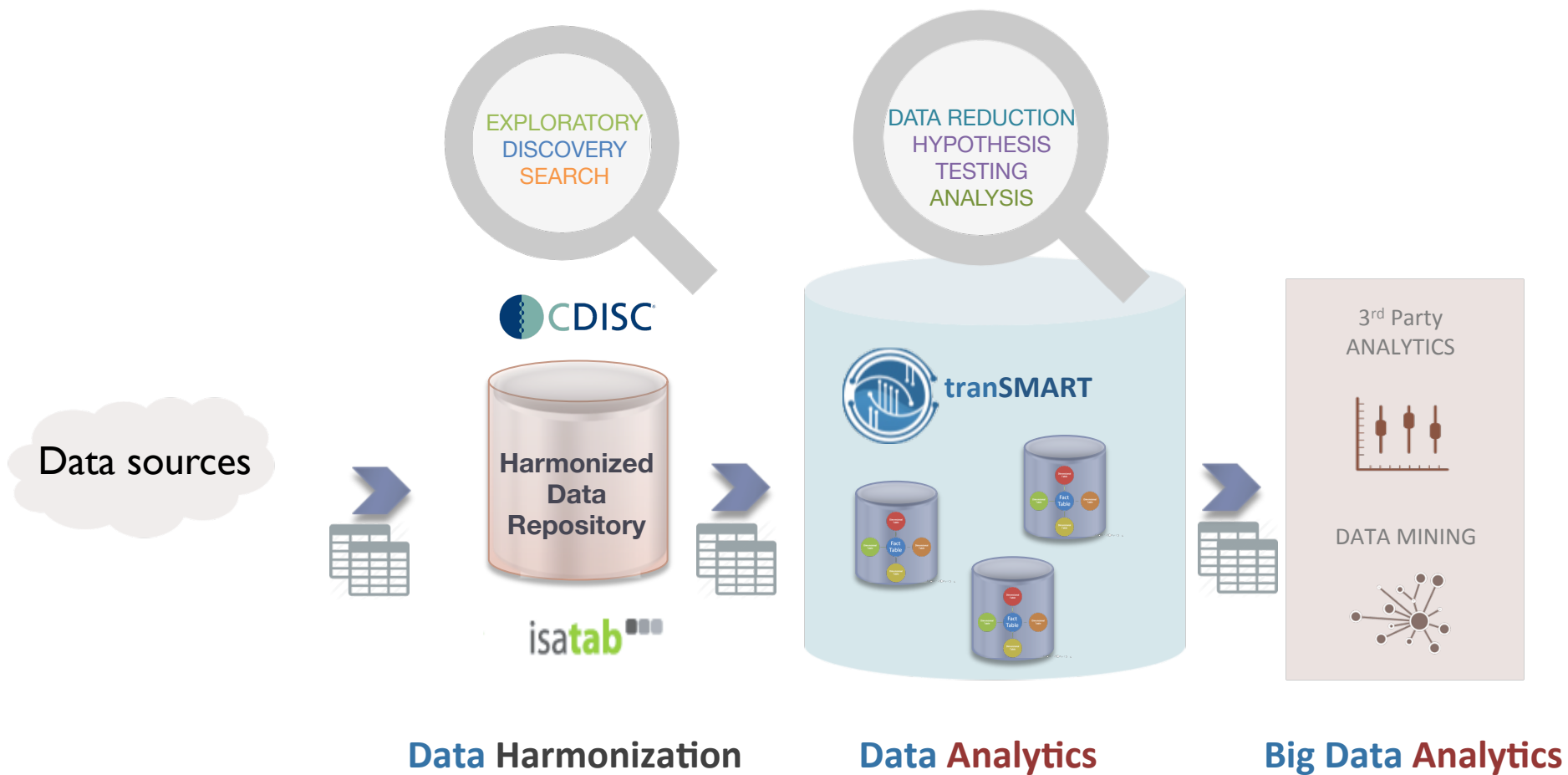
Delivery Package: 2+ WPs. Client Delivery.

- eTRIKS/tranSMART 1.1, GUI
- ABI Risk, Oncotrack
- Public Server

Task Force: 3+ WPs

Functional Domain Teams: WP2/WP6

eTRIKS Platform Architecture



eTRIKS Supported Projects

Oncology



Safety



RA-Map



Approach (OA)



Neuro



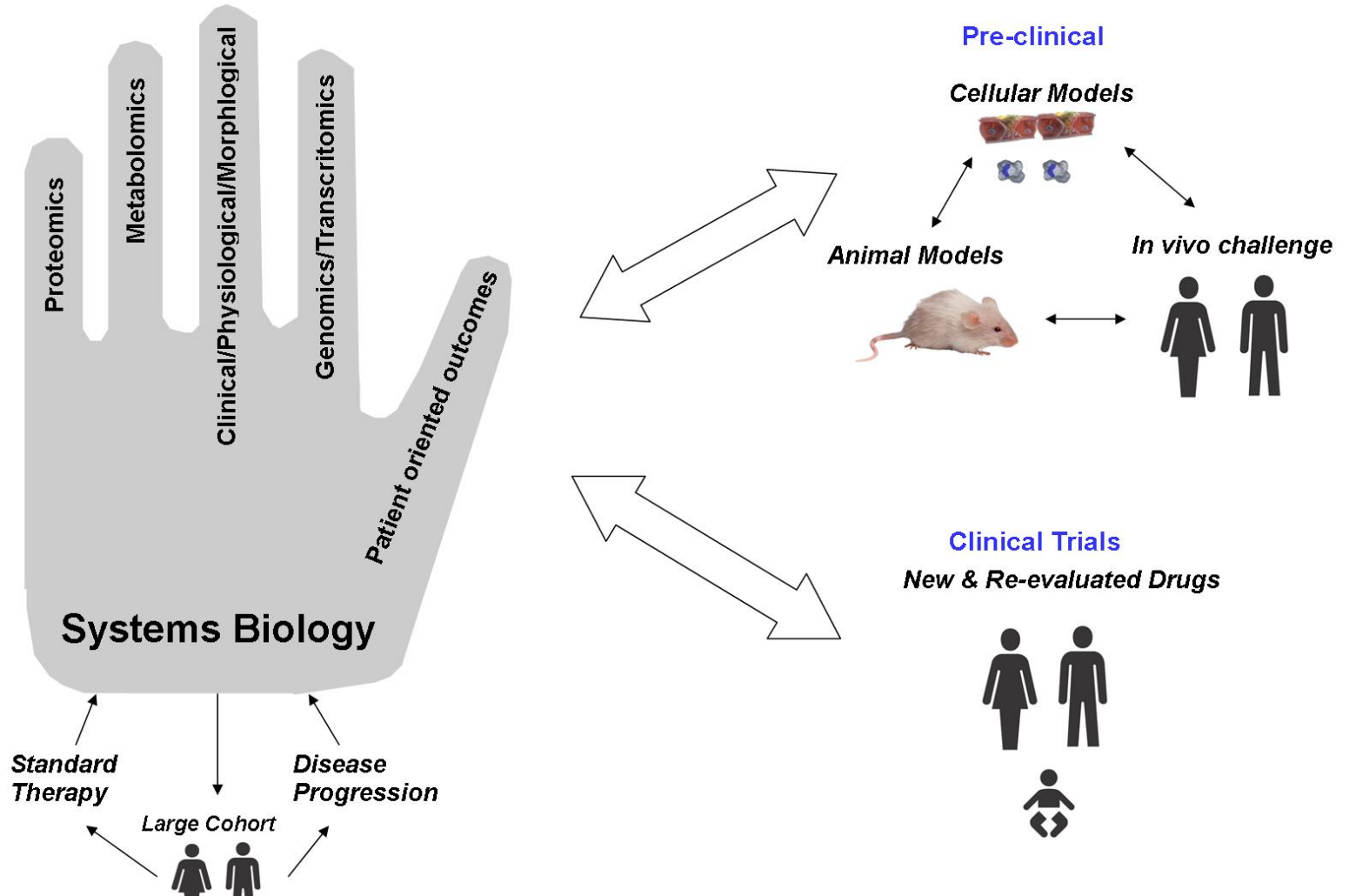
Infection

Inflammation / Immunology

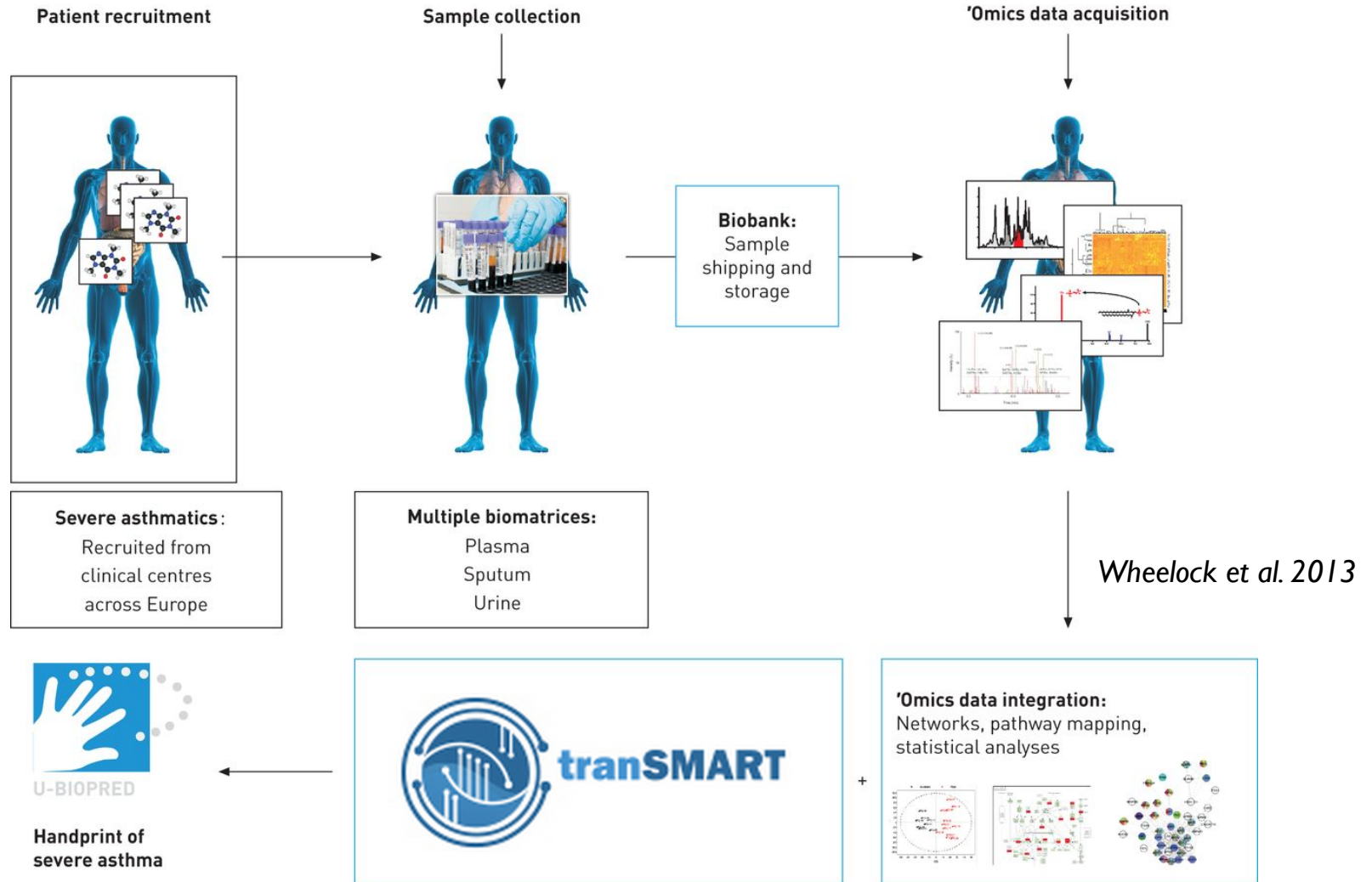


Unbiased BIOmarkers in PREDiction of
respiratory disease outcomes

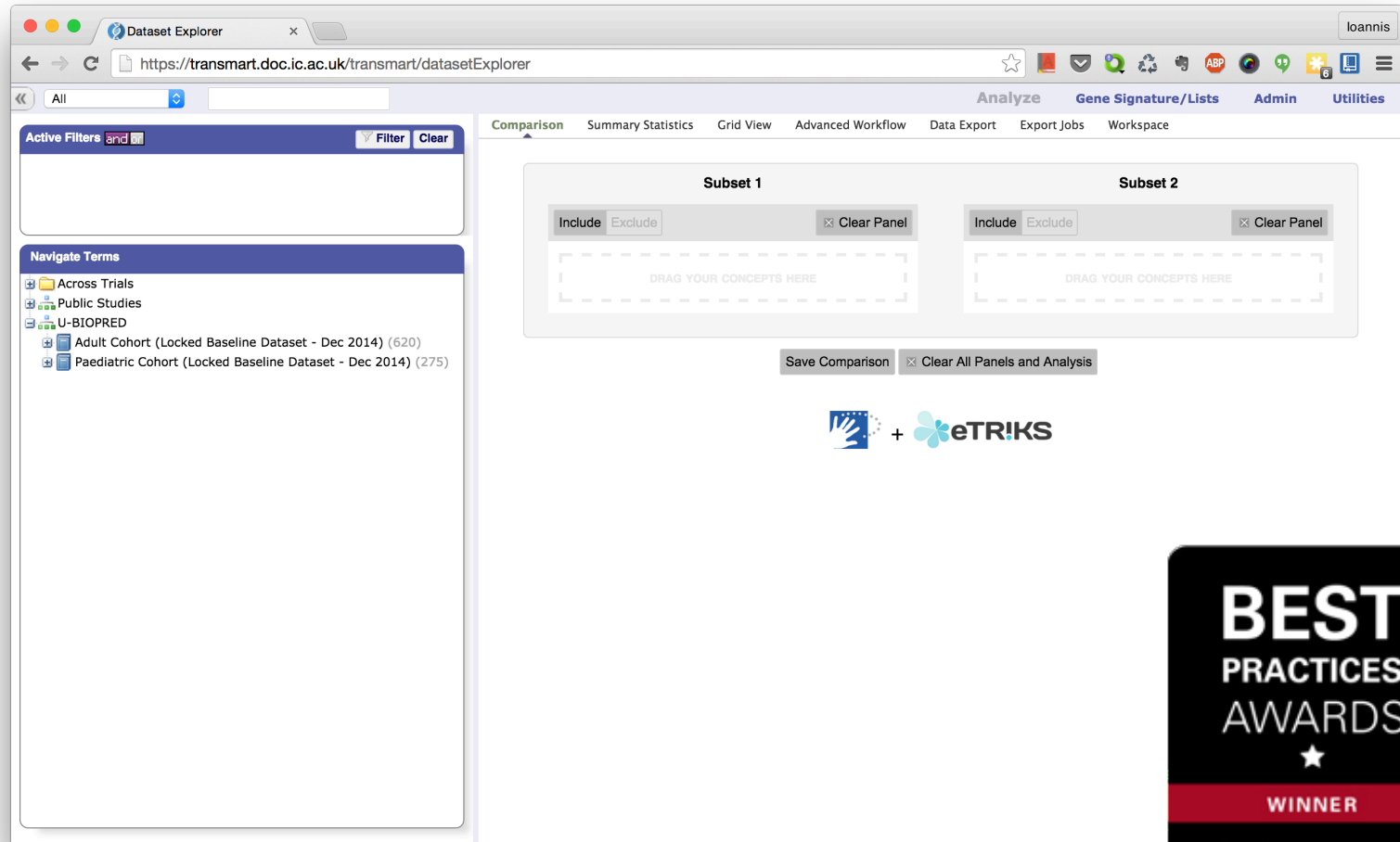
U-BIOPRED project



Sample and Data Flow



U-BIOPRED eTRIKS-tranSMART



The screenshot displays the tranSMART Dataset Explorer web application. The browser address bar shows the URL <https://transmart.doc.ic.ac.uk/transmart/datasetExplorer>. The user is logged in as 'loannis'. The interface includes a navigation menu with options like 'Analyze', 'Gene Signature/Lists', 'Admin', and 'Utilities'. A left sidebar contains 'Active Filters' and 'Navigate Terms', with a tree view showing 'U-BIOPRED' and its sub-cohorts: 'Adult Cohort (Locked Baseline Dataset - Dec 2014) (620)' and 'Paediatric Cohort (Locked Baseline Dataset - Dec 2014) (275)'. The main workspace is titled 'Comparison' and features two panels, 'Subset 1' and 'Subset 2', each with 'Include', 'Exclude', and 'Clear Panel' buttons and a dashed box for 'DRAG YOUR CONCEPTS HERE'. Below the panels are 'Save Comparison' and 'Clear All Panels and Analysis' buttons. The eTRIKS logo is visible at the bottom center of the interface.

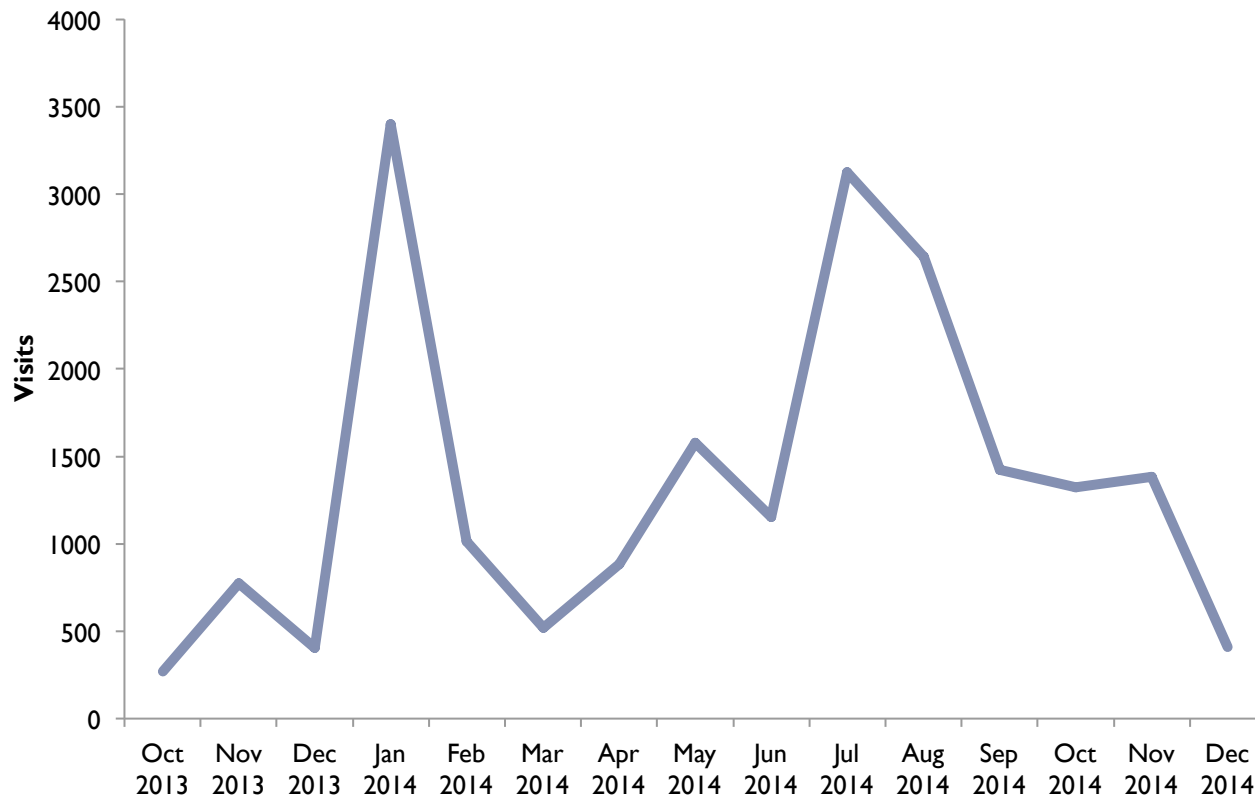


U-BIOPRED eTRIKS-tranSMART content to date

Cohort	Omics Type	Datasest	Number of Subjects/ Sample	Number of features	Total Features (x10 ³)
Adult	Gene expression	Blood	312	54675	17058.6
		Bronchial Brushings	149	54675	8146.575
		Bronchial Biopsies	108	54675	5904.9
		Nasal Brushings	89	54675	4866.075
		Sputum	120	54675	6561
	Protein abundance	Serum Somalogic	579	1129	653.691
		Sputum Somalogic	121	1129	136.609
		Serum MS ^{^e}	131	130	17.03
		Sputum MS ^{^e}	100	3231	323.1
	Lipid abundance	Sputum MS ^{^e}	128	15622	1999.616
		Urine Eicosanoids	611	13	7.943
		Sputum Eicosanoids	321	13	4.173
	Breath metabolites	eNOSE	106	190	20.14
		GC/MS	62	7036	436.232
Paediatric	Lipid abundance	Urine Eicosanoids	35	13	
		Paediatric eNOSE	107	190	20.33
	Breath metabolites	Paediatric GC/MS	101	7036	710.636
Adult		Clinical / Phenotypic	620	2886	1789.32
Paediatric		Clinical / Phenotypic	275	1210	332.75
Total					92656.3

Visit Statistics

▶ Number of users: 227



Thank you!



Imperial College
London Data Science Institute

www.etriks.org

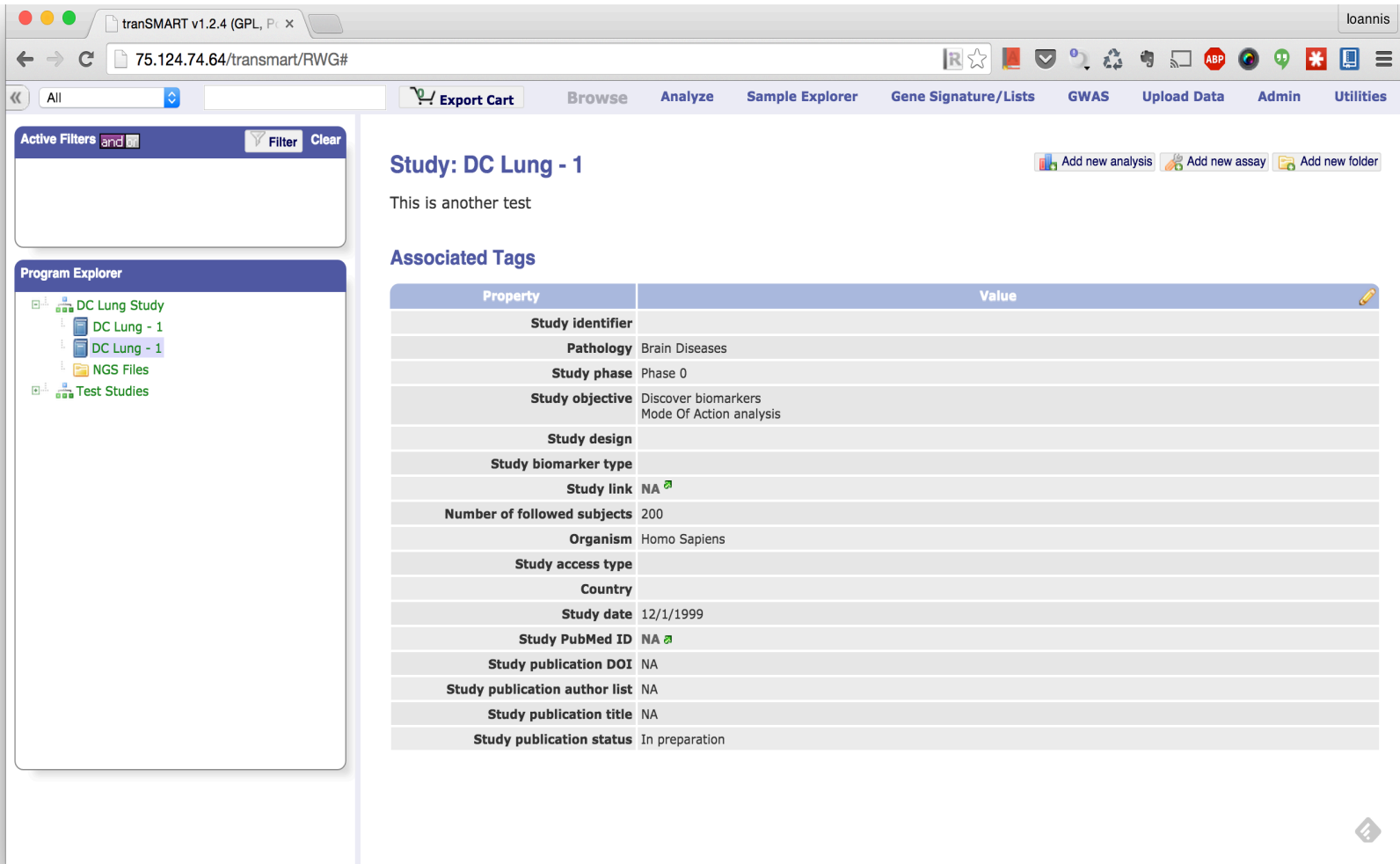
transmartfoundation.org



efpia



Browse and Search – Storing study information, metadata and associated raw files



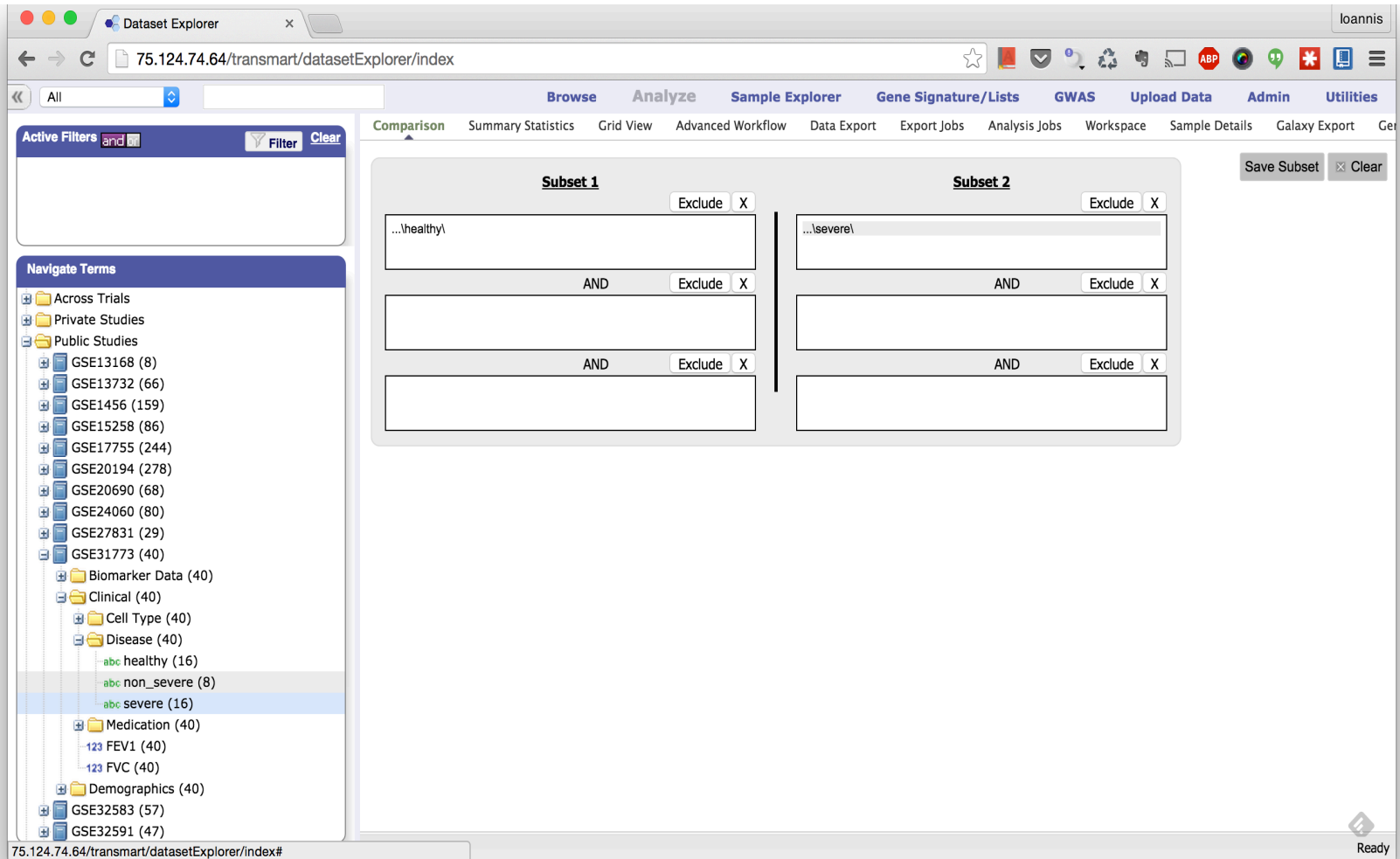
The screenshot shows the tranSMART v1.2.4 web interface. The browser address bar displays '75.124.74.64/transmart/RWG#'. The top navigation bar includes 'Export Cart', 'Browse', 'Analyze', 'Sample Explorer', 'Gene Signature/Lists', 'GWAS', 'Upload Data', 'Admin', and 'Utilities'. The user 'Ioannis' is logged in.

The main content area displays 'Study: DC Lung - 1' with the text 'This is another test'. Below this is the 'Associated Tags' section, which is a table with the following data:

Property	Value
Study identifier	
Pathology	Brain Diseases
Study phase	Phase 0
Study objective	Discover biomarkers Mode Of Action analysis
Study design	
Study biomarker type	
Study link	NA
Number of followed subjects	200
Organism	Homo Sapiens
Study access type	
Country	
Study date	12/1/1999
Study PubMed ID	NA
Study publication DOI	NA
Study publication author list	NA
Study publication title	NA
Study publication status	In preparation

On the left side, the 'Program Explorer' shows a tree view with 'DC Lung Study' expanded to show 'DC Lung - 1' and 'DC Lung - 1' expanded to show 'NGS Files' and 'Test Studies'.

Analyze - Cohort Selection Tool



The screenshot displays the 'Dataset Explorer' web application interface. The browser address bar shows the URL `75.124.74.64/transmart/datasetExplorer/index`. The user is logged in as 'Ioannis'.

The interface features a top navigation bar with tabs: **Browse**, **Analyze** (active), **Sample Explorer**, **Gene Signature/Lists**, **GWAS**, **Upload Data**, **Admin**, and **Utilities**. Below this is a secondary navigation bar with options: **Comparison**, **Summary Statistics**, **Grid View**, **Advanced Workflow**, **Data Export**, **Export Jobs**, **Analysis Jobs**, **Workspace**, **Sample Details**, **Galaxy Export**, and **Gen**.

On the left side, there is a 'Navigate Terms' tree view. The 'Disease' category is expanded, showing sub-categories: 'healthy' (16), 'non_severe' (8), and 'severe' (16). The 'severe' category is currently selected.

The main workspace is titled 'Comparison' and is used for defining two subsets for analysis:

- Subset 1**: Contains a text input field with the value `...healthy`. Below it are buttons for 'AND', 'Exclude', and 'X'.
- Subset 2**: Contains a text input field with the value `...severe`. Below it are buttons for 'Exclude' and 'X'.

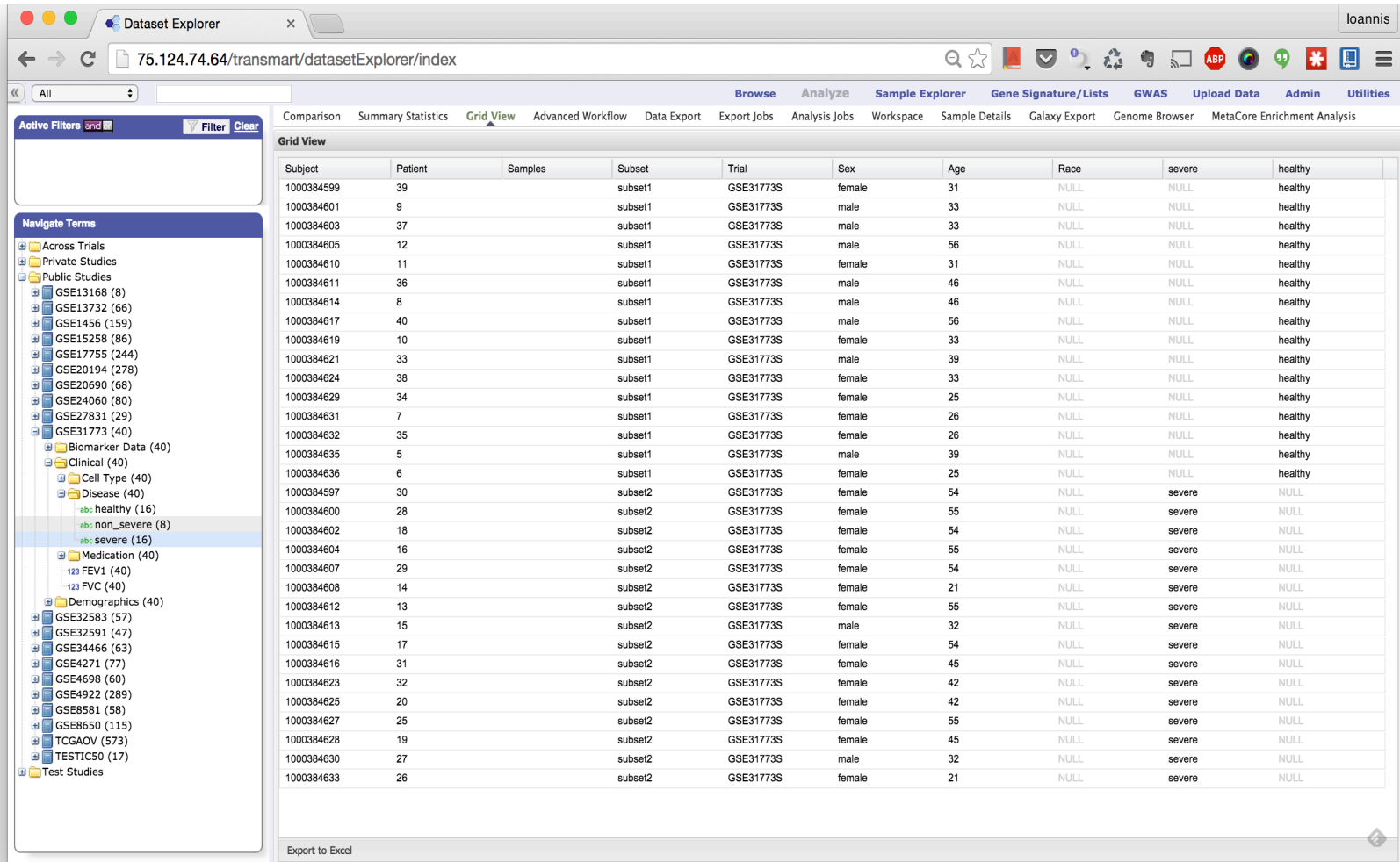
Below the subset definitions, there are three rows of 'AND' logic connectors, each with 'Exclude' and 'X' buttons, allowing for complex logical combinations of the subsets.

At the bottom right of the workspace, there are buttons for 'Save Subset' and 'Clear'. The status bar at the bottom right indicates 'Ready'.

Analyze - Cohort Summary Statistics & clinical variable statistical analysis



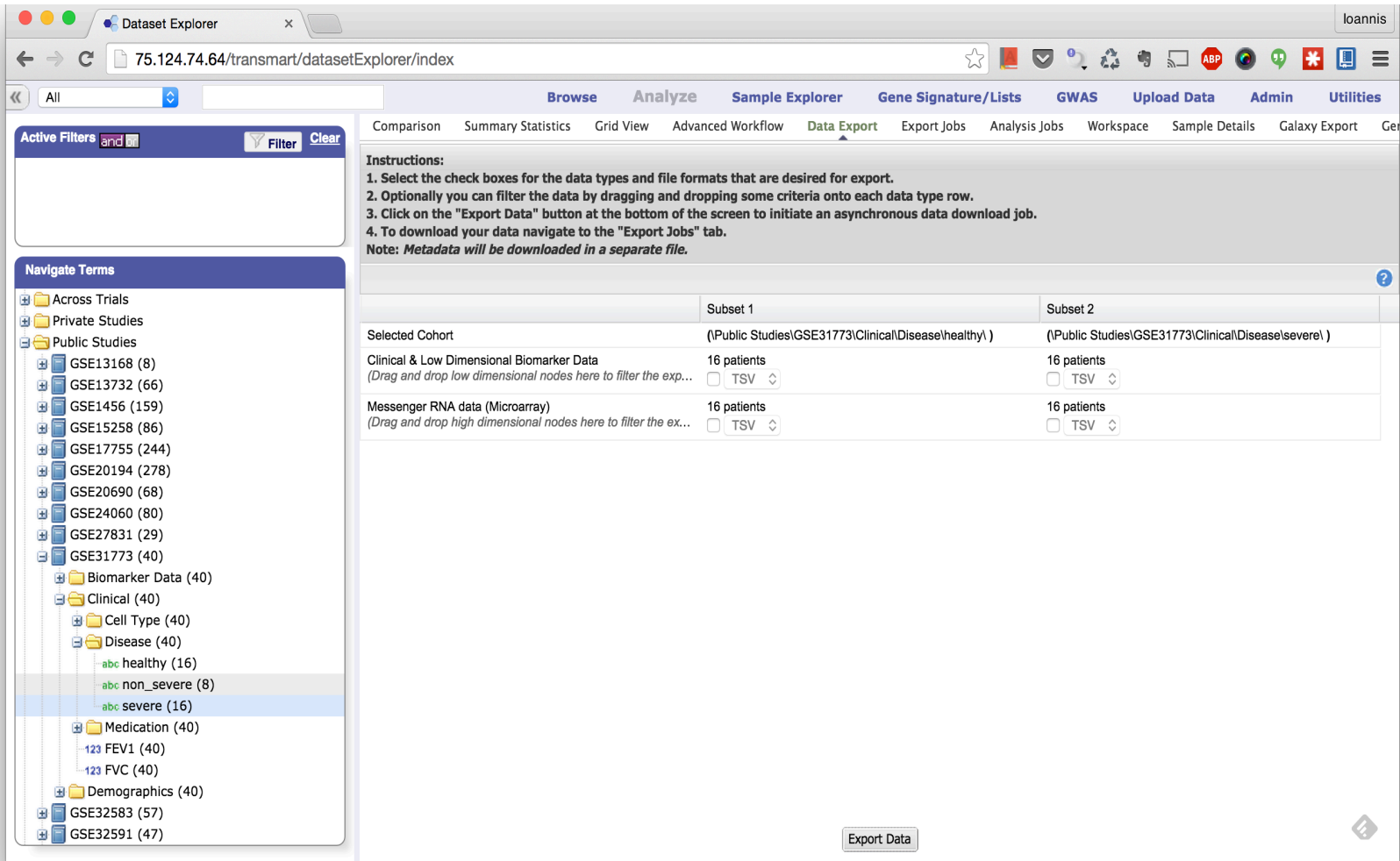
Analyze – Grid View for visual inspection and export of selected data



The screenshot displays the eTRIKS Dataset Explorer web application. The browser address bar shows the URL `75.124.74.64/transmart/datasetExplorer/index`. The user is logged in as 'loannis'. The interface includes a navigation menu with options like 'Browse', 'Analyze', 'Sample Explorer', and 'Grid View' (which is currently selected). A sidebar on the left contains a 'Navigate Terms' tree view with categories such as 'Public Studies', 'Biomarker Data', 'Clinical', 'Medication', and 'Demographics'. The main area shows a 'Grid View' table with columns for Subject, Patient, Samples, Subset, Trial, Sex, Age, Race, severe, and healthy. The table contains 30 rows of data. At the bottom of the grid, there is an 'Export to Excel' button.

Subject	Patient	Samples	Subset	Trial	Sex	Age	Race	severe	healthy
1000384599	39		subset1	GSE31773S	female	31	NULL	NULL	healthy
1000384601	9		subset1	GSE31773S	male	33	NULL	NULL	healthy
1000384603	37		subset1	GSE31773S	male	33	NULL	NULL	healthy
1000384605	12		subset1	GSE31773S	male	56	NULL	NULL	healthy
1000384610	11		subset1	GSE31773S	female	31	NULL	NULL	healthy
1000384611	36		subset1	GSE31773S	male	46	NULL	NULL	healthy
1000384614	8		subset1	GSE31773S	male	46	NULL	NULL	healthy
1000384617	40		subset1	GSE31773S	male	56	NULL	NULL	healthy
1000384619	10		subset1	GSE31773S	female	33	NULL	NULL	healthy
1000384621	33		subset1	GSE31773S	male	39	NULL	NULL	healthy
1000384624	38		subset1	GSE31773S	female	33	NULL	NULL	healthy
1000384629	34		subset1	GSE31773S	female	25	NULL	NULL	healthy
1000384631	7		subset1	GSE31773S	female	26	NULL	NULL	healthy
1000384632	35		subset1	GSE31773S	female	26	NULL	NULL	healthy
1000384635	5		subset1	GSE31773S	male	39	NULL	NULL	healthy
1000384636	6		subset1	GSE31773S	female	25	NULL	NULL	healthy
1000384597	30		subset2	GSE31773S	female	54	NULL	severe	NULL
1000384600	28		subset2	GSE31773S	female	55	NULL	severe	NULL
1000384602	18		subset2	GSE31773S	female	54	NULL	severe	NULL
1000384604	16		subset2	GSE31773S	female	55	NULL	severe	NULL
1000384607	29		subset2	GSE31773S	female	54	NULL	severe	NULL
1000384608	14		subset2	GSE31773S	female	21	NULL	severe	NULL
1000384612	13		subset2	GSE31773S	female	55	NULL	severe	NULL
1000384613	15		subset2	GSE31773S	male	32	NULL	severe	NULL
1000384615	17		subset2	GSE31773S	female	54	NULL	severe	NULL
1000384616	31		subset2	GSE31773S	female	45	NULL	severe	NULL
1000384623	32		subset2	GSE31773S	female	42	NULL	severe	NULL
1000384625	20		subset2	GSE31773S	female	42	NULL	severe	NULL
1000384627	25		subset2	GSE31773S	female	55	NULL	severe	NULL
1000384628	19		subset2	GSE31773S	female	45	NULL	severe	NULL
1000384630	27		subset2	GSE31773S	male	32	NULL	severe	NULL
1000384633	26		subset2	GSE31773S	female	21	NULL	severe	NULL

Analyze – Data Export for mass data (clinical and omics) export



The screenshot shows the Dataset Explorer web application interface. The browser address bar displays the URL `75.124.74.64/transmart/datasetExplorer/index`. The application has a navigation menu with tabs: **Browse**, **Analyze**, **Sample Explorer**, **Gene Signature/Lists**, **GWAS**, **Upload Data**, **Admin**, and **Utilities**. The **Analyze** tab is active, and the **Data Export** sub-tab is selected.

On the left side, there is a **Navigate Terms** tree view showing a hierarchical structure of data categories. The **Clinical** category is expanded, showing sub-categories like **healthy** (16), **non_severe** (8), and **severe** (16). The **severe** category is currently selected.

The main content area displays **Instructions** for data export:

1. Select the check boxes for the data types and file formats that are desired for export.
2. Optionally you can filter the data by dragging and dropping some criteria onto each data type row.
3. Click on the "Export Data" button at the bottom of the screen to initiate an asynchronous data download job.
4. To download your data navigate to the "Export Jobs" tab.

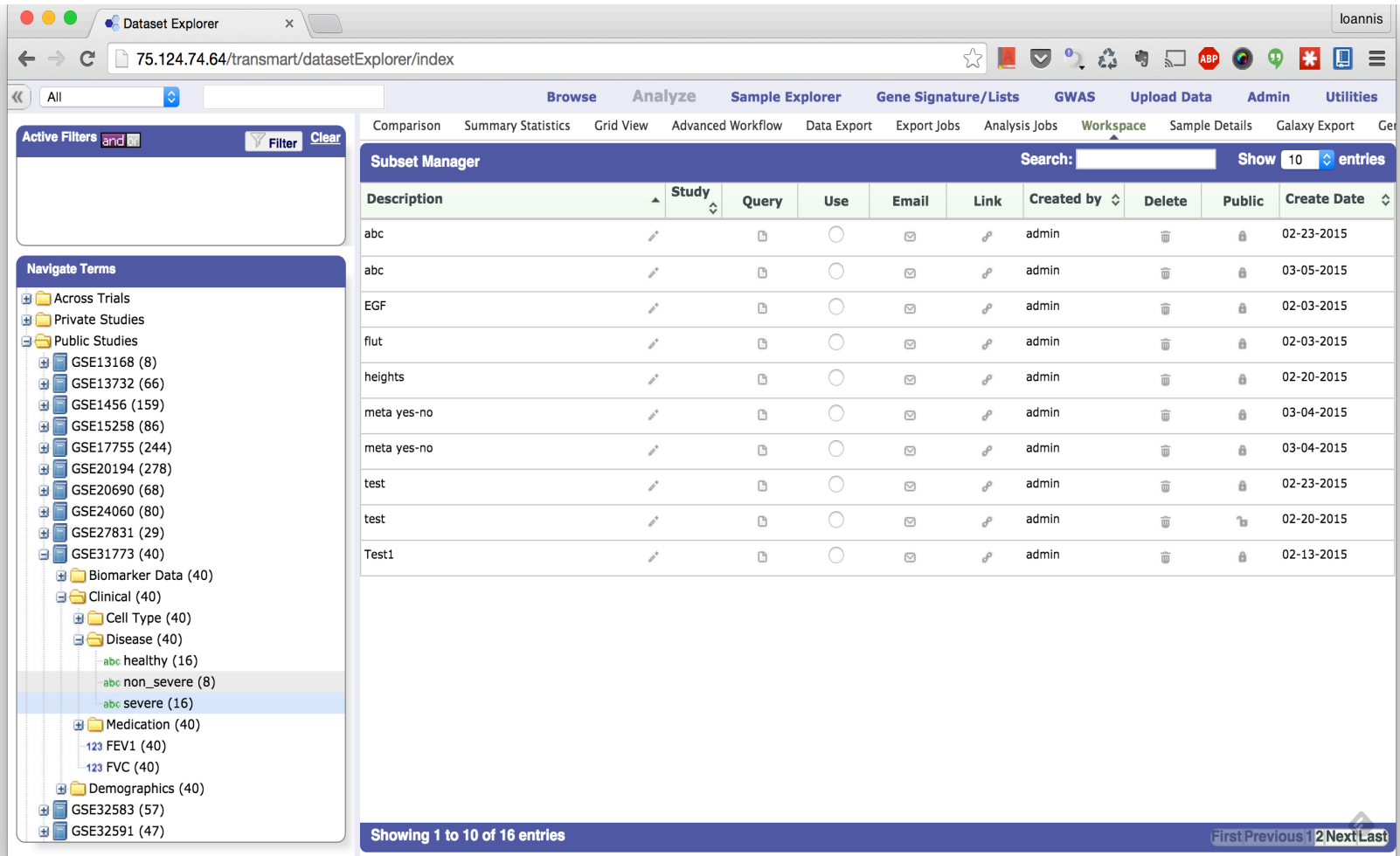
Note: Metadata will be downloaded in a separate file.

Below the instructions is a table for selecting data subsets:

	Subset 1	Subset 2
Selected Cohort	(\Public Studies\GSE31773\Clinical\Disease\healthy\)	(\Public Studies\GSE31773\Clinical\Disease\severe\)
Clinical & Low Dimensional Biomarker Data <i>(Drag and drop low dimensional nodes here to filter the exp...)</i>	16 patients <input type="checkbox"/> TSV	16 patients <input type="checkbox"/> TSV
Messenger RNA data (Microarray) <i>(Drag and drop high dimensional nodes here to filter the ex...)</i>	16 patients <input type="checkbox"/> TSV	16 patients <input type="checkbox"/> TSV

At the bottom right of the interface, there is an **Export Data** button.

Analyze – Workspace for query saving and reuse

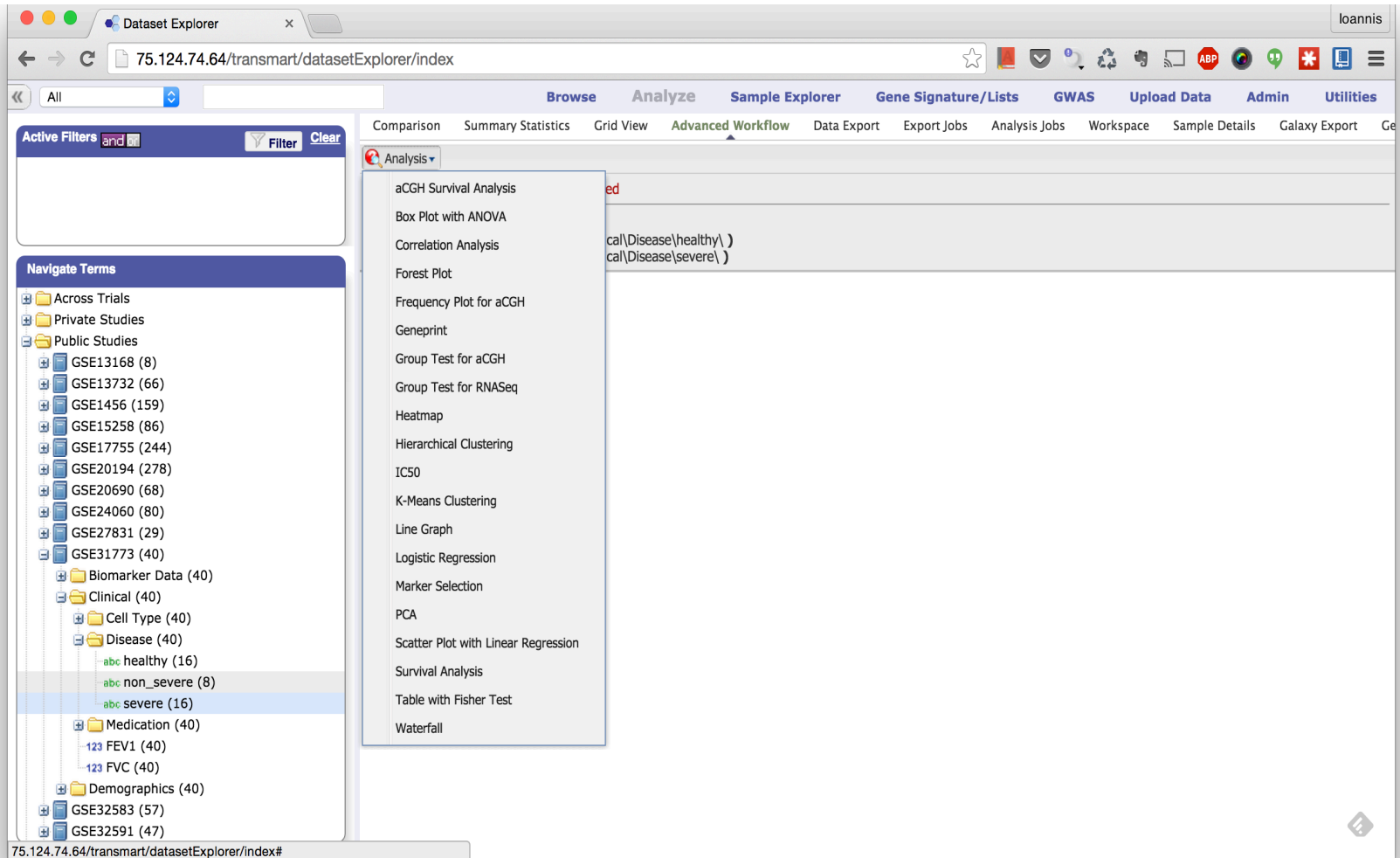


The screenshot displays the Dataset Explorer web application interface. The main content area shows a 'Subset Manager' table with the following data:

Description	Study	Query	Use	Email	Link	Created by	Delete	Public	Create Date
abc						admin			02-23-2015
abc						admin			03-05-2015
EGF						admin			02-03-2015
flut						admin			02-03-2015
heights						admin			02-20-2015
meta yes-no						admin			03-04-2015
meta yes-no						admin			03-04-2015
test						admin			02-23-2015
test						admin			02-20-2015
Test1						admin			02-13-2015

The interface also includes a 'Navigate Terms' sidebar on the left with a tree view of study categories, and a 'Showing 1 to 10 of 16 entries' status bar at the bottom.

Analyze – Advanced workflow for omics and advanced analysis



The screenshot displays the eTRIKS Dataset Explorer web application. The browser address bar shows the URL `75.124.74.64/transmart/datasetExplorer/index`. The user is logged in as 'Ioannis'. The main navigation bar includes 'Browse', 'Analyze', 'Sample Explorer', 'Gene Signature/Lists', 'GWAS', 'Upload Data', 'Admin', and 'Utilities'. The 'Analyze' menu is open, showing a list of analysis options:

- aCGH Survival Analysis
- Box Plot with ANOVA
- Correlation Analysis
- Forest Plot
- Frequency Plot for aCGH
- Geneprint
- Group Test for aCGH
- Group Test for RNASeq
- Heatmap
- Hierarchical Clustering
- IC50
- K-Means Clustering
- Line Graph
- Logistic Regression
- Marker Selection
- PCA
- Scatter Plot with Linear Regression
- Survival Analysis
- Table with Fisher Test
- Waterfall

The left sidebar contains 'Active Filters' and 'Navigate Terms'. The 'Navigate Terms' tree is expanded to show a hierarchy of data categories, with 'Disease' selected. Under 'Disease', the sub-categories 'healthy', 'non_severe', and 'severe' are visible, with 'severe' currently selected. The 'severe' category contains 16 samples.

Sample Explorer – Biobank data visualiser

Sample Explorer :: transM/ x ioannis

75.124.74.46:5880/transmart/sampleExplorer/list#

Browse Analyze **Sample Explorer** Gene Signature/Lists GWAS Upload Data Admin Utilities

Search Filters

By BioBank

- 105 (1)
- 160 (1)
- 249 (1)
- 292 (1)
- 326 (1)
- 484 (1)
- 494 (1)
- 695 (1)
- 843 (1)
- 929 (1)

By Subject Treatment

- (10)

By Source Organism

- human (10)

By Data Type

- (10)

By Sample Treatment

- (10)

By Tissue

- blood (6)
- brain (4)

By Pathology

- MS (10)

By Data Set

- EXP:GSE4382 (10)

By Sample ID

- 1 (1)
- 10 (1)
- 2 (1)
- 3 (1)
- 4 (1)
- 5 (1)
- 6 (1)
- 7 (1)
- 8 (1)
- 9 (1)

Clear Search ?

Comparison

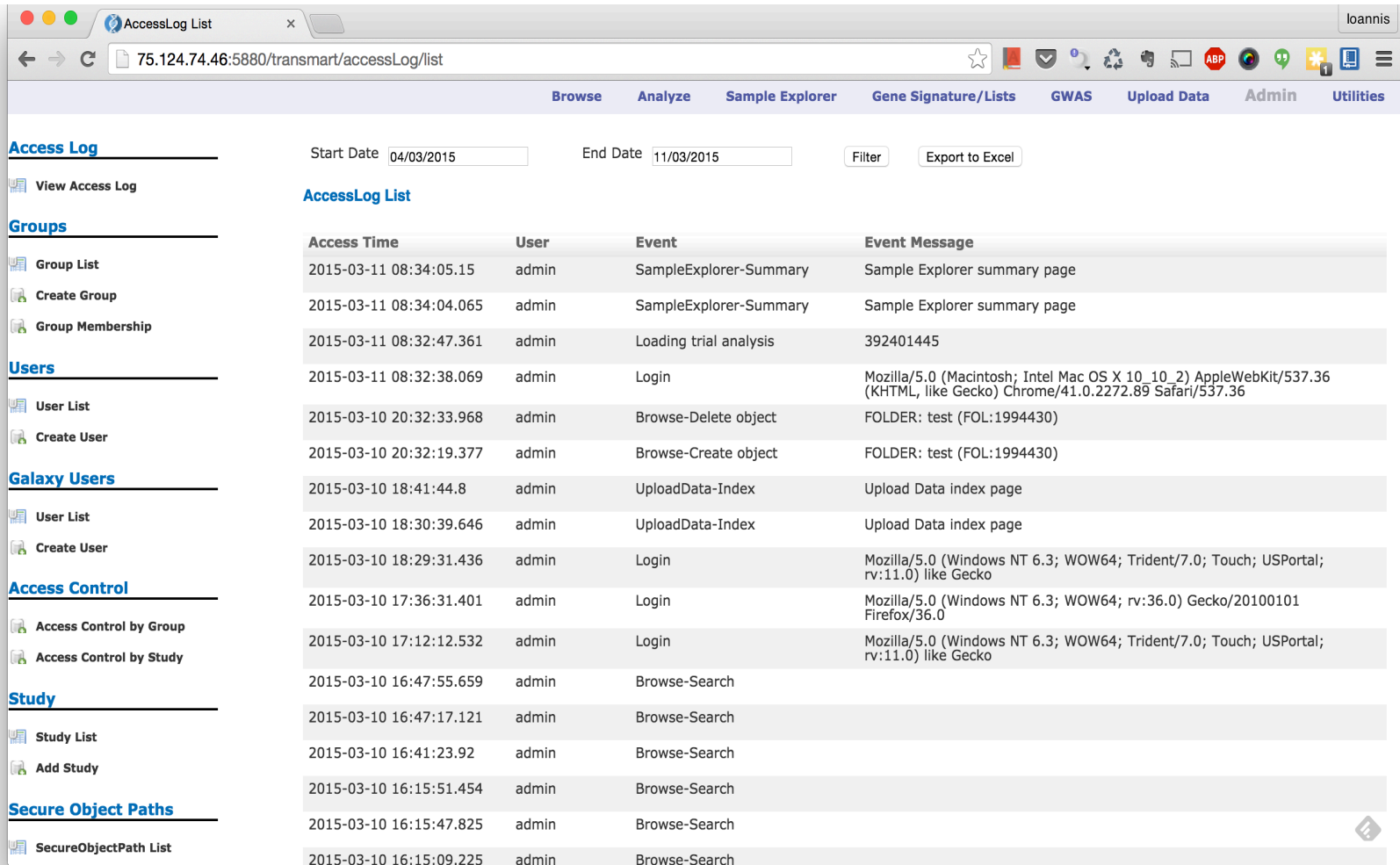
Subject Treatment: (-)

Sample Contact Information Collapse All Expand All

BioBank	Subject Treatment	Source Organism	Data Type	Tissue	Pathology	Aliquot Count
249	-	human	-	blood	MS	1
843	-	human	-	blood	MS	1
695	-	human	-	blood	MS	1
484	-	human	-	blood	MS	1
160	-	human	-	blood	MS	1
929	-	human	-	brain	MS	1
292	-	human	-	brain	MS	1
105	-	human	-	brain	MS	1
326	-	human	-	blood	MS	1
494	-	human	-	brain	MS	1

75.124.74.46:5880/transmart/sampleExplorer/list#

Admin Panel – Allowing granular access policy definition



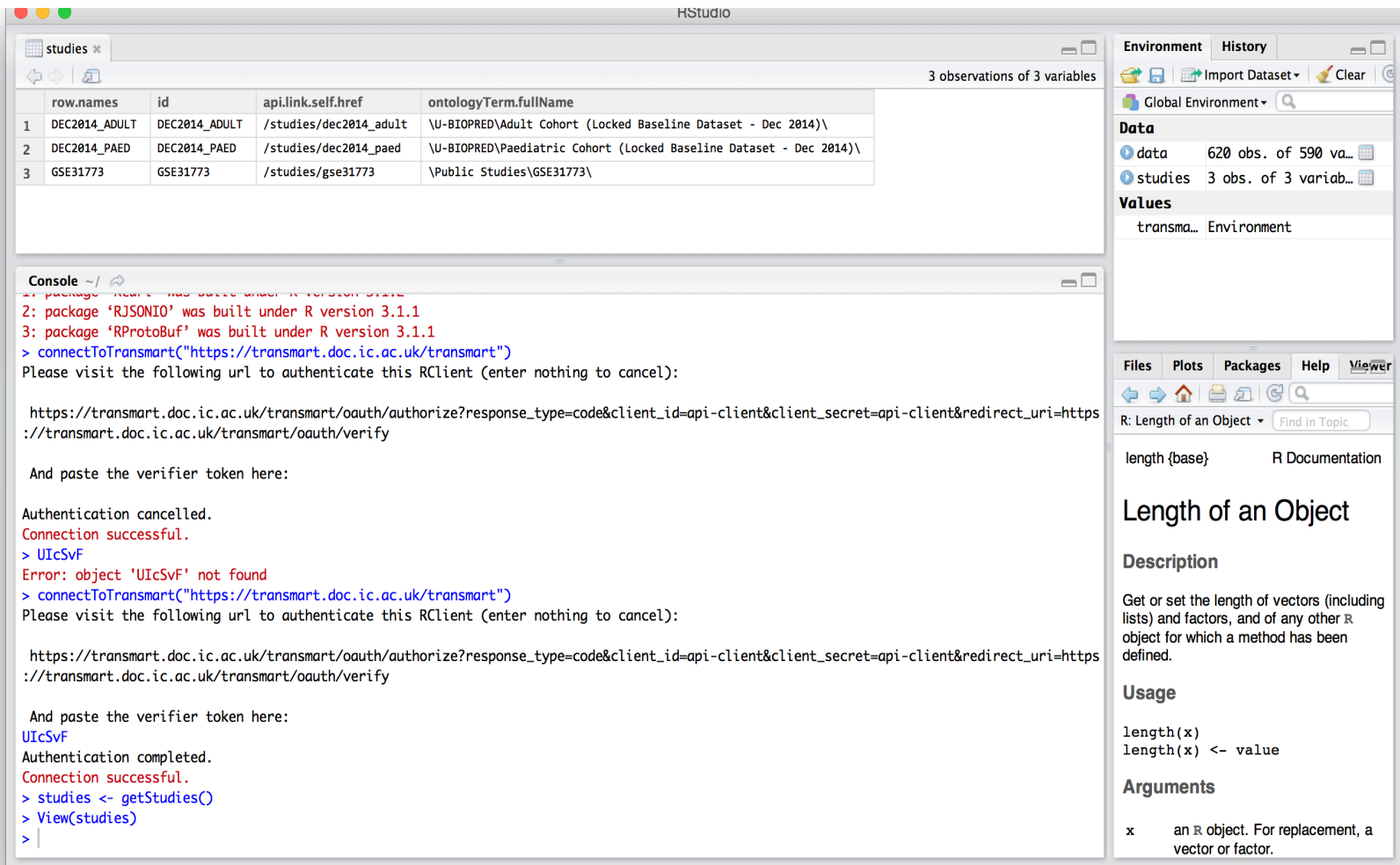
The screenshot displays the 'Access Log List' page in a web browser. The browser's address bar shows the URL '75.124.74.46:5880/transmart/accessLog/list'. The page features a top navigation bar with tabs for 'Browse', 'Analyze', 'Sample Explorer', 'Gene Signature/Lists', 'GWAS', 'Upload Data', 'Admin', and 'Utilities'. The 'Admin' tab is currently active.

On the left side, there is a sidebar menu with sections for 'Access Log', 'Groups', 'Users', 'Galaxy Users', 'Access Control', 'Study', and 'Secure Object Paths'. Each section contains sub-links for viewing, creating, and managing items.

The main content area is titled 'AccessLog List' and includes a search filter with 'Start Date' (04/03/2015) and 'End Date' (11/03/2015) fields, along with 'Filter' and 'Export to Excel' buttons. Below this is a table of access events:

Access Time	User	Event	Event Message
2015-03-11 08:34:05.15	admin	SampleExplorer-Summary	Sample Explorer summary page
2015-03-11 08:34:04.065	admin	SampleExplorer-Summary	Sample Explorer summary page
2015-03-11 08:32:47.361	admin	Loading trial analysis	392401445
2015-03-11 08:32:38.069	admin	Login	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_2) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/41.0.2272.89 Safari/537.36
2015-03-10 20:32:33.968	admin	Browse-Delete object	FOLDER: test (FOL:1994430)
2015-03-10 20:32:19.377	admin	Browse-Create object	FOLDER: test (FOL:1994430)
2015-03-10 18:41:44.8	admin	UploadData-Index	Upload Data index page
2015-03-10 18:30:39.646	admin	UploadData-Index	Upload Data index page
2015-03-10 18:29:31.436	admin	Login	Mozilla/5.0 (Windows NT 6.3; WOW64; Trident/7.0; Touch; USPortal; rv:11.0) like Gecko
2015-03-10 17:36:31.401	admin	Login	Mozilla/5.0 (Windows NT 6.3; WOW64; rv:36.0) Gecko/20100101 Firefox/36.0
2015-03-10 17:12:12.532	admin	Login	Mozilla/5.0 (Windows NT 6.3; WOW64; Trident/7.0; Touch; USPortal; rv:11.0) like Gecko
2015-03-10 16:47:55.659	admin	Browse-Search	
2015-03-10 16:47:17.121	admin	Browse-Search	
2015-03-10 16:41:23.92	admin	Browse-Search	
2015-03-10 16:15:51.454	admin	Browse-Search	
2015-03-10 16:15:47.825	admin	Browse-Search	
2015-03-10 16:15:09.225	admin	Browse-Search	

Rclient – Programmatic database querying



The screenshot shows the RStudio interface with the following components:

- Environment:** Shows 3 observations of 3 variables.
- Data:** Lists 'data' (620 obs. of 590 va...) and 'studies' (3 obs. of 3 variab...).
- Values:** Lists 'transma...' and 'Environment'.
- Files, Plots, Packages, Help, Viewer:** Navigation and toolbars.
- Console:** Shows the execution of R code to connect to the Transmart API and retrieve the 'studies' data.
- Help Panel:** Displays the documentation for the `length()` function.

Console Output:

```

1: package 'RJSONIO' was built under R version 3.1.1
2: package 'RJSONIO' was built under R version 3.1.1
3: package 'RProtoBuf' was built under R version 3.1.1
> connectToTransmart("https://transmart.doc.ic.ac.uk/transmart")
Please visit the following url to authenticate this RClient (enter nothing to cancel):

https://transmart.doc.ic.ac.uk/transmart/oauth/authorize?response_type=code&client_id=api-client&client_secret=api-client&redirect_uri=https://transmart.doc.ic.ac.uk/transmart/oauth/verify

And paste the verifier token here:

Authentication cancelled.
Connection successful.
> UIcSvF
Error: object 'UIcSvF' not found
> connectToTransmart("https://transmart.doc.ic.ac.uk/transmart")
Please visit the following url to authenticate this RClient (enter nothing to cancel):

https://transmart.doc.ic.ac.uk/transmart/oauth/authorize?response_type=code&client_id=api-client&client_secret=api-client&redirect_uri=https://transmart.doc.ic.ac.uk/transmart/oauth/verify

And paste the verifier token here:
UIcSvF
Authentication completed.
Connection successful.
> studies <- getStudies()
> View(studies)
>
  
```

Help Panel: Length of an Object

Description

Get or set the length of vectors (including lists) and factors, and of any other R object for which a method has been defined.

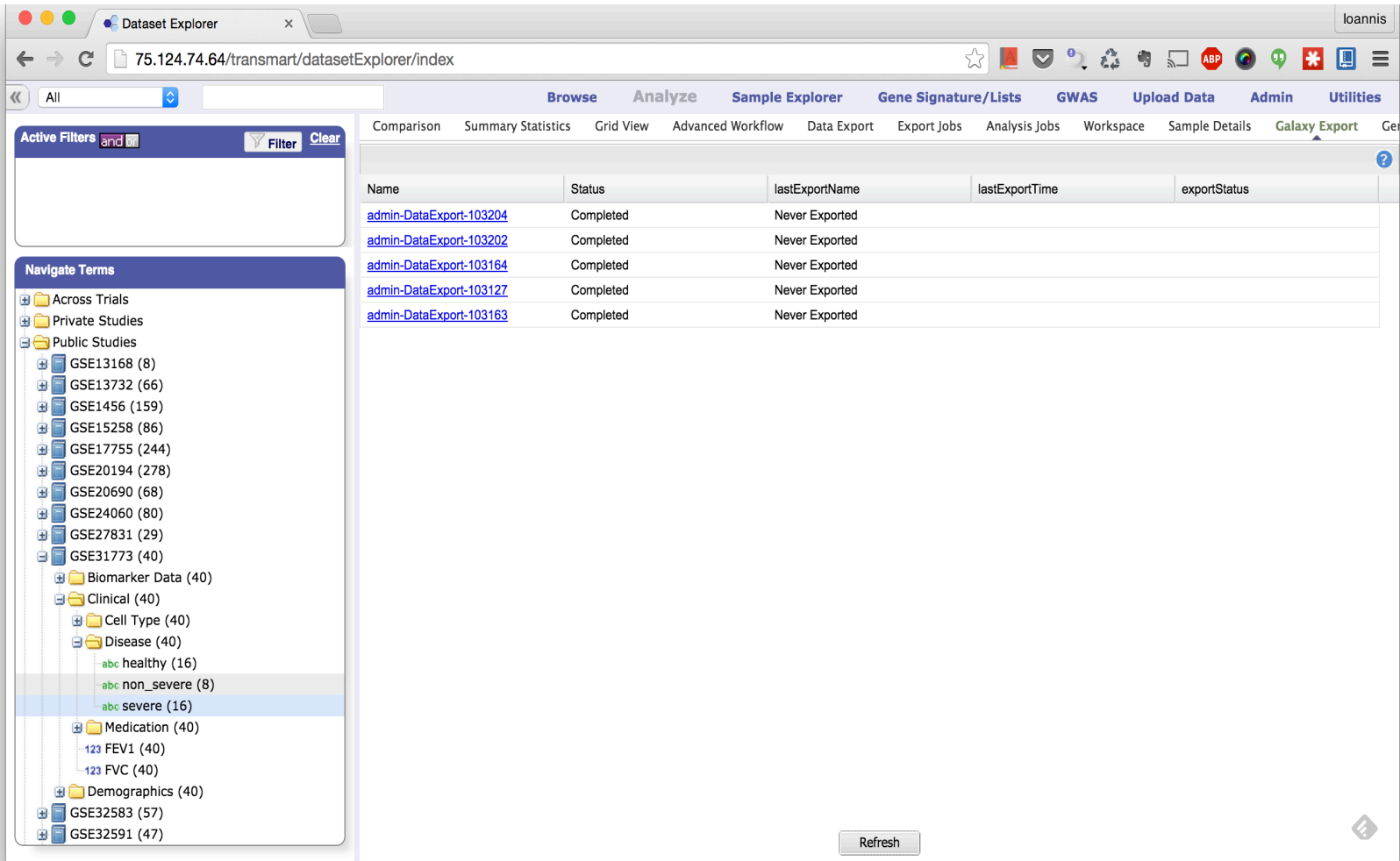
Usage

```
length(x)
length(x) <- value
```

Arguments

`x` an R object. For replacement, a vector or factor.

Galaxy Plugin – Connection to Galaxy and Advanced workflow development



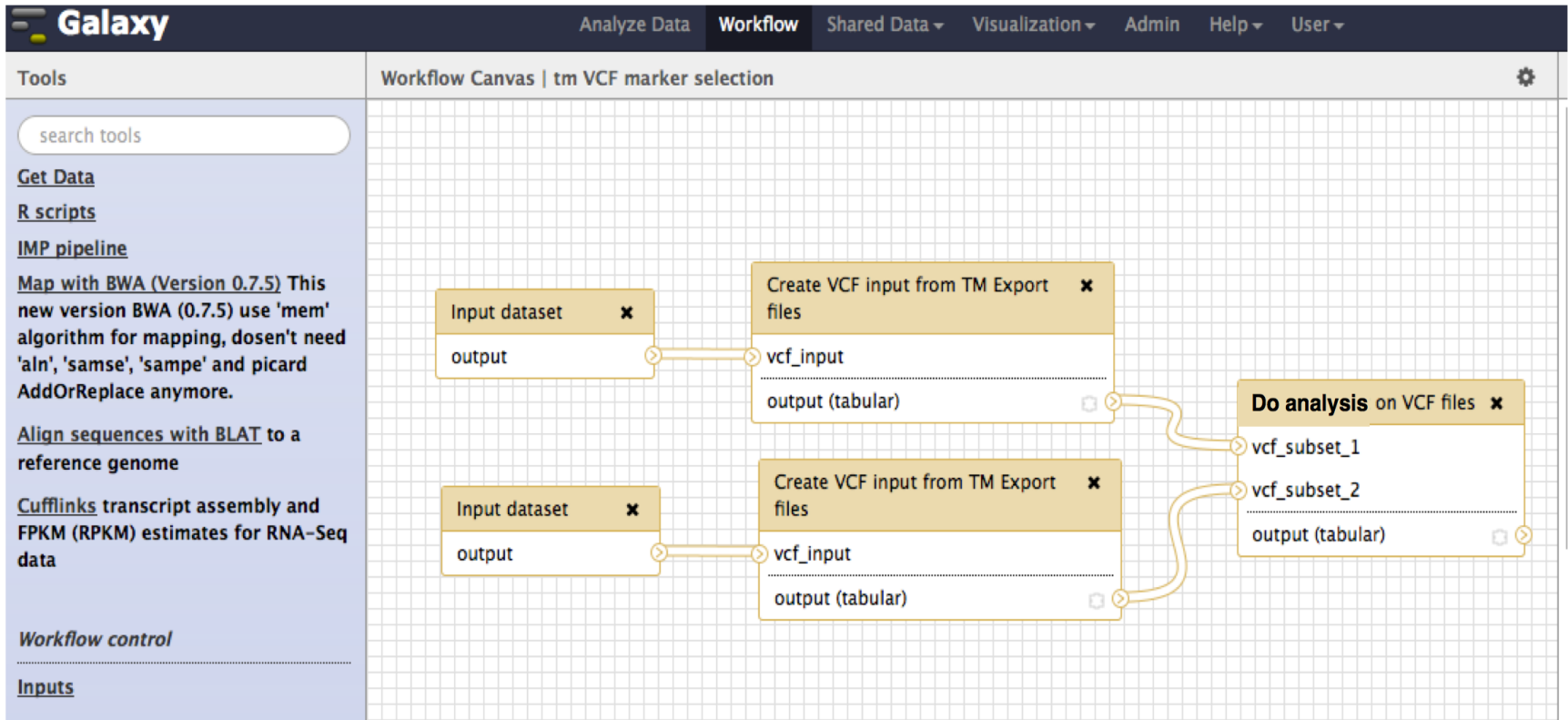
The screenshot shows the Dataset Explorer web interface. The browser address bar displays `75.124.74.64/transmart/datasetExplorer/index`. The interface includes a navigation menu with options like **Browse**, **Analyze**, **Sample Explorer**, **Gene Signature/Lists**, **GWAS**, **Upload Data**, **Admin**, and **Utilities**. Below the menu, there are tabs for **Comparison**, **Summary Statistics**, **Grid View**, **Advanced Workflow**, **Data Export**, **Export Jobs**, **Analysis Jobs**, **Workspace**, **Sample Details**, and **Galaxy Export**. The main content area features a table with the following data:

Name	Status	lastExportName	lastExportTime	exportStatus
admin-DataExport-103204	Completed	Never Exported		
admin-DataExport-103202	Completed	Never Exported		
admin-DataExport-103164	Completed	Never Exported		
admin-DataExport-103127	Completed	Never Exported		
admin-DataExport-103163	Completed	Never Exported		

On the left side, there is a **Navigate Terms** panel with a tree view of categories and sub-categories, including **Public Studies** (with various GSE IDs), **Biomarker Data**, **Clinical**, **Cell Type**, **Disease** (with sub-categories like **healthy**, **non_severe**, **severe**), **Medication**, and **Demographics**.

A **Refresh** button is located at the bottom right of the interface.

Galaxy Plugin – Connection to Galaxy and Advanced workflow development

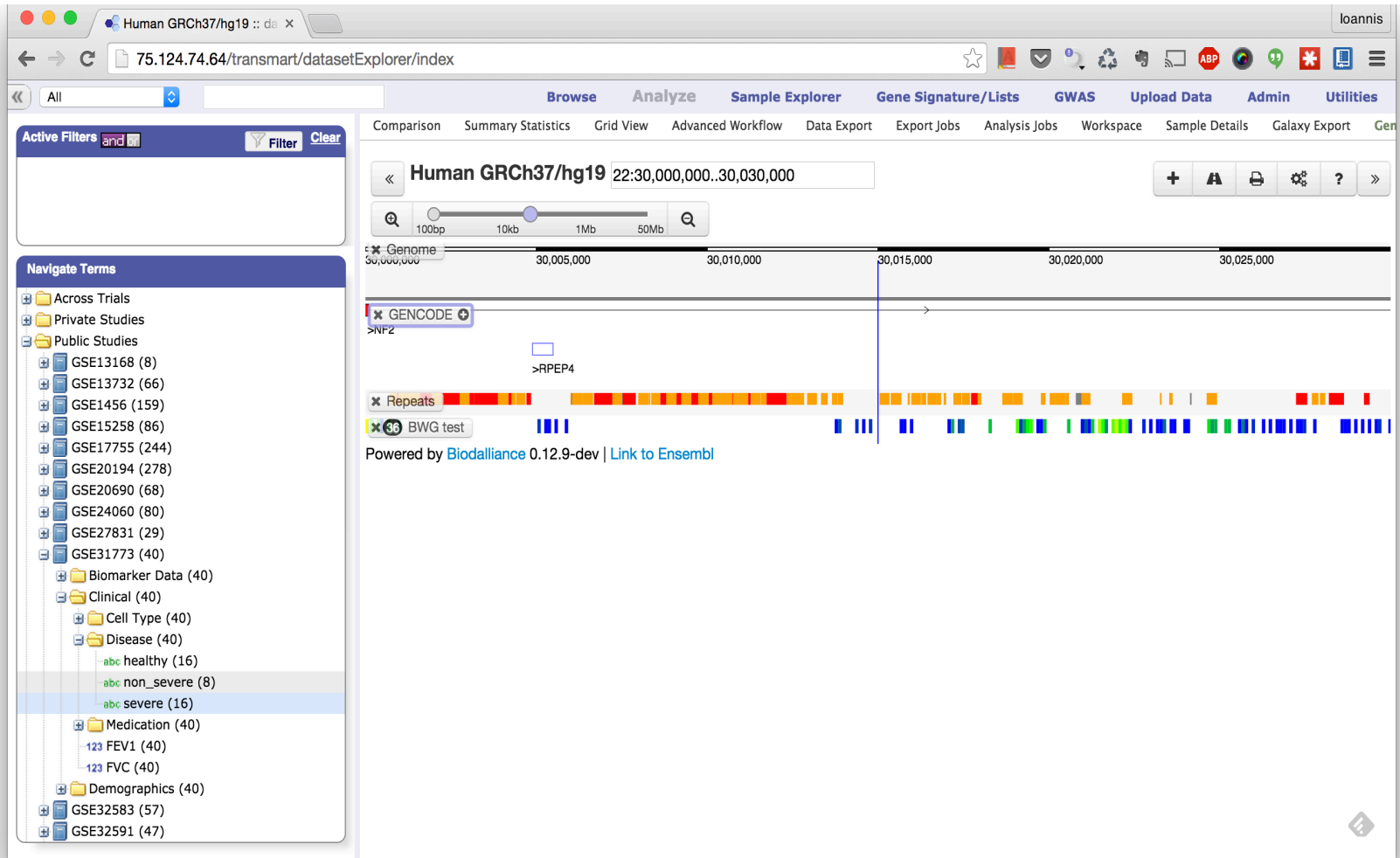


The screenshot displays the Galaxy interface with the 'Workflow Canvas' for a workflow titled 'tm VCF marker selection'. The interface includes a top navigation bar with 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User' menus. On the left, a 'Tools' sidebar lists categories like 'Get Data', 'R scripts', and 'IMP pipeline', with a search bar and a 'Map with BWA (Version 0.7.5)' tool description. The main canvas shows a workflow with three steps:

- Input dataset** (yellow box) with an 'output' port.
- Create VCF input from TM Export files** (yellow box) with a 'vcf_input' port connected to the 'output' of the first 'Input dataset' and an 'output (tabular)' port.
- Do analysis on VCF files** (yellow box) with two 'vcf_subset_1' and 'vcf_subset_2' ports connected to the 'output (tabular)' ports of both 'Create VCF input' steps, and an 'output (tabular)' port.

```
graph LR; A[Input dataset] --> B[Create VCF input from TM Export files]; C[Input dataset] --> D[Create VCF input from TM Export files]; B --> E[Do analysis on VCF files]; D --> E;
```


Genome Browser (Dalliance – HTML 5) – for omics data visualisation



The screenshot displays the eTRIKS Genome Browser interface. The browser window title is "Human GRCh37/hg19 :: da x" and the address bar shows "75.124.74.64/transmart/datasetExplorer/index". The user is logged in as "loannis".

The interface includes a navigation menu with tabs: "Browse", "Analyze", "Sample Explorer", "Gene Signature/Lists", "GWAS", "Upload Data", "Admin", and "Utilities". Below this is a sub-menu with options: "Comparison", "Summary Statistics", "Grid View", "Advanced Workflow", "Data Export", "Export Jobs", "Analysis Jobs", "Workspace", "Sample Details", and "Galaxy Export".

The main content area shows a genomic track for "Human GRCh37/hg19" at coordinates 22:30,000,000..30,030,000. A zoom slider is set to 100bp. The track displays several layers:

- Genome**: A black line representing the reference genome.
- GENCODE**: A track showing gene models, including a gene labeled ">RPEP4".
- Repeats**: A track showing repetitive elements in red and yellow.
- BWG test**: A track showing blue vertical bars representing data points.

At the bottom of the track, it says "Powered by Biodalliance 0.12.9-dev | [Link to Ensembl](#)".

On the left side, there is a "Navigate Terms" panel with a tree view of categories and their associated study counts:

- Across Trials
- Private Studies
- Public Studies
 - GSE13168 (8)
 - GSE13732 (66)
 - GSE1456 (159)
 - GSE15258 (86)
 - GSE17755 (244)
 - GSE20194 (278)
 - GSE20690 (68)
 - GSE24060 (80)
 - GSE27831 (29)
 - GSE31773 (40)
- Biomarker Data (40)
- Clinical (40)
- Cell Type (40)
- Disease (40)
 - abc healthy (16)
 - abc non_severe (8)
 - abc severe (16)
- Medication (40)
 - 123 FEV1 (40)
 - 123 FVC (40)
- Demographics (40)
- GSE32583 (57)
- GSE32591 (47)