Big data in health care: the distributed learning solution

Prof. Philippe Lambin

U.H. Maastricht

Disclosures

- Research collaborations incl. funding
  - Varian (VATE, chinaCAT, euroCAT), Siemens (euroCAT), SohPhilips (EURECA, TraIT, BIONIC), Xerox (EURECA), ptTheragnostic, OncoRadiomics
- Public research funding
  - Radiomics (USA-NIH/U01CA143062), euroCAT(EU-Interreg), duCAT & StraTegY (NL-STW), EURECA (EU-FP7), BD2decide (Horizon2020), Bionic (NWO)

Why did we start the a Big Data project (CAT*) project?

*Computer Assisted Theragnostic: CAT=euroCAT, duCAT, VATE, chinaCAT etc.

Evidence based medicine

Conventional Clinical Research

High data quality

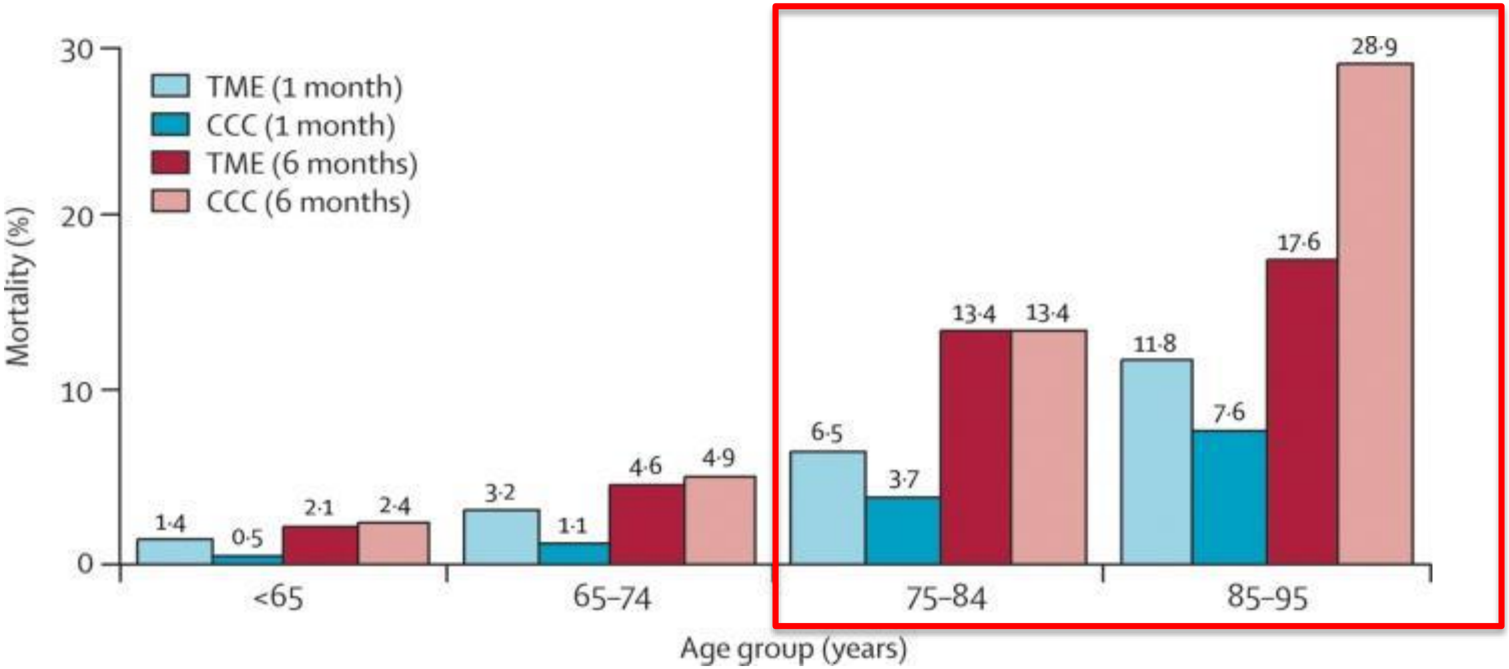
Low data quantity

Controlled

- Assigned patients
- “EORTC-RTOG grade” QA/Protocol
- Biobanking, translational research

- Less then 3% of the patients
- Highly biased population
- Randomized trials rarely done for new technologies

Example: having *no evidence* can have dramatic consequences

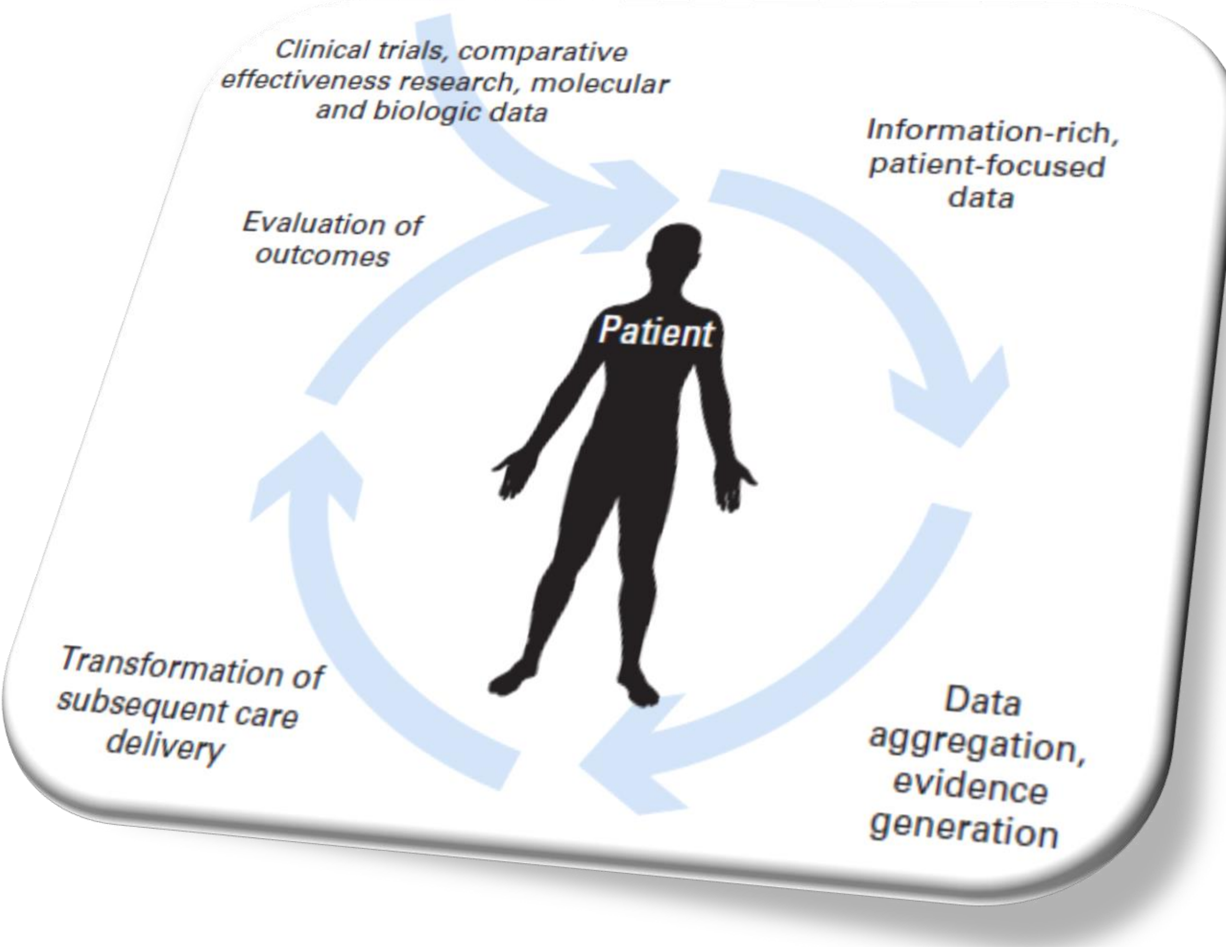


Rutten et al. Lancet Oncology 2008; 9: 494

The solution? Use the 97%: Rapid Learning Health Care or “Big data in health care”

- In [..] rapid-learning [..] data routinely generated through **patient care and clinical research** feed into an ever-growing [..] set of **coordinated databases**.
- *J Clin Oncol 2010;28:4268*

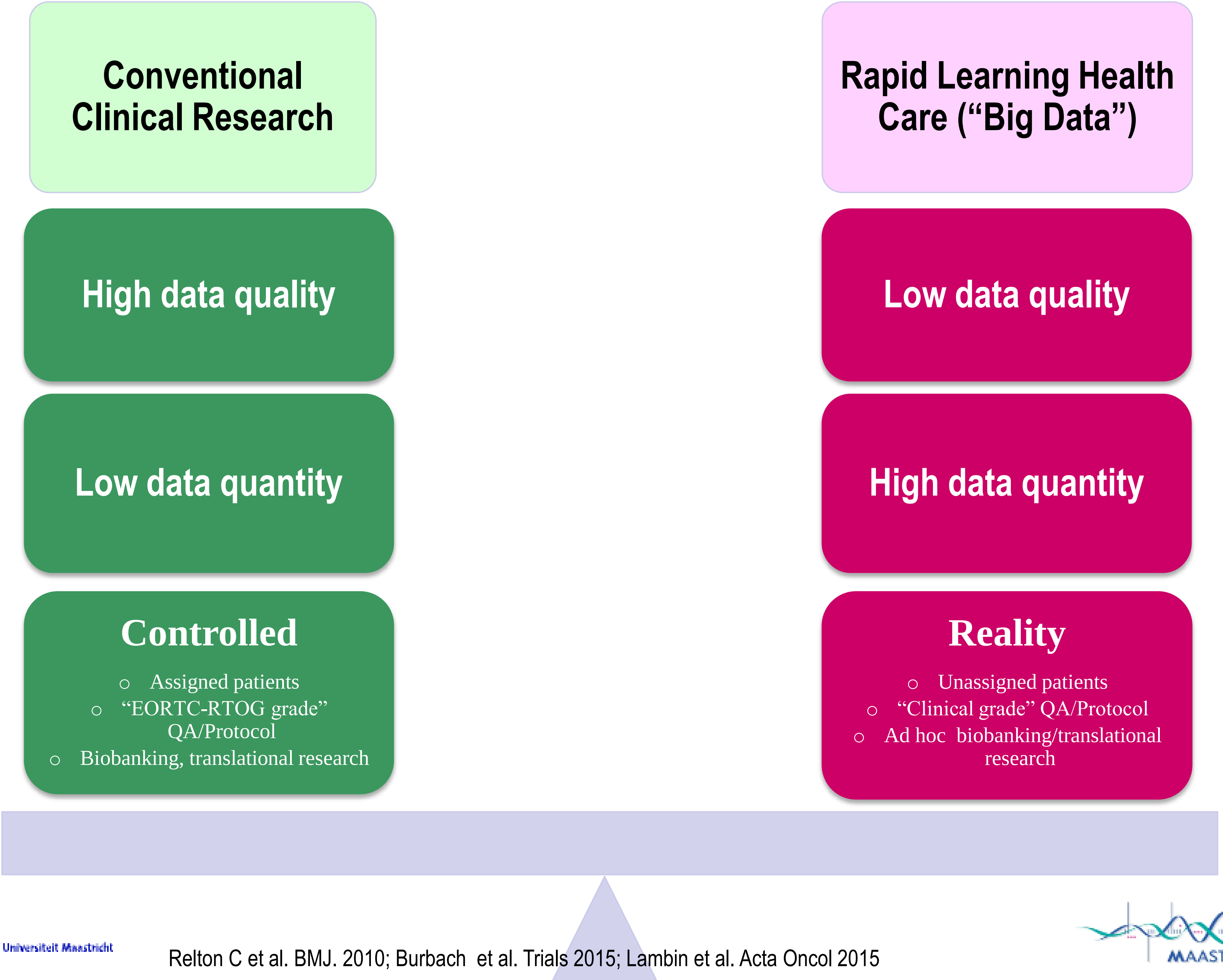
- [..] rapid learning [..] where we can **learn from each patient** to guide practice, is [..] crucial to guide rational health policy and to contain costs [..].
- *Lancancet Oncol 2011;12:933*


Clinical trials, comparative effectiveness research, molecular and biologic data
Information-rich, patient-focused data
Evaluation of outcomes
Patient
Transformation of subsequent care delivery
Data aggregation, evidence generation

Examples:

1. *Radiotherapy CAT (www.eurocat.info)*
2. *ASCO’s CancerLinQ*

Lambin *et al.* Aug 28, Radiother Oncol 2013, Adv. Drug Dev 2016



Conventional Clinical Research

High data quality

Low data quantity

Controlled

- ○ Assigned patients
- ○ "EORTC-RTOG grade" QA/Protocol
- ○ Biobanking, translational research

Rapid Learning Health Care ("Big Data")

Low data quality

High data quantity

Reality

- ○ Unassigned patients
- ○ "Clinical grade" QA/Protocol
- ○ Ad hoc biobanking/translational research

Relton C et al. BMJ. 2010; Burbach et al. Trials 2015; Lambin et al. Acta Oncol 2015

Example of clinically relevant questions

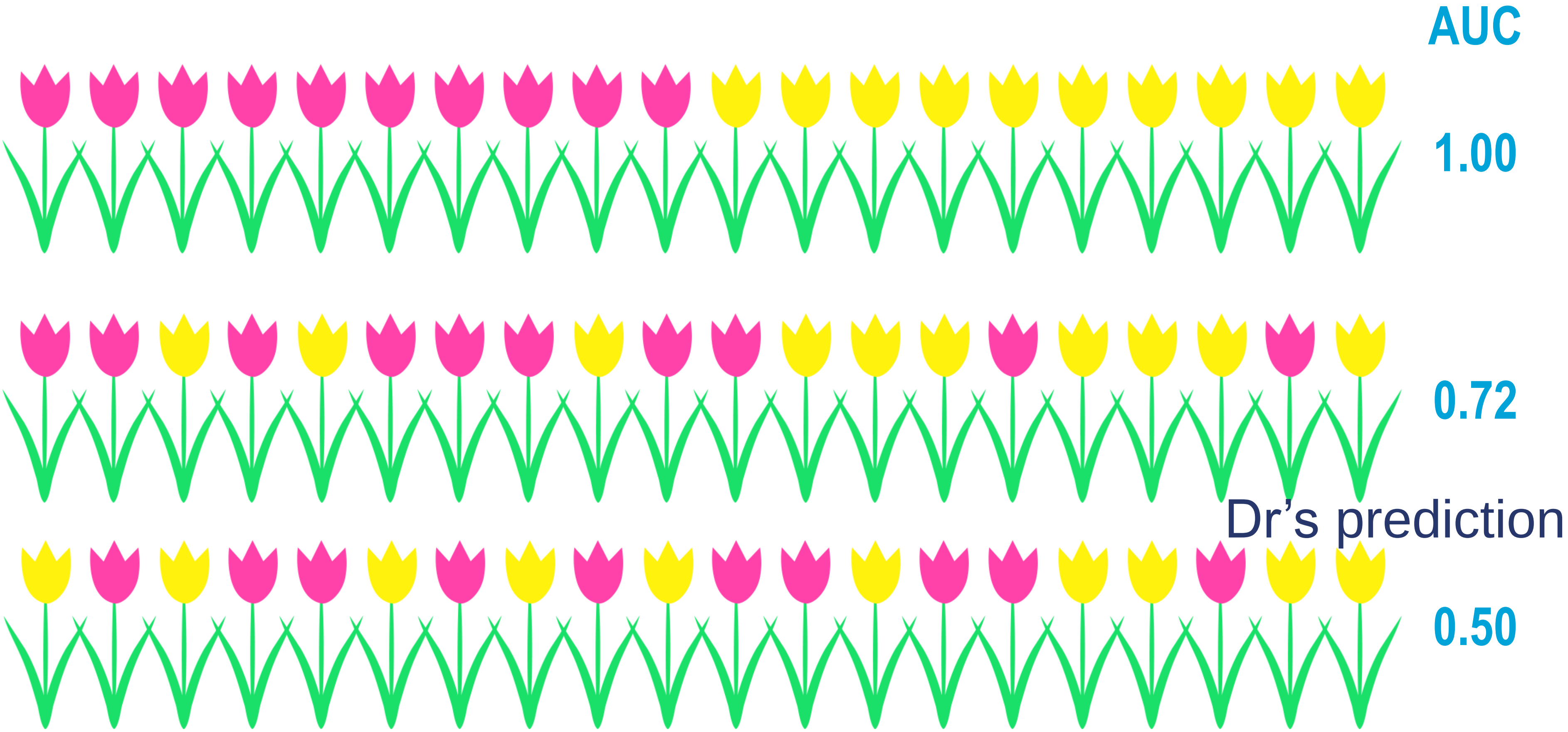
Treatment of

- 80 years old rectal cancer?
- 70 years old Stage IIIB NSCLC?
- 60 years old prostate cancer with oligometastasis?
- Local relapse of a stage 3 oropharynx?
- Cervix cancer stage 3, HIV+
- …

Can we predict a tulip’s color by looking at the bulb?

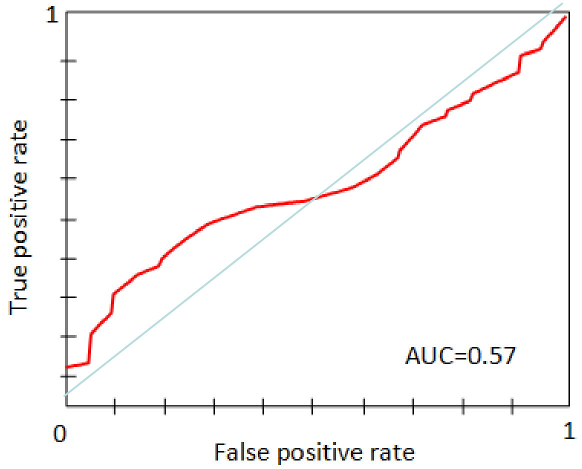
http://www.amystewart.com

Predicting the tulip color

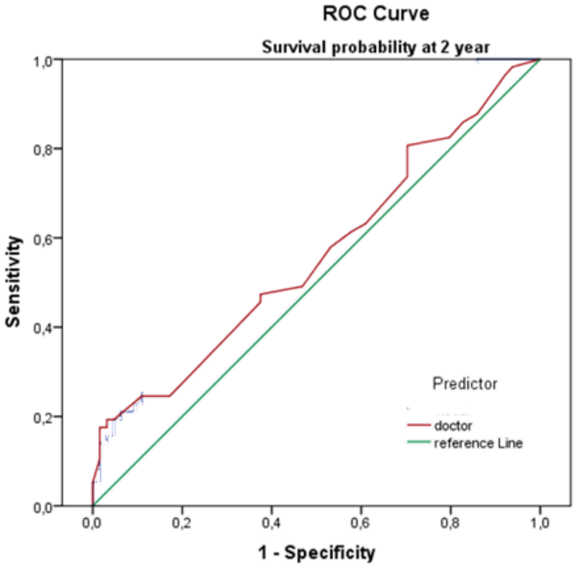


Oberije *et al.* Radiother Oncol. 2014

Prediction by MDs? *Unskilled* (Prospective trial)



- •NSCLC inoperable M0
- •2 year survival
- •30 patients
- •8 MDs
- •Retrospective
- •**AUC: 0.57**



NSCLC

2 year survival

158 patients

5 MDs

Prospective

AUC: 0.56

Oberije et al. Radiother Oncol. 2014

Prediction by MDs? Unskilled *and* *unaware* of it

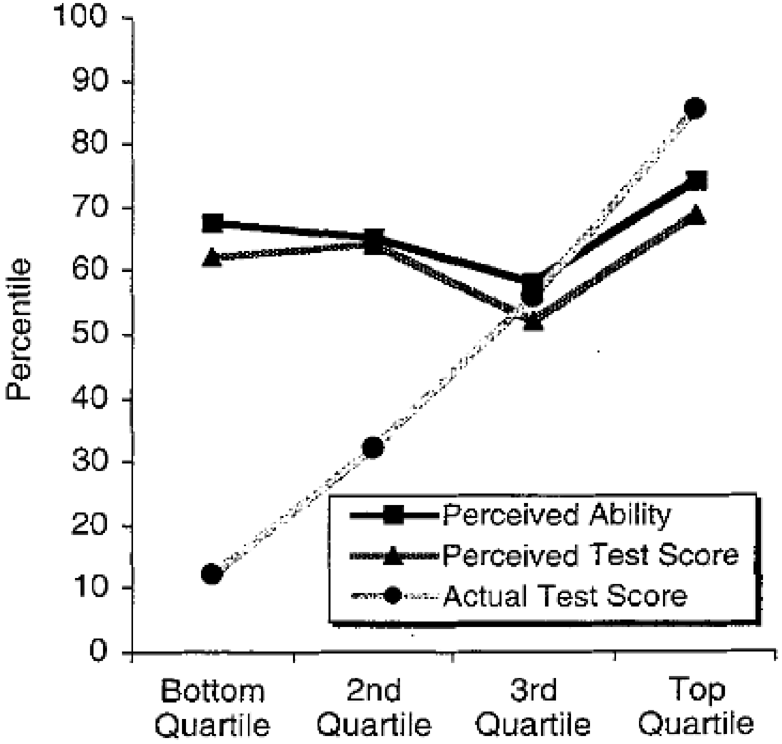
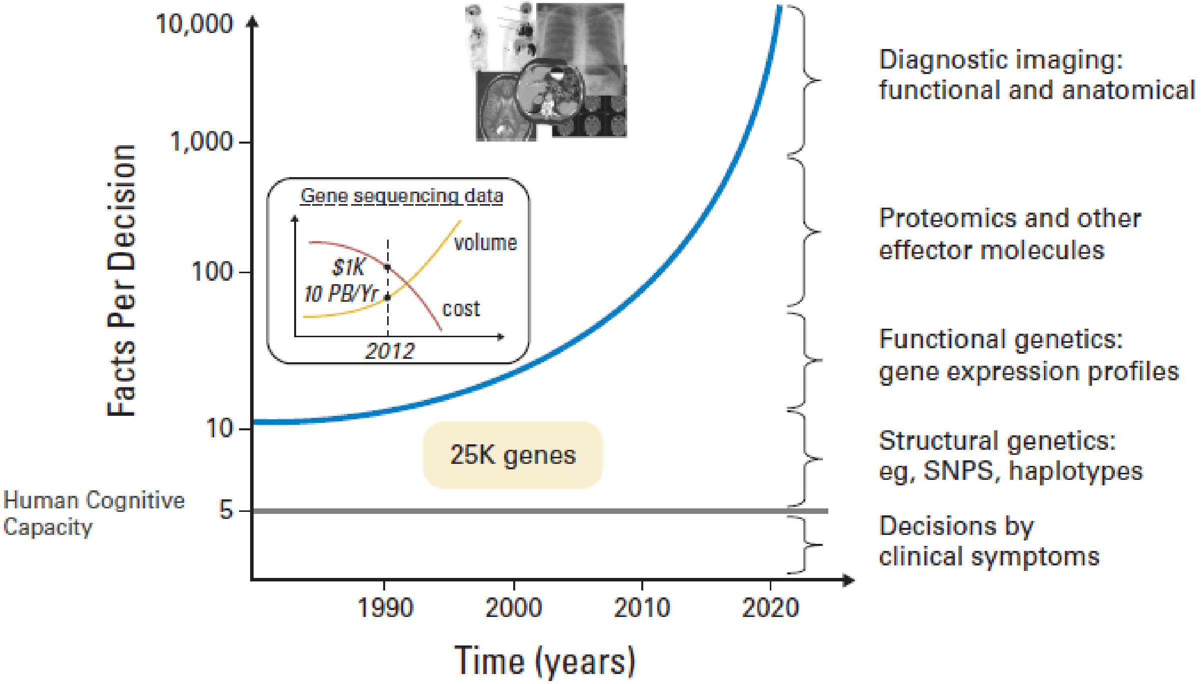


Figure 2. Perceived logical reasoning ability and test performance as a function of actual test performance (Study 2).

Unskilled and unaware of it: *How difficulties in recognizing one's own incompetence leads to inflated self-assessments.* J Pers Soc Psych

Kruger et al. 1999

No guiltiness! The doctor is drowning



- Explosion of data
- Explosion of decisions
- Explosion of 'evidence'*

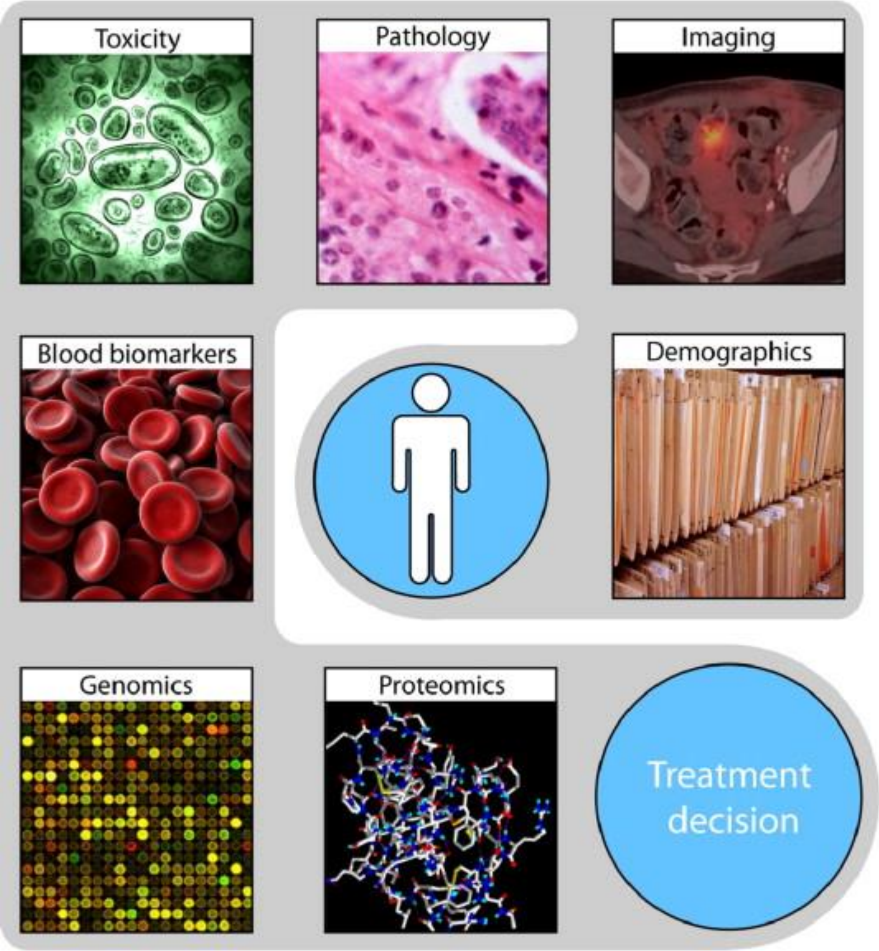
Our vision in 2 min

*2010: 1574 & 1354 articles on lung cancer & radiotherapy = 7.5 per day

Half-life of knowledge estimated at 7 years (in young students)

J Clin Oncol 2010;28:4268
JMI 2012 Friedman, Rigby

Multifactorial Decision Support System



But we need Data, preferably *most* of them

Lambin et al. Nature Rev. Clin. Oncol.

What are

the barriers

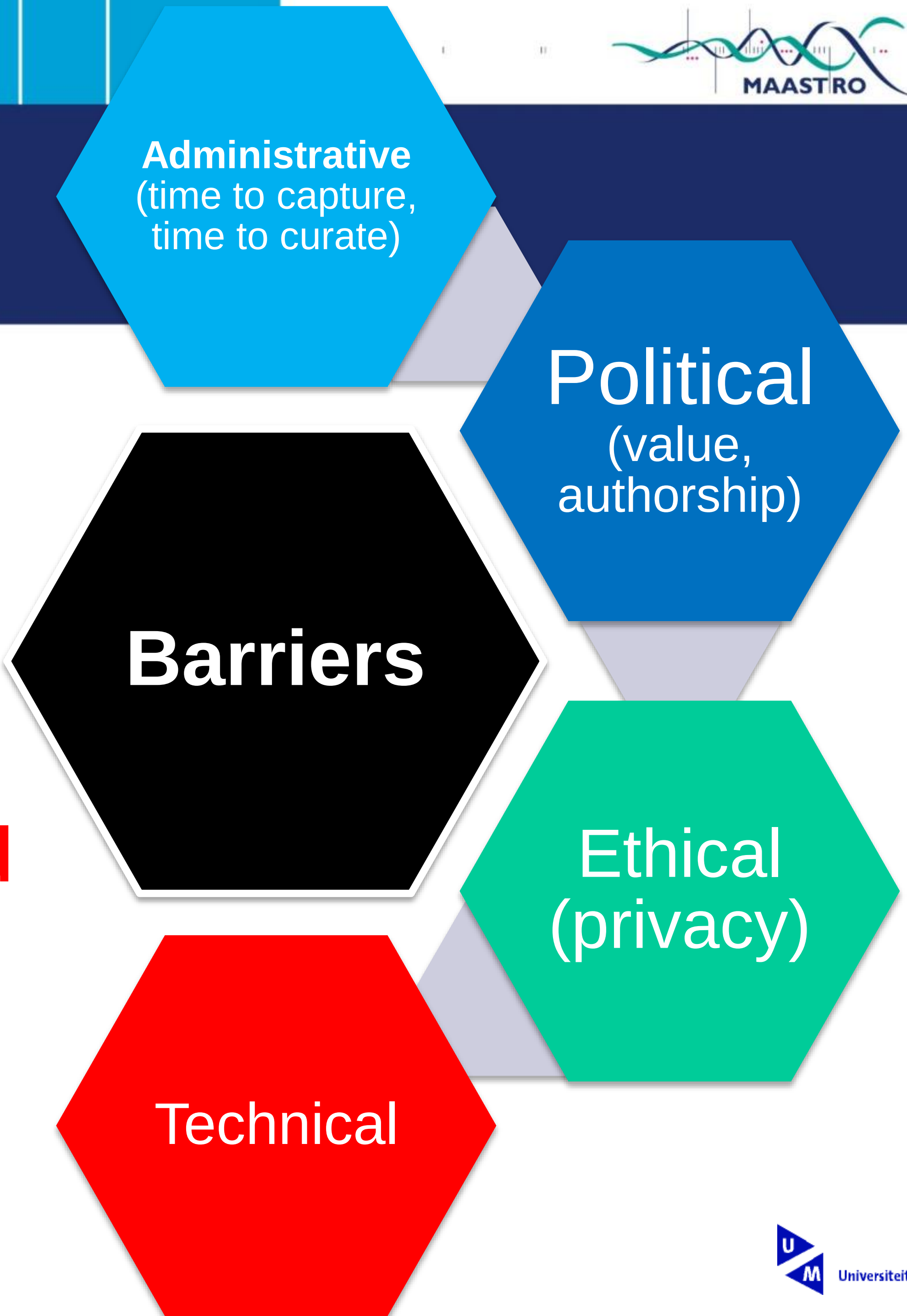
to share the data?

What is the solution?

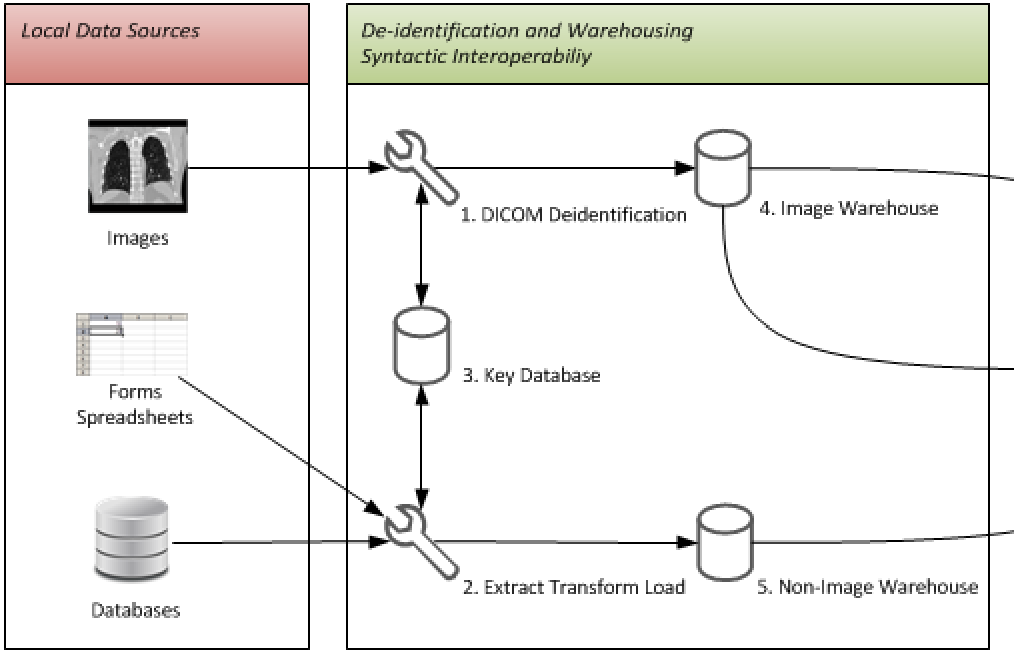
Sharing data

[..] the problem is not really technical […]. Rather, the problems are **ethical, political, and administrative**.
Lancet Oncol 2011;12:933

Solutions: Distributed learning from federated databases



In-hospital infra & de-identification


Local Data Sources
Images
Forms
Spreadsheets
Databases
De-identification and Warehousing
Syntactic Interoperabiliy
1. DICOM Deidentification
2. Extract Transform Load
3. Key Database
4. Image Warehouse
5. Non-Image Warehouse

SPARQL = “Simple Protocol And RDF Query Language”
a query language for databases

Deidentification:

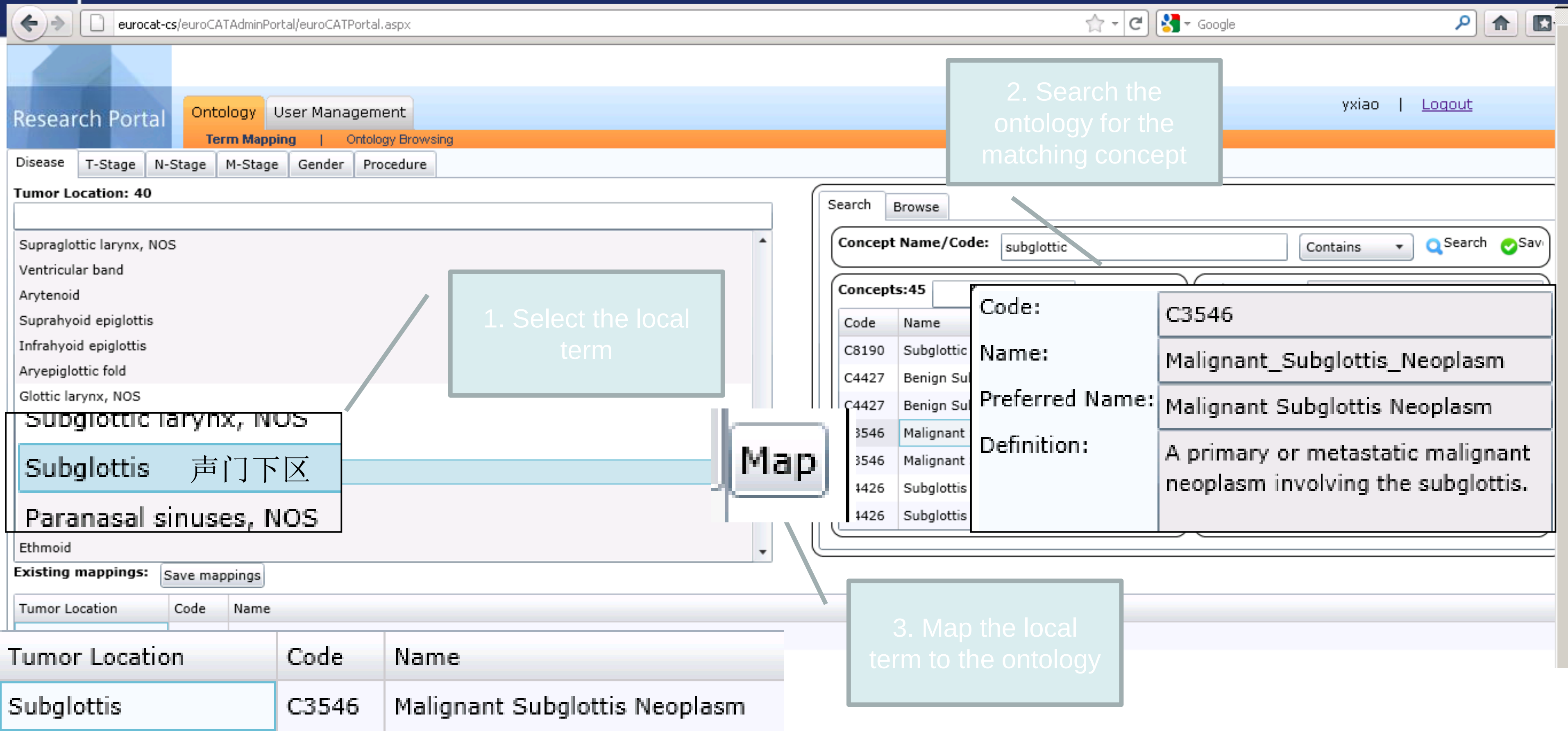
- Removal of obvious patient identifiers (name, MRN, social security number, email etc.)
- Assign a persistent token pseudonym
- Change (data banding) of obvious but required patient identifiers (everyone born and died on the 15th of the month, part of the postal code)
- No individual patient data leaves the hospital

The Semantic Web

- The **Semantic Web** is an extension of the Web through standards by the World Wide Web Consortium (W3C). The standards promote common data formats and exchange protocols on the Web.
- According to the W3C, "The Semantic Web provides a common framework* that allows data to be shared and reused across application, enterprise, and community boundaries". The term was coined by Tim Berners-Lee for a *web of data that can be processed by machines*.

*SPARQL is a semantic query language for databases

Ontology – International Coding System



Tumor Location: 40

- Supraglottic larynx, NOS
- Ventricular band
- Arytenoid
- Suprahyoid epiglottis
- Infrahyoid epiglottis
- Aryepiglottic fold
- Glottic larynx, NOS
- Subglottic larynx, NOS
- Subglottis 声门下区
- Paranasal sinuses, NOS
- Ethmoid

Concept Name/Code: subglottic

Concepts:45

Code	Name
C8190	Subglottic
C4427	Benign Sul
C4427	Benign Sul
3546	Malignant
3546	Malignant
4426	Subglottis
4426	Subglottis

Code: C3546

Name: Malignant_Subglottis_Neoplasm

Preferred Name: Malignant Subglottis Neoplasm

Definition: A primary or metastatic malignant neoplasm involving the subglottis.

Existing mappings:

Tumor Location	Code	Name
Subglottis	C3546	Malignant Subglottis Neoplasm

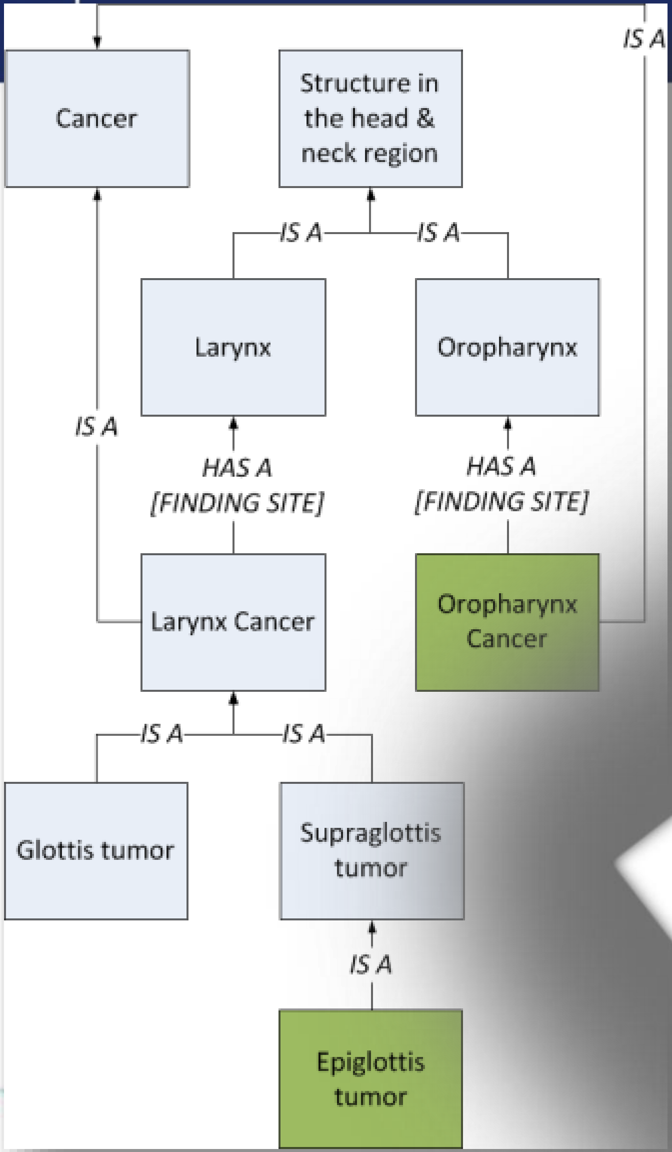
1. Select the local term

2. Search the ontology for the matching concept

3. Map the local term to the ontology

4. See the result of your mapping

An ontology is more than a dictionary

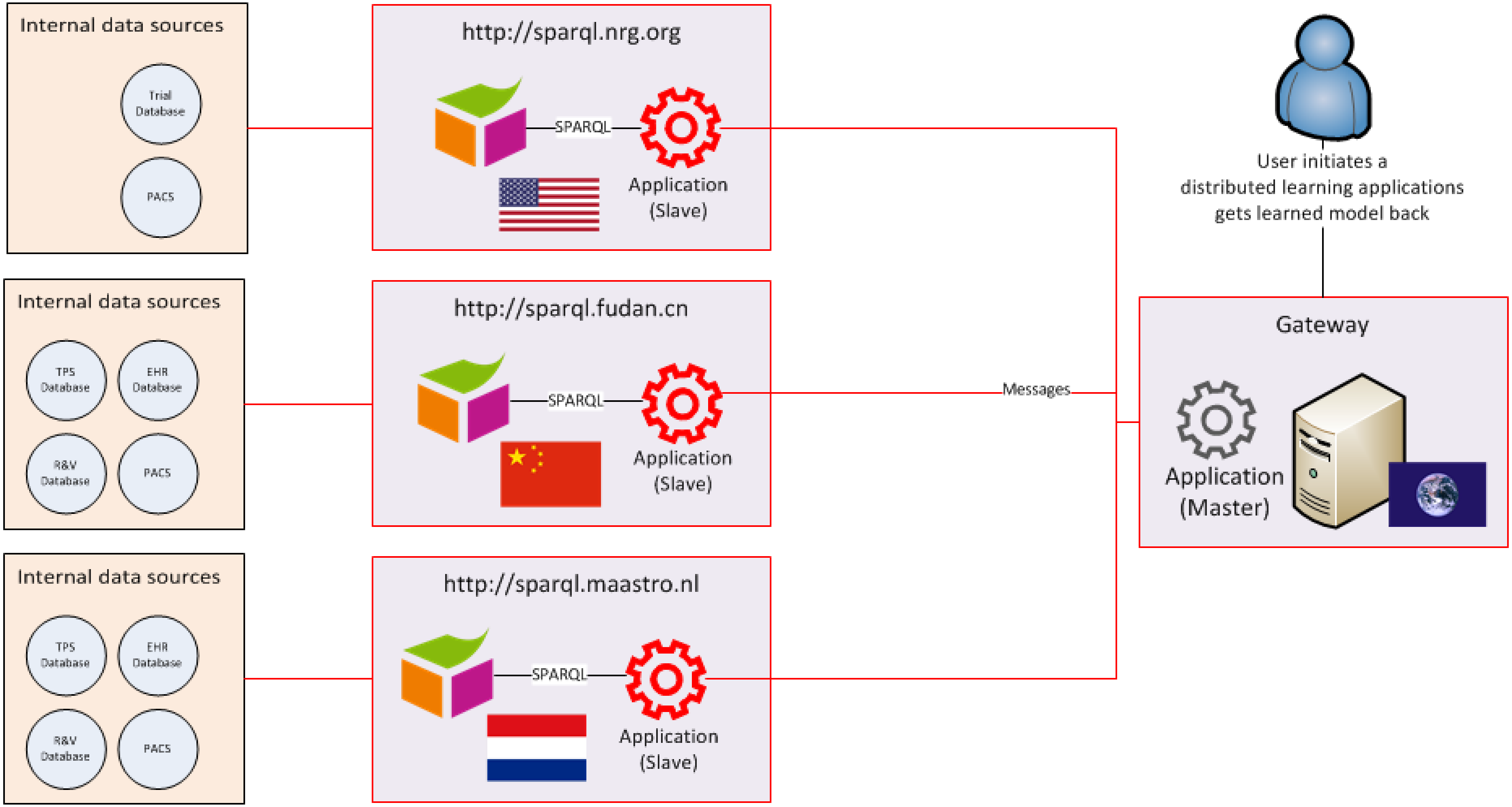


Ontology is a set terms & th **relationships**.

Then we have "machine readable data" accessible to Artificial Intelligence

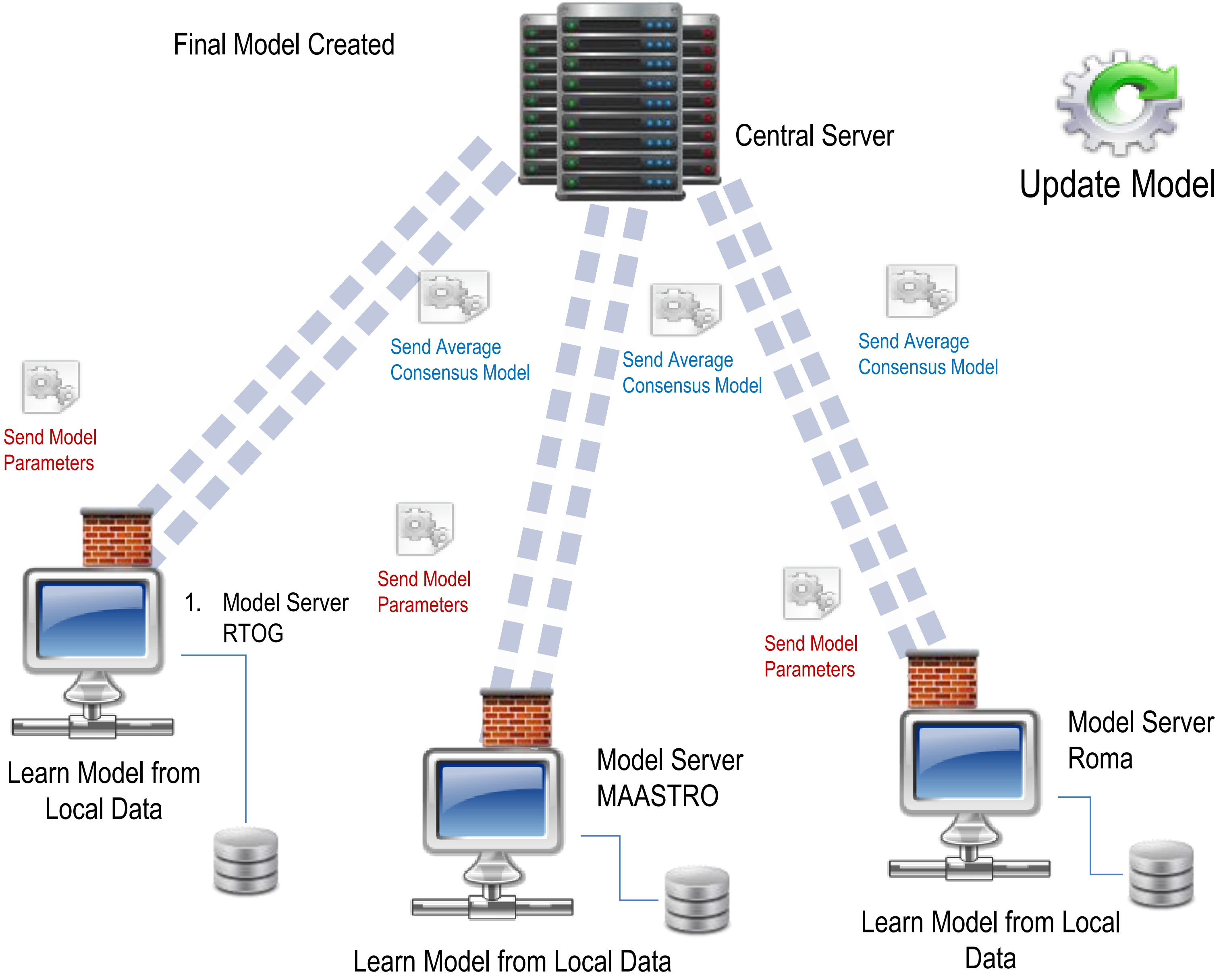
Hospitals

Semantic box *(the secondary research database*



SPARQL : Query language for application

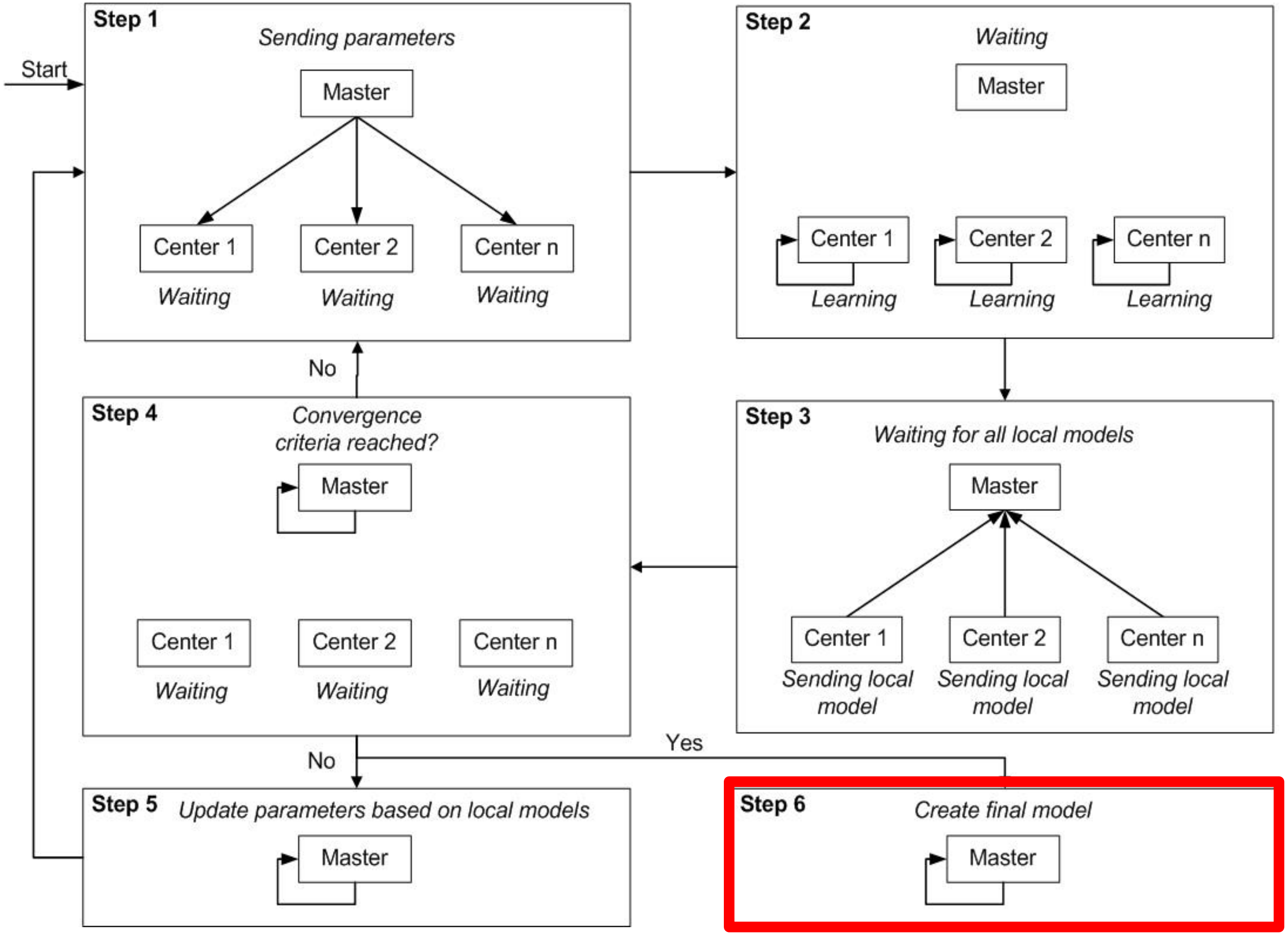
Distributed Learning Architecture


Final Model Created
Central Server
Update Model
Send Average Consensus Model
Send Average Consensus Model
Send Average Consensus Model
Send Model Parameters
Send Model Parameters
Send Model Parameters
1. Model Server RTOG
Learn Model from Local Data
Model Server MAASTRO
Learn Model from Local Data
Model Server Roma
Learn Model from Local Data

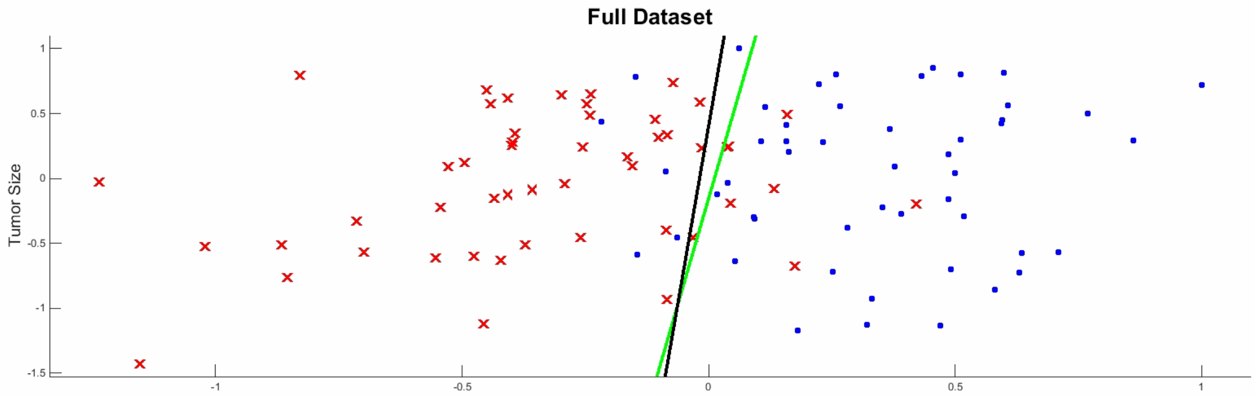
Only aggregate data is exchanged between the Central Server and the local Servers

Distributed learning: more real

>500x



Visualization of Distributed Learning: Support Vector Machines



✖ Event Patient

● Non-Event Patient

╲ Distributed Learning Solution

╲ Centralized Learning Solution

Simulated Data

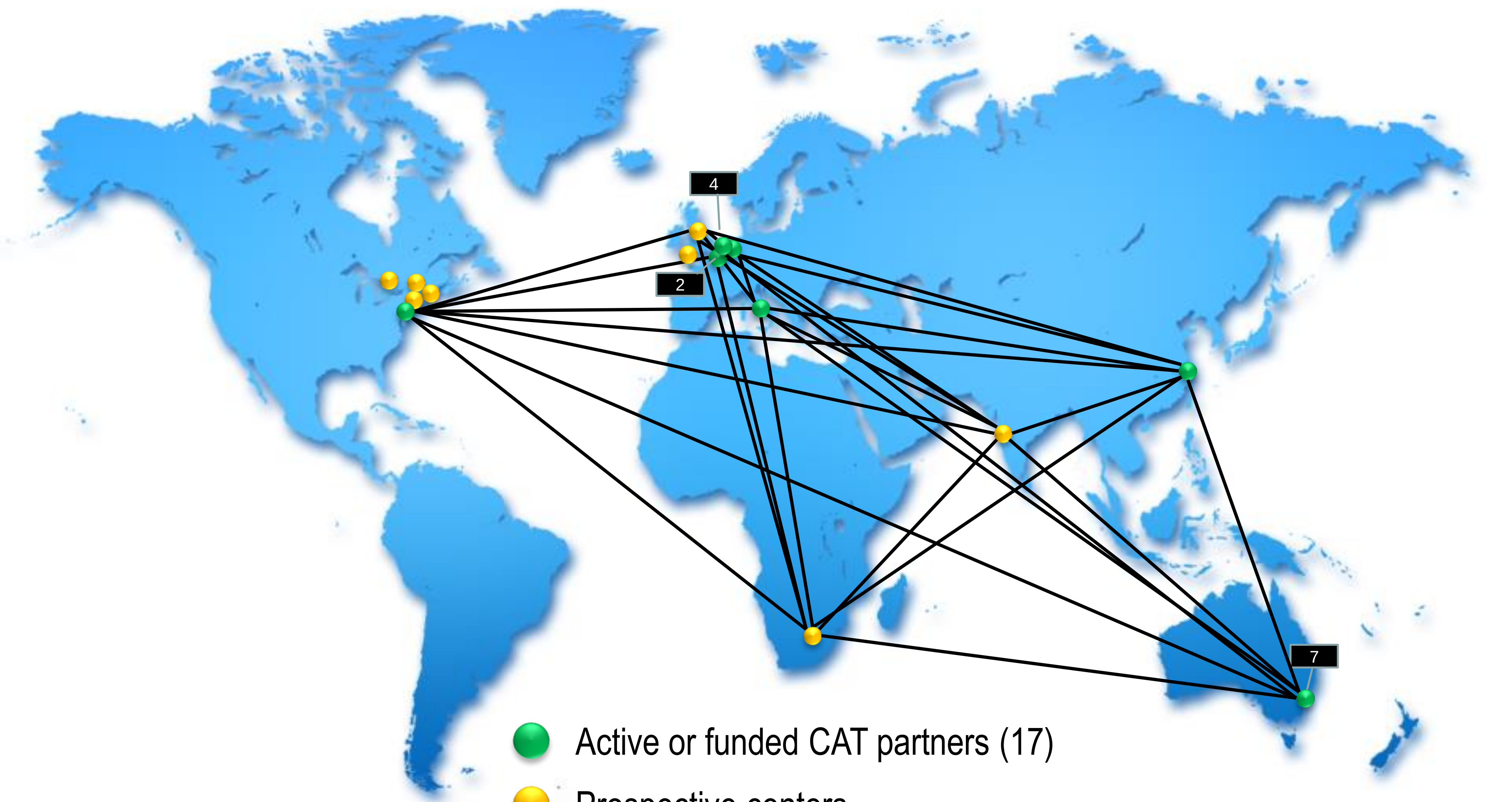
Does all of that work ? euroCAT's example

- Distributed learning = Centralized learning
- Distributed learning better than learning on individual data

Learn in	Validate in	AUC
Aachen (n=7)	Liège (n=186)	0.61
Eindhoven (n=32)	Liège (n=186)	0.72
Hasselt (n=45)	Liège (n=186)	0.68
Maastricht (n=52)	Liège (n=186)	0.75
Alle 4 samen (n=136)	Liège (n=186)	**0.77**
Alle 5 samen (n=322)	World (n=*inf*)	?

- 550 iterations, two hours (centralized < 1 min)

Funded: euroCAT, duCAT, chinaCAT, VATE, ozCAT
New: ukCAT, indiaCAT



● Active or funded CAT partners (17)

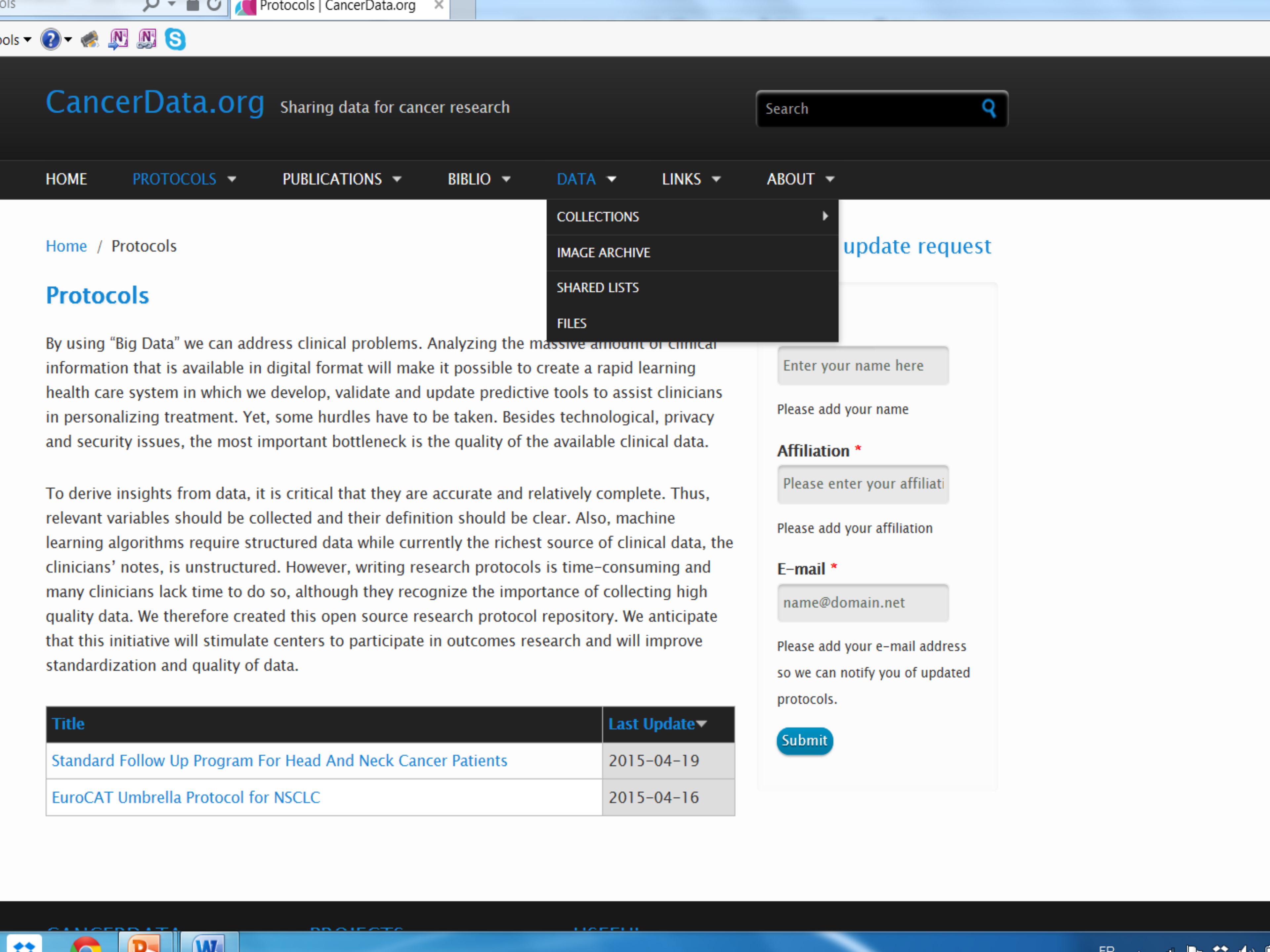
● Prospective centers

Map from cgadvertising.com

Can we improve the quality of the data?

Yes 1) with automated check,
2) validated procedure to imputate missing data (« Amazone type ») and
3) with standardized follow-up protocol

Meldolesi E. et al. Radiother Oncol. 2014



Protocols

By using “Big Data” we can address clinical problems. Analyzing the massive amount of clinical information that is available in digital format will make it possible to create a rapid learning health care system in which we develop, validate and update predictive tools to assist clinicians in personalizing treatment. Yet, some hurdles have to be taken. Besides technological, privacy and security issues, the most important bottleneck is the quality of the available clinical data.

To derive insights from data, it is critical that they are accurate and relatively complete. Thus, relevant variables should be collected and their definition should be clear. Also, machine learning algorithms require structured data while currently the richest source of clinical data, the clinicians’ notes, is unstructured. However, writing research protocols is time-consuming and many clinicians lack time to do so, although they recognize the importance of collecting high quality data. We therefore created this open source research protocol repository. We anticipate that this initiative will stimulate centers to participate in outcomes research and will improve standardization and quality of data.

Title	Last Update▼
Standard Follow Up Program For Head And Neck Cancer Patients	2015-04-19
EuroCAT Umbrella Protocol for NSCLC	2015-04-16

update request

Enter your name here

Please add your name

Affiliation *

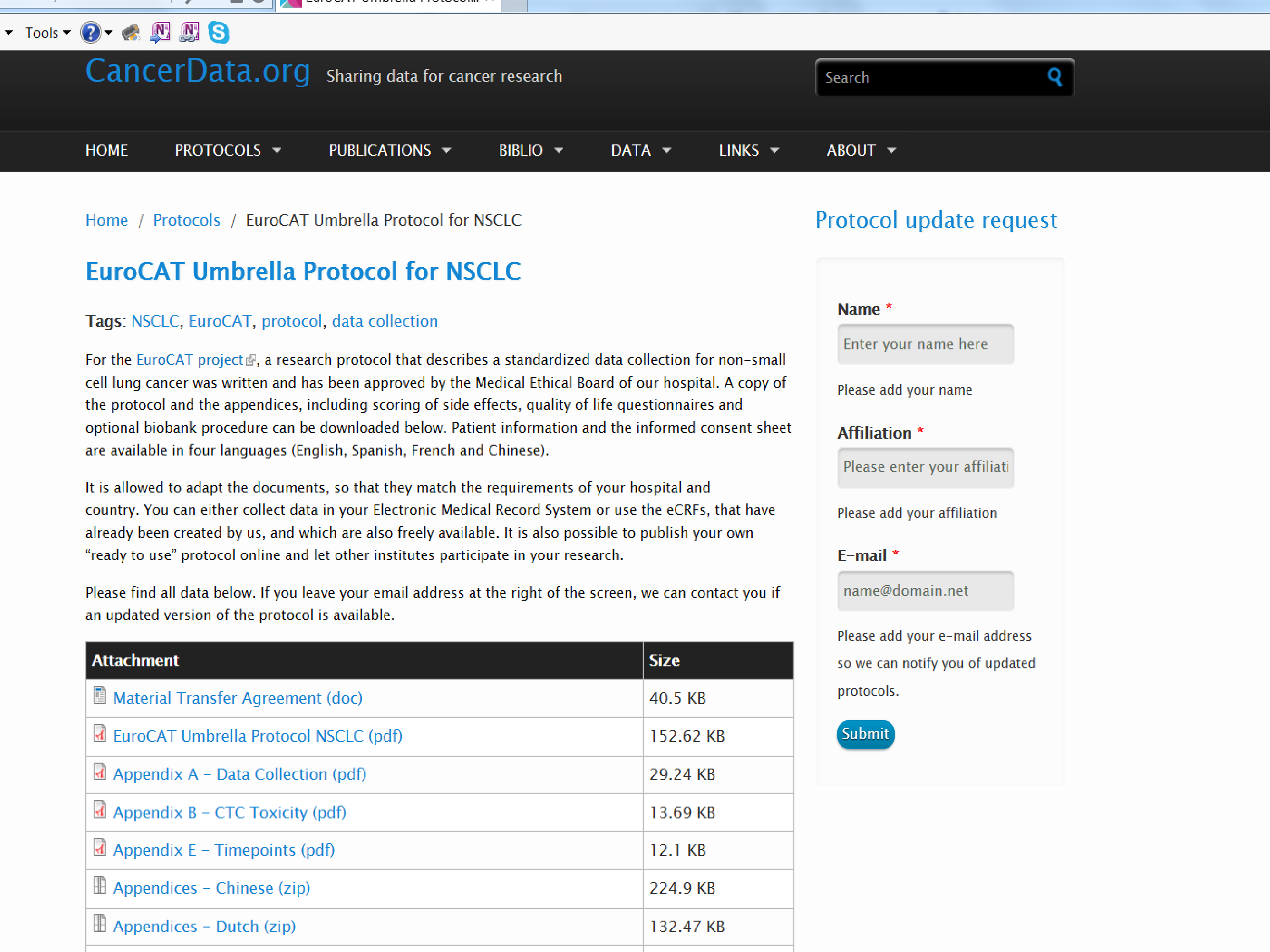
Please enter your affiliati

Please add your affiliation

E-mail *

name@domain.net

Please add your e-mail address so we can notify you of updated protocols.



Home / Protocols / EuroCAT Umbrella Protocol for NSCLC

EuroCAT Umbrella Protocol for NSCLC

Tags: NSCLC, EuroCAT, protocol, data collection

For the EuroCAT project, a research protocol that describes a standardized data collection for non-small cell lung cancer was written and has been approved by the Medical Ethical Board of our hospital. A copy of the protocol and the appendices, including scoring of side effects, quality of life questionnaires and optional biobank procedure can be downloaded below. Patient information and the informed consent sheet are available in four languages (English, Spanish, French and Chinese).

It is allowed to adapt the documents, so that they match the requirements of your hospital and country. You can either collect data in your Electronic Medical Record System or use the eCRFs, that have already been created by us, and which are also freely available. It is also possible to publish your own "ready to use" protocol online and let other institutes participate in your research.

Please find all data below. If you leave your email address at the right of the screen, we can contact you if an updated version of the protocol is available.

Attachment	Size
Material Transfer Agreement (doc)	40.5 KB
EuroCAT Umbrella Protocol NSCLC (pdf)	152.62 KB
Appendix A – Data Collection (pdf)	29.24 KB
Appendix B – CTC Toxicity (pdf)	13.69 KB
Appendix E – Timepoints (pdf)	12.1 KB
Appendices – Chinese (zip)	224.9 KB
Appendices – Dutch (zip)	132.47 KB

Protocol update request

Name *

Enter your name here

Please add your name

Affiliation *

Please enter your affiliati

Please add your affiliation

E-mail *

name@domain.net

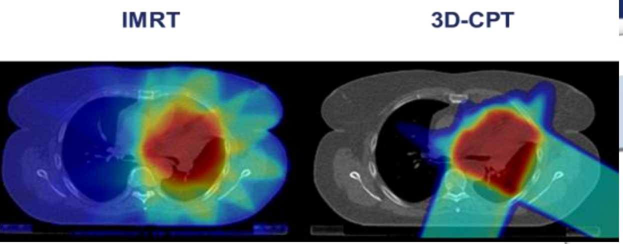
Please add your e-mail address so we can notify you of updated protocols.

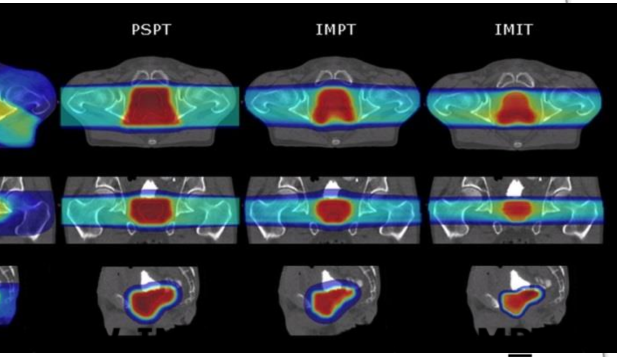
What about

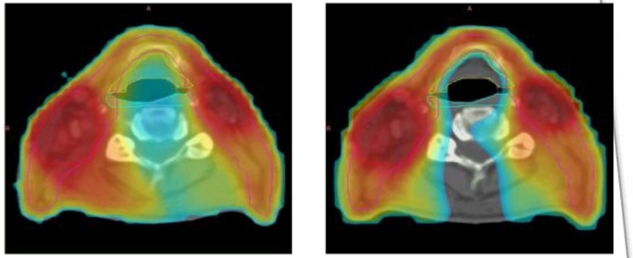
the costs of treatments?

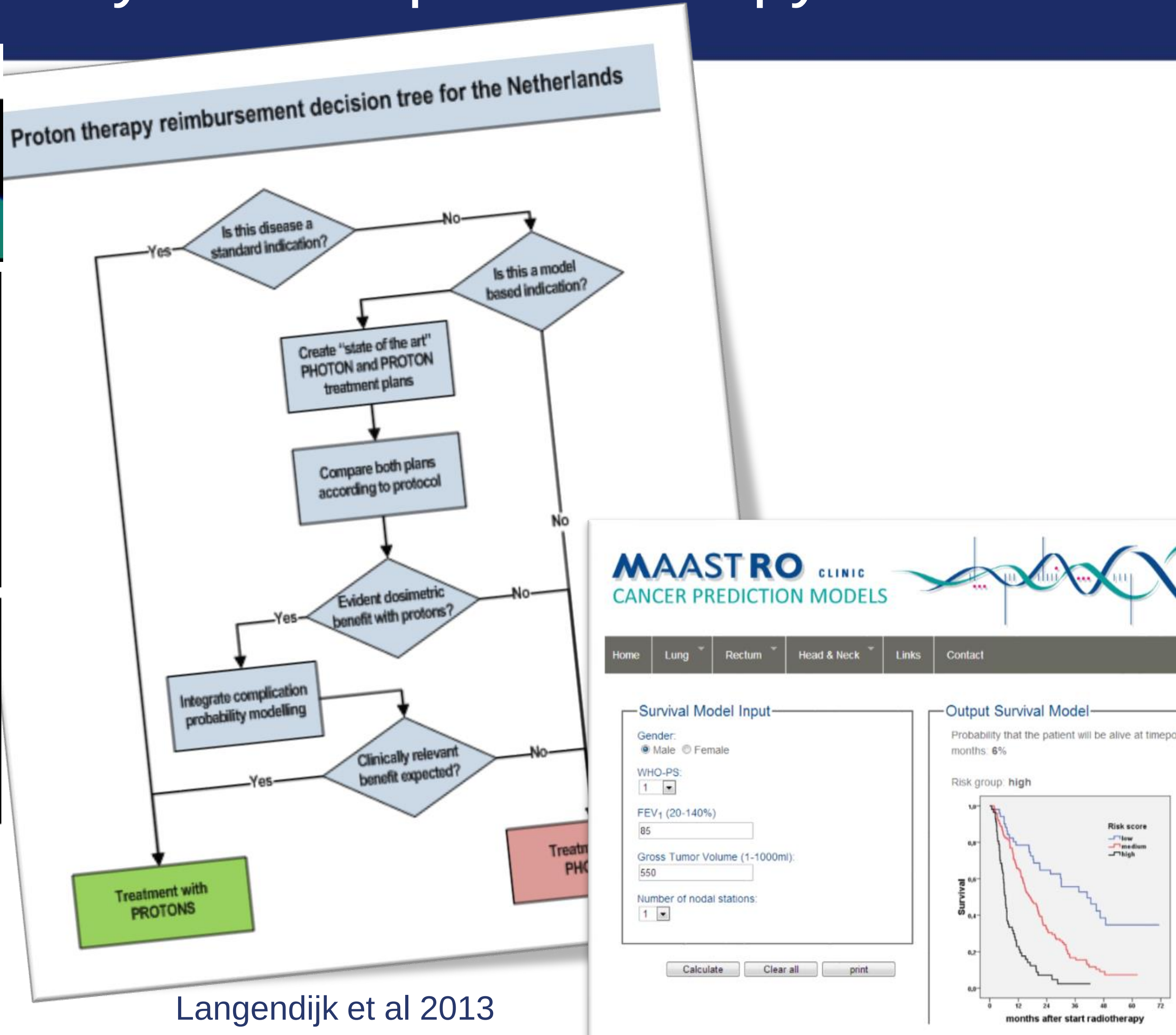
e.g. Protontherapy…

Protons *without* dose escalation: model-based decision support system for protontherapy








Proton therapy reimbursement decision tree for the Netherlands
Is this disease a standard indication?
Yes
No
Is this a model based indication?
Create "state of the art" PHOTON and PROTON treatment plans
Compare both plans according to protocol
No
Evident dosimetric benefit with protons?
Yes
No
Integrate complication probability modelling
Clinically relevant benefit expected?
Yes
No
Treatment with PROTONS
Treatm PHO
MAASTRO CLINIC
CANCER PREDICTION MODELS
Home
Lung
Rectum
Head & Neck
Links
Contact
Calculate
Clear all
print
Risk score
low
medium
high
Survival
months after start radiotherapy
1,0
0,8
0,6
0,4
0,2
0,0
0
12
24
36
48
60
72

Survival Model Input

Gender:
● Male ○ Female

WHO-PS:
1

FEV$_1$ (20-140%)
85

Gross Tumor Volume (1-1000ml):
550

Number of nodal stations:
1

Output Survival Model

Probability that the patient will be alive at timepo
months: 6%

Risk group: high

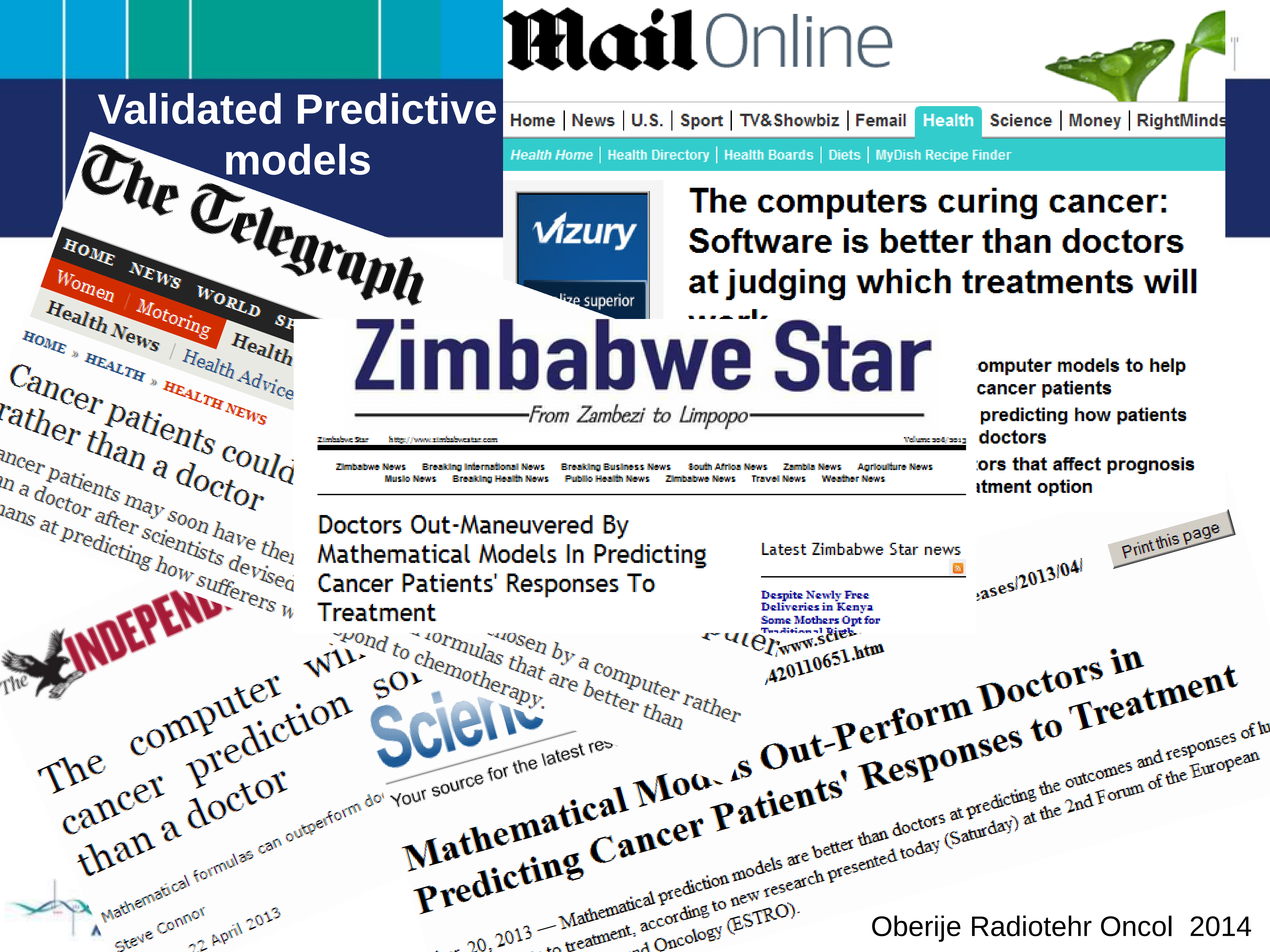
Langendijk et al 2013

Can you give me

examples

of new knowledge coming from Big Data approaches

Validated Predictive models



Mail Online

Home | News | U.S. | Sport | TV&Showbiz | Femail | Health | Science | Money | RightMinds

Health Home | Health Directory | Health Boards | Diets | MyDish Recipe Finder

Vizury

...lize superior

The computers curing cancer: Software is better than doctors at judging which treatments will work

...omputer models to help cancer patients

...predicting how patients ...doctors

...tors that affect prognosis ...atment option

Print this page

...eases/2013/04/

The Telegraph

HOME NEWS WORLD SP...

Women | Motoring | Health

Health News | Health Advice

HOME » HEALTH » HEALTH NEWS

Cancer patients could ... rather than a doctor

...ancer patients may soon have the... ...n a doctor after scientists devised ...ians at predicting how sufferers w...

Zimbabwe Star

From Zambezi to Limpopo

Zimbabwe Star http://www.zimbabwestar.com Volume 208/2013

Zimbabwe News Breaking International News Breaking Business News South Africa News Zambia News Agriculture News Music News Breaking Health News Public Health News Zimbabwe News Travel News Weather News

Doctors Out-Maneuvered By Mathematical Models In Predicting Cancer Patients' Responses To Treatment

Latest Zimbabwe Star news

Despite Newly Free Deliveries in Kenya Some Mothers Opt for Traditional Birth...

...spond to chemotherapy. ...formulas that are better than ...chosen by a computer rather

...puter www.scien... ...420110651.htm

The INDEPEND...

The computer wi... cancer prediction so... than a doctor

Mathematical formulas can outperform do...

Steve Connor

22 April 2013

Scienc...

Your source for the latest res...

Mathematical Mod...s Out-Perform Doctors in Predicting Cancer Patients' Responses to Treatment

...r 20, 2013 — Mathematical prediction models are better than doctors at predicting the outcomes and responses of lu... ...to treatment, according to new research presented today (Saturday) at the 2nd Forum of the European ...nd Oncology (ESTRO).

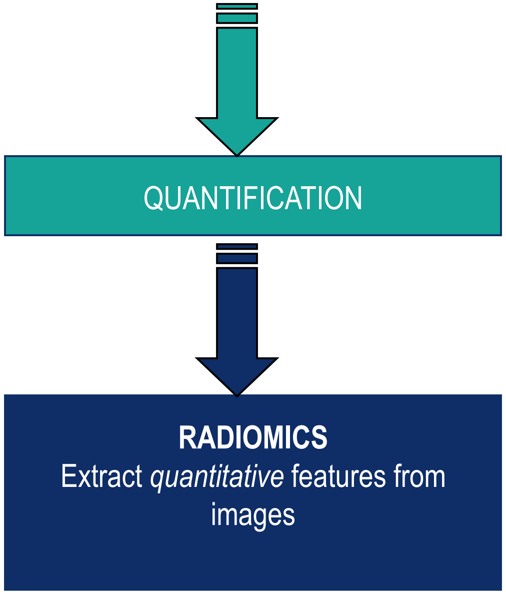
Oberije Radiotehr Oncol 2014

The Radiomic hypothesis

One can extract *more* quantitative information from standard imaging

Radiology:

- Implicit knowledge
- Interpretability

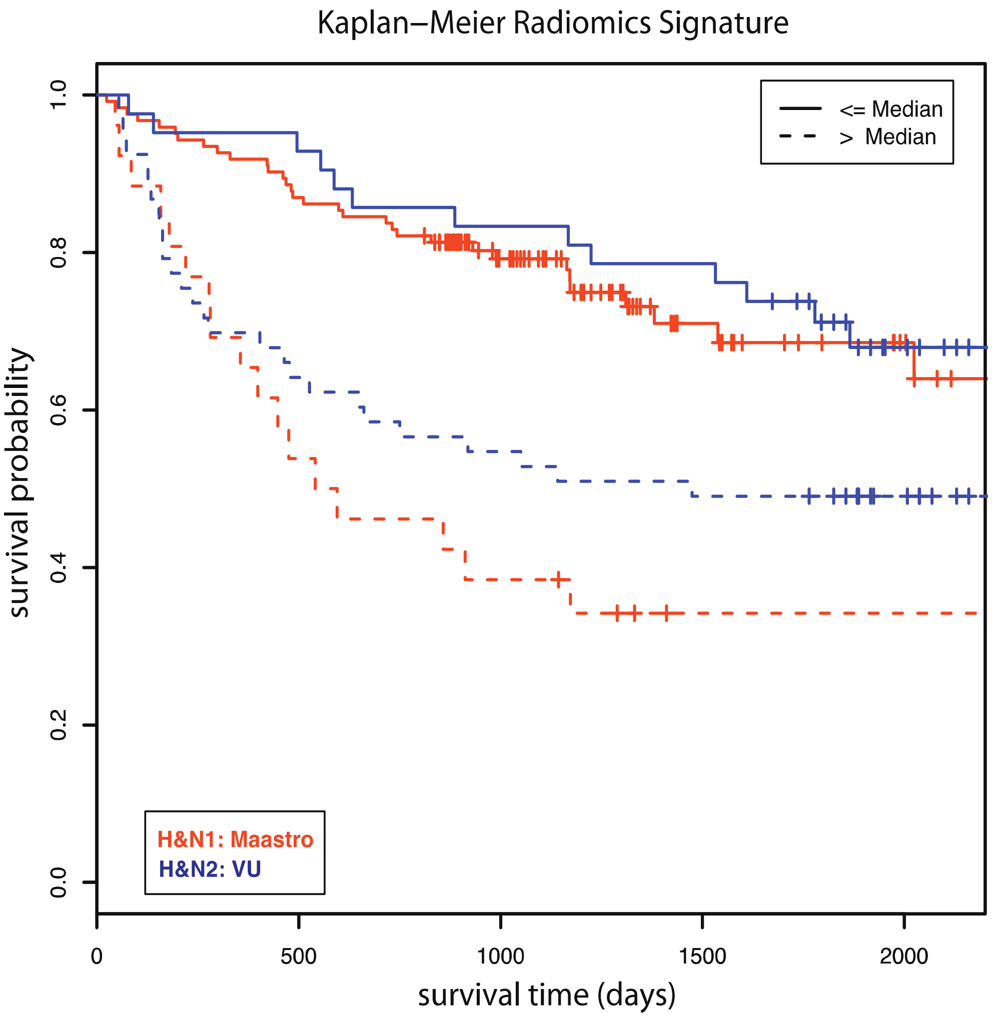


QUANTIFICATION

RADIOMICS
Extract *quantitative* features from images

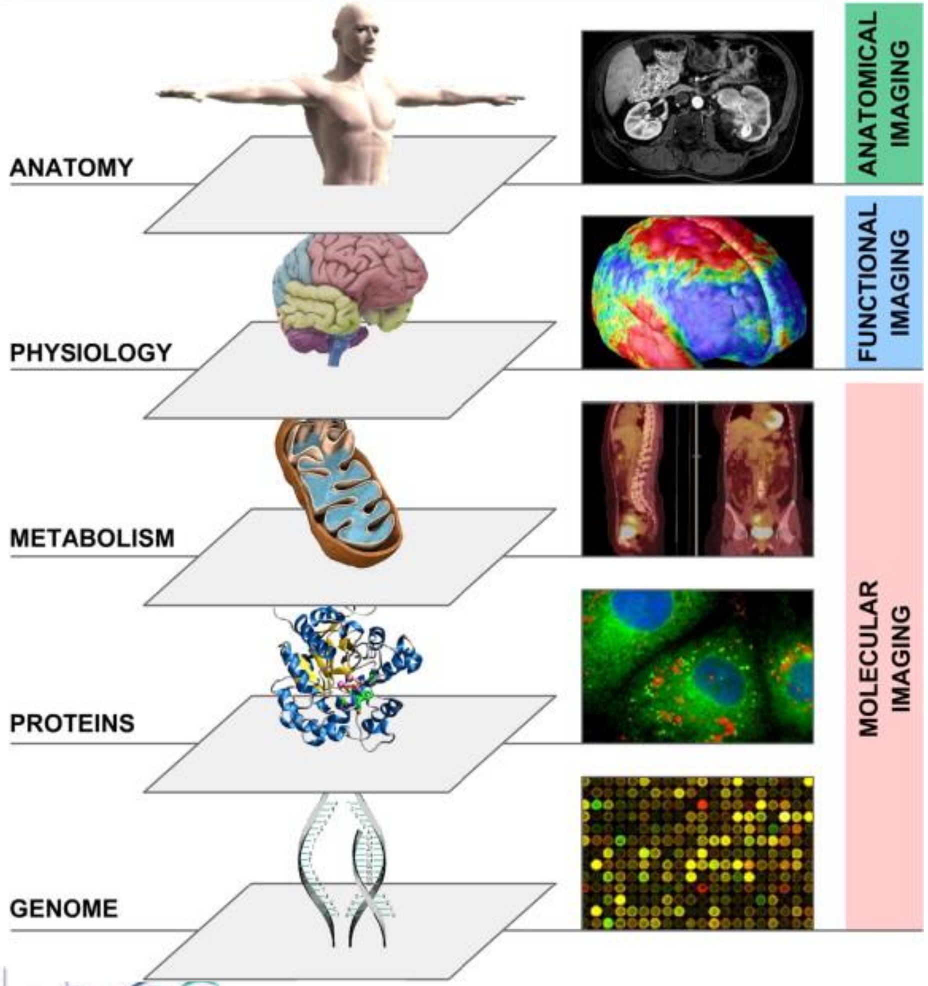
Lambin et al. EJC, 2012; Aerts, Lambin et al. Nature Commun 2014

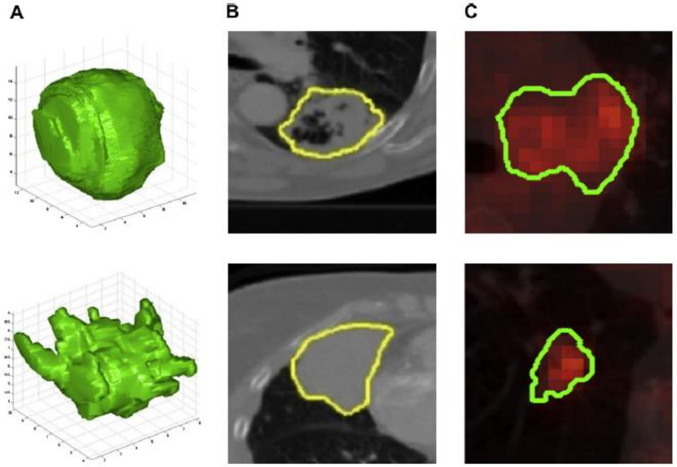
Predict survival in Lung and Head & neck cancer better then TNM

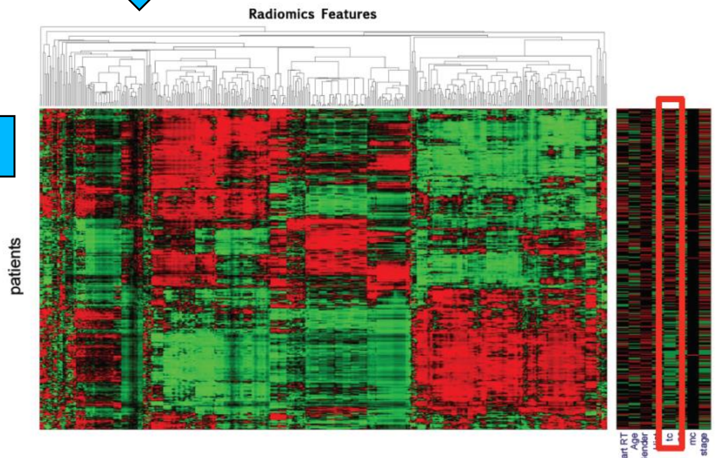


Aerts…Lambin, Nature Commun 2014; Leijenaar et al. Acta Oncol 2015

Entering the OMICS era… Radiomics



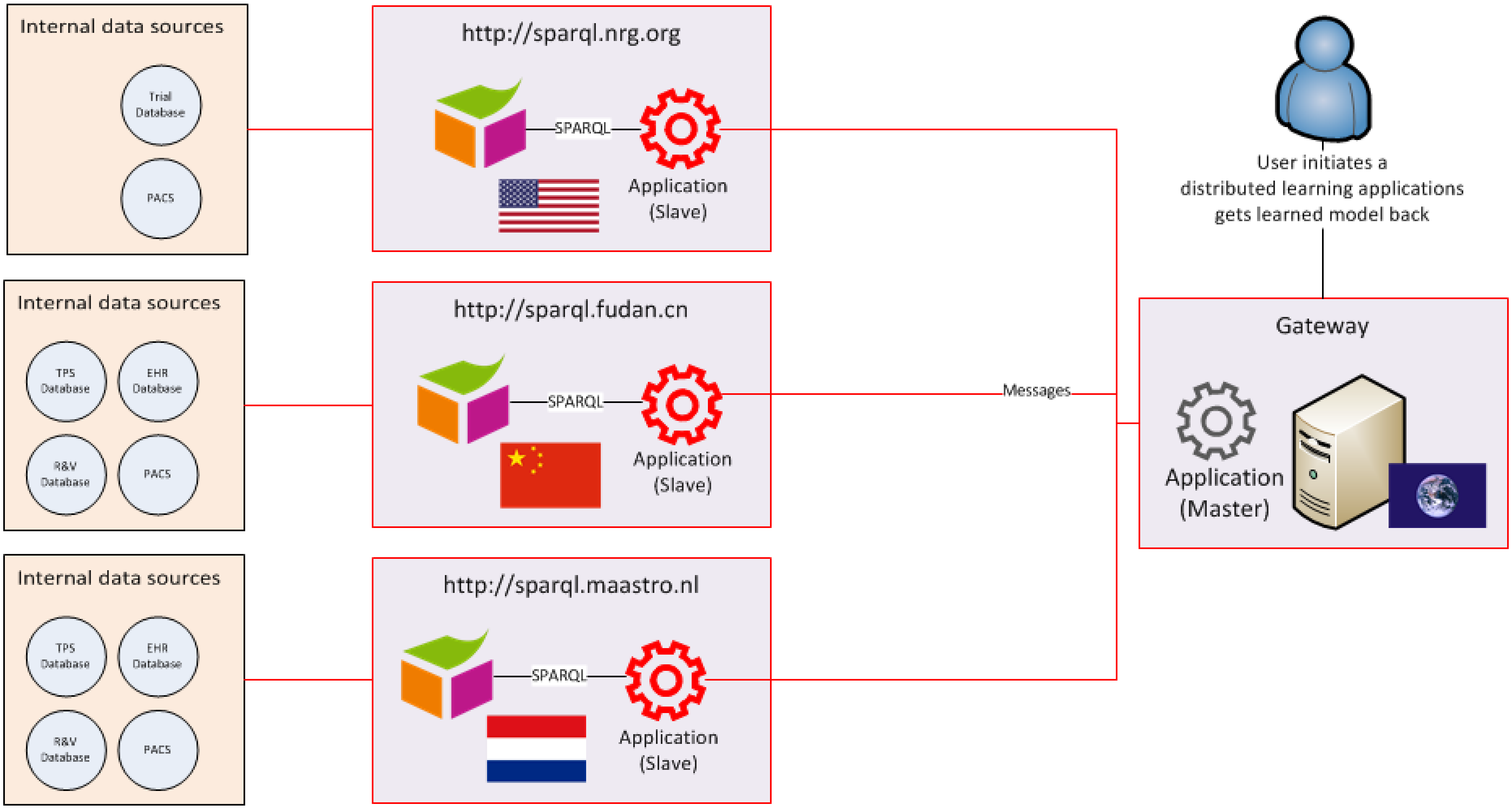




Lambin et al. EJC, 2012; Aerts, Lambin et al. Nature Commun 2014

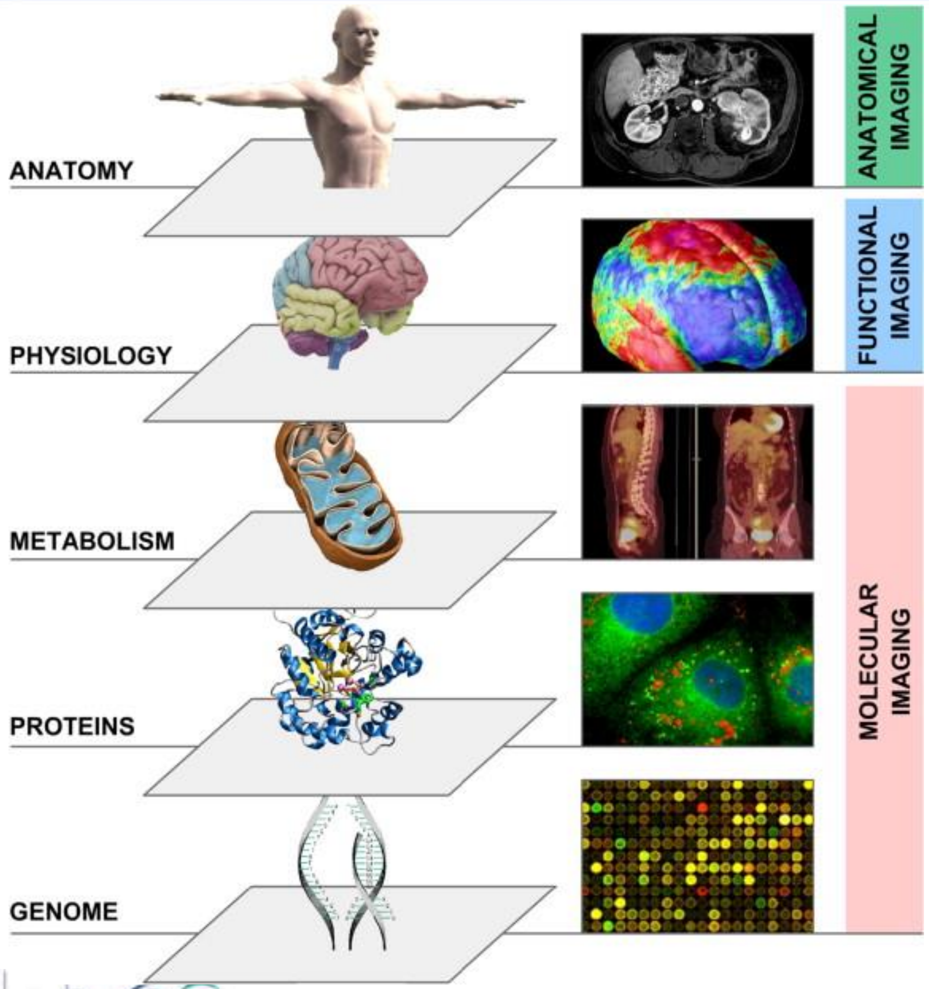
Distributed learning for Radiomics

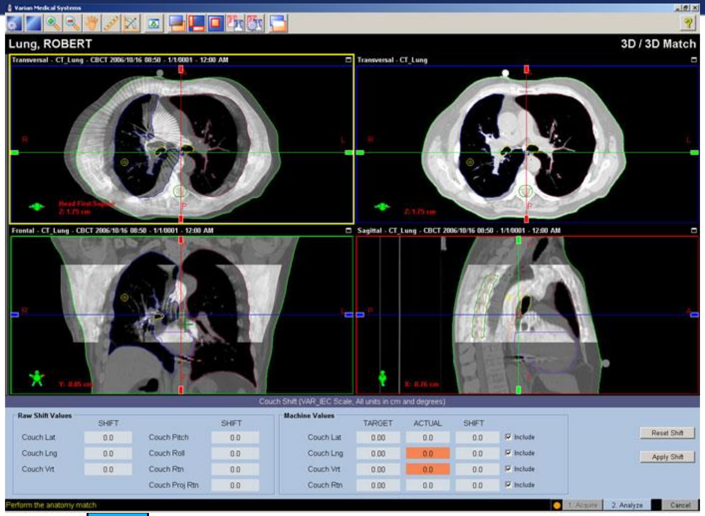
Hospitals *Semantic box*



SPARQL : Query language for application initiating an action

What’s new in Radiomics? Quantify tumour biology on *Cone Beam CT*




Varian Medical Systems
Lung, ROBERT
3D / 3D Match
Transversal - CT_Lung - CBCT 2006/10/16 08:50 - 1/1/0001 - 12:00 AM
Transversal - CT_Lung
Head First Supine
Z: 1.75 cm
Z: 1.75 cm
Frontal - CT_Lung - CBCT 2006/10/16 08:50 - 1/1/0001 - 12:00 AM
Sagittal - CT_Lung - CBCT 2006/10/16 08:50 - 1/1/0001 - 12:00 AM
Y: 0.05 cm
X: 0.76 cm

Couch Shift (VAR_IEC Scale, All units in cm and degrees)

Raw Shift Values

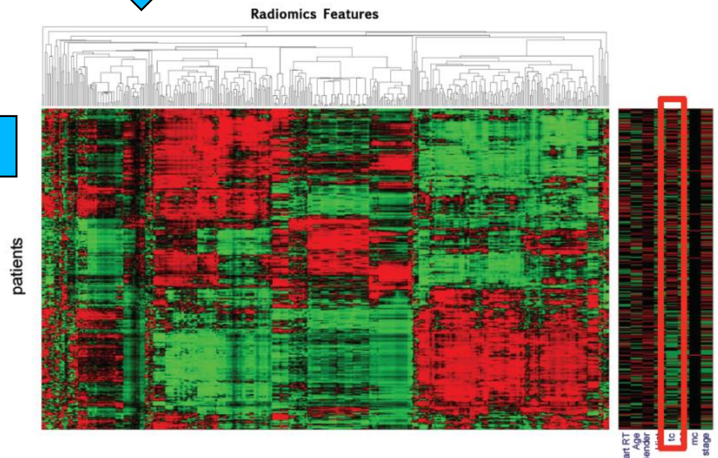
	SHIFT		SHIFT
Couch Lat	0.0	Couch Pitch	0.0
Couch Lng	0.0	Couch Roll	0.0
Couch Vrt	0.0	Couch Rtn	0.0
		Couch Proj Rtn	0.0

Machine Values

	TARGET	ACTUAL	SHIFT	
Couch Lat	0.00	0.0	0.0	☑ Include
Couch Lng	0.00	0.0	0.0	☑ Include
Couch Vrt	0.00	0.0	0.0	☑ Include
Couch Rtn	0.00	0.0	0.0	☑ Include

Reset Shift
Apply Shift
1. Acquire
2. Analyze
Cancel

Perform the anatomy match



Janita van Timmeren et al, Oral commun. Thursday ICTR

What about the patient?

The 5 P’s of modern medicine

(from Leroy Hood)

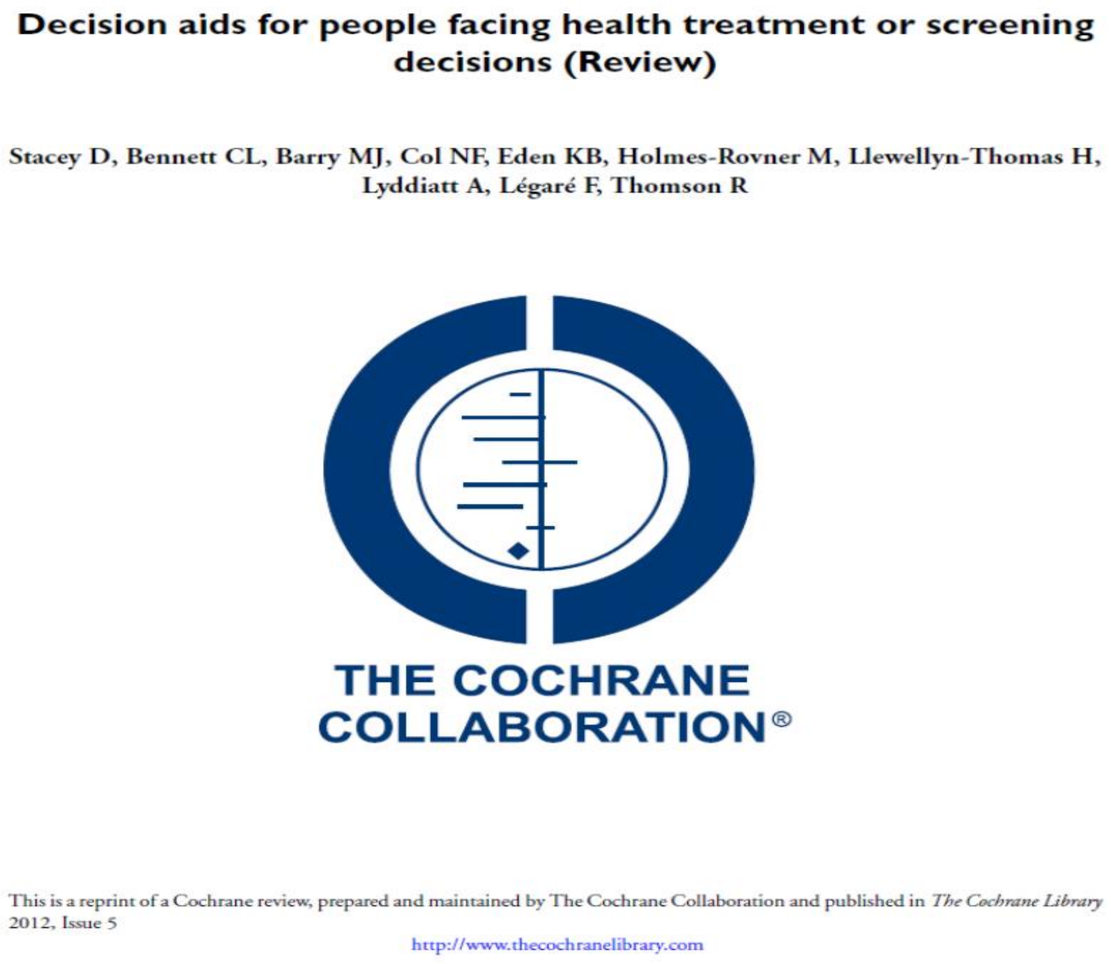
« P » for Personalized

« P » for Preventive

« P » for Predictive

« P » for *Participatory*

Shared Decision Making 1.0 with Decision aids



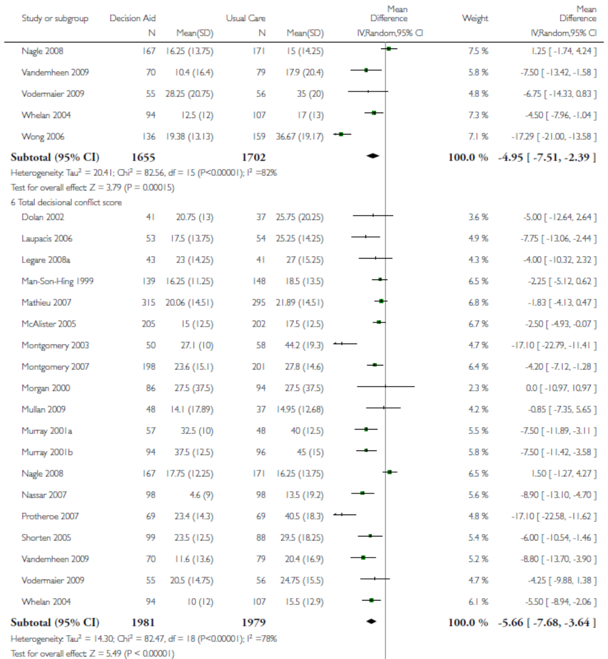
Decision aids for people facing health treatment or screening decisions (Review)

Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt A, Légaré F, Thomson R

THE COCHRANE COLLABORATION®

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2012, Issue 5

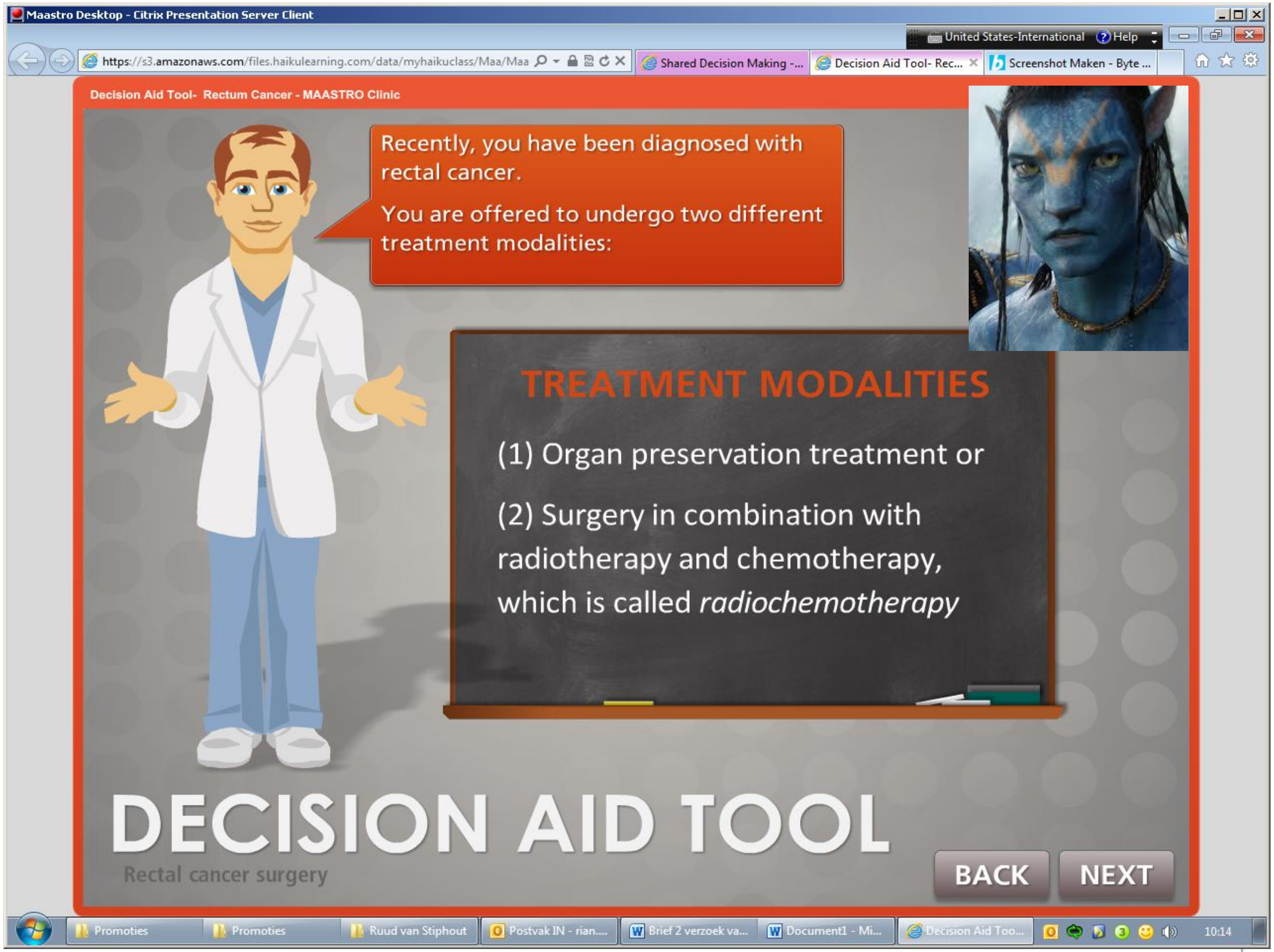
http://www.thecochranelibrary.com



Study or subgroup	Decision Aid		Usual Care		Mean Difference	Weight	Mean Difference
	N	Mean(SD)	N	Mean(SD)	IV,Random,95% CI		IV,Random,95% CI
Nagle 2008	167	16.25 (13.75)	171	15 (14.25)		7.5 %	1.25 [-1.74, 4.24]
Vandemheen 2009	70	10.4 (16.4)	79	17.9 (20.4)		5.8 %	-7.50 [-13.42, -1.58]
Vodermaier 2009	55	28.25 (20.75)	56	35 (20)		4.8 %	-6.75 [-14.33, 0.83]
Whelan 2004	94	12.5 (12)	107	17 (13)		7.3 %	-4.50 [-7.96, -1.04]
Wong 2006	136	19.38 (13.13)	159	36.67 (19.17)		7.1 %	-17.29 [-21.00, -13.58]
Subtotal (95% CI)	**1655**		**1702**			**100.0 %**	**-4.95 [-7.51, -2.39]**
Heterogeneity: Tau² = 20.41; Chi² = 82.56, df = 15 (P<0.00001); I² =82%							
Test for overall effect: Z = 3.79 (P = 0.00015)							
6 Total decisional conflict score							
Dolan 2002	41	20.75 (13)	37	25.75 (20.25)		3.6 %	-5.00 [-12.64, 2.64]
Laupacis 2006	53	17.5 (13.75)	54	25.25 (14.25)		4.9 %	-7.75 [-13.06, -2.44]
Legare 2008a	43	23 (14.25)	41	27 (15.25)		4.3 %	-4.00 [-10.32, 2.32]
Man-Son-Hing 1999	139	16.25 (11.25)	148	18.5 (13.5)		6.5 %	-2.25 [-5.12, 0.62]
Mathieu 2007	315	20.06 (14.51)	295	21.89 (14.51)		6.8 %	-1.83 [-4.13, 0.47]
McAlister 2005	205	15 (12.5)	202	17.5 (12.5)		6.7 %	-2.50 [-4.93, -0.07]
Montgomery 2003	50	27.1 (10)	58	44.2 (19.3)		4.7 %	-17.10 [-22.79, -11.41]
Montgomery 2007	198	23.6 (15.1)	201	27.8 (14.6)		6.4 %	-4.20 [-7.12, -1.28]
Morgan 2000	86	27.5 (37.5)	94	27.5 (37.5)		2.3 %	0.0 [-10.97, 10.97]
Mullan 2009	48	14.1 (17.89)	37	14.95 (12.68)		4.2 %	-0.85 [-7.35, 5.65]
Murray 2001a	57	32.5 (10)	48	40 (12.5)		5.5 %	-7.50 [-11.89, -3.11]
Murray 2001b	94	37.5 (12.5)	96	45 (15)		5.8 %	-7.50 [-11.42, -3.58]
Nagle 2008	167	17.75 (12.25)	171	16.25 (13.75)		6.5 %	1.50 [-1.27, 4.27]
Nassar 2007	98	4.6 (9)	98	13.5 (19.2)		5.6 %	-8.90 [-13.10, -4.70]
Protheroe 2007	69	23.4 (14.3)	69	40.5 (18.3)		4.8 %	-17.10 [-22.58, -11.62]
Shorten 2005	99	23.5 (12.5)	88	29.5 (18.25)		5.4 %	-6.00 [-10.54, -1.46]
Vandemheen 2009	70	11.6 (13.6)	79	20.4 (16.9)		5.2 %	-8.80 [-13.70, -3.90]
Vodermaier 2009	55	20.5 (14.75)	56	24.75 (15.5)		4.7 %	-4.25 [-9.88, 1.38]
Whelan 2004	94	10 (12)	107	15.5 (12.9)		6.1 %	-5.50 [-8.94, -2.06]
Subtotal (95% CI)	**1981**		**1979**			**100.0 %**	**-5.66 [-7.68, -3.64]**
Heterogeneity: Tau² = 14.30; Chi² = 82.47, df = 18 (P<0.00001); I² =78%							
Test for overall effect: Z = 5.49 (P < 0.00001)							

Shared Decision Making 2.0: model-based virtual patient or

Avatar-based Shared Decision making



Maastro Desktop - Citrix Presentation Server Client
United States-International
Help
https://s3.amazonaws.com/files.haikulearning.com/data/myhaikuclass/Maa/Maa
Shared Decision Making -...
Decision Aid Tool- Rec...
Screenshot Maken - Byte ...
Promoties
Promoties
Ruud van Stiphout
Postvak IN - rian....
Brief 2 verzoek va...
Document1 - Mi...
Decision Aid Too...
10:14

Decision Aid Tool- Rectum Cancer - MAASTRO Clinic

Recently, you have been diagnosed with rectal cancer.

You are offered to undergo two different treatment modalities:

TREATMENT MODALITIES

(1) Organ preservation treatment or

(2) Surgery in combination with radiotherapy and chemotherapy, which is called *radiochemotherapy*

DECISION AID TOOL

Rectal cancer surgery

BACK NEXT

Universiteit Maastricht

www.treatmentchoice.info

MAASTRO

What next?: The patient managing its own data

Data =

Our vision in 2 min:

“from hospital to patient”

Personal Health Train

Take home message

1. We need Decision Support Systems (DSS = a “meta TPS”) to manage the large quantity of data and implement Personalized medicine in radiotherapy in particular for protontherapy due to its costs.

2. Two complementary approaches: conventional clinical trials (+ data reuse) + “Big Data approach” (Rapid Learning Health Care).

3. Building cancer informatics tools to enable analysis, exploration, and rapid evaluation of novel therapies or stratification e.g. Distributed learning based on semantic web technology.

4. DSS facilitate Share Decision Making, participative precision medicine and cost effective Health care (the 4th & 5th “P”). One key example could be protontherapy.

Lambin *et al.* Aug 28, Radiother Oncol 2013, Adv. Drug Dev 2016

Acknowledgements

- Policlinico Gemelli, Roma, Italy
- UH Ghent, Belgium
- UH Leuven, Belgium
- UH Nijmegen, Netherlands
- …

- CHU Liege, Belgium
- Uniklinikum Aachen, Germany
- LOC Genk/Hasselt, Belgium
- Catherina Zkh Eindhoven, Netherlands

Main MAASTRO collaborators

- Andre Dekker
- Cary Oberije
- Timo Deist
- Erik Roelofs
- Arthur Jochems
- Sean Walsh
- Ralph Leijenaar
- Janita van Timmeren

Thank you for your attention

More :

www.predictcancer.org

www.eurocat.info

www.cancerdata.org

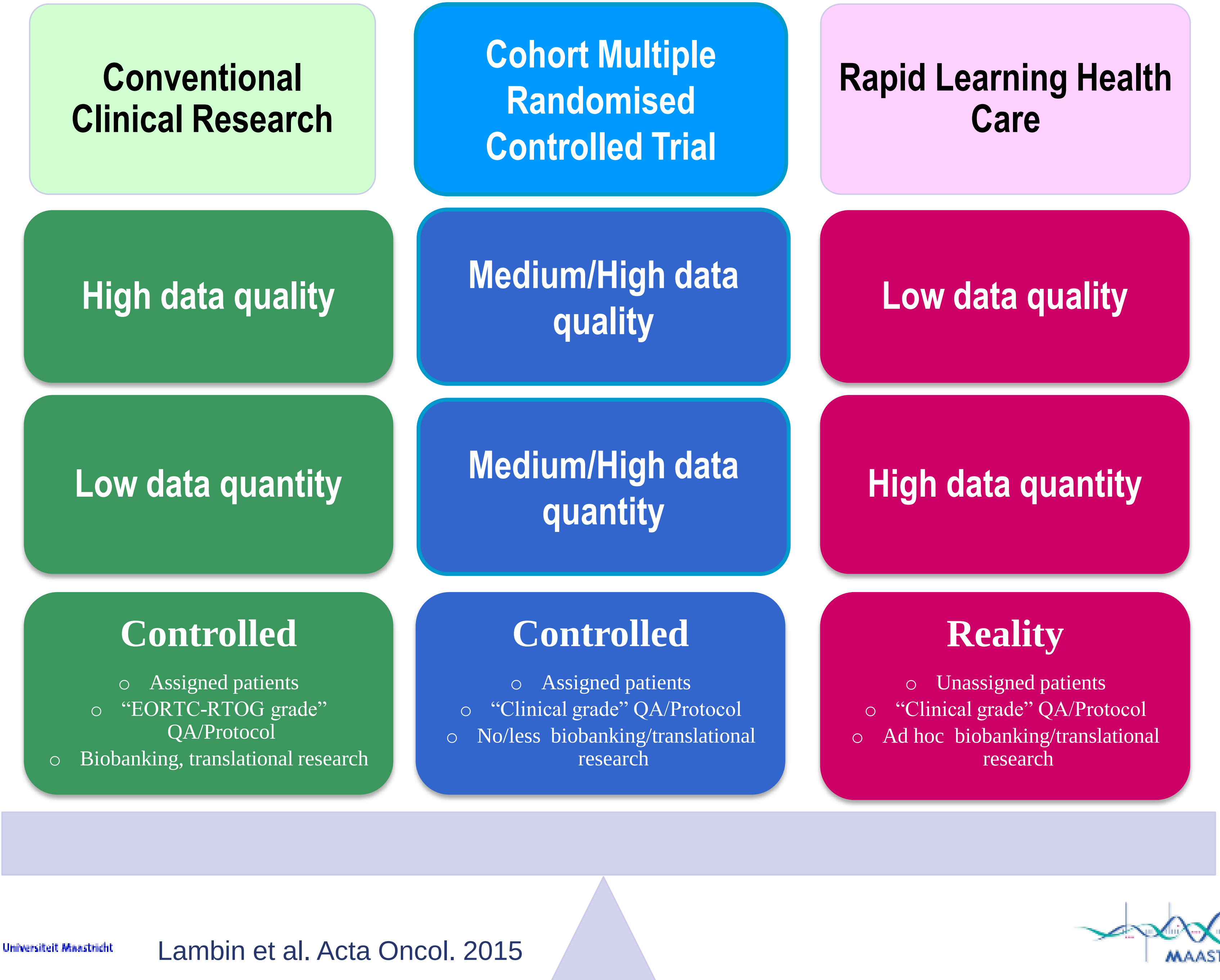
www.mistir.info

www.predictcancer.org

What about

expensive new treatment?

e.g. Protontherapy, Immunotherapy…
We *need* randomized trials to convince the payers and the 2d line specialists.



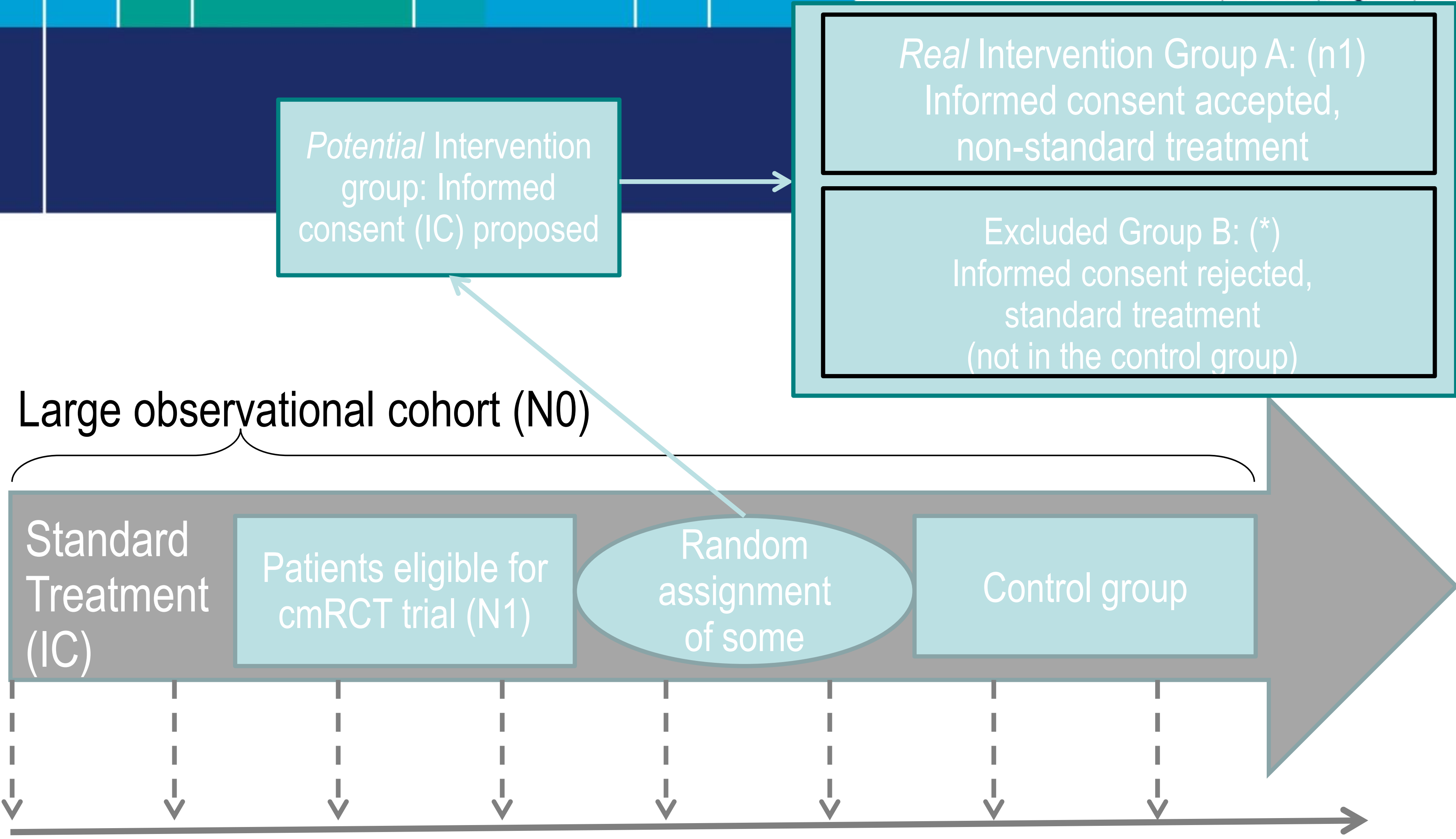
Conventional Clinical Research	Cohort Multiple Randomised Controlled Trial	Rapid Learning Health Care
High data quality	Medium/High data quality	Low data quality
Low data quantity	Medium/High data quantity	High data quantity
Controlled ○ Assigned patients ○ “EORTC-RTOG grade” QA/Protocol ○ Biobanking, translational research	**Controlled** ○ Assigned patients ○ “Clinical grade” QA/Protocol ○ No/less biobanking/translational research	**Reality** ○ Unassigned patients ○ “Clinical grade” QA/Protocol ○ Ad hoc biobanking/translational research

Lambin et al. Acta Oncol. 2015

Protons *with* dose escalation: *potential* solution = The cohort multiple randomised controlled trial design

The cohort multiple randomised controlled trial design is a *pragmatic method* taking advantage of the standardized follow-up approaches.

Relton C et al. BMJ. 2010; Burbach et al. Trials 2015; Lambin et al. Acta Oncol 2015



Real Intervention Group A: (n1)
Informed consent accepted,
non-standard treatment

Excluded Group B: (*)
Informed consent rejected,
standard treatment
(not in the control group)

Potential Intervention group: Informed consent (IC) proposed

Large observational cohort (N0)

Standard Treatment (IC)

Patients eligible for cmRCT trial (N1)

Random assignment of some

Control group

Standardized Follow-up Protocol (ideally multicentric): Dr and Patient reported outcomes, imaging…

Relton C et al. BMJ. 2010; Burbach et al. Trials 2015; Lambin et al. Acta Oncol 2015

40 Years After Tuskegee: Reuniting Medical Research and Practice

Ruth Faden (Bioethics) Jan 16 2013, 10:44 AM ET

the Atlantic

Guidelines to protect human research subjects impede efficient generation and exchange of knowledge.

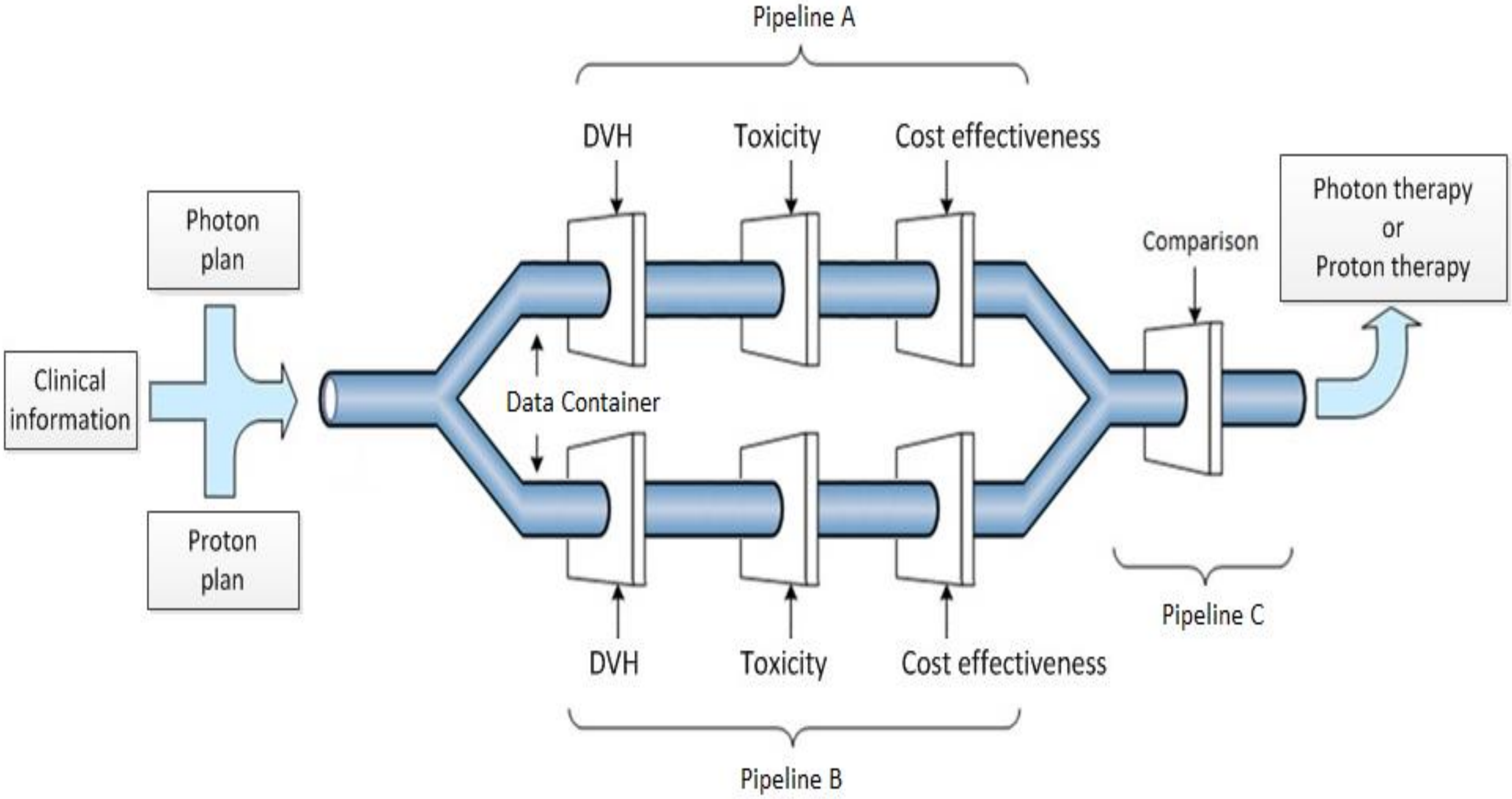
..each episode of care we receive, should generate data and evidence that improve the care of patients who come after us; we then, in turn, benefit from what is systematically learned from the care received by patients who come before us.

care of patients who come after us; we then, in turn, benefit from what is systematically learned from the care received by patients who come before us. Through continuous, real-time learning, we can provide better care to more people, save lives, become smarter, and wring every dollar of value from the system. This is what the Institute of Medicine has dubbed the "learning healthcare system."

MAASTRO

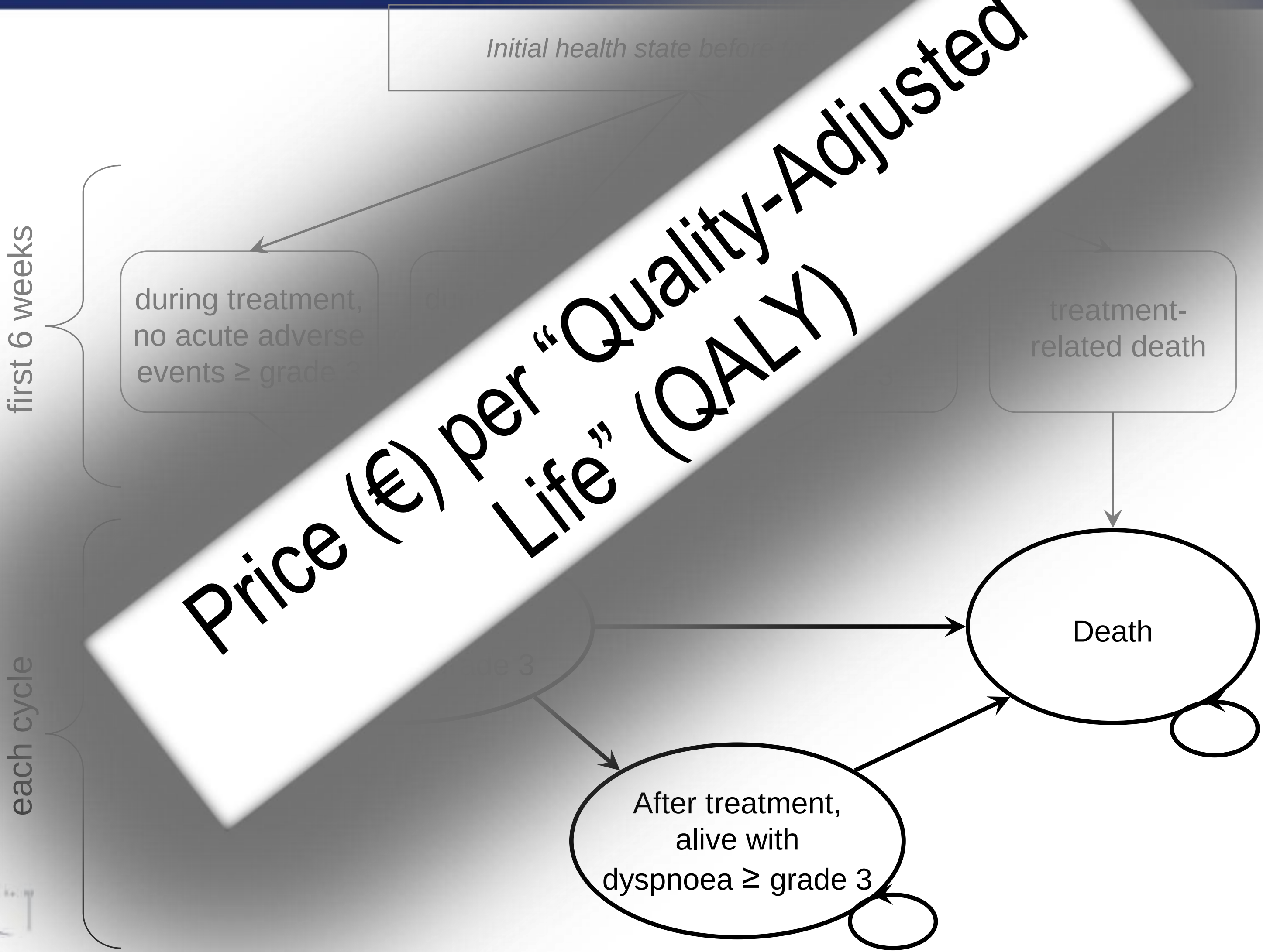
Universiteit Maastricht

PRODECIS: working for head & neck cancer



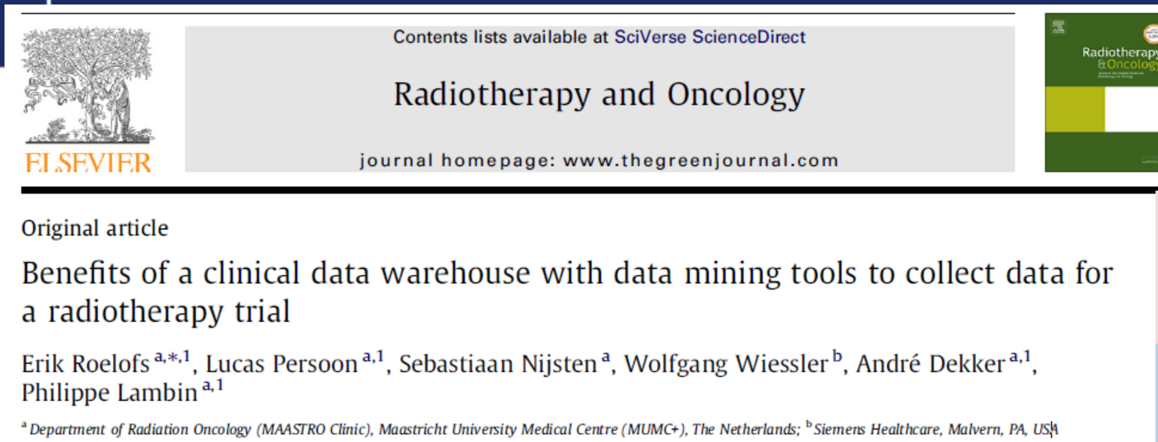
Cheng Q, Roelofs E, Lambin P et al. Submitted

Costs of the whole care cycle: Markov Model


Initial health state befo
first 6 weeks
during treatment, no acute adverse events ≥ grade 3
treatment-related death
each cycle
de 3
Death
After treatment, alive with dyspnoea ≥ grade 3

Price (€) per “Quality-Adjusted Life” (QALY)

Data warehousing for research



Contents lists available at SciVerse ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

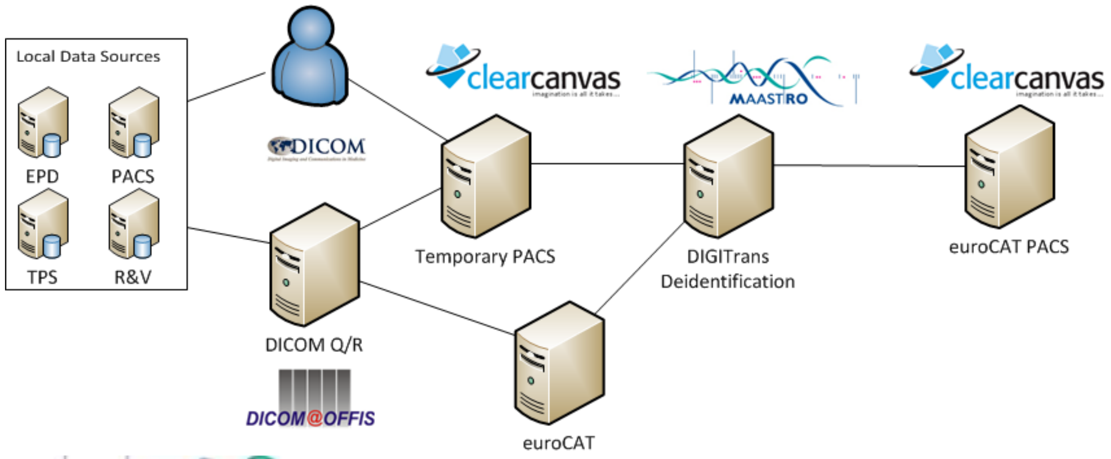
Original article

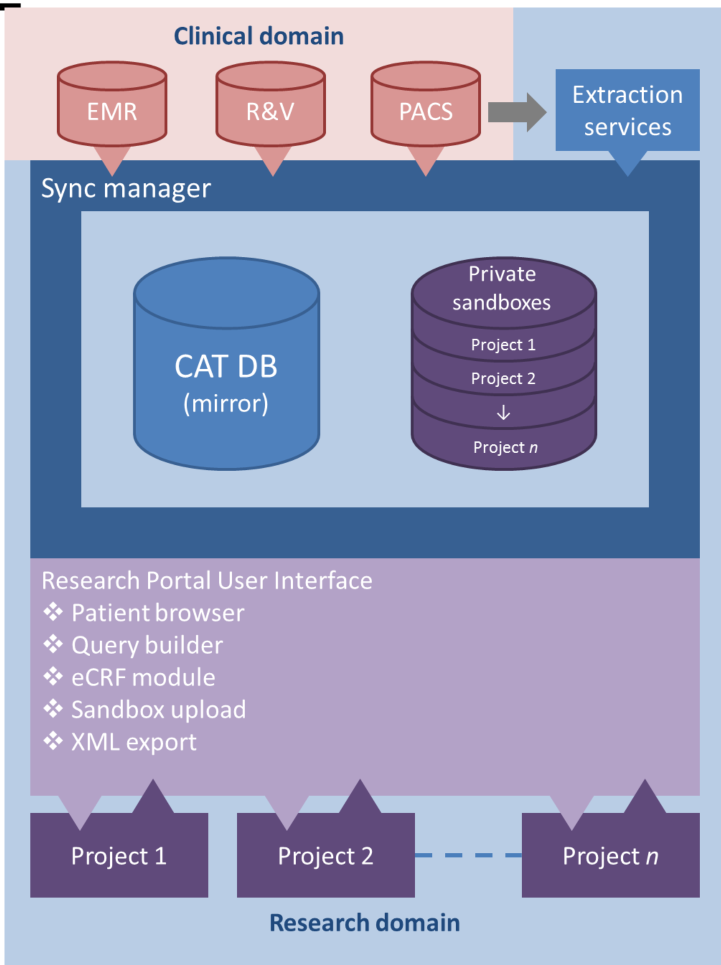
Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial

Erik Roelofs [a,*,1], Lucas Persoon [a,1], Sebastiaan Nijsten [a], Wolfgang Wiessler [b], André Dekker [a,1], Philippe Lambin [a,1]

[a] Department of Radiation Oncology (MAASTRO Clinic), Maastricht University Medical Centre (MUMC+), The Netherlands; [b] Siemens Healthcare, Malvern, PA, USA

euroCAT





Clinical domain

EMR · R&V · PACS → Extraction services

Sync manager

CAT DB (mirror)

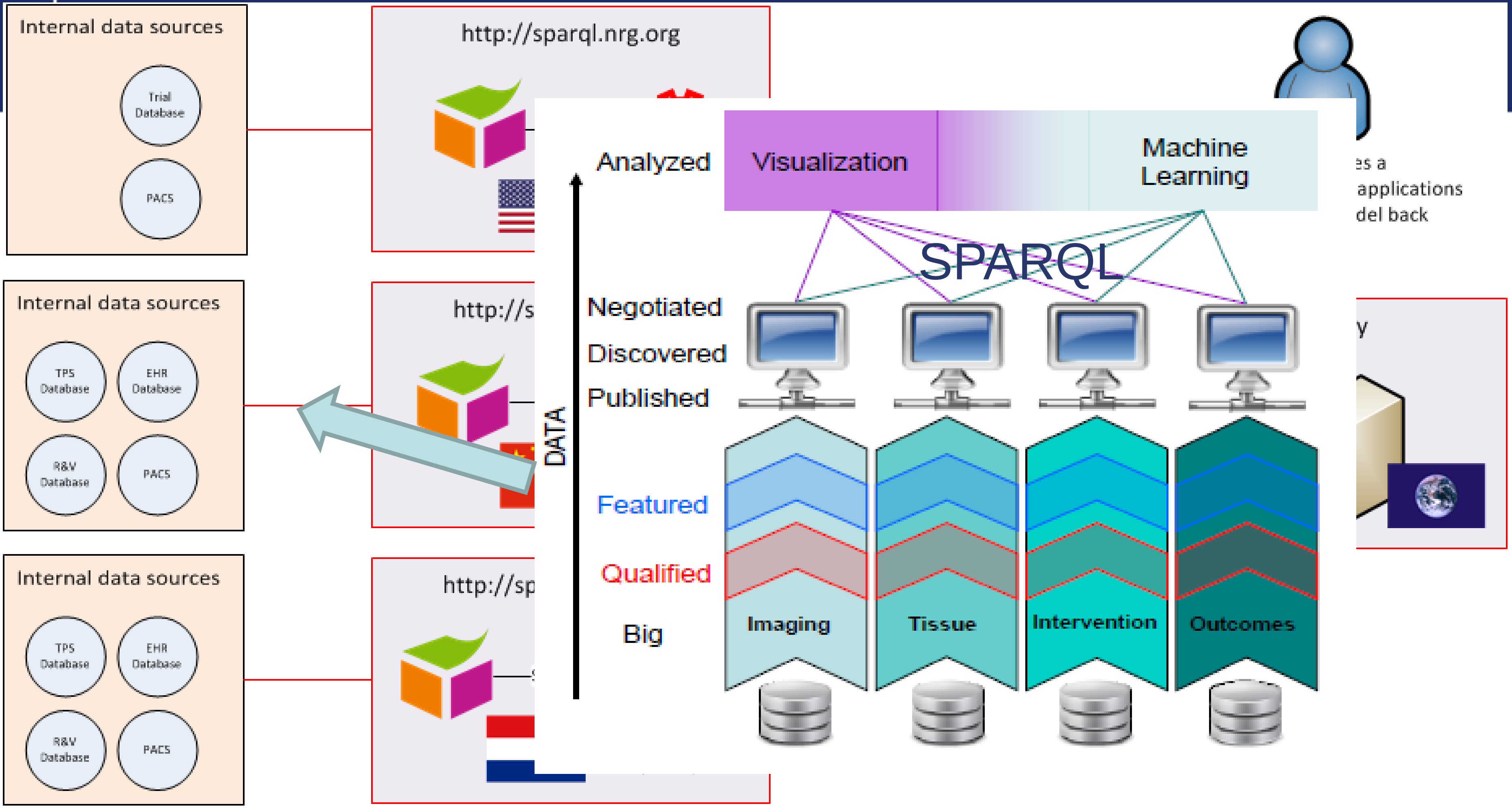
Private sandboxes: Project 1, Project 2, ↓, Project *n*

Research Portal User Interface
- ❖ Patient browser
- ❖ Query builder
- ❖ eCRF module
- ❖ Sandbox upload
- ❖ XML export

Project 1 · Project 2 · Project *n*

Research domain

Under the hood – Publish on Semantic Web



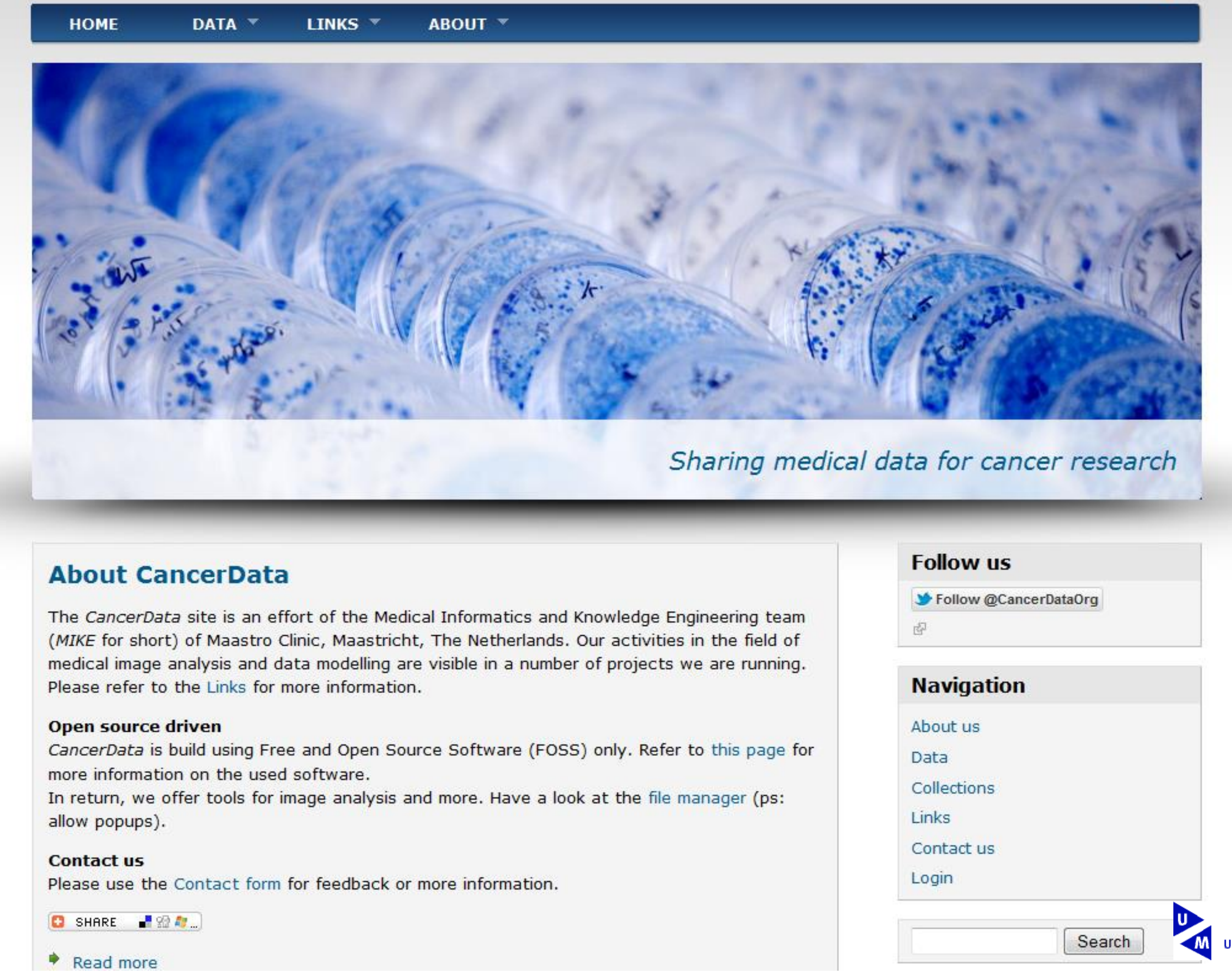
Take home message: Questions?

1. We need Decision Support Systems (DSS = a “meta TPS”) to manage the large quantity of data and implement Personalized medicine

2. Two complementary approaches: conventional clinical trials (+ data reuse) + Rapid Learning Health Care

3. Building cancer informatics tools to enable analysis, exploration, and rapid evaluation of novel therapies or stratification e.g. Distributed learning, Radiomics...

Open source data of publications: www.cancerdata.org



Sharing medical data for cancer research

About CancerData

The *CancerData* site is an effort of the Medical Informatics and Knowledge Engineering team (*MIKE* for short) of Maastro Clinic, Maastricht, The Netherlands. Our activities in the field of medical image analysis and data modelling are visible in a number of projects we are running. Please refer to the Links for more information.

Open source driven
CancerData is build using Free and Open Source Software (FOSS) only. Refer to this page for more information on the used software.
In return, we offer tools for image analysis and more. Have a look at the file manager (ps: allow popups).

Contact us
Please use the Contact form for feedback or more information.

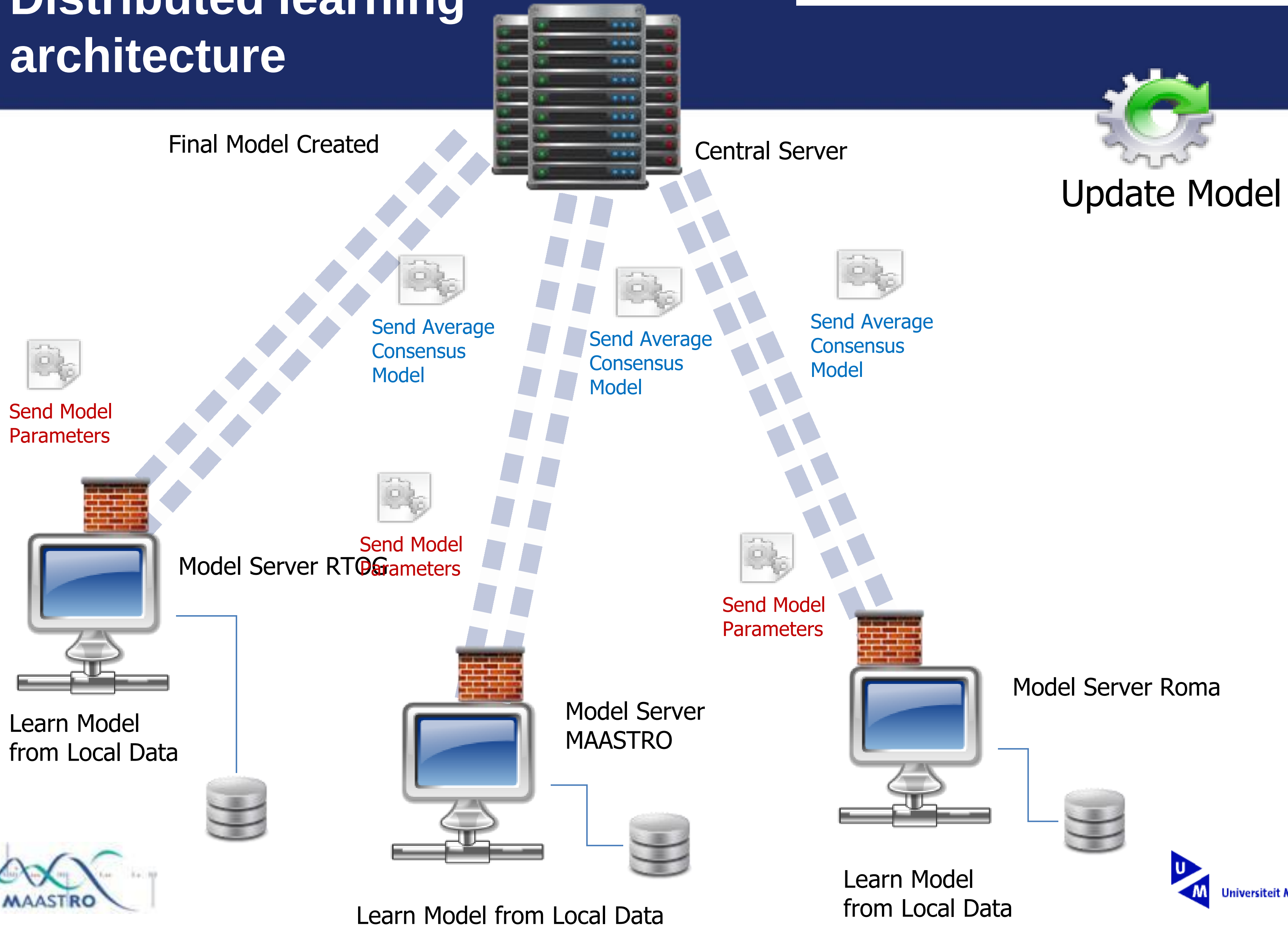
Follow us

Navigation

Universiteit Maastricht

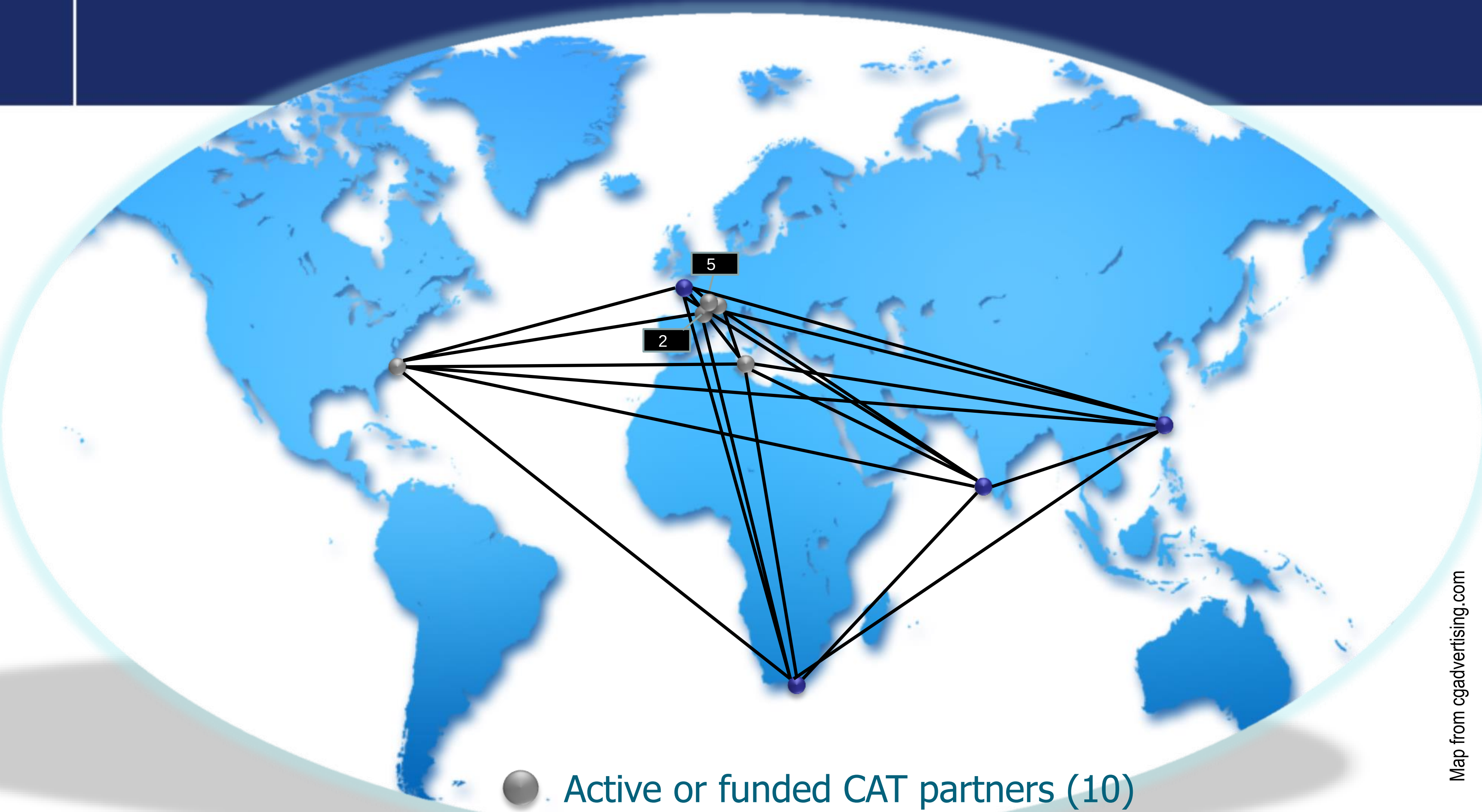
Distributed learning architecture

Only aggregate data is exchanged between the Central Server and the local Servers



Final Model Created
Central Server
Update Model
Send Average Consensus Model
Send Average Consensus Model
Send Average Consensus Model
Send Model Parameters
Send Model Parameters
Send Model Parameters
Model Server RTOG
Learn Model from Local Data
Model Server MAASTRO
Learn Model from Local Data
Model Server Roma
Learn Model from Local Data

Network euroCAT + in 9/2013



Map from cgadvertising.com

● Active or funded CAT partners (10)

● Prospective centers (4)

Herceptin:

Δ « Companion biomarker »

Tirapazamin:

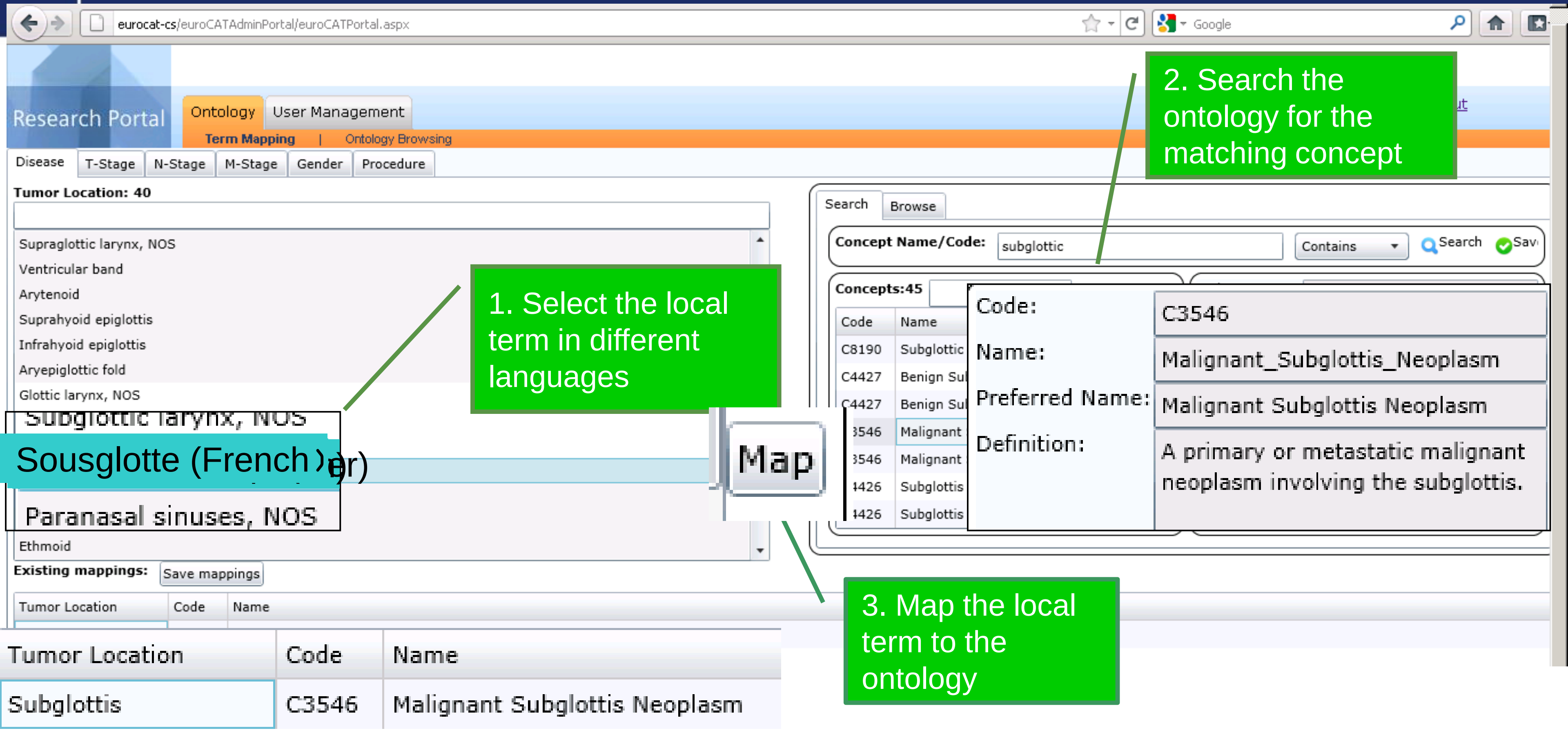
Biomarker +

Protontherapy

Biomarker -

Ontology mapping

(To be done once)



eurocat-cs/euroCATAdminPortal/euroCATPortal.aspx

Research Portal

Ontology | User Management

Term Mapping | Ontology Browsing

Disease | T-Stage | N-Stage | M-Stage | Gender | Procedure

Tumor Location: 40

- Supraglottic larynx, NOS
- Ventricular band
- Arytenoid
- Suprahyoid epiglottis
- Infrahyoid epiglottis
- Aryepiglottic fold
- Glottic larynx, NOS
- Subglottic larynx, NOS
- Sousglotte (French)er)
- Paranasal sinuses, NOS
- Ethmoid

Existing mappings: Save mappings

Map

Search | Browse

Concept Name/Code: subglottic | Contains | Search | Sav

Concepts:45

Code	Name
C8190	Subglottic
C4427	Benign Su
C4427	Benign Su
3546	Malignant
3546	Malignant
4426	Subglottis
4426	Subglottis

Code: C3546

Name: Malignant_Subglottis_Neoplasm

Preferred Name: Malignant Subglottis Neoplasm

Definition: A primary or metastatic malignant neoplasm involving the subglottis.

Tumor Location	Code	Name
Subglottis	C3546	Malignant Subglottis Neoplasm

1. Select the local term in different languages

2. Search the ontology for the matching concept

3. Map the local term to the ontology

4. See the result of your mapping

Our hypothesis

Protontherapy model der the *strict conditions* th population has been sele

... wi actorial Decision

The “one size fits all philosophy will not work with protons

The Dutch aproach for protontherapy

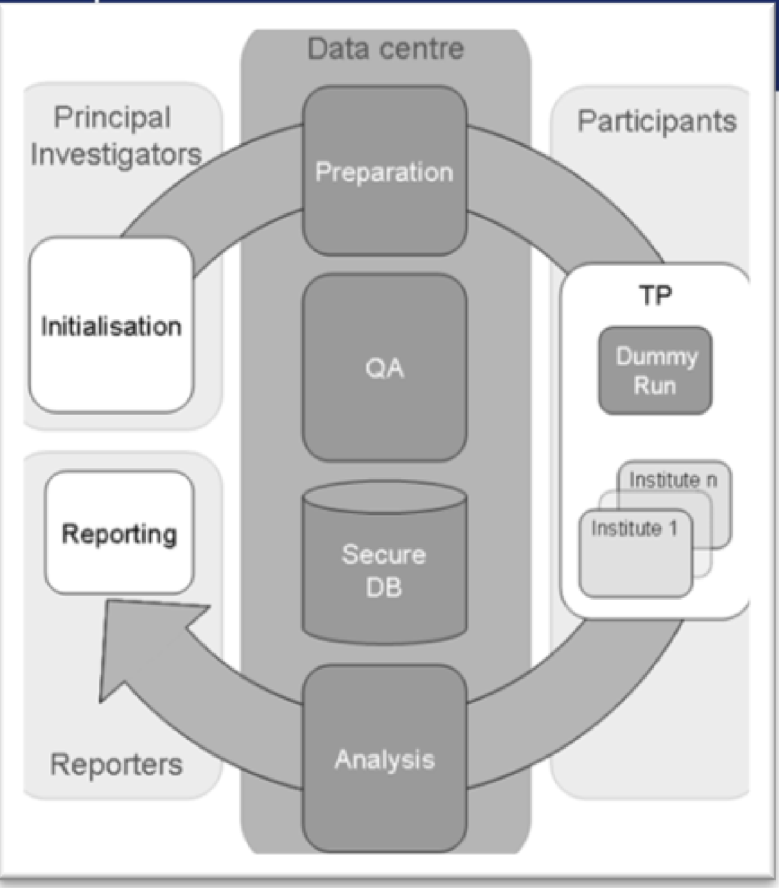
1. **The standard indications*** (pediatric, melanoma of the eyes…): fully reimbursed

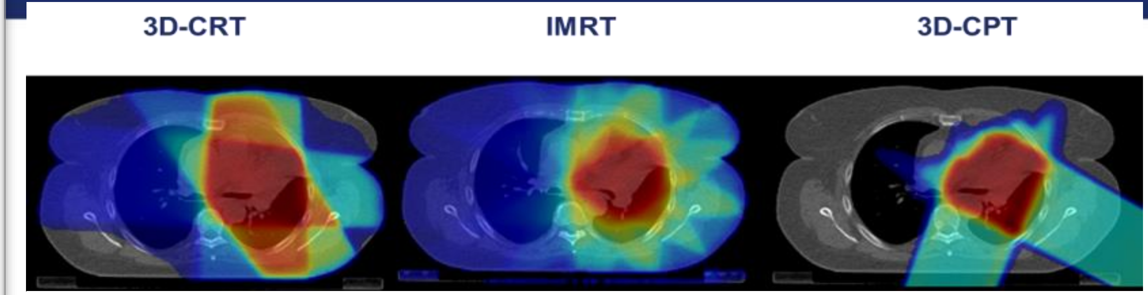
2. **The trial patients**: externally funded

3. **The model based indications*** (head & neck, lung, breast, prostate, *reirradiations*…): need an accredited Decision Support System (DSS)

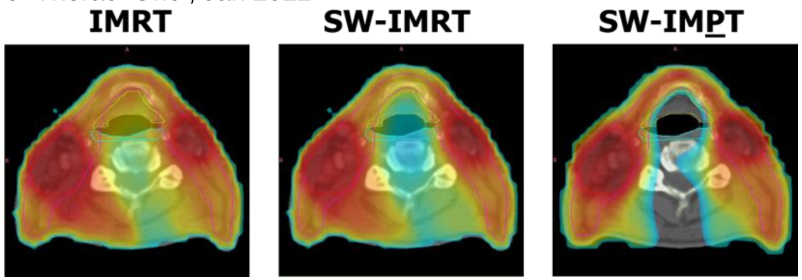
* Equipoise, ALARA… Only if there is no Dose escalation

Past: In Silico clinical trials

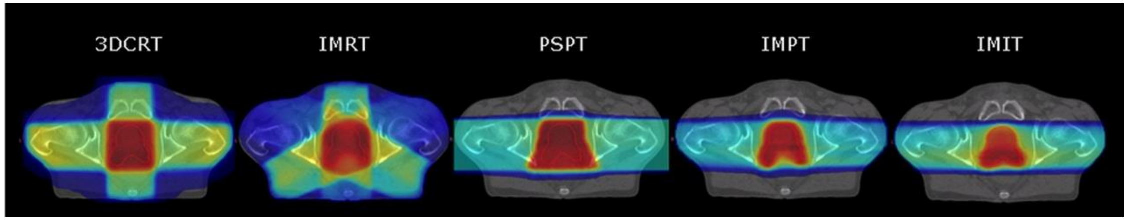




Roelofs, et al. J. Thorac. Onc., Jan 2012

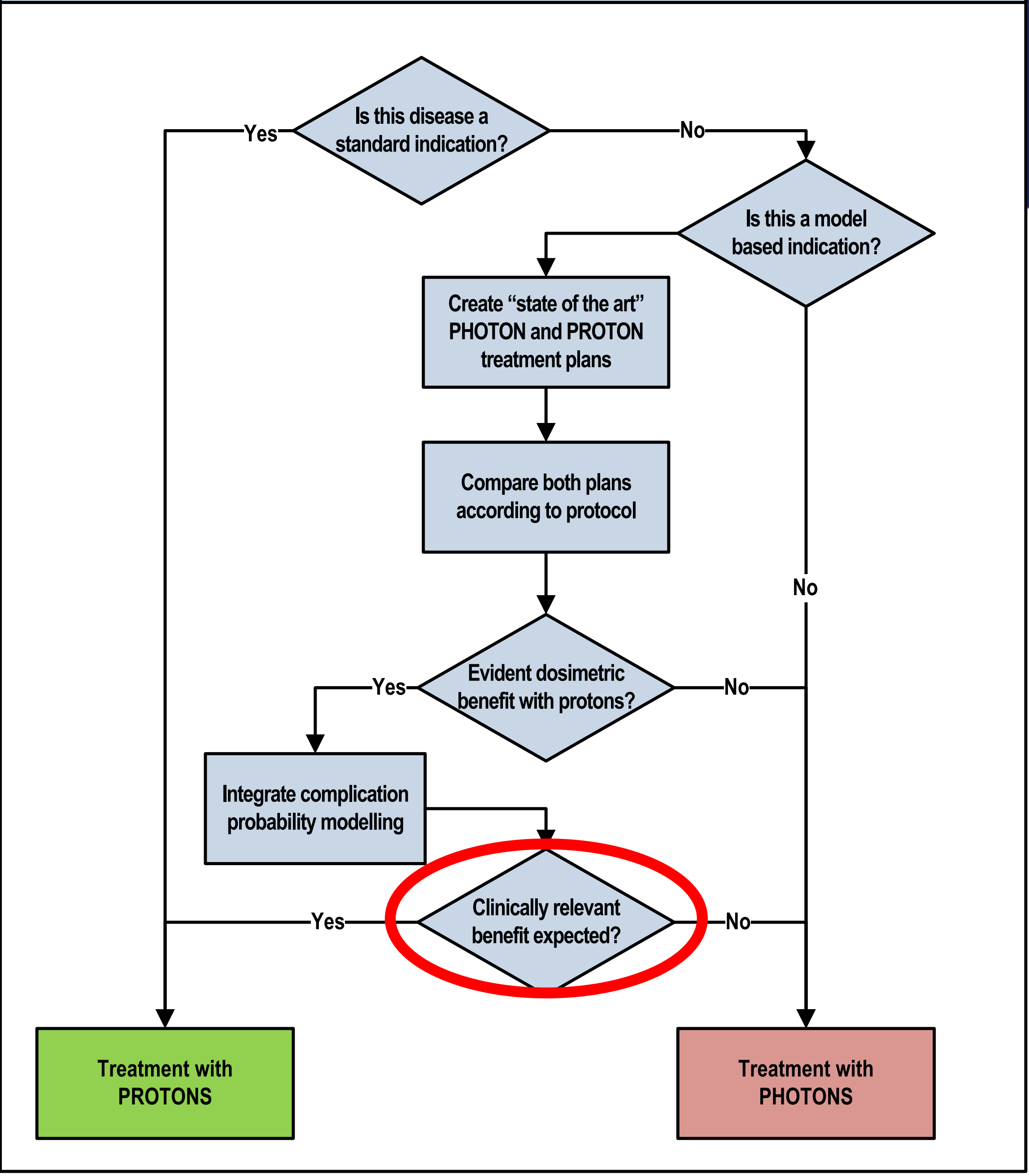


Van der Laan, et al. Acta Oncol. Apr 2013



Roelofs et al,, *2015*

Proton therapy reimbursement decision tree for the Netherlands



Is this disease a standard indication?

Yes

No

Is this a model based indication?

Create "state of the art" PHOTON and PROTON treatment plans

Compare both plans according to protocol

No

Evident dosimetric benefit with protons?

Yes

No

Integrate complication probability modelling

Clinically relevant benefit expected?

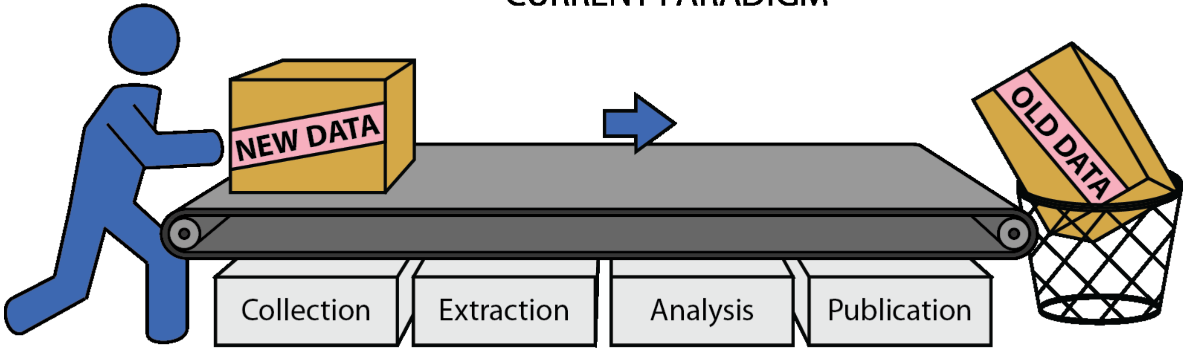
Yes

No

Treatment with PROTONS

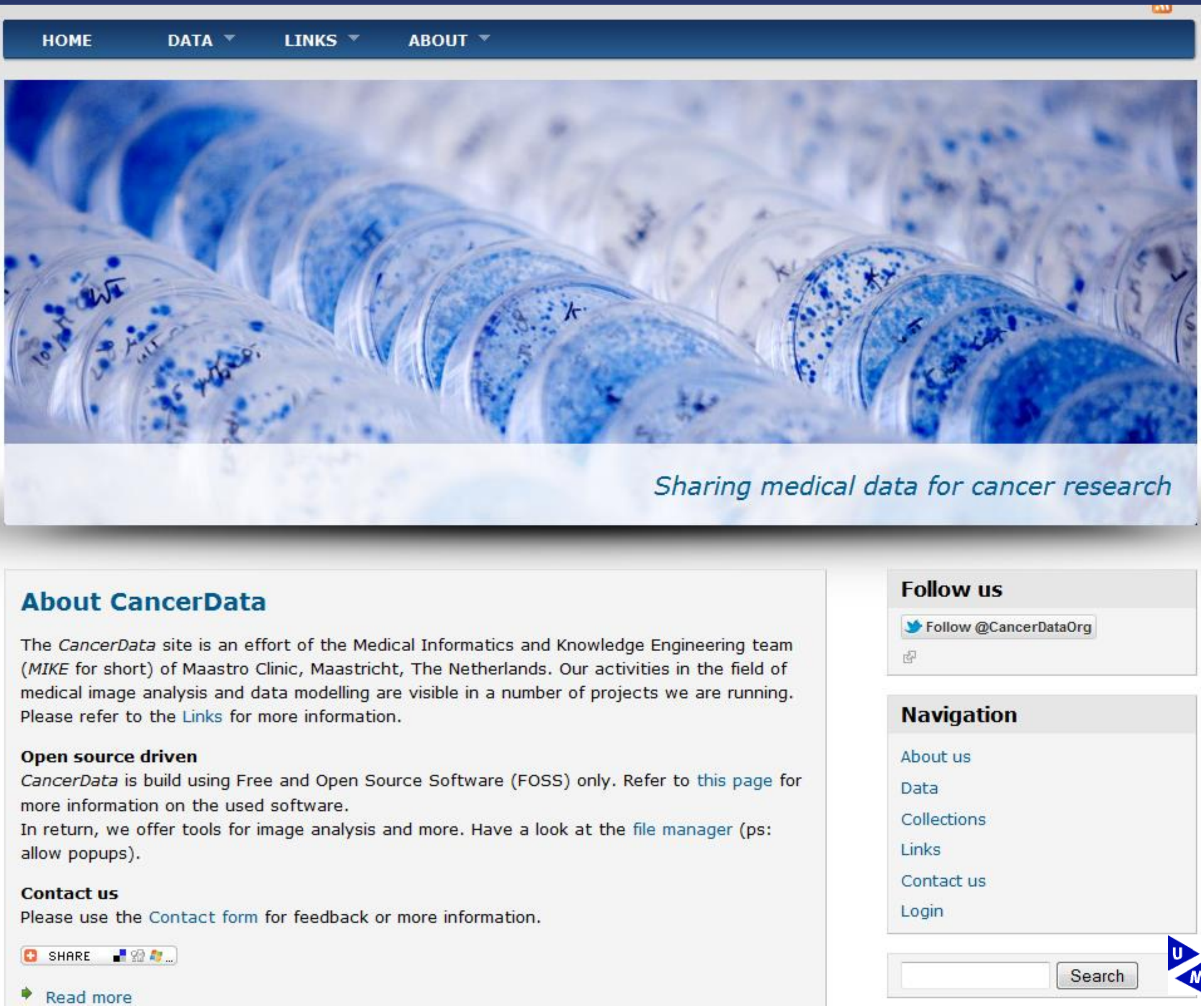
Treatment with PHOTONS

CURRENT PARADIGM



Modified from
Deasy et al.

Open source data of publications: www.cancerdata.org



About CancerData

The *CancerData* site is an effort of the Medical Informatics and Knowledge Engineering team (*MIKE* for short) of Maastro Clinic, Maastricht, The Netherlands. Our activities in the field of medical image analysis and data modelling are visible in a number of projects we are running. Please refer to the Links for more information.

Open source driven
CancerData is build using Free and Open Source Software (FOSS) only. Refer to this page for more information on the used software.
In return, we offer tools for image analysis and more. Have a look at the file manager (ps: allow popups).

Contact us
Please use the Contact form for feedback or more information.

Follow us

Navigation

- About us
- Data
- Collections
- Links
- Contact us
- Login

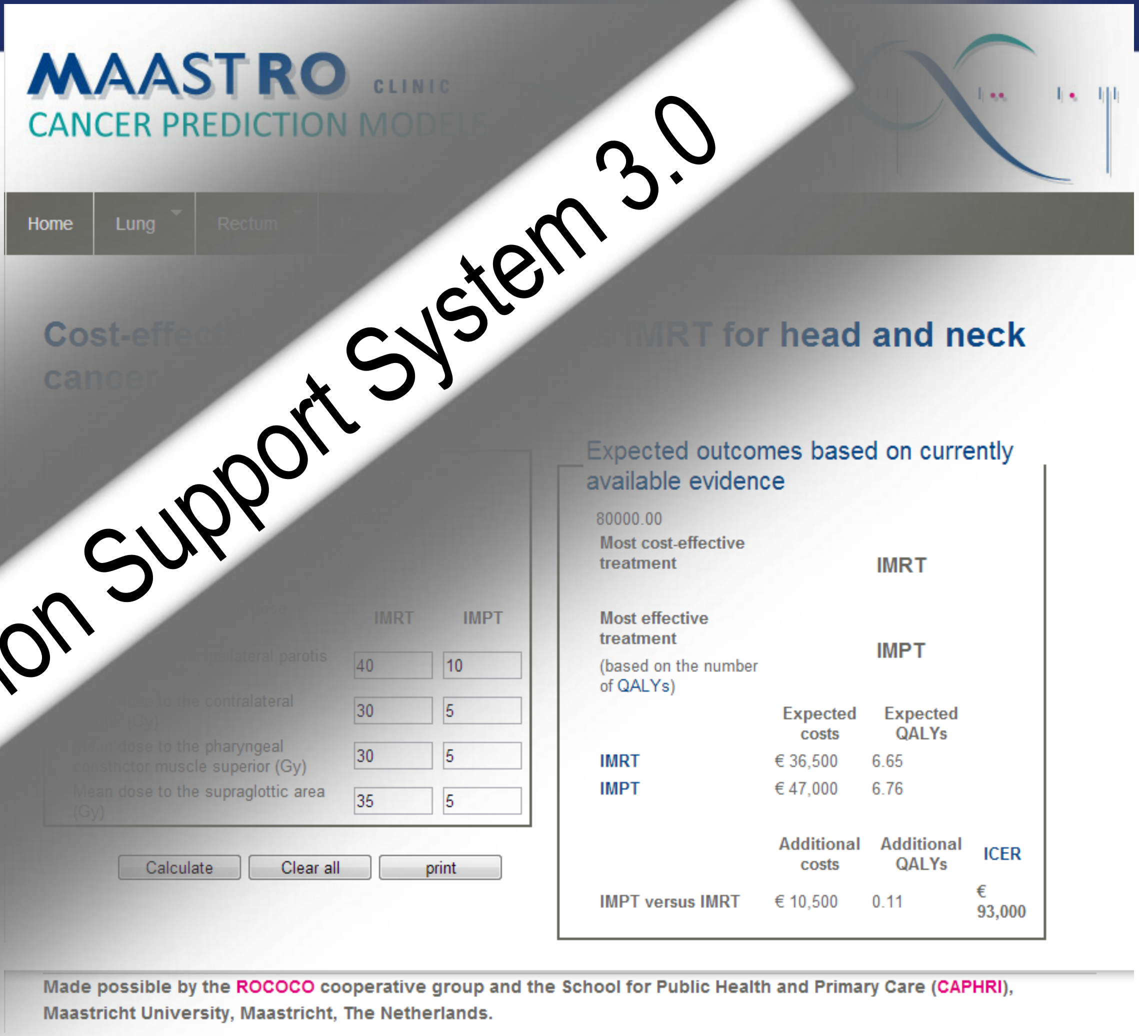
- The **Semantic Web** is an extension of the Web through standards by the World Wide Web Consortium (W3C). The standards promote common data formats and exchange protocols on the Web, most fundamentally the Resource Description Framework (RDF).
- According to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries". The term was coined by Tim Berners-Lee for a web of data that can be processed by machines.

Will this approach increase **the cost of care?**

No, it could even decrease them if you look at the cost of the *whole care cycle.*

Cost effectiveness: www.predictcancer.org

Decision Support System 3.0



MAASTRO CLINIC

CANCER PREDICTION MODE

Cost-effecti IMRT for head and neck cancer

	IMRT	IMPT
ipsilateral parotis	40	10
dose to the contralateral (Gy)	30	5
dose to the pharyngeal constrictor muscle superior (Gy)	30	5
Mean dose to the supraglottic area (Gy)	35	5

Expected outcomes based on currently available evidence

80000.00

Most cost-effective treatment	IMRT
Most effective treatment (based on the number of QALYs)	IMPT

	Expected costs	Expected QALYs
IMRT	€ 36,500	6.65
IMPT	€ 47,000	6.76

	Additional costs	Additional QALYs	ICER
IMPT versus IMRT	€ 10,500	0.11	€ 93,000

Made possible by the ROCOCO cooperative group and the School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands.