Statusbericht WISDOM & ANEURIST

Fraunhofer

SCAI

Institute
Algorithms and
Scientific Computing

Martin Hofmann-Apitius
Abteilung Bioinformatik
Fraunhofer Institut für Algorithmen und
Wissenschaftliches Rechnen (SCAI)
Sankt Augustin

Fraunhofer-Campus Schloss Birlinghoven

Institutes

- Algorithms and Scientific Computing SCAI
- Intelligent Analysis and Information Systems IAIS
- Applied Information Technology FIT

600 Scientists, 200 Students
Linked to Universities Bonn, Aachen and Cologne

Fraunhofer Institute Algorithms and Scientific Computing

b-it

Beteiligung von Fraunhofer SCAI an Grid-Projekten mit biomedizinischen Inhalten bzw. Szenarios

WISDOM (virtual screening)

@neurIST (text- and data mining)

BRIDGE (EU-China Grid; virtual screening)

SIMDAT (Pharma activity; data integration)

Phosphorus (optical testbed; virtual screening scenario)

SCAI
Fraunhofer Institute
Algorithms and
Scientific Computing

Virtual Screening on Grids: WISDOM et al.

Fraunhofer Institute Algorithms and Scientific Computing

Virtual High Throughput Screening (vHTS) on an Optical High Speed Testbed

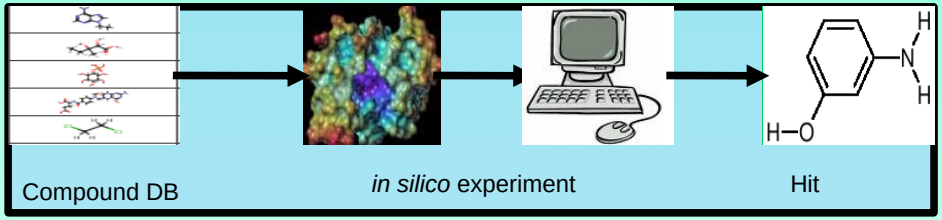
Martin Hofmann-Apitius
for Shahid Mohammad
Bonn-Aachen International Center for Information Technology (B-IT)
and
Fraunhofer Institute SCAI,
Department of Bioinformatics

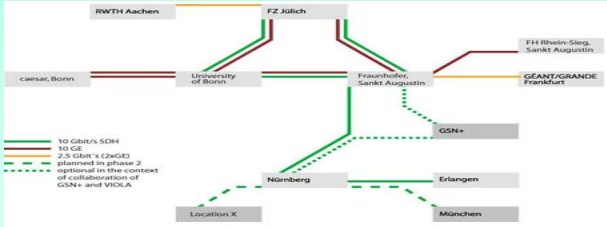
Fraunhofer Institute Algorithms and Scientific Computing

HEALTHGRID 2008 CONFERENCE CHICAGO JUNE 2-4, 2008

Goals

- Deploy an *in silico* vHTS workflow
- Test the high bandwidth optical Grid environment (VIOLA)
- Testing the UNICORE Grid middleware for advanced bioinformatics applications
- Design an efficient compound library against the advanced malarial targets.





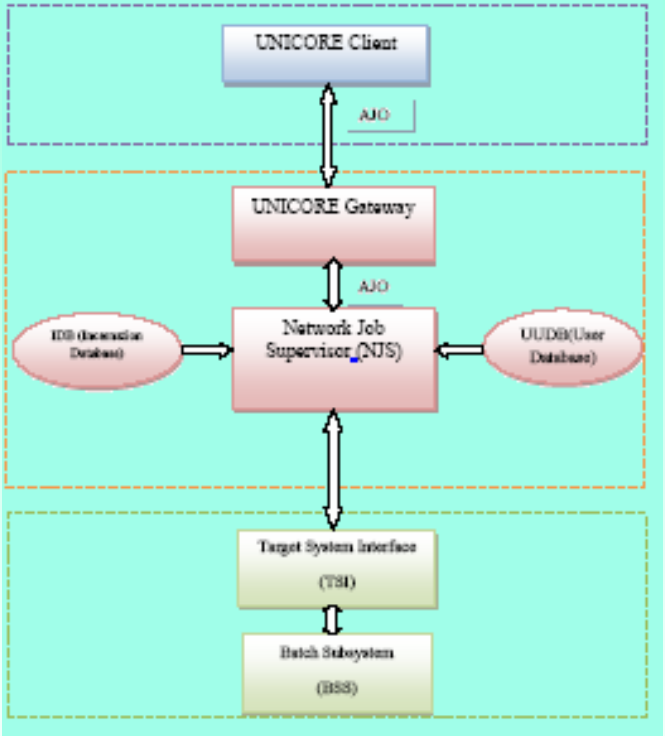
	203.24	C9H17NO4	(2-acetoxy-3-carboxylato-
	156.14	C7H8O4	5,6-dihydroxycyclohexa-1,3-
	75.11	C3H9NO	1-aminopropan-2-ol
	169.07	C3H8NO5P	(3-amino-2-oxo-propoxy)
	202.55	C6H3ClN2O4	1-chloro-2,4-dinitro-benzene
	163.18	C7H9N5	9-ethyl-9H-purin-6-amine
	148.16	C6H12O4	2,3-dihydroxy-3-methyl-pentano
	260.14	C6H13O9P	(2,3,4,5,6-pentahydroxycycl
	473.44	C20H23N7O7	2-[4-[(2-amino-4-hydroxy-
	98.96	C2H4Cl2	1,2-dichloroethan

SCAI
Fraunhofer Institute Algorithms and Scientific Computing

HEALTHGRID 2008 CONFERENCE CHICAGO
JUNE 2-4, 2008

UNICORE

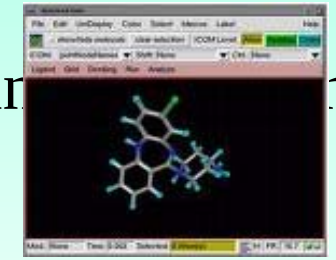
- Seamless, secure and intuitive access to distributed resources and data,
- Transparent access to Grids and reduces complexity of the Grid environment, applications and hardware,
- Security and trust mechanisms are based on X.509v3 certification.
- Ease of use
  - GUI job submission, job monitoring



Three layer architecture of the UNICORE middleware system.

Molecular docking

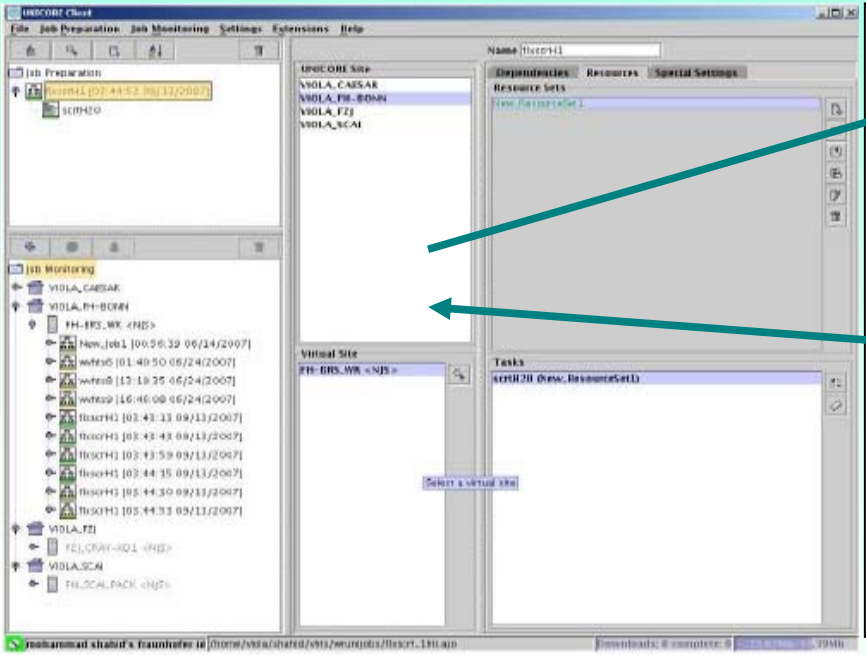
- Two molecular docking applications used in this vHTS
  - FlexX and AutoDock
- At large scale both are CPU and data intensive.
- **FlexX (BioSolveIT)**
  - FlexX takes ~ 1-3 minutes on a standard in[illegible]nine.
  - One executable: Can be parallelized
  - Embarrassingly parallel

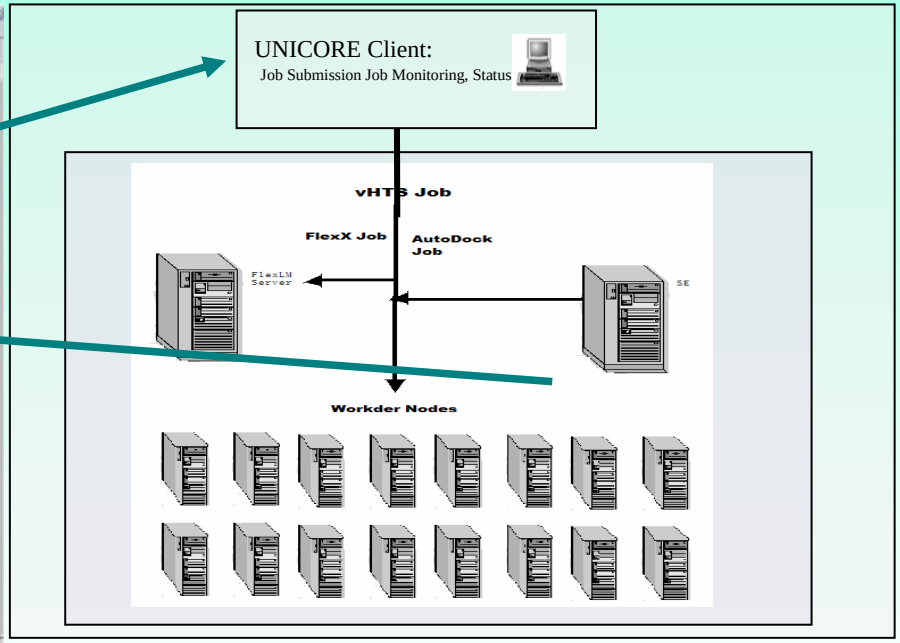


- **AutoDock (Scripps Research Institute)**
  - Comparatively time-consuming, a single docking job takes ~ 30-60 minutes on a standard CPU.
  - One executable: Can be parallelized
  - Embarrassingly parallel or massively parallel

Simplified vHTS Workflow on the Grid

UNICORE Client: prepare, submit jobs





Collect results...

Execute jobs

Fraunhofer Institute Algorithms and Scientific Computing

HEALTHGRID 2008 CONFERENCE CHICAGO JUNE 2-4, 2008

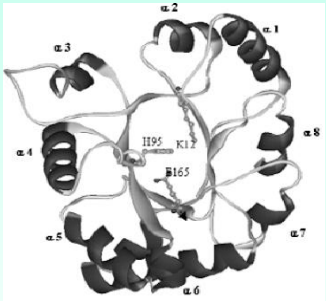
Drug Targets Selection

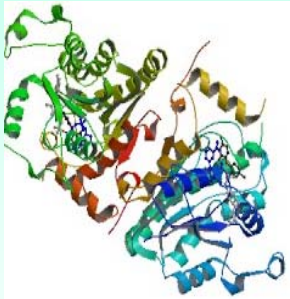
Drug Targets:

- Two classes of malarial targets
  - PfTIM: significant in the energy metabolism pathway
  - PfENR: significant in the lipid biosynthesis pathway

Target Validation:

- Selection of valid drug targets
- PDB IDs:1LYX, 1O5X, 1NHG,1NNU
- Crystal structures available …..





Yeh I, Altman R.B. (2006). Mini reviews in medicinal chemistry. 6:177-202.

Compound Library Design-I

Rationale:

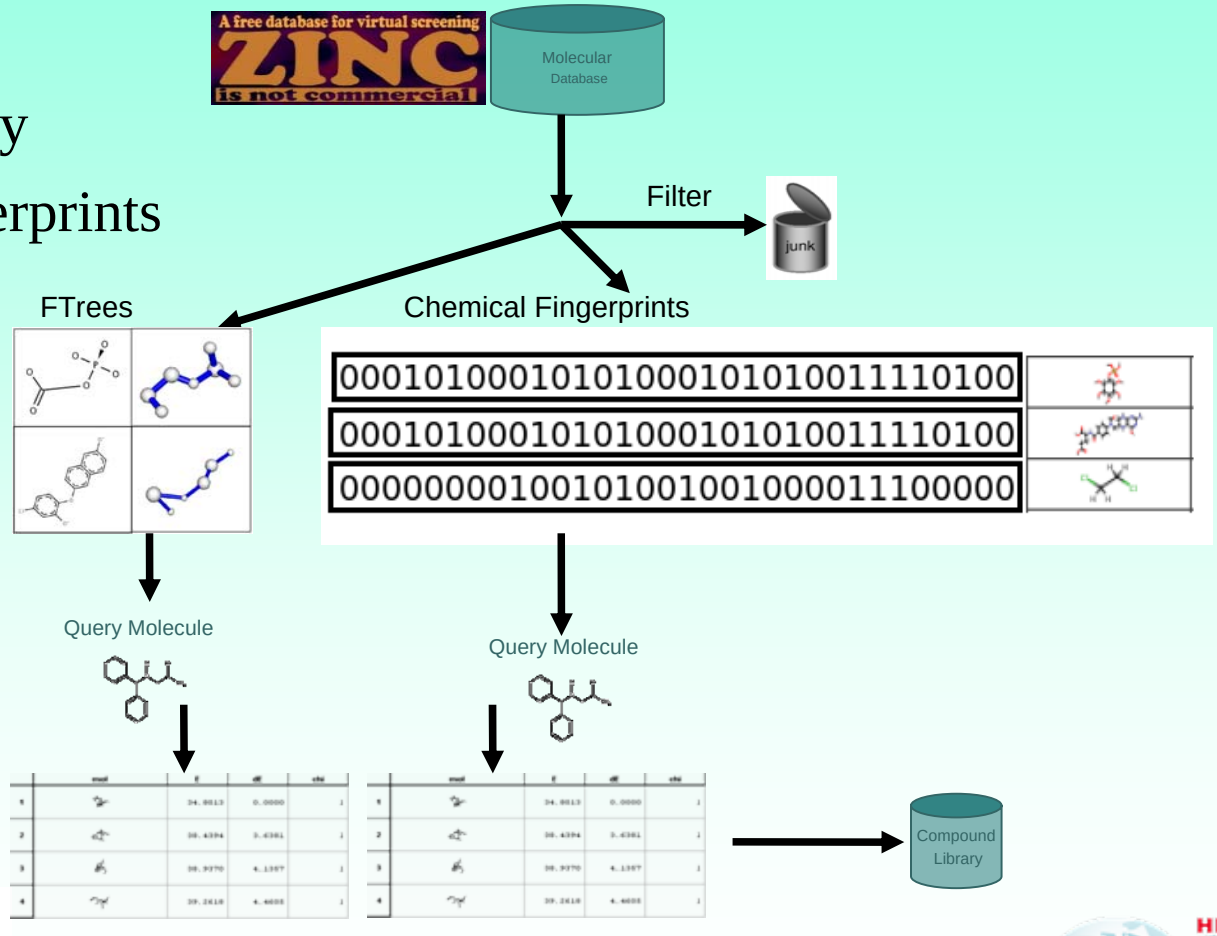
- To screen the whole compound databases is not relevant
- Reduce the “haystack”; increase the “hit” rate...
- Design a focused library to improve efficiency of screening.
- Two compound libraries for each class of target designed by using three different similarity methods
  - Feature Trees
  - MOE Similarity
  - Chemical fingerprints

Compound Library Design-II

Feature Trees

MOE Similarity

Chemical fingerprints



Fraunhofer Institute Algorithms and Scientific Computing

SCAI

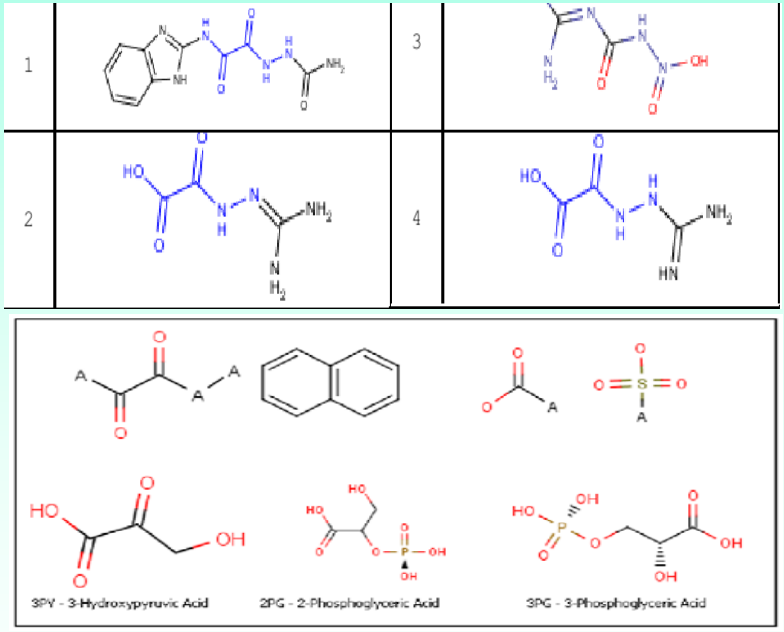
HEALTHGRID 2008 CONFERENCE CHICAGO JUNE 2-4, 2008

Results –I

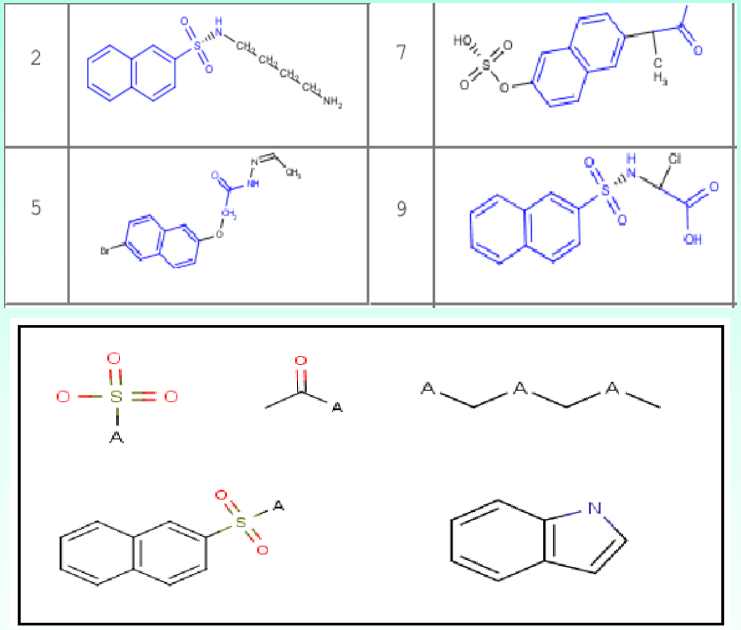
- Post processing and analysis:
  - Extracting the required information from the huge amount of output files
  - Designing a database to filter the results
  - Statistical analyses
  - Visual analysis of the ranking hits
- Several novel candidate compounds were identified
  - Potential competitive inhibitors for the drug targets

Results -III

FlexX ranking compounds for targets PfTIM.



AutoDock ranking compounds for targets PfTIM



Fraunhofer Institute Algorithms and Scientific Computing

HEALTHGRID 2008 CONFERENCE CHICAGO JUNE 2-4, 2008

Conclusions and future work…

Demonstrated once again the significance and relevance of the Grid infrastructure in the DD program

- The VIOLA Optical Testbed
  - production ready infrastructure for high throughput biological experiments
  - high speed connectivity
  - allow efficient data distribution and output collection
  - fast data transfer between the computing sites: improves overall performance

Conclusions and future work…

- The PHOSPHORUS Project
  - Within PHOSPHORUS project, the VIOLA Grid is connected to several other European Grids through the same high speed optical links.
  - Increase the computational base for doing further target-oriented docking simulations.
  - WSRF-compliant services for docking tools

Individualized Risk Assessment on Grids: ANEURIST (@neurIST)

Fraunhofer Institute Algorithms and Scientific Computing

Overview on

aneurIST

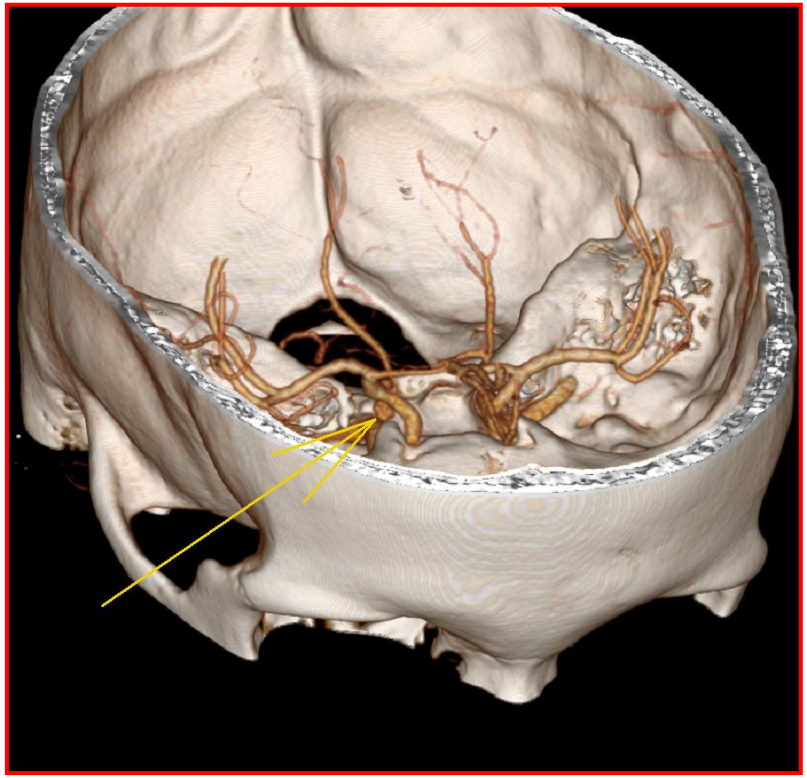
Integrated biomedical informatics for the management of cerebral aneurysms

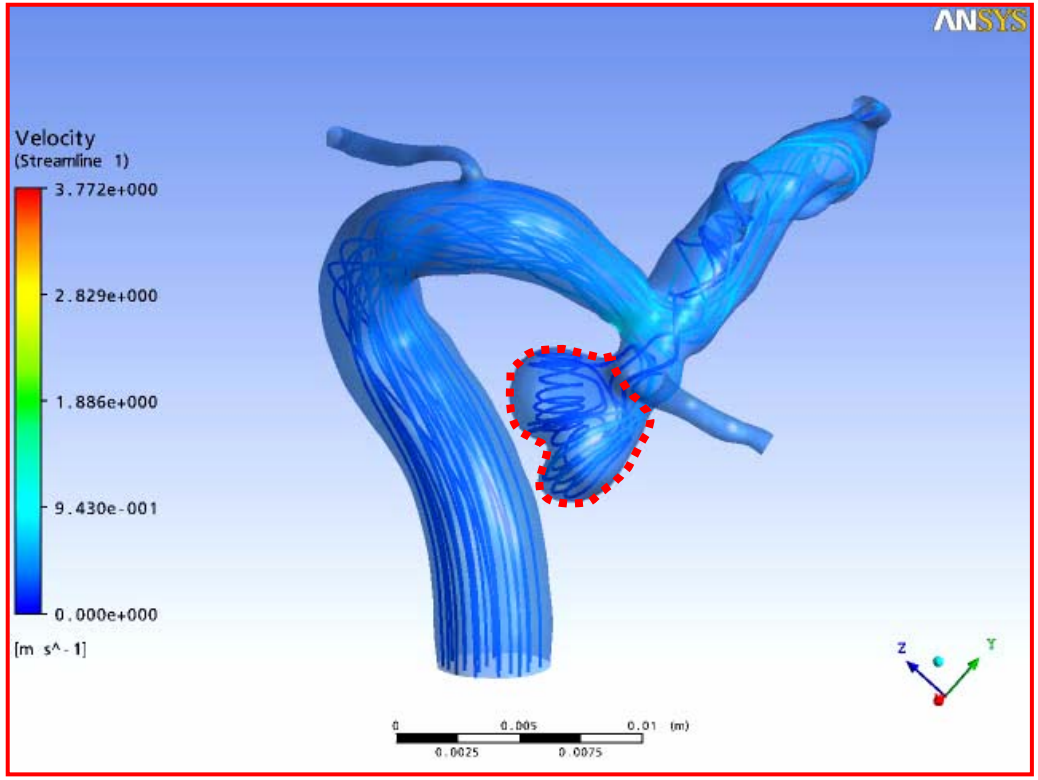
Funded under the Framework 6 programme of the European Commission
Contract No: IST-2004-027703
www.aneurist.org

Martin Hofmann-Apitius

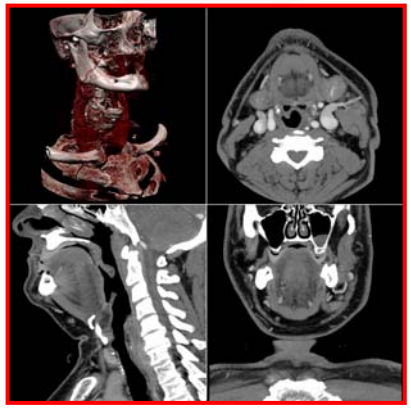
For the entire @neurIST - team

The Integrated Management of Cerebral Aneurysm





Three Common Aneurysm Scenarios…

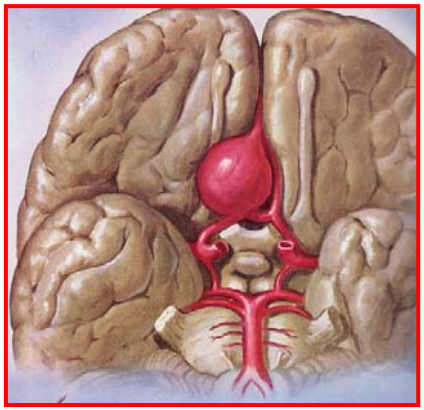


1. Aneurysm suspected on scan, contrast CT requested, aneurysm verified
2. Neck vessel dissection suspected, contrast CT requested, aneurysm verified
3. Initial non-contrast scan negative, patient deteriorates, contrast-CT to exclude vessel dissection or meningitis reveals aneurysm

1

2

3

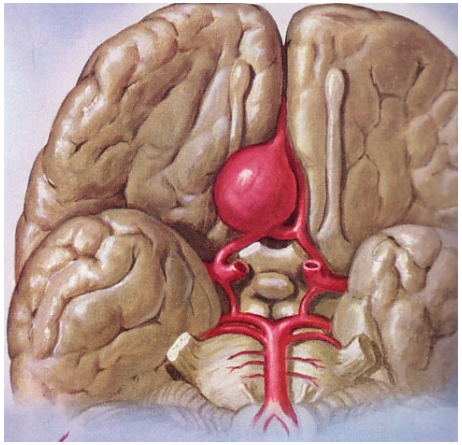


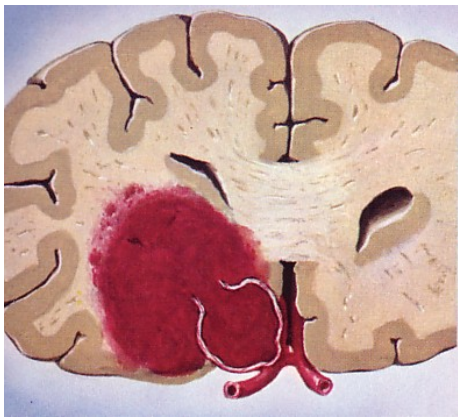
Other possibilities…

- *If the initial scan is MR, the aneurysm may show immediately*
- *Providential use of contrast agent may also reveal the problem*

Clinical background on Cerebral Aneurysms (CAs)

- **Management of unruptured aneurysms is controversial** as the *disease is very complex* and current decision making is solely based on consideration of very limited information like the aneurysms' size and location (ISUIA trial).
- There is evidence that genetic predisposition may be involved in early phases of the natural history of aneurysms (initiation)
- Current minimally invasive (endovascular) treatment is favored over surgical treatment for many aneurysms (ISAT trial), but both are risky, costly and do not prevent recurrence in all cases.
- There is a need for a new generation of therapies treating the cause rather than symptoms of the disease
- There is increasing incidental diagnosis of unruptured aneurysms, due to wider availability of modern imaging
- Better educated population will actually demand availability and discussion of the therapeutic options offered to them





Natural history of CA

@neurIST

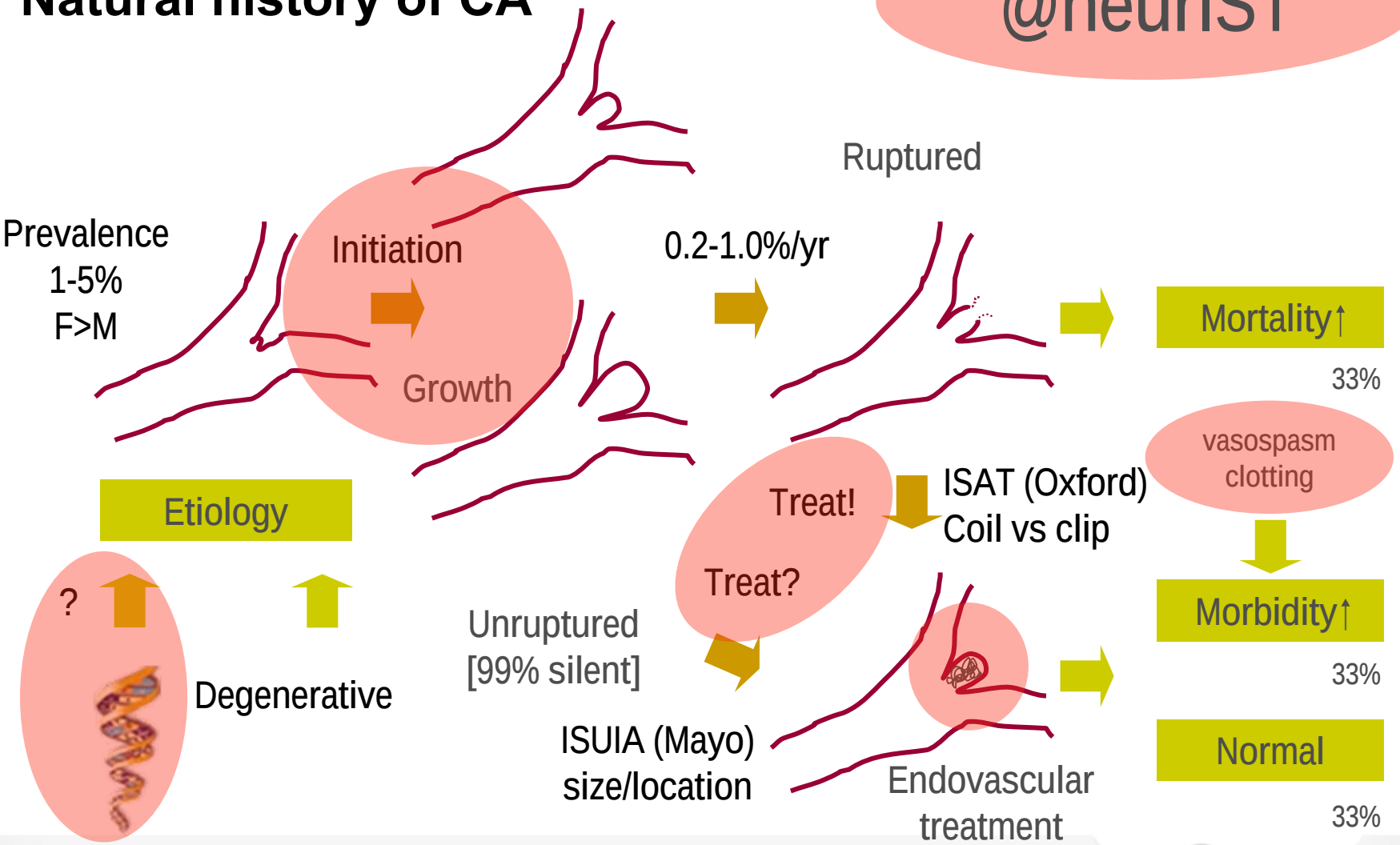
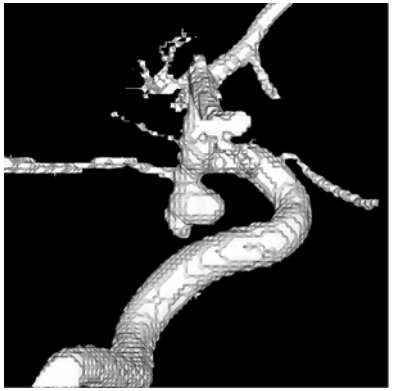
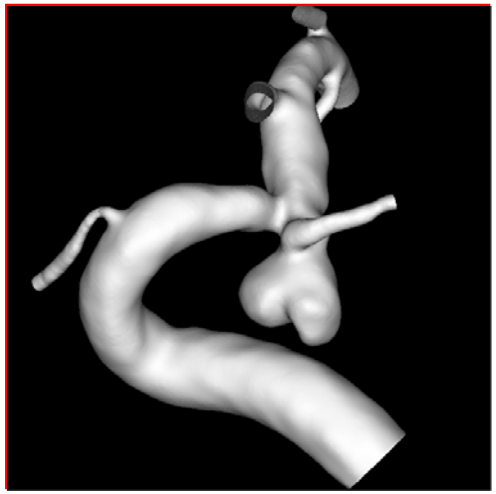
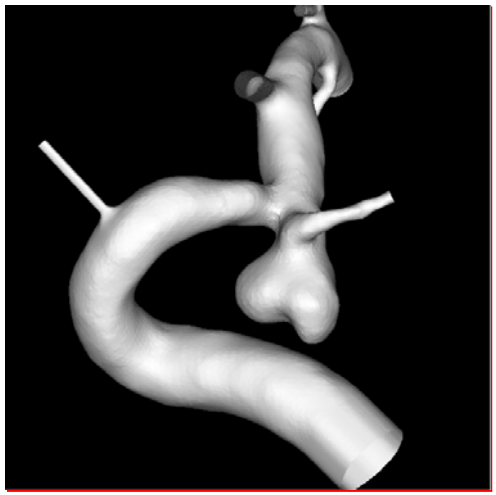


Image data for vascular modelling : The Case Study

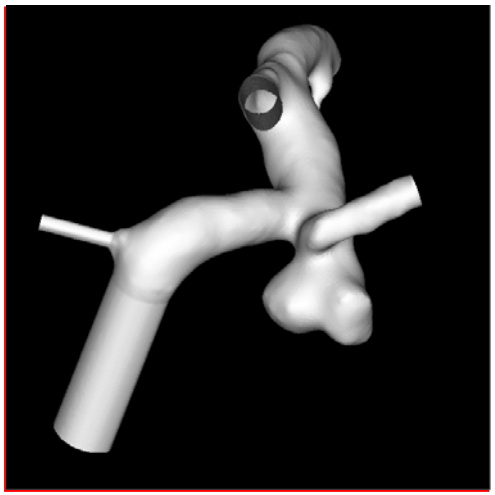




3DRA (512^3)



3DRA (128^3)

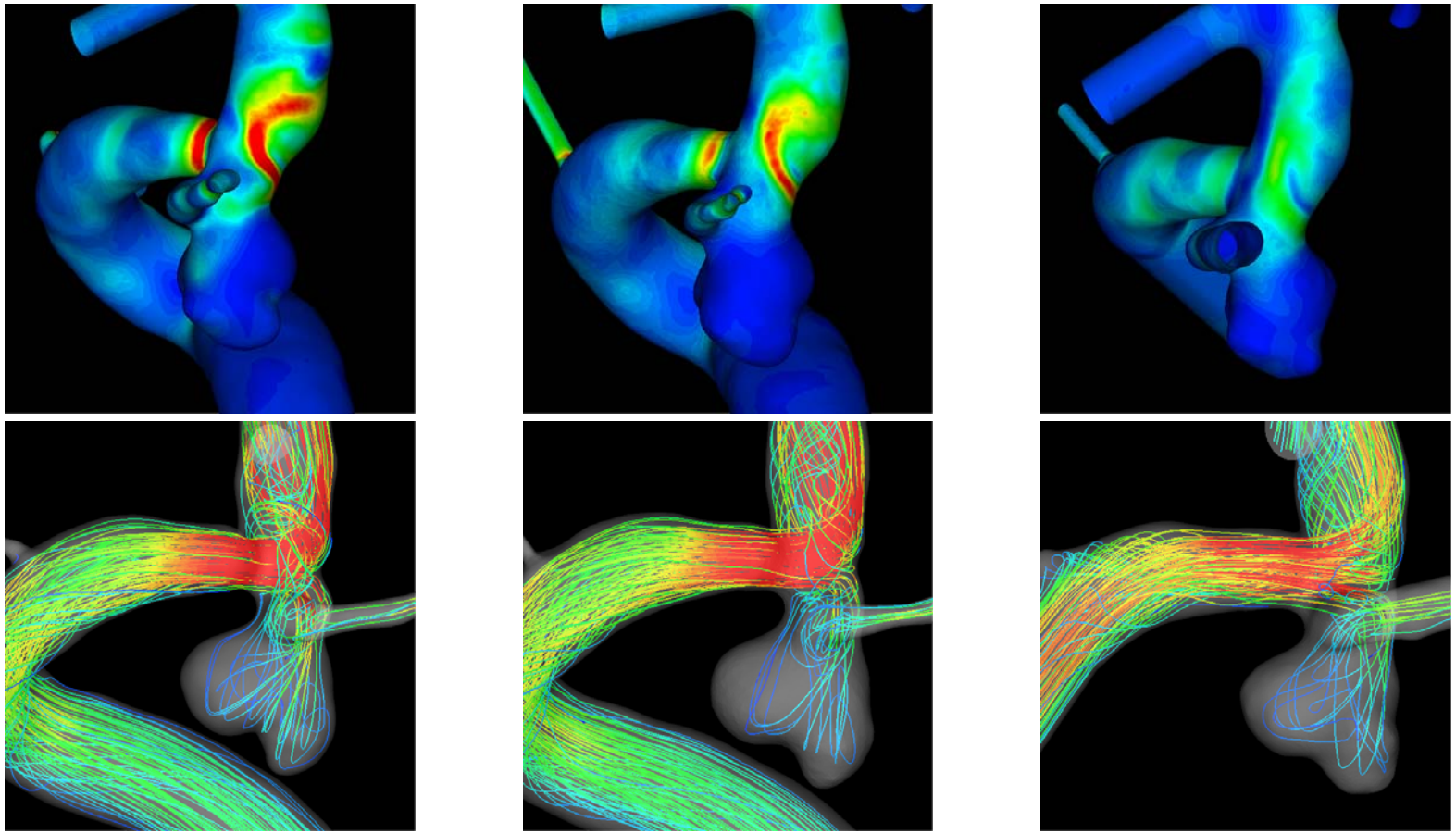


CTA

Image data for vascular modelling : The Case Study

WSS

Flow patterns

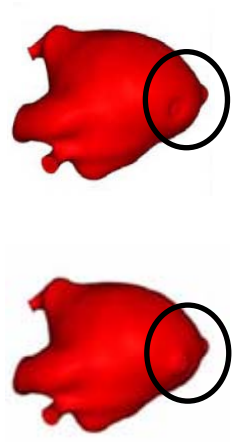


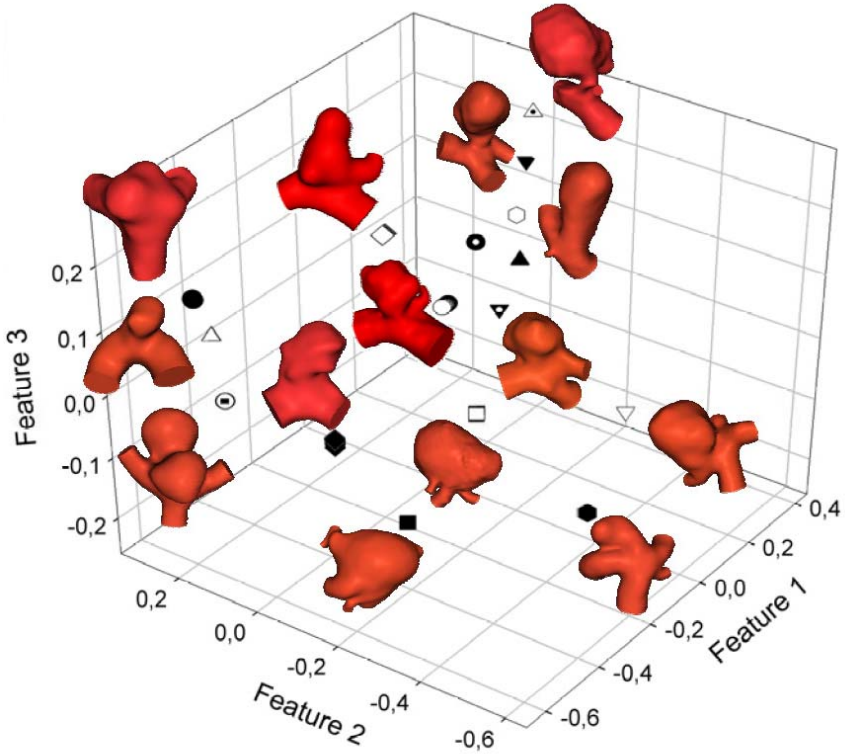
3DRA (512^3) 3DRA (128^3) CTA

Moments as Feature Descriptors

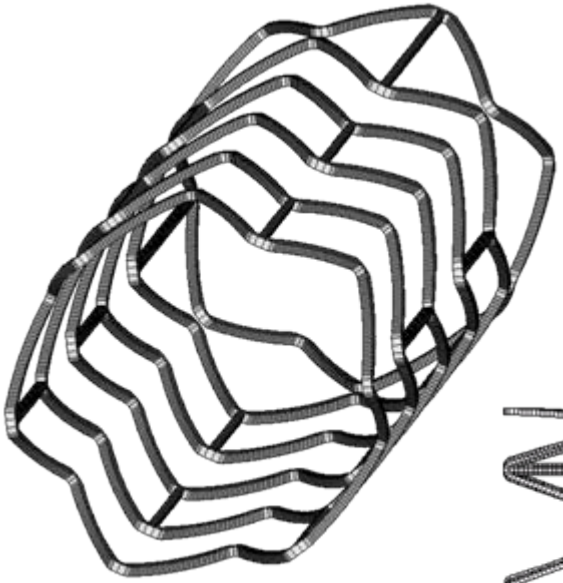
- Complete bases to describe the shapes, from which invariant descriptors to similarity transformations can be extracted

Small variations to increase statistics and account for segmentation variability

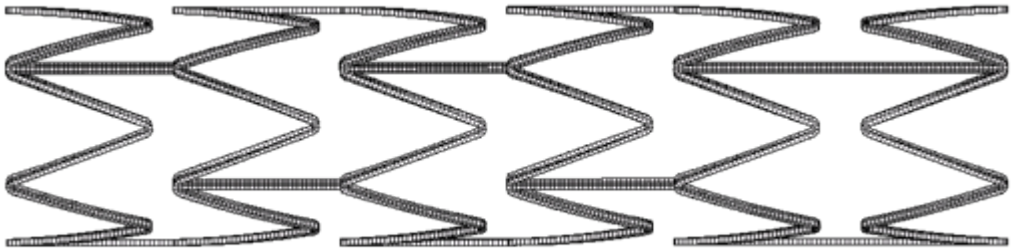




1. Generation of numerical stent model

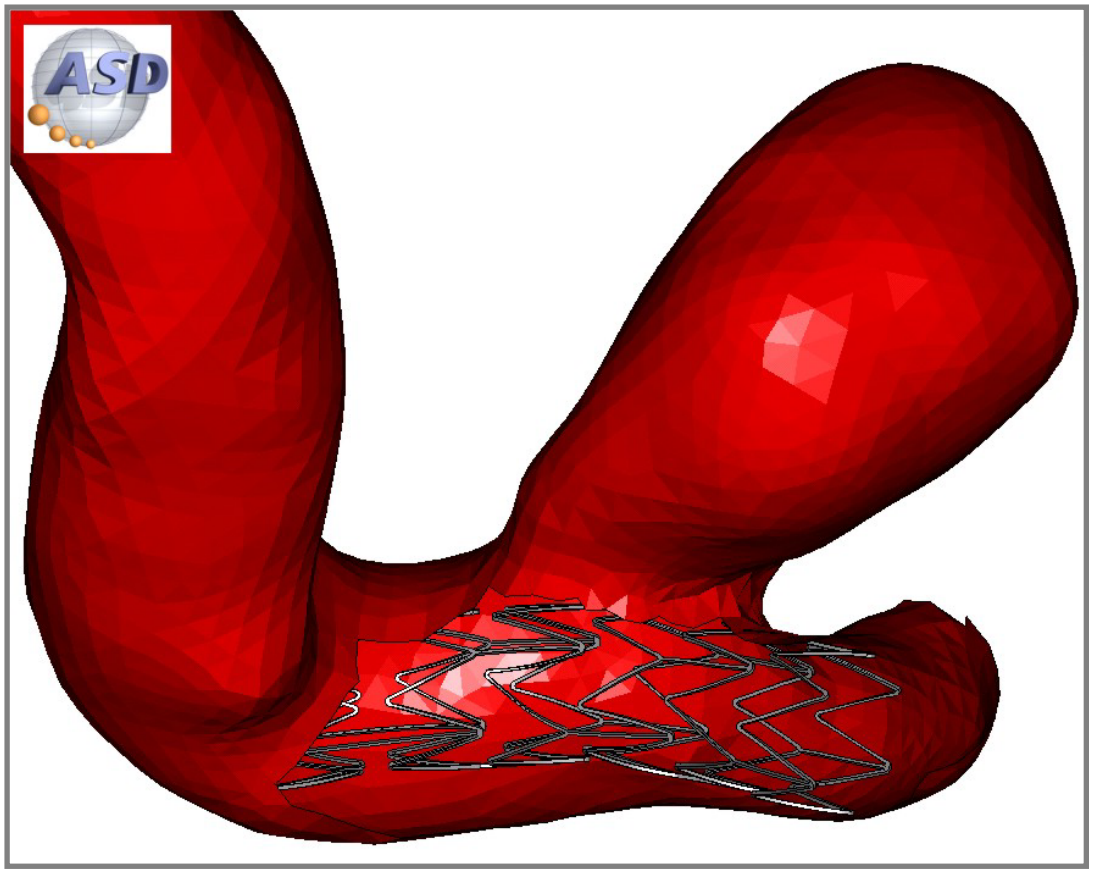


Diameter: 3.0 mm
Length: 12.6 mm
Thickness: 0.08 mm
Element type: solid
Material: titanium alloy

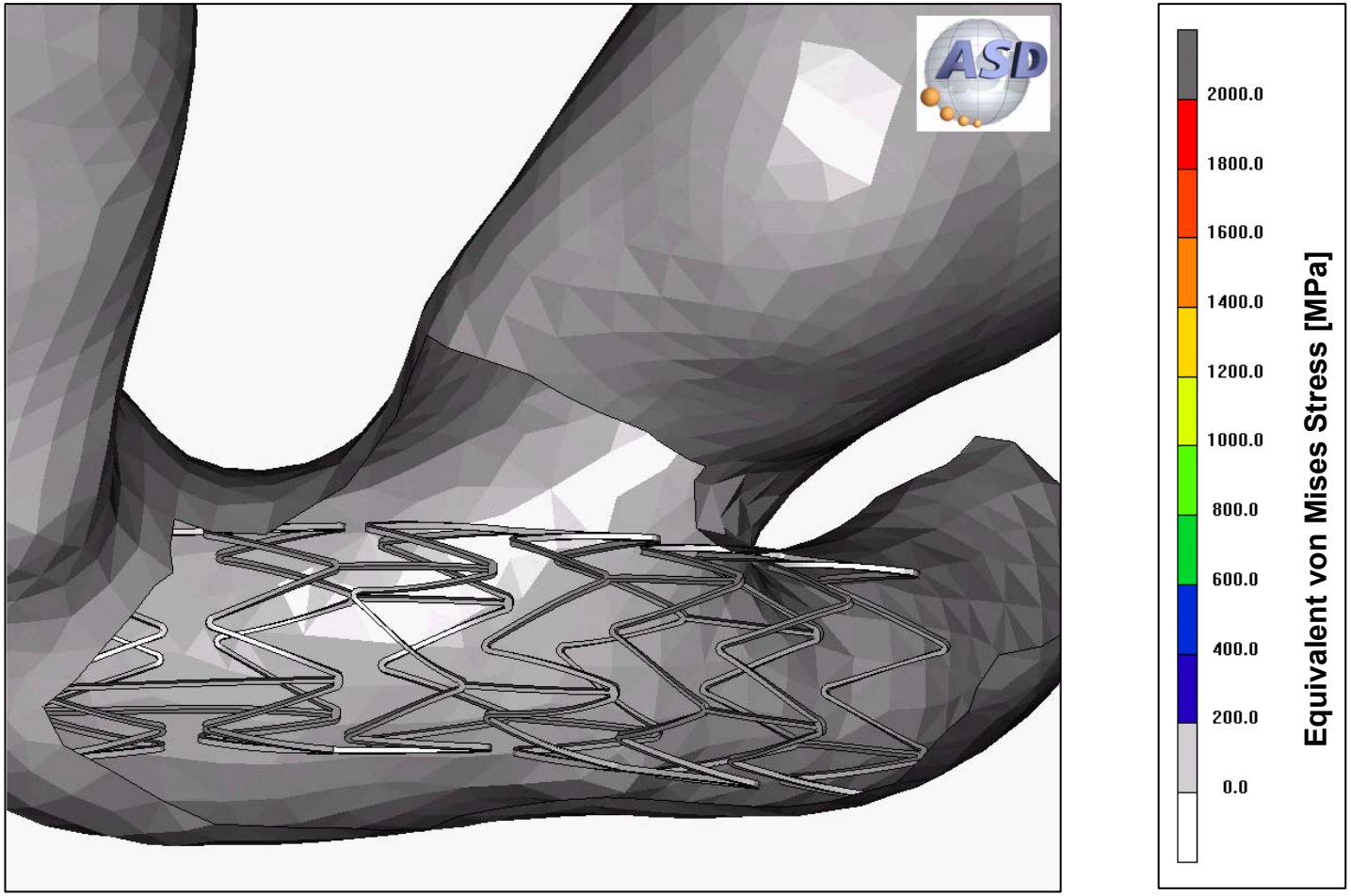


3. Expansion of stent near aneurysm (FEM) 1 of 3

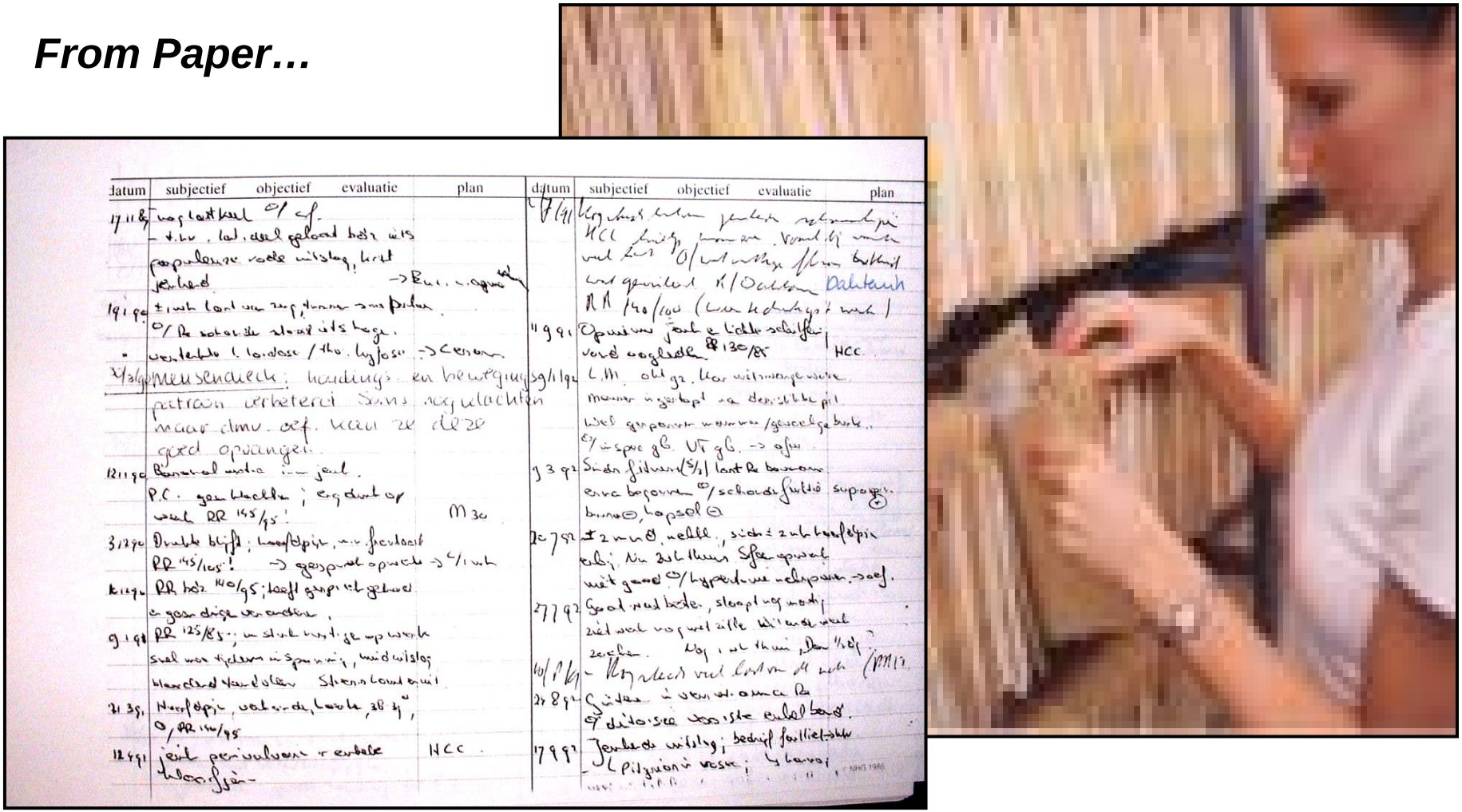
<u>Contact analysis:</u> stent / arterial wall (rigid surface)



3. Expansion of stent near aneurysm (FEM) 3 of 3



From Paper…



datum	subjectief	objectief	evaluatie	plan
17 11 8[illegible]	[illegible]	[illegible]		
19 1 90	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]		
12 1 90	[illegible]	[illegible]		
3 12 90	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]		
9 1 90	[illegible]	[illegible]		
21 3 91	[illegible]	[illegible]		
12 4 91	[illegible]	[illegible]	HCC	

datum	subjectief	objectief	evaluatie	plan
[illegible]	[illegible]	[illegible]		
11 9 91	[illegible]	[illegible]	HCC	
[illegible]	[illegible]	[illegible]		
3 3 92	[illegible]	[illegible]		
20 7 92	[illegible]	[illegible]		
27 7 92	[illegible]	[illegible]		
[illegible]	[illegible]	[illegible]		
21 8 92	[illegible]	[illegible]		
17 9 92	[illegible]	[illegible]		

Distributed data collection

Descriptive Data

@neuQuest

Non structured digital information

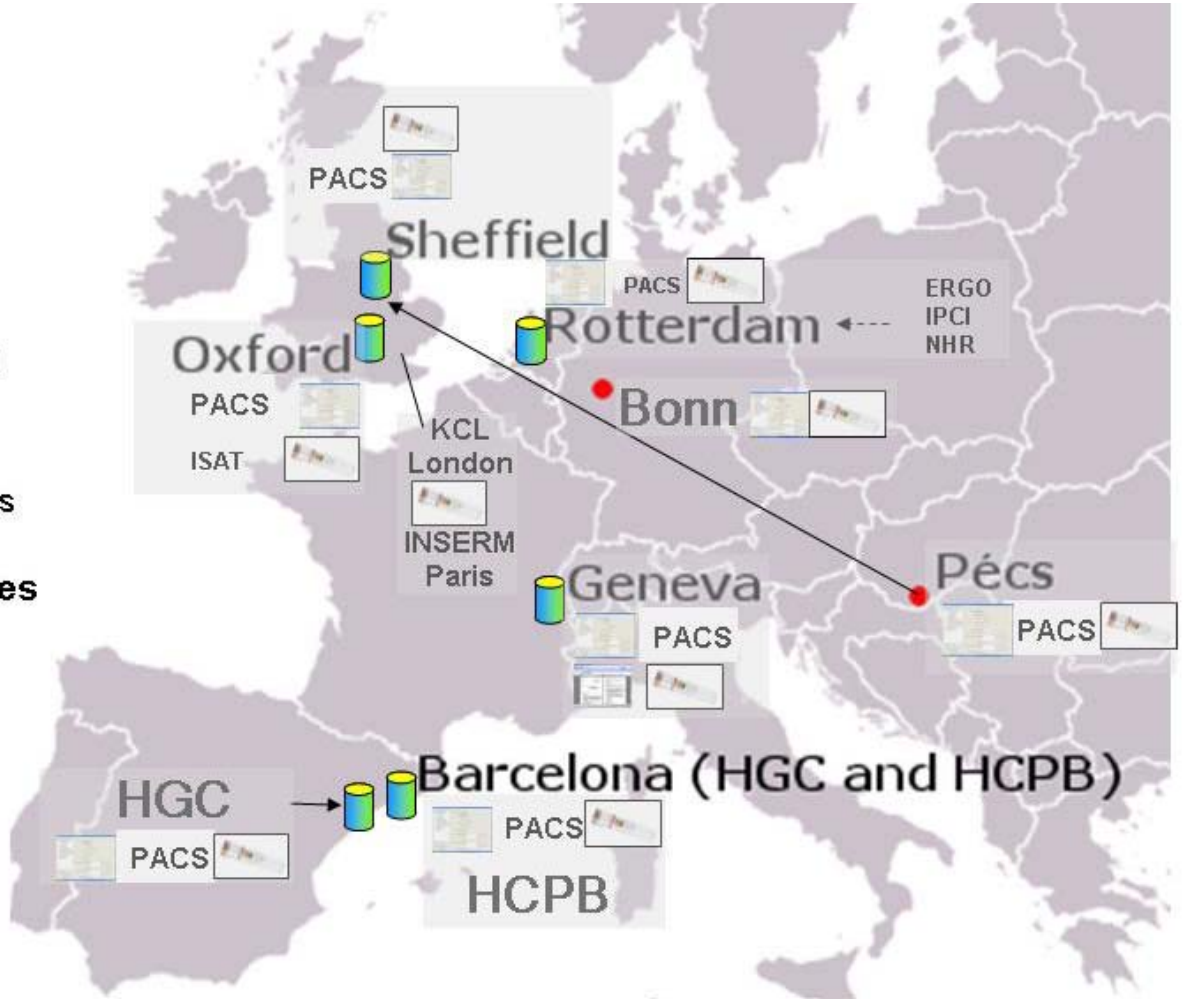
Representative Data

PACS eRadiology Archives

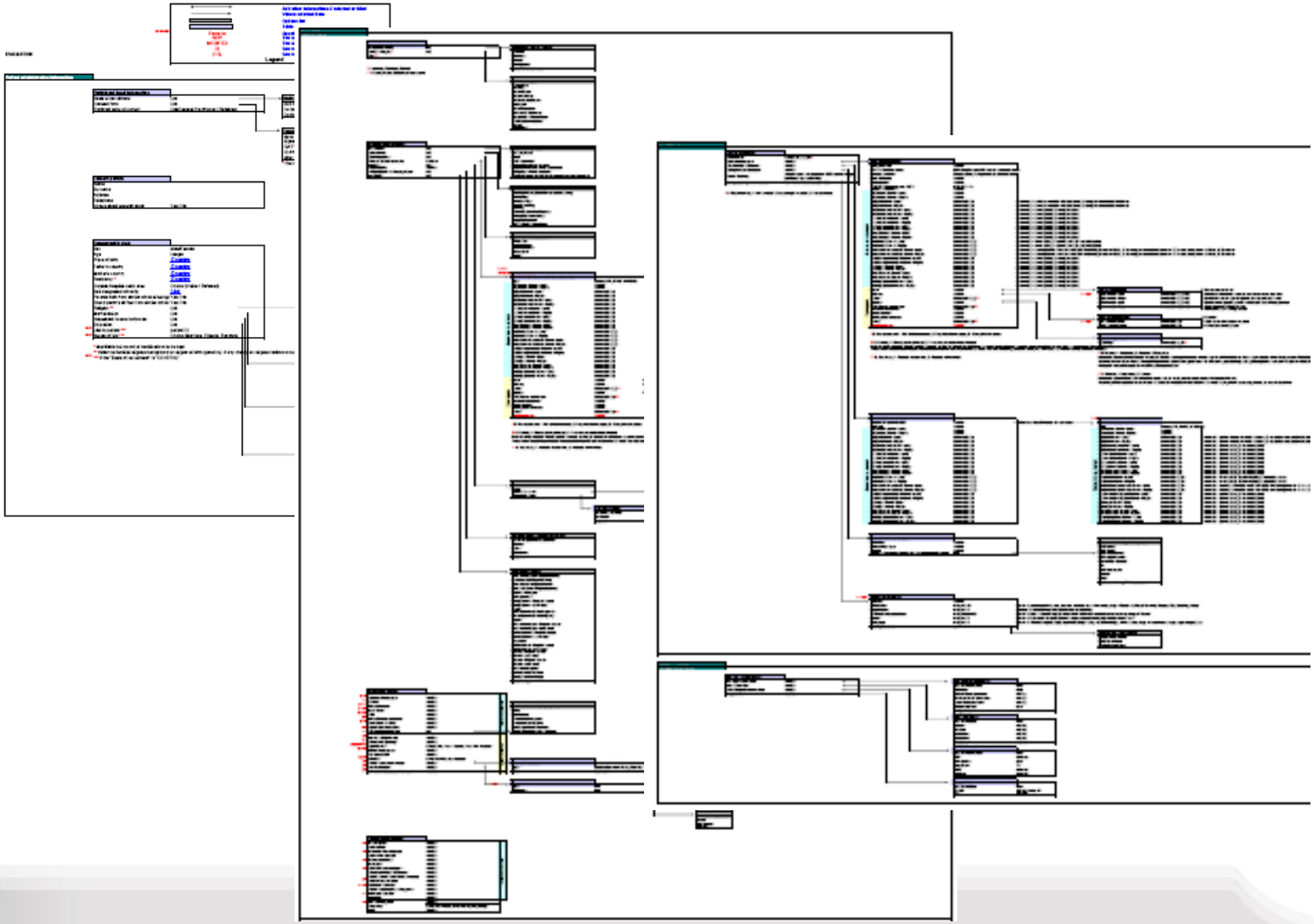
Conservation of samples

Bio Samples

@neurIST BioIS



First prototype of Clinical Reference Information Model (CRIM)



Integrative application suites that are being developed

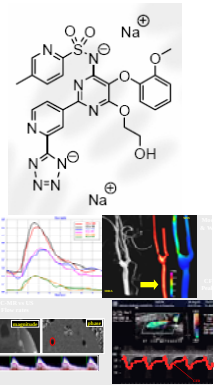
@neurIST Systems

- **@neuRisk**
- **@neuEndo**
- **@neuLink**
- **@neuFuse**
- **@neuCompute/Info**

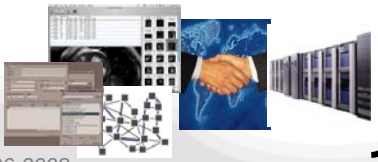
Improve **decision making processes** in the management of **unruptured aneurysms** by providing a score that integrates all the available information for **identifying at-risk patients** and reducing current over treatment

Support **computational design processes** towards a **next generation of smart flow-correcting implants** to treat **ruptured aneurysms** and reduce current treatment costs, side effects and recurrence.

Support the **knowledge discovery for linking genetics to disease**, vasospasm and blood clotting after cerebral hemorrhage



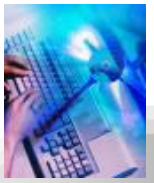
Support the **integration of modeling, simulation and visualization of multimodal data**



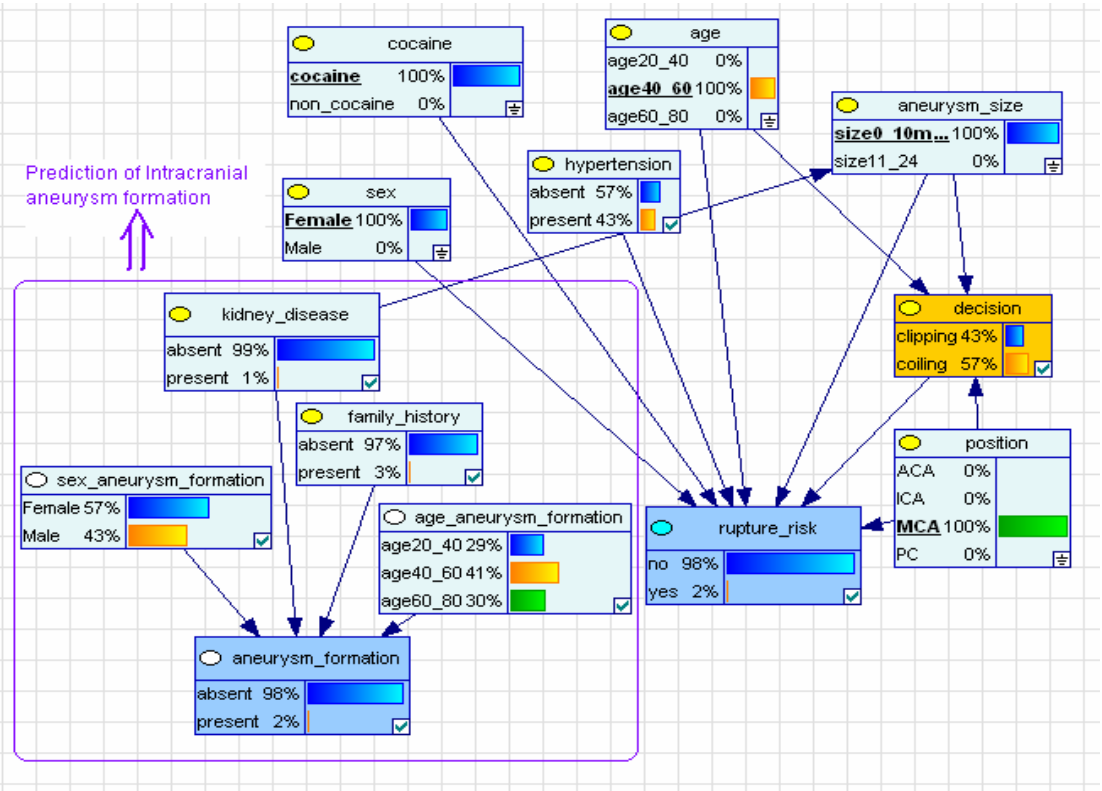
Support **integration of data and computing resources.**

IT Support Suites

Enabling IT

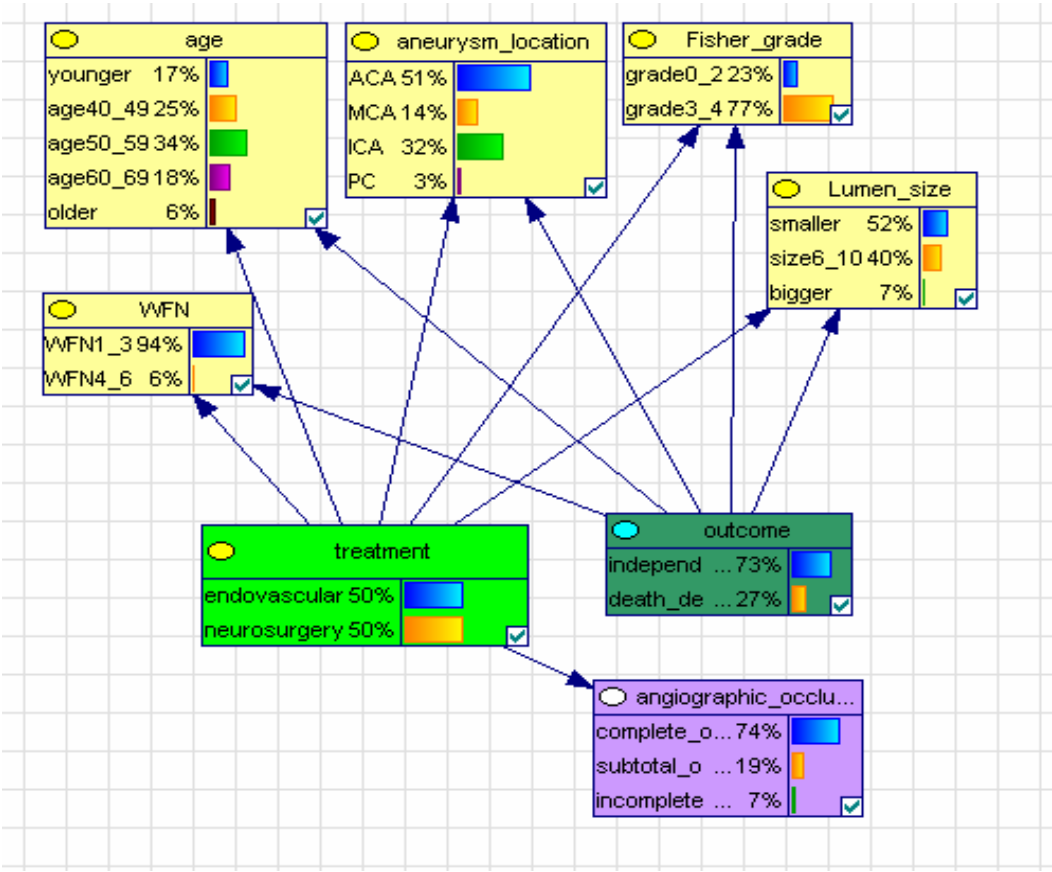


Bayesian Model of Aneurysm Formation and Rupture Risk



Contribution by Dr. Christoph Friedrich, and BeiBei Han, SCAI, to WP4.1 & WP5.1

Bayesian Model of Clinical Treatment Outcome (ISAT Data)



Contribution by Dr. Christoph Friedrich, and BeiBei Han, SCAI, to WP4.1 & WP5.1

Interactive Risk Assessment

- Personalized risk assessment
- Interactive treatment planning
- Risk-Benefit visualisation