



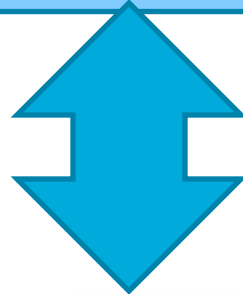
tranSMART

December 2010

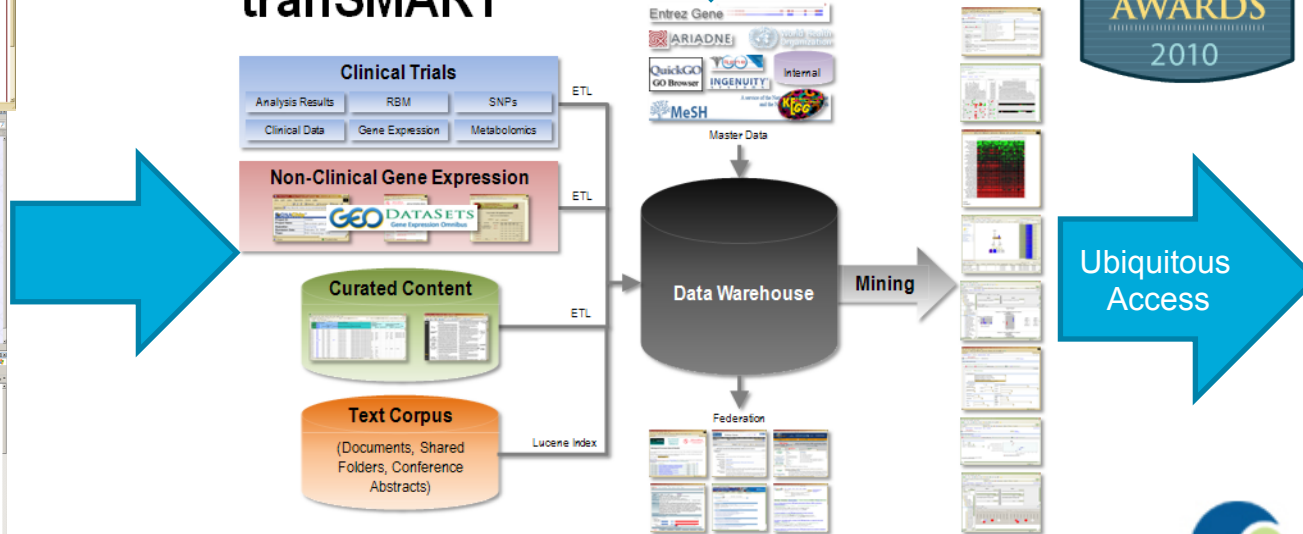
Aim

- Focus - Translational Research
- Sampling of questions to be answered
 - what is the correlation between animal models and human data?
 - what is the best biomarker strategy for a given compound ?
 - what is the best indication for a given compound?
 - how can a disease stratified based on clinical data?
 - is there support for a target of interest based on clinical data?

Precompetitive Sharing



tranSMART

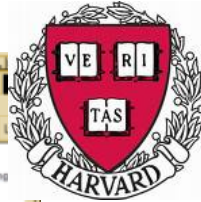


CIO100



- Indication selection
- Target/pathway hypothesis
- Biomarker hypothesis
- Clinical study design

Academic Partnerships



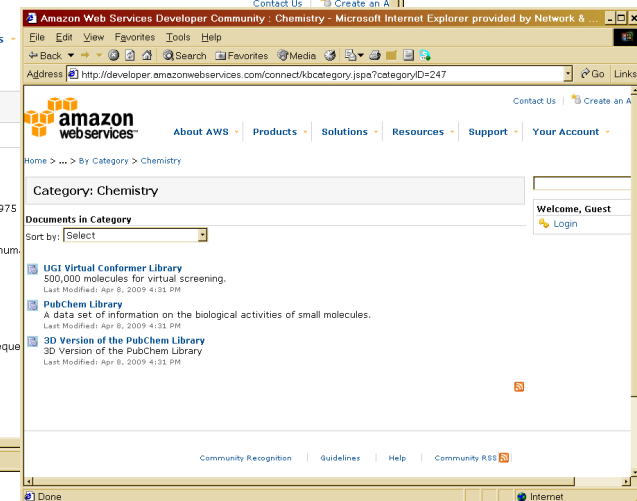
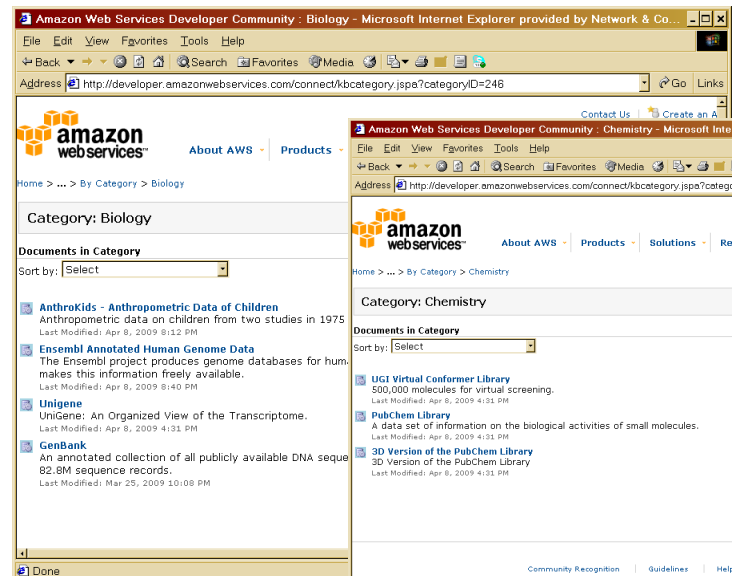
- Based on i2b2 release 1.3

- GDE Consortium is being formed



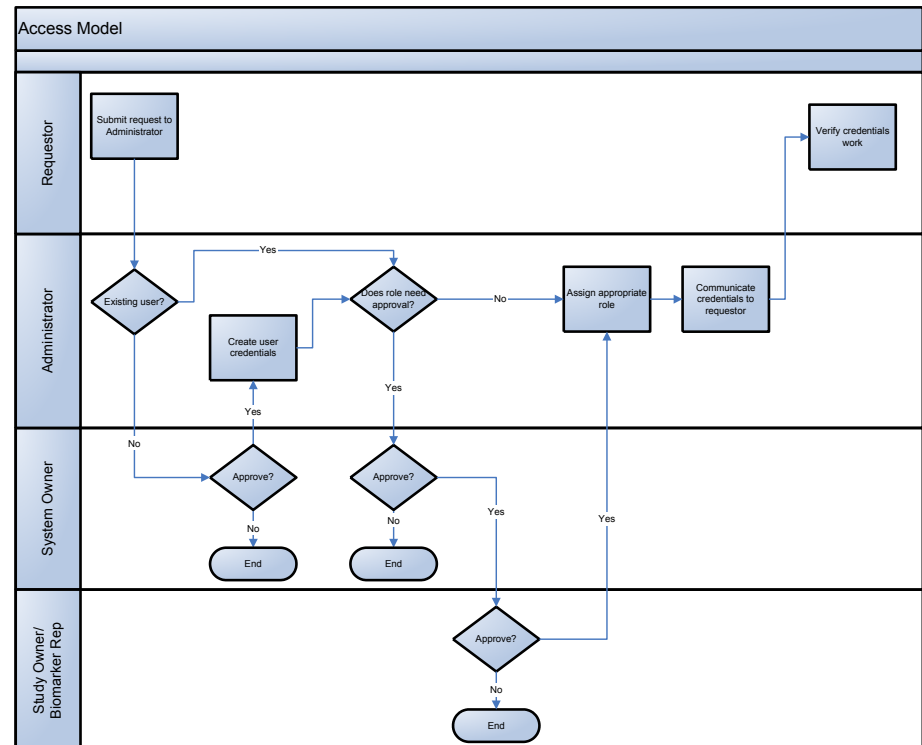
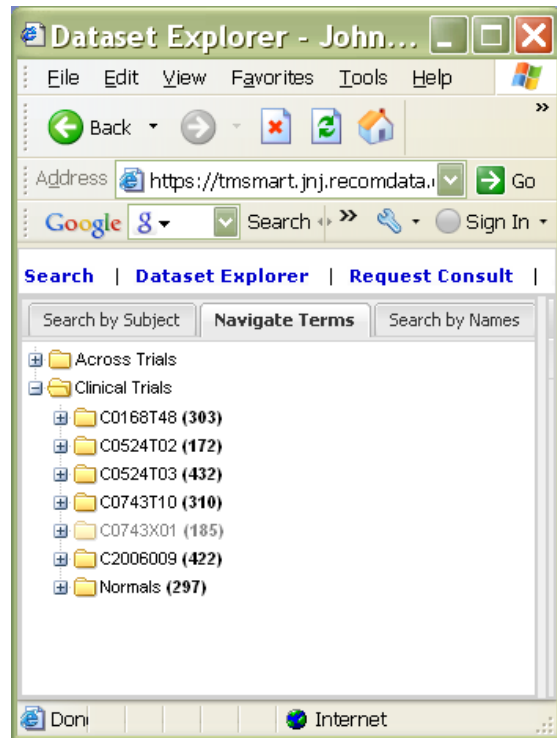
Infrastructure

- Amazon Elastic Compute Cloud
 - Security
 - Cost
 - Scalability
 - HPC
 - Amazon data feeds



Fine-grained Security Model

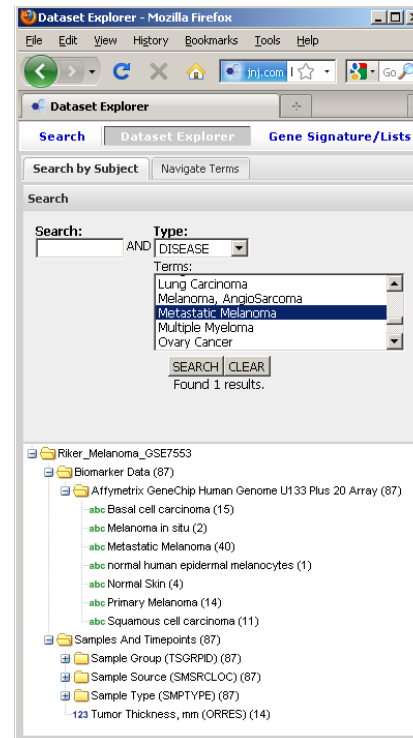
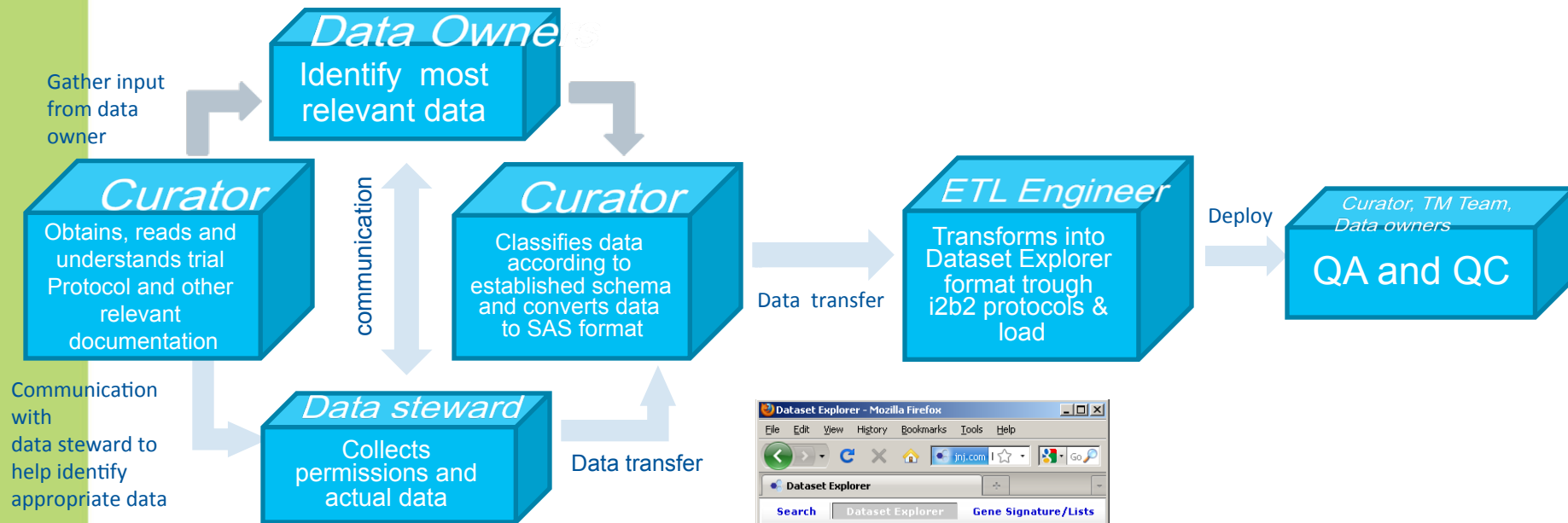
- Strong authentication
- Strong governance – study owners, established process
 - Training
 - Access granting process is established and followed
 - Clinical trial nodes grayed-out if no access is granted



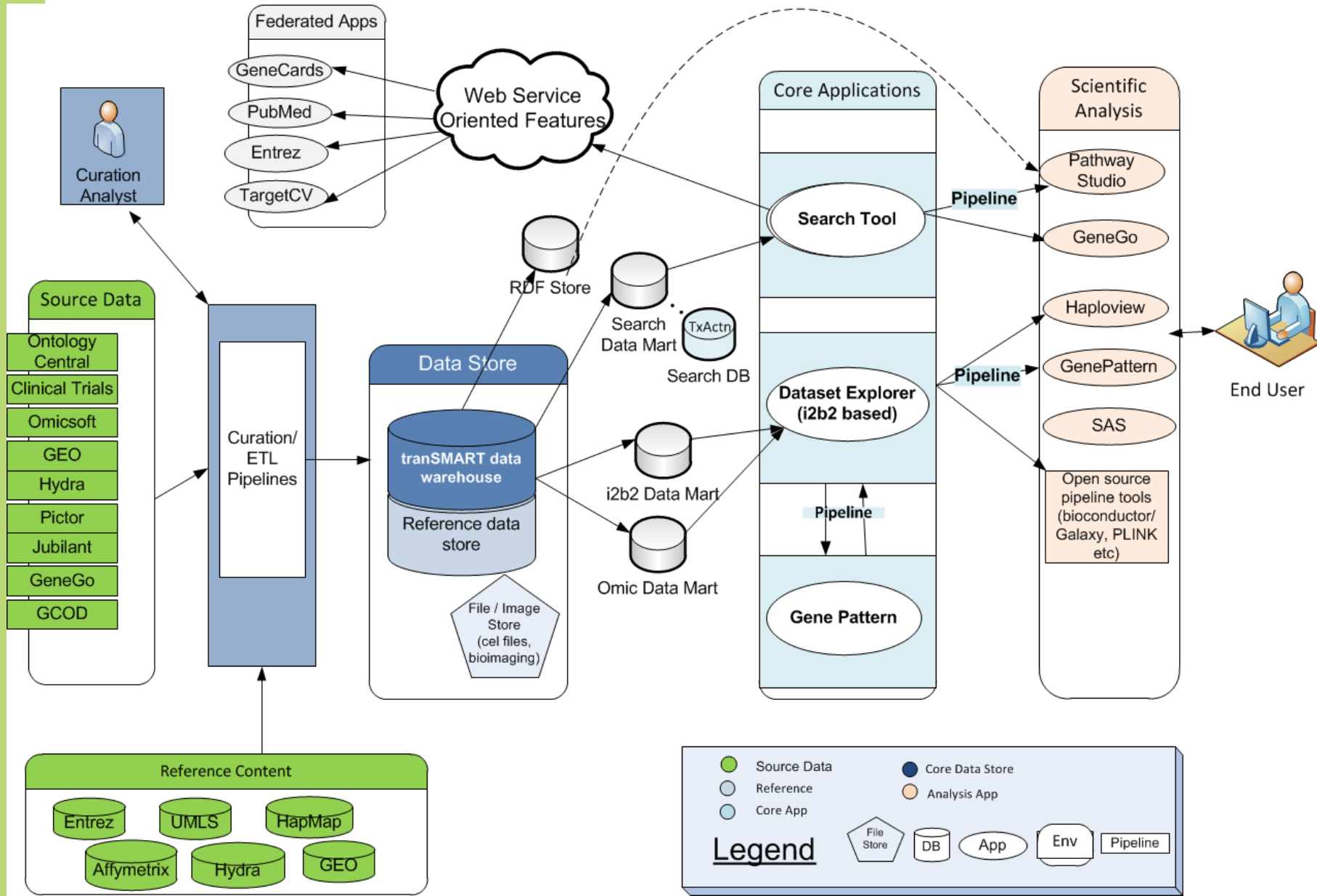
Dictionaries, Ontologies and Master Data

- Gene/protein names and alternatives (Entrez)
- Gene name mapping service (Affy ID, SNP, protein)
- Pathways (GO, KEGG, GeneGo, Ingenuity, Ariadne, MSigDB)
- Internal signatures
- Diseases (MeSH, ICD10)
- Clinical trial observations (MedDRA)
- J&J drugs dictionary
- Clinical trials metadata dictionary
- Cell line dictionary
- Curated inhibitors (CAS id)

Curation Process for Internal Trials



Big Scary Diagram



tranSMART Data Warehouse

- Data

- Structured Data

- Clinical trials, clinical and pre-clinical gene expression, protein profiling (RBM), SNP, PD markers, metabolomics, proteomics
 - In-house – immunology (large and small mol), oncology, psychiatry
 - Public and commercial

- Unstructured Data

- Curated text
 - Text indexing

- Master data, ontologies, vocabularies and metadata
 - Federated sources



- User Interfaces

- Search

- Gene, pathway, disease, compound, trial, and combinations

- Hypothesis testing

- Cohort selection and comparison/analysis

- Hypothesis generation

- Gene signatures

- Analytics workflows



Statistics

- Dataset Explorer
 - 42 internal studies (clinical trials, experimental medicine studies, in vivo and ex vivo experiments)
 - 43 public studies
 - Oncology, immunology, cardiovascular, CNS
- Search
 - 10 internal studies
 - 9,000 GEO/Array Express comparisons
 - 97 DFCI curated studies
 - 100,000 curated biomarker assertions

Platforms and Sources Roadmap

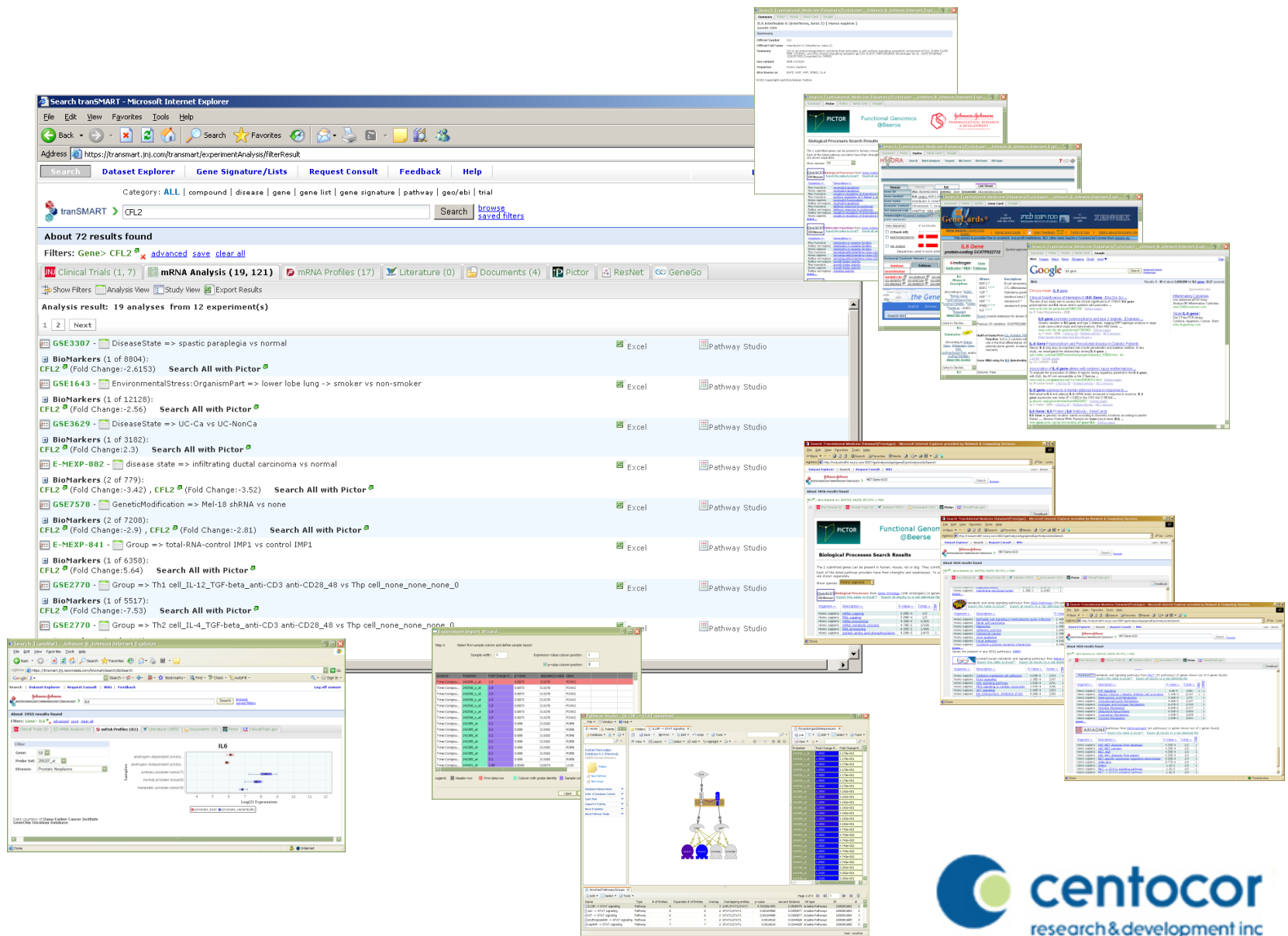
Experimental Platform	Q2 2009	Q3 2009	Q4 2009	Q1 2010	Q2 2010	Q3 2010	Q4 2010	Q1 2011	Q2 2011	Q3 2011
Clinical Trial Data										
Rules-based Medicine										
Affymetrix gene expression										
Candidate SNP										
OmicSoft analyzed mRNA										
ELISA										
Flow cytometry										
Cell counts										
Metabolomics										
Proteomics										
MSD										
Luminex										
MS counts										
Illumina gene expression										
SNP Chip/CNV										
ex vivo										
Histology										
in vivo										
miRNA										
CTC										
Full genome sequencing										
RNA-seq										
PK/Tox										
Epigenetics										
Phosphorylation										
Lipidomics										

Experimental Data Analysis	Q2 2009	Q3 2009	Q4 2009	Q1 2010	Q2 2010	Q3 2010	Q4 2010	Q1 2011	Q2 2011	Q3 2011
Jubilant Oncology										
Pathways										
Beerse preclinical mRNA pipeline										
ClinicalTrials.gov (Link)										
Hydra (Link)										
Entrez (Link)										
GeneCards (Link)										
Google Scholar (Link)										
OmicSoft public mRNA (GEO, EBI)										
La Jolla preclinical mRNA										
Dana-Farber Ct oncology mRNA										
Conference abstracts (indexing)										
DIP folder (indexing)										
Jubilant Asthma										
Biomarker folder (indexing)										
Resnet (text mining)										
GeneGo Prostate Cancer Content										
ClinicalTrials.gov (text mining)										
GeneGo Disease Pages										
dbSNP										
P53 mutation DB										
COSMIC Link										
J&J Biobank Link										
Histology (Aperio) Link										
GVK Bio biomarkers										
Broad Connectivity Map										
Cancer Gene Data Curation										
Prous Integrity										
GeneGo asthma maps										
AsthmaMap										
OASIS										
Cellminer										

Use Cases

- Knowledge Management/Search
- Hypothesis Testing
- Hypothesis Generation

Search



Content in Semantic Networks

jubilant sample article.pdf - Adobe Reader

File Edit View Document Tools Window Help

68.1%

Find

Handwritten notes:

- Run in MET*
- Body weight*
- During Phase 1 (NOT VITAL TO INCLUDE ALL DETAILS)*
- DETAILED CT SCAN ID TAPER (NOT VITAL TO INCLUDE DETAILS)*
- ALLOWED MEDICATION*

Study design

This was a double-blind, placebo-controlled, multicenter, parallel-group trial. On entry into the study, each patient was notified that he or she currently presented ECR to an equivalent dose of BOP. At the end of the second week of the 4- to 6-week run-in period, the dose of BOP was adjusted upward or downward to maintain previous asthma control. The patients were monitored to ensure both the presence of asthma symptoms at levels acceptable to the patients and investigators and patient safety. A stable BOP dose (baseline dose) was maintained for 4 weeks prior to randomization. A mean total symptom score over the last 2 weeks of run-in of 2 or more (range, 0-9) was required for the patient to be eligible for randomization to metformin or placebo administered subcutaneously as baseline (week 0). The dose and dosing intervals were based on the serum half-life of metformin; efficacy results from previous studies, and the subject's body weight and baseline serum IgE. This amount for each patient a dosage approximately equal to 0.016 mg/kg IgE (57mg) per 4 weeks. Subjects thus received 150 mg or 300 mg every 4 weeks or 225 mg, 300 mg, or 375 mg every 2 weeks. The baseline BOP dose was maintained unchanged during the 4-month stable control phase (weeks 1-16). During the subsequent 3-month steroid reduction phase (weeks 16-28), the dose of BOP was reduced by approximately 25% of the baseline dose every 2 weeks for 8 weeks until discontinuation or worsening of asthma symptoms, defined by any one of the following criteria: requirement of an unscheduled physician visit; peak expiratory flow (PEF) < 20% of personal best; decrease in morning PEF > 20% on 22 of 3 consecutive days in comparison with the week before the steroid reduction phase; decrease in PEF, of > 20% from the value at the beginning of the steroid reduction phase; 30% increase in 24-hour observed usage on 22 of 3 consecutive days (exceeding 8 puffs/day); or 22 of 3 consecutive nights with awakening for asthma symptoms requiring rescue medication. If worsening occurred, BOP was increased by 25% and the patient was reevaluated for tapering after a week of improvement. For the final 4 weeks of the steroid reduction phase, each subject was maintained on the lowest effective dose of BOP that did not result in worsening of symptoms.

Concomitant medication

Bronchodilators, 2 puffs (90 µg/act) as needed, (maximum, 8 puffs daily) was allowed for symptomatic episodes of wheezing. Patients exhibiting this different change (including the need to contact the investigator immediately for further evaluation. Treatment with stable doses of immunotherapy and other nonsteroids

Global assessment of efficacy

At the end of the study (week 28), subjects global treatment effectiveness as measured by asthma, good (marked improvement of anti-asthmatic) or placebo administered subcutaneously as baseline (week 0), or worsening.

Serum IgE levels

Serum samples were taken at baseline at weeks 16 and 24 to determine IgE concentration. Serum IgE in baseline (pre-randomization) samples was used to determine the mean expressed flow IgE. Serum concentration of total IgE (free IgE and bound IgE) were measured using a standard IgE assay. Samples were assayed for total IgE (free IgE and bound IgE) were measured using a standard IgE assay. Samples were assayed for total IgE (free IgE and bound IgE) were measured using a standard IgE assay.

Safety and tolerability

Adverse events were recorded and categorized as mild, moderate, or severe. Mild adverse events were defined as those that did not interfere with the study. Moderate adverse events were defined as those that interfered with the study. Severe adverse events were defined as those that were life-threatening or resulted in death.

Statistical analysis

The following variables were analyzed through use of the generalized Cochran-Mantel-Haenszel (van der Tweel) test, stratified by



Microsoft Excel - Data, Disease Breakup.xls							
File Edit View Insert Format Tools Data Window Help							
D1D2 A B C D E F G							
General Information		Disease Information					
Record #	Gene	Molecule Type	Variant Of	Disease (Common Name)	Disease Site (ICD10)		
133	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
134	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
135	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
136	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
137	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
138	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
139	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
140	770	MET	4233	mRNA	MET	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid
141	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
142	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
143	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
144	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
145	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
146	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
147	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
148	MET	4233	mRNA	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
149	MET	4233	mRNA	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
150	MET	4233	mRNA	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
151	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
152	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
153	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
154	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
155	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
156	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
157	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
158	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
159	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
160	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
161	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
162	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
163	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
164	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
165	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
166	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
167	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
168	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
169	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
170	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
171	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
172	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
173	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
174	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
175	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
176	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
177	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
178	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
179	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
180	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
181	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
182	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
183	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
184	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
185	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
186	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
187	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
188	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
189	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
190	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
191	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
192	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
193	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
194	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
195	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
196	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
197	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
198	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
199	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
200	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
201	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
202	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
203	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
204	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
205	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
206	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
207	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
208	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
209	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
210	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
211	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
212	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
213	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
214	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
215	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
216	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
217	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
218	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
219	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
220	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
221	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
222	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
223	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
224	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
225	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
226	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
227	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
228	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
229	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
230	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
231	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
232	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
233	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
234	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
235	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
236	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
237	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
238	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
239	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
240	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
241	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
242	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
243	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
244	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
245	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
246	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
247	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
248	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
249	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
250	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
251	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
252	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
253	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
254	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
255	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
256	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
257	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
258	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
259	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
260	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
261	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
262	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
263	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
264	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
265	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
266	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
267	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
268	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
269	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
270	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
271	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
272	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
273	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
274	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
275	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
276	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
277	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
278	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
279	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
280	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
281	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
282	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
283	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
284	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
285	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
286	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
287	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
288	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
289	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
290	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
291	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
292	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
293	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
294	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
295	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
296	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
297	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
298	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
299	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
300	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
301	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
302	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
303	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
304	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
305	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
306	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
307	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
308	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
309	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
310	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
311	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
312	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
313	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
314	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
315	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
316	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
317	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
318	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
319	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
320	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
321	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
322	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
323	MET	4233	Protein	Colorectal Cancer	Malignant Neoplasm of Rectosigmoid		
32							

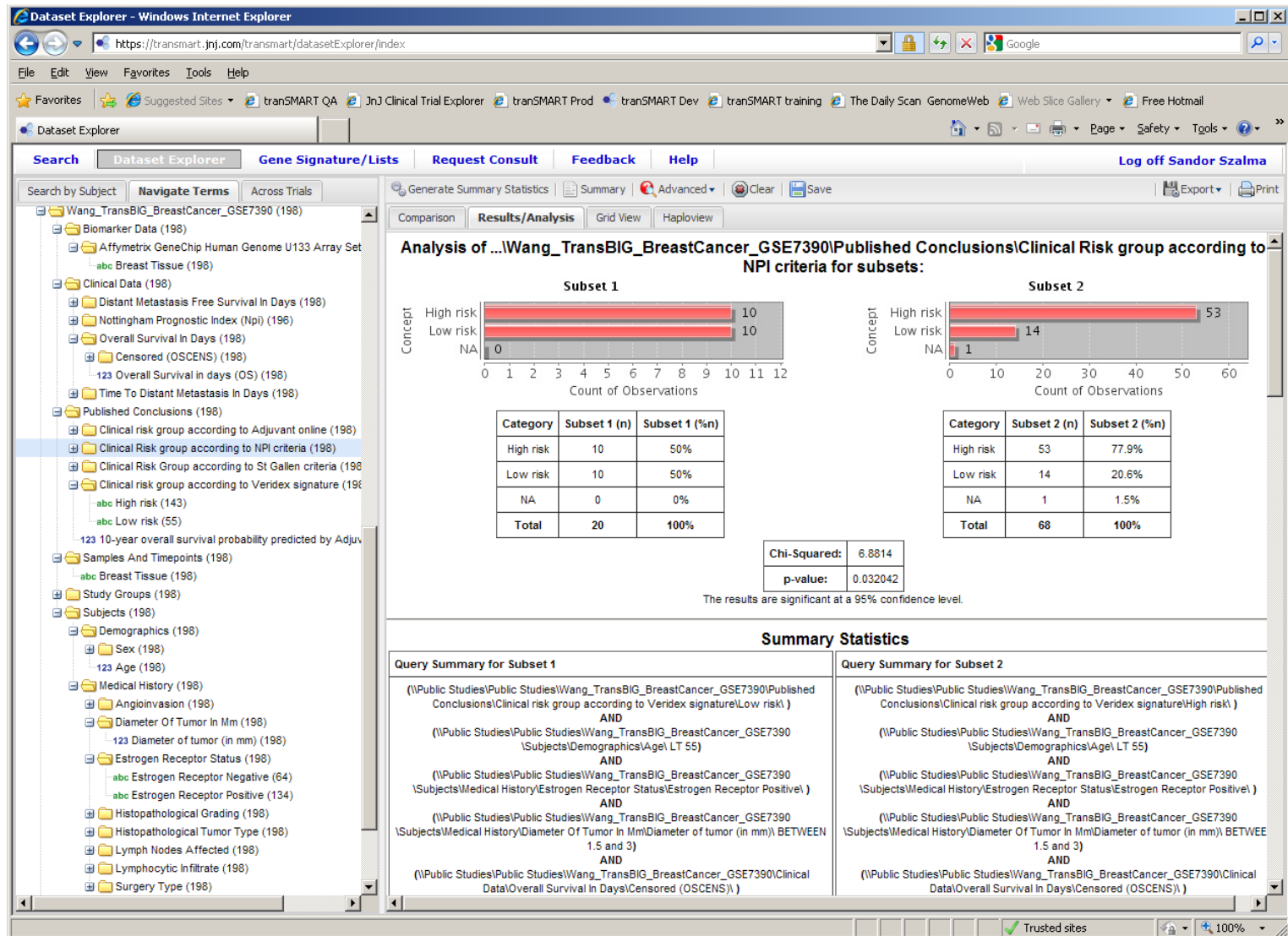
Hypothesis Testing – Flexible Queries

The screenshot displays the 'Dataset Explorer' web application running in a Windows Internet Explorer browser. The address bar shows the URL: <https://transmart.jnj.com/transmart/datasetExplorer/index>. The browser's toolbar includes standard navigation buttons and a search bar.

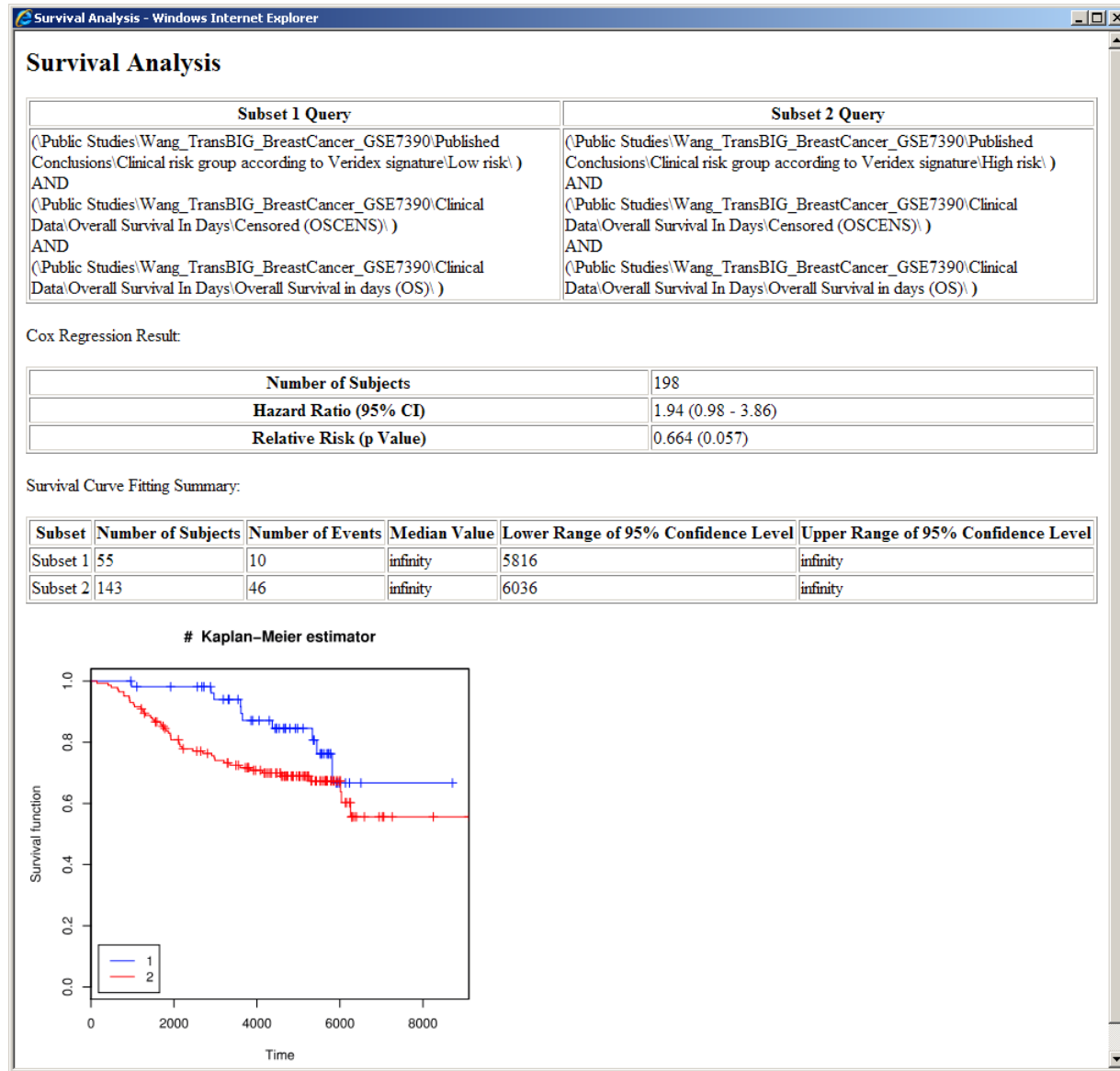
The application interface features a top navigation bar with tabs: **Search**, **Dataset Explorer** (active), **Gene Signature/Lists**, **Request Consult**, **Feedback**, and **Help**. A 'Log off Sandor Szalma' link is located on the right. Below the navigation bar, there are tabs for 'Search by Subject', 'Navigate Terms', and 'Across Trials'. The 'Search by Subject' tab is active, showing a hierarchical tree of datasets on the left. The tree includes categories like 'Wang_TransBIG_BreastCancer_GSE7390 (198)', 'Biomarker Data (198)', 'Clinical Data (198)', 'Published Conclusions (198)', 'Samples And Timepoints (198)', 'Study Groups (198)', 'Subjects (198)', 'Demographics (198)', 'Medical History (198)', 'Histopathological Grading (198)', 'Histopathological Tumor Type (198)', 'Lymph Nodes Affected (198)', 'Lymphocytic Infiltrate (198)', and 'Surgery Type (198)'. The '123 Overall Survival in days (OS) (198)' dataset is selected.

The main content area is divided into two panels: 'Subset 1' and 'Subset 2'. Each panel contains a query builder interface with text input fields and buttons for 'Exclude' and 'X'. The queries are structured using logical operators (AND) to combine conditions. For example, in Subset 1, the query is:
...Low risk
AND
...Age < 55
AND
...Estrogen Receptor Positive
AND
...Diameter of tumor (in mm) between 1.5 and 3
AND
...Censored (OSCEMS)
AND
...Overall Survival in days (OS)
AND
...
In Subset 2, the query is:
...High risk
AND
...Age < 55
AND
...Estrogen Receptor Positive
AND
...Diameter of tumor (in mm) between 1.5 and 3
AND
...Censored (OSCEMS)
AND
...Overall Survival in days (OS)
AND
...
The bottom of the browser window shows the status bar with 'Trusted sites' and a zoom level of 100%.

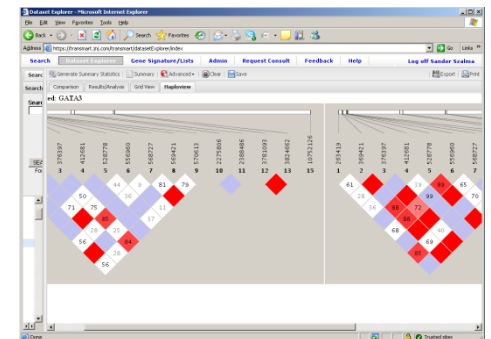
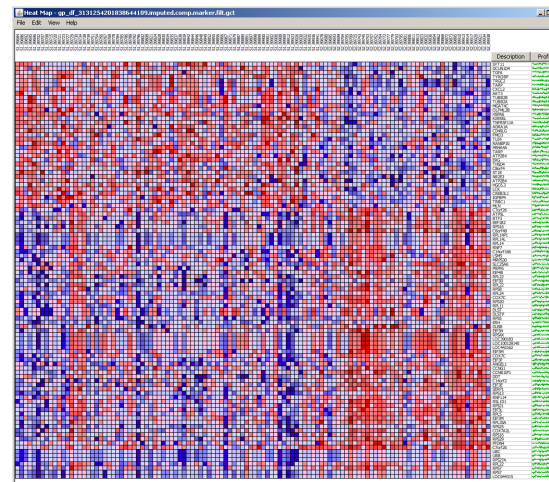
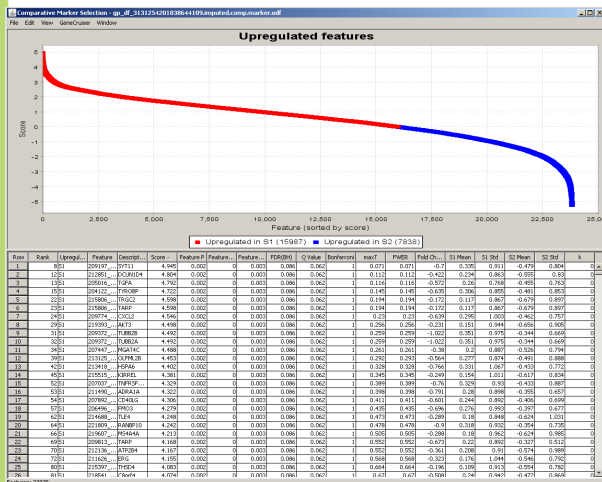
Hypothesis Testing – Simple Statistics



Hypothesis Testing – Advanced Statistics



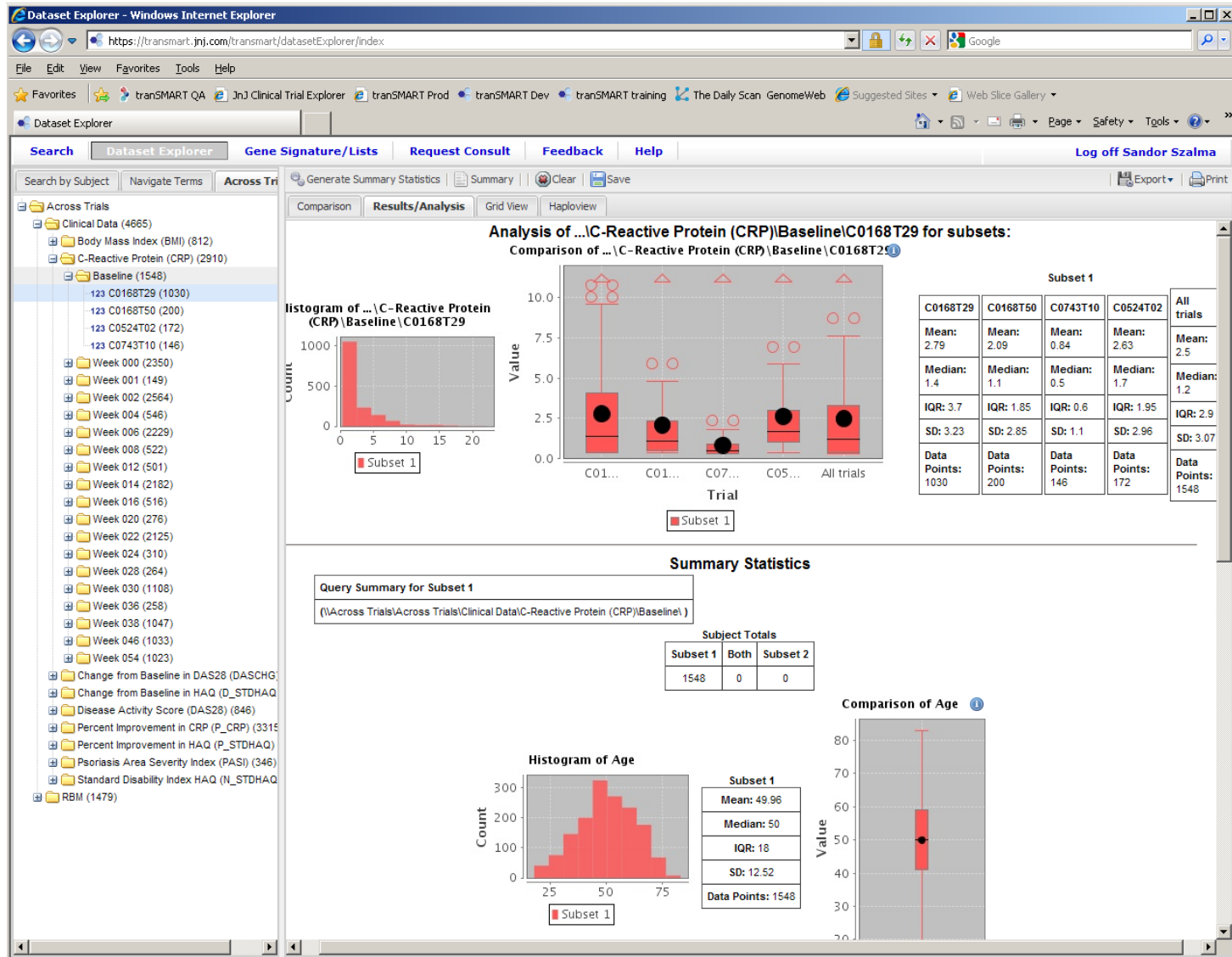
Heatmap visualization of gene expression data. The dendrogram on the left indicates hierarchical clustering of samples, and the dendrogram on the top indicates hierarchical clustering of genes. The color scale represents expression levels, ranging from blue (low) to red (high).



The screenshot shows the Dataset Explorer web application. The browser address bar indicates the URL: <https://transmartdev.jnj.com/transmart/datasetExplorer/index>. The application has a top navigation bar with links: Search, Dataset Explorer, Gene Signature/Lists, Admin, Request Consult, Feedback, Help, and Log off Sandor Szalma. Below this is a search bar and a 'Navigate Terms' section. The left sidebar displays a tree of datasets, with 'Veridex_LungCancer_2003 (63)' selected. The main content area shows a 'Comparison' view with two subsets, 'Subset 1' and 'Subset 2', each containing a list of tissues and a 'Cancer Group' section. The 'Comparison' view includes buttons for 'Exclude' and 'X' for each item.



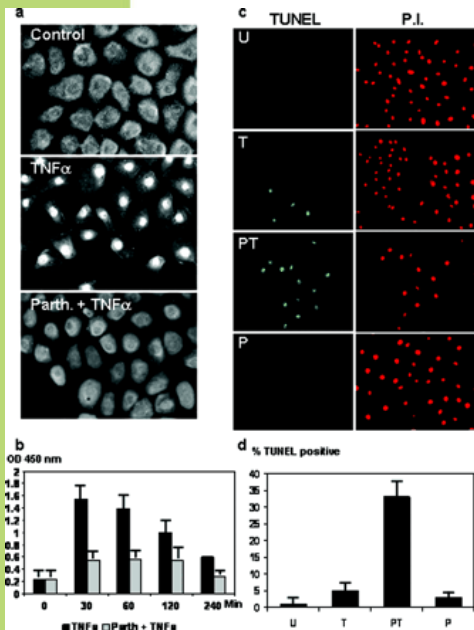
Cross-Study Queries



Indication Hypothesis Generation

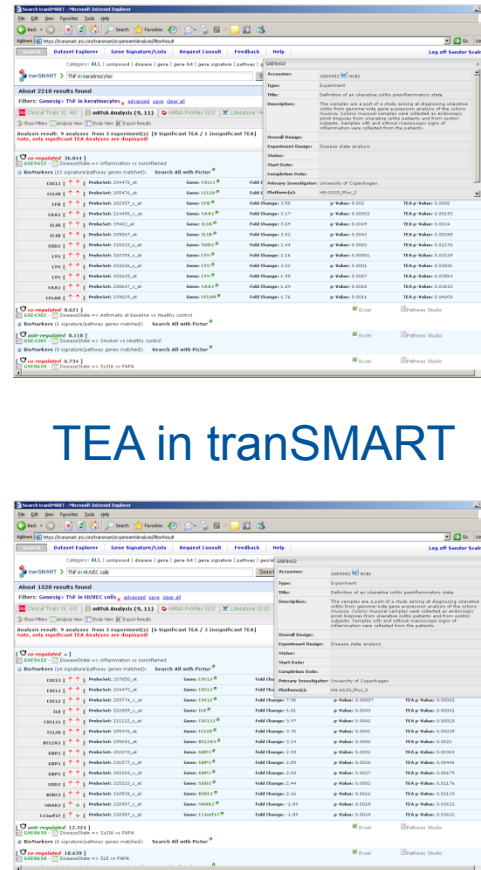
- Disease Profiles Database
 - Well curated
 - Well annotated
 - QCd
 - Database of 9,000 public comparisons from NCBI GEO and EBI Array Express
 - Subset is molecular profiles of disease vs. normal samples
 - Also, proprietary data!
- Drug or Pathway Modulation Signatures
 - In vitro, in vivo or clinical samples
 - Treatment vs. placebo or before vs. after or ...
 - Drug, siRNA, Ab, KO, ...
 - Measure gene expression, protein expression, ...
 - Statistical analysis and derivation of statistically significant list of genes (proteins, ...) with direction

Indication Hypothesis Generation



in vitro
ex vivo
in vivo

TEA in tranSMART



Disease	Keratinocyte	HUVEC	Rat
TNF Stimulation	40	40	40
IBD	40	40	40
Crohn's disease	40	40	40
Ulcerative colitis	40	40	22.7
TNF Stimulation	40	40	40
Crohn's disease	40	40	40
TNF Stimulation	40	40	40
Inflammation	40	40	40
Mycobacterium tuberculosis	40	12.6	
Influenza A	33.4	40	30.9
RA	32.7	40	27.5
TNF Stimulation	32.2	35.6	25.7
Mycobacterium tuberculosis	32.2	33.8	13.1
Ulcerative colitis	30.2	18	19.8
RA	27.9	40	20.2
TNF Stimulation	24.8	23.6	12.1
TNF Stimulation	23.4	22.4	11.4
Infectious Colitis	22.4	36.7	15.6
Dermatomyositis	18.4	20.6	6
Crohn's disease	17.3	34.4	15.3
Osteoarthritis	16.8	13.4	
RA	16.1	14	
Asthma	16.1		
Influenza A	13	24.3	15
Ulcerative colitis	12.4	8	6.2
RA		21.6	19.4
TNF Stimulation		21.3	

Implemented Disruptive Technology

- Open source
- Cloud computing
- Pre-competitive sharing



- Informatics/IT
- Recombinant Data
- Business partners across multiple TAs (Biology, Biomarkers, Translational Medicine, Clinical)
- Academic partners

JOURNAL OF TRANSLATIONAL MEDICINE

IMPACT FACTOR 3.41

Welcome Sandor Szalma (Log off)

BioMed Central home | Journals A-Z | Feedback | Support | My details

Home | Browse articles | Series | Search | Weblinks | Submit article | My JTM | About JTM

Top

Abstract

Background

Methods

Results

Discussion

Methodology

Effective knowledge management in translational medicine

Sándor Szalma¹, Venkata Koka², Tatiana Khasanova² and Eric D Perakslis³

¹ Centocor R&D, Inc. 3210 Merryfield Row, San Diego, CA 92130, USA

² GeneGo, 169 Saxony Road, #104, Encinitas, CA 92024, USA

³ Centocor R&D, Inc. 145 King of Prussia Rd., Radnor, PA 19087, USA

✉ author email ✉ corresponding author email

Journal of Translational Medicine 2010, **8**:68 doi:10.1186/1479-5876-8-68

Highly accessed Open Access

Journal of Translational Medicine
Volume 8

Viewing options:

- Abstract
- Full text
- PDF (3MB)

Associated material:

- Readers' comments
- PubMed record

DISCOVERY

nature publishing group

How Informatics Can Potentiate Precompetitive Open-Source Collaboration to Jump-Start Drug Discovery and Development

ED Perakslis¹, J Van Dam² and S Szalma³

Much enthusiasm and energy are being directed toward open-source software approaches, precompetitive data sharing, and external innovation in the biopharmaceutical industry. At Informatics & Informatics (I&I), we have undertaken aggressive approaches to optimizing productivity and progress to address unmet medical needs. In this article, transMART, a fully-translational data warehouse based on the open-source I&I platform, serves as a case study and the basis for discussing the positive role that informatics can play in accelerating translational research.

UNMET MEDICAL NEEDS It is well documented that productivity in biopharmaceutical R&D, as defined by most measures, has not grown in proportion to the increased investment by industry, academia, and government regulatory agencies¹. A significant challenge is the lack of interoperability of practical results: meaningful biological knowledge, although the problem is being approached, is actually fragmented, mostly disparate (real, digital), and with no regulatory approaches (the US Food and Drug Administration's Clinical Path), these strategies require increased data availability, depth, and interoperability^{2,3}. Furthermore, in drug discovery and development, the decision involving the greatest expense occurs later in the development cycle (such as manufacturing investment decisions and "go/no-go" decisions for phase III trials). Considered translational data are the primary drivers of these decisions, but the urgent requirement for such data remains largely unmet.

There has recently been a great deal of interest in the concept of pre-competitive data sharing in disease biology. The idea here is that a single legal entity can afford to generate and manage the necessary depth of content required for adequate biological study to ensure that data can be shared and managed without devolving the data or eliminating the ability to patent results⁴. The challenge is engaging, with several specific, business models being proposed, and there are indications that this strategy may be more cost effective^{5,6,7}. At I&I, we achieved a 12-fold reduction

R&D INFORMATIC AT I&I Within the Immunology, Oncology, and Biotechnology R&D sections of I&I pharmaceuticals, a novel and innovative approach has been taken to informatics, open-source collaboration, precompetitive data sharing, organizational partnering, and translational research. First, drug discovery and development are treated as a single process and system for optimization (Figure 1). Experimental data are collected and managed in a way that ensures connectivity and utility for key decisions across the entire process. The organizational silos of informatics have been united under a centralized leadership and a flourishing community of practitioners has emerged. Regulatory compliance and overall efficiency have been improved by consistent approaches and increased transparency of data, processes, and policy.

Early on, the I&I Informatics group was assigned the task of developing a knowledge management platform that would provide access to all I&I data as well as to external analysis. Approximately 1 year later, the transMART data warehouse was launched, and it has provided a major step forward in translational capability⁸.

OPEN-SOURCE APPROACH Careful selection of the core technologies is essential. After rigorous evaluation, it was decided to build transMART on the open-source I&I Framework to enable data sharing and partnering (http://www.i2i2.org). Alternatives included long-term coding and several small-scale open-source solutions. The result has been an excellent and robust system, which is an outstanding community of academic partners, and the ability to share and propagate our success (Figure 2). In addition, transMART is the first I&I application to be hosted externally on the Amazon Elastic Compute Cloud (EC2). Cloud computing offers the advantages of low cost, a pay-per-use system, high speed, and ready scalability^{9,10}. At I&I, we achieved a 12-fold reduction

© Informatics & Informatics, Inc., San Diego, CA, USA. Correspondence: ED Perakslis (ed@i2i2.org)

Received: 15 March 2010 Accepted: 20 May 2010 Published online: 20 May 2010

44

CITATION: Szalma S, Koka V, Khasanova T, Perakslis E: How Informatics Can Potentiate Precompetitive Open-Source Collaboration to Jump-Start Drug Discovery and Development. *Journal of Translational Medicine* 2010, **8**:68