e-data: harmonisation of collected data for storage, usage and sharing

Franco-British bilateral symposium - October, 16th 2012

Quân Tran – IT Manager of IFCT, French Intergroup of Thoracic Oncology
Member of AcaDM, French academic data management network

Thank you!

- Claire Mouchot, The science and technology department of the French Embassy in the United Kingdom

- AcaDM, French academic data management network
  - Brice Dubois – CLCC Baclesse, Caen
  - Lilian Laborde – Institut Paoli Calmettes, Marseille

Contents

- Context in France - Oncology
- Need of harmonisation
  - Examples with the e-data of a patient
  - e-data of a patient in and off a hospital
  - Focus : Biomarqueurs France (epidemiology)
- Standards in clinical research
- Conclusion

Observation of the context in France Oncology

Great vitality of public research funding in France

- Many call for projects, ex. epidemiology studies
- French Network of Cancer Registries FRANCIM
- PHRC budgets (Hospital Clinical Research Programs)
- INCa (French National Cancer Institute), cancer plans:
  - First : 2003-2007
  - Second : 2009-2013 (1.9 billion €)

http://www.plan-cancer.gouv.fr/

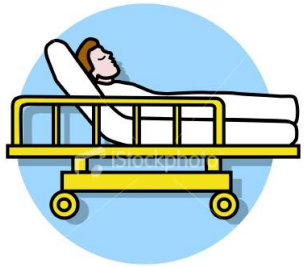
Observation of the context in France Oncology

However

- Difficulty to work together
  - in France
  - between european countries
- Budget
  - Focus on visible needs (CRAs, statistician)
  - But technical part often underestimated (equipment, software and HR). Data Management has to adapt to a given situation
- IT systems are selected
  - without harmonisation in the national scale
  - by managers (hospital direction, not research) and not data managers

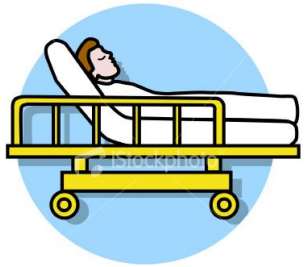
The e-data life of a patient

Example of a patient with lung cancer:



- Electronic Health record in a hospital
- Cancer registries
- Clinical study
- Biological resources development project
  - Ex. TVN pilot project (National Virtual Biobank)
- Epidemiology
  - Ex. Biomarqueurs France

The e-data life of a patient



Electronic Health record in a hospital:

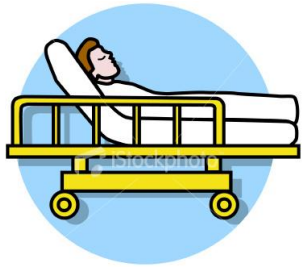
- File #: 201200048 (NIP, IPP ?)
- Monsieur DUPONT Jean-Pierre, né le 8 mars 1948 à Caen, a été diagnostiqué d'un Adénocarcinome du poumon le 28 septembre 2012 dans notre centre.

- Free text entry and/or codes (PMSI, medicalisation program of Information System) → Medico-economic evaluation

The e-data life of a patient

Regional Cancer Registries:

French Network of Cancer Registries FRANCIM



File #	Sexe	Prénom	Nom	Date naissance	Commune naissance
123456789	1	Jean-Pierre	DUPONT	08/03/1948	14118

Topo	Morpho	Date diagnostic	Lieu diagnostic
8255/3	C34.3	28/09/2012	14118

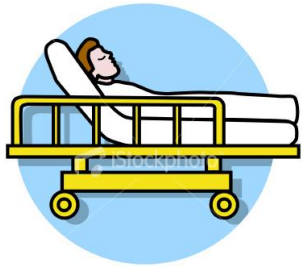
INSEE nomenclature for Caen (National Statistics Agency)

By the way, 08/03/1948:

Is it 3rd of august 1948, or 8th of march 1948 ?

The e-data life of a patient

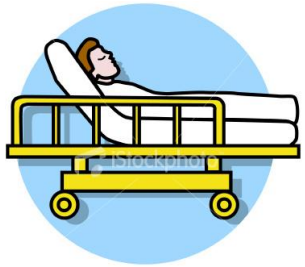
Clinical study (lung cancer)



Inclusion #	051-003
Last name	D
First name	J
Sex	Masculin (Code = 1, or = M ?, or = H ?)
Birthdate	08/03/1948
Birthplace zipcode	14000
Diagnosis date	28/09/2012
Histology	Adénocarcinome

The e-data life of a patient

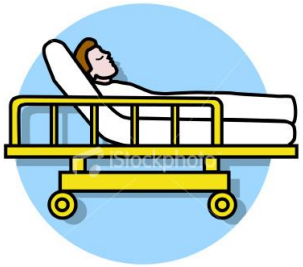
National Virtual Biobank (TVN project):
Tumorothèque Virtuelle Nationale



ID Patient	1315153
Last name	
First name	
Sexe patient	M
Date naissance patient	19480308
Birthplace zipcode	
Date du diagnostic	20120928
Diagnostic principal de la tumeur initiale CIM 10	C34.3
Code organe CIMO	C34.3
Type lésionnel histopathologique CIMO	8255/3
Code organe ADICAP	RP
Type lésionnel histopathologique ADICAP	A7V0

Date format :
YYYMMDD

The e-data life of a patient



The same information is stored several times, but in different ways

The e-data life of a patient INSIDE

Within a hospital, different & independent IT systems:

- Electronic Health record in hospital
- DIAMIC (anatomopathology)
- STARE (hematology)
- BIOBASE II (frozen samples management)
- APOR (biomarkers)
- …

The e-data life of a patient OUTSIDE

- French Network of Cancer Registries FRANCIM
- Clinical study
- Biomarqueurs France
- TVN pilot project (National referential)
  - 6 biobanks for pilot phase

The e-data life of a patient
Links between IN and OUT

IT systems specific to
Paris Hospitals (AP-HP)

Electronic
Health record

DIAMIC

STARE

BIOBASE II

APOR

?

Cancer registries

Clinical study

Biomarqueurs
France

TVN

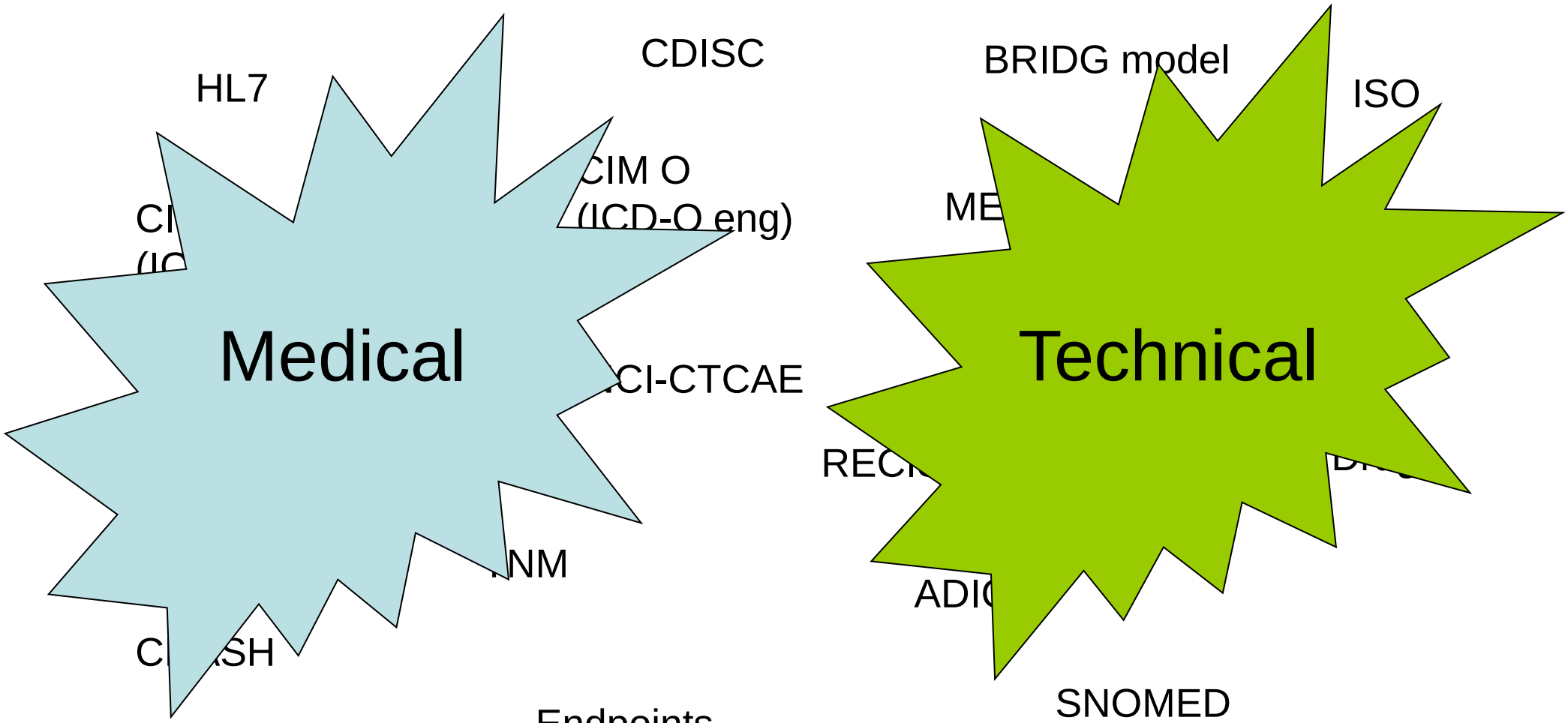
INSTITUT NATIONAL DU CANCER

The e-data life of a patient

- Knowledge of the field is essential to solve the question mark
  - Identification of patient
  - Identification of tumor sample
  - Name of patient (what is stored: married/maiden name ?)
  - Date format (cultural differences)
  - Duplicates identification & management
  - File formats (XML, CSV, Excel...)
  - Units
  - …

The e-data life of a patient
Standards may help

There are standards in Oncology, but many abbreviations and acronyms…



Version of each?

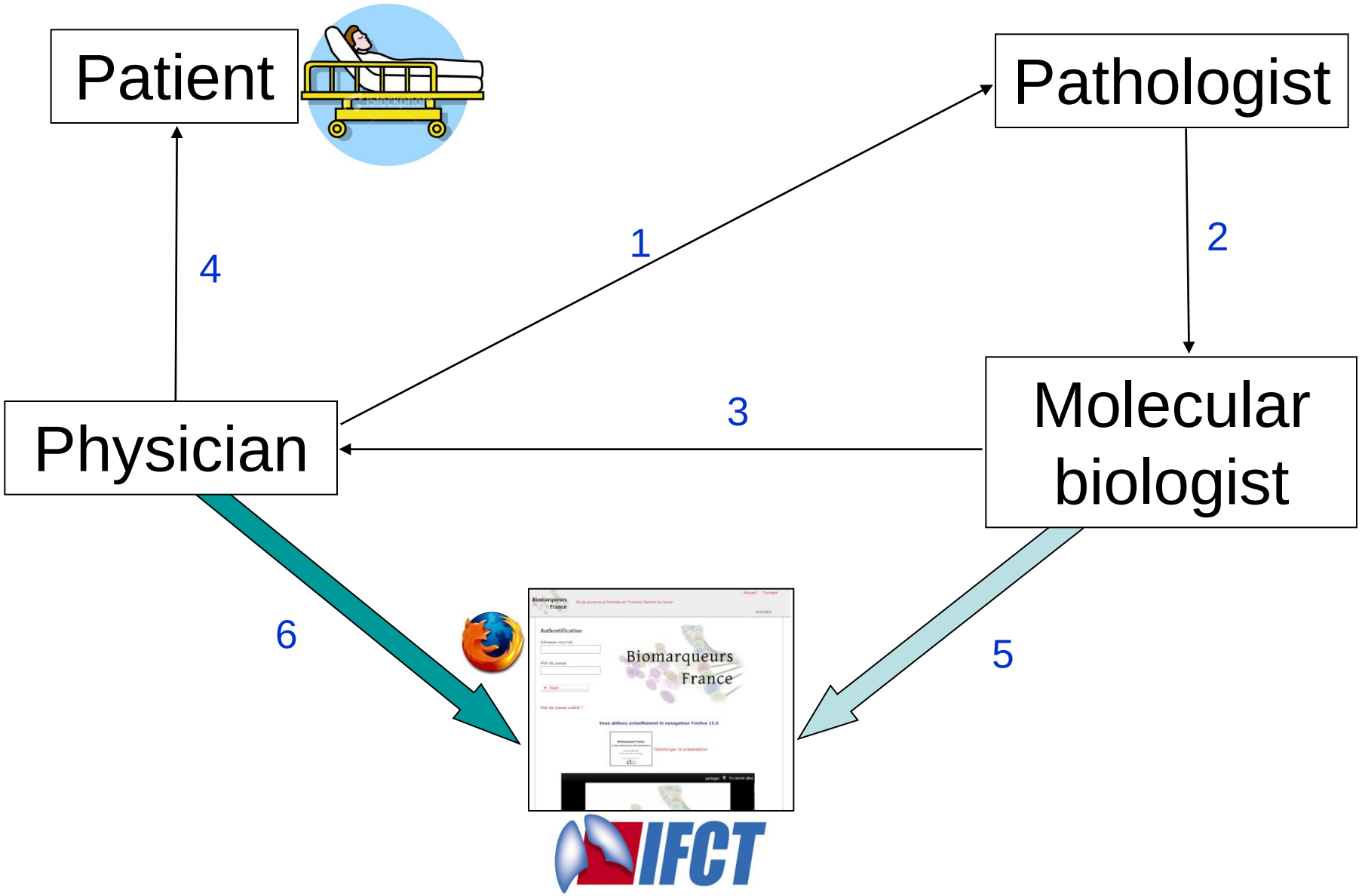
Focus: Biomarqueurs France

- 28 molecular biology platforms
- 17,000 expected analyses in 1 year
- Financed by INCa (3.5 million € /year for lung)
- But

What for?

Biomarqueurs France

Focus: Biomarqueurs France



Focus: Biomarqueurs France

- Guidelines and recommendations by INCa
  - Which biomarkers
  - Which data should be on report
- 28 molecular biology platforms

  → 28 « standard » reports for analyses


Mesure 21

SOINS ET VIE DES MALADES

AOÛT 2010

Bonnes pratiques pour la recherche à visée théranostique de mutations somatiques dans les tumeurs solides

LE PRÉLÈVEMENT
QUELLES MUTATIONS RECHERCHER ?
LA RECHERCHE DES MUTATIONS
LES RÉSULTATS

DESTINÉ À L'USAGE DES PROFESSIONNELS DE SANTÉ

L'identification d'altérations génétiques au sein des cellules cancéreuses a débouché sur la mise en évidence de nouveaux biomarqueurs moléculaires. En complément du diagnostic et des facteurs pronostiques anatomocytopathologiques, ces paramètres sont aujourd'hui indispensables pour le diagnostic, la classification, le choix ou/et la surveillance du traitement de beaucoup d'hémopathies et de tumeurs solides malignes.

En particulier, alors que le développement des anticorps monoclonaux anti-récepteurs à l'EGF (*EGFR*) a constitué une avancée importante dans la prise en charge des patients atteints de cancer colorectal métastatique, plusieurs études ont montré que seuls les patients dont la tumeur ne présentait pas de mutation du gène *KRAS* étaient susceptibles de bénéficier de ce traitement. Dans ce contexte, l'Agence européenne du médicament a autorisé l'utilisation du cetuximab (Erbitux®) et du panitumumab (Vectibix®)

www.e-cancer.fr

Institut National du Cancer

Focus: Biomarqueurs France

- 28 different IT departments and systems
  - No possible interoperability
  - No consensus on a common format for data: no possible import in a central system
  - Manual work of coding. HR, time and cost +++

Focus: Biomarqueurs France

DATE DE DEMANDE DE DIAGNOSTIC	DATE DE RECEPTION DU PRELEVEMENT	DATE DE DEBUT DE MANIPULATION	DATE DE RESULTAT	DATE ADDENDUM 1	DATE ADDENDUM 2
28/03/2012	16/07/2012	23/07/2012	25/07/2012	06/08/2012	21/08/2012

DATE DE PRELEVEMENT	TYPE DE TUMEUR	STADE DE LA MALADIE
14/03/2012	Adénocarcinome	IV

TYPE DE PRELEVEMENT	SITE DE PRELEVEMENT
Biopsie	Surrénale

Statut tabagique : Ancien Fumeur 40PA

Conditionnement : Formol;

Tumorothèque :

N° Bloc : C12/1197

Focus: Biomarqueurs France

RESULTATS

	EXON 19	EXON 21 (L858R)
TECHNIQUE	HRM-Analyse de fragment	PCR ASO
RESULTATS	NON MUTE	NON MUTE

	EXON 20 (T790M)	EXON 18	KRAS
TECHNIQUE	PCR ASO	PCR - Séquençage	HRM-SEQUENCAGE
RESULTATS	NON MUTE	NON MUTE	c.34G>T/p.G12C

	EML4-ALK	BRAF (V600E)	HER2 (EXON 20)
TECHNIQUE	Immunohistochimie	PCR ASO	non réalisé
RESULTATS	IMMUNOHISTOCHIMIE NEGATIVE	NON MUTE	NON REALISE

CONCLUSION

ABSENCE DE MUTATION ACTIVATRICE DES EXONS 19 ET 21 D'EGFR

Focus: Biomarqueurs France

DEMANDE DE RECHERCHE DE MUTATION DU GENE EGFR

MOTIF DE LA DEMANDE : Non précisé (NP)

ONCOLOGUE REFERENT : NP

CONTEXTE CLINIQUE : Métastase d'un adénocarcinome broncho-pulmonaire mucosécrétant

TTF-1+ stade pTNM : NP

PRELEVEMENTS : Prélèvement N° : **C030** , reçu le **20/04/2012**, correspondant à **bloc d'inclusion en paraffine d'une ponction de la surrénale droite, prélevée le 12/04/2012. Réf. BH29815**

Bloc paraffine n° : **C030** **/ BH 29** fixé en Formol : **NP**

POURCENTAGE DE LA TUMEUR SUR L'ECHANTILLON : 60%

MACRO DISSECTION NECESSAIRE SUR BLOC : **NON**

Demande transmise le 23/04/2012. Demande transmise le 23/04/2012.

- Requires more time for the physician
- Prior work of data identification and transcodification before entry

RECHERCHE DE TRANSLOCATION ALK/EML4

MOTIF DE LA DEMANDE : NP

ONCOLOGUE REFERENT : CF.

CONTEXTE CLINIQUE : cf. – Statuts EGFR et KRAS : non mutés

ETUDE IMMUNOHISTOCHIMIQUE ANTI ALK

Technique réalisée après démasquage EDTA pH8 – AC Clône 5A4 – Clinisciences – 1/50ème
Contrôle positif externe = +

RESULTAT :

Absence d'expression de ALK, en faveur d'une absence de réarrangement de ALK/EML4

Analyse par FISH : non réalisée en l'absence de demande expresse et argumentée du clinicien/oncologue

Prélèvements communiqués :
- lame HES (contrôle histologique) + 2 LBSF+ archivées dans le Département de Pathologie
- Bloc réf. BH29815 renvoyé au Dr GEORGES, ACP Paris 75012

Compte rendu fait le 29/05/2012

Absence of ALK expression, favoring the absence of ALK/EML4 rearrangement

« Negative »

Focus: Biomarqueurs France

- What the physician needs:
  - EGFR mutational status : Yes or No, Positive or Negative
    → Adapted treatment

- What he gets:
  - Absence of mutation
  - Mutation not detectable
  - Absence of detectable mutation
  - c.2303G>T, p.Ser768 Ileu
  - Not amplifiable
  - Not amplified
  - Not interpretable
  - non-contributive analysis
  - …

+

Or ?

-

Date du prélèvement :

Code prélèvement / N° Bloc :

Date de la demande :

Modalité d'obtention du prélèvement :

○ Fibroscopie ○ Chirurgie ○ Ponction transthoracique sous scan ○ Autre

Type histologique CBNPC : ○ Epidermoïde ○ Adénocarcinome ○ Grandes Cellules ○ Autre

Résultats :

EGF-R	○ Muté activ ○ Muté résist ○ Muté autre ○ Non muté ○ Ind. ○ Non fait
KRas	○ Muté ○ Non muté ○ Indéterminé ○ Non fait
BRaf	○ Muté ○ Non muté ○ Indéterminé ○ Non fait
Réarrangement ALK	○ Réarrangé ○ Non réarrangé ○ Indéterminé ○ Non fait
PI3K	○ Muté ○ Non muté ○ Indéterminé ○ Non fait
HER2	○ Muté ○ Non muté ○ Indéterminé ○ Non fait

Traitement, 1ère Ligne

Le traitement de **1ère ligne** a-t-il été adapté au résultat des analyses biologiques ?

- ◯ **Oui** ◯ **Non**
  - ◯ Délai trop long
  - ◯ Autre raison

Traitement réellement administré en **1ère ligne** :

- ◯ Chimiothérapie basée sur taxane
- ◯ Chimiothérapie basée sur vinorelbine
- ◯ Chimiothérapie basée sur pemetrexed
- ◯ AMM EGFR-TKI
- ◯ Essai clinique. Type de molécule :
- ◯ Autre :
- ◯ Pas de traitement de 1ère ligne

Le patient a-t-il reçu une radiothérapie ?

- ◯ Non
- ◯ Oui, à visée palliative
- ◯ Oui, thoracique à visée curative

Evaluation de la réponse au traitement de **1ère ligne** : ◯ RC ◯ RP ◯ ST ◯ PRG ◯ NE ◯ NF

En cours à ce jour ?

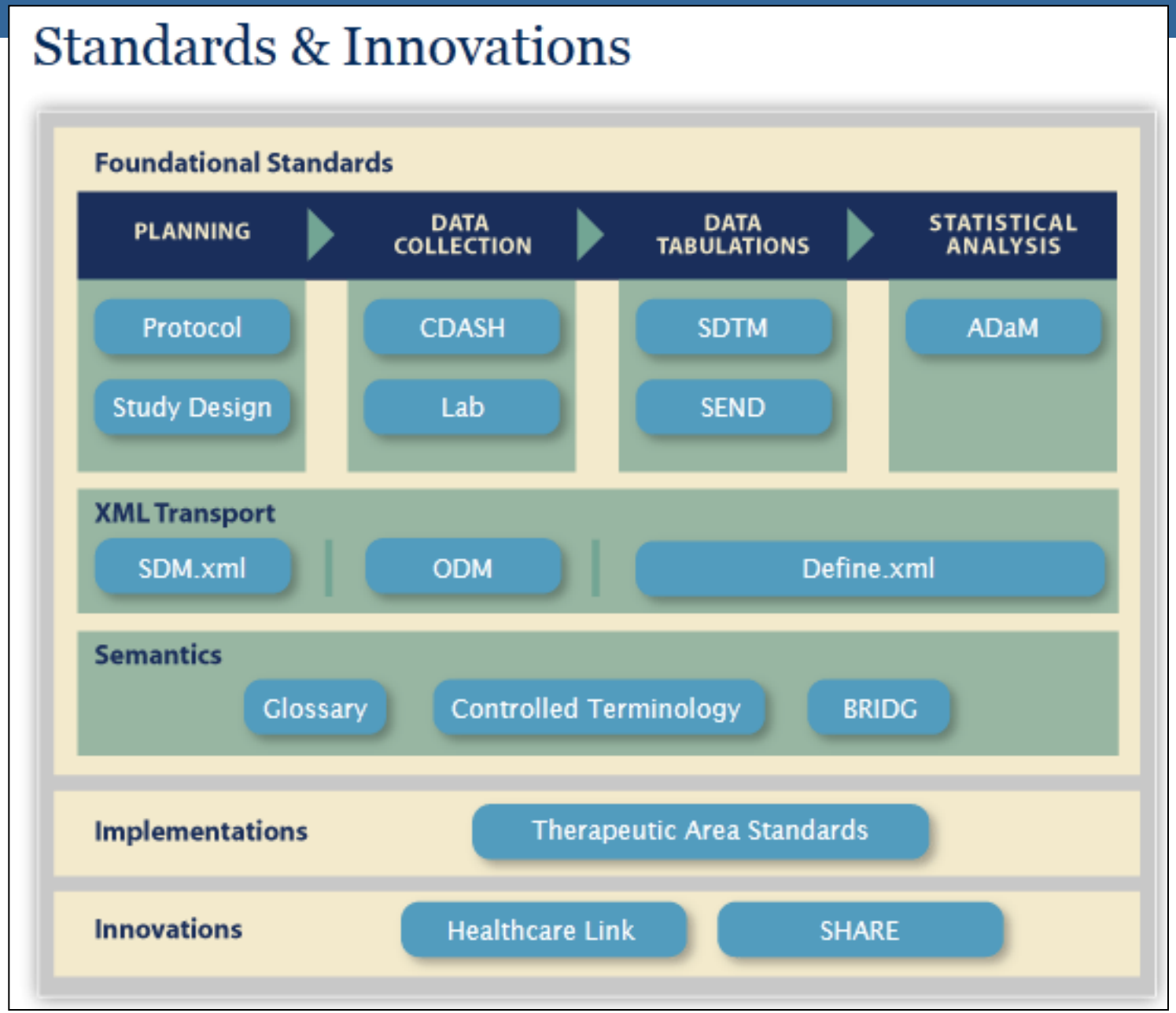
- ◯ Oui, en date du :
- ◯ Non, progression en date du :
- ◯ Non, toxicité en date du :
- ◯ Ne sais pas
- ◯ Autre

Date de fin Ligne 1 : (jj/mm/aaaa)

Standards in clinical research

CDISC

Clinical Data Interchange Standards Consortium

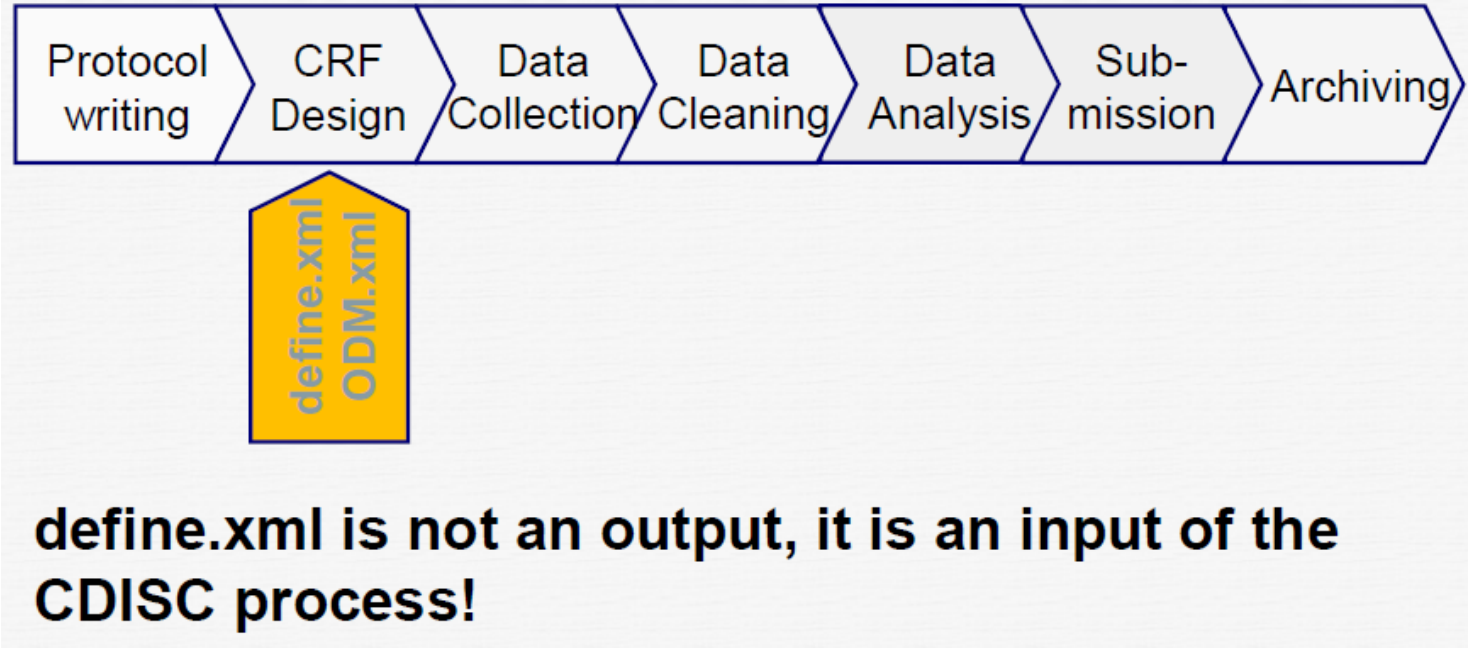


Standards in clinical research
CDISC

- IFCT chose a product with CDISC ODM certification
  - The Operational Data Model *is designed to facilitate the archive and interchange of the metadata and data for clinical research*
  - *The XML-based content and format standard for the acquisition, exchange, reporting or submission, and archival of case report form (CRF)-based clinical research data*
  - Independent from databases, storing of ODM is independent from hard- and software

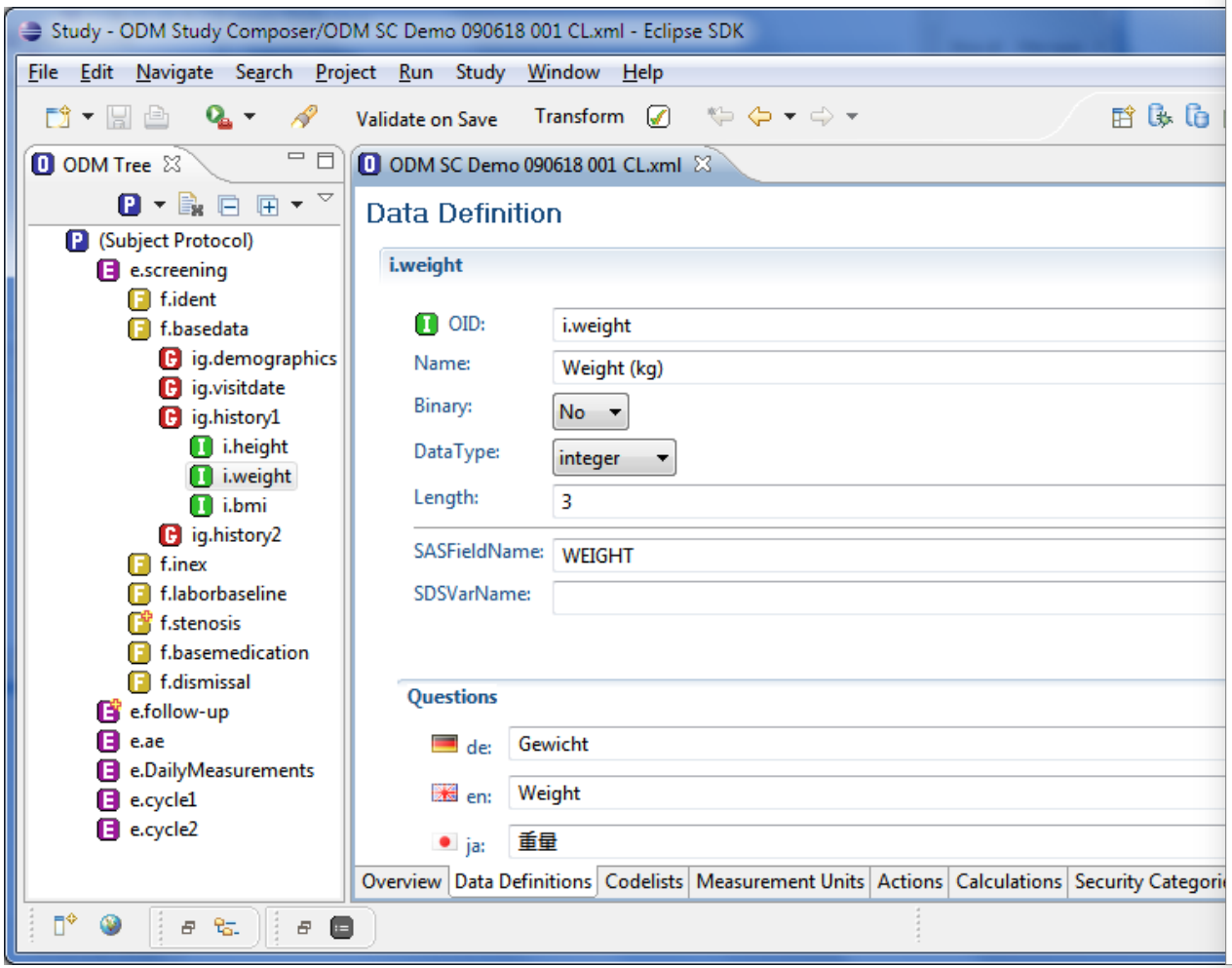
Standards in clinical studies
CDISC

- Their concept: inject CDISC at the start of the process

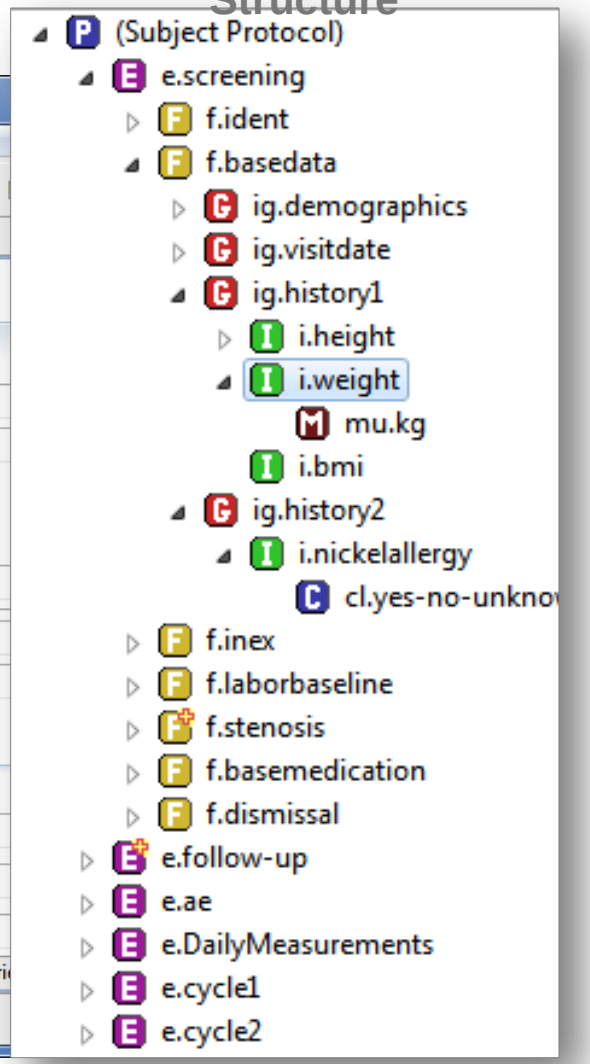


Standards in clinical studies
CDISC

Use the ODM structure to…



… describe the Study Structure



Standards in clinical research
CDISC

- Metadata in XML format connected to Data in XML format
  - Unambiguous interpretation of data
  - Easy transformation into other formats with little programming
  - More easy exchange between databases of different research groups
  - Possibility to use different EDC systems simultaneously and to switch between EDC systems during long-term projects
  - Robust archiving format including audit trail

- International controlled terminology → easier aggregation of different databases for cross-trial analysis

Conclusion (1/2)

Ideal full interoperability may not be feasible, but some improvements may help:

- Collaborative vision planned in the beginning of the project
  - RYTHMIC and ITMIG (thymoma cancer)
- Before developing a new system/project, design from the beginning to adapt it to standard
- Datamanagement department financing should not be underestimated

Conclusion (2/2)

- Clinical research: CDISC → Expensive IT and EDC systems for FDA submissions
- Academic world:
  - Not concerned with submissions, most of the time
  - We do not have the same financial means
  - Still, many groups begin to integrate CDISC, but stay PRAGMATIC → « CDISC like »
  - Active workgroup within AcaDM, project **YacaCDISC**
  - Must continue the effort of CDISC implementation for future european and international collaborations